Simulationen am Computer als hochschuldidaktisches Instrument in der Stochastik – Entwicklung eines Konzeptes und empirische Untersuchung seines Nutzens

Dissertation

zur Erlangung des akademischen Grades
Dr. rer. nat.

eingereicht an der
Mathematisch-Naturwissenschaftlich-Technischen Fakultät
der Universität Augsburg

von
Franziska Maria Wandtner-Reichmann

im Juni 2020

Universität
Augsburg
University

Erstgutachter: Prof. Dr. Reinhard Oldenburg

Zweitgutachter: Prof. Dr. Götz Kersting

Tag der mündlichen Prüfung: 07.09.2020

Danksagung

Mein herzlicher Dank gilt an erster Stelle meinen beiden Betreuern Prof. Dr. Götz Kersting und Prof. Dr. Reinhard Oldenburg für die Idee zu dieser Arbeit, ihre hilfreichen Ratschläge und die wertschätzende Unterstützung, die Freiheit, die sie mir gelassen haben, und ihren langen Atem.

Der AG Stochastik an der Goethe Universität in Frankfurt danke ich für die fruchtbaren Diskussionen, und hierbei besonders Prof. Dr. Anton Wakolbinger für seinen fachlichen Rat und seinen Enthusiasmus sowie Dr. Brooks Ferebee und Prof. Dr. Gaby Schneider für ihre wertvollen Hinweise zur statistischen Auswertung. Die Unterstützung und Förderung, die ich in dieser Gruppe besonders von Prof. Dr. Götz Kersting und Prof. Dr. Anton Wakolbinger bereits während meines Studiums genießen durfte, sind für mich einzigartig und in verschiedener Hinsicht prägend gewesen.

Den Studierenden an der Goethe Universität in Frankfurt, die sich zum Ausfüllen der Fragebögen und zu Interviews bereit erklärt haben, danke ich für ihre wichtigen Beiträge zur Evaluation des Projektes.

Der Stiftung Polytechnische Gesellschaft danke ich für ihre finanziellen und ideellen Zuwendungen während meines Studiums und zu Beginn dieses Promotionsprojektes.

Außerdem danke ich Katharina Som und Bastian Werth für unsere Freundschaft, die uns seit Beginn unseres Mathematik-Studiums verbindet, und die fachlichen und nicht-fachlichen Gespräche diese Arbeit betreffend.

Schließlich bin ich meiner Familie zu tiefem Dank verpflichtet: Meinen Eltern für ihre bedingungslose Unterstützung, die sie mir Zeit meines Lebens schenken, meinem Bruder und meinem Mann, den ich ohne das Projekt, das zu dieser Arbeit geführt hat, vermutlich niemals getroffen hätte. Seine Leidenschaft für die Mathematik, der inspirierende Austausch und seine genauen Nachfragen sowie sein unerschütterliches Vertrauen in meine Fähigkeiten haben mich beflügelt und ermutigt, diese Arbeit zu vollenden. Mit ihm und unseren Kindern durch das Leben gehen zu dürfen, macht mich sehr glücklich.

Zusammenfassung

Fehlvorstellungen im Gebiet der Stochastik sind auch bei mathematisch versierten Menschen weit verbreitet und hartnäckig – das zeigen zahlreiche Untersuchungen aus Vergangenheit und Gegenwart. Inwiefern Simulationen am Computer zur Korrektur beitragen können und welchen Nutzen sie insgesamt für die Entwicklung stochastischer Fertigkeiten haben, ist Gegenstand dieser Arbeit. Hierzu wurde ein Konzept entwickelt, wie in einer Universitätsveranstaltung zur Elementaren Stochastik für Bachelor- und Gymnasiallehramts-studierende Simulationsaufgaben eingesetzt werden können. Dabei wurde das Programm R verwendet. Der didaktische Wert dieser „R Aufgaben" wird in der vorliegenden Arbeit analysiert. In einer groß angelegten explorativen Studie, die sich über zwei Semester erstreckte, wird anschließend durch selbst entwickelte Fragebögen untersucht, wie sich die Bearbeitung der Simulationsaufgaben auf die Entwicklung stochastischer Fertigkeiten sowie die Selbsteinschätzungen der Studierenden bezüglich ihrer Fähigkeiten und Einstellungen dem Fach Stochastik gegenüber auswirken. Die empirische Untersuchung unterteilt sich in eine Vorstudie mit 44 Teilnehmern aus einem Durchgang, in dem keine Simulationsaufgaben eingesetzt wurden, und eine Hauptstudie mit 91 Teilnehmern aus dem Semester, in dem das entwickelte Konzept erprobt wurde. In die Auswertung wurden nur diejenigen Studierenden einbezogen, die die Fragebögen sowohl zu Beginn als auch am Ende des Semesters ausgefüllt hatten, sodass individuelle Entwicklungen untersucht werden konnten. Zusätzlich erfolgte in der Hauptstudie auch eine explizite Evaluation des Konzeptes durch die Studierenden sowie als Orientierungshilfe für weitergehende mögliche qualitative Schlussfolgerungen die Befragung von insgesamt 15 Studierenden in halbstrukturierten Interviews.

Auch wenn sich die Fehlvorstellungen, wie erwartet, als tief verwurzelt und damit schwer korrigierbar erwiesen, konnten durch die regelmäßige Bearbeitung der „R Aufgaben" Erfolge erzielt werden – insbesondere bei expliziter Thematisierung der Fehlvorstellungen in den Aufgaben. Das Konzept scheint dabei insbesondere für tendenziell leistungsschwächere und eher unsichere Studierende gewinnbringend, aber auch bei den leistungsstärkeren Studierenden gibt es Hinweise auf eine Verbesserung ihrer stochastischen Fertigkeiten.

Der Einsatz der Aufgaben wird von den Studierenden mehrheitlich befürwortet, wobei für sie auch der Nutzen im Hinblick auf den Erwerb technischer Fertigkeiten des Programmierens maßgeblich zu sein scheint. Insgesamt ist auf technischer Ebene aber ein breites Unterstützungsangebot erforderlich. Das für die Erprobung des Konzeptes entwickelte Angebot, bestehend aus Templates und einer zusätzlichen „Fragestunde", sollte ggf. noch durch die Einrichtung eines Vorkurses ergänzt werden.

INHALTSVERZEICHNIS

I. VORÜBERLEGUNGEN UND BEGRÜNDUNG FÜR DAS VORHABEN

Die Stochastik als die mathematische Lehre des Zufalls ist aus vielen Bereichen der modernen Wissenschaften nicht mehr wegzudenken. So hat sie beispielsweise nicht nur die Wirtschaftswissenschaften maßgeblich beeinflusst, es gibt auch eine Reihe von Anwendungen in der Informatik, den Naturwissenschaften und der Medizin. Trotz der hohen Präsenz stochastischer Fragestellungen und Konzepte finden sich bisher gerade im hochschuldidaktischen Bereich kaum Untersuchungen über eine günstige Vermittlung.

Die Lehrbücher beschränken sich im Wesentlichen auf einen theoretischen Zugang, der zwar das wichtige technische und formale Handwerkszeug für eine wissenschaftliche Beschäftigung mit der Stochastik bereitstellt, jedoch wenig Möglichkeiten bietet, so etwas wie eine „stochastische Intuition" zu entwickeln. Diese scheint jedoch für ein tieferes Verständnis der Zusammenhänge ebenso wichtig zu sein. Der Wissenschaftsautor Leonard Mlodinow bemerkt hierzu in seinem Buch *The Drunkard's Walk – How randomness rules our lives* treffend:

> *The theory of randomness is fundamentally a codification of common sense. But it is also a field of subtlety, a field in which great experts have been famously wrong and expert gamblers infamously correct. What it takes to understand randomness and overcome our misconceptions is both experience and a lot of careful thinking.*[1]

Die Beobachtung mehrerer Hochschullehrer, dass das Fach Stochastik vielen Studierenden schwer fällt, und die Beschreibung hartnäckiger Fehlvorstellungen in der Literatur gaben den Anstoß für diese Arbeit. In didaktischer Literatur wird häufig postuliert, dass Simulationen das Verständnis von Stochastik vertiefen könnten, es existieren aber wenige Studien dazu. Während es im Bereich der Schuldidaktik bereits Untersuchungen zu stärker experimentell orientierten Einführungen in die Stochastik gibt, die auch die Möglichkeiten der neuen Medien, insbesondere Simulationen am Computer, einbeziehen[2], findet man im Bereich der Hochschuldidaktik bisher kaum solche Ansätze, insbesondere solche, die über einzelne Fallstudien hinausgehen[3]. Gerade im Hochschulbereich bietet sich aber der Einsatz von Simulationen am Computer an, um stochastische Phänomene plastisch zu machen, da

[1] Leonard Mlodinow, The Drunkard's Walk. How randomness rules our lives, New York 2008, 21.

[2] Vgl. etwa Thorsten Meyfarth, Die Konzeption, Durchführung und Analyse eines simulationsintensiven Einstiegs in das Kurshalbjahr Stochastik der gymnasialen Oberstufe. Eine explorative Entwicklungsstudie, Dissertation, Universität Kassel 2008, erschienen in: Kasseler Online-Schriften zur Didaktik der Stochastik, Bd. 6.

[3] Eine Ausnahme bilden die Forschungsarbeiten von Rolf Biehler, teilweise in Kooperation mit Carmen Maxara, vgl. hierzu Kap. I.A.2. dieser Arbeit.

Studierende der Mathematik oft schon zu Beginn ihres Studiums erste Erfahrungen mit mathematischer Software machen und so meist bereits über Grundkenntnisse des Programmierens verfügen, die genutzt und vertieft werden können.

I.A. FORSCHUNGSSTAND

I.A.1. Forschungsstand zu Fehlvorstellungen und Schwierigkeiten beim Umgang mit stochastischen Zusammenhängen

Den Grundstein für die Erforschung von Fehlvorstellungen auf dem Gebiet der Stochastik legten der spätere Nobelpreisträger Daniel Kahneman und Amos Tversky. Im Rahmen der Entwicklung ihrer Prospect-Theorie zum Treffen von Entscheidungen in Risikosituationen beschäftigten sie sich in umfangreichen empirischen Studien auch mit kognitiven Verzerrungen. Sie unterscheiden in diesem Zusammenhang zwischen subjektiver und objektiver Wahrscheinlichkeit. Während die objektive Wahrscheinlichkeit jenes Ergebnis beschreibt, das in wahrscheinlichkeitstheoretischer Berechnung richtig wäre, ist die subjektive Wahrscheinlichkeit der Wert, der von einem Individuum in einer bestimmten Situation für die Entscheidung zugrunde gelegt wird. Die beiden Werte sind dabei nicht unbedingt identisch. Die Bezeichnung der subjektiven Wahrscheinlichkeit kommt dem in dieser Arbeit verwendeten Begriff der stochastischen Intuition nahe.[4]

> *We use the term "subjective probability" to denote any estimate of the probability of an event, which is given by a subject, or inferred from his behaviour. These estimates are not assumed to satisfy any axioms or consistency requirements. We use the term "objective probability" to denote values calculated, on the basis of stated assumptions, according to the laws of the probability calculus.*[5]

Als eine grundlegende hartnäckige Fehlvorstellung, die zu falschen intuitiven Einschätzungen von Wahrscheinlichkeiten in unterschiedlichen Kontexten führe, beschreiben sie die Vorstellung, dass eine Stichprobe stets alle wesentlichen Eigenschaften der Population teile und sie in diesem Sinne repräsentativ sei. Die Ursache dafür, dass viele Resultate in Stochastik als kontraintuitiv empfunden würden, sehen sie in dieser Fehlvorstellung begründet.[6] Aus dieser Grund-Fehlvorstellung folgern sie zwei abgeleitete Fehlvorstellungen:

---

[4] Vgl. Kapitel I.A.3. dieser Arbeit.
[5] Daniel Kahneman & Amos Tversky, Subjective Probability: A Judgement of Representativeness, in: Cognitive Psychology 3/3 (1972), 431.
[6] Vgl. Kahneman/Tversky (1972), 435f.

Erstens die Vorstellung, dass das Gesetz der großen Zahlen auch im Kleinen gelte und zweitens, dass die Stichprobengröße keinen Einfluss auf die Genauigkeit von Schätzungen habe, insbesondere nicht auf die Variabilität des Mittelwerts. Diese beiden Fehlvorstellungen sollen im Folgenden beschrieben werden.

Zur Veranschaulichung der Fehlvorstellung von Ausgeglichenheit bzw. „lokaler Repräsentativität", d.h. die Überzeugung, dass das „Gesetz der großen Zahlen" auch im Kleinen gelte, ziehen die Autoren u.a. das Beispiel einer Folge von fairen Münzwürfen heran:

> *People view chance as unpredictable but essentially fair. Thus, […] they expect even short sequences of coin tosses to include about the same number of heads and tails. More generally, a representative sample is one in which the essential characteristics of the parent population are represented not only globally in the entire sample, but also locally in each of its parts. A sample that is locally representative, however, deviates systematically from chance expectations: it contains too many alternations and too few clusters.*
> *The law of large numbers ensures that very large samples are highly representative of the populations from which they are drawn. Elsewhere (Tversky & Kahneman, 1971), we have characterized the expectancy of local representativeness as a belief in the law of small numbers, according to which "the law of large numbers applies to small numbers as well." This belief, we suggest, underlies the erroneous intuitions about randomness, which are manifest in a wide variety of contexts.*[7]

Dass Zufall unvorhersehbar sei, bringen die Autoren mit den Ergebnissen aus anderen Studien[8] in Zusammenhang, dass nämlich Muster wie z.B. das abwechselnde Auftreten von Kopf und Zahl als unwahrscheinlich eingeschätzt würden. Die Fairness zeige sich darin, dass auch in kurzen Sequenzen die Anzahlen von Kopf- und Zahl-Würfen als ausgeglichen erwartet werde.[9]

Auch die bekannte "Gambler's Fallacy", die Annahme, dass sich die Wahrscheinlichkeit für ein gewünschtes Ereignis durch ein vorheriges Eintreten oder Nicht-Eintreten auch dann verringere oder erhöhe, wenn die Ereignisse unabhängig sind, erklären die Autoren mit dieser Fehlvorstellung der „lokalen Repräsentativität". Ihre Erkenntnis reihe sich in die Ergebnisse vielfältiger vorausgegangener Forschung ein, sogar William Feller habe sie schon beschrieben. Die Hartnäckigkeit dieser falschen Einschätzung sehen sie darin bestätigt, dass diese Fehlvorstellung sogar bei erfahrenen Psychologen vorzufinden sei.[10]

---

[7] Kahneman/Tversky (1972), 435.
[8] Z.B. G. S. Tune, Response preferences: A review of some relevant literature, Psychological Bulletin 61 (1964), 286-302 sowie Willam A. Wagenaar, Subjective randomness and the capacity to generate information, in: Acta Pychologica 33 (1970), 233-242, beide zitiert nach: Kahneman/ Tversky (1972), 434.
[9] Vgl. Kahneman/Tversky (1972), 434f.
[10] Vgl. Kahneman/Tversky (1972), 433.

Als eine zweite Fehlvorstellung, die ebenfalls aus der Repräsentationsvorstellung resultiere, beschreiben sie das mangelnde Bewusstsein für den Einfluss der Stichprobengröße in unterschiedlichen Zusammenhängen. So zeigten sie, dass Studierende einerseits nicht in der Lage seien, den Einfluss der Stichprobengröße auf die Verteilungen von binomialverteilten Zufallsvariablen zu erkennen, wenn sie selbst solche Verteilungen skizzieren sollen, andererseits sei ihnen auch nicht bewusst, dass sich die Stichprobengröße auf die Variabilität von zu berechnenden Statistiken, z.B. des Mittelwerts, auswirke.

> *Thus, the event of finding more than 600 boys in a sample of 1000 babies, for example, is as representative as the event of finding more than 60 babies in a sample of 100 babies. The two events, therefore, would be judged equally probable, although the latter, in fact, is vastly more likely.*[11]

Zur Untersuchung der Fehlvorstellung bezüglich der Variabilität des Mittelwerts bei unterschiedlichen Stichprobengrößen verwendeten die beiden Autoren in einer Studie mit Studenten u.a. folgende Aufgabe:

> *A certain town is served by two hospitals. In the larger hospital about 45 babies are born each day, and in the smaller hospital about 15 babies are born each day. As you know, about 50% of all babies are boys. The exact percentage of baby boys, however, varies from day to day. Sometimes it may be higher than 50%, sometimes lower.*
> *For a period of 1 year, each hospital recorded the days on which (more/less) than 60% of the babies born were boys. Which hospital do you think recorded more such days?*[12]

Sie stellten fest, dass die meisten Befragten den Einfluss der Stichprobengröße missachteten und somit keinen Unterschied zwischen den beiden Krankenhäusern erwarteten[13], was sie als Ausdruck der oben beschriebenen Fehlvorstellung deuten. Bemerkenswert ist an dieser Stelle, dass die Autoren in ihrem Paper zwar als Überschrift für die Untersuchungen dieser Fehlvorstellung “Sampling Distributions“, also Stichprobenverteilungen, wählen, in den Ausführungen aber nicht explizit schreiben, dass es ihnen um Aufgaben geht, zu deren Lösung ein Vergleich von Mittelwertsberechnungen bei unterschiedlichen Stichprobengrößen nötig ist. Das veranlasste Peter Sedlmeier und Gerd Gigerenzer zu einer Präzisierung der Fehlvorstellung, wodurch Kahnemans und Tverskys Ergebnisse auch mit scheinbar gegenteiligen Resultaten anderer Autoren vereinbar würden. So seien die meisten Menschen sehr wohl intuitiv in der Lage, die Rolle der Stichprobengröße zu erkennen, allerdings nur so

[11] Kahneman/Tversky (1972), 437.
[12] Kahneman/Tversky (1972), 443.
[13] Vgl. Kahneman/Tversky (1972), 444.

lange, wie es um einzelne Stichproben gehe, d.h. Häufigkeitsverteilungen zugrunde lägen.[14] Anders sehe es aus, wenn Stichproben-/ Kennwerteverteilungen, also Verteilungen von Kenngrößen, in diesem Fall solche des Mittelwerts, beurteilt werden müssten. Die Unterscheidung zwischen Häufigkeitsverteilungen auf der einen und Stichproben-/ Kennwerteverteilungen auf der anderen Seite ist wie folgt zu verstehen: Im Falle von Häufigkeitsverteilungen liegt in der Weise, wie Sedlmeier und Gigerenzer den Begriff verwenden, eine einzelne Stichprobe zugrunde, die auf (relative oder absolute) Häufigkeiten bestimmter Merkmale untersucht wird. Anders als der Begriff vielleicht vermuten lässt, deuten die Ausführungen von Sedlmeier und Gigerenzer darauf hin, dass die entsprechenden Zufallsvariablen auch kontinuierlich sein können. Entscheidend ist, dass der Ausgangspunkt für die Überlegungen auf der Ebene einzelner Stichprobe liegt, die betrachtete Zufallsvariable kann also beispielsweise den (zufälligen) Anteil eines bestimmten Merkmals in der Stichprobe angeben. In den zugehörigen Aufgaben kann dann z.B. verlangt werden, die Situation bei zwei Stichproben unterschiedlicher Größe zu vergleichen. Bei Stichproben-/ Kennwerteverteilungen liegen dagegen von vornherein viele Stichproben zugrunde, wobei jeweils bestimmte Kennwerte (z.B. die Mittelwerte) dieser Stichproben betrachtet werden, die selbst ja auch Zufallsvariablen darstellen. In diesem Fall könnte in den zugehörigen Aufgaben beispielsweise verlangt werden, die Verteilungen der Mittelwerte zu vergleichen, wenn aus derselben Ausgangspopulation Stichproben unterschiedlicher Größe gezogen und jeweils Mittelwerte berechnet werden.[15] In die zweite Kategorie fielen laut Sedlmeier und Gigerenzer die Aufgaben von Kahneman und Tversky, in die erste dagegen Aufgaben etwa aus Studien von Jean Piaget und Bärbel Inhelder[16], aber auch einer Reihe weiterer Autoren, die die Fähigkeit, den Einfluss der Stichprobengröße in den vorgelegten Kontexten richtig zu beurteilen, belegten.[17] Ein Beispiel einer solchen Aufgabe aus der ersten Kategorie, das Gigerenzer und Sedlmeier aus einer Studie von Arnold D. Well, Alexander Pollatsek und Susan J. Boyce aus dem Jahr 1990 zitieren, ist folgendes:

---

[14] Vgl. Peter Sedlmeier & Gerd Gigerenzer, Intuitions About Sample Size: The Empirical Law of Large Numbers, in: Journal of Behavioral Decision Making 10 (1997), 33-51. Bei ihrer Analyse gehen Sedlmeier und Gigerenzer allerdings nicht auf Kahneman und Tverskys Ergebnisse bei der Aufgabe, Verteilungen zu skizzieren, ein. Sie beschränken sich auf Textaufgaben im Stil der „Krankenhausaufgabe".

[15] Vgl. Sedlmeier/ Gigerenzer (1997), 35ff. Der Begriff der Stichproben-/ Kennwerteverteilung ist in der psychologischen Statistik gebräuchlich, vgl. hierzu bspw. Jürgen Bortz, Statistik für Human- und Sozialwissenschaftler, 6. Aufl., Heidelberg 2005, 89f.

[16] Vgl. Jean Piaget & Bärbel Inhelder, The Origin of the Idea of Chance in Children (L. Leake et al. trans.), New York 1975, zitiert nach: Sedlmeier/ Gigerenzer (1997), 34.

[17] Vgl. Sedlmeier/ Gigerenzer (1997), 33ff.

When they turn 18, American males must register for the draft at a local post office. In addition to other information, the height of each male is obtained. The national average height of 18-year-old males is 5 feet, 9 inches.
Yesterday, 25 men registered at a post office A and 100 men registered at a post office B. At the end of the day, a clerk at each post office computed and recorded the average height of the men who registered there that day. Which would you expect to be true? (circle one)

1. *The average height at post office A was closer to the national average than was the average height at post office B.*
2. *The average height at post office B was closer to the national average than was the average height at post office A.*
3. *There is no reason to think that the average height was closer to the national average at one post office than at the other.*[18]

Der wesentliche Unterschied zwischen der „Post-Aufgabe“ und der „Krankenhaus-Aufgabe“ liege darin, dass im einen Fall die Situation an einem einzigen Tag, also auf Grundlage jeweils einer Stichprobe und somit im Kontext von Häufigkeitsverteilungen, beurteilt werden müsse, im anderem Fall dagegen auf der Grundlage von jeweils 365 Stichproben und folglich im Kontext von Stichprobenverteilungen.

Als Ursache dafür, dass im Kontext von Häufigkeitsverteilungen der Einfluss der Stichprobengröße auf die Genauigkeit einer Schätzung, etwa des Mittelwerts, vielen Personen bewusst sei, postulieren Sedlmeier und Gigerenzer die alltägliche Erfahrung im Umgang mit solchen Schätzungen – die Autoren sprechen in Anlehnung an Hans Freudenthal auch von dem „Empirischen Gesetz der großen Zahlen“. Die Einsicht, dass eine größere Stichprobe die Genauigkeit im Allgemeinen erhöhe, spiegele sich auch in dem kulturübergreifenden Phänomen der Wertschätzung von Lebenserfahrung wider:

For instance, many cultures value the knowledge of older men and women, presumably because they can draw on a larger sample of observations than younger people in making predictions about the behaviour of nature and humans. So the intuition that estimates and predictions based on larger samples tend to be more accurate might just be the cumulative result of millennia of experience.[19]

Im Gegensatz dazu seien im Alltag die Gelegenheiten, sich mit Eigenschaften von Stichproben-/ Kennwerteverteilungen auseinanderzusetzen, sehr begrenzt. Auf diese mangelnde Erfahrung führen die Autoren dann auch die schlechtere Intuition zurück:

[18] Arnold D. Well, Alexander Pollatsek & Susan J. Boyce, Understanding the effects of sample size on the variability of the mean, in: Organizational Behavior and Human Decision Processes 47/2 (1990), 297, zitiert nach: Sedlmeier/ Gigerenzer (1997), 37f.
[19] Sedlmeier/ Gigerenzer (1997), 46.

Why is the role of sample size in sampling distributions so hard to grasp? One consideration is that frequency distributions are involved in everyday problems of estimation and prediction, whereas the rule that the variability of a sampling distribution decreases with increasing sample size seems to have only a few applications in ordinary life. In general, taking repeated samples and looking at the distribution of their means is rare in the everyday and only recent in scientific practice. [...] Thus with the exception of statisticians and their kin, humans may have experienced little selective pressure to develop intuitions about the impact of sample size on the variance of sampling distributions.[20]

Anzumerken ist an dieser Stelle, dass die Unterscheidung zwischen den Kontexten der Häufigkeits- und Stichproben-/ Kennwerteverteilungen nicht so eindeutig ist, wie von Sedlmeier und Gigerenzer dargestellt. Im Fall der „Krankenhaus-Aufgabe" würde beispielsweise aufgrund der Tatsache, dass die Geschlechterverhältnisse an allen Tagen unabhängig und identisch verteilt sind, auch die Betrachtung eines einzelnen Tages für die Beantwortung der Fragestellung reichen, was wiederum zur Begutachtung auf der Ebene von Häufigkeits- statt auf Stichproben-/ Kennwerteverteilungen führen würde. Für den Gegenstand dieser Arbeit sei bemerkt, dass ebenso die im Zuge von Simulationen auftretende Frage nach einer günstigen Anzahl von Wiederholungen zur Reduzierung von Unsicherheiten bei der Schätzung von Parametern je nach Betrachtungsweise sowohl als Einschätzung von Variabilität im Rahmen von Häufigkeits- als auch von Stichproben- /Kennwerteverteilungen aufgefasst werden kann. Daher wird im Folgenden nur dann eine Unterscheidung zwischen den beiden Kontexten vorgenommen, wenn es eine eindeutige Zuordnung gibt.
An dieser Stelle sei auch noch bemerkt, dass mit den vorgestellten Aufgaben zur Diagnose der untersuchten Fehlvorstellungen keine Aussage getroffen werden kann, ob die Personen, die die Fehlvorstellungen nicht teilen, eine richtige Vorstellung von der Größenordnung des Einflusses der Stichprobengröße besitzen.[21]

Neben den bereits erwähnten Fehlvorstellungen im Gebiet der Stochastik werden in der Literatur noch weitere beschrieben, die aber weniger detailliert untersucht worden sind. Zwei sehr grundlegende Fehlvorstellungen beziehen sich auf den Begriff „Zufall". Es sind die Gleichsetzungen von Zufall mit Unabhängigkeit sowie von Zufall mit uniformen

[20] Sedlmeier/ Gigerenzer (1997), 46f.
[21] So kann etwa zu der sich anschließenden Frage nach dem Verständnis und der intuitiven Anwendung des „$\sqrt{n}$-Gesetzes" z.B. bei der Abschätzung von Standardfehlern, also der Standardabweichungen von Mittelwerten, keine Aussage getroffen werden, vgl. hierzu bspw. Götz Kersting & Anton Wakolbinger, Elementare Stochastik, in Reihe: Mathematik kompakt, hrsg von Martin Brokate et. al., Basel 2008, 70.

Verteilungen.[22] Bemerkenswerterweise wird trotz des Auftretens von solchen Fehlvorstellungen der Begriff des Zufalls in kaum einem Lehrbuch zur Stochastik explizit definiert, sondern das Verständnis seiner Bedeutung vorausgesetzt. Eine Ausnahme bildet ein altes Buch von Emanuel Czuber aus dem Jahr 1908, in dem es heißt:

> *„Wenn wir den Erfolg oder das Ereignis [...] als ein zufälliges bezeichnen, so negieren wir damit die Notwendigkeit seines Eintretens […].“*[23]

Diese Vorstellung von Zufall als dem Gegenteil von Notwendigkeit beschreibt das Schlüsselkonzept der Variabilität, das für das Verständnis stochastischer Zusammenhänge maßgeblich ist.[24]

Die sich anschließende Frage, ob sich Fehlvorstellungen korrigieren lassen, wird in der Literatur sehr unterschiedlich beantwortet. Kahneman und Tversky schließen aus ihren Untersuchungen, in denen sich die Fehlvorstellung bezüglich der Stichprobengröße sowohl bei Schülerinnen und Schülern sowie Studierenden als auch bei erfahrenen Psychologinnen und Psychologen nachweisen ließ[25], dass sich die Fehlvorstellung weder durch theoretische Einsicht noch durch statistische Erfahrung ausräumen lasse:

> *The notion that sampling variance decreases in proportion to sample size is apparently not part of man's repertoire of intuitions. […]*
> *We surely do not mean to imply that man is incapable of appreciating the impact of sample size on sampling variance. People can be taught the correct rule, perhaps even with little difficulty. The point remains that people do not follow the correct rule, when left to their own devices. Furthermore, the study of the conduct of research psychologists (Cohen 1962, Tversky & Kahneman 1971) reveals that a strong tendency to underestimate the impact of sample size lingers on despite knowledge of the correct rule and extensive statistical training. For anyone who would wish to view man as a reasonable intuitive statistician, such results are discouraging.*[26]

In diesem Punkt widersprechen Kahneman und Tversky dem zuvor zitierten Feller. Für letzteren ist stochastische Intuition sehr wohl formbar, und zwar in seinen Augen durch die

---

[22] Vgl. z.B. Hermann Dinges, Variables, in particular random variables, in: Activity and Sign. Grounding Mathematics Education, hrsg. von Michael H.G. Hoffmann, Johannes Lenhard und Falk Seeger, New York 2005, 308.

[23] Emanuel Czuber, Wahrscheinlichkeitsrechnung und ihre Anwendung auf Fehlerausgleichung, Statistik und Lebensversicherung, 1. Bd., Leipzig und Berlin 1908, 7f.

[24] Woher diese Variabilität rührt, ob es also „objektive“ Zufälligkeit in der Welt gibt oder den Menschen aufgrund mangelnder Fülle an Information bestimmte Sachverhalte nur zufällig erscheinen, ist Gegenstand philosophischer Diskussion. Ein Überblick über die erkenntnistheoretischen Positionen und Argumente hierzu findet sich z.B. in Thierry Martin, Wahrscheinlichkeit – ein mehrdeutiger Begriff, in: Zufall und Chaos, Spektrum der Wissenschaft Spezial 1 (2010), 12-16.

[25] Vgl. Amos Tversky & Daniel Kahneman, Belief in the Law of Small Numbers, Psychological Bulletin 76/2 (1971), 105-110.

[26] Kahneman/Tversky (1972), 444f.

Beschäftigung mit stochastischer Theorie und ihren Anwendungen, wie er in seinem 1950 erstmals erschienen Werk „An Introduction to Probability Theory and Its Applications“ ausführt:

> *[...] the modern student has no appreciation of the modes of thinking, the prejudices, and other difficulties against which the theory of probability had to struggle when it was new. Nowadays newspapers report on public opinion, and the magic of statistic embraces all phases of life to the extent that young girls watch the statistics of their chances to get married. Thus everyone has acquired a feeling for the meaning of statements such as "the chances are three in five". Vague as it is, this intuition serves as background and guide for the first step. It will be developed as the theory progresses and acquaintance is made with more sophisticated applications.*[27]

Auch Sedlmeier und Gigerenzer finden im Kontext ihrer Metaanalysen Hinweise darauf, dass sich die Fehlvorstellung korrigieren lassen könnte. Anders als Feller vermutet, zeigten sich in den Studien allerdings nur bei Personen, die „hands-on“ Erfahrungen mit den in den Aufgaben präsentierten Problemstellungen machen konnten (in diesem Fall ging es um das Galton-Brett[28]), eine Verbesserung ihrer Intuition, jedoch nur im Kontext von Häufigkeitsverteilungen, also bei dem Aufgabentyp, den Kahneman und Tversky gar nicht untersuchten, und bei dem die Intuition von vornherein in den meisten Fällen zutreffend zu sein scheint.[29]

Dass Fehlvorstellungen hartnäckig und insbesondere durch einen theoretischen Zugang kaum auszuräumen seien, sehen auch die Autoren einer Überblicksstudie aus dem Jahr 2008 zum Forschungsstand der Didaktik von Einführungskursen in Statistik an Universitäten bestätigt. Andrew Zieffler et al. fassen in einer groß angelegten Überblicksarbeit zum Forschungsstand auf dem Gebiet des Lernens und Lehrens von (elementarer) Statistik wesentliche Ergebnisse unterschiedlicher qualitativer und quantitativer Untersuchungen zusammen. Bemerkenswert sei an dieser Stelle insbesondere, dass auch Studierende mit guten Ergebnissen in den Kursen kaum tragfähige Vorstellungen von stochastischen Konzepten wie dem Mittelwert, der Standardabweichung oder dem Zentralen Grenzwertsatz vorweisen konnten. Beispielsweise

[27] William Feller, An Introduction to Probability Theory and Its Applications, Vol. I, Third Edition, New York, London, Sydney 1968, 2.
[28] Vgl. Christopher J. Jones & Paul L. Harris, Insight into the law of large numbers: A comparison of Piagetian and judgement theory, Quarterly Journal of Experimental Psychology 34A (1982), 479-488, zitiert nach: Sedlmeier/ Gigerenzer (1997), 39.
[29] Vgl. Sedlmeier/ Gigerenzer (1997), 39.

konnte der Mittelwert von vielen Studierenden ausschließlich mit dem zugrundeliegenden Kalkül verknüpft werden[30].

> *Studies that focus on assessing students' statistical reasoning have typically revealed that even after formal instruction, many people remain unable to reason correctly about important statistical ideas and concepts [...].[31]*
>
> *Most students were found to have relatively unsophisticated understandings of the concepts of mean and standard deviation and only fragmentary recall, at best, of the Central Limit Theorem. These studies suggest that even high performing students may not be able to reason about even basic statistical concepts, such as the mean and standard deviation, much less about abstract concepts, such as the Central Limit Theorem.[32]*

Die Autoren folgern aus diesen Resultaten, dass intuitive Zugänge stärker in den Vordergrund gerückt werden sollten.[33] Auch wenn die didaktische Wahl der Zugangsweise ohne Zweifel den Lernerfolg der Studierenden beeinflusst, stellen Zieffler et al. in ihrer Überblicksarbeit fest, dass die Schwierigkeiten im Umgang mit Konzepten der Statistik bedeutsam und hartnäckig blieben, und bei vielen Studierenden nach Einführungsveranstaltungen selbst elementare Konzepte des Faches kaum vorhanden seien.[34]

In dieser Arbeit soll – basierend auf den oben ausgeführten Forschungsergebnissen – ein besonderer Fokus auf den folgenden (bereits verhältnismäßig gut erforschten) Fehlvorstellungen liegen, deren Veränderlichkeit durch den Einsatz von Simulationen untersucht werden soll:

(F1) Zufälligkeit ließe sich durch lokale Ausgeglichenheit charakterisieren, d.h. das „Gesetz der großen Zahlen" wäre auch im Kleinen gültig. Beispiele sind die Unterschätzung der erwarteten Länge von Runs in Wurfserien mit fairer Münze und die „Gambler's Fallacy".

(F2) Die Genauigkeit von Schätzungen bzw. die Variabilität von Schätzwerten hinge nicht von der Größe der zugrunde liegenden Stichproben ab,

(a) auf der Ebene von Häufigkeitsverteilungen z.B. durch Ignorieren der Anzahlen von Beobachtungen, die in Mittelwertsberechnungen eingehen,

[30] Vgl. Andrew Zieffler, Joan Garfield, Shirley Alt, Danielle Dupuis, Kristine Holleque & Beng Chang , What Does Research Suggest About the Teaching and Learning of Introductory Statistics at the College Level? A Review of the Literature, Journal of Statistics Education 16/2 (2008), 4. Zitiert nach: DOI (pdf), abrufbar unter: https://doi.org/10.1080/10691898.2008.11889566.
[31] Zieffler et al. (2008), 5.
[32] Zieffler et al. (2008), 6.
[33] Vgl. Zieffler et al. (2008), 6.
[34] Zieffler et al. (2008), 14.

(b) auf der Ebene von Stichproben- / Kennwerteverteilungen z.B. durch fehlende Beachtung von Standardfehlern bzw. noch allgemeiner der völligen Vernachlässigung der Variabilität von Kennwerten.

Außer Fehlvorstellungen scheinen auch nicht-kognitive Faktoren das Erlernen von stochastischen Zusammenhängen bzw. einen verständigen Umgang mit ihnen zu erschweren, nämlich insbesondere Ängstlichkeit und negative Einstellungen dem Fach gegenüber. Der Einfluss dieser Faktoren wurde im Gebiet der Statistik erforscht, wie Zieffler et al. darstellen. Dabei stellte sich nicht nur heraus, dass beispielsweise die Einschätzung der Schwierigkeit von Aufgaben oder auch der Nützlichkeit des Studiums dieses Faches die Leistungsfähigkeit beeinflusste, sondern auch, dass die Einstellungen sich im Laufe von Statistikkursen kaum veränderten. Inwiefern etwa Ängstlichkeit im Kontext von Statistikaufgaben mit einer allgemeinen Prüfungsangst korrelliert, wurde dabei nicht untersucht, was die Belastbarkeit der Ergebnisse verringert.[35] Nichtsdestotrotz sollten auch nicht-kognitive Einflussgrößen im Blick behalten werden, wenn Schwierigkeiten im Umgang mit stochastischen Fragestellungen untersucht werden.

I.A.2. Forschungsstand zum Einsatz von Simulationen bei der Lehre des Fachs Stochastik

Fast selbstverständlich wird in didaktischer Literatur häufig angenommen, dass Simulationen ein geeignetes Mittel seien, das Verständnis von Stochastik zu vertiefen und Fehlvorstellungen entgegen zu wirken. So werden beispielsweise in einem der Standardwerke zur Didaktik, „Mathematikunterricht in der Sekundarstufe II – Didaktik der Stochastik“ von Uwe-Peter Tietze, Manfred Klika und Hans Wolpers herausgegeben, diesbezügliche Thesen führender Didaktiker so formuliert:

- *Die in der realen Welt vorhandenen Zufallsgeneratoren wie Münze, Würfel, Glücksrad, Zufallszahlen usw. sollten zum Aufbau reichhaltiger und beziehungshaltiger Erfahrungen mit stochastischen Situationen genutzt werden.*
- *Insbesondere Simulationen dienen dem Aufbau solcher Erfahrungen und der Entwicklung eines sicheren intuitiven Fundaments für die zentralen Begriffe und Verfahren der Stochastik.*[36]

[35] Vgl. Zieffler et al. (2008), 7ff.
[36] Hans Wolpers, Didaktik der Stochstik, in Reihe: Mathematikunterricht in der Sekundarstufe II, hrsg. von Uwe-Peter Tietze, Manfred Klika, Hans Wolpers, Bd. 3, Braunschweig/ Wiesbaden 2002, 102.

Noch prägnanter wird an späterer Stelle resümiert:

Ganz wesentlich für den Lernerfolg sind eigene experimentelle Erfahrungen.[37]

Aus den Zitaten lässt sich ableiten, dass unter Simulationen nicht zwangsläufig computerbasierte Experimente verstanden werden. Wolpers erklärt den Begriff folgendermaßen:

> *Unter Simulation versteht man in der Stochastik Verfahren, mit Hilfe von geeigneten Zufallsgeneratoren eine stochastische Situation `nachzuspielen´, um so ein Modell für diese Situation zu erhalten, das dann zur weiteren Analyse und zur Prognose eingesetzt werden kann.*[38]

Dass Simulationen am Computer aufgrund ihrer spezifischen Mehrdimensionalität dabei einen eigenständigen Typus mathematischer Modellierung bilden, arbeitet Johannes Lenhard aus wissenschaftsphilosophischer Perspektive heraus.[39] In dieser Arbeit soll im Folgenden der Fokus auf computerbasierten Simulationen liegen und der Forschungsstand bezüglich des Einsatzes solcher Experimente in der Lehre der Stochastik analysiert werden.

Joachim Engel und Rolf Biehler beschreiben im „Handbuch der Mathematikdidaktik" den „Common-sense" in der Mathematikdidaktik zu Simulationen folgendermaßen:

> *Auch wenn manche Details zum effizienten Einsatz von Simulationen in der Lehre von Stochastik noch empirisch weiter erforscht werden müssen (Mills 2002), besteht unter Mathematikdidaktikern und Kognitionspsychologen ein breiter Konsensus dahingehend, dass Simulationen herausragende Vorzüge bieten, um bei Lernenden das Verstehen abstrakter Konzepte der Stochastik verbessern (Biehler 1991; Burill 2002; Maxara und Biehler 2006; Mills 2002; Zieffler und Garfield 2007).*[40]

---

[37] Wolpers (2002), 111.

[38] Wolpers (2002), 129. Kütting und Sauer liefern eine ähnliche Beschreibung und nennen Beispiele, u.a. auch für ein Monte Carlo Verfahren zur Schätzung des Wertes der Zahl π. Vgl. Kütting/ Sauer (2014), 213ff.

[39] Vgl. Johannes Lenhard, Mit allem rechnen – zur Philosophie der Computersimulation, in Reihe: Ideen&Argumente hrsg. von Wilfried Hinsch & Lutz Wingert, Berlin/Boston 2015, dort insbes. 112ff.

[40] Rolf Biehler & Joachim Engel, Stochastik: Leitidee Daten und Zufall, in: Handbuch der Mathematikdidaktik, hrsg. von Regina Bruder, Lisa Hefendehl-Hebeker, Barbara Schmidt-Thieme & Hans-Georg Weigand, Berlin, Heidelberg 2015, 243. Eine identische Einschätzung findet sich in Joachim Engel & Rudolf Grübel, Bootstrap – oder die Kunst, sich selbst aus dem Sumpf zu ziehen, Mathematische Semesterberichte 55 (2008), zitiert nach Preprint: https://www.researchgate.net/publication /225158275_Bootstrap__oder_die_Kunst_sich_selbst_aus_dem_Sumpf_zu_ziehen/link/0912f5125f57c4da6c000000/download, 4 (zuletzt aufgerufen am 02.01.2020). Untersuchungen aller Didaktiker, auf die in dem Zitat verwiesen wird, werden in diesem Kapitel genauer vorgestellt, außer solche von Gail Burrill, die sich explizit auf schulischen Stochastikunterricht beziehen. Er schlägt u.a. vor, besonderen Wert auf das Verständnis von Stichprobenverteilungen als Vorbereitung für das Konzept von Hypothesentests zu legen und dabei mit Simulationsergebnissen zu arbeiten, vgl. bspw. Gail Burrill, Simulation as a tool to develop statistical understanding, ICOTS-6 (2002), abrufbar unter: https://iase-web.org/Conference_Proceedings.php?p=ICOTS_6_2002 (zuletzt aufgerufen am 05.01.2020).

Tragfähige Belege für diese Einschätzung zu finden, ist allerdings schwierig. Insbesondere für den Bereich der Vermittlung von Konzepten aus der Wahrscheinlichkeitstheorie gibt es äußerst wenige Forschungsarbeiten dazu, die meisten Studien konzentrieren sich auf den Einsatz von Simulationen in Statistik-Kursen, in denen also Methoden für die Analyse von Datensätzen gelehrt werden. Aber auch die Ergebnisse dieser Studien sind häufig schwer verallgemeinerbar, da sie sich auf sehr spezifische Anwendungssituationen beziehen.[41] Einen breiten Überblick über den Forschungsstand hierzu bilden die Arbeiten von Jamie D. Mills aus dem Jahr 2002[42] und die bereits oben zitierte von Zieffler et. al aus dem Jahr 2008. Viele der dort beschriebenen Arbeiten befassen sich nicht mit Konzepten, die den Einsatz von Simulationen während eines ganzen Kurses zur Statistik vorsehen, sondern nur für die Vermittlung einzelner Inhalte bzw. für einzelne Fallstudien zu bestimmten Themen.[43] Zu der Frage, ob Simulationen für das Verständnis statistischer Zusammenhänge hilfreich sind, finden sich bei Zieffler et al. teilweise inkonsistente Ergebnisse[44] und Mills kommt zu dem Schluss, dass die in der Literatur häufig anzutreffende Empfehlung für den Einsatz von Simulationen (die sie selbst auch teilt) insgesamt viel zu schwach empirisch abgesichert sei, was sie sowohl auf praktische Schwierigkeiten bei der Durchführung solcher Studien als auch den Arbeitsaufwand zurückführt[45].

Die Situation hat sich diesbezüglich kaum geändert. Eine Ausnahme bilden die Forschungsarbeiten von Rolf Biehler und seinen Mitarbeitenden, die sich allerdings im Bereich der Hochschuldidaktik auf Veranstaltungen für Lehramtsstudierende (Sekundarstufe I) beschränken.[46] Sie setzen die für didaktische Zwecke entwickelte Software Fathom in Veranstaltungen zur Elementaren Stochastik für Lehramtsstudierende ein. Die Aufgaben

---

[41] Vgl. Zieffler et al. (2008), 13.

[42] Vgl. Jamie D. Mills, Using Computer Simulation Methods to Teach Statistics: A Review of the Literature, Journal of Statistics Education 10/1 (2002). Zitiert nach: DOI (pdf), abrufbar unter: https://doi.org/10.1080/10691898.2002.11910548.

[43] Vgl. z.B. zum Thema Stichprobenverteilungen Beth Chance, Robert delMas & Joan Garfield, Reasoning About Sampling Distributions, in: The Challenge of Developing Statistical Literacy, Reasoning and Thinking, hrsg. von Dani Ben-Zvi & Joan Garfield, Dordrecht, The Netherlands 2004, 295-323, zitiert nach: Zieffler et al. (2008), 11, und zum Thema Standardabweichung Robert delMas & Yan Liu, Exploring Students' Conceptions of the Standard Deviation, Statistics Education Research Journal [Online], 4/1(2005), 55-82, zitiert nach: Zieffler et al. (2008), 11. Außerdem: Mills (2002), 4ff.

[44] Vgl. Zieffler et. al (2008), 10f.

[45] Vgl. Mills (2002), 13f.

[46] Ein weiterer Schwerpunkt der Arbeitsgruppe sind zahlreiche Forschungsarbeiten zum Einsatz von Simulationen am Computer in der Schule, z.B. Rolf Biehler, Working styles and obstacles: Computer-supported collaborative learning in statistics, ICOTS-7 (2006), abrufbar unter: https://iase-web.org/Conference_Proceedings.php?p=ICOTS_7_2006 (zuletzt aufgerufen am 04.01.2020).
Eine weitere Ausnahme bildet eine Studie von Zieffler und Garfield aus dem Jahr 2007, die sich allerdings nicht auf Mathematikstudierende bezieht und zudem ausschließlich mit für Lehrzwecke entwickelter Software bzw. Web-Applets arbeitet, vgl. Andrew Zieffler & Joan B. Garfield, Studying the Role of Simulation in Developing Students' Statistical Reasoning, Bulletin of the International Statistical Institute 56th Session (2007).

zielen dabei explizit auf Simulationen von stochastischen Zusammenhängen auf elementarem fachlichem Niveau (z.B. Simulation von Münzwürfen oder des Geburtstagsproblems mit Hilfe eines Urnenmodells)[47]. Sowohl bei ihren Konzepten für den Einsatz von Simulationen in der Schule als auch bei denen für den Einsatz an der Universität arbeiten die Autoren mit einem eng gefassten Simulationsplan, in dem die Schülerinnen und Schüler bzw. die Studierenden ihre Simulationen mit Fathom planen, dokumentieren und auswerten sollen. Das Ziel dessen ist, dass sich die Schülerinnen und Schüler bzw. die Studierenden nicht in technischen Details verlieren, sondern stets das Ziel ihrer Simulation im Blick behalten und strukturiert vorgehen.[48] Dieser recht basale Zugang zum Einsatz von Simulationen wird in diesem Fall auch deshalb gewählt, weil sich in Voruntersuchungen von Biehler und Maxara gezeigt hat, dass die Studierenden, die die Zielgruppe der vorgestellten Kurse bilden, Experimente bzw. Simulationen gar nicht als methodische Option zur Lösung von stochastischen Problemsituationen in Betracht ziehen.[49] In sich anschließenden Studien wurde untersucht, wie sich der Einsatz der Software auf die Entwicklung stochastischer und technischer Kompetenzen auswirkt. Dass technische Simulationskompetenzen und stochastische Kompetenzen nicht unbedingt zusammenhängen, deutet sich z.B. in einer explorativen Fallstudie von Carmen Maxara an, bei der in einer Vorlesung zur Elementaren Stochastik begleitende Simulationsaufgaben mit der Software Fathom gestellt wurden und anschließend die Kompetenzen in beiden Bereichen bei acht ausgewählten Studierenden untersucht wurden.[50] Bezüglich der stochastischen Kompetenzen stellte sich heraus, dass Fehlvorstellungen hartnäckiger waren als erwartet und sich selbst durch unmittelbare Simulationen der betreffenden Zusammenhänge kaum ausräumen ließen. Die Autoren

[47] Vgl. Carmen Maxara & Rolf Biehler, Students' probabilistic simulation and modeling competence after a computer-intensive elementary course in statistics and probability, ICOTS-7 (2006), abrufbar unter: https://iase-web.org/Conference_Proceedings.php?p=ICOTS_7_2006 (zuletzt aufgerufen am 04.01.2020).

[48] Vgl. Rolf Biehler & Andreas Prömmel, Developing students' computer-supported simulation and modelling competencies by means of carefully designed working environments, Proceedings of ICOTS 8 (2010), abrufbar unter: https://www.researchgate.net/publication/257927054_Developing_Students'_Computer-Supported_Simulation_and_Modelling_Competencies_by_Means_of_Carefully_Designed_Working_Enviroments (zuletzt aufgerufen am 04.01.2020) sowie Carmen Maxara & Rolf Biehler, Constructing stochastic simulations with a computer tool – students' competencies and difficulties, Proceedings of the Fifth Congress of the European Society for Research in Mathematics Education (2007), 762-771, abrufbar unter: https://www.researchgate.net/publication/257927049_Constructing_stochastic_simulations_with_a_computer_tool_-_students'_competencies_and_difficulties (zuletzt aufgerufen am 04.01.2020).

[49] Vgl. Rolf Biehler & Carmen Maxara, Eingangstest Stochastik – Vorkenntnisse von Lehramtsstudierenden, Beiträge zum Mathematikunterricht 2005, Hildesheim, 91-94, abrufbar unter: https://eldorado.tu-dortmund.de/handle/2003/30643 (zuletzt aufgerufen am 04.01.2020).

[50] Vgl. Carmen Maxara, Simulationskompetenzen und stochastische Kompetenzen – Ergebnisse einer explorativen Fallstudie, Vortrag auf der 43. Jahrestagung der GDM in Oldenburg (2009), abrufbar unter: https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=2ahUKEwjxwvK01unmAhXLzoUKHe5iCmEQFjAAegQIAxAC&url=https%3A%2F%2Fcore.ac.uk%2Fdownload%2Fpdf%2F46913768.pdf&usg=AOvVaw3Oj_jQX1ETG7qctFu6uYb- (zuletzt aufgerufen am 04.01.2020).

schlagen deshalb vor, in den Aufgabenstellungen eine explizite Auseinandersetzung mit den Fehlvorstellungen, im Sinne der Formulierung von Hypothesen und einem Abgleich mit den Ergebnissen, zu verlangen.[51] Aus ihren Untersuchungen im Rahmen der Lehre an der Universität und auch der Schule schließen die Autoren um Biehler zudem, dass im Kontext von Fehlvorstellungen insbesondere die Variabilität von Schätzungen in Abhängigkeit von der Stichprobengröße bearbeitet werden sollte, explizit beziehen sie sich an dieser Stelle auf das „$\sqrt{n}$– Gesetz" etwa beim Standardfehler. Sie schlagen vor, hier theoretische und simulierte Ergebnisse stärker zu vernetzen.[52]

Die Simulationsaufgaben, die in den Studien verwendet werden, dienten einerseits der Generierung von Hypothesen zu stochastischen Zusammenhängen, andererseits auch zur Überprüfung theoretischer Lösungen.[53] Zu den Anwendungsfeldern von Simulationen schreiben Maxara und Biehler:

> *Ideally the use of simulation can serve two different pedagogical purposes (Biehler, 1991). Students are to develop "modeling competence" related to elementary probability, and simulation competence can be part of it. Instead of setting up a formal model and determining unknown probabilities mathematically they can set up a model, implement it in a computer and develop estimates of unknown probabilities by means of simulation. In this sense simulation replaces mathematics. A different purpose is to use simulation for making probabilistic situations more experiential. The observed phenomenon can be used as a reference for concept development, for elaborating probabilistic intuitions, and for overcoming misconceptions. In this sense simulation replaces the real world, which is to be modeled.*[54]

Zur ersten Verwendungsart von Simulationen, dem Ersatz für theoretische Lösungen, bemerken die Autoren an anderer Stelle, dass diese insbesondere in Situationen nützlich sei, in denen Studierende keine theoretischen Lösungsstrategien zur Verfügung hätten. Theoretische und simulierte Lösungswege könnten aber auch kombiniert werden. Die zweite Verwendungsart von Simulationen als Mittel zur Sammlung von Erfahrungen mit stochastischen Konzepten könne sowohl der Begriffsbildung als auch einem Ausbau von Intuitionen dienen.[55] Dass Simulationen noch grundsätzlicher für einen konstruktivistischen Lernansatz des entdeckenden Lernens[56] hilfreich sein könnten, führt Michael Wood in einer Forschungsarbeit zum Einsatz von Simulationen in Statistik-Kursen aus. Gerade Studierende

---

[51] Vgl. Biehler/ Maxara (2006), 5f.
[52] Biehler/ Prömmel (2010), 2 und Maxara/ Biehler (2007), 770f.
[53] Vgl. Maxara/ Biehler (2006), 5.
[54] Maxara/ Biehler (2006), 1.
[55] Vgl. zu den beschriebenen Verwendungsarten von Simulationen auch Maxara/ Biehler (2007), 762.
[56] Der Begriff des entdeckenden Lernens im Mathematikunterricht wurde von Heinrich Winter geprägt.

mit wenig fachlichem Hintergrund könnten mit Hilfe von Simulationen Einsichten in statistische Zusammenhänge gewinnen, ohne sich zuerst in die theoretischen Hintergründe einarbeiten zu müssen.[57]

Simulationen sind aber nicht nur ein didaktisches Mittel, sondern ein mächtiges und vielgenutztes Werkzeug der Statistik. Joachim Engel und Rudolf Grübel beleuchten am Beispiel von Bootstrap-Verfahren ihre Bedeutung und weisen darauf hin, dass durch Simulationen Approximationen für Problemstellungen möglich werden, die analytisch kaum zugänglich oder sehr unübersichtlich wären. Durch computer-basierte Verfahren rückten auch nicht-parametrische Verfahren stärker in den Fokus, die weniger oder gar keine Annahmen über die Daten voraussetzten. Durch den Einsatz von Simulationen werde also auch die Praxis moderner Statistik abgebildet.[58] Dabei übersehen Engel und Grübel auch den didaktischen Nutzen nicht und schließen aus einer breiten Analyse der hierzu getätigten Forschungsarbeiten, dass Simulationen zur Auseinandersetzung mit Fehlvorstellungen und dem Aufbau tragfähiger stochastischer Konzepte geeignet zu sein scheinen:

> *In der Stochastik sind Simulationstechniken zugleich ein mächtiges Werkzeug zum Problemlösen wie auch ein herausragendes Medium der Visualisierung und Unterstützung von Lernprozessen. Mathematisch basierend auf dem Gesetz der großen Zahl und einem frequentistischen Wahrscheinlichkeitsbegriff lassen sich Wahrscheinlichkeiten durch Simulationen annäherungsweise bestimmen. Simulationen können auch genutzt werden, um das Verstehen von Zufallsprozessen zu fördern, indem sie verdeutlichen, dass Zufallsvariable im Einzelergebnis unvorhergesehene Ergebnisse produzieren, langfristig aber vorhersehbare Muster erzeugen. Für den Lernprozess bedeutsam erlauben Simulationen einen experimentellen Arbeitsstil, der neue Zusammenhänge entdecken lässt, indem Lernende neue Szenarien unter der Leitfrage „Was wäre wenn…“ untersuchen. Bei Simulationen ersetzen wir eine reale Situation durch ein Experiment, das ein Modell des Originals ist, das aber leicht manipuliert und analysiert werden kann. Da das Experiment mittels Zufallsgeneratoren am Computer durchgeführt wird, lassen sich ohne nennenswerten Aufwand Replikationen in hoher Zahl durchführen sowie Auswirkungen von alternativen Modellannahmen und Parameterfestlegungen untersuchen. Simulationen unterstützen valide Vorstellungen bezüglich Zufall und Wahrscheinlichkeit und konfrontieren Fehlvorstellungen. Da allerdings Simulationen lediglich eine (angenäherte) Problemlösung und nicht unmittelbare Gründe für deren Richtigkeit liefern, haben Simulationen per se keine explanative Kraft. Sie erlauben aber einen entdeckenden Arbeitsstil, indem Lernende selbst Daten produzieren und analysieren und mit Zufallsstichproben einer Population experimentieren, deren Parameter bekannt sind.*[59]

[57] Vgl. Michael Wood, The Role of Simulation Approaches in Statistics, Journal of Statistics Education 13/3 (2005), zitiert nach: DOI (pdf), abrufbar unter: https://doi.org/10.1080/10691898.2005.11910562.
[58]Engel/Grübel (2008), 1ff.
[59] Engel/ Grübel (2008), 4.

Basierend auf den beschriebenen Überlegungen wird in dieser Arbeit der Einsatz von Simulationen in drei Anwendungsfeldern unterschieden:

(A1) als Hilfsmittel zur Überprüfung von theoretischen Ergebnissen oder Verfahrensweisen,
(A2) als Hilfsmittel zur Exploration und Generierung neuer Hypothesen sowie
(A3) als Hilfsmittel zur Annäherung nicht oder nur sehr schwer berechenbarer Größen.[60]

I.A.3. Überlegungen zu den Begriffen Intuition und Verständnis

In dieser Arbeit soll der Begriff der (stochastischen) Intuition von seiner ursprünglichen Wortherkunft, abgeleitet vom mittellateinischen Wort intuitio, als „unmittelbare Anschauung"[61] verstanden werden bzw. in Anlehnung an Wolpers als *„unmittelbares Wissen […], das Verstehensprozesse und Handlungen steuert"*[62]. Mit „Unmittelbarkeit" soll in diesem Zusammenhang ein Ausbleiben bewusster Modellierungsschritte, theoretischer Schlussfolgerungen bzw. (stochastischer) Kalkülanwendungen gemeint sein. Für diesen Kontext bietet es sich an, ein von Herbert Kütting und Martin J. Sauer vorgeschlagenes Modell zur Veranschaulichung der Herangehensweise an stochastische Probleme zu erweitern (vgl. Abb. I.A.1.). Die beiden Autoren beschreiben das Vorgehen ausgehend von einer stochastischen Problemstellung, die sich im alltäglichen, nicht unbedingt mathematischen Sachzusammenhang ergibt, wie folgt: Zunächst wird durch eine Modellbildung das reale stochastische Problem in ein stochastisches Modell übersetzt, wobei gewisse – möglichst im Sachzusammenhang begründbare – Modellannahmen getroffen werden müssen. Innerhalb dieses stochastischen Modells wird dann durch die Anwendung stochastischer Theorie eine Lösung gefunden, die dann noch für das reale Ausgangsproblem unter Berücksichtigung der Modellannahmen interpretiert werden muss.[63] Erweitert man dieses Modell im Hinblick auf intuitive Lösungen realer stochastischer Probleme, so würden bei einer intuitiven Lösung die Schritte Modellbildung, stochastische Theorie (Kalkül) sowie Interpretation übersprungen bzw. unbewusst durchlaufen werden. Bei einer intuitiven Lösung eines innermathematischen

[60] Die beschriebenen Anwendungsfelder lassen sich auch mit den Untersuchungen von Wolpers zum Einsatz von Computern im Stochastikunterricht in Einklang bringen, vgl. Woplers (2002), 16f. und 120ff.
[61] Dudenredaktion (Bibliographisches Institut), Duden online, Eintrag zum Stichwort „Intuition", https://www.duden.de/rechtschreibung/Intuition (zuletzt aufgerufen am 15.04.2020).
[62] Wolpers (2002), 138.
[63] Vgl. Herbert Kütting & Martin J. Sauer, Elementare Stochastik. Mathematische Grundlagen und didaktische Konzepte, in Reihe: Mathematik Primarstufe und Sekundarstufe I + II, hrsg. von Friedhelm Padberg, 3. Aufl., korrigierter Nachdruck, Berlin, Heidelberg 2014, 74f.

Problems würde dann entsprechend die bewusste Anwendung der stochastischen Theorie ausbleiben.

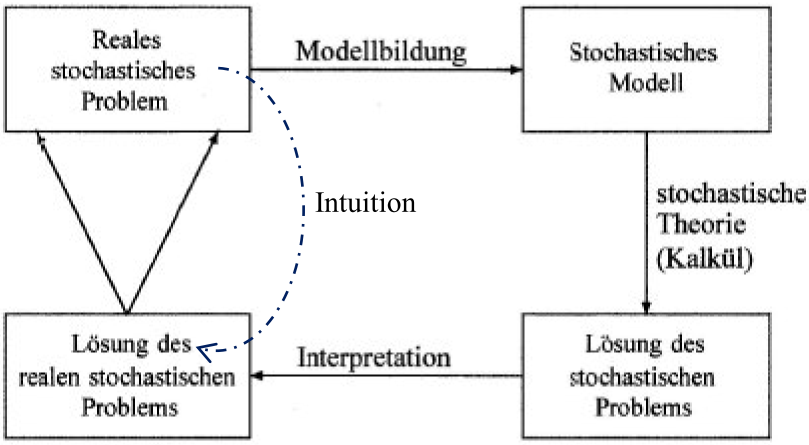


Abb. I.A.1. Im Schaubild werden die Denkschritte bei der Bearbeitung stochstischer Probleme nach Kütting und Sauer mit eigener Erweiterung zur intuitiven Herangehensweise dargestellt.[64]

Wenngleich Intuitionen in dieser Lesart nicht mit bewussten Denkprozessen operieren, sind sie dennoch durch (theoretische und praktische) Beschäftigung mit den Sachverhalten, auf die sie sich beziehen, veränderbar. Wolpers unterscheidet daher zwischen primären und sekundären Intuitionen: Während primäre Intuitionen aus Alltagswissen und -erfahrungen resultierten, entwickelten sich sekundäre Intuitionen durch die Beschäftigung mit entsprechenden theoretischen Erkenntnissen – idealerweise unter Berücksichtigung der primären Intuitionen, da ein rein formaler Zugang durch die Präsenz stochastischer Fragestellungen in außermathematischen Kontexten nicht tragfähig genug sei.[65] Im Fall der Stochastik seien sekundäre Intuitionen besonders wichtig, da die primären Intuitionen auf diesem Gebiet häufig in die Irre leiteten.[66] Dies führt Wolpers auf die Besonderheit stochastischer Phänomene und die Schwierigkeit ihrer Reflexion zurück, auf die die im Alltag präsenteren und strukturell einfacheren kausal-deterministischen Denkweisen („Ursache –

[64] Das ursprüngliche Schaubild wurde entnommen aus: Kütting/ Sauer (2014), 75, ergänzt um den gestrichelten Pfeil mit der Beschriftung „Intuition“

[65] Zu einem ähnlichen Schluss kommt auch Petra Hauer-Typpelt von der Universität Wien, die deswegen dafür plädiert, im schulischen Stochastikunterricht statt eines stark kalkülorientierten Zugangs viel Wert auf eine inhaltliche Füllung der grundlegenden Begriffe der Stochastik zu legen, damit Schülerinnen und Schüler tragfähige Grundvorstellungen etwa zu Konzepten wie „Wahrscheinlichkeit“ und „Zufall“ aufbauen können, die sich nicht allein aus der (teils irreführenden) alltagssprachlichen Verwendung der Begriffe speisen. Vgl. Petra Hauer-Typpelt, Angemessene Grundvorstellung zu Wahrscheinlichkeit und Zufall entwickeln – Vorschläge für den Stochastikunterricht, in: Didaktikhefte der Österreichischen Mathematischen Gesellschaft, Heft 43 (2010), 75-87, abrufbar unter: https://www.oemg.ac.at/DK/Didaktikhefte (zuletzt aufgerufen am 26.01.20).

[66] Falsche primäre Intuitionen werden in dieser Arbeit als „Fehlvorstellungen“ bezeichnet.

Wirkung") nicht anwendbar seien, wie umfangreiche Forschungen zu „stochastischem Denken" ergeben habe.[67]

> *Auch zu stochastischen Phänomenen wird eine Fülle von Alltagserfahrungen gesammelt. […] Ganz anders aber als bei der gedanklichen Aufarbeitung von kausal deterministischen Phänomenen verhält es sich mit der gedanklichen Ordnung stochastischer Phänomene. Kinder entwickeln zwar durch ihre Erfahrungen mit und in stochastischen Situationen „primäre Intuitionen" für deren Erklärung und kognitive Einordnung. Diese primären Intuitionen werden aber in der Regel nicht reflektiert und aufgearbeitet, so dass Fehlschlüsse und sogar Aberglaube die Deutungen von Phänomenen und Handlungen in stochastischen Situationen prägen können.*[68]

Um brauchbare sekundäre Intuitionen entwickeln zu können, sei ein handlungsorientierter Unterricht mit einem großen Anteil an Schüleraktivität von zentraler Bedeutung. Wolpers schlägt in diesem Zusammenhang explizit von den Lernenden selbst durchgeführte (und idealerweise auch selbst geplante) Simulationen als Methode vor, die auch am Computer durchgeführt werden könnten. Er führt einschränkend aber auch aus, dass Fehlvorstellungen auf diesem Gebiet schwer zu korrigieren seien.[69]

Von „stochastischen Intuitionen", für die charakteristisch ist, dass die stattfindenden Denkprozesse weitgehend unbewusst ablaufen, soll in dieser Arbeit „stochastisches Verständnis" unterschieden werden. Während der Begriff der (stochastischen) Intuition in der Literatur häufig ohne nähere Bestimmung verwendet wird, gibt es zu der Frage, was (mathematisches) Verständnis ausmacht, eine Vielzahl an kognitionspsychologischen und didaktischen Überlegungen, die hier nicht im Detail wiedergegeben werden sollen[70]. Vielmehr wird vorgeschlagen, sich bezüglich des Begriffes des mathematischen Verständnisses an den Arbeiten von Zalman Usiskin von der Universität Chicago zu orientieren, der sich im Rahmen des University of Chicago School Mathematics Project intensiv mit der Frage, wie mathematische Kompetenzen von Kindern optimal vermittelt werden können, beschäftigt hat.[71] Er schlägt vor, das Verständnis mathematischer Konzepte

[67] Vgl. Wolpers (2002), 138ff.
[68] Wolpers (2002), 140.
[69] Vgl. Wolpers (2002), 108, 140f., 170f.
[70] Grundlegend hierfür sind u.a. die Modelle von Jean Piaget und Jerome S. Bruner sowie die Unterscheidung zwischen deklarativem und prozeduralem Wissen, vgl. hierzu z.B. Kristina Reiss & Christoph Hammer, Grundlagen der Mathematikdidaktik. Eine Einführung für den Unterricht in der Sekundarstufe, in Reihe: Mathematik kompakt, Basel 2013, 28ff. und 38. Ein anderer Ansatz, der sich auf kognitionspsychologische Forschung stützt, ist die Unterscheidung zwischen intuitiven und analytischen Denkmodi, beschrieben von Roland W. Scholz. Vgl. hierzu bspw. Wolpers (2002), 144f.
[71] Besonders bekannt ist das im Rahmen des Projekts entwickelte Konzept *Everyday Mathematics*, vgl. z.B. http://everydaymath.uchicago.edu/ (zuletzt aufgerufen am 07.01.2020).

mit einem fünfdimensionalen Modell zu erklären, in dem fachliches Wissen und fachbezogene Kompetenzen kombiniert werden.[72] Dabei nimmt er bezüglich des Begriffs „Verständnis“ eine vermittelnde Position zwischen der behavoristischen und kognitivistischen Ansicht ein:

> *We in education act both as behaviourists and cognitivists. We view “understanding” as something that goes on in the brain without external actions yet we want students to exhibit their understanding by responding to tasks we present before them.*[73]

In seiner Arbeit untersucht er als Beispiele für mathematische Konzepte die Multiplikation von Brüchen und geometrische Kongruenz. Sein Begriff von mathematischen Konzepten ist aber weit gefasst und beinhaltet sowohl sehr allgemeine Konzepte wie z.B. Zahl oder Funktion, als auch spezielle Konzepte wie z.B. den Mittelwert einer Menge, den Satz des Pythagoras oder das Lösen linearer Gleichungen.[74] Mathematische Konzepte können für ihn also sowohl Begriffe, Theoreme als auch Handlungsstrategien umfassen, wenn sie multidimensional betrachtet werden.[75]

Als Voraussetzung für das Verständnis eines Konzepts benennt Usiskin die Beherrschung der dem Konzept zugehörigen Fachsprache.

> *Mathematics is, among its many other attributes, a language of discourse. It is both a written language and a spoken language, for – particulary in school mathematics – we have words for virtually all the symbols. Familiarity with this language is a precursor to all understanding.*[76]

So könne man etwa die Multiplikation von Brüchen nicht verstehen, ohne die Begriffe „Bruch“ und „Multiplikation“ zu kennen.

> *In general, dealing with the written and spoken vocabulary of a concept is an essential part of its understanding that transcends all aspects of that understanding.*[77]

Die fünf Dimensionen, die seiner Ansicht nach (aufbauend auf der entsprechenden mathematischen Fachsprache) das Verständnis mathematischer Konzepte konstituieren, benennt er wie folgt:

---

[72] Vgl. Zalman Usiskin, What does it mean to understand some mathematics, in: Selected Regular Lectures from the 12th International Congress on Mathematical Education, hrsg. von Sung Je Cho, Springer Switzerland 2015, 821-841.
[73] Usiskin (2015), 822.
[74] Vgl. Usiskin (2015), 823.
[75] Vgl. Usiskin (2015), 836f.
[76] Usiskin (2015), 824.
[77] Usiskin (2015), 824.

(1) *skill-algorithm dimension*
(2) *property-proof dimension*
(3) *use-application (modeling) dimension*
(4) *representation-metaphor dimension*
(5) *history-culture dimension*[78]

Im Folgenden soll auf Grundlage der Beispiele, die Usiskin wählt, eine inhaltliche Erläuterung dieser fünf Dimensionen erfolgen, bevor Usiskins Rechtfertigung des Begriffs „Dimensionen" wiedergegeben werden soll.

(1) Bei der *skill-algorithm dimension* geht es um Handlungsstrategien, die es ermöglichen, effizient zu einem richtigen Ergebnis zu kommen. Diese Dimension umfasst sowohl die Kenntnis von Algorithmen als auch (in einer höheren Ausprägung) die Fähigkeit, aus mehreren möglichen Algorithmen den in der gegebenen Situation effizientesten auszuwählen und umzusetzen. Man spricht diesbezüglich auch von *prozeduralem Verständnis*.[79] In einer noch höheren Ausprägung umfasst diese Dimension die Entwicklung eigener Algorithmen. Außerdem kann zu dieser Dimension ggf. auch der richtige Umgang mit Software, die die Ausführung der Algorithmen übernimmt, gezählt werden.[80]
(2) Die *property-proof dimension* beinhaltet sowohl die Formulierung mathematischer Beweise als auch das Identifizieren von mathematischen Eigenschaften, die einem Konzept zugrunde liegen. Diese Eigenschaften werden häufig von Eigenschaften allgemeinerer Konzepte abgeleitet.[81]
(3) Die *use-application (modeling) dimension* bezeichnet die Kenntnis unterschiedlicher inner- und außermathematischer Anwendungszusammenhänge eines Konzepts.[82]
(4) Bei der *representation-metaphor dimension* geht es um Repräsentationen des Konzeptes anhand konkreter Objekte, bildlicher Darstellungen oder sprachlicher Metaphern. Möglichkeiten zur Repräsentation von Konzepten böten auch Graphen

[78] Usiskin (2015), 821.
[79] Usiskin betont an dieser Stelle, dass er in diesem Sinne auch schon das korrekte Ausführen eines Algorithmus der Sphäre des Verständnisses des entsprechenden Konzepts zurechnet, und grenzt sich damit von anderen Sichtweisen ab, die algorithmisches Ausführen und Verständnis kontrastieren würden. Vgl. Usiskin (2015), 824f.
[80] Vgl. Usiskin (2015), 832.
[81] Vgl. Usiskin (2015), 826f. sowie 832f.
[82] Vgl. Usiskin (2015), 827f. sowie 833.

und Diagramme. Außerdem gebe es innermathematische Repräsentationen (etwa algebraische von geometrischen Konzepten). [83]

(5) Die *history-culture dimension* handelt von der historischen und kulturellen Bedeutung mathematischer Konzepte sowie ihrer Entstehungssituation. Diese Dimension beinhaltet auch kulturelle Unterschiede bei der Verwendung von Konzepten.[84]

Usiskin spricht von „Dimensionen“, da diese Aspekte von Verständnis eines Konzepts relativ unabhängig voneinander seien, und zwar in dem Sinne, dass sie nicht zwangsläufig aufeinander aufbauten. Sie könnten im Extremfall sogar isoliert voneinander gelernt werden.[85]

> *The four dimensions of understanding[86] are relatively independent in the sense that they can be, and are often, learned in isolation of each other, and no particular dimension need precede any of the others. Some believe mathematics should begin with real world situations; others with skills; others with concrete materials; and still others believe you should develop mathematical theory first and let everything else come from that.[87]*

Der Begriff „Dimension“ biete sich außerdem an, da das Verständnis bezüglich eines Aspekts in den verschiedenen Dimensionen unterschiedlich ausgeprägt sein könne.

> *Each dimension has aspects that can be memorized. Skills, names of properties, connections between mathematics and the real world, and even work with representations can be memorized. They also have potential for highest level of creative thinking: the invention of algorithms, the proofs that things work, the discovery of new application for old mathematics, the development of new representations or metaphors.[88]*

Im Gegensatz zur vorher skizzierten stochastischen Intuition steht bei Usiskins Begriff des mathematischen Verständnisses also der bewusste Umgang mit den Konzepten im Zentrum – der sich in den beschriebenen fünf Dimensionen manifestiert. Angewendet auf die Fragestellung dieser Arbeit kann folglich untersucht werden, inwiefern Simulationen am Computer zu einem (vertieften) Verständnis stochastischer Konzepte beitragen können.

[83] Diese Dimension von Verständnis wird insbesondere in der Kognitionspsychologie hervorgehoben, vgl. Usiskin (2015), 828ff. sowie 833f.
[84] Vgl. Usiskin (2015), 834f.
[85] Vgl. Usiskin (2015), 834. Diesbezüglich ist zu bemerken, dass etwa die Unabhängigkeit der Dimensionen (1) und (2) insbesondere bei nicht konstruktiven Beweisformen denkbar ist, weshalb das von Usiskin gewählte Beispiel der Multiplikation von Brüchen nicht ganz optimal gewählt zu sein scheint.
[86] Hier bezieht sich Usiskin auf die Dimensionen (1) bis (4). Aus den folgenden Ausführungen lässt sich schließen, dass er auch die Dimension (5) in ähnlicher Weise charakterisieren würde.
[87] Usiskin (2015), 834.
[88] Usiskin (2015), 834.

Vorstellbar sind hierbei insbesondere Beiträge zu den ersten vier genannten Dimensionen: Bezüglich der *skill-algorithm dimension* könnten Simulationen am Computer insbesondere als Überprüfung dienen (Anwendungsfeld (A1)). Hinzu kommt, dass auch die technischen Fertigkeiten selbst, die für Simulationen benötigt werden, als Bestandteil dieser Dimension verstanden werden können. Ein Beitrag zur *property-proof dimension* könnte sein, durch die Simulationen Eigenschaften von Konzepten zu identifizieren, die durch einen formalen Zugang schwer ersichtlich sind (Anwendungsfeld (A2)). Bezüglich der *use-application (modeling) dimension* sind Beiträge auf zwei Ebenen denkbar: Einerseits sind Simulationen selbst ein Anwendungsfeld stochastischer Konzepte, man denke beispielsweise an Monte Carlo Simulationen (Anwendungsfeld (A3)), andererseits beinhalten Simulationsaufgaben in der Regel auch Modellierungsschritte. Was die *representation-metaphor dimension* betrifft, so liegt die Stärke von Simulationen am Computer in der leichten Visualisierbarkeit der Ergebnisse, die die Programme mitliefern. Die Visualisierung z.B. in Form von Diagrammen, kann sowohl im Kontext von Anwendungsfeld (A1) als auch von Anwendungsfeld (A2) erfolgen.

I.B. Zielsetzung der Arbeit

Die Dissertation soll eine hochschuldidaktische Arbeit mit stoffdidaktischer Ausrichtung auf das Gebiet der Stochastik sein. Es soll untersucht werden, inwiefern Simulationen am Computer das Verständnis elementarer stochastischer Zusammenhänge verbessern können. Die Zielgruppen sind sowohl Studienanfänger des Bachelor in Mathematik als auch Lehramtsstudierende für das Fach Mathematik an Gymnasien.

Zunächst soll ein Konzept entwickelt werden, wie Simulationsaufgaben nicht nur in Fallstudien, sondern semesterbegleitend in die Lehre der Stochastik integriert werden können. Es bietet sich an, die Simulationsaufgaben in den regulären (fakultativen) Übungsbetrieb einzugliedern, wobei die Studierenden im Sinne einer didaktisch wertvollen Problem- und Handlungsorientierung einerseits möglichst eigenständig und selbsttätig[89] an der Planung, Durchführung und Interpretation der Simulationen arbeiten, andererseits aber auch binnendifferenziert ausreichend (technische) Unterstützung erhalten können sollen.

Als Programmiersprache soll „R“ verwendet werden. Dafür spricht, dass diese Programmiersprache inzwischen zu einem Standardwerkzeug im Bereich Stochastik,

[89] Vgl. hierzu z.B. Wolpers (2002), 116ff., sowie zum Begriff und Wert selbstständigen Arbeitens (bereits im schulischen Mathematikunterricht) Regina Bruder, Timo Leuders & Andreas Büchter, Mathematikunterricht entwickeln. Bausteine für kompetenzorientiertes Unterrichten, 2. Aufl., Berlin 2012, 103ff.

insbesondere Statistik, geworden ist (Realitätsnähe[90]) und sie überdies auch schon in verschiedenen Veranstaltungen des Fachbereichs Mathematik an der Johann Wolfgang Goethe Universität Frankfurt eingesetzt wird, wo das Konzept erprobt werden soll. Zudem sind die Entwicklungsumgebung sowie zahlreiche Pakete frei verfügbar. Eine Alternative wäre, eine speziell für Lehrzwecke entwickelte Software zu benutzen, z.B. Fathom. Diese didaktisch motivierten Programme stoßen jedoch rasch an technische Grenzen, wenn es um fortgeschrittenere Simulationen geht, da sie für das Erlernen von grundlegenden statistischen Verfahren und Experimentieren mit einfachen Zufallsexperimenten ausgelegt sind, und sind daher für den Zweck dieser Arbeit weniger geeignet.[91] Neben der Vertiefung stochastischer Kompetenzen könnten durch den Einsatz von R zudem technische Simulationskompetenzen erworben werden, die sowohl für Lehramts- als auch für Bachelorstudierende im folgenden Ausbildungskontext und späteren Berufsbereich nützlich sein könnten.
Inhaltlich sollen durch die Simulationen theoretische Erkenntnisse im Sinne eines entdeckenden, problemorientierten Lernens[92] „erfahrbar“ gemacht werden und so einerseits ein erweiterter Blick auf stochastische Phänomene ermöglicht sowie andererseits der Feinsinn für stochastische Zusammenhänge geschärft werden. Dabei soll die Simulation am Computer nicht als Konkurrenz zu theoretischen Zugängen aufgefasst werden, sondern für ein Zusammenspiel von „Theorie“ und „Praxis“, von „Beweis“ und „Ausprobieren“ geworben werden. Als Inspiration kann auf Material aus dem bereits mehrfach an der Goethe Universität durchgeführten Seminars „Stochastik mit dem Rechner“ für Lehramtsstudierende zurückgegriffen werden. Dieses kann aufbereitet und um geeignete Themen erweitert werden. Insbesondere sollen aber Aufgaben entwickelt werden, die sich inhaltlich stark an den Themen der Vorlesung „Elementare Stochastik“ orientieren.

Durch eine Evaluation des zu entwickelnden Ansatzes, für die geeignete Instrumente entwickelt werden müssen, soll explorativ untersucht werden,

- welche Gruppen von Studierenden welchen Bearbeitungsmodus wählen,
- inwiefern durch den Einsatz von Simulationen möglicherweise vorhandene Fehlvorstellungen besser abgebaut werden können als durch eine „klassische“ Vorlesung mit theoretischen Übungsaufgaben, d.h. wie die Bearbeitung von

[90] Der motivationale Aspekt von Anwendungsorientierung wird bspw. von Wolpers beschrieben. Er bezieht sich in erster Linie auf den Inhalt von Aufgaben, die Erkenntnisse lassen sich aber m. E. auf den methodischen Zugang übertragen. Vgl. Wolpers (2002), 118ff.
[91] Zu didaktisch günstigen Softwarelösungen vgl. z.B. Rolf Biehler (1997), Software for Learning and for Doing Statistics, International Statistical Review 65/2, 167-189. Eine weitere Übersicht findet sich bei Wolpers (2002), 133f. Letzterer bezieht sich bei seiner Analyse allerdings ausschließlich auf den Einsatz in der Schule.
[92] Vgl. hierzu z.B. Wolpers (2002), 114 f.

Simulationsaufgaben die Fertigkeiten der Studierenden im Gebiet der Stochastik beeinflusst,

- inwiefern sich der Einsatz von Simulationen auf die Einstellungen der Studierenden dem Fach Stochastik gegenüber auswirkt sowie auf ihre Selbsteinschätzung bzgl. Verständnis und Intuition,
- inwiefern Studierende selbst den Einsatz von Simulationen für lohnenswert halten,
- welche Unterstützungen von den Studierenden benötigt werden.

I.C. Einbettung in das Konzept der Vorlesung „Elementare Stochastik"

Im Gegensatz zu Vermittlungsansätzen im schulischen Kontext basiert die Lehre der Stochastik an der Universität hauptsächlich auf einer maßtheoretischen Modellierung von Zufallsexperimenten.[93] So hilfreich dieser Zugang auch für fortgeschrittene Arbeiten auf diesem Gebiet ist, so wenig dockt er an alltagspraktische Erfahrungen und Vorstellungen von Zufall und Wahrscheinlichkeit an und ist auch für zahlreiche Anwendungen der Stochastik wenig zentral. Der Mathematiker Edward Nelson bemerkt in seinem Buch „Radically elementary probability theory":

> *The foundations of probability theory were laid just over fifty years ago, by Kolmogorov. I am sure that many other probabilists teaching a beginning graduate course have also had the feeling that these measure-theoretic foundations serve more to salve our mathematical consciences than to provide an incisive tool for the scientist who wishes to apply probability theory.*[94]

Das Konzept der Vorlesung „Elementare Stochastik" an der Johann Wolfgang Goethe Universität in Frankfurt[95], das von Götz Kersting und Anton Wakolbinger unter Mitarbeit von Brooks Ferebee entwickelt und vielfach erprobt wurde, versucht eine Alternative zu dem maßtheoretischen Ansatz zu bieten, indem Begriffe wie etwa der von Zufallsvariablen stärker mit intuitiven Vorstellungen verbunden werden und trotzdem ein formales Werkzeug entwickelt wird, mit dem auch fortgeschrittene Probleme der Stochastik gelöst werden können.

[93] Dass nichtsdestotrotz sogar auch in Schulbüchern teilweise ein axiomatischer Zugang zum Konzept der Zufallsvariablen gewählt wird, was für die Entwicklung von „stochastischem Denken" hinderlich sein könnte, wird bei Wolpers (2002), 104f. beschrieben.

[94] Edward Nelson, Radically elementary probability theory, Annals of Mathematics Studies 117, Princeton 1987, vii (Preface).

[95] Eine ausführliche Darlegung des Konzepts findet sich in Kersting/ Wakolbinger (2008).

Aus den Überlegungen von Kersting und Wakolbinger lässt sich als Grundvorstellung[96], die sich mit dem Begriff der Zufallsvariablen verbinden soll, die zufällige Wahl eines Elements aus einer Menge (nämlich dem Zielbereich/ Wertebereich S der Zufallsvariable) ableiten. Die zu einer Zufallsvariablen gehörende Wahrscheinlichkeitsverteilung gibt dann die Chancen für die verschiedenen Elemente an.[97] Im Folgenden wird auf einen gewissen mathematischen Formalismus zurückgegriffen, den Kersting und Wakolbinger verwenden, um die Analogie zwischen stetigen und diskreten Zufallsvariablen zu veranschaulichen. Die Verteilungen diskreter Zufallsvariablen lassen sich durch Gewichte beschreiben:

> *Eine Zufallsvariable heißt diskret, falls ihr Zielbereich abzählbar ist oder allgemeiner eine abzählbare Menge S enthält mit der Eigenschaft* $P\,(X \in S) = 1.$ *Die Abbildung* ρ,
>
> $$A \mapsto \rho[A] := P(X \in A), \quad A \subset S,$$
>
> *heißt die Verteilung von* X*. Die Zahlen*
>
> $$\rho(a) := P(X = a), \quad a \in S,$$
>
> *sind die Verteilungsgewichte.*[98]

Dabei ist

$$P(X \in A) = \sum_{a \in A} P(X = a), \quad A \subset S,$$
$$P(X \in S) = 1.$$[99]

Bei kontinuierlichen Zufallsvariablen bieten die Autoren eine anschauliche Interpretation von Dichten als infinitesimale Gewichte[100] an[101], um eine Analogie zu den Gewichten von diskreten Zufallsvariablen konstruieren zu können. Im Fall von Intervallen der reellen Achse als Zielbereichen wird dies folgendermaßen ausformuliert:

---

[96] Der Begriff wird hier in Anlehnung an die Verwendung bei Rudolf vom Hofe gebraucht, der als drei charakteristische Komponenten von Grundvorstellungen visuelle Repräsentationen, Anknüpfung an bekannte Handlungsmuster und prototypische Anwendungszusammenhänge benennt. Vgl. bspw. Rudolf vom Hofe, Grundvorstellungen mathematischer Inhalte als didaktisches Modell, Journal für Mathematik-Didaktik 13/4 (1992), 345–364 und Rudolf vom Hofe, Grundvorstellungen mathematischer Inhalte, Heidelberg 1995.
[97] Vgl. Kersting/ Wakolbinger (2008), 1.
[98] Kersting/ Wakolbinger (2008), 19.
[99] Kersting/ Wakolbinger (2008), 18.
[100] Unter infinitesimalen Größen sind in Anlehnung an Isaac Newtons Überlegungen Objekte zu verstehen, die beliebig klein, aber größer als Null sind. Eine mathematische Präzisierung des Begriffes findet man in der Nichtstandardanalysis. Vgl. hierzu bspw. Joseph W. Dauben, Abraham Robinson and Nonstandard Analysis – History, Philosophy, and Foundations of Mathematics, in: History and Philosophy of Modern Mathematics, hrsg. von William Aspray & Philip Kitcher, Univ. of Minnesota 1988, 177–200. Als pdf abrufbar unter: http://www.mcps.umn.edu/philosophy/11_7Dauben.pdf (zuletzt aufgerufen am 15.04.2020). Zu infinitesimalen Größen im Kontext von Stochastik vgl. bspw. Nelson (1987), 16ff.
[101] Vgl. Kersting/ Wakolbinger (2008), 17.

Sei $S \subset \mathbb{R}$ ein Intervall mit Endpunkten l, r, $-\infty \leq l \leq \infty$, und sei $f: S \to \mathbb{R}$ eine nicht-negative integrierbare Funktion mit der Eigenschaft

$$\int_l^r f(a)\, da = 1.$$

Gilt dann für eine Zufallsvariable X mit Zielbereich S für alle Intervalle $[c, d] \subset S$ die Gleichung

$$P\,(X \in [c, d]) = \int_c^d f(a)\, da,$$

so sagt man, dass X die Dichte $f(a)da$ besitzt. Wir schreiben dann auch kurz

$$P\,(X \in da) = f(a)\, da, \quad a \in S.$$ [102]

Der Ziel- bzw. Wertebereich der Zufallsvariablen entspricht dem Ereignisraum bzw. Zustandsraum. Auf die formale Definition eines Wahrscheinlichkeitsraums zur Modellierung von Zufallsexperimenten wird verzichtet, insbesondere wird kein Grundraum bzw. keine Grundmenge Ω definiert. Folglich wird auch die Ereignismenge nicht als σ-Algebra über dem Grundraum beschrieben und auf die klassische Definition eines Wahrscheinlichkeitsmaßes in diesem Kontext verzichtet. Eine Zufallsvariable wird also zunächst nicht als messbare Funktion verstanden, sondern es kann bei der intuitiven Vorstellung der zufälligen Wahl aus einer Menge bleiben. Sie rückt damit stärker in den Fokus als bei maßtheoretischen Zugängen, wo sie sozusagen ein abgeleitetes Objekt[103] ist, basierend auf einem Wahrscheinlichkeitsraum und einem Messraum, der den Wertebereich der Funktion darstellt (dem Ereignisraum). Damit könnten Schwächen des „klassischen" Ansatzes vermieden werden, wie etwa, dass die Modellierung im maßtheoretischen Kontext teils zu unnötigen oder sogar irreführenden Abstraktionen führe. Als Beispiel nennt Kersting in einem Paper, in dem er eine Axiomatik von Zufallsvariablen und Ereignissen rechtfertigt, u.a. die Elemente des Grundraumes Ω:

> *The 'small omegas' do not show up in any relevant result of probability theory, and one could do without them, if they were not needed to* define *measurable functions. [...] All these observations indicate that measurable spaces and mappings indeed make up a model of events and random variables. This is not to say that such models should be avoided, but one should not overlook that they might mislead. [...] Also one should be cautious in giving the elements of Ω some undue relevance ('state of the world'), which may create misconceptions.*[104]

[102] Kersting/ Wakolbinger (2008), 37.
[103] Vgl. Götz Kersting, G, Random Variables – without Basic Space, in: Trends in Stochastic Analysis, hrsg. von Jochen Blath, Peter Mörters, Michael Scheutzow, Cambridge 2009, 13-34.
[104] Kersting (2009), 16.

Die Diskussion darüber, wie zentral und eigenständig Zufallsvariablen in der Stochastik behandelt werden sollen, hat eine historische Vorgeschichte, in der auch die Frage nach der Vereinbarkeit mit intuitiven Vorstellungen aufgeworfen wurde, wie Kersting herausarbeitet:

> *Today it is customary to adapt random variables to a context from measure theory. Yet the feeling has persisted that random variables are objects in their own right. This was manifest, when measure theory took over in probability: According to J. Doob (interviewed by Snell) "it was a shock for probabilists to realize that a function is glorified into a random variable as soon as its domain is assigned a probability distribution with respect to which the function is measurable." Later the experts insisted that it is the idea of random variables, which conforms the intuition. Legendary is L. Breiman's statement: "Probability theory has a right and a left hand. On the right is the rigorous foundational work using the tools of measure theory. The left hand 'thinks probabilistically', reduces problems to gambling situations, coin-tossing, and motions of a physical particle." In applications of probability the concept of a random variable never lost its appeal. We may quote D. Mumford: "There are two approaches to developing the basic theory, eliminating the probabilistical language ... The other is to put the concept of 'random variable' on center stage and work with manipulations of random variables wherever possible". And, "for my part, I find the second way ... infinitely clearer".*
> *From an architectural point of view these considerations and statements suggest to try and start from random variables in the presentation of probability theory and therewith to bring intuition and methods closer together – rather than to gain random variables as derived quantities in the accustomed measure-theoretic manner.*[105]

Der zitierte David Mumford äußert sich sogar noch weitergehend, indem er Zufallsvariablen mit den grundlegenden Objekten der Geometrie, Algebra und Analysis vergleicht:

> *The basic object of study in probability is the* random variable *and I will argue that it should be treated as a basic construct, like spaces, groups and functions, and it is artificial and unnatural to define it in terms of measure theory.*[106]

Der Ansatz, Zufallsvariablen in den Mittelpunkt zu rücken, ist nicht einzig didaktischer Natur. Es kann eine eigene mathematische Theorie dazu entwickelt werden, die auch für Forschungszwecke dienlich ist. Durch diese alternative Zugangsweise zu stochastischen Fragestellungen können andere Beweisstrategien verfolgt werden, die gegenüber den

[105] Kersting (2009), 14.

[106] David Mumford, The Dawning of the Age of Stochasticity, in: Mathematics. Frontiers and Perspectives, hrsg. von Vladimir I. Arnold, Michael Atiyah, Peter D. Lax & Barry Mazur, AMS 2000, 198.
Mumford geht noch weiter und sieht in Zufallsvariablen sogar die grundlegenden Objekte unseres Denkens im Allgemeinen: *My overall conclusion is that I believe stochastic methods will transform pure and applied mathematics in the beginning of the third millennium. Probability and statistics will come to be viewed as the natural tools to use in mathematical as well as scientific modelling. The intellectual world as a whole will come to view logic as a beautiful elegant idealization but to view statistics as the standard way in which we reason and think.* (Mumford (2000), 217 und vgl. Mumford (2000), 200.) Diese philosophischen Überlegungen zum Stellenwert von Zufallsvariablen in unseren Denkprozessen sind allerdings nicht Gegenstand dieser Arbeit.

maßtheoretischen vorteilhaft sein können. Kersting zeigt dies am Beispiel von Beweisen des Zentralen Grenzwertsatzes, bei denen in der klassischen maßtheoretischen Verfahrensweise charakteristische Funktionen benutzt werden und durch rein analytische Schlüsse das Resultat hergeleitet wird. Mit der alternativen Modellierung kann man dagegen inner-stochastisch argumentieren, indem mit einer Coupling Methode standardnormalverteilte Zufallsvariablen durch die unabhängig, identisch verteilten Ausgangs-Zufallsvariablen ersetzt werden, und gezeigt wird, dass die Abweichung von der Normalverteilung asymptotisch verschwindet. Anders als bei der analytischen Methode kommen hier stochastische Argumente zum Einsatz.[107] Auch Mumford zeigt an einem Beispiel, dass der Ansatz, mit Zufallsvariablen zu argumentieren, zu tieferen Einsichten in die stochastischen Zusammenhänge führen kann:

> *[…] making the random variables explicit tells you the real stochastic meaning of the result.*[108]

Bemerkenswert ist, dass neben Mumford auch ein anderer mit der Fields Medaille ausgezeichneter Mathematiker Zufallsvariablen stärker in den Fokus rückt. Terence Tao, einer der bedeutendsten Mathematiker der Gegenwart, der auch auf dem Gebiet der Stochastik wegweisende Arbeit leistet, schreibt in seinem Buch „Topics in random matrix theory":

> *In measure theory, the underlying measure space plays a prominent foundational role, with the measurable sets and measurable functions (the analogues of the events and the random variables) always being viewed as somehow being attached to that space. In probability theory, in contrast, it is the events and their probabilities that are viewed as being fundamental, with the sample space being abstracted away as much as possible, and with the random variables and expectations being viewed as derived concepts. […]*
> *However, it is possible to take the abstraction process one step further, and view the algebra of random variables and their expectations as being the foundational concept, and ignoring both the presence of the original sample space, the algebra of events, or the probability measure.*[109]

Ähnlich wie Kersting wirbt auch Tao dafür, neben der klassischen maßtheoretischen Zugangsweise eine alternative Definition von Zufallsvariablen zuzulassen und damit alternative Modellierungen stochastischer Zusammenhänge zu ermöglichen, um eine Perspektiverweiterung gewinnen zu können:

[107] Den dataillierten Beweis findet man in Kersting (2009), 14f. und noch ausführlicher in Kersting/ Wakolbinger (2008), 74f.

[108] Mumford (2000), 206. Er illustriert seine Aussage am Beispiel eines Beweises des Levy-Khintchine Theorems.

[109] Terence Tao, Topics in random matrix theory, in Reihe: Graduate Studies in Mathematics, Vol. 132, AMS 2012, zitiert nach Preprint (pdf), abrufbar unter: https://terrytao.wordpress.com/category/teaching/254a-random-matrices, 155 (zuletzt aufgerufen am 10.04.2020).

Note that this foundational preference is to some extent a metamathematical one rather than a mathematical one; in many cases it is possible to rewrite the theory in a mathematically equivalent form so that some other mathematical structure becomes designated as the foundational one, much as probability theory can be equivalently formulated as the measure theory of probability measures. However, this does not negate the fact that a different choice of foundations can lead to a different way of thinking about the subject, and thus to ask a different set of questions and to discover a different set of proofs and solutions. Thus it is often of value to understand multiple foundational perspectives at once, to get a truly stereoscopic view of the subject.[110]

Der mathematische Ansatz zur Definition von Zufallsvariablen, der in dem zitierten Paper von Kersting vorgestellt wird, stellt somit eine alternative Herangehensweise an die Untersuchung von Zufallsexperimenten dar, der in der Vorlesung „Elementare Stochastik“ ohne die mathematische Herleitung verwendet wird. Das wesentliche Charakteristikum, Zufallsvariablen und ihre Interpretationen in den Mittelpunkt der Betrachtungen zu rücken, ist für den Einsatz von Simulationen ein geeigneter Anknüpfungspunkt, da bei ihnen ja Realisationen von Zufallsvariablen erzeugt werden, die den Ausgangspunkt für Interpretationen bilden. Zudem ist das mit diesem Konzept verbundene didaktische Ziel, die stochastische Intuition zu schärfen und die stochastischen Prinzipien selbst so direkt wie möglich zu untersuchen, mutmaßlich durch Simulationen noch besser zu erreichen.

Die fachlichen Ziele der Vorlesung orientieren sich an einer Einführung in grundlegende Konzepte der Wahrscheinlichkeitstheorie sowie der Statistik. Sie umfassen diskret und kontinuierlich verteilte Zufallsvariablen, Erwartungswert und Varianz, stochastische Abhängigkeit und Unabhängigkeit, den Zentralen Grenzwertsatz, bedingte Wahrscheinlichkeiten, den Satz von Bayes, Hypothesentests, Konfidenzintervalle und lineare Modelle in der Statistik sowie Markovketten.
Für den Einsatz von Simulationen eignen sich dabei insbesondere Themen, bei denen die Variabilität von Schätzungen untersucht werden kann. Zusätzlich zu einer möglichen Vertiefung der genannten Konzepte können die Studierenden durch sie auch weitere stochastische Phänomene bzw. Verfahrensweisen erforschen, z.B. Größenverzerrungen, Regression zur Mitte, Erzeugung von Zufallszahlen oder Monte Carlo Verfahren.

[110] Tao (Preprint), 156 (in der Fußnote Nr. 38).

II. DER MODELLVERSUCH „R AUFGABEN“ ZUR ELEMENTAREN STOCHASTIK

In diesem Kapitel soll der Modellversuch zum Einsatz von Simulationsaufgaben bei der Vorlesung „Elementare Stochastik“ an der Goethe Universität in Frankfurt beschrieben werden. Die Vorlesung richtet sich an Bachelor- und Gymnasiallehramtsstudierende zu Beginn des Studiums und besteht aus einer vierstündigen Vorlesung und einer zweistündigen Übung. Jede Woche werden den Studierenden insgesamt vier Übungsaufgaben gestellt, die sie als Hausaufgaben lösen sollen, wobei in der Übungsgruppe Hinweise zu Lösungsstrategien gegeben werden. Die Lösungen werden von Tutoren korrigiert und in den Übungsgruppen besprochen. Durch die Lösung der Aufgaben (und eine Präsentation in der Übungsgruppe) können Bonuspunkte für die Klausur am Ende des Semesters gesammelt werden.

II.A. DER AUFBAU DES VERSUCHS

Für den Modellversuch wurde eine der vier wöchentlich zu lösenden Übungsaufgaben durch eine Simulationsaufgabe ersetzt, die „R Aufgabe“ genannt wurde. Die Lösungen wurden von den Tutoren korrigiert, die Hinweise zu Lösungsstrategien allerdings anders als bei den theoretischen Aufgaben nicht in den Übungsgruppen gegeben, sondern – zur Vereinheitlichung der Voraussetzungen – in Form eines Templates auf der Homepage der Veranstaltung zur Verfügung gestellt.[111] Zusätzlich wurde eine „Fragestunde“ angeboten, in der Lösungen vergangener Simulationsaufgaben detailliert diskutiert und (insbesondere technische) Fragen zu der aktuellen Aufgabe gestellt werden konnten. Im Anschluss daran konnten Studierende in diesem Rahmen auch gemeinsam an der aktuellen „R Aufgabe“ arbeiten. Der Besuch der Fragestunde war freiwillig und bot keine Möglichkeit zum Erwerb weiterer Bonuspunkte für die Klausur. Die Lösungen der Simulationsaufgaben wurden nach der Abgabe auch auf der Homepage veröffentlicht. Zusätzlich zu den „R Aufgaben“ wurden auch in der Vorlesung an einigen Stellen Simulationen bzw. deren Ergebnisse vorgeführt, so dass auch die Studierenden, die die „R Aufgaben“ nicht bearbeiteten, Kontakt zu Simulationen haben konnten.

Die Simulationsaufgaben behandelten eine Auswahl an fachlichen Inhalten der Vorlesung, teilweise in unmittelbarem Anschluss an theoretische Aufgaben, teilweise unabhängig von ihnen. Eine didaktische Analyse findet sich im folgenden Kapitel II.B.

[111] Die Alternative, einen festen Simulationsplan wie ihn bspw. Rolf Biehler empfiehlt (vgl. Kap. I.A.2.), zu verwenden, wird in diesem Fall aufgrund der Komplexität der Simulationsaufgaben als weniger hilfreich eingeschätzt als die Templates.

Der Einsatz dieser Simulationsaufgaben wurde durch Fragebögen und Interviews evaluiert. Als Kontrollgruppe wurden die Teilnehmerinnen und Teilnehmer der Vorlesung ein Jahr zuvor herangezogen, die dieselbe Vorlesung ohne die Simulationsaufgaben belegt hatten. Sie wurden – ebenso wie die Experimentalgruppe – zu Beginn und am Ende des Semesters mit Hilfe von Fragebögen zu stochastischen Kompetenzen und Einstellungen dem Fach gegenüber befragt.

II.B. Didaktische Analyse der „R Aufgaben“

II.B.1. Vorbemerkungen und Überblick über die Aufgaben

Im Folgenden werden die entwickelten „R Aufgaben“ vorgestellt und im Hinblick auf die Fragen, welche fachlichen und methodischen Kompetenzerweiterungen sie bei den Studierenden leisten können, analysiert. Auf der fachlichen Ebene sollen die Aufgaben den Studierenden Lernchancen bezüglich der drei herausgearbeiteten Anwendungsfelder von Simulationen ermöglichen:

(A1) als Hilfsmittel zur Überprüfung von theoretischen Ergebnissen oder Verfahrensweisen,
(A2) als Hilfsmittel zur Exploration und Generierung neuer Hypothesen sowie
(A3) als Hilfsmittel zur Annäherung nicht oder nur sehr schwer berechenbarer Größen.

Dabei liegt ein Schwerpunkt auf den Anwendungsfeldern (A1) und (A2). Zu ihnen lässt sich im Fall von (A1) in 7 und im Fall von (A2) in 10 Fällen ein Bezug herstellen. In das Anwendungsfeld (A3) wird durch Monte Carlo Simulationen in zwei „R Aufgaben“ ein Einblick ermöglicht. Die „R Aufgaben“ sind als Lernaufgaben konzipiert, die ein eigenständiges Weiterforschen ermöglichen. Außerdem bieten die Aufgaben anders als die meisten theoretischen Übungen des Kurses einen Anlass zur Versprachlichung[112] stochastischer Zusammenhänge, da die Ergebnisse teilweise aufgrund von in den Aufgaben vorgesehenen Visualisierungen leichter in Worte gefasst werden können oder auch für eine Interpretation im gegebenen Sachzusammenhang sogar in Alltagssprache übersetzt werden müssen. In den „R Aufgaben“ sollen wesentliche Themen und Konzepte einer Einführungsvorlesung in die Stochastik vertieft und auch gängige Fehlvorstellungen der Stochastik, wie etwa die „Gambler’s Fallacy“[113], aufgegriffen werden. Neben den klassischen diskreten und kontinuierlichen Verteilungen (Bernoulli-, Binomial-, Normal- und

[112] Die Frage, inwiefern mit diesen Versprachlichungsanreizen auch Lernchancen für Stochastik verbunden sind, insbesondere was den Abbau von Fehlvorstellungen betrifft, steht zwar nicht im expliziten Fokus der Arbeit, soll aber bei der Untersuchung trotzdem im Blick behalten werden.
[113] Vgl. Kahneman/ Tversky (1972).

Exponentialverteilungen, uniforme Verteilungen) thematisieren sie Monte Carlo Verfahren, Verzerrungsprobleme bei Schätzungen aus der Statistik (illustriert am Beispiel von Belegzeiten von Betten im Krankenhaus und am Phänomen Regression zur Mitte), Konfidenzintervalle, aber auch Fragen zur „Erkennung" von Zufall am Beispiel des Münzwurfs und die Erzeugung von Zufallszahlen mit der Verwerfungsmethode.

Eine Übersicht der adressierten stochastischen Konzepte und Fehlvorstellungen (vgl. Kap. I.A.1.) findet sich in der folgenden Tabelle Tab.II.B.1. Bei der Fehlvorstellung (F2) zur Variabilität von Schätzungen wird nur dann zwischen den Fällen (a) und (b) unterschieden, wenn die Adressierung in der Aufgabe eindeutig ist. Wird beispielsweise implizit der Einfluss der Wiederholungsanzahl auf die Variabilität von Schätzungen in der Aufgabe behandelt, kann das je nach Perspektive als Adressierung der Fehlvorstellung (F2) auf der Ebene von Häufigkeits- oder aber von Stichproben-/ Kennwerteverteilungen verstanden werden.

Zwischen den Adressierungen der Fehlvorstellungen F1 (lokale Repräsentativität) und F2 (Unabhängigkeit zwischen der Variabilität von Schätzungen und den zugrundeliegenden Stichprobengrößen) gibt es einen wesentlichen Unterschied: Während bei F1 sowohl in der Vorlesung als auch in den Übungsaufgaben explizit thematisiert wurde, worin die Fehlvorstellung besteht und dass sie bei vielen Menschen vorzufinden ist, wurde dies bei F2 nicht gemacht. Dadurch sollten Anhaltspunkte für die Beantwortung der Frage, ob eine explizite Behandlung von Fehlvorstellungen einer impliziten überlegen ist, selbst wenn die implizite häufiger vorgenommen wird. Der Einfluss von Stichprobengrößen auf die Genauigkeit von Schätzungen (F2) wird in gewisser Weise bei (fast) allen Simulationsaufgaben adressiert, da Wiederholungsanzahlen festgelegt werden müssen. Die Fälle, in denen diese Adressierung die einzige in der Aufgabe darstellt, werden in der folgenden Tabelle als „indirekt" gekennzeichnet. Wird der Inhalt von F2 dagegen ausführlicher thematisiert, etwa indem die Aufgabe auf eine Untersuchung eines $\sqrt{n}$ – Zusammenhangs abzielt, wird dies in der Tabelle als „implizit" bezeichnet, da bei F2 auch eine solche Aufgabe nicht explizit die Auseinandersetzung mit entsprechenden fehlerhaften Vorstellungen vorsieht.

Nr.	Titel der Aufgabe	Stochastische Konzepte	Fehlvor-stellungen	Aufgabenteile zur Versprachlichung / Interpretation
1	Zum Warmwerden mit R	Zufällige Münzwurfreihen, Summen unabhängiger Bernoulli-verteilter ZV, Erwartungswert, implizit: Einfluss der Wiederholungsanzahl auf die Variabilität von (Kennwerte-) Schätzungen[114]	F2 (implizit)	
2	Crabs (Fortsetzung)	Mehrstufige Zufallsexperimente, Binomialverteilung, implizit: Einfluss der Wiederholungsanzahl auf die Variabilität von Schätzungen	F2 (implizit)	Teil iii)
3	Monte Carlo Simulation zur Bestimmung von π	Monte Carlo Simulation, Standardfehler, $\sqrt{n}$ – Gesetz	F2(b) (indirekt)	
4/5	Schätzung der Belegzeiten von Betten in einem Krankenhaus	Exponentialverteilung, Größenverzerrung, Erwartungswert, implizit: Einfluss der Wiederholungsanzahl auf die Variabilität von Schätzungen	F2 (implizit)	Aufg. 4: Teil iii) Aufg. 5: Teile i) und ii)
6	Vorzeichenwechsel bei Münzwürfen	Vorzeichenwechsel, Länge von Runs, Erwartungswert, statistischer Test, implizit: Einfluss der Wiederholungsanzahl auf die Variabilität von Schätzungen	F1 (explizit) F2 (implizit)	Teil iii)
7	Regression zur Mitte	Normalverteilung, Korrelation, Variabilität des Mittelwerts, statistische Tests, implizit: Einfluss der Wiederholungsanzahl auf die Variabilität von Schätzungen	F2 (implizit)	Teile ii) und iii)
8	Rejection Sampling und Monte Carlo Simulation	Erzeugung von Zufallszahlen, Monte Carlo Simulation, implizit: Einfluss der Wiederholungsanzahl auf die Variabilität von Schätzungen	F2 (implizit)	Teil i)
9	Fehlstände bei einer rein zufälligen Permutation	Uniforme Verteilung, Erwartungswert, Varianz, Normalverteilung, Zentraler Grenzwertsatz		Teile i) und ii)
10	Konfidenzintervalle	Binomialverteilung, Zentraler Grenzwertsatz, Zufälligkeit von Konfidenzintervallen, Überdeckungswahrscheinlichkeit, Standardfehler, $\sqrt{n}$ - Gesetz	F2(b) (indirekt)	Teil iii)
11	Bedingte Wahr-scheinlichkeiten – Code erklären	Bedingte Wahrscheinlichkeiten, Schätzen, implizit: Satz von Bayes,		inhaltlich: Teil vi)
12	Klausuraufgabe – Code erklären	Polya-Urne, Schätzen, uniforme Verteilung		inhaltlich: Teil iv)

Tab.II.B.1. In der Tabelle werden die Inhalte der „R-Aufgaben“ kurz bezüglich der behandelten stochastischen Konzepte und Fehlvorstellungen sowie ihrer Versprachlichungsanteile charakterisiert.

[114] Die Anzahl der Wiederholungen ist nicht in der Aufgabenstellung vorgegeben, sondern muss selbst gewählt werden.

In allen „R Aufgaben" wird ein intuitiver Zugang zu Zufallsvariablen gewählt, indem sie als „Variablen", die einen zufälligen[115] Wert annehmen können, aufgefasst werden. Auf maßtheoretische Zugänge wird vollständig verzichtet. Ein besonderes Augenmerk wird zudem auf die Variabilität als wesentliche Eigenschaft stochastischer Zusammenhänge gelegt und in diesem Zusammenhang der Einfluss der Wiederholungsanzahl eines Zufallsexperiments als wesentlicher Faktor für die Güte einer Schätzung auf vielfältige Weise erfahrbar gemacht.

Auf der methodischen Ebene sollen die Aufgaben den Studierenden die Möglichkeit eröffnen, sich mit grundlegenden Ideen des Programmierens und konkret mit basalen Befehlen und Strukturen der Programmiersprache R vertraut zu machen. Da nicht davon ausgegangen werden kann, dass alle Studierenden Vorkenntnisse im Programmieren besitzen, werden in den Templates nicht nur grundlegende Befehle des Programms R, sondern auch allgemeine strukturelle Konzepte wie for-Schleifen oder if-Abfragen und Objekte wie Vektoren oder Matrizen schrittweise eingeführt. Da zu erwarten ist, dass besonders bei der Syntax von Befehlen zur Simulation von Zufallszahlen und der Erzeugung von Graphiken Schwierigkeiten auftauchen könnten, wird diese zunächst weitgehend vorgegeben. Um den Umgang mit dem Programm insgesamt zu erleichtern, wurden die Templates so konstruiert, dass sie jeweils stückweise in die Konsole kopiert werden können und nicht als Ganzes eingelesen werden müssen. Dadurch können die Teilaufgaben einzeln nacheinander bearbeitet oder auch nur einzelne Teile des Templates als Hilfestellung verwendet werden, wobei alle Befehle in R bzw. Strukturen, die bis dahin nicht vorkamen, für die Lösung der Aufgabe aber nötig sind, explizit gezeigt werden. Ein Lösungsvorschlag der aktuellen Aufgabe auf Grundlage des Templates wurde jede Woche nach erfolgter Abgabe auf der Homepage veröffentlicht.

Den Studierenden wurde bei der Bearbeitung der Aufgaben freigestellt, ob sie eigenständig ein Programm in R schreiben oder ob sie auf ein Template zurückgreifen wollten, das auf der Homepage zur Verfügung gestellt wurde und als eine Art „Lückentext mit Anleitung" konzipiert war, bei dem eine Struktur zur Lösung[116] vorgegeben ist und nur einzelne Teile

[115] Die Tatsache, dass durch ein deterministisches Programm nur Pseudozufallszahlen erzeugt werden können, wird dabei ausgeklammert.

[116] Biehler und Maxara verwenden in ihren Untersuchungen einen festen Simulationsplan als Strukturierungshilfe für die Studierenden (vgl. Kap. I.A.2. dieser Arbeit). Aufgrund der Komplexität und strukturellen Vielfältigkeit der Simulationen, die in dem hier beschriebenen Kurs verwendet werden sollen, bietet sich die Verwendung eines solchen Plans nicht an. Durch die Templates sollen den Studierenden aber Orientierungshilfen für den Aufbau der für die Lösung der Aufgaben erforderlichen Simulationen geboten werden.

verändert oder ergänzt werden mussten. Dabei stand methodische Hilfestellung im Sinne von technischen Hilfestellungen zur Programmierung im Vordergrund. Die Templates wurden so konstruiert, dass Studierende ohne Vorkenntnisse im Programmieren die „R Aufgaben“ lösen können sollten. Zur Vermeidung übermäßiger Frustration wurde insbesondere bei Graphiken ausführliche Hilfestellung geleistet, da bei diesen durch die zahlreichen Parameter, die in R vorgesehen sind, besonders viele Fehlerquellen lauern. Im Laufe des Kurses stieg sowohl der Anteil als auch die Schwierigkeit der Teile, die von den Studierenden selbst erarbeitet werden mussten, an.

Bei der Erstellung der Templates stand keine effiziente Programmierung im Vordergrund, sondern es sollten die einzelnen Denkschritte hin zur Lösung sichtbar gemacht werden. Exemplarisch wird im folgenden Kapitel im Rahmen der didaktischen Analyse der einzelnen Aufgaben das Template zur Aufgabe 3 (Monte Carlo Simulation) analysiert, da in der dritten Woche die grundlegenden Funktionen in R bereits eingeführt waren und das Template somit einen auch für die folgenden Aufgaben beispielhaften Charakter besitzt. Alle Templates finden sich ebenso wie die Lösungsvorschläge im Anhang. Bei der didaktischen Analyse werden wesentliche Elemente der Templates dann beschrieben, wenn sie charakteristische Methodiken bzw. für das Programmieren zentrale Konzepte erstmals enthalten.

Zusätzlich zu diesen Templates gab es das Angebot einer wöchentlichen „R Fragestunde“ ergänzend zum regulären Tutorium, bei der Fragen zur (aktuellen und vorherigen) „R Aufgabe“ diskutiert und die „R Aufgaben“ auch gemeinsam gelöst werden konnten. Der Besuch der Fragestunde war freiwillig und nicht mit weiteren Anreizen (etwa dem möglichen Erwerb zusätzlicher Bonuspunkte für die Klausur o.ä.) versehen. Pro Woche gab es zwei feste Termine für die „R Fragestunde“, zusätzlich konnte auch noch eine individuelle Beratung persönlich oder per Email in Anspruch genommen werden.

Insgesamt wurde bei den „R Aufgaben“ somit eine Binnendifferenzierung auf drei Stufen angeboten (keine technische Unterstützung, Template, Fragestunde), womit insbesondere aus geringeren Vorkenntnissen oder Fertigkeiten beim Programmieren resultierende Lernschwierigkeiten abgefedert werden sollten.

II.B.2. Didaktische Analyse der einzelnen Aufgaben

1. Einstiegsaufgabe: Fairer Münzwurf

Zum Warmwerden mit R. Wir untersuchen das Spiel *Kopf oder Zahl*: Eine faire Münze wird n mal geworfen. Jedes Mal, wenn die Münze 'Kopf' zeigt, gewinnt Peter einen Euro von Paul, und wenn die Münze 'Zahl' zeigt, verliert Peter einen Euro an Paul.

i) Simulieren Sie mit Hilfe von R 10 Runden des Spiels für $n = 40$ und dokumentieren Sie die Gewinne bzw. Verluste von Peter am Ende der Spiele.

ii) Stellen Sie für einen Spielverlauf graphisch dar, wieviel Peter nach jedem Münzwurf gewonnen oder verloren hat.

iii) Wiederholen Sie das Spiel häufig und stellen Sie die Gewinne und Verluste von Peter am Ende der Spiele mit einem Stabdiagramm graphisch dar. Welchen Wert erwartet man am häufigsten?

Das Hauptanliegen dieser Aufgabe ist es, die Studierenden mit Grundlagen des Programmierens und wichtigen Befehlen in R (z.B. Erzeugung von Zufallszahlen, Definition und Gebrauch von Vektoren sowie Funktionen, Verwendung von for-Schleifen und if-Abfragen, Erstellen von Graphiken) vertraut zu machen sowie ihnen erste Erfahrungen mit dem Programm als Hilfsmittel für stochastische Experimente zu ermöglichen. Die Hilfestellungen im Template sind daher sehr ausführlich kommentiert und es bietet auch eine allgemeine Einführung in R. Bei der Aufgabe selbst soll ein explorativer Kontakt mit Simulationen im Vordergrund stehen (A2), der Lernangebote auf unterschiedlichen Ebenen ermöglicht.

Das Spiel, das im Einleitungstext vorgestellt wird, ist denkbar einfach und übersichtlich. Die folgenden Aufgaben beziehen sich nur auf die Perspektive eines Spielers (nämlich Peter), um die Modellierung der Spielergebnisse zu erleichtern. Aus der Sicht eines Spielers kann je Münzwurf zwischen „Gewinn" und „Verlust" unterschieden werden, und diese intuitiv mit den Werten 1 und -1 belegt werden. Im ersten Teil der Aufgabe wird zehnmal eine Kette unabhängiger Bernoulli-verteilter Zufallsvariablen erzeugt, was in R durch den Befehl `sample`, also Ziehen mit Zurücklegen aus zwei Elementen, einfach umsetzbar ist. Die Zufallsvariablen werden dabei mit einem anschaulichen inhaltlichen Bezug zum Spiel versehen, wodurch ihre Bedeutungen unmittelbar einsichtig sind. Die Summe dieser Zufallsvariablen gibt das Endergebnis eines Spiels wieder und ist aus der Sicht des Spielers von höherem Interesse als die Einzelergebnisse der Münzwürfe, die in der Programmvorlage aber auch ausgegeben werden, um das Endergebnis eines Spiels nachvollziehbar zu machen.

Dass die Summe dieser Zufallsvariablen selbst wieder eine Zufallsvariable ist, soll durch die zehnfache Wiederholung verdeutlicht werden. Durch diese Wiederholung wird den Studierenden die Variabilität als wesentliche Eigenschaft stochastischer Experimente vor Augen geführt.

Die Visualisierung eines einzelnen Spielverlaufs im zweiten Teil der Aufgabe soll ebenfalls die Variabilität stochastischer Experimente veranschaulichen – hier bezogen auf die Ergebnisse der einzelnen Münzwürfe innerhalb eines Spieles. Aus Gründen der Übersichtlichkeit wird in dem Template vorgeschlagen, in der Graphik eine waagrechte Linie bei der Null einzufügen, um den Bereich von „Gewinn" und „Verlust" optisch abzutrennen. Es ist möglich, dass Studierende sich in diesem Teil auch schon mit der als „Spielerfehlschluss" („Gambler's Fallacy") bekannten Fehlvorstellung auseinandersetzen, dass das Gesetz der großen Zahl auch „im Kleinen" gelten müsse, also bei unabhängigen Wiederholungen von Zufallsexperimenten länger nicht mehr vorgekommene Ergebnisse mit einer höheren Wahrscheinlichkeit auftreten würden als kürzer zurückliegende. Die Fehlvorstellung wird aber nicht explizit thematisiert.

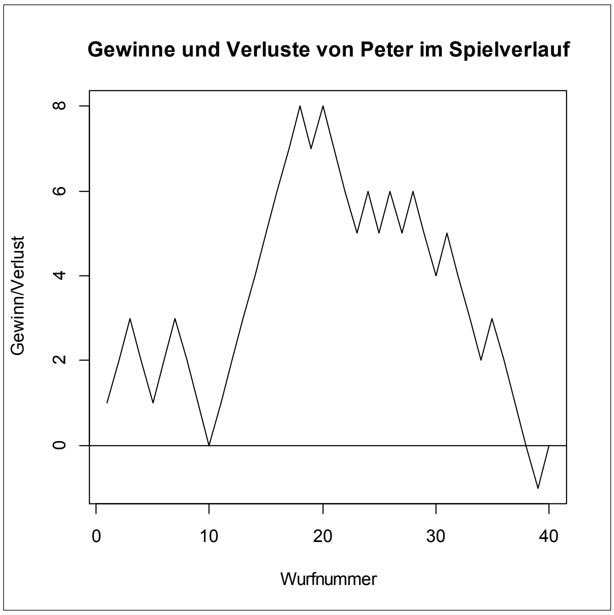


Abb.II.B.1. Mit dem Lösungstemplate simulierter Spielverlauf von Peter (Aufg. 1, ii).

Im dritten Teil der Aufgabe wird die Beobachtung aus Teilaufgabe i) noch einmal aufgegriffen, dass das Spielergebnis selbst eine Zufallsvariable ist. Intuitiv ist klar, dass bei diesem Spiel Peter „im Mittel" nichts gewinnt oder verliert. Die Studierenden sollen in dieser Teilaufgabe diese Intuition bestätigen (A1). Durch die Formulierung „Wiederholen Sie das Spiel häufig […]" soll bei den Studierenden die Frage aufgeworfen werden, wie viele

Wiederholungen für eine zuverlässige Schätzung des – hier nicht näher spezifizierten – Mittelwerts, der in diesem Fall im Modalwert des Säulendiagramms abgebildet wird, nötig sind. Durch Experimentieren mit der Anzahl der Wiederholungen können die Studierenden bereits ein Gefühl dafür entwickeln bzw. ihre Intuition festigen, dass bei häufigeren Wiederholungen der erwartete Wert zuverlässiger vorhergesagt wird.

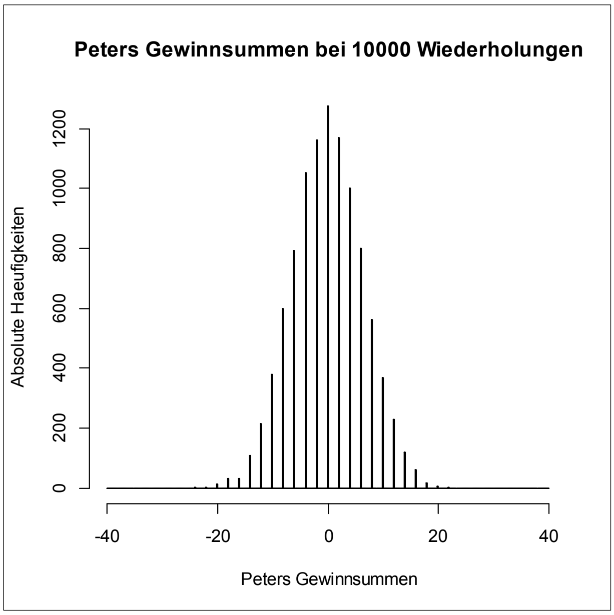


Abb.II.B.2. Mit dem Lösungstemplate simulierte Gewinnsummen von Peter bei 10000 Wdh. (Aufg. 1, iii).

2. <u>Crabs</u>

Das Glücksspiel Crabs geht so: Es werden zwei Würfel geworfen, die Augensumme sei X. Ist X = 7 oder 11, so gewinnt der Spieler, ist X = 2, 3 oder 12, so verliert er. Andernfalls werden die beiden Würfel erneut geworfen, und zwar so oft, bis die erzielte Augensumme wieder gleich X, oder aber 7 ist. Im ersten Fall gewinnt der Spieler, im zweiten Fall verliert er. Wir groß ist die Gewinnwahrscheinlichkeit?

Fortsetzung (R-Aufgabe):

i) Schreiben Sie ein Programm, das den Spielverlauf von Crabs simuliert. Spielen Sie mit diesem Programm Crabs 10 Mal und dokumentieren Sie die Ergebnisse.
ii) Simulieren Sie viele Runden dieses Spiels. Wie hoch schätzen Sie die Gewinnwahrscheinlichkeit des Spielers?
iii) Untersuchen Sie mit einer Simulation: Kann ein Spieler, der 100 Runden Crabs spielt, einen Unterschied zu einem fairen Spiel mit einiger Sicherheit erkennen? Vergleichen Sie die Gewinnhäufigkeiten bei 100 Runden mit Hilfe eines R Programms.

Crabs ist ein (vor allem in den USA beliebtes) Würfelspiel, bei dem mit zwei Würfeln gewürfelt und in jedem Zug jeweils die Augensumme berechnet wird. Beträgt die

Augensumme im ersten Zug 7 oder 11, dann gewinnt der Spieler. Bei den Augensummen 2, 3 oder 12 im ersten Zug verliert der Spieler. In allen anderen Fällen wird so lange erneut geworfen, bis die Augensumme wieder mit der ersten übereinstimmt (dann gewinnt der Spieler) oder 7 beträgt (dann verliert der Spieler). Die Anzahl der Züge variiert somit von Spiel zu Spiel.

Bei dieser „R Aufgabe“, die als „Fortsetzung“ einer Aufgabe auf demselben Übungsblatt eingeleitet wird, wird das Ziel verfolgt, den Studierenden einerseits Simulationen als Instrument zur Bestätigung von theoretischen Ergebnissen (A1), andererseits aber auch als Instrument zur Exploration neuer Fragestellungen (A2) vorzustellen. In der auf dem Übungsblatt vorangegangen Aufgabe sollten die Studierenden nämlich die Gewinnwahrscheinlichkeit beim Würfelspiel Crabs berechnen. Sie liegt bei etwa 49,3%. Durch die Visualisierung einzelner Spielverläufe in Teil i) wird zunächst verdeutlicht, dass die Spielrunden unterschiedlich lange dauern. Außerdem dient sie als Kontrolle, ob das Spiel richtig modelliert wurde, da die Ergebnisse überprüft werden können. Die Modellierung des Spiels kann durch zweifaches Ziehen mit Zurücklegen aus den Zahlen 1 bis 6 erreicht werden und knüpft somit an die Idee aus der vorherigen „R Aufgabe“ zum Münzwurf an. Im Teil ii) sollen sich die Studierenden erneut mit der für Simulationen stets wichtigen Frage auseinandersetzen, wie viele Wiederholungen für eine zuverlässige Schätzung nötig sind. Da es sich bei diesem Aufgabenteil um eine Überprüfung eines berechneten Ergebnisses durch Simulation handelt, liegt hier auch ein Testproblem vor, an dem die Anzahl der Wiederholungen als Einflussfaktor auf die Güte der Schätzung untersucht werden kann. Auf einer grundlegenderen Ebene ermöglicht die Aufgabe zunächst die Erkenntnis, dass Simulationen zu Schätzungen von Größen führen, die aufgrund des eingeführten Zufalls nicht exakt sind, sondern Schwankungen unterliegen. Im dritten Teil der Aufgabe geht es darum, mit Hilfe von Simulationen die Perspektive eines Spielers einzunehmen und zu untersuchen, ob für ihn ein Unterschied zu einem fairen Spiel erkennbar ist. Trotz einer sehr hohen Anzahl von Spielrunden (100 Stück) ist dies kaum möglich, wie die Interpretation entsprechender Säulendiagramme (vgl. Abb.II.B.3) zeigt. Eine alternative Visualisierung dieses Sachverhalts findet sich in Abb.II.B.4., in der zur besseren Übersichtlichkeit gestufte Glockenkurven verwendet werden, die die erwarteten Häufigkeiten der entsprechenden Gewinnanzahlen bei 100 Runden Crabs abbilden. Die beiden Kurven überlagern sich fast – für den Spieler ist eine Unterscheidung zwischen Crabs und einem fairen Spiel somit unmöglich.

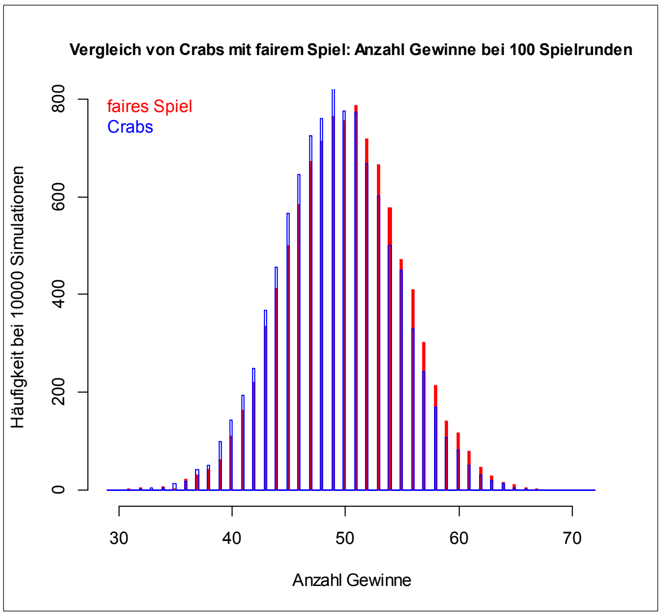


Abb.II.B.3. Mit dem Lösungstemplate erstelltes Säulendiagramm zum Vergleich eines fairen Spiels mit Crabs (Aufg. 2, iii).

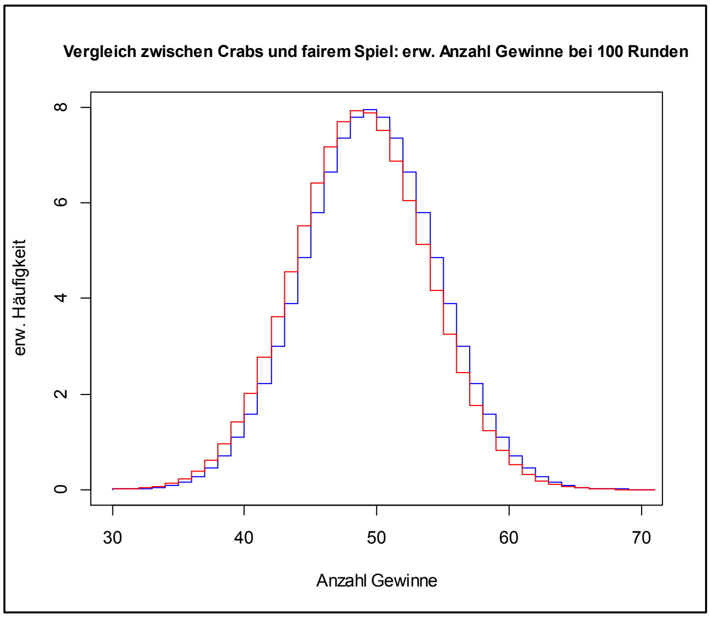


Abb.II.B.4. Eine alternative Darstellung des Sachverhalts mit gestuften Glockenkurven: Hier werden die erwarteten Häufigkeiten der entsprechenden Anzahlen an Gewinnen bei 100 Runden Crabs dargestellt.

Die Simulationen werden bei diesem Aufgabenteil als Hilfsmittel zur Exploration einer neuen Fragestellung verwendet. Im Template wird vorgeschlagen, die Spielergebnisse bei Crabs und einem fairen Spiel direkt mit Hilfe binomialverteilter Zufallsvariablen zu vergleichen. Zur Modellierung können in diesem Fall Binomialverteilungen mit den Erfolgswahrscheinlichkeiten 0.5 und 0.493 zu Grunde gelegt werden, da die Spielverläufe selbst ja für diese Fragestellung nicht von Interesse sind. Dadurch wird bei den Studierenden auch eine inhaltliche Auseinandersetzung mit der Binomialverteilung angeregt. Außerdem bietet die Aufgabe einen Anlass zur Versprachlichung stochastischer Zusammenhänge, da das Ergebnis der Simulation in Worte gefasst werden soll.

Auf der methodischen Ebene dient die Aufgabe in erster Linie einer Festigung der erworbenen Fähigkeiten beim Programmieren (Zufallszahlen mit dem Befehl `sample` erzeugen, Vektoren und Funktionen definieren, if-Abfragen und for-Schleifen verwenden sowie Graphiken erzeugen). Im Template wird neben dem Befehl zur Erzeugung binomialverteilter Zufallsvariablen lediglich eine „while"-Schleife als neues Element eingeführt.

3. Monte Carlo Simulation zur Bestimmung von π

Schätzen Sie π mit Hilfe einer *Monte Carlo Simulation*:
Betrachten Sie dazu den Kreis mit Mittelpunkt $\left(\frac{1}{2}\middle|\frac{1}{2}\right)$ mit Radius $\frac{1}{2}$ im Einheitsquadrat $[0,1]x[0,1]$. Simulieren Sie dann zunächst 10000 Punkte, die uniform auf dem Einheitsquadrat verteilt sind und berechnen Sie die relative Häufigkeit der Punkte, die im Kreis liegen. (Tipp: Die x- und y-Werte der Punkte sind dann auch jeweils uniform auf $[0,1]$ verteilt.)

i) Welchen Schätzer für den Flächeninhalt des Kreises erhalten Sie auf diese Weise?
ii) Welchen Wert für π würde man damit schätzen? (Vergleichen Sie mit der exakten Formel für den Flächeninhalt.)
iii) Wiederholen Sie das Verfahren 100 Mal. In welchem Bereich liegen Ihre Schätzungen für π?
iv) Wieviele Punkte sollte man mindestens simulieren, um π bis auf zwei Stellen hinter dem Komma genau schätzen zu können?

Bei Monte Carlo Simulationen werden Größen, die nicht oder nur sehr schwer berechenbar sind, mit Hilfe übersichtlicher Zufallsexperimente numerisch angenähert (A3). Das Setting, zu dem die Größe gehört, muss dabei selbst nicht unbedingt stochastisch sein. Bei dieser Aufgabe wird eine Monte Carlo Simulation anhand eines Testproblems eingeführt, wodurch die Güte der Methode in diesem Fall untersucht werden kann. Da der Flächeninhalt des Kreises aus dem relativen Anteil der Punkte, die zufällig innerhalb des Kreises lagen, geschätzt wird und in der Formel π als Faktor enthalten ist, ermöglicht diese einfache Simulation durch Umstellen der Formel eine Monte Carlo Schätzung von π. Die Variabilität der Schätzungen wird im Teil iii) untersucht, was in die Frage von Teil iv) mündet, wie viele Punkte denn mindestens simuliert werden müssen, um eine zuverlässige Schätzung von π zu erreichen. Bei dieser Teilaufgabe ist es den Studierenden freigestellt, ob sie die Anzahl zunächst berechnen und ggf. eine Simulation als Kontrolle nutzen (A1) oder eine geeignete Zahl durch Experimentieren herausfinden (A2).

Im zu dieser Aufgabe angebotenen Template werden als Hilfestellungen eine Funktion `abstand` definiert, die den Abstand eines Punktes vom Mittelpunkt des Einheitsquadrates ausgibt, gezeigt, wie man uniform verteilte Zufallsvariablen erzeugt, eine Zählvariable `imkreis` eingeführt und eine if-Abfrage vorgestellt, mit der überprüft werden kann, ob ein bestimmter Punkt im Kreis liegt:

```
# Monte Carlo Methode zur Schaetzung von pi

# Ein Punkt
abstand<-function(x,y)
{
abstand<-sqrt((x-0.5)^2+(y-0.5)^2)
abstand
}

x<-runif(1,min=0,max=1)
y<-runif(1,min=0,max=1)

# im Kreis?
imkreis<-0
if(abstand(x,y)<=0.5){imkreis<-imkreis+1}
```

Wie die Studierenden diese Hilfestellungen nutzen, bleibt ihnen überlassen. Das Template schlägt vor, diese Bausteine zusammensetzen, um eine Funktion `montecarlo` zu schreiben, die *w* uniform auf dem Einheitsquadrat verteilte Punkte erzeugt und zählt, wie viele davon in dem Kreis liegen. Die Anzahl der Punkte soll dabei nicht festgelegt sein, sondern es gibt eine Input-Variable, damit diese Funktion auch noch für Teilaufgabe iv) verwendet werden kann. Die Stellen, an denen die Studierenden etwas einfügen müssen, wenn sie dem Template folgen möchten, sind mit dem auskommentierten Hinweis „`# Hier müssen Sie etwas ergaenzen`" markiert.

Dass der relative Anteil der Punkte im Kreis einen Schätzer für den Flächeninhalt des Kreises darstellt, sollen die Studierenden selbst herausfinden, um Teilaufgabe i) und damit auch ii) lösen zu können:

```
# Wiederhole mit w=10000 Punkten: am besten definiert man hier eine
Funktion mit Input w. Wie bekommt man daraus einen Schaetzer für den
Flaecheninhalt? Errechnen Sie damit einen Schaetzer für pi.
```

```
montecarlo<-function(w)
{
# Hier müssen Sie etwas ergaenzen
}
```

Für die Teilaufgabe iii) wird die Struktur vorgegeben, dass Einträge eines Vektors am besten sukzessive mit einer for-Schleife überschrieben werden. Die R-Befehle `min()`, `max()` bzw. `summary()` werden gezeigt, da sie in den vorherigen Aufgaben nicht vorkamen. Die Schätzung ist bei diesem Verfahren in den meisten Fällen bis auf eine Nachkommastelle genau.

```
# Verfahren 100 Mal wiederholen und Bereich für pi-Schaetzungen
angeben

schaetzerpivektor<-rep(0,100)

for(j in 1:100)
{
# Hier muss etwas ergaenzt werden
schaetzerpivektor[j]<-schaetzerpi
}

# Bereich, in dem Schaetzungen liegen:
min(schaetzerpivektor)
max(schaetzerpivektor)

# oder
summary(schaetzerpivektor)

# Wie genau ist die Schaetzung?
```

Zur Teilaufgabe iv) wird lediglich ein Tipp gegeben, da für diese die bereits entwickelte Funktion `montecarlo` verwendet werden kann, um die nötige Anzahl zu ermitteln. Mit dem Standardfehler, der Standardabweichung des Mittelwerts, $f = \frac{s}{\sqrt{n}}$ lässt sich ermitteln, dass eine Verhundertfachung der Stichprobenanzahl n die Genauigkeit des Schätzers um das Zehnfache erhöht. Bei 1 000 000 Wiederholungen ist die Schätzung in den meisten Fällen auf zwei Nachkommastellen genau, bei mehr Wiederholungen wird die Schätzung aber natürlich noch zuverlässiger. Dieses Ergebnis kann sowohl theoretisch (über den Standardfehler) als auch explorativ (durch Ausprobieren verschiedener Wiederholungsanzahlen und Ausgabe der deskriptiven Kenngrößen wie bei Aufgabenteil ii) ermittelt werden.

```
# Wieviele Punkte braucht man, um bis auf 2 Nachkommastellen genau
zu schaetzen? Das können Sie mit Ihrer Funktion montecarlo
herausfinden oder theoretisch herleiten!
```

4. Schätzung der Belegzeiten von Betten in einem Krankenhaus

In einer Stadt hat ein großes Krankenhaus mit 800 Betten eröffnet. Nach 100 Tagen soll untersucht werden, wie lange die Betten durchschnittlich am Stück belegt sind. An diesem Tag werden alle Patienten notiert, die im Krankenhaus liegen, und nach ihrer Entlassung vermerkt, wie lange sie im Krankenhaus waren.
Wir wollen untersuchen, wie gut diese Methode ist, um zu schätzen, wie lange die Betten im Mittel belegt sind.

i) Angenommen, die tatsächliche Liegezeit der Patienten ist *exp(0.1)*-verteilt. Simulieren Sie für alle 800 Betten Belegzeiten von 100 Patienten (wir nehmen an, dass es keine Unterbrechungen gibt).

ii) Speichern Sie in einem Vektor ab, wie lange die Patienten, die zum Zeitpunkt $t = 100$ in den Betten lagen, diese belegt haben. Was ist der Mittelwert?

iii) Stellen Sie die Belegzeiten der Patienten, die zum Zeitpunkt t = 100 im Bett lagen, mit Hilfe eines Histogramms dar und vergleichen Sie mit einer *exp(0.1)*-Verteilung.

Bei dieser Aufgabe, die in der folgenden Woche fortgesetzt wird, soll ein klassisches Verzerrungsproblem aus der Statistik untersucht werden, illustriert an Belegzeiten von Betten in einem Krankenhaus. Dabei wird davon ausgegangen, dass die Betten durchgehend belegt sind, nach der Entlassung eines Patienten also unmittelbar ein neuer aufgenommen wird. Die Studierenden sollen erkennen, dass die Belegzeiten (pro Patient) verzerrt geschätzt werden (A2), wenn man zu einem festen Zeitpunkt die Liegezeiten der aktuell die Betten belegenden Patienten als Schätzer verwendet, da länger im Krankenhaus verbleibende Patienten eine höhere Wahrscheinlichkeit haben, in diese Studie einzugehen. Um die Verzerrung zu erkennen, sollen Liegezeiten von Patienten simuliert werden (wodurch die tatsächliche Verteilung ja bekannt ist) und mit den durch das oben geschilderte Verfahren in die Studie eingehenden Liegezeiten verglichen werden. Der graphische Vergleich der Dichte einer Exponentialverteilung (Parameter $\lambda = 0.1$) mit einem Histogramm der eingehenden Liegezeiten, die keine repräsentative Auswahl der simulierten Liegezeiten darstellt, erfolgt in Teil iii) der Aufgabe. Die Modellierung der Liegezeiten wird im Teil i) der Aufgabe vorgegeben. Die geschilderte Situation im Krankenhaus ist analog zu Problemlagen, die in der Warteschlangentheorie behandelt werden. Modelliert werden sollen die „Wartezeiten“ der Patienten auf ihre Entlassung, was den Liegezeiten in den Betten entspricht. Als einfachstes Modell bietet sich für diese Klasse von Problemen ein Poisson-Prozess an. Ein Poisson-Prozess ist ein Zählprozess, klassischerweise z.B. für die Modellierung von Atomzerfall eingesetzt. Die Ereignisse, die in dem Poisson-Prozess in diesem Fall „gezählt“ werden sollen, sind die Entlassungen von Patienten aus dem Krankenhaus. Aus der Theorie der Stochastischen Prozesse ist bekannt, dass die Wartezeiten (hier Liegezeiten) bei einem

Poisson-Prozess exponentialverteilt sind.[117] Da lediglich die Liegezeiten von Interesse sind, müssen nur sie simuliert werden, und nicht der zugrunde liegende Poisson-Prozess, was die Aufgabe so weit erleichtert, dass sie auch mit den Kenntnissen der Elementaren Stochastik lösbar wird. Der mathematische Hintergrund dieser Modellierung kann sich den Studierenden noch nicht vollständig erschließen, da hierzu Kenntnisse aus dem Themengebiet der Stochastischen Prozesse notwendig wären. Dennoch kann sie den Studierenden auch ohne dieses Hintergrundwissen plausibel erscheinen: Da die Exponentialverteilung eine stetige Verteilung über den nicht-negativen reellen Zahlen ist, bietet sie sich für die Modellierung von Zeiten an. Die Dichte ist eine fallende Funktion, was für die Modellierung von Liegezeiten ebenfalls sinnvoll ist.

Als erwartete Länge der Liegezeit wird in der Aufgabe durch die Wahl von $\lambda = 0.1$ der Wert von 10 Tagen festgelegt. Die Berechnung des Erwartungswertes einer exponentialverteilten Zufallsvariablen muss bekannt sein, um im Teil ii) eine Verzerrung zu erkennen.[118] Im dritten Aufgabenteil soll die Ursache der Verzerrung, die in einer nicht repräsentativen Auswahl der Liegezeiten begründet ist, erkannt und versprachlicht werden. Erkennbar ist die Verzerrung daran, dass beim Histogramm viel Fläche im Vergleich zur Dichte der Exponentialverteilung nach rechts verschoben ist (vgl. Abb.II.B.5). Eine genauere Analyse der Ursache für die Verzerrung erfolgt aber erst in der nächsten Aufgabe, da dafür sichergestellt sein muss, dass die Studierenden die Verzerrung überhaupt erkennen.

Durch diese Aufgabe wird den Studierenden zudem eine Möglichkeit aufgezeigt, wie Schätzverfahren in der Statistik mit Hilfe von Simulationen im Hinblick auf mögliche Verzerrungen getestet werden können (A1).

Für die methodische Umsetzung der Aufgabe stehen den Studierenden bereits alle wesentlichen Hilfsmittel in R zur Verfügung, einzig der Befehl für das Erzeugen exponentialverteilter Zufallsvariablen wird im Template neu eingeführt. Neben strukturellen Hilfen zur Programmierung enthält das Template auch noch die Hilfestellung, wie die Liegezeit des Patienten, der zum Zeitpunkt t = 100 das Bett belegt hat, aus dem Vektor ermittelt werden kann.

---

[117] Vgl. Götz Kersting & Anton Wakolbinger, Stochastische Prozesse, in Reihe: Mathematik kompakt, hrsg. von Martin Brokate et. al, Basel 2014, 95ff.

[118] Zum theoretischen Hintergrund dieser Verzerrung und Möglichkeiten ihrer Berechnung vgl. Kersting/ Wakolbinger (2014), 100.

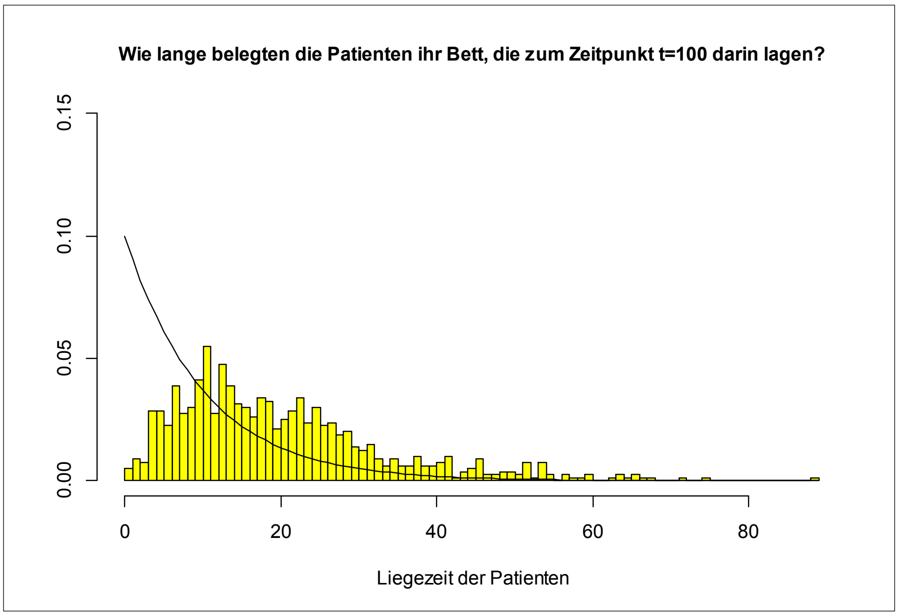


Abb.II.B.5. Mit dem Lösungstemplate hergestellte Graphik zum Vergleich des Histogramms, das sich aus den geschätzten Liegezeiten ergibt, mit der tatsächlichen Dichtefunktion der Exponentialverteilung (Aufg. 4, iii)

5. <u>Fortsetzung der letzten R Aufgabe: Belegzeiten von Betten</u>

Fortsetzung von Aufgabe 16: In einer Stadt hat ein großes Krankenhaus mit 800 Betten eröffnet. Nach 100 Tagen soll untersucht werden, wie lange die Betten durchschnittlich am Stück belegt sind. An diesem Tag werden alle Patienten notiert, die im Krankenhaus liegen, und nach ihrer Entlassung vermerkt, wie lange sie im Krankenhaus waren.
Wir wollen untersuchen, wie gut diese Methode ist, um zu schätzen, wie lange die Betten im Mittel belegt sind.

i) Machen Sie (unter Zuhilfenahme der Ergebnisse von Aufgabe 16) plausibel, warum die oben beschriebene Methode nicht gut ist, um die durchschnittliche Belegzeit der Betten zu untersuchen. Wie kommt es zu der Verzerrung?
ii) Schlagen Sie eine Methode vor, wie man die Belegzeiten unverzerrt schätzen kann und wenden Sie diese mit R auf simulierte Daten für die Belegzeiten an.

Der erste Teil der Aufgabe, der eine unmittelbare inhaltliche Fortsetzung der letzten „R Aufgabe“ darstellt, dient zunächst der Klärung, wie es zu der Verzerrung kommt (A2). Die Studierenden sollen erkennen, dass aufgrund der längeren Liegezeiten Patienten, die mehr Tage im Krankenhaus bleiben als andere, eher ausgewählt werden, da sie mit einer höheren Wahrscheinlichkeit zu dem gewählten Zeitpunkt in einem Krankenhausbett liegen. Dadurch wird die mittlere Belegzeit eines Bettes durch einen Patienten überschätzt. Im Lösungstemplate zu der Aufgabe wurde eine Veranschaulichung dieser Tatsache angeboten (vgl. Abb.II.B.6.).

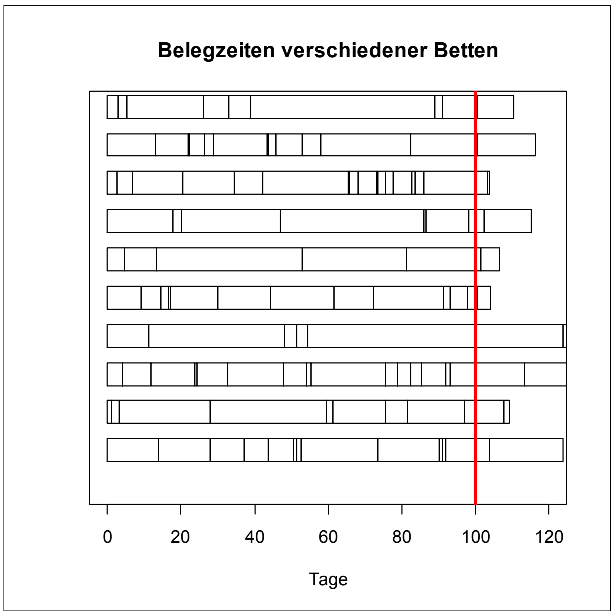


Abb.II.B.6. Mit dem Lösungstemplate hergestellte Graphik zur Veranschaulichung der Verzerrung (Aufg. 5, i).

Um die Verzerrung zu verhindern, muss demnach ein Verfahren gewählt werden, bei dem die Länge der Liegezeit keinen Einfluss auf die Wahrscheinlichkeit hat, für die Studie ausgewählt zu werden. Eine Möglichkeit wäre es, die Liegezeiten der nächsten x Patienten als Grundlage für die Schätzung zu wählen, die ab einem bestimmten Zeitpunkt (z.B. t = 100) im Krankenhaus neu aufgenommen werden. Die Teilaufgabe ii) verlangt von den Studierenden damit eine Transferleistung. Die von ihnen vorgeschlagene Methode können sie selbst testen, indem sie die von ihnen aus den simulierten Daten geschätzte mittlere Liegezeit mit dem tatsächlichen Erwartungswert von 10 Tagen vergleichen (A1). Hierzu findet sich ein Hinweis im Template. Außerdem wird eine analoge Darstellung zur Abbildung Abb.II.B.5 angeboten, mit der die mit der neuen Methode geschätzten Liegezeiten in Form eines Histogramms wiederum mit der Dichte einer entsprechenden Exponentialverteilung verglichen werden können.

Wie ein alternatives Schätzverfahren aussehen könnte, wird im Template dagegen nicht angedeutet, da es ja methodische Hilfestellung im Umgang mit R bieten, nicht aber den inhaltlichen Transferanspruch der Aufgabe aufheben soll.

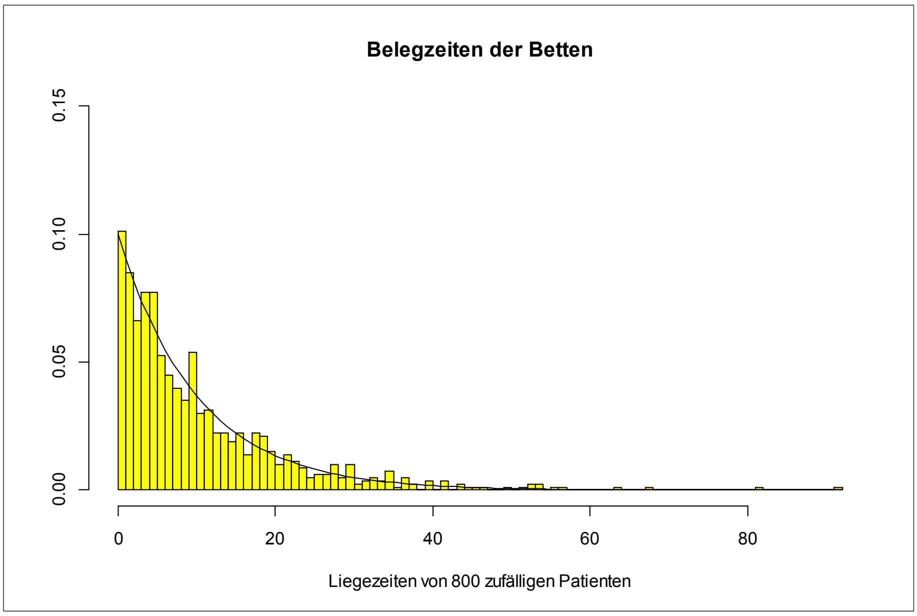


Abb.II.B.7. Mit dem Lösungstemplate hergestellte Graphik zum Vergleich des Histogramms mit der tatsächlichen Dichtefunktion der Exponentialverteilung bei einer unverzerrten Schätzmethode (Aufg. 5, ii).

6. Vorzeichenwechsel bei Münzwürfen

Wir untersuchen die Anzahl Vorzeichenwechsel beim fairen und unfairen Münzwurf.

i) Versuchen Sie das Ergebnis eines 20fachen Münzwurfs zu fälschen, indem Sie einfach eine Folge von *Kopf* und *Zahl* notieren. Eine andere Person sollte die Fälschung möglichst schwierig erkennen können. Notieren Sie eine solche Fälschung und zählen Sie die Vorzeichenwechsel.

ii) Simulieren Sie mit Hilfe von R 3 faire Münzwürfe der Länge 20 und stellen Sie den Verlauf mitsamt der Vorzeichenwechsel graphisch dar. Simulieren Sie anschließend viele faire Münzwürfe und stellen Sie die Anzahlen der Vorzeichenwechsel in einem Stabdiagramm dar.

iii) Vergleichen Sie die Anzahl der Vorzeichenwechsel aus dem von Ihnen gefälschten Münzwurf und notieren Sie, wie häufig eine mindestens so große Abweichung vom Erwartungswert bei den simulierten Münzwürfen vorkam. Wenn Sie selbst einschätzen müssten, ob die von Ihnen notierte Folge eine Fälschung war oder nicht, wie würden Sie sich entscheiden? Begründen Sie Ihre Antwort.

iv) Simulieren Sie nun für $p = 0$, $p = 0.1$, … , $p = 1$ viele p-Münzwürfe der Länge 20 und stellen Sie die daraus geschätzte erwartete Anzahl Vorzeichenwechsel in Abhängigkeit von p graphisch dar. Wie könnte eine Formel für die erwartete Anzahl Vorzeichenwechsel aussehen?

Diese „R Aufgabe“ befasst sich explizit mit der Fehlvorstellung der Ausgeglichenheit (bzw. „Gambler’s Fallacy“) von Zufallsergebnissen und bezieht sich unmittelbar auf eine Aufgabe aus den Fragebögen, die den Studierenden in den Umfragen vor und nach dem Kurs vorgelegt wurden, in der die Studierenden bewerten sollen, welche Münzwürfe sie als „echt zufällig“ und welche als „gefälscht“ ansehen würden (vgl. Kap. II.C.1.).

Die Studierenden sollen die Fehlvorstellung anhand eines von ihnen selbst gefälschten Münzwurfs enttarnen, wobei das Maß an Ausgeglichenheit durch die Anzahl an Vorzeichenwechseln abgebildet wird. Bei einem fairen Münzwurf der Länge 20 liegt die erwartete Anzahl an Vorzeichenwechseln bei 9,5, was dem in Aufgabenteil ii) geforderten Stabdiagramm entnommen oder als Mittelwert mit R berechnet werden kann. Durch die graphische Abbildung einzelner Münzwürfe im Aufgabenteil ii) wird veranschaulicht, dass bei dieser erwarteten Anzahl an Vorzeichenwechseln durchaus längere Folgen von Kopf oder Zahl vorkommen können, die der Fehlvorstellung einer Ausgeglichenheit „im Kleinen“ entgegenstehen (vgl. z.B. Abb.II.B.8.).

Im dritten Teil der Aufgabe sollen die Studierenden ihre Einschätzung des eigenen Münzwurfs versprachlichen, was zu einer Auseinandersetzung mit der eigenen Fehlvorstellung anregen soll. Im Aufgabentext ist hierzu die inhaltliche Hilfestellung enthalten, dass die Studierenden für eine Einschätzung die relative Häufigkeit berechnen können, wie oft eine (mindestens) so große Abweichung vom Erwartungswert wie in ihrem eigenen Beispiel bei den simulierten Münzwürfen vorkommt. Dieses Vorgehen entspricht der Grundidee statistischer Tests und regt zu einer Reflexion über die Frage an, wie viel Variabilität durch Zufall erklärbar ist, und ab wann man „stutzig“ werden sollte, weil auch eine Fälschung vorliegen könnte.

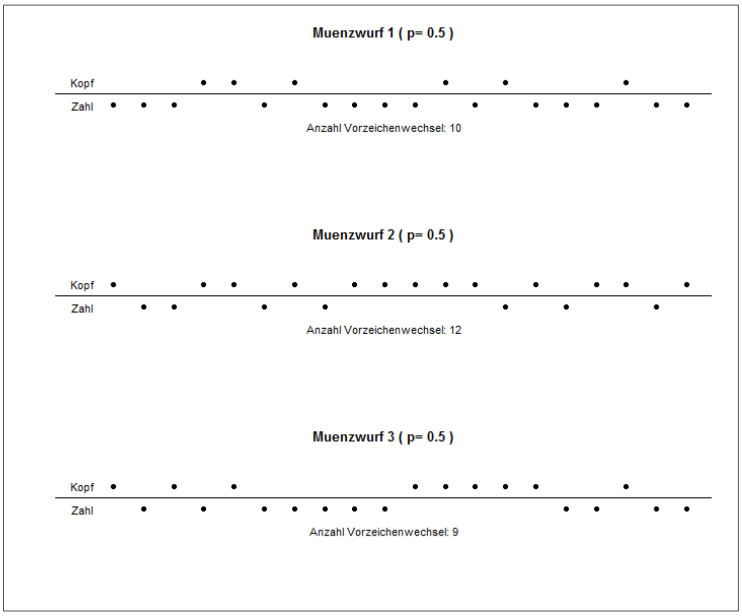


Abb.II.B.8. Mit dem Lösungstemplate simulierte Münzwürfe (Aufgabe 6, ii).

Die Anzahl der Vorzeichenwechsel als Kriterium für die „Echtheit" der Münzwürfe wird im vierten Teil der Aufgabe auf unfaire Münzwürfe übertragen und die Studierenden sollen anhand der graphischen Darstellung die Formel für die erwartete Anzahl an Vorzeichenwechseln bei einem p-Münzwurf der Länge *20* (nämlich $e(p) = p \cdot (1-p) \cdot (20-1) \cdot 2 = 38 \cdot p \cdot (1-p)$) als Hypothese aufstellen, wodurch sie in expliziter Weise mit dem Anwendungsfeld (A2) von Simulationen konfrontiert werden. Die Faktoren p und $(1-p)$ können aus den Nullstellen der Parabel abgelesen werden, der Scheitelpunkt $(0{,}5 \mid 9{,}5)$ kann mit Hilfe des Aufgabenteils ii) abgeleitet werden. Durch Einzeichnen der vermuteten Funktion kann die Hypothese schließlich bestätigt werden (vgl. Abb. II.B.9.).

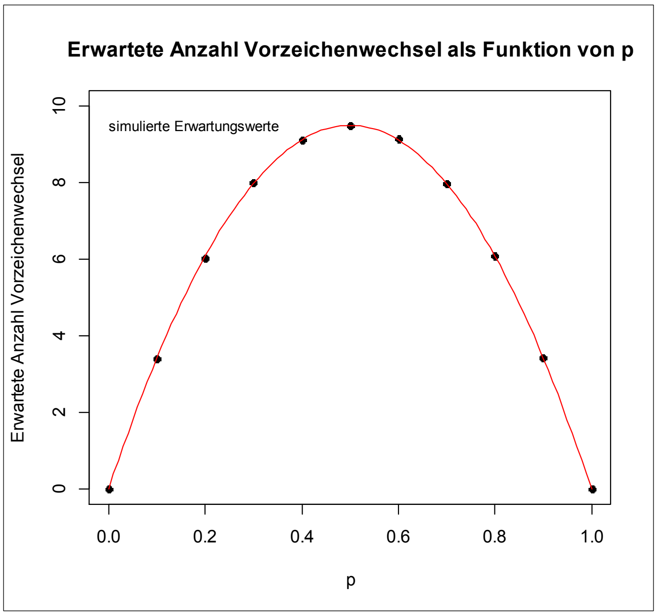


Abb.II.B.9. Mit dem Lösungstemplate erstellte Graphik zur Kontrolle der vermuteten Formel zur Berechnung der erwarteten Anzahl an Vorzeichenwechsel bei einem p-Münzwurf (Aufgabe 6, iii).

7. Regression zur Mitte

Regression zur Mitte. In der Zeitung lesen Sie die Überschrift: *Pfunde verlieren durch Senderwechsel: Übergewichtige, die ihren Radiosender wechseln, verlieren im Mittel ca. 5 kg nach 6 Monaten!*

Der Hintergrund ist die folgende Studie: Es wurden 200 Männer, die mehr als 100 kg wogen, untersucht. Alle gaben an, dass sie beim Autofahren immer einen bestimmten Radiosender hörten. Ihr Gewicht wurde zu Beginn der Studie gemessen, dann sollten sie für 6 Monate den Radiosender wechseln. Ansonsten gab es keine Anweisungen. Nach 6 Monaten wurde ihr Gewicht erneut gemessen. Es stellte sich heraus, dass die Personen im Mittel etwa 5 kg abgenommen hatten. Daraus wurde gefolgert, dass Personen, die abnehmen möchten, nur ihren Radiosender wechseln brauchen.

Sie hegen Zweifel an diesem Ergebnis.

Machen Sie mit Hilfe einer R Simulation plausibel, dass der Wechsel des Radiosenders keinen Einfluss auf die Gewichtsabnahme bei übergewichtigen Personen hat.

Wir nehmen an, dass das Körpergewicht von Männern (annähernd) normalverteilt ist mit Mittelwert 82 kg und Standardabweichung 8. Außerdem gehen wir davon aus, dass das Körpergewicht einer Person vor und nach 6 Monaten korreliert ist ($\kappa = 0.7$). Für jeden Mann ergibt sich so ein korreliertes Paar von Zufallszahlen *(X, Y)*.

i) Simulieren Sie das Körpergewicht von 10000 Männern mit den Vorgaben von oben und stellen Sie die Daten mit einem Streudiagramm dar.
ii) Stellen Sie die Werte x_i der schweren Männer (über 100 kg) in einem Histogramm dar und vergleichen Sie es mit einem Histogramm der Werte y_i. Was können Sie schlussfolgern?
iii) Was sollte man am Aufbau der Studie ändern, um wirklich herausfinden zu können, ob der Wechsel des Radiosenders einen Einfluss auf das Gewicht hat?

Diese „R Aufgabe“ befasst sich erneut mit einem klassischen Problem der Statistik, nämlich der sogenannten „Regression zur Mitte“. Der Hintergrund dieses Verzerrungsproblems ist, dass bei manchen Studien nur die Entwicklung von Individuen verfolgt wird, die bezüglich eines (annähernd normalverteilten) Merkmals weit vom Mittelwert entfernt sind. Es soll durch eine Intervention eine Annäherung an den Mittelwert erreicht werden. Da Merkmalsausprägungen von Individuen aber ja mit höherer Wahrscheinlichkeit näher am Mittelwert liegen als weiter weg, kommt eine Annäherung an den Mittelwert häufig einfach durch Zufall zustande und bildet gar nicht den Erfolg der Intervention ab, was bei einem Vergleich der Ergebnisse der Individuen, die bezüglich des untersuchten Merkmals zu Beginn weit vom Mittelwert entfernt waren, aber nicht ersichtlich ist. So lassen sich etwa klassische Fehlschlüsse zum pädagogischen Wert von Strafen und Loben erklären.[119]

Bei dieser Aufgabe wird eine Schlussfolgerung aus einer (erfundenen) Studie vorgestellt, die offensichtlich nicht stimmen kann. In den ersten beiden Teilaufgaben soll das Ergebnis anhand simulierter Daten rekonstruiert werden, wobei im Aufgabentext eine weitgehende Hilfestellung zum Aufbau der Simulation enthalten ist, da die Idee schwer selbst herzuleiten zu sein scheint und die Aufgabe ohne die simulierten Daten nicht weiter bearbeitet werden könnte. Im Template wird auf das R-Paket „MASS“ verwiesen, mit dem leicht korrelierte normalverteilte Zufallsvariablen erzeugt werden können, aber auch eine Alternative aufgezeigt, wie die Daten simuliert werden können, falls dieses Paket nicht zur Verfügung

[119] Vermeintlich führt Loben guter Leistungen von Schülerinnen und Schüler zu schlechteren Ergebnissen bei der nächsten Wiederholung einer Übung, während Tadeln schlechter Leistungen eine Verbesserung zur Folge hat. Dieser Fehlschluss, der auf dem Phänomen der Regression zur Mitte beruht, wurde bspw. von Kahneman im Zuge einer Psychologie-Vorlesung für Ausbilder der israelischen Luftwaffe aufgedeckt, vgl. Mlodinow (2008), 7f.

steht. Durch den graphischen Vergleich der simulierten Gewichte der schweren Männer vorher und nachher wird das Phänomen der „Regression zur Mitte“, also der rein auf dem Zufall basierenden Verlagerung hin zum Mittelwert, anschaulich gemacht. In der Schlussfolgerung, die im zweiten Aufgabenteil gefordert ist, sollen die Studierenden diese Erkenntnis in Worte fassen und damit den Fehlschluss der Studie erkennen. Durch das Bewusstmachen des Effekts wird die Grundlage für die Bearbeitung des dritten Teils gelegt, in dem die Studierenden einen methodischen Ansatz für eine Studie entwickeln sollen, mit dem zwischen diesem Zufallseffekt und einem tatsächlichen Einfluss einer Intervention auf das Gewicht unterschieden werden kann.

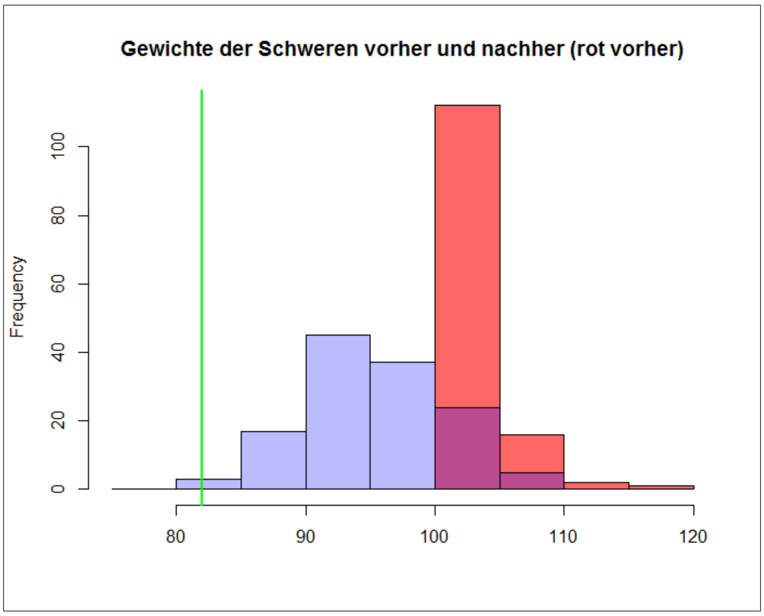


Abb.II.B.10. Mit dem Lösungstemplate erstellte Graphik zur Veranschaulichung des Effekts „Regression zur Mitte (Aufgabe 7,ii).

Der explorative Teil iii) ist sehr offen gestellt und lässt unterschiedliche Zugangsweisen zu, die eigenen Überlegungen sind am simulierten Datensatz aber überprüfbar. Die Studierenden sollen erkennen, dass durch eine Einbeziehung aller Personen aus dem Datensatz die Verzerrung des Ergebnisses vermieden werden kann. Eine andere Möglichkeit wäre, die schweren Personen in eine Kontroll- und eine Interventionsgruppe aufzuteilen und die Ergebnisse beider Gruppen zu vergleichen. Die Aufgabe kann dem Anwendungsbereich (A2) zugeordnet werden.

8. Rejection Sampling und Monte Carlo Simulation

Rejection Sampling. Wir wollen den relativen Flächenanteil des Kugeldreiecks auf der Erdoberfläche bzw. auf der Einheitssphäre S^2 bestimmen, dessen Eckpunkte Frankfurt, Kapstadt und Peking bilden.

Das Dreieck hat das Innere $\Delta = \{\lambda_1 x_1 + \lambda_2 x_2 + \lambda_3 x_3 \in S^2 : \lambda_1,\ \lambda_2, \lambda_3 > 0\}$, wobei die Einheitsvektoren $x_1,\ x_2,\ x_3$ die Positionen von Frankfurt, Kapstadt und Peking auf S^2 angeben. Für $u \in S^2$ gilt $u \in \Delta$, falls die Determinanten $\det(u, x_2,\ x_3)$, $\det(x_1, u, x_3)$ und $\det(x_1, x_2, u)$ alle dasselbe Vorzeichen haben wie $\det(x_1, x_2, x_3)$.

i) Erzeugen Sie uniform verteilte Punkte $U_1, U_2, \ldots, U_n$ auf S^2 nach folgendem Schema: Wähle 10000 V_i uniform in $[-1, 1]^3$. Gilt $|V_i| > 1$, verwerfe den Wert, falls $|V_i| \leq 1$, setze $U_i = V_i / |V_i|$. Wiederholen Sie dieses Verfahren unabhängig. Wieviel Prozent der V_i werden etwa verworfen? Erklären Sie das Verfahren.

ii) Schätzen Sie den relativen Flächenanteil des Kugeldreiecks auf der Erdoberfläche bzw. auf der Einheitssphäre S^2, dessen Eckpunkte Frankfurt, Kapstadt und Peking bilden. Hinweis: Die Koordinaten sind Frankfurt 50° N, 8° O, Kapstadt 34° S, 18° O, Peking 40° N, 116° O. Die Umrechnung in kartesische Koordinaten findet sich im R Code auf der Homepage.

Bei dieser „R Aufgabe" wird die Verwerfungsmethode zur Simulation von Zufallszahlen einer auf direktem Weg schwierig zu erzeugenden Verteilung auf der Grundlage einer anderen, leichter handhabbaren Verteilung veranschaulicht. In diesem Fall sollen uniform auf der Einheitssphäre verteilte Punkte erzeugt werden. Dazu werden Punkte V_i simuliert, die uniform im Würfel $[-1, 1]^3$ verteilt sind, und dann alle verworfen, die nicht innerhalb der Einheitskugel liegen. Die Punkte, die in der Einheitskugel liegen (für die also $|V_i| \leq 1$ gilt), werden dann auf die Oberfläche der Kugel projiziert, wodurch eine uniforme Verteilung auf der Einheitssphäre erreicht wird. Dieses Verfahren, das im ersten Aufgabenteil beschrieben ist, sollen die Studierenden zunächst versprachlichen. Im zweiten Teil der Aufgabe sollen diese Zufallszahlen für eine Monte Carlo Simulation genutzt werden (A3), bei der der relative Anteil des Flächeninhalts eines Kugeldreiecks an der Gesamtfläche einer Sphäre bestimmt werden soll. Da nur der relative Anteil gefragt ist, kann – obwohl es in der Aufgabe eigentlich um ein Kugeldreieck auf der Erdoberfläche geht – die Einheitssphäre verwendet werden, auf der im Teil i) bereits uniform verteilte Zufallszahlen erzeugt wurden. Dies wird in der Aufgabe jedoch nicht explizit erklärt, sondern soll von den Studierenden erkannt werden.

Um eine mögliche Lernschwierigkeit, die durch die Verwendung der geographischen Koordinaten hervorgerufen werden könnte, zu vermeiden, findet sich die Umrechnung der geographischen Koordinaten in kartesische Koordinaten im Template. Eine weitere Hilfestellung wird im Template dadurch gegeben, dass methodische Hinweise zur Erzeugung von auf der Kugelsphäre uniform verteilter Zufallsvariablen enthalten sind, wobei die bereits in der Aufgabe beschriebenen Schritte des Rejection Samplings durch technische Hinweise zur Umsetzung in R ergänzt werden.

9. <u>Fehlstände bei einer rein zufälligen Permutation</u>

Wir wollen die Fehlstände einer rein zufälligen Permutation mit einer R Simulation untersuchen. Schreiben Sie ein Programm, das für beliebiges n eine rein zufällige Permutation der Zahlen $\{1, \dots , n\}$ simuliert, die Anzahl der Fehlstände berechnet, an denen j für $j = 2, \dots , n$ mit einem kleineren Element beteiligt ist, und schließlich die Anzahl aller Fehlstände ausgibt.

i) Wiederholen Sie das oben beschriebene Verfahren w = 10000 Mal für $n = 5$.
Stellen Sie jeweils die Anzahlen der Fehlstände, an denen j, $j = 2, \dots , 5$, in den w Wiederholungen mit einem kleineren Partner beteiligt war, in einem Stabdiagramm dar. Was ist der Bezug zu Aufgabe 35?

ii) Erstellen Sie nun für $n = 5$, $n = 10$ und $n = 15$ Histogramme der Gesamtanzahl der Fehlstände (w = 10000 Wiederholungen). Gegen welche Verteilung scheinen die Gesamtanzahlen zu konvergieren? Hätten Sie das erwartet? Begründen Sie Ihre Antwort.

Diese „R Aufgabe“, bei der ein expliziter Bezug zu einer theoretischen Übungsaufgabe auf demselben Blatt hergestellt wird, bietet zunächst eine Veranschaulichung von in der theoretischen Aufgabe definierten Zufallsvariablen Y_i, $i = 2,\dots,n$. Sie werden dort als uniform verteilt auf $\{0, 1,\dots, i\text{-}1\}$ eingeführt. Die Studierenden sollten in der theoretischen Aufgabe zunächst den Erwartungswert und die Varianz der Zufallsvariablen berechnen. Durch diese „R Aufgabe“ wird ein Anwendungszusammenhang dieser Zufallsvariablen hergestellt, da sie sich als Anzahl der Fehlstände einer Permutation der Zahlen $\{1, \dots , n\}$ interpretieren lassen, an denen i (bzw. in der R-Aufgabe j) mit einem kleineren Partner beteiligt ist. Diese Interpretation sollen die Studierenden im ersten Teil der Aufgabe versprachlichen. Ihre theoretisch berechneten Ergebnisse können die Studierenden nach dem Erkennen des Zusammenhangs durch die Simulation überprüfen, weshalb im ersten Teil der Aufgabe ein Bezug zum Anwendungsfeld (A1) gegeben ist.

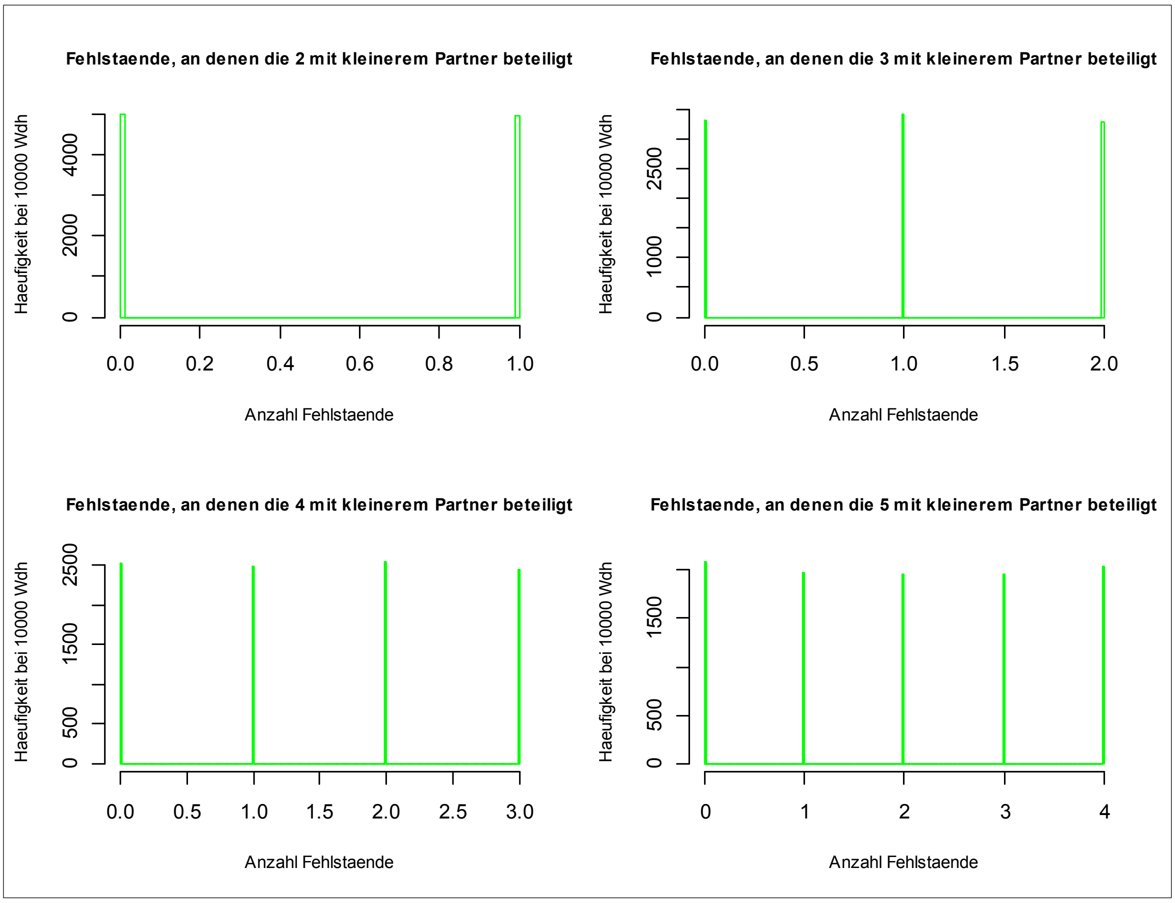


Abb.II.B.11. Veranschaulichung der uniformen Verteilungen der Anzahlen der Fehlstände, an denen j bei einer Permutation der Zahlen $\{1, \ldots, n\}$ mit einem kleineren Partner beteiligt ist (Aufgabe 9,i).

Im zweiten Teil geht es dagegen um eine explorative Anwendung von Simulationen (A2), da eine Hypothese aufgestellt werden soll, gegen welche Verteilung die Gesamtanzahlen der Fehlstände einer Permutation konvergieren. Das Histogramm deutet auf eine Normalverteilung hin, was insofern interessant ist, als es sich zwar um eine Summe unabhängiger Zufallsvariablen handelt (nämlich der Summe der Y_i), die Summanden aber keine identisch verteilten Zufallsvariablen sind. Trotzdem scheint eine Art Zentraler Grenzwertsatz vorzuliegen, was hier auch tatsächlich der Fall ist[120]. Mit diesem Phänomen, das sich vermutlich nicht mit der Erwartung der Studierenden deckt, sollen sie sich im zweiten Teil der Aufgabe sprachlich auseinander setzen. Die Dichte der Normalverteilung kann mit Hilfe der in der theoretischen Aufgabe vorgegebenen Formeln zur Berechnung des Erwartungswerts und der Varianz der Summe der Y_i in das Histogramm eingezeichnet werden, um die Hypothese auch graphisch zu untermauern (vgl. Abb.II.B.12).

[120] Feller (1968), 256f.

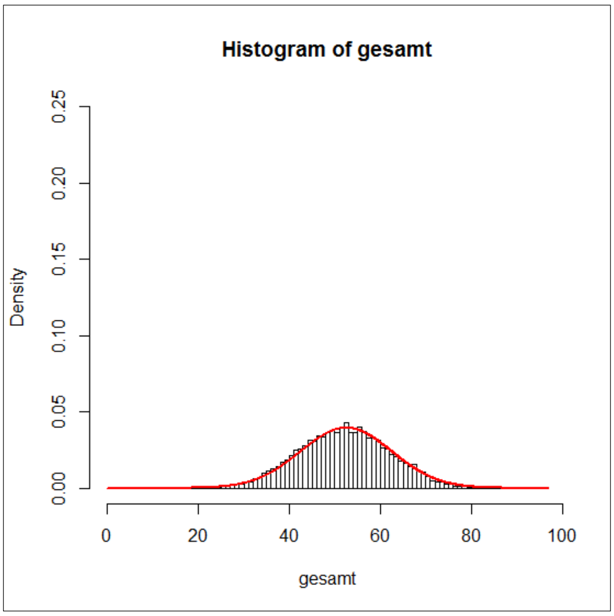


Abb.II.B.12. Mit dem Lösungstemplate erstelltes Histogramm der Gesamtanzahlen der Fehlstände für n=15 mit eingezeichneter Dichte der Normalverteilung (Aufgabe 9,ii).

Das Template leistet bei dieser Aufgabe hauptsächlich Hilfestellungen bei der Erzeugung der Graphiken sowie zum Auffinden von Fehlständen für die Lösung des ersten Teils.

10. <u>Konfidenzintervalle</u>

Konfidenzintervalle. Der Besitzer einer Schokoladenfabrik möchte herausfinden, zu welchen Anteilen er Zartbitter- und Milchschokolade produzieren sollte, um möglichst genau den Geschmack seiner Kunden zu treffen. Dazu ließ er eine Umfrage unter repräsentativ ausgewählten Personen machen, die jeweils angaben, ob sie lieber Zartbitter- oder Milchschokolade mögen.

Sei p der tatsächliche Anteil der Menschen in der Bevölkerung, die Zartbitter-Schokolade vorziehen. In einer Stichprobe der Größe n können die Präferenzen als Zufallsvariablen $X_1, \dots, X_n$ aufgefasst werden, mit $X_i = 1$, falls die befragte Person lieber Zartbitter mag und $X_i = 0$, wenn sie lieber Milchschokolade isst. Wenn n klein ist im Vergleich zur Größe der gesamten Population ist $S_n = \sum_{i=1}^{n} X_i$ annähernd Bin(n,p)-verteilt.

Sei $\hat{p} = \frac{S_n}{n}$ der geschätzte Anteil der Personen, die Zartbitter-Schokolade vorziehen.

i) Simulieren Sie das Ergebnis einer solchen Umfrage für $p = 1/3$ und $n = 100$. Berechnen Sie anschließend mit Hilfe des Zentralen Grenzwertsatzes ein Konfidenzintervall für $\hat{p}$.

ii) Wiederholen Sie das Verfahren anschließend $w = 1000$ mal und stellen Sie die Konfidenzintervalle graphisch dar. Welcher relative Anteil der Konfidenzintervalle überdeckt das wahre p? Wie breit sind die Konfidenzintervalle im Mittel?

iii) Wie hängt die mittlere Breite der Konfidenzintervalle von n ab? Berechnen Sie jeweils $w = 1000$ Konfidenzintervalle für $n = 10, 20, \dots, 10000$ und errechnen Sie für jedes n die mittlere Länge l_n. Tragen Sie anschließend $\log(n)$ auf der x-Achse gegen $\log(l_n)$ auf. Welche Beziehung zwischen n und l_n können Sie daraus ableiten? Was ist die theoretische Begründung dafür?

Bei dieser „R Aufgabe“ steht zunächst eine Veranschaulichung der (in der Theorie bekannten) Eigenschaft von Konfidenzintervallen im Zentrum, dass sie selbst zufällig sind. Sie überdecken daher mit einer gewissen Wahrscheinlichkeit den (in statistischen Anwendungssituationen unbekannten) wahren Parameter, der geschätzt werden soll. In diesem Fall geht es um den Erwartungswert der Personen, die Zartbitter-Schokolade bevorzugen, der durch das arithmetische Mittel in einer Stichprobe geschätzt wird. Da das arithmetische Mittel laut Zentralem Grenzwertsatz approximativ normalverteilt ist, können Konfidenzintervalle berechnet werden, die zu einer festgelegten Wahrscheinlichkeit (üblicherweise 95%) den wahren Wert überdecken. Da in diesem Fall die tatsächliche Wahrscheinlichkeit p, dass eine Person Zartbitter-Schokolade vorzieht, bekannt ist, kann durch die Simulation der Stichprobenergebnisse im Teil i) dieser relative Überdeckungsanteil zunächst bestätigt werden (A1), indem die Lage der aus der Simulation berechneten Konfidenzintervalle mit dem tatsächlichen Wert von p verglichen werden (vgl. Abb.II.B.13.). Im zweiten und dritten Teil wird die Simulation dann noch für einen explorativen Zweck genutzt (A2), nämlich um die Beziehung zwischen der Stichprobengröße n und der mittleren Breite l_n der resultierenden Konfidenzintervalle zu untersuchen, wobei als Hilfestellung die Erstellung eines log-log-Plots gefordert ist, der die Hypothesenbildung erleichtert, da durch die Logarithmierung exponentielle Zusammenhänge linear werden (vgl. Abb.II.B.14). Durch die anschließende theoretische Reflexion sollen die Studierenden das Ergebnis auf die Konstruktion der Konfidenzintervalle zurückführen: Die Breite verringert sich bei steigendem n um einen Faktor der der Größenordnung $\frac{1}{\sqrt{n}}$, es gilt nämlich $l_n = c \cdot n^{-0{,}5}$ mit einer Konstanten $c = 2\sqrt{\hat{p}(1-\hat{p})}$, da die Konfidenzintervalle in diesem Fall durch die (zufälligen) Werte $\hat{p} \pm \sqrt{\hat{p}(1-\hat{p})/n}$ begrenzt sind und somit eine Breite von $2\sqrt{\hat{p}(1-\hat{p})/n}$ besitzen.

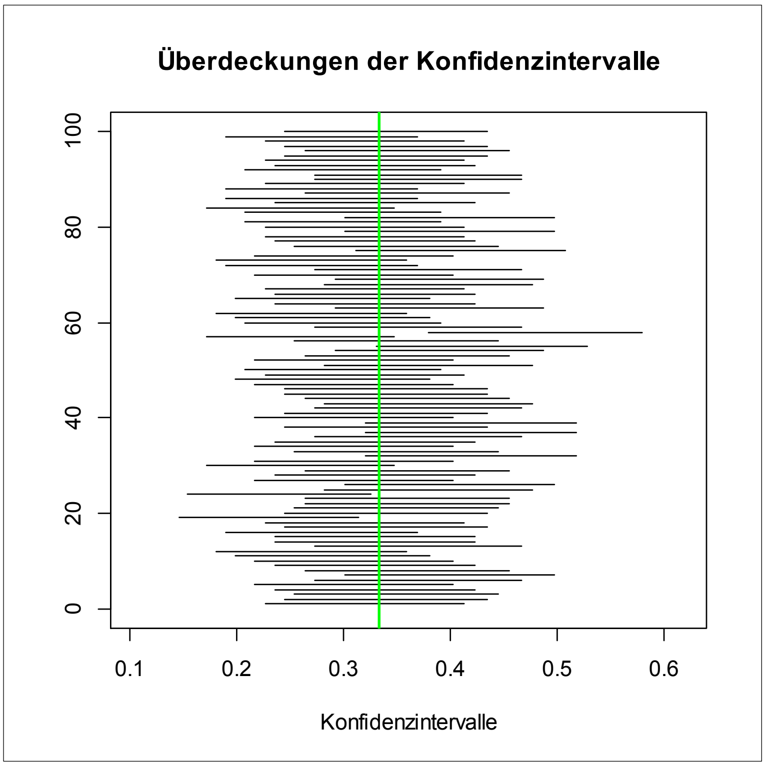


Abb.II.B.13. Mit dem Lösungstemplate erstellte Veranschaulichung des Überdeckungsanteils der aus der Simulation resultierenden Konfidenzintervalle, wobei der wahre Wert p als grüne Linie eingezeichnet ist (Aufgabe 10, ii).

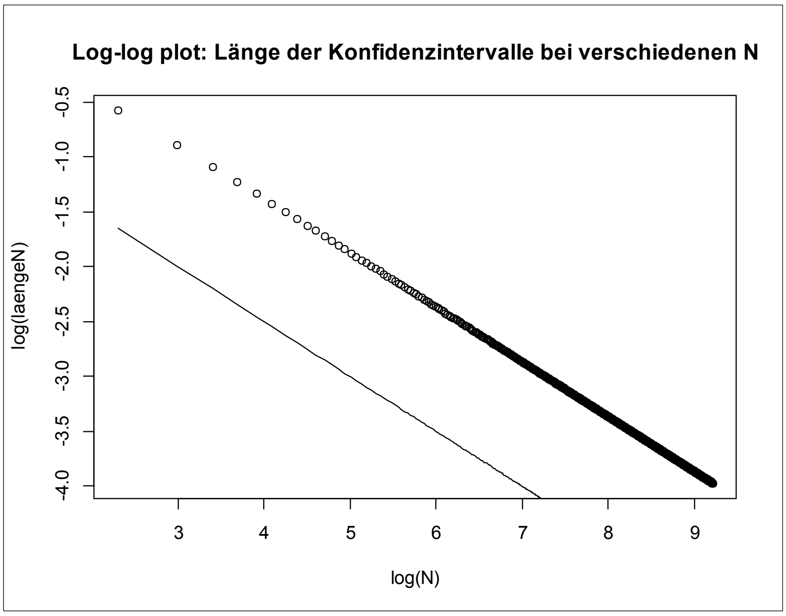


Abb.II.B.14. Mit dem Lösungstemplate erstellter log-log-Plot zur Hypothesenbildung über den Zusammenhang zwischen der Stichprobengröße n und der mittleren Breite l_n der resultierenden Konfidenzintervalle, wobei zusätzlich eine Gerade mit Steigung -0,5 eingezeichnet wurde (Aufgabe 10, iii).

Im Template wird neben technischen Hilfestellungen zum strukturellen Aufbau der Simulation und zur Erstellung der Graphiken für den dritten Teil der inhaltliche Tipp gegeben, die im Log-log-Plot entstehende Funktion mit einer Geraden zu vergleichen, deren Steigung man im Code vorgeben soll, was in der Aufgabenstellung zwar nicht gefordert, für die Hypothesenbildung aber hilfreich sein kann.

11. Bedingte Wahrscheinlichkeiten - Code erklären

Erklären Sie den Code, indem Sie die Fragen (i) bis (vi) beantworten!

```
# Aufgabe: Erklaeren Sie den Code und beantworten Sie die Fragen!


rm(list=ls())

w<-10000

urne<-sample(c("black","blue"),w,prob=c(0.1,0.9),replace=T)
kugel<-rep("help",w)

# i) Was ist bis hierher passiert?

for(i in 1:w)
      {
       if(urne[i]=="black"){kugel[i]<-
sample(c("red","green"),1,prob=c(0.9,0.1))}
       if(urne[i]=="blue"){kugel[i]<-
sample(c("red","green"),1,prob=c(0.1,0.9))}
      }

# ii) Was ist bis hierher passiert?

gruenausblau<-0
rotausschwarz<-0

for(i in 1:w)
      {
       if(kugel[i]=="green" & urne[i]=="blue"){gruenausblau<-
gruenausblau+1}
       if(kugel[i]=="red" & urne[i]=="black"){rotausschwarz<-
rotausschwarz+1}
      }

# iii) Was ist bis hierher passiert?

# Nehmen wir an, die ersten 10 Einträge von "urne" und "kugel"
wären:
# urne[1:10]
# [1] "blue"  "blue"  "blue"  "blue"  "blue"
# [6] "blue"  "blue"  "blue"  "blue"  "black"
```

```
# und
# kugel[1:10]
# [1] "green" "green" "green" "red" "green"
# [6] "green" "green" "green" "green" "red"

# iv) Was waeren dann die Werte von "gruenausblau" und
"rotausschwarz" nach den ersten 10 Schritten der for-Schleife?

gruengesamt<-length(kugel[kugel=="green"])
rotgesamt<-length(kugel[kugel=="red"])

gruenausblau/gruengesamt
rotausschwarz/rotgesamt

# v) Was ist hier berechnet worden?

# Ein Ergebnis war:
# gruengesamt
# [1] 8170
# gruenausblau/gruengesamt
# [1] 0.9892289
# und
# rotgesamt
# [1] 1830
# rotausschwarz/rotgesamt
# [1] 0.4945355

# vi) Erklaeren Sie, warum rotgesamt so viel kleiner ist als
gruengesamt. Stellen Sie sich folgendes Spiel vor: Zunächst wird
aus einer blauen und schwarzen Urne eine mit den
Wahrscheinlichkeiten von oben ausgewählt. Dann wird aus dieser Urne
eine Kugel (mit den Wahrscheinlichkeiten von oben) gezogen und Ihnen
die Farbe gezeigt. Sie müssen raten, aus welcher Urne die Kugel kam.
In welchem Fall ist es schwer und in welchem leicht?
```

Die letzte „R Aufgabe“ des Kurses unterscheidet sich strukturell von den vorherigen, da die Studierenden hier nicht selbst programmieren, sondern einen vorgegeben Code kommentieren und mit Hilfe möglicher Ausgaben interpretieren sollen. Der Hintergrund ist, dass es sich hierbei um eine Klausurvorbereitung handelt, in der auch eine „R Aufgabe“ in Form einer Kommentierung von Code vorkommen sollte, da ein selbstständiges Erstellen von R Code ohne die Möglichkeit der Kontrolle am Computer einerseits als zu anspruchsvoll, andererseits aber vor allem als Testaufgabe eines Stochastik-Kurses als wenig sinnvoll erachtet wurde, da dann ja auch keine Ergebnisse interpretiert hätten werden können, worin sich ein stochastisches Verständnis zeigt. Inhaltlich wird hier ein Spiel mit zwei Urnen simuliert, bei dem bedingte Wahrscheinlichkeiten geschätzt werden sollen. Das Spiel kann z.B. als Modellierung der Situation bei einem medizinischen Test zur Erkennung einer seltenen Krankheit interpretiert werden. In der ersten Stufe des simulierten Zufallsexperiments wird zunächst zufällig eine der beiden Urnen (schwarz oder blau) ausgewählt, aus der im zweiten Schritt zufällig eine Kugel gezogen wird. Die Wahrscheinlichkeiten für die Auswahl der Urnen unterscheiden sich dabei beträchtlich: Mit einer Wahrscheinlichkeit von 10% wird die schwarze und mit einer Wahrscheinlichkeit von 90% wird die blaue Urne ausgewählt. In beiden Urnen befinden sich grüne und rote Kugeln. In der schwarzen Urne beträgt der relative Anteil an roten Kugeln 90%, in der blauen Urne beträgt er dagegen nur 10%. Dem „Spieler" wird anschließend die Farbe der Kugel gezeigt und er muss raten, aus welcher Urne gezogen wurde. Die Simulation lässt bereits die Hypothesenbildung (A2) zu, dass es im Fall einer grünen Kugel leicht ist, die Urne zu erraten (im Beispiel wurde sie zu fast 99% aus der blauen Urne gezogen), im Fall einer roten Kugel jedoch sehr schwer, da hier die Wahrscheinlichkeit für beide Urnen bei etwa 50% lag. Die entsprechenden bedingten Wahrscheinlichkeiten können mit dem Satz von Bayes berechnet werden, was in der Aufgabe aber nicht gefordert ist. Durch die Kommentierung des Codes sollen die Studierenden einerseits Grundkenntnisse im Programmieren mit R nachweisen, andererseits aber auch eine inhaltliche Interpretation der Simulation leisten, indem sie die Struktur des Zufallsexperiments beschreiben und die Ergebnisse im Hinblick auf die Frage, wann das Erraten der Urne für den Spieler einfach und wann schwierig ist, deuten.

12. Klausuraufgabe - Code erklären (Polya-Urne)

Erklären Sie den Code, indem Sie die Fragen (i) bis (iv) beantworten!

```
rm(list =ls())

w<-100
n<-1000

rotinurnevektor <-rep (0,w)

# i) Was ist bis hier passiert?

for (j in 1:w)
        {
         urne <-c("blau" ,"rot" )
         rotinurne <-1

         for (i in 1:n)
                  {
                   zug <-sample (urne ,1,replace =T)
                   urne <-c(urne ,zug )
                   if(zug =="rot" ){rotinurne <-rotinurne+1}
                   }
```

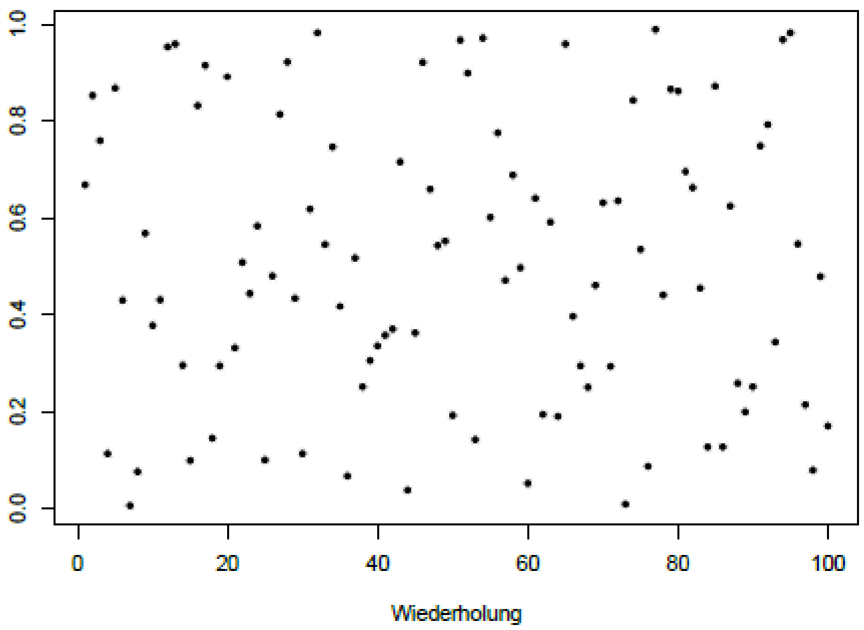
```
        rotinurnevektor [j]<-rotinurne /n
        }

# ii) Was bewirkt die innere for -Schleife?

# iii) Was bewirkt die äußere for -Schleife? Was wird in dem Vektor
rotinurnevektor gespeichert?

plot (1:w,rotinurnevektor ,xlab ="Wiederholung" ,ylab ="",pch =16)

# iv) Was wird in der Graphik abgebildet? Was kann man
schlussfolgern?
```



In Anlehnung an die letzte „R-Aufgabe" des Kurses werden die Studierenden auch in der Klausur aufgefordert, einen vorgelegten Code zu kommentieren, der typische Objekte wie Vektoren, for-Schleifen und if-Abfragen sowie den im Kurs häufig verwendeten Befehl „sample" enthält. Zudem soll durch die Interpretation einer durch das Programm erzeugten Graphik die im Laufe des Kurses zu erwerbende Fähigkeit, Ergebnisse von Simulationen am Computer in Worte zu fassen, überprüft werden.

Durch das Programm wird mit w=100 Wiederholungen eine Polya-Urne simuliert: Pro Wiederholung wird insgesamt n=1000 Mal aus einer Urne mit roten und blauen Kugeln jeweils eine Kugel gezogen und diese Kugel anschließend zusammen mit einer weiteren Kugel derselben Farbe zurückgelegt. Zu Beginn befinden sich pro Wiederholung eine rote und eine blaue Kugel in der Urne. Am Ende jeder Wiederholung wird der relative Anteil der roten Kugeln in der Urne gespeichert. In der Graphik werden diese relativen Anteile der roten Kugeln in der Urne, die in den 100 Simulationen nach jeweils vielen Zügen erzielt wurden, abgebildet. Das Ergebnis legt den Schluss nahe, dass diese relativen Anteile der roten Kugeln zufällig sind, genauer: uniform auf [0,1] verteilt sein könnten.[121]

[121] Der theoretische Hintergrund dieser Aufgabe muss den Studierenden für die Lösung nicht bekannt sein. Tatsächlich lässt sich mit theoretischen Erkenntnissen aus dem Gebiet der Stochastischen Prozesse zeigen, dass der relative Anteil der roten Kugeln in einer Polya-Urne einen zufälligen Grenzwert besitzt und diese Grenzwerte uniform auf [0,1] verteilt sind. Vgl. Kersting/Wakolbinger (2014), 20.

II. C. Explorative Studie zum Nutzen der „R Aufgaben“ anhand von Fragebögen und halbstrukturierten Interviews

Insgesamt wurden in zwei Jahrgängen (SoSe 2011 und SoSe 2012) die Einstellungen zur Stochastik sowie die stochastischen Fertigkeiten zu Beginn und am Ende des Semesters untersucht. Dazu wurden die Studierenden jeweils mit Hilfe von Fragebögen befragt, die in der Vorlesung bzw. den Übungsgruppen ausgefüllt werden konnten. Die Teilnahme war freiwillig und anonym. Um individuelle Veränderungen erfassen zu können, wurde für jeden Teilnehmer ein persönlicher Code erstellt[122]. In die Auswertung wurden nur diejenigen Studierenden einbezogen, die die Fragebögen sowohl zu Beginn als auch am Ende des Semesters ausgefüllt hatten. Die Befragung der Studierenden aus dem SoSe 2011 diente als Vorstudie, da die Veranstaltung „Elementare Stochastik“ in diesem Semester zwar denselben Inhalt wie im SoSe 2012 hatte, jedoch weder in der Vorlesung noch in den Übungen Simulationen am Computer eingesetzt wurden.

Für die Auswertung des Modellprojekts aus dem SoSe 2012 im Hinblick auf die Frage, welchen Nutzen der Einsatz von Simulationen in der Vorlesung und vor allem die Beschäftigung mit den „R Aufgaben“ für die Studierenden hatte, wurden einerseits Fragebögen eingesetzt, um die „R Aufgaben“ zu evaluieren sowie (ebenso wie im SoSe 2011) Einstellungen zur Stochastik und stochastische Fertigkeiten zu Beginn und am Ende des Semesters zu vergleichen, andererseits aber auch halbstrukturierte Interviews geführt. Zunächst sollen die Ergebnisse aus den Fragebögen analysiert werden.

II.C.1. Der Aufbau der Fragebögen und die Methodik ihrer Auswertung

Die Fragebögen zu Beginn des Semesters bestanden neben einem kurzen Einleitungstext aus drei Teilen:

- einem Abschnitt mit Angaben zur Person (Studiengang, Abiturjahrgang, der Frage nach dem Besuch eines Mathematik-Leistungskurses und der Frage, ob die Vorlesung zum ersten Mal gehört wird),

[122] Er bestand aus Teilen des eigenen Vornamens und denen der Eltern sowie der Adresse, wobei bei einem Umzug die Studierenden zu Beginn und am Ende des Semesters denselben Code verwenden sollten.

- einem Aufgabenteil, bei dem die Studierenden nach der Bearbeitung jeder Aufgabe auf einer sechsstufigen Skala ankreuzen sollten, wie sicher sie sich sind, dass ihre Antwort richtig ist,

- einem Teil, in dem die Einschätzung der eigenen Kompetenz im Fach Stochastik sowie Einstellungen zur Stochastik auf sechsstufigen Skalen erfasst wurden.

Da der im SoSe 2011 eingesetzte Fragebogen zu Beginn des Semesters zu viel Zeit für die Bearbeitung in Anspruch genommen hatte, musste er gekürzt werden. So wurde im SoSe 2012 ein Teil, in dem Vorkenntnisse zu Begriffen abgefragt wurden, gestrichen und im Aufgabenteil nur noch 5 der ursprünglich 7 Aufgaben verwendet. Bei den 5 Aufgaben, die beiden Jahrgängen gestellt wurden, handelte es sich in 4 Fällen um Multiple-Choice Aufgaben und in einem Fall um eine Freitext-Aufgabe. Die Aufgaben in den Fragebögen am Ende des Semesters waren bei beiden Jahrgängen identisch.

Um die Ergebnisse der beiden Durchgänge sinnvoll vergleichen zu können, werden im Folgenden nur die Aufgaben ausgewertet, die beiden Jahrgängen gestellt wurden.

Die Fragebögen am Ende des Semesters bestanden neben einem kurzen Einleitungstext aus folgenden Teilen:

- einem Teil mit Angaben weiteren Angaben zur Person (Geschlecht und der Frage, ob zusätzlich auch das Proseminar Stochastik besucht wurde),
- einem Aufgabenteil, wie am Anfang des Semesters kombiniert mit einer sechsstufigen Skala zur Angabe der Sicherheit
- einem Teil, in dem dieselben Einschätzungen der eigenen Kompetenz im Fach Stochastik sowie Einstellungen zur Stochastik auf sechsstufigen Skalen erfasst wurden wie zu Beginn des Semesters,

sowie im SoSe 2012

- einem Teil, mit dem der Einsatz von Simulationen in der Vorlesung und die „R Aufgaben“ auf sechsstufigen Skalen beurteilt werden sollten.

Durch den **AUFGABENTEIL** sollten das Auftreten und ggf. die Auflösung von in der Literatur als typisch beschriebenen Fehlvorstellungen untersucht werden. Die Aufgaben aus diesem Teil wurden unterschiedlichen Quellen entnommen. Der Hintergrund der einzelnen Aufgaben soll im Folgenden skizziert werden. Aufgaben, die in der Vorstudie mindestens ¾ der Befragten bereits zu Beginn richtig beantworteten, wurden am Ende des Semesters nicht mehr gestellt, da bei ihnen geringe Entwicklungen zu beobachten gewesen wären. Das war bei drei der fünf Aufgaben der Fall (*Aufgabe zur Wahrscheinlichkeit von Münzwurfergebnissen, Aufgabe zum Testen eines Würfels* und *Post-Aufgabe*). Am Ende des Semesters wurde dann eine weitere, neue Aufgabe hinzugefügt (*Multiple-Choice-Test-Aufgabe*).

1.) Die ***Krankenhaus-Aufgabe*** wurde zu Beginn des Semesters gestellt und am Ende des Semesters durch die den mathematischen Inhalt betreffend identische ***Grundschule-Aufgabe*** ersetzt:

In einer Stadt gibt es zwei Krankenhäuser. Im größeren werden jeden Tag etwa 45 Kinder geboren, im kleineren etwa 15 Kinder. Wir gehen davon aus, dass etwa 50% aller neugeborenen Kinder Mädchen sind. Der exakte Prozentsatz der Mädchen in den beiden Krankenhäusern variiert aber natürlich von Tag zu Tag.
Für den Zeitraum von einem Jahr wurden in beiden Krankenhäusern jeweils die Tage gezählt, an denen mehr als 60% der Neugeborenen weiblich waren. In welchem Krankenhaus erwarten Sie mehr solche Tage?

a **im kleineren Krankenhaus**
b **im größeren Krankenhaus**
c **in beiden sind etwa gleich viele zu erwarten**

Wie sicher sind Sie, dass Ihr Ergebnis richtig ist?

absolut unsicher	ziemlich unsicher	eher unsicher	eher sicher	ziemlich sicher	absolut sicher
○	○	○	○	○	○

bzw.

In einer Stadt gibt es zwei Grundschulen. In der größeren werden jedes Jahr etwa 100 Kinder eingeschult, in der kleineren etwa 50 Kinder. Wir gehen davon aus, dass etwa 50% aller eingeschulten Kinder Mädchen sind. Der exakte Prozentsatz der Mädchen in den beiden Grundschulen variiert aber natürlich von Jahr zu Jahr.
Für den Zeitraum von 20 Jahren wurden in beiden Grundschulen jeweils die Jahre gezählt, an denen mehr als 60% der eingeschulten Kinder weiblich waren. In welcher Grundschule erwarten Sie mehr solche Jahre?

a **in der der kleineren Grundschule**
b **in der größeren Grundschule**
c **in beiden sind etwa gleich viele zu erwarten**

Wie sicher sind Sie, dass Ihr Ergebnis richtig ist?

absolut unsicher	ziemlich unsicher	eher unsicher	eher sicher	ziemlich sicher	absolut sicher
○	○	○	○	○	○

Die *Krankenhaus-Aufgabe* wurde bereits in unterschiedlichen Versionen in mehreren Studien eingesetzt, wie Sedlmeier und Gigerenzer[123] darlegen. Sie zielt auf die Fehlvorstellung F2(b) an, da es bei dieser Aufgabe im Wesentlichen darum geht, zu erkennen, dass das arithmetische Mittel Schwankungen unterliegt, die abhängig von der Stichprobengröße sind: Der „Standardfehler", definiert als die Standardabweichung in der Grundgesamtheit dividiert durch die Wurzel der Stichprobengröße, ist durch seine Definition abhängig von der Stichprobengröße.[124] Je größer die Stichprobe, desto kleiner erwartet man den Standardfehler, wenn von derselben Grundgesamtheit ausgegangen wird. Anders formuliert: Je größer die Stichproben, desto kleiner die Varianz der berechneten arithmetischen Mittelwerte. Bei kleineren Stichproben wird die Variabilität größer. Da der Erwartungswert des Anteils an Mädchen bei 50% liegt, sollten im kleineren Krankenhaus mehr Tage im Jahr gezählt werden, an denen die arithmetischen Mittelwerte davon (deutlich) abweichen. Die Antwort a ist somit korrekt. Diese Version der Aufgabe zielt auf die Form der Kennwerteverteilungen und hat somit nach Sedlmeier und Gigerenzer einen höheren Schwierigkeitsgrad als beispielsweise die *Post-Aufgabe*, da hierfür die menschliche Intuition, die sich aus Alltagserfahrungen entwickelt, nicht anwendbar sei.[125]

2.) Die ***Aufgabe zur Wahrscheinlichkeit von Münzwurfergebnissen*** wurde den Befragten nur zu Beginn des Semesters vorgelegt:

Sie werfen eine faire Münze 20 mal und notieren „Kopf" als 0 und „Zahl" als 1. Welche der folgenden 01-Folgen ist wahrscheinlicher?
Folge A: 0 0 1 0 0 1 0 1 0 0 0 1 1 1 1 1 0 1 1 1
Folge B: 0 0 0 0 0 0 0 0 0 1 1 1 1 1 1 1 1 1 1 1

a	**Folge A ist wahrscheinlicher.**
b	**Folge B ist wahrscheinlicher.**
c	**Beide Folgen sind gleich wahrscheinlich.**

Wie sicher sind Sie, dass Ihr Ergebnis richtig ist?

absolut unsicher	ziemlich unsicher	eher unsicher	eher sicher	ziemlich sicher	absolut sicher
○	○	○	○	○	○

[123] Vgl. Sedlmeier/Gigerenzer (1997).
[124] Hintergrund ist der Zentrale Grenzwertsatz für Summen von unabhängigen und identisch verteilten Zufallsvariablen: Die arithmetischen Mittelwerte bei wiederholtem Ziehen einer Stichprobe mit festem Umfang aus einer bestimmten Grundgesamtheit sind approximativ normalverteilt, wobei der Erwartungswert dieser Kennwerteverteilung dem Erwartungswert in der Grundgesamtheit und die Standardabweichung dem Standardfehler entspricht.
[125] Vgl. Sedlmeier/Gigerenzer (1997), 37f. und 45ff.

Diese Aufgabe zielt auf die Fehlvorstellung der Ausgeglichenheit/ lokalen Repräsentativität[126] ab (F1), die davon ausgeht, dass bei unabhängigen Münzwürfen unregelmäßige Folgen wahrscheinlicher wären als regelmäßige und dass auch in kurzen Teilsequenzen die relative Häufigkeit für Kopf bzw. Zahl ungefähr der tatsächlichen Wahrscheinlichkeit, also hier 50%, entsprächen. Da die Übertragung des Gesetzes der großen Zahl auf kurze Teilsequenzen jedoch unzulässig ist, ist die Antwort c richtig.

3.) Sowohl zu Beginn als auch am Ende des Semester enthielten die Fragebögen die ***Aufgabe zur Erkennung möglicherweise gefälschter Münzwürfe*** (hier die kürzere Version vom Ende des Semesters):

Eine Person A behauptet, Sie habe 20 mal eine faire Münze[1] geworfen und „Kopf“ als 0 und „Zahl“ als 1 notiert. Sind bei den folgenden 01-Folgen welche dabei, bei denen Sie daran zweifeln, dass sie auf diese Weise zustande gekommen sind? Falls ja, kreuzen Sie diese bitte an.

☐ **1 1 1 0 1 0 0 0 1 1 0 1 1 1 0 0 0 0 1 0**
☐ **0 1 0 0 1 0 1 0 0 1 1 0 1 0 0 1 1 0 1 0**
☐ **1 1 0 0 0 0 0 0 0 1 0 0 0 1 0 0 0 0 0 0**
☐ **1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1**

Ggf. Begründung?

Wie sicher sind Sie, dass Ihre Einschätzung richtig ist?

absolut unsicher	ziemlich unsicher	eher unsicher	eher sicher	ziemlich sicher	absolut sicher
○	○	○	○	○	○

[1] Bei einer „fairen“ Münze ist die Wahrscheinlichkeit, dass sie bei einem Wurf „Kopf“ zeigt, genauso groß wie die Wahrscheinlichkeit, dass sie „Zahl“ zeigt, nämlich jeweils 1/2.

Die Schwierigkeit dieser Aufgabe besteht insbesondere darin, den Unterschied zur vorherigen Aufgabe zu erkennen: In diesem Fall handelt es sich um eine statistische Fragestellung, bei der aus gegebenen Beobachtungen ein Rückschluss auf einen Parameter (nämlich die

[126] Vgl. Kahneman/Tversky (1972), 434.

Wahrscheinlichkeit für „Kopf“ bzw. Zahl“ bei der verwendeten Münze) bzw. auf die Art der Erzeugung erfolgen soll.

Bei drei der vier Reihen scheint eine Schlussfolgerung zuverlässig möglich zu sein: Die erste Reihe erscheint zufällig: Nullen und Einsen kommen etwa gleich oft vor (in diesem Fall sogar jeweils genau 10 Mal) und die Reihe hat 9 Vorzeichenwechsel (die erwartete Anzahl ist 9.5, vgl. „R Aufgabe“ Nr. 6). Die zweite Reihe dagegen scheint durch auffällig viele Vorzeichenwechsel zu ausgeglichen und wird dadurch verdächtig. Bezieht man die von Kahneman und Tversky beschriebene Fehlvorstellung der Ausgeglichenheit/ lokalen Repräsentativität[127] (F1) mit ein, könnte sie auch per Hand gefälscht worden sein. Die dritte Reihe könnte zufällig entstanden sein, auch wenn verhältnismäßig viele Nullen enthalten sind. Es ist aber auch vertretbar, an ihrer Zufälligkeit zu zweifeln. Bei der letzten Reihe ist der Anteil an Nullen und Einsen dagegen sehr auffällig unterschiedlich, was gegen eine faire Münze spricht.

4.) Zur ***Aufgabe zum Testen eines Würfels*** wurden die Studierenden nur zu Beginn des Semesters befragt (hier die Version aus dem SoSe 2012):

Nehmen wir an, Sie haben einen Würfel, bei dem Sie den Eindruck haben, dass er nicht fair[2] ist. Es soll am nächsten Tag das folgende Spiel gespielt werden: Der Würfel wird einmal geworfen. Fällt eine gerade Augenzahl, gewinnt Spieler A. Bei einer ungeraden Augenzahl gewinnt Spieler B. Wie können Sie vor dem Spiel herausfinden, ob Sie lieber A oder B sein möchten?

Wie sicher sind Sie, dass Ihr Vorgehen richtig ist?

absolut unsicher	ziemlich unsicher	eher unsicher	eher sicher	ziemlich sicher	absolut sicher
○	○	○	○	○	○

[2] Bei einem „fairen“ Würfel ist die Wahrscheinlichkeit für alle Augenzahlen 1 bis 6 gleich groß, nämlich jeweils 1/6.

[127] Vgl. Kahneman/Tversky (1972), 434.

Der Hintergrund dieser Aufgabe ist es, zu ermitteln, ob die Befragten (wiederholte) Experimente als mögliches Verfahren zur Schätzung von Wahrscheinlichkeiten ansehen und somit auch eine grundlegende Idee des Anwendungsfeldes (A3) von Simulationen, vgl. Kap. II.B. zur Didaktischen Analyse der „R Aufgaben", besitzen. Personen, denen die Idee zum Einsatz dieses Hilfsmittels nicht zur Verfügung steht, werden bei dieser Aufgabe kein Lösungsverfahren vorschlagen können. Biehler und Maxara fanden bei ihren Untersuchungen der Eingangsvoraussetzungen von Lehramtsstudierenden heraus, dass bei nur etwa 35% der Studierenden diese Idee vorhanden ist. Die von Biehler und Maxara befragten Personen hatten als Studienziele allerdings nur die Lehrämter an Grundschulen, Hauptschulen oder Realschulen.[128] Ob diese Erkenntnis auch auf die Besucher des Kurses „Elementare Stochastik" übertragbar ist, die sich hauptsächlich aus Studierenden mit dem Ziel Bachelor Mathematik oder Gymnasiallehramt zusammensetzen, sollte mit dem Einsatz dieser Aufgabe evaluiert werden.

5.) Die ***Post-Aufgabe*** wurde den Studierenden nur zu Beginn des Semesters vorgelegt:

Mit 18 Jahren müssen sich alle männlichen Amerikaner für die Einberufung zum Militär bei ihrem Postamt registrieren lassen. Dabei wird neben anderen Informationen auch ihre Größe vermerkt. Die nationale Durchschnittsgröße von 18 Jahre alten Männern beträgt 5 Fuß und 9 Inches.
Gestern ließen sich 25 Männer im Postamt A registrieren und 100 im Postamt B. Am Ende des Tages wurden die durchschnittlichen Körpergrößen in beiden Postämtern berechnet.
Von welcher der folgenden Aussagen erwarten Sie, dass sie richtig ist?

a **Die durchschnittliche Körpergröße im Postamt A war näher am nationalen Durchschnitt als die durchschnittliche Größe im Postamt B.**

b **Die durchschnittliche Körpergröße im Postamt B war näher am nationalen Durchschnitt als die durchschnittliche Größe im Postamt A.**

c **Es gibt keinen Grund, anzunehmen, dass die durchschnittliche Größe in einem der beiden Postämtern näher am nationalen Durchschnitt war als im anderen.**

Wie sicher sind Sie, dass Ihr Ergebnis richtig ist?

absolut unsicher	ziemlich unsicher	eher unsicher	eher sicher	ziemlich sicher	absolut sicher
○	○	○	○	○	○

Bei der *Post-Aufgabe* werden anders als bei der *Krankenhaus-* bzw. *Grundschule-Aufgabe* nur die Schätzungen aus zwei unterschiedlich großen Stichproben aus derselben Grundgesamtheit zu Grunde gelegt, sie befasst sich folglich mit der Fehlvorstellung F2(a). Dass die Schätzung

[128] Vgl. Biehler/ Maxara (2005).

aus einer größeren Stichprobe genauer sein sollte als aus einer kleineren, folgt nach Einschätzung von Sedlmeier und Gigerenzer der menschlichen Intuition basierend auf Alltagserfahrungen und sollte daher von den meisten Personen richtig beantwortet werden.[129]

5.) Zur ***Multiple-Choice-Test-Aufgabe*** wurden die Studierenden nur am Ende des Semesters befragt:

Am Ende einer Vorlesung soll ein Multiple-Choice Test geschrieben werden. Test A besteht aus 10 Fragen, Test B aus 20 Fragen. Pro Frage gibt es zwei Antwortmöglichkeiten, von denen genau eine richtig ist. Beide Tests gelten als bestanden, wenn mindestens 60% der Fragen richtig beantwortet sind. Bei welchem der beiden Tests ist die Wahrscheinlichkeit größer, dass ein Prüfling, der bei jeder Frage rein zufällig eine Antwort ankreuzt, besteht?

a **Bei Test A ist die Wahrscheinlichkeit größer.**
b **Bei Test B ist die Wahrscheinlichkeit größer.**
c **Bei beiden Tests ist die Wahrscheinlichkeit gleich groß.**

Wie sicher sind Sie, dass Ihr Ergebnis richtig ist?

absolut unsicher	ziemlich unsicher	eher unsicher	eher sicher	ziemlich sicher	absolut sicher
○	○	○	○	○	○

Diese *Multiple-Choice-Test-Aufgabe*[130] ähnelt inhaltlich der *Post-Aufgabe*, stellt aber ein höheres Anforderungsniveau dar. Der Erwartungswert des Anteils richtiger Antworten liegt bei 50%, wenn der Prüfling rein zufällig ankreuzt. Je größer die Stichprobe ist (also je mehr Aufgaben gestellt werden), desto genauer sollte die Schätzung sein. Die Aufgabe testet also wieder das Vorhandensein der Fehlvorstellung F2(a). Nun erfordert die Lösung der Aufgabe aber noch den Denkschritt, dass zum Bestehen ja sozusagen eine ungenaue Schätzung des Erwartungswerts erforderlich ist (es müssen ja 60% statt der erwarteten 50% richtig geraten werden). Die Chance hierfür ist somit bei einem kürzeren Test größer (Antwort a). Ob die Intuition, die Sedlmeier und Gigerenzer als auf Alltagserfahrungen basierend beschrieben haben[131], flexibel genug ist, um auch in diesem Fall noch zur richtigen Lösung zu führen, ist nicht unmittelbar klar.

[129] Sedlmeier/ Gigerenzer (1997), 45.
[130] Die Aufgabe wurde in dieser Formulierung auch von Biehler und Maxara zur Evaluation der Vorkenntnisse von Lehramtsstudierenden mit dem Ziel Grundschule, Hauptschule oder Realschule eingesetzt, vgl. Biehler/ Maxara (2005).
[131] Sedlmeier/Gigerenzer (1997), 45.

Auf den Aufgabenteil folgte ein Teil, in dem die **Einschätzung der eigenen Kompetenz im Fach Stochastik** sowie **Einstellungen zur Stochastik** auf sechsstufigen Skalen erfasst wurden. Alle Skalen, die in den Fragebögen verwendet wurden, hatten sechs Stufen und somit keine wenig aussagekräftige „Mitte“.[132] In diesem Teil waren die Stufen folgendermaßen bezeichnet:

trifft überhaupt nicht zu	trifft weitgehend nicht zu	trifft eher nicht zu	trifft eher zu	trifft weitgehend zu	trifft vollständig zu
○	○	○	○	○	○

Die Einschätzung der eigenen Kompetenz wurde mit insgesamt 6 Items abgefragt, wobei die Befragten die Stärke ihrer Zustimmung angeben sollten. Die Formulierungen waren außer bei E1 zu Beginn und am Ende des Semesters identisch. Bei E1 und E6 deuten hohe Zustimmungswerte auf eine hohe Einschätzung der eigenen Kompetenz, bei E2, E3, E4 und E7 dagegen auf eine niedrige.

E1: „Mein Vorwissen aus der Schule schätze ich als sehr gut ein.“ / „Mein Wissen in Stochastik schätze ich als sehr gut ein.“

E2: „In Stochastik kommt meistens etwas anderes raus, als man vermutet.“

E3: „Bei Stochastik versagt meine Intuition.“

E4: „Ich kann mir nur sehr schwer vorstellen, was Zufall bedeuten soll.“

E6: „Das Lösen von Stochastik-Aufgaben fällt mir im Allgemeinen nicht schwerer als von Aufgaben aus anderen mathematischen Bereichen.“

E7: „Ich traue meinen Lösungen von Stochastik-Aufgaben im Allgemeinen nicht.“

Zur Untersuchung der Einstellungen zur Stochastik wurden den Befragten jeweils drei Aussagen vorgelegt:

E5: „Ich mag Stochastik.“

E8: „Stochastik ist nützlich.“

E9: „Stochastik ist kein typisch mathematisches Fach.“

[132] Zu Antworttendenzen bei Fragebögen vgl. bspw. Markus Bühner, Einführung in die Test- und Fragebogenkonstruktion, 3. Aufl., München 2011, 116.

Bei der **Hauptstudie** sollte zusätzlich der **EINSATZ VON SIMULATIONEN IN DER VORLESUNG SOWIE VON „R AUFGABEN“** bei den Übungsblättern evaluiert werden. Dazu wurden den Befragten die folgenden Aussagen vorgelegt, die sie auf derselben Skala wie bei dem Teil zur Einschätzung der eigenen Kompetenz im Fach Stochastik und den Einstellungen zur Stochastik bewerten sollten.

R1: „Die Simulationen in der Vorlesung waren hilfreich für ein vertieftes Verständnis der Inhalte.“

R2: „Die Simulationen in den Übungen waren hilfreich für ein vertieftes Verständnis der Inhalte.“

R3: „Ich habe den Eindruck, dass sich durch die Simulationen meine stochastische Intuition verbessert hat.“

R4: „Selbst Programme für Simulationen zu schreiben ist besser als nur Simulationsergebnisse gezeigt zu bekommen.“

R5: „Ohne die R-Fragestunde hätte ich die Aufgaben nicht lösen können.“

R6: „Die R-Aufgaben waren interessant.“

R7: „Ich würde empfehlen, auch weiterhin Simulationen in den Übungen zur Elementaren Stochastik zu verwenden.“

Die Befragten der **Hauptstudie** wurden auf Grundlage ihrer Angaben zur Häufigkeit der Bearbeitung der „R Aufgaben“ sowie des Besuchs der „Fragestunde“ den folgenden Experimentalgruppen zugeordnet:

R12: Personen, die die „R Aufgaben“ nie oder selten bearbeitet haben

R34: Personen, die die „R Aufgaben“ häufig oder immer eigenständig (also ohne regelmäßigen Besuch der Fragestunde) bearbeitet haben

Frage: Personen, die die Fragestunde oft oder immer besucht haben

Die **AUSWERTUNG DER FRAGEBÖGEN** erfolgte mit Hilfe des Programms R (Version 3.2.2), wobei ein explorativer Ansatz gewählt wurde, d.h. bei allen angegebenen p-Werten wurde keine Anpassung für multiples Testen vorgenommen. Neben deskriptiven Statistiken (Prozentangaben, Mittelwerte und Standardfehler) wurden t-Tests und Fisher-Tests[133] durchgeführt.

Für die Angaben auf Skalen wurden Mittelwerte und Standardfehler berechnet, obwohl es sich um ordinale Datenformate handelt. Da durch die Angabe von Mittelwerten und Standardfehlern (die im Folgenden jeweils auf die zweite Dezimale gerundet wurden) aber auch kleinere Veränderungen in den Gruppen besser sichtbar werden als durch die Verwendung des Medians, wurde diese Darstellung bevorzugt. Bei den Angaben auf Skalen wurden die Mittelwerte zwischen und innerhalb der Gruppen mit t-Tests[134] verglichen, wobei jeweils angegeben wird, ob es sich um gepaarte oder ungepaarte Vergleiche handelt. Gepaart heißt in diesem Fall, dass die Mittelwerte der individuellen Differenzen gegen Null getestet wurden. Daher wurden Probanden, die mindestens einmal (also am Anfang oder am Ende des Semesters) zu einem Item keine Angabe gemacht hatten, aus dieser Analyse des entsprechenden Items ausgeschlossen.

Für die Wahl von (üblicherweise mit hoher Teststärke verbundenen) t-Tests auch im Falle ordinaler Daten spricht, dass sich diese als recht robust gegenüber Verletzungen der Normalverteilungsannahme erweisen und zudem – anders als Wilcoxon-Tests[135], die die Alternative darstellen würden – die betrachteten Mittelwerte in ihre Teststatistik einbeziehen.[136] Um jedoch zu verhindern, dass aufgrund unzulässiger Anwendung dieser Tests falsche statistische Schlüsse gezogen werden, wurde die Normalverteilungsannahme bei den untersuchten Daten mit Hilfe von Kolmogorov-Smirnov-Tests[137] überprüft und in den Fällen, in denen entweder eine starke Verletzung dieser Annahme angezeigt wurde (hochsignifikanter p-Wert von weniger als 1%) oder aber ein p-Wert von unter 5% im Kolmogorov-Smirnov-Test zusammen mit einem signifikanten Ergebnis bei der

---

[133] Zu den statistischen Hintergründen von t-Tests und Fishers exaktem Test vgl. bspw. Kersting/ Wakolbinger (2008), 122ff.

[134] Um auch in Fällen ungleicher Varianzen in den Stichproben adäquate Resultate zu liefern, führt R in der hier verwendeten Standardeinstellung bei t-Tests die Welch-Korrektur aus. Vgl. hierzu bspw. Rainer Schlittgen, Das Statistiklabor. R leicht gemacht, 2. akt. Aufl., Berlin Heidelberg 2009, 153.

[135] Wilcoxon-Tests existieren ebenso wie auch t-Tests für gepaarte (verbundene) und ungepaarte (unverbundene) Stichproben. Im Fall von ungepaarten Stichproben ist auch die Bezeichnung „Mann-Whitney-U-Test“ gebräuchlich. Vgl. hierzu bspw. Sheldon M. Ross, Statistik für Ingenieure und Naturwissenschaftler (übersetzt von Carsten Heinisch), 3. Aufl., München 2006. 465 und 469ff.

[136] Vgl. zur Robustheit und Teststärke von t-Tests bspw. Bortz (2005), 141 und 144f.

[137] Vgl. Sheldon (2006), 454ff.

Durchführung eines t-Tests, in dem die entsprechenden Daten verwendet wurden, vorlag, zusätzlich ein Wilcoxon-Test angewendet. Sowohl die Ergebnisse der Kolmogorov-Smirnov-Tests als auch die der Wilcoxon-Tests befinden sich im Anhang. Mit den Kolmogorov-Smirnov-Tests konnten in den meisten Fällen keine (starken) Verletzungen der Normalverteilungsannahme festgestellt werden. Bei den Vergleichen, bei denen zusätzlich Wilcoxon-Tests durchgeführt wurden, kam es in keinem Fall zu einem beachtenswert anderen Ergebnis als bei dem entsprechenden t-Test.

II.C.2. Deskription der Studierenden, die an den Umfragen teilgenommen haben

Im SoSe 2011 haben insgesamt 44 Studierende an den Umfragen zu Beginn und am Ende des Semesters teilgenommen, im SoSe 2012 waren es 91 Studierende. Studierende, die sich nur an einer der beiden Umfragen beteiligt haben, wurden aus der Analyse ausgeschlossen. Studierende, die einen Fragebogen unvollständig ausgefüllt hatten, wurden nur bei Untersuchungen zu individuellen Änderungen ausgeschlossen, wenn für die entsprechende Variable nicht beide Angaben (vom Beginn und vom Ende des Semesters) vorlagen.

II.C.2.a. Vorstudie (SoSe 2011)

Die Mehrheit der Personen, die im SoSe 2011 an den Umfragen teilgenommen hatten, studierte das Fach Mathematik mit dem Ziel Bachelor (n=22) oder Gymnasiallehramt (n=19), und nur ein kleiner Anteil ein anderes Fach wie Informatik (n=2) oder Chemie (n=1). Die meisten Befragten waren weiblich (28 weiblich, 16 männlich). Die Abiturjahrgänge reichten von 1994 bis 2010, wobei der Median bei 2009 lag (1. Quartil: 2008, 3. Quartil: 2010). Von den Befragten hatten etwa 68% in der Schule einen Mathematik-Leistungskurs besucht. Für etwa 86,4% der Teilnehmer handelte es sich um die erste Belegung des Kurses „Elementare Stochastik". Fünf der Befragten (etwa 11%) besuchten zusätzlich das Proseminar Stochastik.

II.C.2.b. Hauptstudie (SoSe 2012)

Die meisten der Studierenden, die im SoSe 2012 an den Umfragen teilgenommen hatten, studierten das Fach Mathematik mit dem Ziel Gymnasiallehramt (n=43) oder Bachelor (n=38), gefolgt von den Fächern Informatik (n=3), Physik (n=3), Wirtschaftspädagogik (n=2), Meteorologie (n=1) und Mathematik als Doppelstudium Bachelor/Lehramt (n=1).

Alle drei Experimentalgruppen waren unter den Befragten vertreten, wobei die Mehrheit der Befragten sich regelmäßig mit den „R Aufgaben“ beschäftigt hatten: Von den 91 Studierenden gaben 23 Studierende an, die „R Aufgaben“ selten oder nie bearbeitet zu haben, 36 Studierende hatten die „R Aufgaben“ eigenen Angaben zur Folge häufig oder immer selbstständig bearbeitet und 32 Studierende gaben an, regelmäßig oder immer die „R Fragestunde“ besucht zu haben. Alle Personen, die die Fragestunde oft oder immer besucht hatten, gaben bis auf eine Ausnahme an, die „R Aufgaben“ oft oder immer bearbeitet zu haben. Für die folgenden Auswertungen wird die erste Gruppe mit „R12“ bezeichnet, die zweite Gruppe mit „R34“ und die letzte Gruppe mit „F“. Die Zusammensetzungen der Gruppen im Hinblick auf die Studiengänge (vgl. Abb.II.C.1 und Tab.II.C.1) kann dabei als zufällig angesehen werden (Fisher‘s exakter Test, p=0.1175).

Insgesamt waren etwa gleich viele Befragte weiblich wie männlich (45 weiblich, 46 männlich), wobei die Geschlechterverhältnisse in den unterschiedlichen Experimentalgruppen unterschiedlich waren: In den Gruppen R12 und R34, also jenen, in denen die Befragten die „R Aufgaben“ selten oder nie (R12) oder aber immer oder oft eigenständig (R34) gelöst hatten, war die Mehrheit der Befragten männlich (R12: 14 männlich, 9 weiblich; R34: 26 männlich, 10 weiblich). In der Gruppe Frage, also jener, in der die Befragten die Fragestunde oft oder immer besucht hatten, war die überwiegende Mehrheit weiblich (26 weiblich, 6 männlich). Das Geschlechterverhältnis in dieser Gruppe unterschied sich signifikant von dem in den anderen beiden Gruppen (Fisher’s exakter Test: R12-Frage: p=0.001933; R34-Frage: p=1.107e-05, R12-R34: p=0.4026). Der Median des Abiturjahrgangs der Befragten lag bei 2010 (1. Quartil: 2009, 3. Quartil: 2011), die Jahrgänge reichten von 1988 bis 2011. Von den Befragten hatten etwa 73% einen Mathematik-Leistungskurs besucht. Etwa 86,8% der Teilnehmer belegten den Kurs „Elementare Stochastik“ zum ersten Mal. Von denjenigen, die den Kurs wiederholten, besuchte die Hälfte die Fragestunde, knapp die Hälfte bearbeitete die „R Aufgaben“ selten oder nie und nur eine Person löste die „R Aufgaben“ oft oder immer eigenständig. Von allen Befragten besuchten 12 Studierende (etwa 13%) zusätzlich das Proseminar Stochastik.

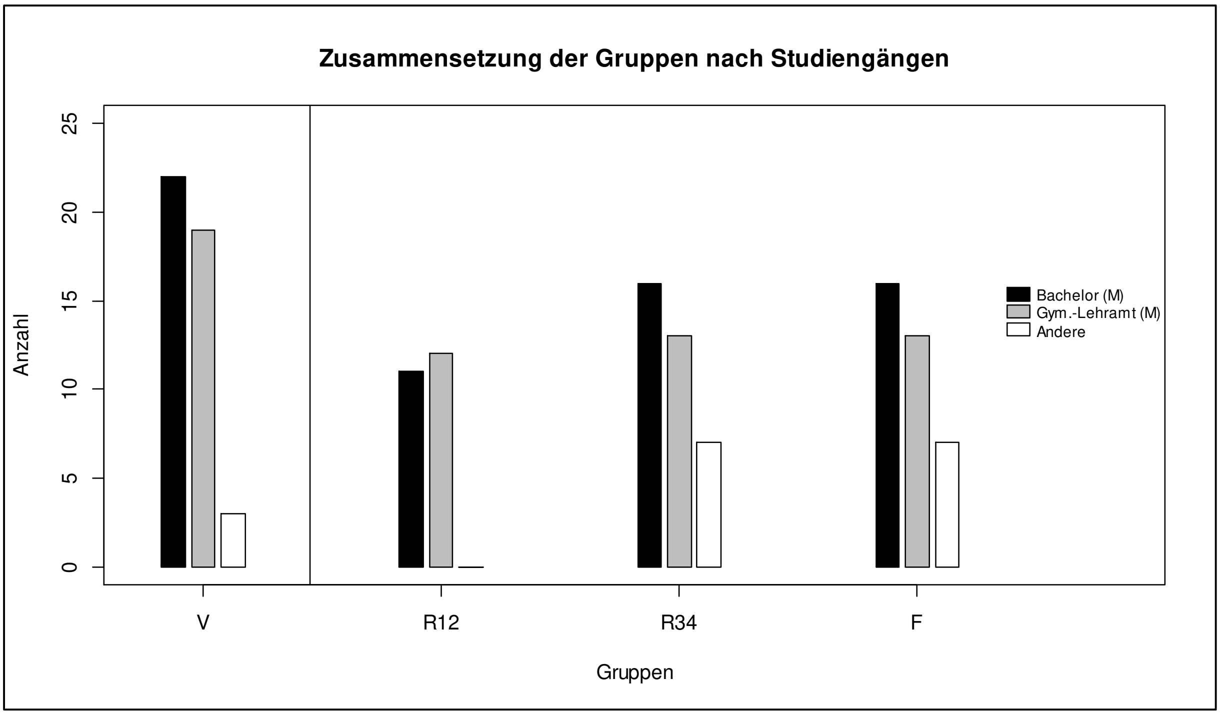


Abb.II.C.1. In der Graphik werden die Zusammensetzungen der Gruppe aus der Vorstudie aus dem SoSe 2011 (n=44) sowie der drei Gruppen aus dem SoSe 2012, „R12“ (n=23), „R34“ (n=36) und „F“ (n=32), im Hinblick auf die Studiengänge Bachelor Mathematik (schwarz), Gymnasiallehramt Mathematik (grau) und Andere (weiß) abgebildet.

	Vorstudie	R12	R34	Frage
Bachelor Mathematik	22	11	16	11
Gymnasiallehramt Mathematik	19	12	13	18
Andere	3	0	7	3

Tab.II.C.1. In der Tabelle finden sich die Gruppenzusammensetzungen im Hinblick auf die Studiengänge in absoluten Zahlen.

II.C.3. Auswertung der Vorstudie (SoSe 2011)

II.C.3.a. Stochastische Fertigkeiten

Die ***„Krankenhaus-Aufgabe“*** zu Beginn des Semesters beantworteten etwa 34% der Befragten richtig; die dazu äquivalente ***„Grundschule-Aufgabe“*** am Ende des Semesters lösten etwa 39% der Befragten korrekt (vgl. Abb.II.C.2). Beide Aufgaben wurden von allen Studierenden bearbeitet.

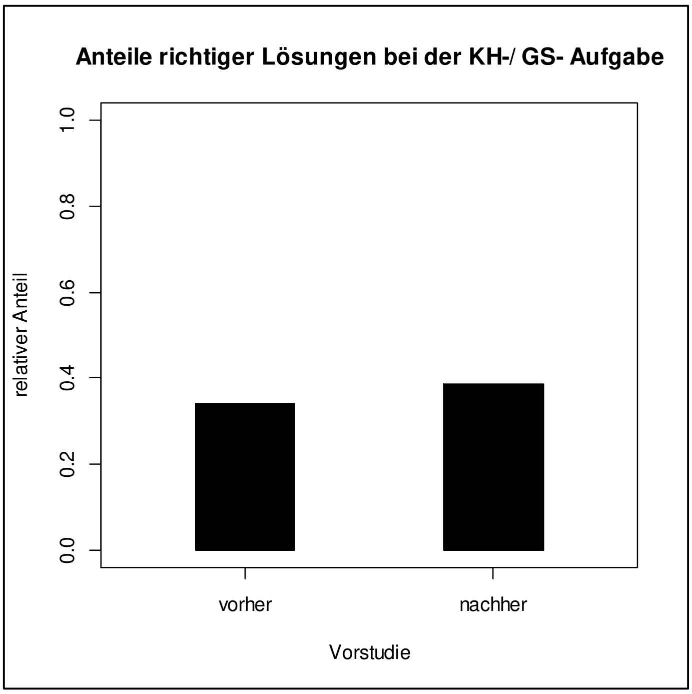


Abb.II.C.2. In der Graphik werden die relativen Anteile richtiger Lösungen der *Krankenhaus-Aufgabe* („vorher“) und der *Grundschule-Aufgabe* („nachher“) abgebildet. Es gab keinen signifikanten Unterschied.

Etwa 70% der Befragten änderten nichts an ihrem Antwortverhalten. Von denen, die die Aufgabe zu Beginn des Semesters richtig gelöst hatten, beantworteten 80% auch am Ende des Semesters die Aufgabe korrekt. Von denen, die die Aufgabe zu Beginn des Semesters falsch gemacht hatten, lösten sie auch am Ende etwa 83% falsch. Die absoluten Zahlen der gegebenen Antworten finden sich in Tab. II.C.2. Zwischen dem Anfang und dem Ende des Semesters gab es keinen signifikanten Unterschied (Fisher’s exakter Test, $p=0.8248$). Die meisten Befragten gaben – der Fehlvorstellung F2(b) folgend – die Antwort c).

		Grundschule		
		Lösung a (richtig)	Lösung b (falsch)	Lösung c (falsch)
Krankenhaus	Lösung a (richtig)	12	0	3
	Lösung b (falsch)	0	3	4
	Lösung c (falsch)	5	1	16

Tab.II.C.2. Die Tabelle bildet die Lösungen der Krankenhaus- und Grundschule-Aufgaben ab.

Zu Beginn des Semesters (*Krankenhaus-Aufgabe*) gab es keinen auffälligen Unterschied bei der Einschätzung der Sicherheit zwischen richtigen und falschen Lösungen. Am Ende des Semesters gaben die Befragten jedoch im Mittel bei richtiger Lösung eine signifikant höhere Sicherheit an als bei falscher Lösung. Zwischen den Zeitpunkten unterschieden sich die Sicherheiten weder bei richtiger noch bei falscher Lösung auffällig. (Vgl. Abb.II.C.3 und Tab.II.C.3)

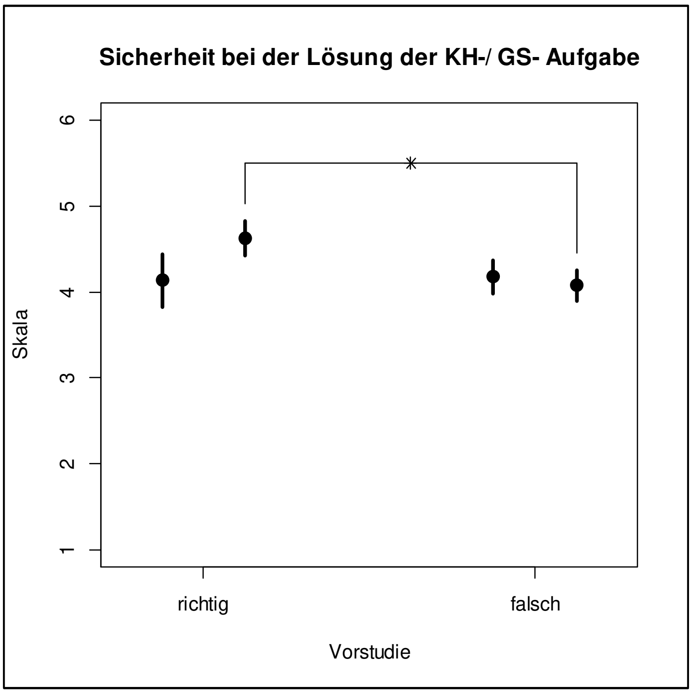


Abb.II.C.3. In der Graphik werden die Sicherheiten, die die Befragten für ihre Lösungen angegeben haben, getrennt nach der Korrektheit der Lösungen abgebildet. Jeweils auf der linken Seite befindet sich der Mittelwert mit Standardfehler der Sicherheit bei der *Krankenhaus-Aufgabe* und rechts der Mittelwert mit Standardfehler der Sicherheit bei der *Grundschule-Aufgabe*. Signifikante Unterschiede sind mit einem Stern markiert.

	Richtig KH	Richtig GS	Falsch KH	Falsch GS
MW	4.13	4.63	4.18	4.07
Stdf.	0.31	0.20	0.20	0.18
Anz.	15	16	28	27

	p-Werte
Richtig KH – Richtig GS	0.1925
Falsch KH – Falsch GS	0.6909

	p-Werte
Richtig KH – Falsch KH	0.9016
Richtig GS – Falsch GS	0.04724

Tab.II.C.3. Mittelwerte (MW) und Standardfehler (Stdf.), jeweils auf die zweite Dezimale gerundet, sowie Anzahlen der Befragten (Anz.) bezüglich der Sicherheiten bei richtiger bzw. falscher Antwort zur *Krankenhaus*(KH)-/*Grundschule*(GS)-*Aufgabe* (sofern Angaben gemacht wurden). Außerdem Vergleiche der Mittelwerte bezüglich der Sicherheiten: p-Werte der t-Tests (ungepaart).

Die ***Aufgabe zur Wahrscheinlichkeit von Münzwurfergebnissen*** wurde nur zu Beginn des Semesters gestellt. Sie wurde von 75% der Befragten richtig beantwortet (vgl. Abb.II.C.4). Alle Befragten gaben eine Lösung an.

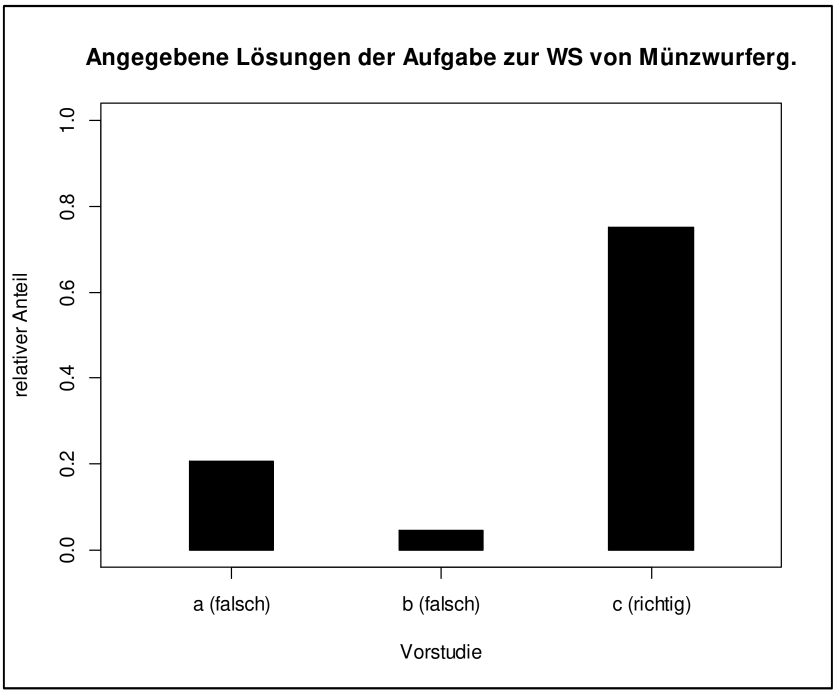


Abb.II.C.4. In der Graphik werden die relativen Anteile der angegebenen Lösungen der *Aufgabe zur Wahrscheinlichkeit von Münzwurfergebnissen* abgebildet.

Die Sicherheiten, die für die eigene Lösung angegeben wurden, unterschieden sich nicht signifikant zwischen richtigen und falschen Ergebnissen (vgl. Abb.II.C.5 und Tab.II.C.4).

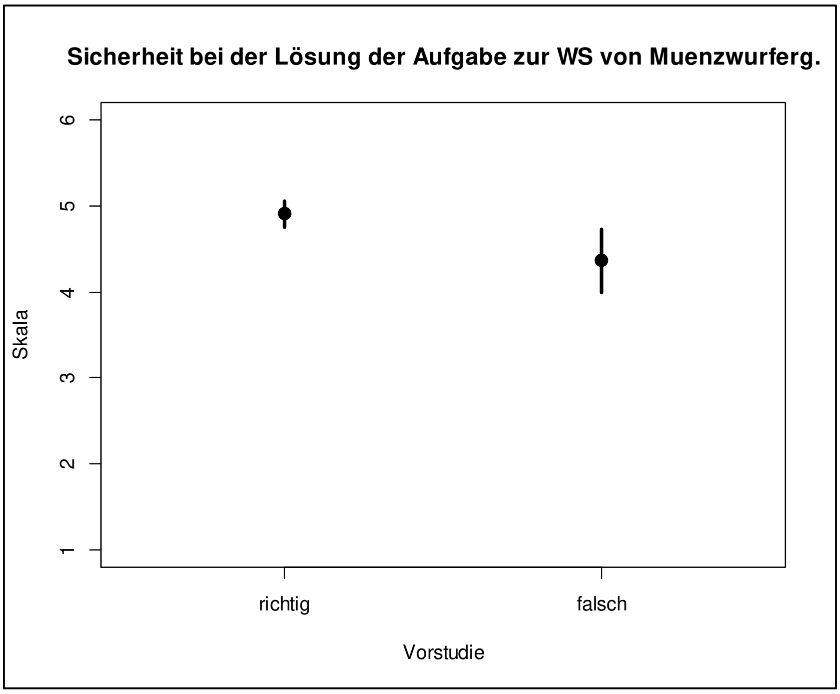


Abb.II.C.5. *Aufgabe zur Wahrscheinlichkeit von Münzwurfergebnissen*: In der Graphik werden die Sicherheiten, die die Befragten für ihre Lösungen angegeben haben, getrennt nach der Korrektheit der Lösungen abgebildet, wobei es keinen signifikanten Unterschied gab. Es sind die Mittelwerte mit Standardfehlern dargestellt.

	Richtig	Falsch
MW	4.91	4.36
Stdf.	0.15	0.36
Anz.	33	11

	p-Wert
Richtig – Falsch	0.1889

Tab.II.C.4. Mittelwerte (MW) und Standardfehler (Stdf.), jeweils auf die zweite Dezimale gerundet, sowie Anzahlen der Befragten (Anz.) bezüglich der Sicherheiten bei richtiger bzw. falscher Antwort zu der *Aufgabe zur Wahrscheinlichkeit von Münzwurfergebnissen*, sowie Vergleiche der Mittelwerte bezüglich der Sicherheiten: p-Wert des t-Tests (ungepaart).

Bei der ***Aufgabe zur Erkennung möglicherweise gefälschter Münzwürfe***, betrachten wir drei Reihen aus dem Fragebogen zu Beginn des Semesters genauer, die in ähnlicher Weise auch in dem Fragebogen am Ende des Semesters verwendet wurden:

1.) 0 0 0 1 0 1 1 1 0 0 1 0 0 0 1 1 1 1 0 1 [138]
2.) 0 1 0 0 1 0 1 0 0 1 1 0 1 0 0 1 1 0 1 0 [139]
3.) 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 [140]

[138] Es handelt sich um die zweite Reihe im Fragebogen zu Beginn des Semesters. Im Fragebogen am Ende des Semesters wurde dieselbe Reihe an erster Stelle verwendet, wobei jeweils die Einsen und Nullen vertauscht wurden.
[139] Diese Reihe stand im Fragebogen zu Beginn des Semesters an der dritten Stelle und am Ende des Semesters an zweiter Stelle.

Die erste Reihe wurde sowohl zu Beginn des Semesters als auch am Ende von niemandem angekreuzt. An der Echtheit der zweiten Reihe zweifelten zu Beginn des Semesters 3 Personen (etwa 7%), am Ende 2 Befragte (etwa 5%). Die dritte Reihe hielten zu Beginn des Semesters 35 Befragte (etwa 80%) für eine Fälschung, am Ende des Semesters waren es 34 (etwa 77%). Eine Person hat die Aufgabe nicht bearbeitet (keine Reihe angekreuzt und zusätzlich keine Angabe zur Sicherheit gemacht).[141] Zwischen den Ergebnissen am Anfang und am Ende des Semesters gab es somit keine bemerkenswerten Unterschiede (vgl. Abb.II.C.6).

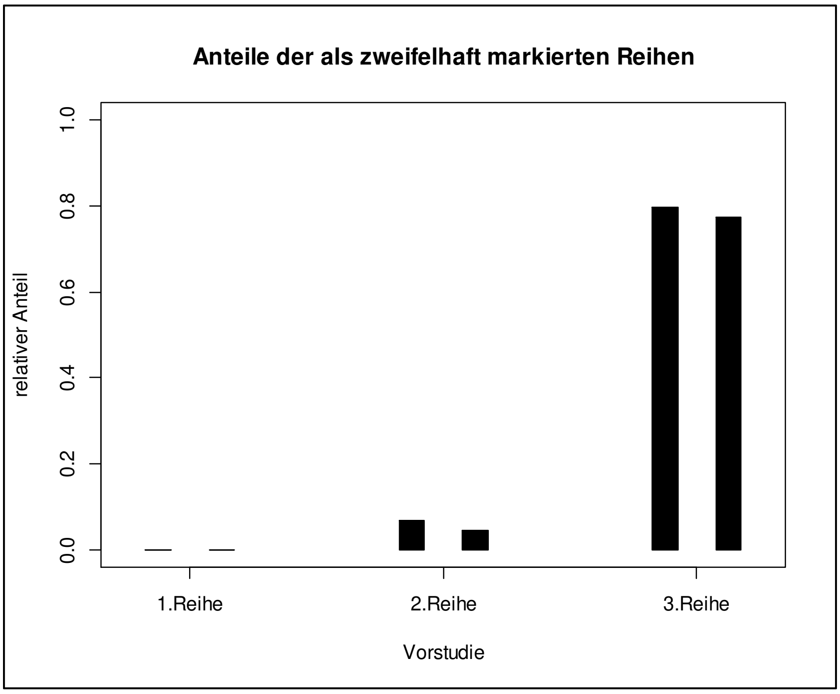


Abb.II.C.6. *Aufgabe zur Erkennung möglicherweise gefälschter Münzwürfe:* Die Graphik zeigt den prozentualen Anteil derer, die die Reihen als „zweifelhaft“ markiert haben. Die linken Säulen geben jeweils die Ergebnisse zu Beginn des Semesters und die rechten Säulen die am Ende des Semesters wieder. Zwischen den beiden Zeitpunkten gab es jeweils keinen signifikanten Unterschied.

Drei der Befragten lösten die Aufgabe am Anfang des Semesters vollständig richtig, d.h. sie kreuzten die erste Reihe nicht an, die zweite und dritte aber schon. Am Ende des Semesters war es nur eine Person. Aufgrund der sehr geringen Anteile richtiger Lösungen wird auf einen Vergleich der Sicherheiten bei richtigen Lösungen verzichtet. Die Sicherheiten, die für falsche

[140] Hierbei handelt es sich sowohl im Fragebogen zu Beginn als auch in dem am Ende des Semesters um die letzte Reihe. Im Fragebogen am Ende des Semesters wurden lediglich die Einsen und Nullen vertauscht.
[141] Da die Nicht-Bearbeitung möglicherweise Information tragen könnte (nämlich dass die Aufgabe nicht verstanden wurde, werden die relativen Anteile in Abb.II.C.6. dennoch auf alle Befragten (n=44) bezogen.

Lösungen angeben wurden, unterschieden sich am Anfang und am Ende des Semesters nicht signifikant (vgl. Tab.II.C.5).

	Falsch vorher	Falsch nachher
MW	4.36	4.4
Stdf.	0.17	0.14
Anz.	39	40

	p-Wert
Vorher - Nachher	0.8498

Tab.II.C.5. Mittelwerte (MW) und Standardfehler (Stdf.), jeweils auf die zweite Dezimale gerundet, der Sicherheiten bei falscher Antwort sowie Anzahlen der Befragten (Anz.), die bezüglich ihrer Sicherheit bei der Lösung der *Aufgabe zur Erkennung möglicherweise gefälschter Münzwürfe* eine Angabe gemacht haben. Außerdem Vergleich der Mittelwerte: p-Wert des t-Tests (ungepaart).

Zu Beginn des Semesters begründeten 30 Studierende ihre Antwort in dem vorgegebenen Freitextfeld. Dabei bezogen sich 17 in ihrer Argumentation auf die Anzahlen von Nullen und Einsen. Zwei Personen argumentierten mit der Länge von Runs, alle anderen lieferten unklare oder falsche Erklärungen (v.a. dass alle Folgen gleich wahrscheinlich seien).

Insgesamt 26 Probanden gaben am Ende des Semesters eine Begründung für ihre Lösung an, wobei ebenfalls 17 von ihnen auf die verschiedenen Anzahlen von Nullen und Einsen Bezug nahmen, zwei mit der Länge der Runs argumentierten und eine Person erkennbare Muster (womit vermutlich die Ausgeglichenheit gemeint war) als ein Indiz für eine Fälschung in die Argumentation einbezog.

Die ***Aufgabe zum Testen eines Würfels*** wurde den Befragten nur zu Beginn des Semesters vorgelegt. Alle Antworten in dem Freitextfeld, die auf die Idee eines wiederholten Würfelns als Testverfahren, um die Fairness des Würfels zu prüfen, schließen ließen, wurden als richtig gewertet. Antworten, dass einmal würfeln genüge, man nur die Anzahlen an geraden und ungeraden Augenzahlen vergleichen müsse, es nicht möglich sei, die Fairness zu prüfen oder unverständliche Antworten, wurden als nicht richtig gewertet. 34 Befragte (etwa 77%) lösten die Aufgabe richtig, 8 Personen (etwa 18%) gaben eine falsche Antwort und 2 Studierende (etwa 5%) bearbeiteten die Aufgabe nicht (vgl. Abb.II.C.7).

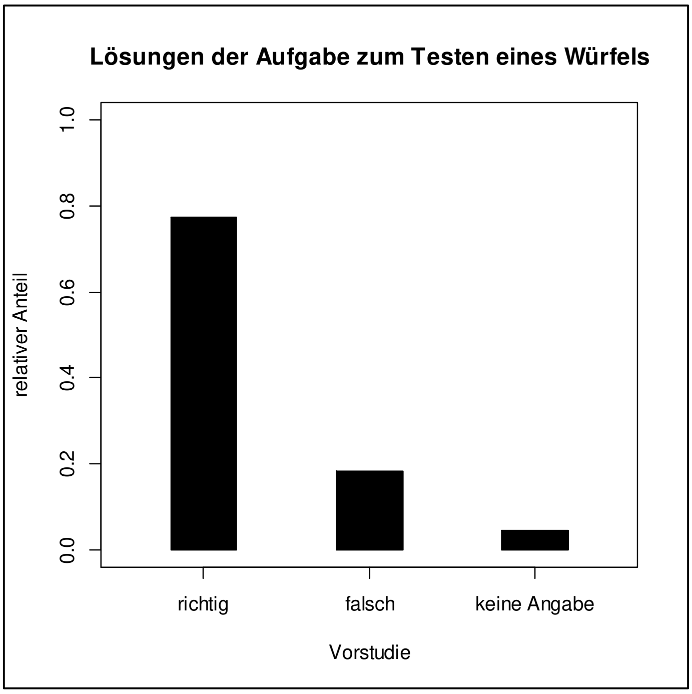


Abb.II.C.7. In der Graphik werden die prozentualen Anteile an richtigen und falschen Lösungen der *Aufgabe zum Testen eines Würfels* angegeben sowie der Anteil der Personen, die keine Angaben gemacht haben.

Die Sicherheiten bei richtigen und falschen Antworten unterschieden sich nicht signifikant, wenngleich der p-Wert mit knapp 9% zumindest auffällig ist (vgl. Abb.II.C.8 und Tab.II.C.6).

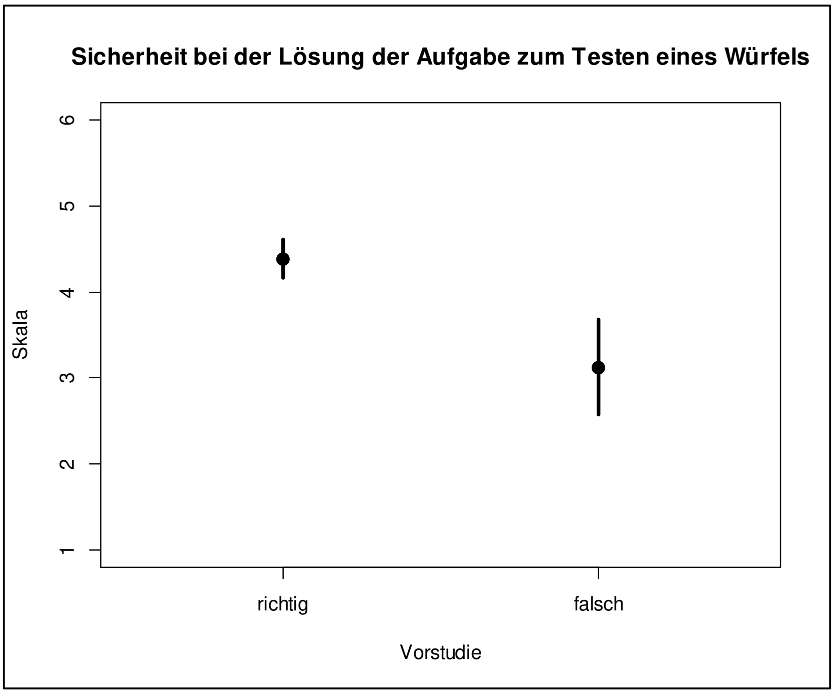


Abb.II.C.8. *Aufgabe zum Testen eines Würfels:* In der Graphik werden die Sicherheiten, die die Befragten für ihre Lösungen angegeben haben, getrennt nach der Korrektheit der Lösungen abgebildet, wobei es keinen signifikanten Unterschied gab. Es sind die Mittelwerte mit Standardfehlern dargestellt.

	Richtig	Falsch
MW	4.38	3.57
Stdf.	0.22	0.37
Anz.	34	7

	p-Wert
Richtig – Falsch	0.08692

Tab.II.C.6. Mittelwerte (MW) und Standardfehler (Stdf.), jeweils auf die zweite Dezimale gerundet, sowie Anzahlen der Befragten (Anz.) bezüglich der Sicherheiten bei richtiger bzw. falscher Antwort zu der *Aufgabe zum Testen eines Würfels*, sowie Vergleiche der Mittelwerte bezüglich der Sicherheiten: p-Wert des t-Tests (ungepaart).

Bei der ***Post-Aufgabe***, die den Befragten ebenfalls nur zu Beginn des Semesters gestellt wurde, gaben 33 Personen (75%) die richtige Antwort, 10 Studierende (etwa 23%) beantworteten sie falsch und eine Person bearbeitete sie nicht (vgl. Abb.II.C.9).

Die Sicherheiten, die die Befragten für ihre Lösung angaben, unterschieden sich nicht signifikant zwischen richtigen und falschen Lösungen (vgl. Abb.II.C.10 und Tab.II.C.7).

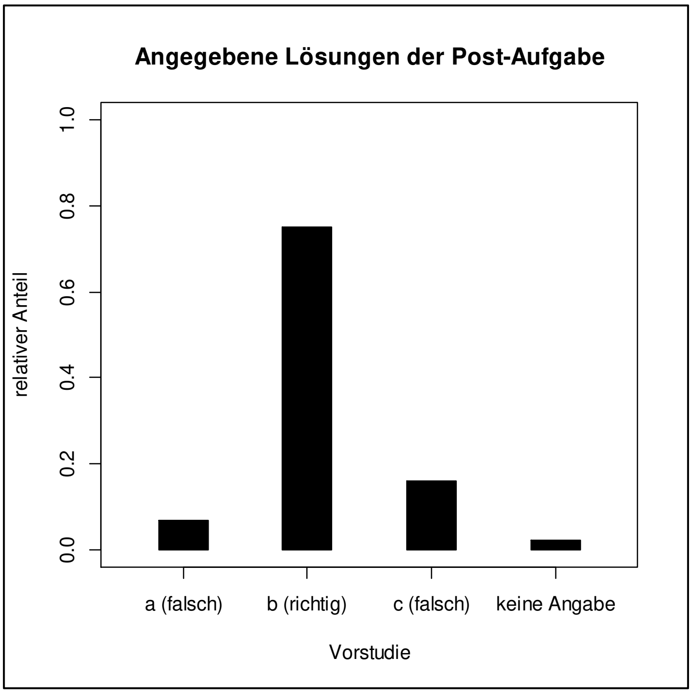


Abb.II.C.9. In der Graphik werden die prozentualen Anteile der angekreuzten Lösungen der *Post-Aufgabe* angegeben sowie der Anteil der Personen, die keine Angaben gemacht haben.

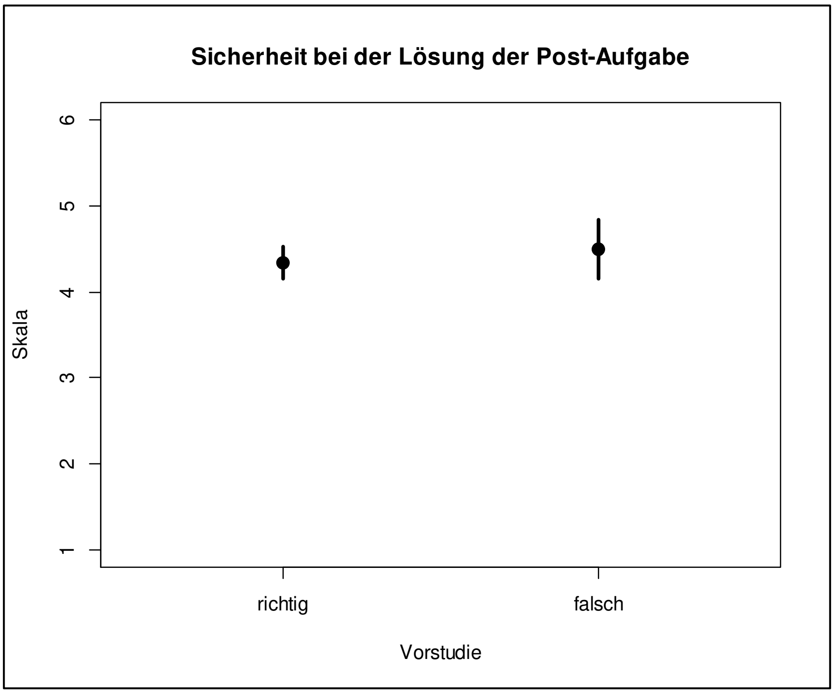


Abb.II.C.10. *Post-Aufgabe:* In der Graphik werden die Sicherheiten, die die Befragten für ihre Lösungen angegeben haben, getrennt nach der Korrektheit der Lösungen abgebildet, wobei es keinen signifikanten Unterschied gab. Es sind die Mittelwerte mit Standardfehlern dargestellt.

	Richtig	Falsch
MW	4.33	4.50
Stdf.	0.19	0.34
Anz.	33	10

	p-Wert
Richtig – Falsch	0.6752

Tab.II.C.7. Mittelwerte (MW) und Standardfehler (Stdf.), jeweils auf die zweite Dezimale gerundet, sowie Anzahlen der Befragten (Anz.) bezüglich der Sicherheiten bei richtiger bzw. falscher Antwort zu der *Post-Aufgabe*, sowie Vergleiche der Mittelwerte bezüglich der Sicherheiten: p-Wert des t-Tests (ungepaart).

Die ***Multiple-Choice-Test-Aufgabe*** wurde den Studierenden nur am Ende des Semesters vorgelegt. Sie wurde lediglich von 15 Befragten (etwa 34%) richtig beantwortet, 28 Befragte (etwa 64%) lösten sie falsch und eine Person machte keine Angabe (vgl. Abb.II.C.11). Bei den falschen Lösungen kann (anders als bei der *Krankenhaus*- bzw. *Grundschule*-Aufgabe) kein so klarer Bezug zu der adressierten Fehlvorstellung hergestellt werden.

Die Sicherheiten, die die Befragten für ihre Lösung angaben, unterschieden sich signifikant zwischen richtigen und falschen Lösungen (vgl. Abb.II.C.12 und Tab.II.C.8). Die Studierenden, die eine falsche Lösung angekreuzt hatten, bezifferten ihre Sicherheit im Mittel signifikant niedriger als jene mit richtiger Lösung.

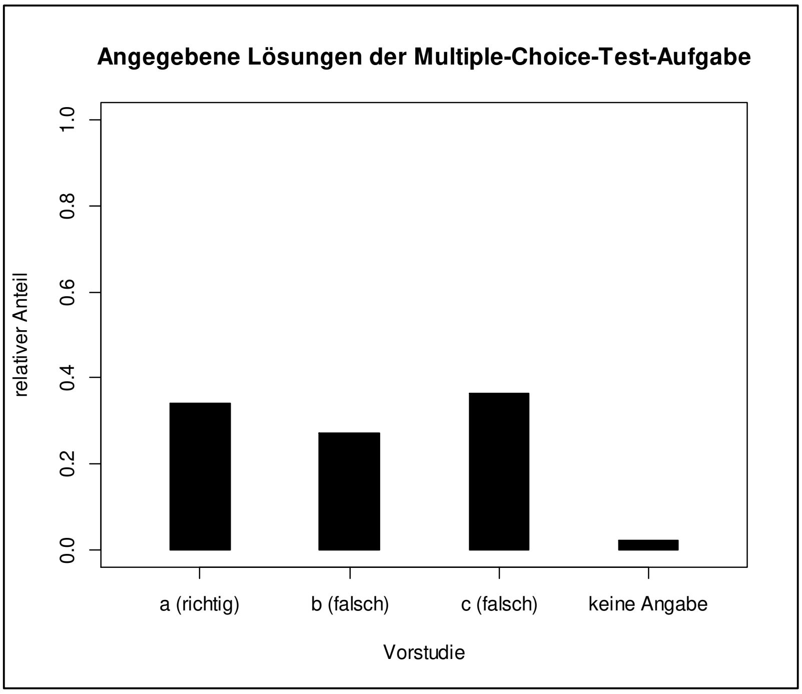


Abb.II.C.11. In der Graphik werden die prozentualen Anteile der angekreuzten Lösungen der *Multiple-Choice-Test-Aufgabe* angegeben sowie der Anteil der Personen, die keine Angaben gemacht haben.

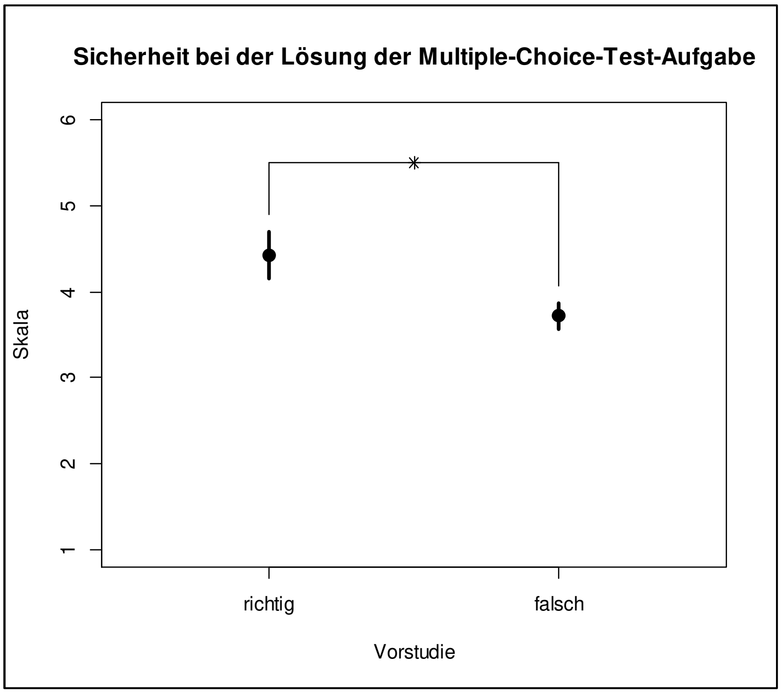


Abb.II.C.12. *Multiple-Choice-Test-Aufgabe:* In der Graphik werden die Sicherheiten, die die Befragten für ihre Lösungen angegeben haben, getrennt nach der Korrektheit der Lösungen abgebildet. Es sind die Mittelwerte mit Standardfehlern dargestellt. Signifikante Unterschiede sind mit einem Stern markiert.

	Richtig	Falsch
MW	4.43	3.71
Stdf.	0.27	0.15
Anz.	15	28

	p-Wert
Richtig – Falsch	0.03215

Tab.II.C.8. Mittelwerte (MW) und Standardfehler (Stdf.), jeweils auf die zweite Dezimale gerundet, sowie Anzahlen der Befragten (Anz.) bezüglich der Sicherheiten bei richtiger bzw. falscher Antwort zu der *Multiple-Choice-Test-Aufgabe* (sofern eine Angabe gemacht wurde). Außerdem Vergleiche der Mittelwerte bezüglich der Sicherheiten: p-Wert des t-Tests (ungepaart).

II.C.3.b. Einschätzungen der eigenen Kompetenz und Einstellungen zur Stochastik

Die ***Einschätzung der eigenen Kompetenz im Fach Stochastik*** sollte von den Befragten im Hinblick auf die Stärke der Zustimmung zu folgenden fünf Aussagen auf einer sechsstufigen Skala (von *„trifft überhaupt nicht zu“*(1) bis *„trifft vollständig zu“*(6)) vorgenommen werden:

E1: „Mein Vorwissen aus der Schule schätze ich als sehr gut ein.“ / „Mein Wissen in Stochastik schätze ich als sehr gut ein.“

E2: „In Stochastik kommt meistens etwas anderes raus, als man vermutet.

E3: „Bei Stochastik versagt meine Intuition.“

E4: „Ich kann mir nur sehr schwer vorstellen, was Zufall bedeuten soll.“

E6: „Das Lösen von Stochastik-Aufgaben fällt mir im Allgemeinen nicht schwerer als von Aufgaben aus anderen mathematischen Bereichen.“

E7: „Ich traue meinen Lösungen von Stochastik-Aufgaben im Allgemeinen nicht.“

Im Folgenden werden die Angaben der Befragten zu diesen Aussagen jeweils zu Beginn und am Ende des Semesters dargestellt und die individuellen Unterschiede mit Hilfe gepaarter t-Tests untersucht (vgl. Abb.II.C.13 und Tab.II.C.9). Zu beachten ist, dass die Höhen der Werte bei den Einschätzungen verschieden zu interpretieren sind: Bei E1 und E6 weisen hohe Werte darauf hin, dass die Befragten ihre Kompetenz als gut ausgeprägt einschätzen, bei E2, E3, E4 und E7 dagegen darauf, dass sie sie als wenig ausgeprägt empfinden. Die Befragten schätzten ihre Kompetenz insgesamt durchschnittlich als mittelmäßig ausgeprägt ein, wobei sie im Mittel davon ausgingen, dass sie sich eher gut vorstellen können, was Zufall bedeutet (E4). Die Einschätzungen zu Beginn und am Ende des Semesters unterschieden sich dabei insgesamt nicht signifikant, wenn man sie auf individueller Ebene vergleicht (vgl. Tab.II.C.9). Einzig bei E7 konnte ein auffälliger Unterschied zwischen dem Beginn und dem Ende des Semesters beobachtet werden, der darauf hindeutet, dass das Vertrauen in die eigenen Lösungen im Laufe des Semesters etwas gesunken sein könnte.

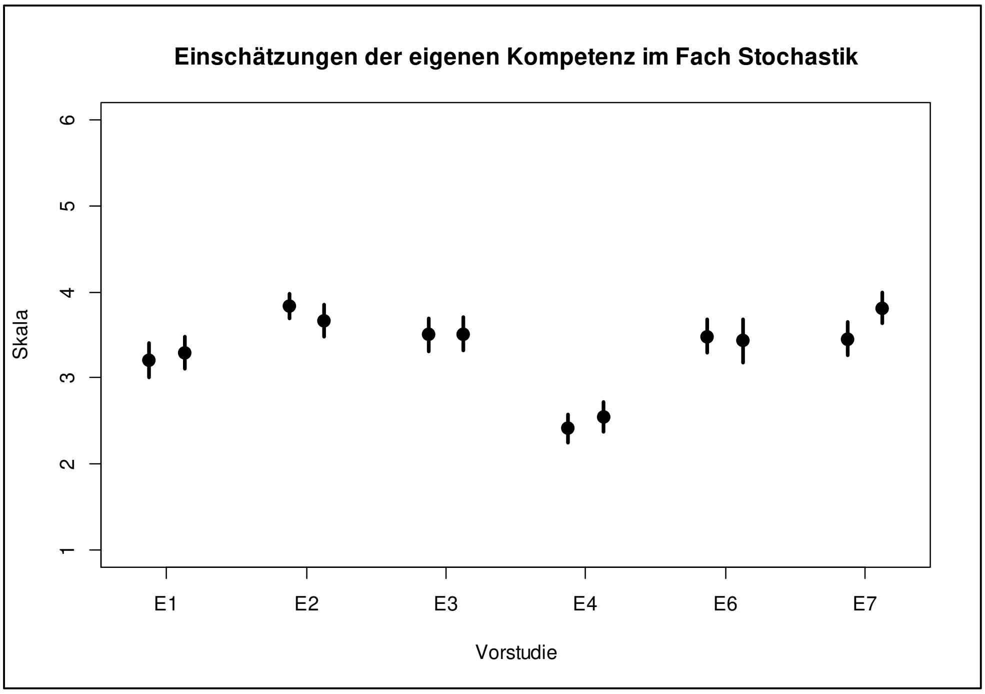


Abb.II.C.13. In der Graphik sind die Selbsteinschätzungen der Befragten im Hinblick auf ihre eigene Kompetenz im Fach Stochastik dargestellt, wobei links jeweils der Mittelwert der Einschätzungen zu Beginn des Semesters mit Standardfehler und rechts der Mittelwert der Einschätzungen am Ende des Semesters mit Standardfehler abgebildet ist. Zwischen den beiden Zeitpunkten gab es jeweils keinen signifikanten Unterschied.
Bei den Aussagen E1 und E6 bedeutet ein hoher Wert, dass die Person ihre Kompetenz als gut einschätzt, bei den Aussagen E2, E3, E4 und E7 bedeutet ein hoher Wert dagegen, dass die Person ihre Kompetenz als niedrig empfindet.

	E1 Diff.	E2 Diff.	E3 Diff.	E4 Diff.	E6 Diff.	E7 Diff.
MW	0.10	-0.20	-0.07	0.10	-0.05	0.36
Stdf.	0.21	0.21	0.19	0.19	0.23	0.19
Anz.	42	41	41	42	42	42

	p-Wert (individuelle Diff.)		p-Wert (individuelle Diff.)
E1	0.6475	E4	0.6173
E2	0.3521	E6	0.8359
E3	0.6962	E7	0.0704

Tab.II.C.9 Mittelwerte (MW) und Standardfehler (Stdf.) der individuellen Differenzen bzgl. E1, E2, E3, E4, E6 und E7 sowie die Anzahlen (Anz.) der Befragten, von denen sowohl die Angabe zu Beginn des Semesters als auch am Ende des Semesters vorlagen. Außerdem p-Werte der gepaarten t-Tests.

Zusätzlich wurde in den Fragebögen auch nach ***Einstellungen zur Stochastik*** gefragt:

E5: „Ich mag Stochastik."

E8: „Stochastik ist nützlich."

E9: „Stochastik ist kein typisch mathematisches Fach."

Die Angaben der Befragten zu Beginn und am Ende des Semesters werden in der folgenden Graphik dargestellt (vgl. Abb.II.C.14) und die individuellen Änderungen untersucht. Dabei ergaben sich keine Hinweise auf eine signifikante Veränderung (vgl. Tab.II.C.10). Auffällig ist, dass die Befragten die Nützlichkeit des Faches (E8) vergleichsweise hoch einschätzten, während sich die anderen Angaben durchschnittlich im Mittelfeld bewegten.

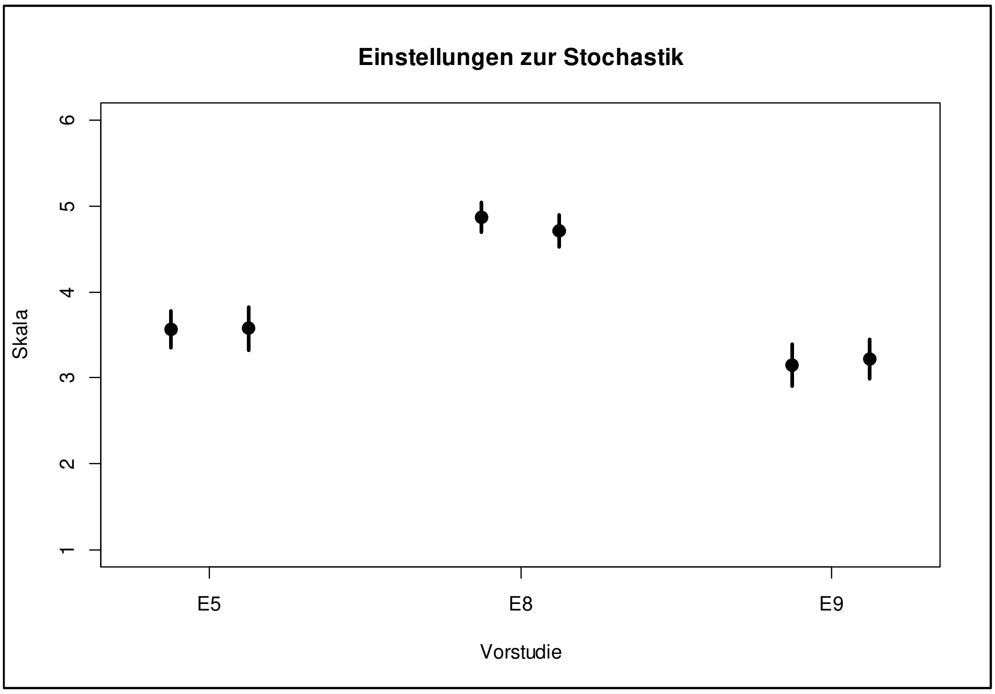


Abb.II.C.14. In der Graphik sind die Angaben der Befragten im Hinblick auf ihre Einstellungen zur Stochastik (E5, E8 und E9) abgebildet, wobei links jeweils der Mittelwert der Angaben zu Beginn des Semesters mit Standardfehler und rechts der Mittelwert der Angaben am Ende des Semesters mit Standardfehler dargestellt ist. Zwischen den beiden Zeitpunkten gab es jeweils keinen signifikanten Unterschied.

	E5 Diff.	E8 Diff.	E9 Diff.
MW	0.02	-0.17	0.08
Stdf.	0.13	0.16	0.28
Anz.	42	41	38

	p-Wert (individuelle Diff.)
E5	0.8501
E8	0.2913
E9	0.7761

Tab.II.C.10. Mittelwerte (MW) und Standardfehler (Stdf.) der individuellen Differenzen bzgl. E5, E8 und E9 sowie die Anzahlen (Anz.) der Befragten, von denen sowohl die Angabe zu Beginn des Semesters als auch am Ende des Semesters vorlagen. Außerdem p-Werte der gepaarten t-Tests.

Zu Beginn des Semesters begründeten 20 der Befragten ihre Einschätzung von E9. Sechs von ihnen stimmten der Aussage „gar nicht“ bis „eher weniger“ zu und begründeten dies damit, dass es auch in Stochastik Definitionen, Sätze und Beweise gebe (1 Mal), das Fach logisch stringent sei (2 Mal) bzw. sie kein Gegenargument hätten (3 Mal). Die restlichen 14 Personen stimmten der Aussage „eher“ bis sogar „vollständig zu“. Als Argumente nannten sie (wobei Mehrfachnennungen vorkamen), dass die Stochastik sehr praxis- und anwendungsorientiert sei (4 Mal), sie stärker diskursiv und auf Anschauung basierend bzw. weniger formal sei als andere mathematische Fächer (4 Mal), man in ihr mit Unsicherheit und Ungenauigkeit umgehen müsse (2 Mal), Berechnungen nicht immer intuitiv seien (2 Mal) und dass sie allgemein andere Denkprozesse/ -ansätze erfordere als es bei sonstigen mathematischen Disziplinen der Fall sei (3 Mal).

Am Ende des Semesters hatten die Studierenden ebenfalls ein Freitextfeld zur Verfügung. In diesem konnten sie ihre Schwierigkeiten bei der Auseinandersetzung mit dem Fach Stochastik benennen. Dies taten insgesamt 19 Befragte. Am häufigsten wurden die Anwendung der Theorie auf die Übungsaufgaben genannt (10 Mal) und die Schwierigkeit, eine geeignete Herangehensweise an die Aufgaben zu finden (4 Mal). Vier Nennungen bezogen sich auf die Struktur der Vorlesung bzw. die Stoffmenge und zweimal wurde das Nachvollziehen von Beweisen und Formeln als Schwierigkeit formuliert. Als inhaltlich schwierig wurde von drei Befragten die Kombinatorik angesehen und von einer Person das Durchdringen der Eigenschaften unterschiedlicher Verteilungen.

II.C.3.c. Zusammenfassung der Ergebnisse aus der Vorstudie

Insgesamt wurde bei drei Aufgaben, die den Studierenden zur Abbildung der bereits zu Beginn des Semesters vorhandenen Grundvorstellungen vorgelegt wurden, ein hoher Anteil richtiger Lösungen erreicht, sodass diese im Fragebogen am Ende des Semesters nicht erneut verwendet wurden:

- Die Aufgabe zur ***Wahrscheinlichkeit von Münzwurfergebnissen*** wurde von 75% der Befragten richtig beantwortet, wobei sich die mittleren angegebenen Sicherheiten nicht auffällig zwischen richtigen und falschen Lösungen unterschieden.
- Die ***Aufgabe zum Testen eines Würfels*** lösten etwa 77% der Befragten richtig. Bei dieser Aufgabe lagen die angegebenen Sicherheiten bei falscher Lösung im Mittel zumindest auffällig unter denen für richtige Lösungen.
- Bei der ***Post-Aufgabe*** lag der Anteil der richtigen Lösungen bei 75%. Die Fehlvorstellung F2(a) scheint ein Großteil der Befragten folglich nicht zu teilen. Bezüglich der Sicherheiten ließ sich nichts Auffälliges beobachten.

Zwei Aufgaben wurden den Studierenden sowohl zu Beginn als auch am Ende des Semesters vorgelegt, um eine mögliche Verbesserung der Grundvorstellungen zu untersuchen:

- Bei der ***Krankenhaus- bzw. Grundschule-Aufgabe*** gab es keinen auffälligen Unterschied zwischen den Anteilen richtiger Lösungen zu Beginn und am Ende des Semesters. In beiden Fällen lösten weniger als 40% der Befragten die Aufgaben korrekt, die Mehrheit schien jeweils der adressierten Fehlvorstellung F2(b) zu folgen. Bei den mittleren Sicherheiten gab es dagegen eine Entwicklung: Konnte zu Beginn des Semesters noch kein Unterschied zwischen richtigen und falschen Lösungen beobachtet werden, war der Wert am Ende des Semesters im Fall der richtigen Lösungen signifikant höher als bei falschen Lösungen.
- Bei der Aufgabe zur ***Erkennung möglicherweise gefälschter Münzwürfe*** gab es keine auffälligen Unterschiede zwischen dem Anfang und dem Ende des Semesters, auch nicht bezüglich der Begründungen, die im Freitextfeld angegeben werden konnten. Der überwiegende Teil der Befragten, die Argumente für ihre Lösung anbrachten, bezogen sich auf die Anzahlen von Nullen und Einsen. Zu beiden Zeitpunkten begründeten jeweils äußerst wenige Probanden (zwei bzw. drei) ihre Lösung mit der Länge der Runs (Anzahl der Vorzeichenwechsel) bzw. mit erkennbaren Mustern. Bezüglich der angegebenen Sicherheiten ließ sich nichts Auffälliges beobachten.

Eine Aufgabe wurde den Studierenden nur am Ende des Semesters vorgelegt, um eine mögliche Erweiterung jener der *Post-Aufgabe* zugrunde liegenden Grundvorstellung zu untersuchen:

- Die ***Multiple-Choice-Test-Aufgabe*** konnten nur etwa ein Drittel der Befragten richtig lösen. Die Antworten waren unter den drei vorgegebenen Möglichkeiten nahezu gleichverteilt. Allerdings unterschieden sich die mittleren Sicherheiten: Die Studierenden, die die richtige Antwort ankreuzten, gaben im Mittel eine signifikant höhere Sicherheit an als jene, die sich für eine falsche Lösung entschieden hatten.

Bei allen Aufgaben bewegten sich die angegebenen Sicherheiten durchschnittlich sowohl bei richtigen als auch bei falschen Antworten im mittleren oder oberen Bereich der Skala.

Die **Selbsteinschätzungen der Befragten bezüglich ihrer Kompetenzen** im Fach Stochastik unterschieden sich lediglich bezüglich des Items E7 (Vertrauen in die eigenen Lösungen) zu Beginn und am Ende des Semesters auffällig im individuellen Vergleich. Das Vertrauen in die eigenen Lösungen von Stochastik-Aufgaben sank im Laufe des Semesters. Der Unterschied war allerdings nicht signifikant. Die Einschätzungen ihrer Kompetenzen lagen bezüglich fast aller Items durchschnittlich im mittleren Bereich der Skala, lediglich die Fähigkeit, sich Zufall vorstellen zu können, wurde tendenziell als ausgeprägter angesehen.

Die **Einstellungen zur Stochastik** unterschieden sich nicht auffällig zwischen den beiden Zeitpunkten. Bemerkenswert ist an dieser Stelle, dass die Studierenden die Nützlichkeit des Fachs (E8) insgesamt als groß bewerteten.

Diejenigen, die ihre Einschätzung der Stochastik als typisch oder untypisch mathematisch (E9) in eigenen Worten begründeten, bezogen sich überwiegend auf die Formalität, aber auch auf die Anwendungs-/ Praxisorientierung des Faches sowie seine Anschaulichkeit.

Bei der Abfrage von Schwierigkeiten, die bei der Beschäftigung mit dem Fach im Laufe des Semesters aufgetreten waren, wurden hauptsächlich die Anwendung der Theorie auf die Übungsaufgaben und das Finden einer geeigneten Herangehensweise für ihre Lösung genannt.

HAUPTSTUDIE SOSE 2012:

Basierend auf den Ergebnissen der Vorstudie sollte in der Hauptstudie untersucht werden, inwiefern der Einsatz von Simulationen in den Übungsaufgaben und in der Vorlesung zum Aufbau bzw. zu einer Erweiterung stochastischer Grundvorstellungen und einem Abbau von Fehlvorstellungen beitragen sowie Einstellungen zur Stochastik verbessern könnte. Daneben sollte eruiert werden, ob die Studierenden in der Beschäftigung mit Simulationen und dem damit verbundenen Einblick in Programmierungszusammenhänge einen Gewinn erkennen können – auch im Hinblick auf ihre berufliche Zukunft.

II.C.4. Auswertung der Hauptstudie (SoSe 2012)

Für die Auswertung wurden die Befragten in drei Experimentalgruppen aufgeteilt. Mit **R12** wird die Gruppe der Personen bezeichnet, die die „R Aufgaben" nie oder selten bearbeitet haben. In der Gruppe **R34** sind diejenigen Studierenden, die die „R Aufgaben" häufig oder immer eigenständig (also ohne regelmäßigen Besuch der Fragestunde) bearbeitet haben und mit der Gruppe **Frage** sind jene Befragte gemeint, die die Fragestunde oft oder immer besucht haben (wobei diese Studierenden alle bis auf eine Ausnahme die „R Aufgaben" auch oft oder immer bearbeitet haben).

II.C.4.a. Stochastische Fertigkeiten

Die ***„Krankenhaus-Aufgabe"*** zu Beginn des Semesters beantworteten insgesamt etwa 48% der Befragten richtig; die dazu äquivalente ***„Grundschule-Aufgabe"*** am Ende des Semesters lösten etwa 60% der Befragten korrekt. Bei den falschen Lösungen überwog sehr deutlich die der Fehlvorstellung F2(b) zuzuordnende Antwort c). Der prozentuale Anteil richtiger Antworten war in der Gruppe der Befragten, die regelmäßig die Fragestunde besucht hatten, sowohl zu Beginn als auch am Ende des Semesters signifikant oder zumindest auffällig niedriger als in den anderen beiden Gruppen (vgl. Abb.II.C.15 und Tab.II.C.11). Innerhalb der Experimentalgruppen gab es bezüglich der Anteile richtiger und falscher Lösungen keine auffälligen Änderungen (vgl. Tab.II.C.12). Beide Aufgaben wurden von allen Studierenden bearbeitet.

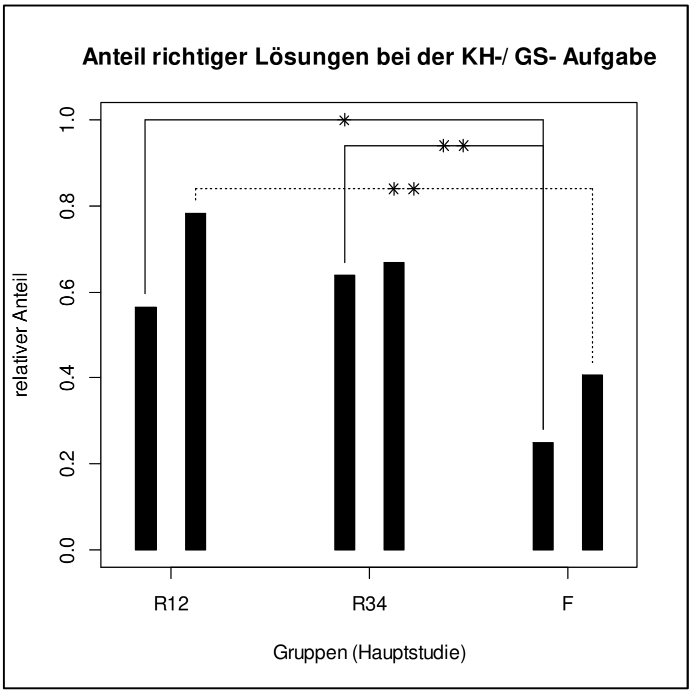


Abb.II.C.15. In der Graphik werden die relativen Anteile richtiger Lösungen der *Krankenhaus-Aufgabe* (jeweils links) und der *Grundschule-Aufgabe* (jeweils rechts), getrennt nach den Experimentalgruppen abgebildet. Signifikante Unterschiede sind mit einem Stern markiert (wobei ein Stern einen p-Wert kleiner als 5% anzeigt und zwei Sterne auf einen p-Wert kleiner als 1% hinweisen).

R12			
	a (richtig)	b (falsch)	c (falsch)
KH	13	2	8
GS	18	0	5

R34			
	a (richtig)	b (falsch)	c (falsch)
KH	23	3	10
GS	24	1	11

Frage			
	a (richtig)	b (falsch)	c (falsch)
KH	8	7	17
GS	13	2	17

R12		
	richtig	falsch
KH	0.57	0.43
GS	0.78	0.22

R34		
	richtig	falsch
KH	0.64	0.36
GS	0.67	0.33

Frage		
	richtig	falsch
KH	0.25	0.75
GS	0.41	0.59

	R12 – R34	R12 – Frage	R34 – Frage
KH	p=0.5956	p=0.02521	p=0.001649
GS	p=0.3906	p=0.006785	p=0.05043

Tab.II.C.11. Angaben zur *Krankenhaus*(KH)- bzw. *Grundschule*(GS)-*Aufgabe*, aufgeteilt nach den drei Experimentalgruppen, in absoluten und relativen Zahlen, sowie Ergebnisse der Untersuchungen auf Unterschiede bezüglich der Richtigkeit der Antworten zwischen den Gruppen mit Hilfe von Fisher-Tests. Die Kategorien der für den Fisher-Test zugrunde gelegten Kontingenztafeln waren dabei „Aufg. richtig / Aufg. falsch" und „Gruppe x / Gruppe y".

	R12	R34	Frage
KH – GS	p=0.2078	p=1	p=0.2869

Tab.II.C.12. Ergebnisse der Untersuchungen auf Unterschiede innerhalb der Gruppen bezüglich der Richtigkeit der Antworten vorher und nachher mit Hilfe von Fisher-Tests. Die Kategorien der für den Fisher-Test zugrunde gelegten Kontingenztafeln waren dabei „KH richtig/ KH falsch" und „GS richtig / GS falsch".

Insgesamt änderten etwa 62% der Befragten nichts an ihrem Antwortverhalten. Von den Personen, die zu Beginn des Semesters die *Krankenhaus-Aufgabe* richtig beantwortet hatten, lösten etwa 80% auch die *Post-Aufgabe* am Ende des Semesters korrekt. Von denjenigen, die die *Krankenhaus-Aufgabe* zu Beginn des Semesters falsch gelöst hatten, gaben etwa 57% auch bei der *Post-Aufgabe* am Ende des Semesters eine falsche Lösung an. Bei der Gruppe derer, die regelmäßig die Fragestunde besucht hatten, ließ sich mit Hilfe eines Fisher-Tests eine signifikante Verbesserung im individuellen Antwortverhalten beobachten (vgl. Tab.II.C.13), bei allen anderen Gruppen gab es keine Hinweise auf Veränderungen.

R12		*GS*		
		a	b	c
KH	a	12	0	1
	b	2	0	0
	c	4	0	4

R34		*GS*		
		a	b	c
KH	A	17	0	6
	B	0	1	2
	C	7	0	3

Frage		*GS*		
		a	B	c
KH	a	6	0	2
	b	1	2	4
	c	6	0	11

R12: KH – GS	p=0.1269
R34: KH – GS	p= 0.2811
Frage: KH – GS	p= 0.03831

Tab.II.C.13. Die Tabelle bildet die Lösungen der *Krankenhaus*(KH)- und *Grundschule*(GS)-*Aufgaben*, aufgeteilt nach den drei Experimentalgruppen ab. Außerdem sind die p-Werte aus Fisher's exakten Tests zur Untersuchung von möglichen Unterschieden (richtig/falsch) im Antwortverhalten (KH/GS) für jede einzelne Experimentalgruppe aufgeführt.

Vergleicht man die angegebenen Sicherheiten bei richtiger Lösung zwischen den Experimentalgruppen, lassen sich nur am Ende des Semesters signifikante Unterschiede zwischen der Gruppe der Personen, die regelmäßig die Fragestunde besucht hatten, und den anderen beiden Gruppen beobachten: Diese Studierenden trauten ihren Lösungen durchschnittlich weniger. Zu Beginn des Semesters war lediglich die mittlere Einschätzung dieser Gruppe im Vergleich zu denen, die die „R Aufgaben" nie oder selten bearbeitet hatten, auffällig niedriger (aber nicht signifikant). Vergleicht man die Mittelwerte innerhalb der Gruppen zu Beginn und am Ende des Semesters, lässt sich kein Hinweis auf eine Veränderung feststellen. Die Angaben aller Gruppen bewegten sich durchschnittlich in der oberen Hälfte der Skala oder nur sehr knapp darunter. (Vgl. Abb.II.C.16 und Tab.II.C.14 sowie Tab.II.C.15.)

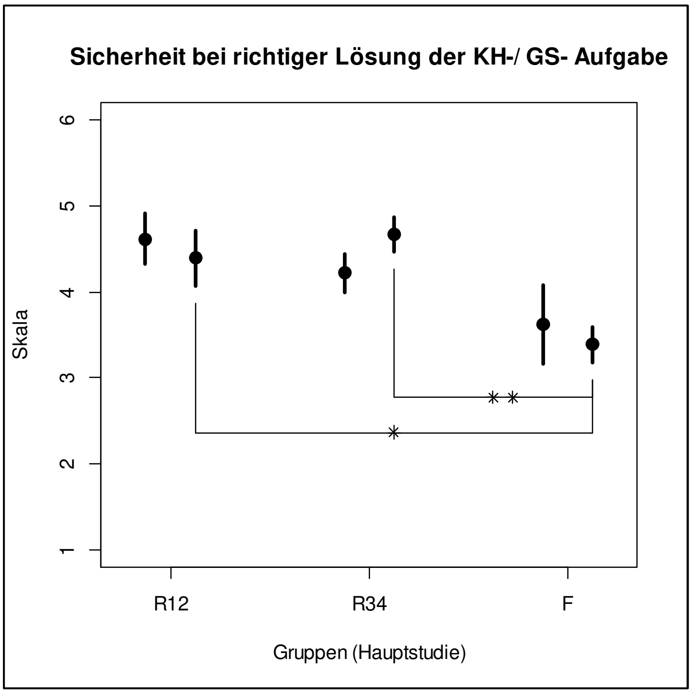


Abb.II.C.16. In der Graphik werden die Sicherheiten, die die Befragten bei richtiger Lösung der *Krankenhaus*(KH)- bzw. *Grundschule*(GS)-*Aufgabe* angegeben haben, getrennt nach den Experimentalgruppen abgebildet. Links befindet sich jeweils der Mittelwert der Sicherheiten bei der *Krankenhaus-Aufgabe* mit Standardfehler und rechts der Mittelwert der Sicherheiten bei der *Grundschule-Aufgabe* mit Standardfehler. Signifikante Unterschiede sind mit einem Stern markiert (wobei ein Stern einen p-Wert kleiner als 5% anzeigt und zwei Sterne auf einen p-Wert kleiner als 1% verweisen).

	R12 KH	R12 GS	R34 KH	R34 GS	Frage KH	Frage GS
MW	4.62	4.39	4.22	4.67	3.63	3.38
Stdf.	0.29	0.32	0.23	0.2	0.46	0.21
Anz.	13	18	23	24	8	13

	p-Werte
R12 KH – R 34 KH	0.2888
R12 KH – Frage KH	0.09268
R34 KH – Frage KH	0.2737

	p-Werte
R12 GS – R34 GS	0.4703
R12 GS – Frage GS	0.01531
R34 GS – Frage GS	0.0001191

Tab.II.C.14. Mittelwerte (MW), Standardfehler (Stdf.) und Anzahl der Befragten (Anz.) für die drei Gruppen bezüglich der Sicherheit bei richtiger Antwort zur *Krankenhaus*(KH)- bzw. *Grundschule*(GS)-*Aufgabe*, sowie Vergleiche der Mittelwerte zwischen den Gruppen bezüglich der Sicherheiten: p-Werte der t-Tests (ungepaart).

	p-Werte (ungepaart)
R12 V – R12 N	0.6066
R34 V – R34 N	0.1411
Frage V – Frage N	0.6458

Tab.II.C.15. Vergleiche der Mittelwerte innerhalb der Gruppen bezüglich der Sicherheit bei richtiger Antwort zur *Krankenhaus*(KH)- bzw. *Grundschule*(GS)-*Aufgabe*: p-Werte der t-Tests (ungepaarte Werte).

Bezüglich der Sicherheiten bei falscher Lösung gab es keine auffälligen Unterschiede zwischen den Gruppen und auch nicht innerhalb der Gruppen, wenn man die Mittelwerte zu Beginn des Semesters mit denen am Ende des Semesters vergleicht. Die Angaben aller

Gruppen bewegten sich im Mittel in der oberen Hälfte der Skala. (Vgl. Abb.II.C.17 und Tab.II.C.16 sowie Tab.II.C.17.)

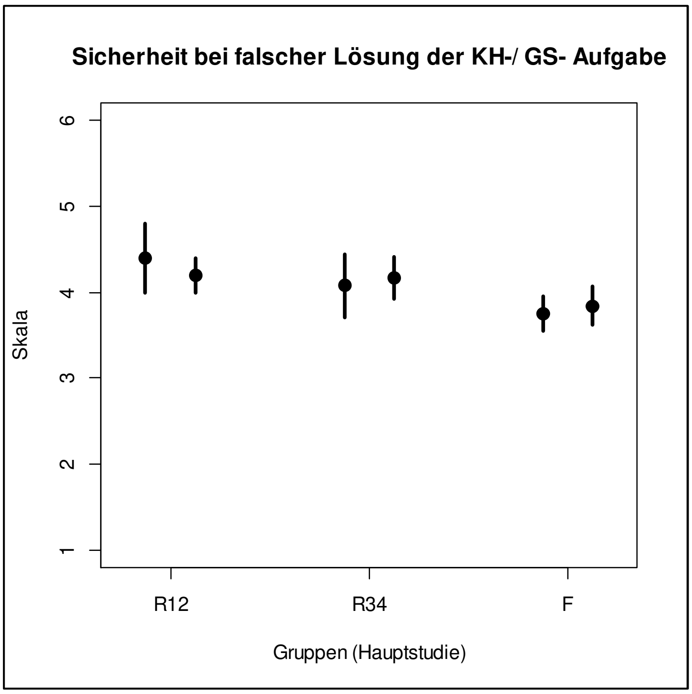


Abb.II.C.17. In der Graphik werden die Sicherheiten, die die Befragten bei falscher Lösung der *Krankenhaus*(KH)- bzw. *Grundschule*(GS)-*Aufgabe* angegeben haben, getrennt nach den Experimentalgruppen abgebildet. Links befindet sich jeweils der Mittelwert der Sicherheiten bei der *Krankenhaus-Aufgabe* mit Standardfehler und rechts der Mittelwert der Sicherheiten bei der *Grundschule-Aufgabe* mit Standardfehler. Es konnten keine signifikanten Unterschiede festgestellt werden.

	R12 KH	R12 GS	R34 KH	R34 GS	Frage KH	Frage GS
MW	4.4	4.2	4.08	4.17	3.75	3.84
Stdf.	0.4	0.2	0.37	0.24	0.2	0.22
Anz.	10	5	13	12	24	19

	p-Werte
R12 KH – R 34 KH	0.5581
R12 KH – Frage KH	0.1692
R34 KH – Frage KH	0.4438

	p-Werte
R12 GS – R34 GS	0.9168
R12 GS – Frage GS	0.2475
R34 GS – Frage GS	0.3289

Tab.II.C.16. Mittelwerte (MW), Standardfehler (Stdf.) und Anzahl der Befragten (Anz.) für die drei Gruppen bezüglich der Sicherheit bei falscher Antwort zur *Krankenhaus*(KH)- bzw. *Grundschule*(GS)-*Aufgabe*, sowie Vergleiche der Mittelwerte zwischen den Gruppen bezüglich der Sicherheiten: p-Werte der t-Tests (ungepaart).

	p-Werte (ungepaart)
R12 KH – R12 GS	0.6625
R34 KH – R34 GS	0.8398
Frage KH – Frage GS	0.7593

Tab.II.C.17. Vergleiche der Mittelwerte innerhalb der Gruppen bezüglich der Sicherheit bei falscher Antwort zur *Krankenhaus*(KH)- bzw. *Grundschule*(GS)-*Aufgabe*: p-Werte der t-Tests (ungepaarte Werte).

Vergleicht man die Angaben innerhalb der Gruppen im Hinblick darauf, ob die Lösungen richtig oder falsch waren, lässt sich kein Unterschied zwischen den mittleren Sicherheiten bei richtigen und falschen Lösungen feststellen (vgl. Tab.II.C.18).

	R12	R34	Frage
KH richtig – KH falsch	p=0.6681	p=0.7474	p=0.8088
GS richtig – GS falsch	p=0.6258	p=0.1204	p=0.1458

Tab.II.C.18. Vergleiche der Mittelwerte innerhalb der Gruppen bezüglich der Sicherheiten bei richtiger und bei falscher Antwort zur *Krankenhaus*(KH)- bzw. *Grundschule*(GS)-*Aufgabe*: p-Werte der t-Tests (ungepaarte Werte).

Die ***Aufgabe zur Wahrscheinlichkeit von Münzwurfergebnissen*** wurde nur zu Beginn des Semesters gestellt. Sie wurde von etwa 75% der Befragten richtig beantwortet. Alle Befragten gaben eine Lösung an. Vergleicht man die Antworten der unterschiedlichen Gruppen, stellt man fest, dass sich der Anteil korrekter Lösungen zwischen der Gruppe der Personen, die regelmäßig die Fragestunde besucht hatten, signifikant von den anderen beiden Gruppen unterscheidet (vgl. Abb.II.C.18 und Tab.II.C.19).

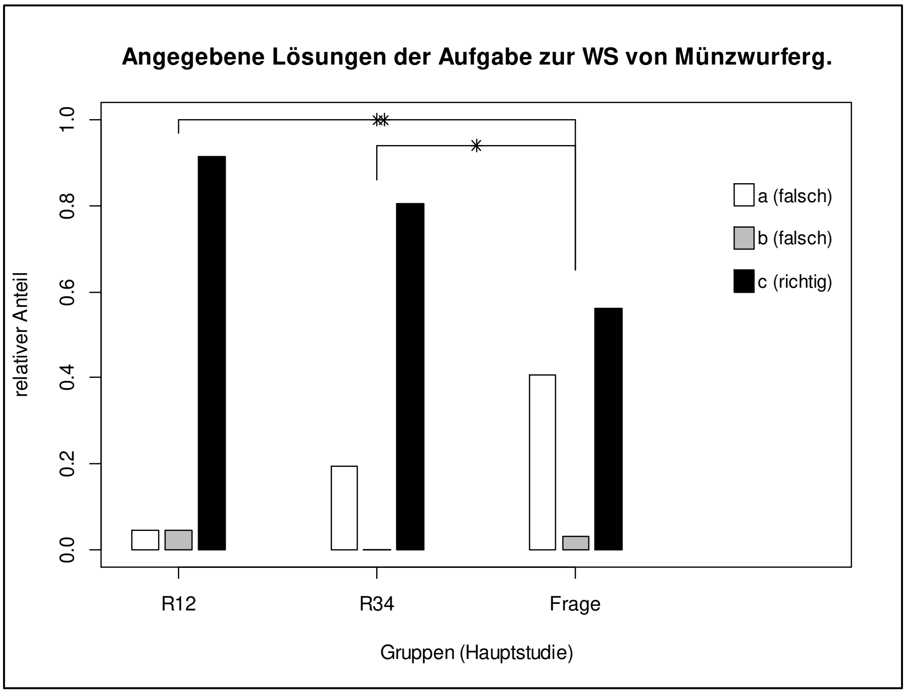


Abb.II.C.18. In der Graphik werden die relativen Anteile der angegebenen Lösungen zur *Aufgabe zur Wahrscheinlichkeit von Münzwurfergebnissen*, getrennt nach den Experimentalgruppen abgebildet. Signifikante Unterschiede zwischen den Gruppen bezüglich der Korrektheit der Lösungen sind mit einem Stern markiert (wobei ein Stern einen p-Wert kleiner als 5% anzeigt und zwei Sterne auf einen p-Wert kleiner als 1% hinweisen).

	a (falsch)	b (falsch)	c (richtig)
R12	1	1	21
R34	7	0	29
Frage	13	1	18

R12 - R34	p = 0.4598
R12 – Frage	p = 0.006306
R34 – Frage	p = 0.03799

Tab.II.C.19. Angegebene Lösungen zur *Aufgabe zur Wahrscheinlichkeit von Münzwurfergebnissen*, getrennt nach den Experimentalgruppen, außerdem p-Werte aus Fisher's exakten Tests zur Untersuchung von möglichen Unterschieden (richtig/falsch) im Antwortverhalten der Gruppen (Gruppe x /Gruppe y).

Die Sicherheiten, die die Befragten bei richtiger Lösung angegeben hatten, unterschieden sich im Mittel nicht signifikant zwischen den drei Experimentalgruppen. Die Angaben aller Gruppen bewegten sich innerhalb der oberen Hälfte Skala. (Vgl. Abb.II.C.19 und Tab.II.C.20.)

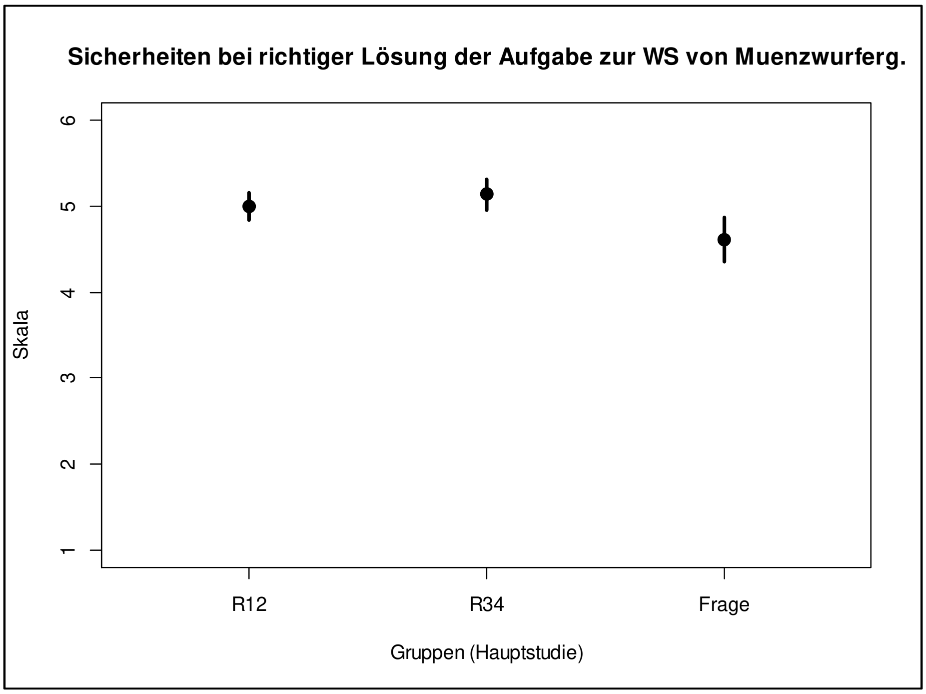


Abb.II.C.19. In der Graphik werden die Sicherheiten, die die Befragten bei richtiger Lösung der *Aufgabe zur Wahrscheinlichkeit von Münzwurfergebnissen* angegeben haben, getrennt nach den Experimentalgruppen abgebildet. Es ist jeweils der Mittelwert mit Standardfehler dargestellt. Es konnten keine signifikanten Unterschiede festgestellt werden.

	R12	R34	Frage
MW	5	5.14	4.61
Stdf.	0.15	0.18	0.26
Anz.	21	29	18

	p-Werte
R12– R 34	0.5597
R12– Frage	0.2055
R34– Frage	0.1013

Tab.II.C.20. Mittelwerte (MW), Standardfehler (Stdf.), jeweils gerundet auf die zweite Dezimale, sowie Anzahlen der Befragten (Anz.) für die drei Gruppen bezüglich der Sicherheit bei richtiger Antwort zur *Aufgabe zur Wahrscheinlichkeit von Münzwurfergebnissen*. Außerdem Vergleiche der Mittelwerte zwischen den Gruppen bezüglich der Sicherheiten: p-Werte der t-Tests (ungepaart).

Vergleicht man die Sicherheiten, die die Befragten bei falscher Lösung der Aufgabe angaben, lassen sich ebenfalls keine signifikanten Unterschiede zwischen den beiden Gruppen derjenigen, die die „R Aufgaben“ regelmäßig bearbeitet hatten, beobachten. Da sich nur zwei Personen in der Gruppe der Befragten, die die „R Aufgaben“ nie oder selten bearbeitet hatten, für eine falsche Lösung entschieden hatten, sind für sie der Mittelwert und Standardfehler nicht aussagekräftig. (Vgl. Abb.II.C.20 und Tab.II.C.21.)

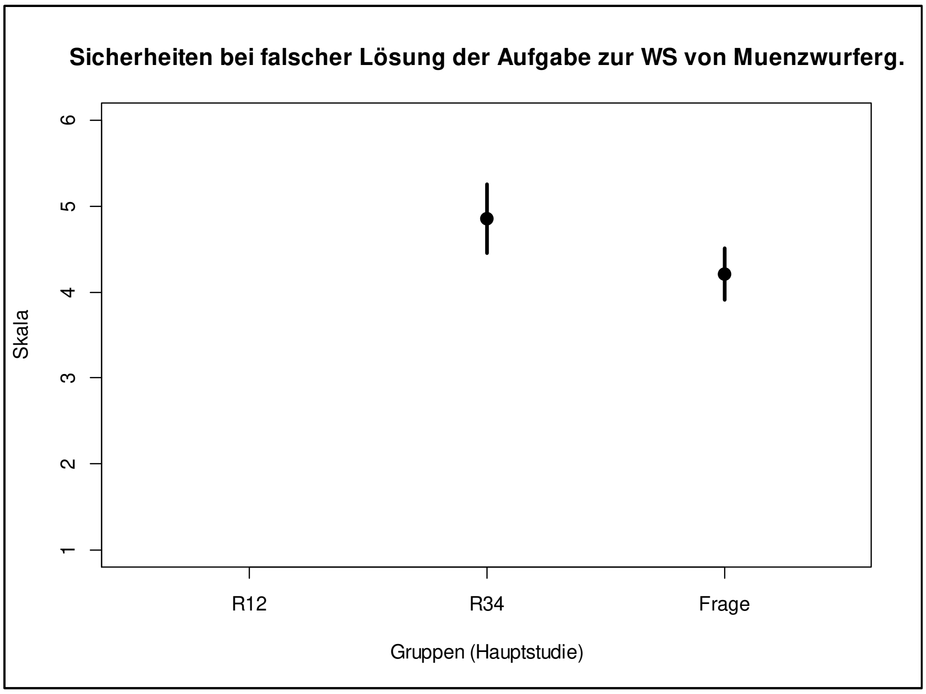


Abb.II.C.20. In der Graphik werden die Sicherheiten, die die Befragten bei falscher Lösung der *Aufgabe zur Wahrscheinlichkeit von Münzwurfergebnissen* angegeben haben, getrennt nach den Experimentalgruppen abgebildet. Es ist jeweils der Mittelwert mit Standardfehler dargestellt. Da in der Gruppe R12 nur zwei Angaben vorlagen, wird für diese Gruppe auf eine Darstellung verzichtet. Es konnten keine signifikanten Unterschiede gefunden werden.

	R12	R34	Frage
MW	(3.5)	4.86	4.21
Stdf.	(1.5)	0.40	0.30
Anz.	2	7	14

	p-Werte
R12– R 34	-
R12– Frage	-
R34– Frage	0.2243

Tab.II.C.21. Mittelwerte (MW), Standardfehler (Stdf.), jeweils gerundet auf die zweite Dezimale, sowie Anzahlen der Befragten (Anz.) für die drei Gruppen bezüglich der Sicherheit bei falscher Antwort zur *Aufgabe zur Wahrscheinlichkeit von Münzwurfergebnissen*. Außerdem Vergleiche der Mittelwerte zwischen den Gruppen bezüglich der Sicherheiten: p-Werte der t-Tests (ungepaart).

Innerhalb der Gruppen unterschieden sich die im Mittel angegebenen Sicherheiten nicht zwischen richtigen und falschen Antworten (vgl. Tab.II.C.22).

	R12	R34	Frage
richtig – falsch	-	p= 0.5413	p= 0.3239

Tab.II.C.22. Vergleiche der Mittelwerte innerhalb der Gruppen bezüglich der Sicherheiten bei richtiger und bei falscher Antwort zur *Aufgabe zur Wahrscheinlichkeit von Münzwurfergebnissen*: p-Werte der t-Tests (ungepaart).

Bei der ***Aufgabe zur Erkennung möglicherweise gefälschter Münzwürfe*** betrachten wir wieder drei Reihen aus dem Fragebogen zu Beginn des Semesters genauer, die in ähnlicher Weise auch in dem Fragebogen am Ende des Semesters verwendet wurden:

1.) 0 0 0 1 0 1 1 1 0 0 1 0 0 0 1 1 1 1 0 1 [142]
2.) 0 1 0 0 1 0 1 0 0 1 1 0 1 0 0 1 1 0 1 0 [143]
3.) 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 [144]

Die erste Reihe wurde zu Beginn des Semesters von niemandem, am Ende des Semesters von 2 Personen (etwa 2%) angekreuzt. An der Echtheit der zweiten Reihe zweifelte zu Beginn des Semesters niemand, am Ende 23 Befragte (etwa 25%). Die dritte Reihe hielten zu Beginn des Semesters 67 Befragte (etwa 74%) für eine Fälschung, am Ende des Semesters waren es 83 (etwa 91%). Vier Personen hatten die Aufgabe zu Beginn des Semesters nicht bearbeitet (keine Reihe angekreuzt und zusätzlich keine Angabe zur Sicherheit), am Ende des Semesters gaben alle Studierenden eine Einschätzung ab.[145] In Abb.II.C.21 sind die Angaben getrennt nach den Experimentalgruppen dargestellt. Bemerkenswert sind bei der zweiten Reihe die Unterschiede zwischen dem Anfang und dem Ende des Semesters. Bei allen Gruppen gab es eine signifikante Änderung des Ankreuzverhaltens. Ebenso kreuzten die beiden Gruppen derer, die die „R Aufgaben" regelmäßig bearbeitet hatten, die 3. Reihe am Ende des Semesters signifikant häufiger an als zu Beginn des Semesters. (Vgl. Tab.II.C.23.)

[142] Es handelt sich um die zweite Reihe im Fragebogen zu Beginn des Semesters. Im Fragebogen am Ende des Semesters wurde dieselbe Reihe an erster Stelle verwendet, wobei jeweils die Einsen und Nullen vertauscht wurden.
[143] Diese Reihe stand im Fragebogen zu Beginn des Semesters an der dritten Stelle und am Ende des Semesters an zweiter Stelle.
[144] Hierbei handelt es sich sowohl im Fragebogen zu Beginn als auch in dem am Ende des Semesters um die letzte Reihe. Im Fragebogen am Ende des Semesters wurden lediglich die Einsen und Nullen vertauscht.
[145] Da die Nicht-Bearbeitung möglicherweise Information tragen könnte (nämlich dass die Aufgabe nicht verstanden wurde, werden die relativen Anteile dennoch auf alle Befragten (n=91) bezogen.

Vergleicht man die Gruppen untereinander, sind zu Beginn des Semesters keine auffälligen Unterschiede zu finden. Am Ende des Semesters markierten jedoch in den Gruppen derer, die die „R Aufgaben“ regelmäßig bearbeitet hatten, die 3. Reihe signifikant häufiger als die Gruppe derer, die die „R Aufgaben“ nie oder selten gelöst hatten (vgl. Tab.II.C.24).

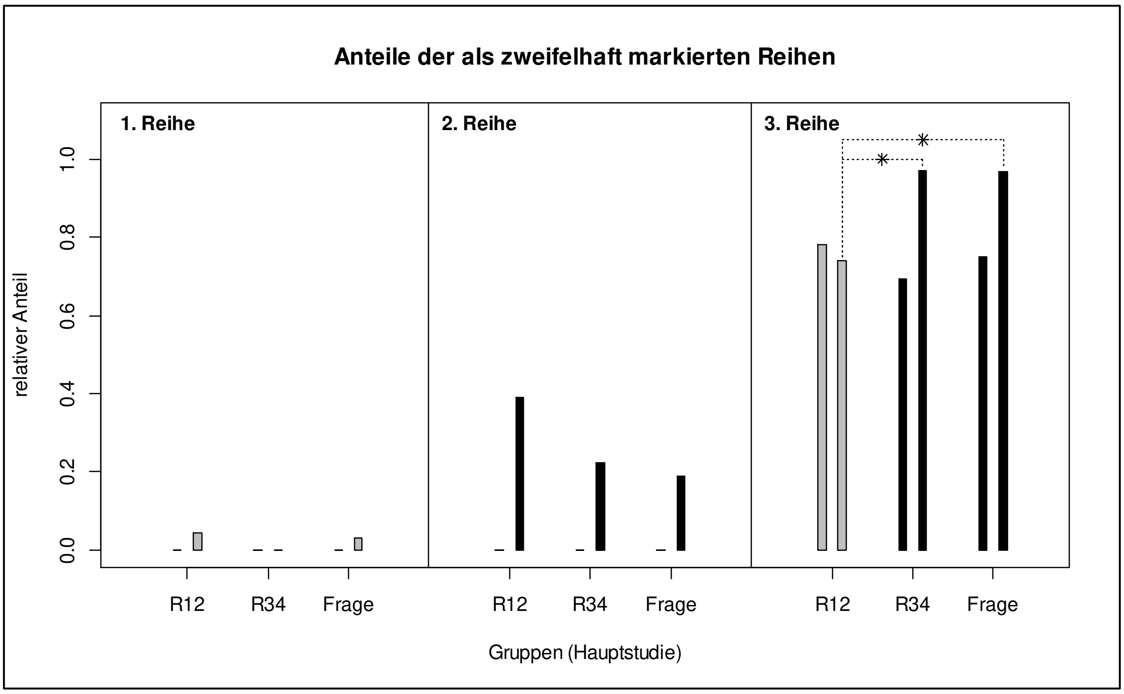


Abb.II.C.21. Die Graphik zeigt den prozentualen Anteil derer, die die Reihen als „zweifelhaft“ markiert haben, aufgeteilt nach den drei Experimentalgruppen. Die linken Säulen geben jeweils die Ergebnisse zu Beginn des Semesters und die rechten Säulen die am Ende des Semesters wieder. In den Fällen, in denen sich innerhalb der Gruppen eine signifikante Änderung des Antwortverhaltens zu Beginn und am Ende des Semesters beobachten lässt, sind die Balken schwarz, in den anderen Fällen grau. Signifikante Unterschiede zwischen den Gruppen sind mit einem Stern markiert.

1. Reihe	R12	R34	Frage
Vorher	0	0	0
Nachher	1	0	1
Anz.	23	36	32

2.Reihe	R12	R34	Frage
Vorher	0	0	0
Nachher	9	8	6
Anz.	23	36	32

3. Reihe	R12	R34	Frage
Vorher	18	25	24
Nachher	17	35	31
Anz.	23	36	32

	R12	R34	Frage
1.Reihe vorher-nachher	p=1	p=1	p=1
2.Reihe vorher-nachher	p=0.001483	p=0.005056	p=0.02417
3.Reihe vorher-nachher	p=1	p=0.002979	p=0.02648

Tab.II.C.23. Anzahlen der als „zweifelhaft“ markierten Reihen zu Beginn und am Ende des Semesters, aufgeteilt nach den Experimentalgruppen. Außerdem p-Werte von Fisher’s exakten Tests zum Vergleich der Anzahlen angekreuzter und nicht angekreuzter Reihen zu Beginn und am Ende des Semesters, aufgeteilt nach den Experimentalgruppen.

Vorher	R12 – R34	R12 – Frage	R34 – Frage
3. Reihe	p= 0.5558	p=1	p= 0.7873

Nachher	R12 – R34	R12 – Frage	R34 – Frage
2. Reihe	p=0.2388	p=0.1282	p=0.772
3. Reihe	p=0.01137	p=0.01713	p=1

Tab.II.C.24. p-Werte von Fisher's exakten Tests zum Vergleich der Anzahlen angekreuzter und nicht angekreuzter Reihen zwischen den Experimentalgruppen, einmal zu Beginn und einmal am Ende des Semesters.

Zu Beginn des Semesters löste keine der befragten Personen die Aufgabe vollständig richtig, d.h. niemand kreuzte die erste Reihe nicht an, die zweite und dritte aber schon. Daher konnte bezüglich der Sicherheiten bei richtiger Antwort kein Vergleich bezüglich der Angaben zu Beginn und am Ende des Semesters angestellt werden. Am Ende des Semesters unterschieden sich die Angaben zur Sicherheit nicht signifikant zwischen den Gruppen und alle Angaben bewegten sich durchschnittlich in der oberen Hälfte der Skala (vgl. Tab.II.C.25).

	R12 Richtig nachher	R34 Richtig nachher	Frage Richtig nachher
MW	4.83	5	4.4
Stdf.	0.31	0.38	0.4
Anz.	6	7	5

	p-Wert
R12 – R34	0.7388
R12 – Frage	0.4156
R34 – Frage	0.3029

Tab.II.C.25. Mittelwerte (MW) und Standardfehler (Stdf.), jeweils auf die zweite Dezimale gerundet, der Sicherheiten bei richtiger Antwort sowie Anzahlen der Befragten (Anz.), die bezüglich ihrer Sicherheit bei der Lösung der *Aufgabe zur Erkennung möglicherweise gefälschter Münzwürfe* eine Angabe gemacht haben. Außerdem Vergleiche der Mittelwerte zwischen den Gruppen: p-Wert des t-Tests (ungepaart).

Zu Beginn des Semesters begründeten 57 Personen ihre Entscheidung, wobei 40 von ihnen die Anzahlen von Nullen und Einsen als Erklärung heranzogen und nur eine Person eine zu starke Gleichmäßigkeit als Kriterium für eine Fälschung formulierte. Die anderen Begründungen ließen entweder den Schluss zu, dass nicht erkannt wurde, dass es sich um eine statistische Fragestellung handelte (v.a. die häufig getätigte Aussage, dass alle Folgen gleich wahrscheinlich seien) oder waren unverständlich.

Am Ende des Semesters gaben 60 Befragte eine Begründung für ihre Lösung an, wobei sich 24 Personen explizit auf die Anzahl der Vorzeichenwechsel bzw. Länge der Runs bezogen (R12: 4, R34: 13, F: 7), 14 Personen bei der letzten Reihe mit den auffällig ungleichen Anzahlen an Nullen und Einsen argumentierten (R12: 2, R34: 2, F: 10) und 8 Personen sowohl auf die Anzahl der Vorzeichenwechsel als auch auf die Anzahlen der Nullen und Einsen Bezug nahmen (R12: 0, R34: 6, F:2). Auffällig ist, dass besonders bei der Gruppe derer, die die „R Aufgaben“ häufig oder immer selbstständig bearbeitet hatten, die Anzahl der Vorzeichenwechsel als Kriterium zur Beurteilung der Glaubwürdigkeit von Münzwurfergebnissen im Bewusstsein gewesen zu sein scheint. Diejenigen, die die „R Aufgaben“ in der Fragestunde bearbeitet hatten, nannten dagegen häufiger das schwächere (und möglicherweise auch schon aus der Schule bekannte) Kriterium der Anzahlen von Nullen und Einsen. Von den Personen, die „R Aufgaben“ selten oder nie bearbeitet hatten, gaben auffällig wenige überhaupt eine (richtige) Begründung an.

Vergleicht man die Sicherheiten, die bei falschen Lösungen angegeben wurden, innerhalb der Gruppen, lässt sich bei der Gruppe derer, die die Fragestunde oft oder immer besucht hatten, beobachten, dass sie am Ende des Semesters signifikant höhere Sicherheiten angaben als am Anfang des Semesters. Zu Beginn des Semesters waren die Unterschiede zwischen der Gruppe der Besucher der Fragestunde und denjenigen, die die „R Aufgaben“ oft oder immer selbstständig bearbeitet hatten, auffällig. Ansonsten schien es keine Unterschiede zu geben. Die Angaben aller Gruppen bewegten sich durchschnittlich innerhalb der oberen Hälfte der Skala. (Vgl. Abb. II.C.22 und Tab.II.C.26.)

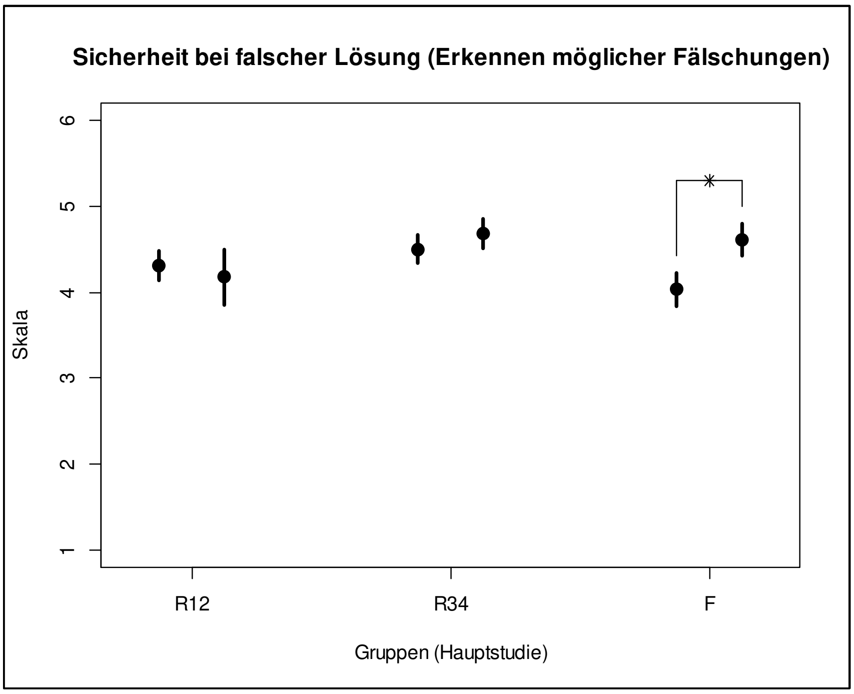


Abb.II.C.22. In der Graphik werden die Sicherheiten, die die Befragten bei falscher Lösung der *Aufgabe zur Erkennung möglicherweise gefälschter Münzwürfe* angegeben haben, getrennt nach den Experimentalgruppen abgebildet. Es ist jeweils der Mittelwert mit Standardfehler dargestellt, wobei links die Angaben zu Beginn des Semesters und rechts die Angaben am Ende des Semesters abgebildet sind. Signifikante Unterschiede sind mit einem Stern markiert.

	R12 Falsch vorher	R12 Falsch nachher	R34 Falsch vorher	R34 Falsch nachher	Frage Falsch vorher	Frage Falsch nachher
MW	4.30	4.18	4.5	4.68	4.03	4.62
Stdf.	0.17	0.32	0.16	0.17	0.19	0.18
Anz.	23	17	32	28	30	26

vorher-nachher	p-Wert
R12	0.7298
R34	0.4514
Frage	0.03446

Vorher	p-Wert
R12 – R34	0.4106
R12 – Frage	0.301
R34 – Frage	0.07047

Nachher	p-Wert
R12 – R34	0.1822
R12 – Frage	0.2492
R34 – Frage	0.8027

Tab.II.C.26. Mittelwerte (MW) und Standardfehler (Stdf.), jeweils auf die zweite Dezimale gerundet, der Sicherheiten bei falscher Antwort sowie Anzahlen der Befragten (Anz.), die bezüglich ihrer Sicherheit bei der Lösung der *Aufgabe zur Erkennung möglicherweise gefälschter Münzwürfe* eine Angabe gemacht haben. Außerdem Vergleiche der Mittelwerte innerhalb der Gruppen zu Beginn und am Ende des Semesters sowie Vergleiche zwischen den Gruppen: p-Werte der t-Tests (ungepaart).

Vergleicht man die Sicherheiten am Ende des Semesters zwischen denjenigen mit richtiger Antwort und denjenigen mit falscher Antwort, lässt sich kein auffälliger Unterschied innerhalb der Gruppen beobachten (vgl. Tab.II.C.27).

	p-Wert
R12: Richtig nachher – Falsch nachher	0.1603
R34: Richtig nachher – Falsch nachher	0.4592
Frage: Richtig nachher – Falsch nachher	0.6427

Tab.II.C.27. Vergleiche der Mittelwerte der Sicherheiten innerhalb der Gruppen bei korrekter und falscher Antwort am Ende des Semesters: p-Wert des t-Tests (ungepaart).

Die ***Aufgabe zum Testen eines Würfels*** wurde den Befragten wieder nur zu Beginn des Semesters vorgelegt. Ebenso wie in der Vorstudie wurden alle Antworten in dem Freitextfeld, die auf die Idee eines wiederholten Würfelns als Testverfahren für die Fairness des Würfels schließen ließen, als richtig gewertet. Antworten, dass einmal würfeln genüge, man nur die Anzahlen an geraden und ungeraden Augenzahlen vergleichen müsse, es nicht möglich sei, die Fairness zu prüfen oder unverständliche Antworten, wurden als nicht richtig gewertet. 73 Befragte (etwa 80%) lösten die Aufgabe richtig, 14 Personen (etwa 15%) gaben eine falsche Antwort und 4 Studierende (etwa 4%) bearbeiteten die Aufgabe nicht. Vergleicht man die Angaben zwischen den verschiedenen Gruppen, lässt sich ein signifikant niedrigerer Anteil richtiger Antworten bei den Besuchern der Fragestunde im Vergleich zur Gruppe derer beobachten, die die „R Aufgaben“ nie oder selten bearbeitet hatten (vgl. Abb.II.C.23 und Tab.II.C.28).

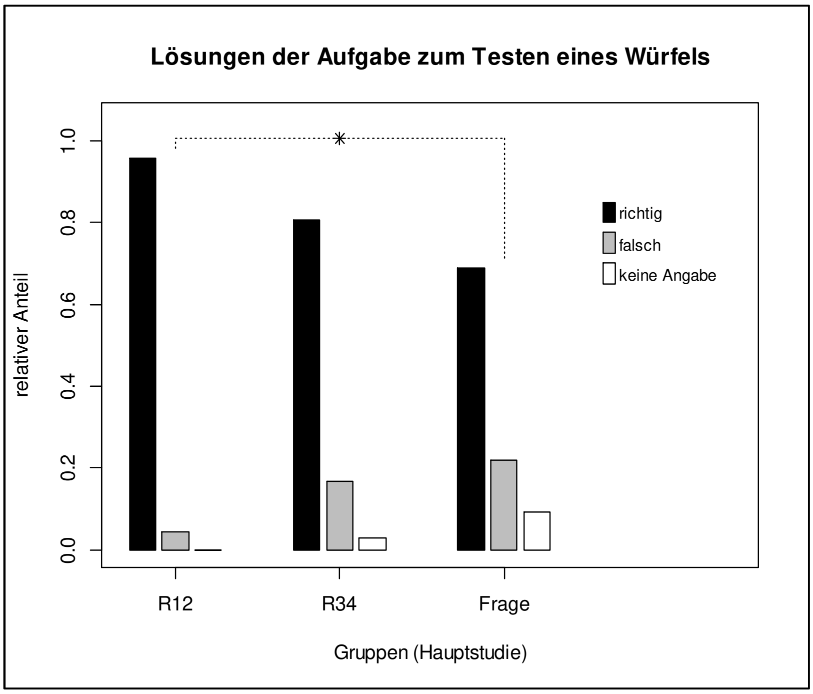


Abb.II.C.23. Die Graphik zeigt die relativen Anteile der Lösungen der *Aufgabe zum Testen eines Würfels*, aufgeteilt nach den Experimentalgruppen. Signifikante Unterschiede sind mit einem Stern markiert.

	Richtig	Falsch	keine Angabe
R12	22	1	0
R34	29	6	1
Frage	22	7	3

R12 - R34	p = 0.1331
R12 – Frage	p = 0.01718
R34 – Frage	p = 0.2794

Tab.II.C.28. Angegebene Lösungen zur *Aufgabe zum Testen eines Würfels*, getrennt nach den Experimentalgruppen, außerdem p-Werte aus Fisher’s exakten Tests zur Untersuchung von möglichen Unterschieden (richtig/falsch bzw. keine Angabe) im Antwortverhalten der Gruppen (Gruppe x /Gruppe y).

Die Sicherheiten, die die Befragten angaben, bewegten sich im Durchschnitt alle innerhalb der oberen Hälfte der Skala. Zwischen richtigen und falschen Lösungen ließen sich innerhalb der Gruppen keine auffälligen Unterschiede beobachten. Vergleicht man die Sicherheiten, die für richtige bzw. falsche Lösungen angegeben wurden, zwischen den Gruppen, kann man bei den richtigen Lösungen zwischen der Gruppe der Besucher der Fragestunde und den anderen beiden Gruppen signifikante Unterschiede feststellen. (Vgl. Abb.II.C.24 und Tab.II.C.29.)

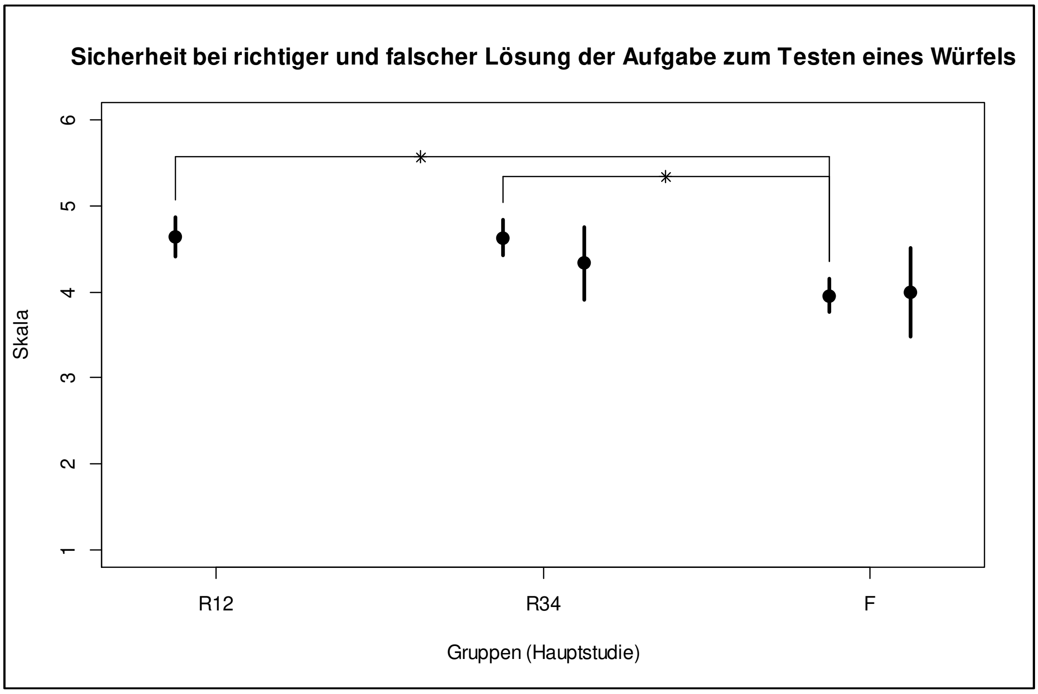


Abb.II.C.24. In der Graphik sind die angegebenen Sicherheiten für richtige Lösungen (jeweils links) und für falsche Lösungen (jeweils rechts) der *Aufgabe zum Testen eines Würfels* dargestellt, wobei jeweils die Mittelwerte und Standardfehler abgebildet sind. Signifikante Unterschiede sind mit einem Stern markiert.

	R12 Richtig	R12 Falsch	R34 Richtig	R34 Falsch	Frage Richtig	Frage Falsch
MW	4.64	(5)	4.63	4.33	3.95	4
Stdf.	0.23	-	0.21	0.42	0.19	0.52
Anz.	22	1	27	6	22	6

	p-Wert
R12: Richtig – Falsch	-
R34: Richtig – Falsch	0.5469
Frage: Richtig - Falsch	0.9367

Richtig	p-Wert
R12 – R34	0.9829
R12 – Frage	0.02926
R34 – Frage	0.02097

Falsch	p-Wert
R12 – R34	-
R12 – Frage	-
R34 – Frage	0.6283

Tab.II.C.29. Mittelwerte (MW) und Standardfehler (Stdf.), jeweils auf die zweite Dezimale gerundet, sowie Anzahlen der Befragten (Anz.) bezüglich der Sicherheiten bei richtiger bzw. falscher Antwort zu der *Aufgabe zum Testen eines Würfels*, sowie Vergleiche der Mittelwerte bezüglich der Sicherheiten innerhalb und zwischen den Gruppen: p-Wert der t-Tests (ungepaart).

Die ***Post-Aufgabe*** bearbeiteten alle Befragten. 72 Studierende (etwa 79%) gaben die richtige Lösung an und 19 Personen (etwa 21%) beantworteten sie falsch. Der Anteil an richtigen Lösungen in der Gruppe derer, die die Fragestunde regelmäßig besucht hatten, lag signifikant unter den Anteilen richtiger Lösungen in den anderen beiden Gruppen (vgl. Abb.II.C.25 und Tab.II.C.30).

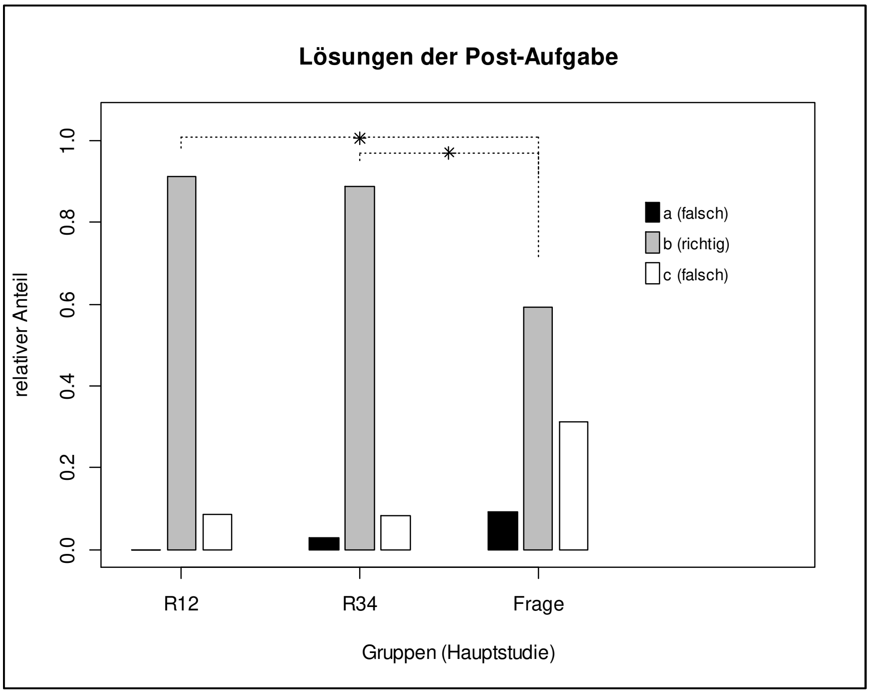


Abb.II.C.25. Die Graphik zeigt die relativen Anteile der Lösungen der *Post-Aufgabe*, aufgeteilt nach den Experimentalgruppen. Signifikante Unterschiede sind mit einem Stern markiert.

	Richtig	Falsch
R12	21	2
R34	32	4
Frage	19	13

R12 - R34	p = 1
R12 – Frage	p = 0.01304
R34 – Frage	p = 0.01024

Tab.II.C.30. Angegebene Lösungen zur *Post-Aufgabe*, getrennt nach den Experimentalgruppen, außerdem p-Werte aus Fisher's exakten Tests zur Untersuchung von möglichen Unterschieden (richtig/falsch) im Antwortverhalten der Gruppen (Gruppe x /Gruppe y).

Die Sicherheiten, die die Befragten für ihre richtige Lösung angaben, unterschieden sich nicht signifikant zwischen den Gruppen (vgl. Abb.II.C.26 und Tab.II.C.31). Die Mittelwerte lagen alle in der oberen Hälfte der Skala. Ein Vergleich zwischen den angegebenen Sicherheiten für richtige und falsche Lösungen war nur bei der Gruppe derer, die die Fragestunde regelmäßig besucht hatten, sinnvoll, da in den anderen beiden Gruppen fast keine falschen Lösungen angegeben wurden. Bei dem Vergleich innerhalb der Gruppe der Befragten, die die Fragestunde regelmäßig besucht hatten, zeigte sich, dass für falsche Lösungen im Mittel eine signifikant niedrigere Sicherheit angegeben wurde als für richtige.

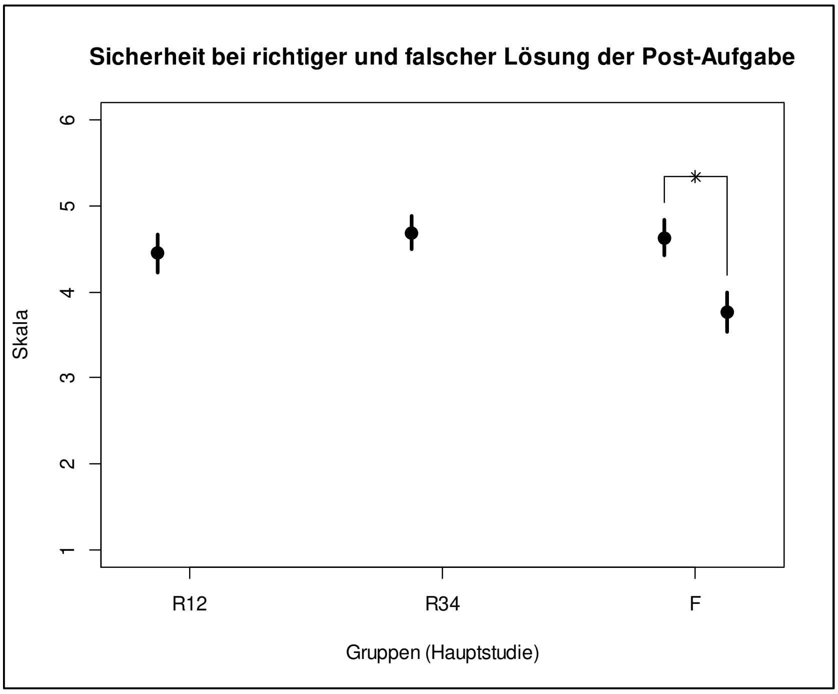


Abb.II.C.26. In der Graphik sind die angegebenen Sicherheiten für richtige Lösungen (jeweils links) und für falsche Lösungen (jeweils rechts) der *Post-Aufgabe* dargestellt, wobei jeweils die Mittelwerte und Standardfehler abgebildet sind. Signifikante Unterschiede sind mit einem Stern markiert. Bei sehr kleinen Fallzahlen wurde auf eine Darstellung verzichtet.

	R12 Richtig	R12 Falsch	R34 Richtig	R34 Falsch	Frage Richtig	Frage Falsch
MW	4.45	(4)	4.69	(4.25)	4.63	3.77
Stdf.	0.22	(1)	0.19	(0.48)	0.21	0.23
Anz.	20	2	32	4	19	13

	p-Wert
R12: Richtig – Falsch	-
R34: Richtig – Falsch	-
Frage: Richtig - Falsch	0.009483

Richtig	p-Wert
R12 – R34	0.4252
R12 – Frage	0.5531
R34 – Frage	0.8435

Falsch	p-Wert
R12 – R34	-
R12 – Frage	-
R34 – Frage	-

Tab.II.C.31. Mittelwerte (MW) und Standardfehler (Stdf.), jeweils auf die zweite Dezimale gerundet, sowie Anzahlen der Befragten (Anz.) bezüglich der Sicherheiten bei richtiger bzw. falscher Antwort zu der *Post-Aufgabe*, die eine Angabe gemacht haben, sowie Vergleiche der Mittelwerte bezüglich der Sicherheiten innerhalb und zwischen den Gruppen: p-Wert der t-Tests (ungepaart).

Die ***Multiple-Choice-Test-Aufgabe*** bearbeiteten 89 der 91 Befragten. Die Mehrheit gab eine falsche Lösung an (50 Studierende, etwa 55%). 39 Studierende (etwa 43%) beantworteten die Aufgabe richtig. Der Anteil an richtigen Lösungen in der Gruppe derer, die die Fragestunde regelmäßig besucht hatten, lag signifikant unter dem Anteil richtiger Lösungen in der Gruppe der Personen, die die „R Aufgaben“ oft oder immer eigenständig gelöst hatten. Zu denen, die die „R Aufgaben“ selten oder nie bearbeitet hatten, ließ sich kein Unterschied feststellen.

(Vgl. Abb.II.C.27 und Tab.II.C.32). Bei den falschen Lösungen kamen die der Fehlvorstellung F2(b) zuzuordnende Antwort c) zwar etwas häufiger vor als die andere falsche Lösung, das Bild ist aber keinesfalls so eindeutig wie bei der *Krankenhaus*- bzw. *Grundschule*-Aufgabe.

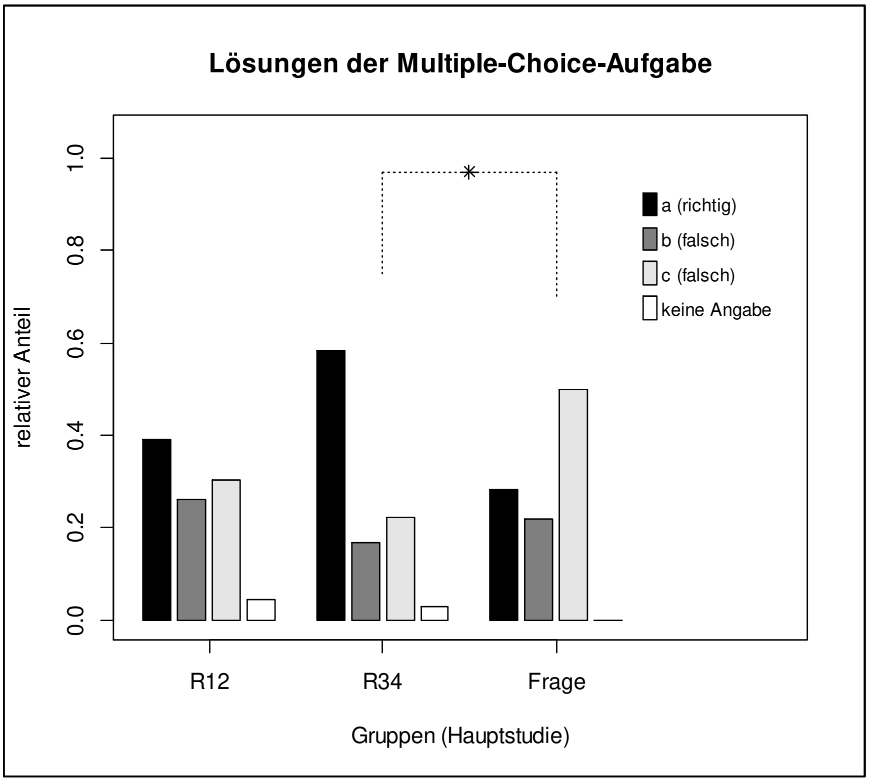


Abb.II.C.27. Die Graphik zeigt die relativen Anteile der Lösungen der *Multiple-Choice-Test-Aufgabe*, aufgeteilt nach den Experimentalgruppen. Signifikante Unterschiede sind mit einem Stern markiert.

	Richtig	falsch/kA
R12	9	14
R34	21	15
Frage	9	23

R12 - R34	p = 0.1872
R12 – Frage	p = 0.5609
R34 – Frage	p = 0.01549

Tab.II.C.32. Angegebene Lösungen zur *Multiple-Choice-Test-Aufgabe*, getrennt nach den Experimentalgruppen, außerdem p-Werte aus Fisher's exakten Tests zur Untersuchung von möglichen Unterschieden (richtig/falsch bzw. keine Angabe) im Antwortverhalten der Gruppen (Gruppe x /Gruppe y).

Die Sicherheiten, die die Befragten für ihre Lösungen angaben, unterschieden sich weder bei den richtigen noch bei den falschen Antworten signifikant zwischen den Gruppen (vgl. Abb.II.C.28 und Tab.II.C.33). Die Mittelwerte lagen alle im mittleren Bereich der Skala. Innerhalb der Gruppen unterschieden sich die angegebenen Sicherheiten für richtige und falsche Lösungen nur in der Gruppe derer, die die Fragestunde regelmäßig besucht hatten, signifikant: Sie gaben im Mittel für falsche Antworten eine geringere Sicherheit an als für richtige.

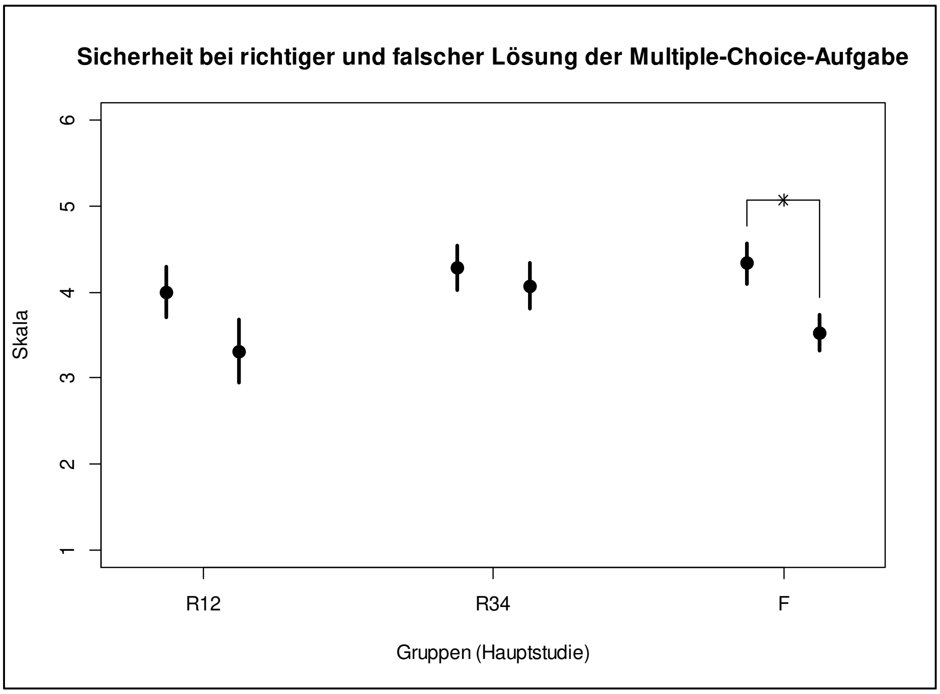


Abb.II.C.28. In der Graphik sind die angegebenen Sicherheiten für richtige Lösungen (jeweils links) und für falsche Lösungen (jeweils rechts) der *Multiple-Choice-Test-Aufgabe* dargestellt, wobei jeweils die Mittelwerte und Standardfehler abgebildet sind. Signifikante Unterschiede sind mit einem Stern markiert.

	R12 Richtig	R12 Falsch	R34 Richtig	R34 Falsch	Frage Richtig	Frage Falsch
MW	4	3.31	4.29	4.07	4.33	3.52
Stdf.	0.29	0.36	0.26	0.27	0.24	0.21
Anz.	9	13	21	14	9	23

	p-Wert
R12: Richtig – Falsch	0.1524
R34: Richtig – Falsch	0.5687
Frage: Richtig - Falsch	0.01738

Richtig	p-Wert
R12 – R34	0.4699
R12 – Frage	0.3849
R34 – Frage	0.8931

Falsch	p-Wert
R12 – R34	0.1049
R12 – Frage	0.6156
R34 – Frage	0.1149

Tab.II.C.33. Mittelwerte (MW) und Standardfehler (Stdf.), jeweils auf die zweite Dezimale gerundet, sowie Anzahlen der Befragten (Anz.) bezüglich der Sicherheiten bei richtiger bzw. falscher Antwort zu der *Multiple-Choice-Test-Aufgabe*, sowie Vergleiche der Mittelwerte bezüglich der Sicherheiten innerhalb und zwischen den Gruppen: p-Wert der t-Tests (ungepaart).

II.C.4.b. Einschätzungen der eigenen Kompetenz und Einstellungen zur Stochastik

Zunächst wird wieder die ***Einschätzung der eigenen Kompetenz im Fach Stochastik***, abgefragt durch die 6 Items E1, E2, E3, E4, E6 und E7, ausgewertet. Die Ergebnisse werden getrennt nach den Experimentalgruppen abgebildet und sowohl auf Unterschiede zwischen den Gruppen als auch auf individuelle Entwicklungen innerhalb der Gruppen untersucht.

E1: „Mein Vorwissen aus der Schule schätze ich als sehr gut ein." / „Mein Wissen in Stochastik schätze ich als sehr gut ein."

Im Mittel schätzten die Befragten ihr (Vor-)Wissen in Stochastik als mittelgut ein (zwischen 3 und 4 auf der Skala). Ein Vergleich der Mittelwerte in den unterschiedlichen Gruppen mit Hilfe von t-Tests wies darauf hin, dass die Studierenden, die die Fragestunde besuchten, eine signifikant niedrigere Selbsteinschätzung bezüglich ihres Vorwissens als die anderen (R12 und R34) hatten. Am Ende des Semesters bestand nur noch zu denen, die die „R Aufgaben" häufig oder immer ohne den Besuch der Fragestunde bearbeitet hatten (R34), ein auffällig niedrigerer Wert bei der Selbsteinschätzung. Innerhalb der Gruppen ließ sich kein signifikanter Hinweis auf eine Veränderung feststellen. (Vgl. Abb.II.C.29 und Tab.II.C.34 sowie Tab.II.C.35)

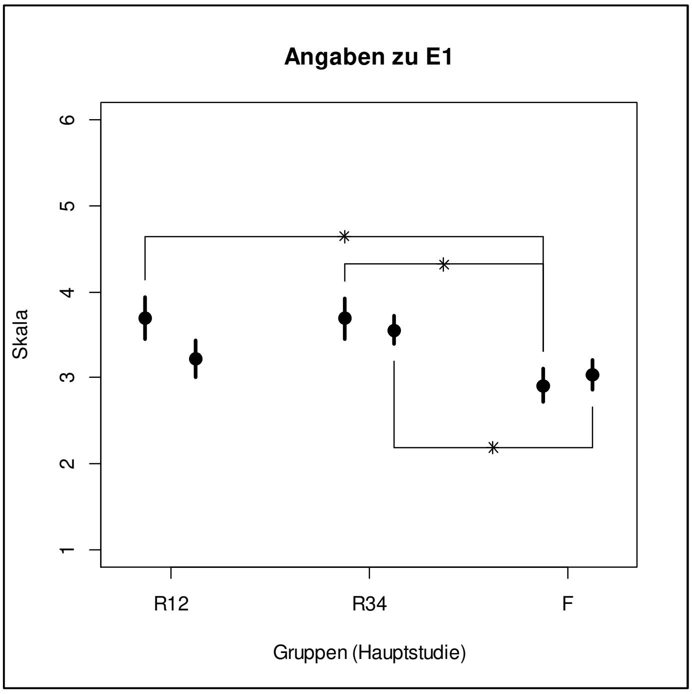


Abb.II.C.29. In der Graphik sind die Selbsteinschätzungen der Befragten, aufgeteilt nach den Experimentalgruppen, im Hinblick auf ihre Zustimmung zu E1 (Einschätzung des eigenen Wissens im Fach Stochastik) dargestellt, wobei links jeweils der Mittelwert der Einschätzungen zu Beginn des Semesters mit Standardfehler und rechts der Mittelwert der Einschätzungen am Ende des Semesters mit Standardfehler abgebildet ist. Bei E1 bedeutet ein hoher Wert, dass die Person ihre Kompetenz im Hinblick auf dieses Item als gut einschätzt. Signifikante Unterschiede sind mit einem Stern markiert.

	R12 V	R12 N	R34 V	R34 N	Frage V	Frage N
MW	3.7	3.22	3.69	3.56	2.91	3.03
Stdf.	0.25	0.22	0.23	0.17	0.19	0.17
Anz.	23	23	35	36	32	32

	p-Werte
R12 V – R34 V	0.9767
R12 V – Frage V	0.0153
R34 V – Frage V	0.01181

	p-Werte
R12 N – R34 N	0.2277
R12 N – Frage N	0.5043
R34 N – Frage N	0.03377

Tab.II.C.34. Mittelwerte (MW) und Standardfehler (Stdf.), jeweils auf die zweite Dezimal gerundet, sowie Anzahlen der Befragten (Anz.), sofern sie bezüglich der Einschätzung E1 eine Angabe gemacht haben. Außerdem Vergleiche der Mittelwerte zwischen den Gruppen bezüglich der Einschätzung E1: p-Werte der t-Tests (ungepaart).

	p-Werte (ungepaart)	Mittelwerte der individuellen Änderung (nachher-vorher)	p-Werte (gepaart)
R12 V – R12 N	0.1533	-0.48	0.11
R34 V – R34 N	0.6523	-0.09	0.6807
Frage V – Frage N	0.6289	0.13	0.525

Tab.II.C.35. Vergleiche der Mittelwerte innerhalb der Gruppen bezüglich der Einschätzung E1: p-Werte der t-Tests (ungepaarte Werte, gerundete Mittelwerte der individuellen Änderungen und gepaarte Werte).

E2: „In Stochastik kommt meistens etwas anderes raus, als man vermutet."

Die Zustimmungswerte zum Item E2 bewegten sich bei allen Gruppen sowohl zu Beginn als auch am Ende des Semesters im mittleren Bereich der Skala, wobei auf der individuellen Ebene keine auffälligen Veränderungen innerhalb der Gruppen zu beobachten waren. Vergleicht man am Ende des Semesters die Einschätzungen derer, die die „R Aufgaben" nie oder selten bearbeitet hatten, mit denen derer, die die Aufgaben regelmäßig bearbeitet hatten, erhält man einen auffällig, aber nicht signifikant höheren Wert bei der zweiten Gruppe. (Vgl. Abb.II.C.30 und Tab.II.C.36 sowie Tab.II.C.37.)

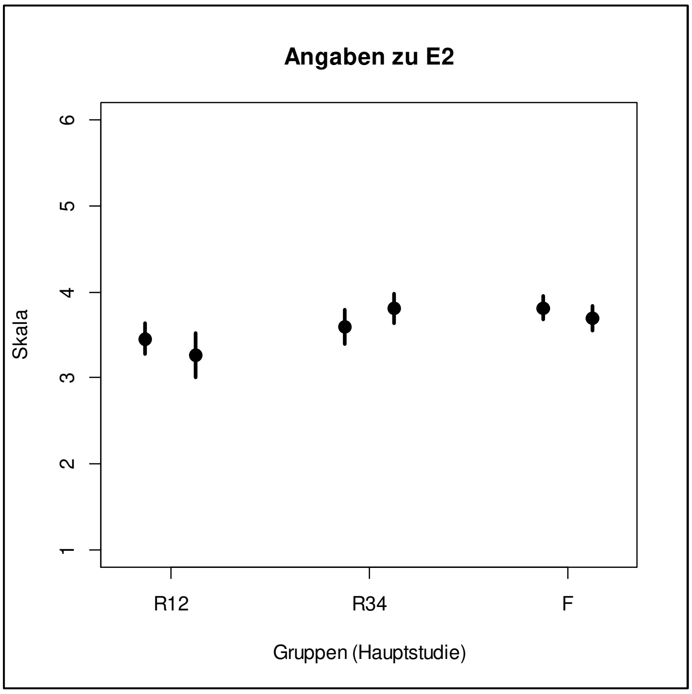


Abb.II.C.30. In der Graphik sind die Selbsteinschätzungen der Befragten, aufgeteilt nach den Experimentalgruppen, im Hinblick auf ihre Zustimmung zu E2 (Aussage, dass in Stochastik meistens etwas anderes herauskommt, als man vermutet) dargestellt, wobei links jeweils der Mittelwert der Einschätzungen zu Beginn des Semesters mit Standardfehler und rechts der Mittelwert der Einschätzungen am Ende des Semesters mit Standardfehler abgebildet ist. Bei E2 bedeutet ein hoher Wert, dass die Person ihre Kompetenz im Hinblick auf dieses Item als schlecht einschätzt. Es konnten keine signifikanten Unterschiede gefunden werden.

	R12 V	R12 N	R34 V	R34 N	Frage V	Frage N
MW	3.45	3.26	3.59	3.81	3.81	3.69
Stdf.	0.18	0.25	0.2	0.17	0.14	0.15
Anz.	22	23	34	36	32	32

	p-Werte
R12 V – R34 V	0.6225
R12 V – Frage V	0.1256
R34 V – Frage V	0.3575

	p-Werte
R12 N – R34 N	0.08056
R12 N – Frage N	0.1524
R34 N – Frage N	0.5968

Tab.II.C.36. Mittelwerte (MW) und Standardfehler (Stdf.), jeweils auf die zweite Dezimal gerundet, sowie Anzahlen der Befragten (Anz.), sofern sie bezüglich der Einschätzung E2 eine Angabe gemacht haben. Außerdem Vergleiche der Mittelwerte zwischen den Gruppen bezüglich der Einschätzung E2: p-Werte der t-Tests (ungepaart).

	p-Werte (ungepaart)	Mittelwerte der individuellen Änderung (nachher-vorher)	p-Werte (gepaart)
R12 V – R12 N	0.5387	-0.09	0.724
R34 V – R34 N	0.4067	0.21	0.2925
Frage V – Frage N	0.5346	-0.13	0.2922

Tab.II.C.37. Vergleiche der Mittelwerte innerhalb der Gruppen bezüglich der Einschätzung E2: p-Werte der t-Tests (ungepaarte Werte, gerundete Mittelwerte der individuellen Änderungen und gepaarte Werte).

E3: „Bei Stochastik versagt meine Intuition."

Der Aussage, dass bei Stochastik ihre Intuition versagt, stimmten die Befragten aller Gruppen durchschnittlich mit mittlerer Stärke zu, wobei sowohl zu Beginn als auch am Ende des Semesters die Personen, die oft oder immer die Fragestunde besucht hatten, im Mittel stärker zustimmten als diejenigen, die die „R Aufgaben" nie oder selten bearbeitet hatten. Auffällige, aber nicht signifikante Unterschiede gab es zwischen jenen, die die „R Aufgaben" oft oder immer eigenständig bearbeitet hatten und den anderen Gruppen zu verschiedenen Zeitpunkten: Zu Beginn des Semesters unterschieden sie sich signifikant von den regelmäßigen Besuchern der Fragestunde, am Ende des Semesters dagegen von der Gruppe, die die „R Aufgaben" selten oder nie bearbeitet hatten. (Vgl. Abb.II.C.31 und Tab.II.C.38 sowie Tab.II.C.39.)

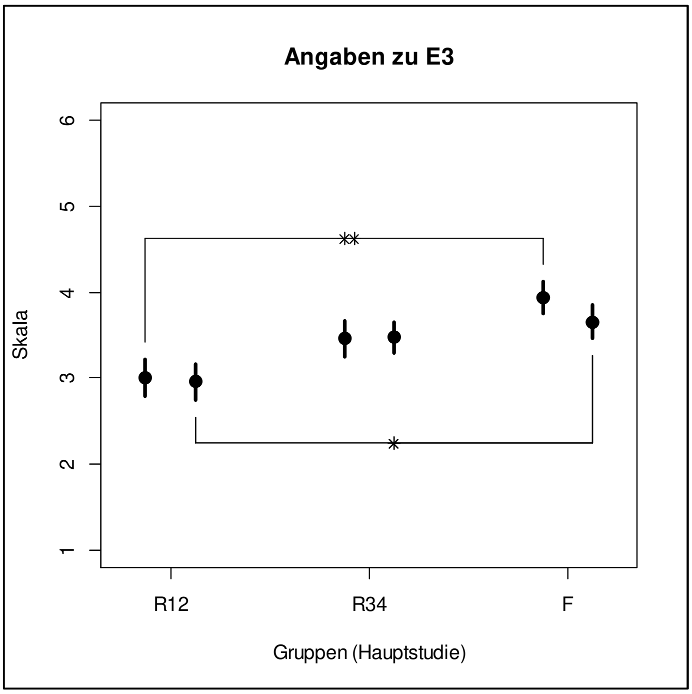


Abb.II.C.31. In der Graphik sind die Selbsteinschätzungen der Befragten, aufgeteilt nach den Experimentalgruppen, im Hinblick auf ihre Zustimmung zu E3 (Einschätzung des Versagens der eigenen Intuition) dargestellt, wobei links jeweils der Mittelwert der Einschätzungen zu Beginn des Semesters mit Standardfehler und rechts der Mittelwert der Einschätzungen am Ende des Semesters mit Standardfehler abgebildet ist. Bei E3 bedeutet ein hoher Wert, dass die Person ihre Kompetenz im Hinblick auf dieses Item als schlecht einschätzt. Signifikante Unterschiede sind mit einem Stern markiert (wobei ein Stern einen p-Wert kleiner als 5% anzeigt und zwei Sterne auf einen p-Wert kleiner als 1% hinweisen).

	R12 V	R12 N	R34 V	R34 N	Frage V	Frage N
MW	3	2.95	3.46	3.47	3.94	3.66
Stdf.	0.22	0.21	0.21	0.18	0.19	0.19
Anz.	23	22	35	36	31	32

	p-Werte
R12 V – R34 V	0.1337
R12 V – Frage V	0.001996
R34 V – Frage V	0.0893

	p-Werte
R12 N – R34 N	0.06982
R12 N – Frage N	0.01728
R34 N – Frage N	0.4829

Tab.II.C.38. Mittelwerte (MW) und Standardfehler (Stdf.), jeweils auf die zweite Dezimal gerundet, sowie Anzahlen der Befragten (Anz.), sofern sie bezüglich der Einschätzung E3 eine Angabe gemacht haben. Außerdem Vergleiche der Mittelwerte zwischen den Gruppen bezüglich der Einschätzung E3: p-Werte der t-Tests (ungepaart).

	p-Werte (ungepaart)	Mittelwerte der individuellen Änderung (nachher-vorher)	p-Werte (gepaart)
R12 V – R12 N	0.8821	-0.09	0.4923
R34 V – R34 N	0.9563	0	1
Frage V – Frage N	0.2946	-0.23	0.2288

Tab.II.C.39. Vergleiche der Mittelwerte innerhalb der Gruppen bezüglich der Einschätzung E3: p-Werte der t-Tests (ungepaarte Werte, gerundete Mittelwerte der individuellen Änderungen und gepaarte Werte).

E4: „Ich kann mir nur sehr schwer vorstellen, was Zufall bedeuten soll.“

Bezüglich des Vorstellungsvermögens von Zufall schätzten sich die regelmäßigen Besucher der Fragestunde sowohl zu Beginn als auch am Ende des Semesters signifikant oder zumindest auffällig schlechter ein als die Personen aus den anderen beiden Gruppen, wobei auch sie sich nach eigener Einschätzung Zufall eher gut vorstellen konnten. Die anderen beiden Gruppen schienen sich diesbezüglich nicht zu unterscheiden. Innerhalb der Gruppen gab es keine auffälligen Veränderungen im Laufe des Semesters. (Vgl. Abb.II.C.32 und Tab.II.C.40 sowie Tab.II.C.41.)

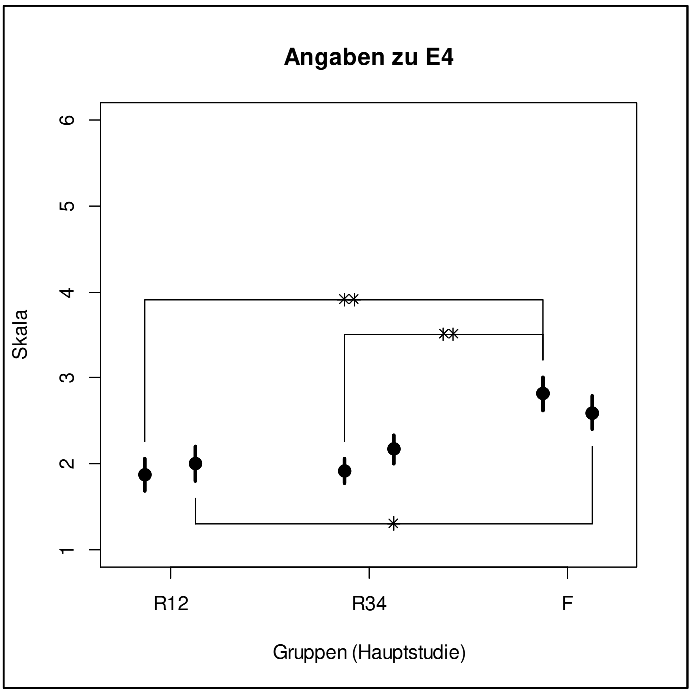


Abb.II.C.32. In der Graphik sind die Selbsteinschätzungen der Befragten, aufgeteilt nach den Experimentalgruppen, im Hinblick auf ihre Zustimmung zu E4 (Schwierigkeit bei der Vorstellung von Zufall) dargestellt, wobei links jeweils der Mittelwert der Einschätzungen zu Beginn des Semesters mit Standardfehler und rechts der Mittelwert der Einschätzungen am Ende des Semesters mit Standardfehler abgebildet ist. Bei E4 bedeutet ein hoher Wert, dass die Person ihre Kompetenz im Hinblick auf dieses Item als schlecht einschätzt. Signifikante Unterschiede sind mit einem Stern markiert (wobei ein Stern einen p-Wert kleiner als 5% anzeigt und zwei Sterne auf einen p-Wert kleiner als 1% hinweisen).

	R12 V	R12 N	R34 V	R34 N	Frage V	Frage N
MW	1.87	2	1.91	2.17	2.81	2.59
Stdf.	0.18	0.2	0.14	0.17	0.2	0.19
Anz.	23	22	35	36	32	32

	p-Werte
R12 V – R34 V	0.8477
R12 V – Frage V	0.0009167
R34 V – Frage V	0.0005348

	p-Werte
R12 N – R34 N	0.522
R12 N – Frage N	0.03498
R34 N – Frage N	0.09572

Tab.II.C.40. Mittelwerte (MW) und Standardfehler (Stdf.), jeweils auf die zweite Dezimal gerundet, sowie Anzahlen der Befragten (Anz.), sofern sie bezüglich der Einschätzung E4 eine Angabe gemacht haben. Außerdem Vergleiche der Mittelwerte zwischen den Gruppen bezüglich der Einschätzung E4: p-Werte der t-Tests (ungepaart).

	p-Werte (ungepaart)	Mittelwerte der individuellen Änderung (nachher-vorher)	p-Werte (gepaart)
R12 V – R12 N	0.6289	0.18	0.1621
R34 V – R34 N	0.2562	0.23	0.1603
Frage V – Frage N	0.4281	-0.22	0.1983

Tab.II.C.41. Vergleiche der Mittelwerte innerhalb der Gruppen bezüglich der Einschätzung E4: p-Werte der t-Tests (ungepaarte Werte, gerundete Mittelwerte der individuellen Änderungen und gepaarte Werte).

E6: „Das Lösen von Stochastik-Aufgaben fällt mir im Allgemeinen nicht schwerer als von Aufgaben aus anderen mathematischen Bereichen."

Mit Hilfe des Items E6 sollte die Kompetenz im Fach Stochastik im Vergleich zur allgemeinen mathematischen Kompetenz erfasst werden. Zu Beginn des Semesters schätzten die regelmäßigen Besucher der Fragestunde ihre Lösungsfertigkeiten von Aufgaben in Stochastik im gesamt-mathematischen Vergleich als geringer ein als die Personen aus den anderen beiden Gruppen. Am Ende des Semesters schien es keinen Unterschied zwischen denen, die die „R Aufgaben" nie oder selten bearbeitet hatten, und den regelmäßigen Besuchern der Fragestunde mehr zu geben sowie nur noch einen auffälligen, aber nicht mehr signifikanten Unterschied zwischen den Personen, die die „R Aufgaben" oft oder immer eigenständig bearbeitet hatten und der letzten Gruppe. Bemerkenswert ist zudem, dass die Personen, die die „R Aufgaben" nie oder selten bearbeitet hatten, sich in ihrer Selbsteinschätzung bezüglich des Items E6 im Laufe des Semesters auffällig (aber nicht signifikant) verschlechterten, wenn man die individuellen Angaben vergleicht. Die Zustimmungswerte bewegten sich insgesamt bei allen Gruppen durchschnittlich im mittleren Bereich. (Vgl. Abb.II.C.33 und Tab.II.C.42 sowie Tab.II.C.43.)

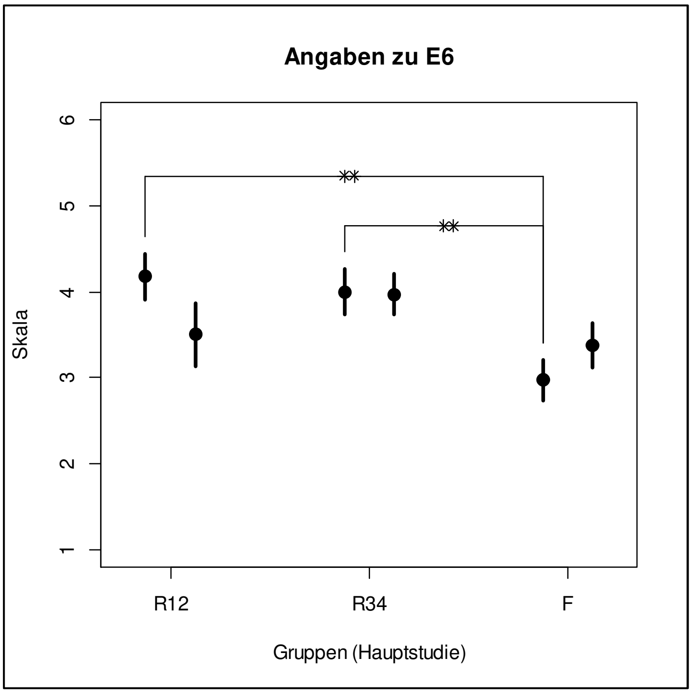


Abb.II.C.33. In der Graphik sind die Selbsteinschätzungen der Befragten, aufgeteilt nach den Experimentalgruppen, im Hinblick auf ihre Zustimmung zu E6 (Einschätzung der Lösungsfertigkeiten in Stochastik verglichen mit denen in anderen mathematischen Bereichen) dargestellt, wobei links jeweils der Mittelwert der Einschätzungen zu Beginn des Semesters mit Standardfehler und rechts der Mittelwert der Einschätzungen am Ende des Semesters mit Standardfehler abgebildet ist. Bei E6 bedeutet ein hoher Wert, dass die Person ihre Kompetenz im Hinblick auf dieses Item als gut einschätzt. Signifikante Unterschiede sind mit einem Stern markiert (wobei zwei Sterne auf einen p-Wert kleiner als 1% hinweisen).

	R12 V	R12 N	R34 V	R34 N	Frage V	Frage N
MW	4.17	3.5	4	3.97	2.97	3.38
Stdf.	0.27	0.37	0.27	0.24	0.23	0.26
Anz.	23	22	35	35	32	32

	p-Werte
R12 V – R 34 V	0.649
R12 V – Frage V	0.001453
R34 V – Frage V	0.004701

	p-Werte
R12 N – R34 N	0.2909
R12 N – Frage N	0.7831
R34 N – Frage N	0.09297

Tab.II.C.42. Mittelwerte (MW) und Standardfehler (Stdf.), jeweils auf die zweite Dezimal gerundet, sowie Anzahlen der Befragten (Anz.), sofern sie bezüglich der Einschätzung E6 eine Angabe gemacht haben. Außerdem Vergleiche der Mittelwerte zwischen den Gruppen bezüglich der Einschätzung E6: p-Werte der t-Tests (ungepaart).

	p-Werte (ungepaart)	Mittelwerte der individuellen Änderung (nachher-vorher)	p-Werte (gepaart)
R12 V – R12 N	0.1505	-0.59	0.05005
R34 V – R34 N	0.9363	-0.03	0.9209
Frage V – Frage N	0.2443	0.41	0.1515

Tab.II.C.43. Vergleiche der Mittelwerte innerhalb der Gruppen bezüglich der Einschätzung E6: p-Werte der t-Tests (ungepaarte Werte, gerundete Mittelwerte der individuellen Änderungen und gepaarte Werte).

E7: „Ich traue meinen Lösungen von Stochastik-Aufgaben im Allgemeinen nicht.“

Im Gegensatz zum Item E6, in dem eine relative Einschätzung des Vertrauens in die eigenen Fertigkeiten beim Lösen von Stochastik-Aufgaben abgefragt wurde, ging es bei diesem Item E7 um eine absolute Einschätzung. Zu Beginn des Semesters hatten die regelmäßigen Besucher der Fragestunde ein signifikant niedrigeres Vertrauen in ihre eigenen Lösungen als die Befragten aus den anderen beiden Gruppen. Am Ende des Semesters gab es nur zwischen dieser Gruppe, die oft oder immer an der Fragestunde teilgenommen hatte, und der Gruppe derer, die die „R Aufgaben“ oft oder immer eigenständig bearbeitet hatten, einen auffälligen (aber nicht signifikanten) Unterschied. Hinzu kommt, dass das Vertrauen in die eigenen Lösungen in der Gruppe der regelmäßigen Besucher der Fragestunde im Laufe des Semesters auffällig (aber nicht signifikant) gestiegen ist. Alle Skalenwerte bewegten sich durchschnittlich im mittleren Bereich. (Vgl. Abb.II.C.34 und Tab.II.C.44 sowie Tab.II.C.45.)

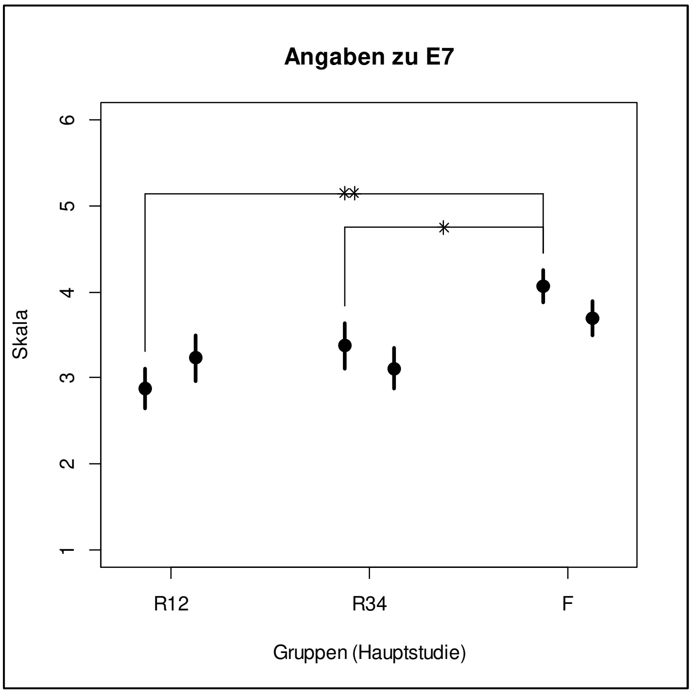


Abb.II.C.34. In der Graphik sind die Selbsteinschätzungen der Befragten, aufgeteilt nach den Experimentalgruppen, im Hinblick auf ihre Zustimmung zu E7 (Vertrauen in eigene Lösungen von Stochastik-Aufgaben) dargestellt, wobei links jeweils der Mittelwert der Einschätzungen zu Beginn des Semesters mit Standardfehler und rechts der Mittelwert der Einschätzungen am Ende des Semesters mit Standardfehler abgebildet ist. Bei E7 bedeutet ein hoher Wert, dass die Person ihre Kompetenz im Hinblick auf dieses Item als schlecht einschätzt. Signifikante Unterschiede sind mit einem Stern markiert (wobei ein Stern einen p-Wert kleiner als 5% anzeigt und zwei Sterne auf einen p-Wert kleiner als 1% hinweisen).

	R12 V	R12 N	R34 V	R34 N	Frage V	Frage N
MW	2.87	3.23	3.37	3.11	4.06	3.69
Stdf.	0.23	0.27	0.26	0.23	0.18	0.20
Anz.	23	22	35	36	32	32

	p-Werte
R12 V – R 34 V	0.1558
R12 V – Frage V	0.0001941
R34 V – Frage V	0.0355

	p-Werte
R12 N – R34 N	0.7458
R12 N – Frage N	0.181
R34 N – Frage N	0.06568

Tab.II.C.44. Mittelwerte (MW) und Standardfehler (Stdf.), jeweils auf die zweite Dezimal gerundet, sowie Anzahlen der Befragten (Anz.), sofern sie bezüglich der Einschätzung E7 eine Angabe gemacht haben. Außerdem Vergleiche der Mittelwerte zwischen den Gruppen bezüglich der Einschätzung E7: p-Werte der t-Tests (ungepaart).

	p-Werte (ungepaart)	Mittelwerte der individuellen Änderung (nachher-vorher)	p-Werte (gepaart)
R12 V – R12 N	0.3192	0.32	0.2005
R34 V – R34 N	0.4599	-0.26	0.1069
Frage V – Frage N	0.1769	-0.38	0.05637

Tab.II.C.45. Vergleiche der Mittelwerte innerhalb der Gruppen bezüglich der Einschätzung E7: p-Werte der t-Tests (ungepaarte Werte, gerundete Mittelwerte der individuellen Änderungen und gepaarte Werte).

Die ***Einstellungen zur Stochastik*** wurden wie in der Vorstudie im Hinblick auf drei Items (E5, E8 und E9) untersucht. Die Ergebnisse werden getrennt nach den Experimentalgruppen abgebildet und sowohl auf Unterschiede zwischen den Gruppen als auch auf individuelle Entwicklung innerhalb der Gruppen untersucht.

E5: „Ich mag Stochastik.“

Die Zustimmungen zu der Aussage „Ich mag Stochastik.“ bewegten sich insgesamt in allen Gruppen durchschnittlich im mittleren Bereich der Skala. Individuelle Veränderungen gab es nur innerhalb der Gruppe der Befragten, die die „R Aufgaben“ nie oder selten bearbeitet hatten: Bei ihr verringerte sich die Zustimmung im Laufe des Semesters signifikant. Zu Beginn des Semesters unterschied sich diese Gruppe signifikant von der Gruppe der regelmäßigen Besucher der Fragestunde, am Ende des Semesters war ihre Zustimmung zu der Aussage durchschnittlich nicht mehr höher. Zwischen den anderen Gruppen gab es keine auffälligen Unterschiede. (Vgl. Abb.II.C.35 und Tab.II.C.46 sowie Tab.II.C.47.)

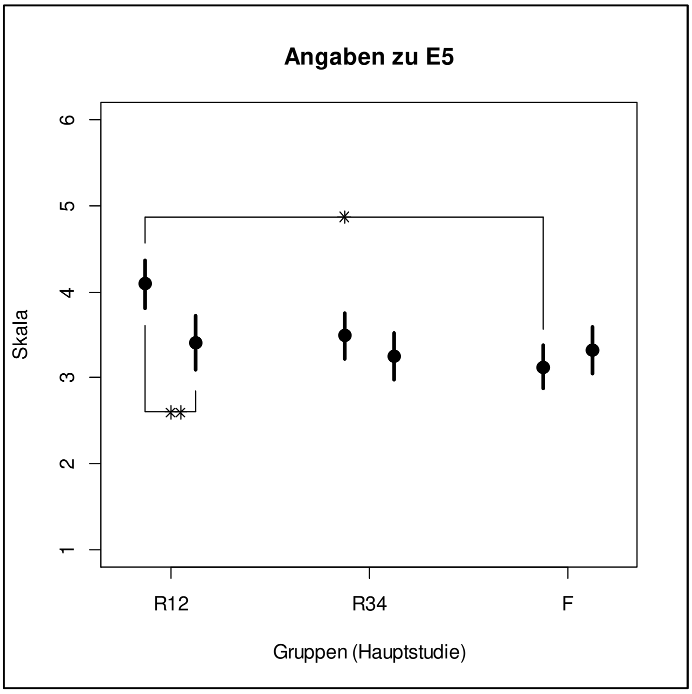


Abb.II.C.35. In der Graphik sind die Selbsteinschätzungen der Befragten, aufgeteilt nach den Experimentalgruppen, im Hinblick auf ihre Zustimmung zu E5 (Aussage „Ich mag Stochastik.“) dargestellt, wobei links jeweils der Mittelwert der Einschätzungen zu Beginn des Semesters mit Standardfehler und rechts der Mittelwert der Einschätzungen am Ende des Semesters mit Standardfehler abgebildet ist. Signifikante Unterschiede sind mit einem Stern markiert (wobei ein Stern einen p-Wert kleiner als 5% anzeigt und zwei Sterne auf einen p-Wert kleiner als 1% hinweisen).

	R12 V	R12 N	R34 V	R34 N	Frage V	Frage N
MW	4.09	3.41	3.49	3.25	3.13	3.32
Stdf.	0.28	0.31	0.26	0.27	0.24	0.27
Anz.	23	22	35	36	32	31

	p-Werte
R12 V – R 34 V	0.1224
R12 V – Frage V	0.01287
R34 V – Frage V	0.3168

	p-Werte
R12 N – R34 N	0.704
R12 N – Frage N	0.8357
R34 N – Frage N	0.8515

Tab.II.C.46. Mittelwerte (MW) und Standardfehler (Stdf.), jeweils auf die zweite Dezimal gerundet, sowie Anzahlen der Befragten (Anz.), sofern sie bezüglich der Einschätzung E5 eine Angabe gemacht haben. Außerdem Vergleiche der Mittelwerte zwischen den Gruppen bezüglich der Einschätzung E5: p-Werte der t-Tests (ungepaart).

	p-Werte (ungepaart)	Mittelwerte der individuellen Änderung (nachher-vorher)	p-Werte (gepaart)
R12 V – R12 N	0.1144	-0.64	0.003342
R34 V – R34 N	0.5352	-0.17	0.4466
Frage V – Frage N	0.5913	0.13	0.5253

Tab.II.C.47. Vergleiche der Mittelwerte innerhalb der Gruppen bezüglich der Einschätzung E5: p-Werte der t-Tests (ungepaarte Werte, gerundete Mittelwerte der individuellen Änderungen und gepaarte Werte).

E8: „Stochastik ist nützlich.“

Die Nützlichkeit von Stochastik wurde in allen Gruppen durchschnittlich als hoch bewertet. Es ließen sich weder zu Beginn noch am Ende des Semesters signifikante Unterschiede zwischen den Gruppen beobachten. Innerhalb der Gruppe derer, die die „R Aufgaben“ oft oder immer eigenständig bearbeitet hatten, ging die Zustimmung zu der Aussage im individuellen Vergleich signifikant zurück, bewegte sich aber auch am Ende noch auf einem hohen Niveau. Bei denen, die die „R Aufgaben“ nie oder selten bearbeitet hatten, war ein ähnlicher Effekt zu sehen, wobei hier der Unterschied nicht signifikant, sondern nur auffällig war. Vergleicht man die Werte innerhalb der Gruppen allerdings nicht individuell, sondern ungepaart, sind die Effekte nicht mehr zu beobachten. (Vgl. Abb.II.C.36 und Tab.II.C.48 sowie Tab.II.C.49.)

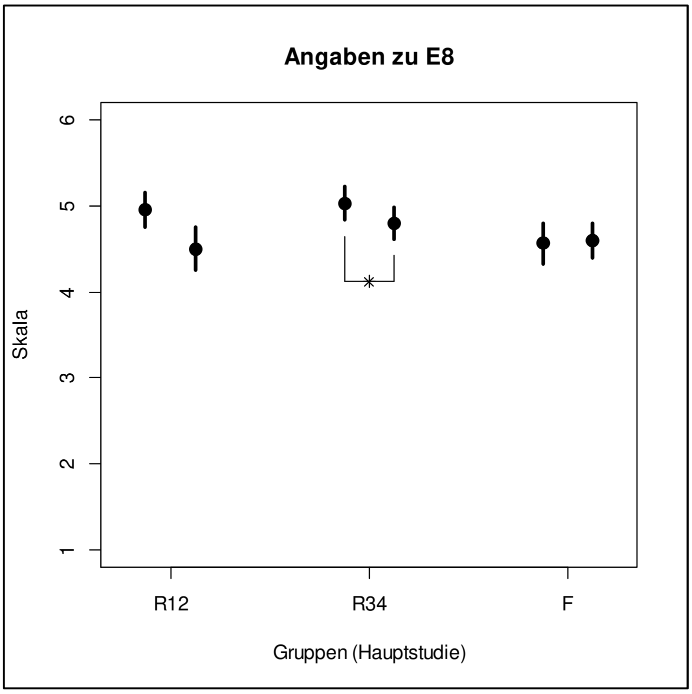


Abb.II.C.36. In der Graphik sind die Selbsteinschätzungen der Befragten, aufgeteilt nach den Experimentalgruppen, im Hinblick auf ihre Zustimmung zu E8 (Einschätzung der Nützlichkeit von Stochastik) dargestellt, wobei links jeweils der Mittelwert der Einschätzungen zu Beginn des Semesters mit Standardfehler und rechts der Mittelwert der Einschätzungen am Ende des Semesters mit Standardfehler abgebildet ist. Signifikante Unterschiede sind mit einem Stern markiert.

	R12 V	R12 N	R34 V	R34 N	Frage V	Frage N
MW	4.96	4.5	5.03	4.8	4.56	4.59
Stdf.	0.20	0.25	0.19	0.18	0.23	0.20
Anz.	23	22	35	35	32	32

	p-Werte
R12 V – R 34 V	0.799
R12 V – Frage V	0.2084
R34 V – Frage V	0.1297

	p-Werte
R12 N – R34 N	0.3413
R12 N – Frage N	0.7725
R34 N – Frage N	0.449

Tab.II.C.48. Mittelwerte (MW) und Standardfehler (Stdf.), jeweils auf die zweite Dezimal gerundet, sowie Anzahlen der Befragten (Anz.), sofern sie bezüglich der Einschätzung E8 eine Angabe gemacht haben. Außerdem Vergleiche der Mittelwerte zwischen den Gruppen bezüglich der Einschätzung E8: p-Werte der t-Tests (ungepaart).

	p-Werte (ungepaart)	Mittelwerte der individuellen Änderung (nachher-vorher)	p-Werte (gepaart)
R12 V – R12 N	0.167	-0.41	0.0829
R34 V – R34 N	0.3942	-0.23	0.04374
Frage V – Frage N	0.9193	0.03	0.8451

Tab.II.C.49. Vergleiche der Mittelwerte innerhalb der Gruppen bezüglich der Einschätzung E8: p-Werte der t-Tests (ungepaarte Werte, gerundete Mittelwerte der individuellen Änderungen und gepaarte Werte).

E9: „Stochastik ist kein typisch mathematisches Fach."

Insgesamt bewegten sich die Angaben der verschiedenen Gruppen durchschnittlich im unteren bzw. unteren mittleren Bereich der Zustimmungsskala. Im Allgemeinen hielten die Befragten das Fach Stochastik also eher für ein typisch mathematisches Fach, wobei zu Beginn des Semesters kein auffälliger Unterschied zwischen den Gruppen zu beobachten war. Bemerkenswert ist in diesem Zusammenhang allerdings, dass sich die Einstellungen diesbezüglich in den beiden Gruppen derer, die die „R Aufgaben" regelmäßig bearbeitet hatten (selbstständig oder mit Hilfe der Fragestunde), im Laufe des Semesters signifikant veränderten: Die Befragten hielten das Fach Stochastik am Ende des Semesters durchschnittlich für weniger typisch als zu Beginn des Semesters. Im Fall der Gruppe, die die „R Aufgaben" oft oder immer eigenständig gelöst hatte, führte das dann auch zu einer signifikant anderen Einschätzung am Ende des Semesters im Vergleich zu der Gruppe derer, die die „R Aufgaben" nie oder selten bearbeitet hatten. Vergleicht man die Gruppe, die regelmäßig die Fragestunde besucht hatte, mit der Gruppe der Personen, die die „R Aufgaben" nie oder selten bearbeitet hatten, ist der Unterschied am Ende des Semesters auffällig, aber nicht signifikant. (Vgl. Abb.II.C.37 und Tab.II.C.50 sowie Tab.II.C.51.)

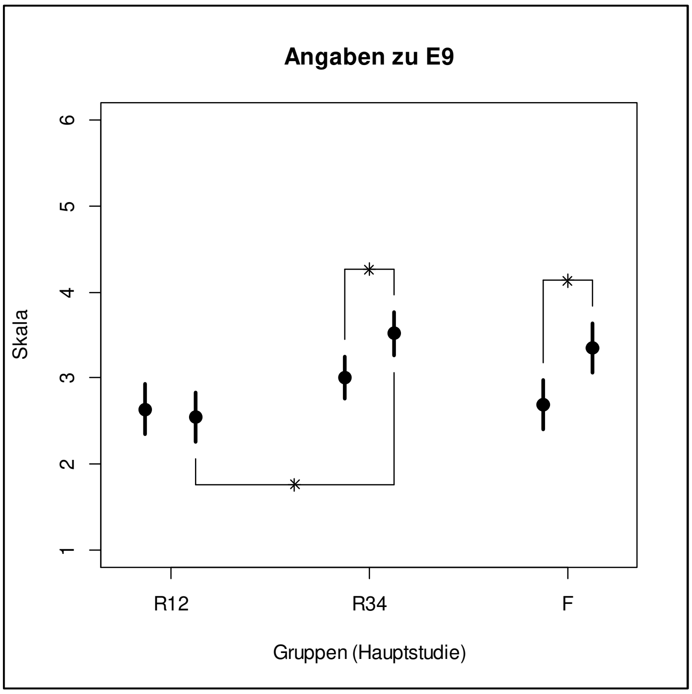


Abb.II.C.37. In der Graphik sind die Selbsteinschätzungen der Befragten, aufgeteilt nach den Experimentalgruppen, im Hinblick auf ihre Zustimmung zu E9 (Einschätzung Stochastik als nicht typisch mathematisch) dargestellt, wobei links jeweils der Mittelwert der Einschätzungen zu Beginn des Semesters mit Standardfehler und rechts der Mittelwert der Einschätzungen am Ende des Semesters mit Standardfehler abgebildet ist. Signifikante Unterschiede sind mit einem Stern markiert.

	R12 V	R12 N	R34 V	R34 N	Frage V	Frage N
MW	2.64	2.55	3	3.51	2.69	3.34
Stdf.	0.30	0.28	0.24	0.25	0.28	0.29
Anz.	22	22	35	35	32	32

	p-Werte
R12 V – R 34 V	0.3491
R12 V – Frage V	0.9014
R34 V – Frage V	0.404

	p-Werte
R12 N – R34 N	0.01337
R12 N – Frage N	0.05348
R34 N – Frage N	0.6536

Tab.II.C.50. Mittelwerte (MW) und Standardfehler (Stdf.), jeweils auf die zweite Dezimal gerundet, sowie Anzahlen der Befragten (Anz.), sofern sie bezüglich der Einschätzung E9 eine Angabe gemacht haben. Außerdem Vergleiche der Mittelwerte zwischen den Gruppen bezüglich der Einschätzung E9: p-Werte der t-Tests (ungepaart).

	p-Werte (ungepaart)	Mittelwerte der individuellen Änderung (nachher-vorher)	p-Werte (gepaart)
R12 V – R12 N	0.8266	-0.05	0.8905
R34 V – R34 N	0.1422	0.51	0.02924
Frage V – Frage N	0.1074	0.66	0.01417

Tab.II.C.51. Vergleiche der Mittelwerte innerhalb der Gruppen bezüglich der Einschätzung E9: p-Werte der t-Tests (ungepaarte Werte, gerundete Mittelwerte der individuellen Änderungen und gepaarte Werte).

Zu Beginn des Semesters gab es – ebenso wie in der Vorstudie – noch die Möglichkeit, eine **Begründung** zu den Einschätzungen bzw. Einstellungsabfragen anzugeben, was 8 Personen in Anspruch nahmen und sich zu E9 äußerten. Drei von ihnen hatten „trifft überhaupt nicht zu“ bzw. „trifft weitgehend nicht zu“ angekreuzt und begründeten das damit, dass Mathematik nicht eindeutig zu definieren sei bzw. damit, dass an der Universität Stochastik anders als in der Schule eher als „angewandte Analysis“ gelehrt werde und somit eine axiomatische Erfassung möglich sei. Die anderen fünf Personen hatten „trifft eher zu“ angekreuzt. Ihre Begründungen waren, dass Begriffe weniger formalisiert würden als in anderen Bereichen bzw. Definitionen fehlten („Wischi-waschi-Mathematik“) und weniger auf das klassische Satz-Beweis-Schema zurückgegriffen werde. Außerdem werde nicht mit „exakten Zahlen“ gerechnet bzw. der Zufall sei nicht in mathematische Formeln zu fassen. Zwei Personen gaben an, dass Stochastik sich darin von anderen mathematischen Bereichen unterscheide, dass ihnen in der Schule nicht so viele oder gar keine Kenntnisse davon vermittelt worden seien.

Am Ende des Semesters hatten die Befragten zusätzlich zu den Angaben auf den Skalen noch die Möglichkeit, besondere Schwierigkeiten bei der Auseinandersetzung mit dem Fach Stochastik zu benennen. Insgesamt notierten 38 Studierende einen oder mehrere Kommentare (R12: 6, R34: 14, F: 18), wobei sich die meisten Äußerungen nicht konkret auf das Fach

Stochastik bezogen, sondern auf allgemeine Schwierigkeiten mit der Hochschulmathematik bzw. der Struktur oder dem Material der Vorlesung (Beweise/Erfassen mathematischer Zusammenhänge/Ansatz finden: 11, formale Schreibweise: 3, Struktur der Vorlesung/Material: 12). Unter den genannten Schwierigkeiten mit dem Fach Stochastik fanden sich am häufigsten Probleme, konkrete Anschauungen und Vorstellungen zu den Konzepten (R12: 1, R34: 2, F: 2) und Verteilungsmerkmalen (F: 3) zu entwickeln bzw. theoretische Sätze oder Beweise zu verstehen, zu denen keine Anschauung vorhanden zu sein schien (R34: 1, F: 3). Begriffe, die in diesem Zusammenhang genannt wurden, waren Kovarianz (R12, R34, F), Zufallsvariable und Indikatorvariable (R34). Zwei Befragte gaben an, dass ihnen besonders eine falsche oder fehlende Intuition das Lösen von Aufgaben erschwerte (F: 2). Auf der methodischen Ebene empfanden drei Studierende eine zu wenig formale Schreibweise als Schwierigkeit (R34: 3), zwei Befragten bereitete das Schreiben von R Codes Probleme (R12, F), außerdem wurden erforderliche Modellierungen (R34) und der Umgang mit Näherungen (R34) als Schwierigkeiten beschrieben.

II.C.4.c. Zusammenfassung der Hauptstudien-Ergebnisse bezüglich der Aufgaben und Kompetenzeinschätzungen sowie Einstellungen zur Stochastik

Bezüglich der Aufgaben, die zu Beginn des Semesters bereits vorhandene Grundvorstellungen abfragten und den Studierenden auch nur zu Beginn des Semesters vorgelegt wurden, ließ sich Folgendes beobachten:

- Die ***Aufgabe zur Wahrscheinlichkeit von Münzwurfergebnissen*** wurde von insgesamt 75% der Befragten richtig beantwortet, wobei die Gruppe der Fragestundenbesucher signifikant schlechter abschnitt als die anderen beiden Gruppen. Bezüglich der angegebenen Sicherheiten gab es weder bei den richtigen noch bei den falschen Antworten auffällige Unterschiede zwischen den Gruppen und innerhalb keiner Gruppe konnte ein Unterschied zwischen den angegebenen Sicherheiten für richtige und falsche Lösungen beobachtet werden.
- Die ***Aufgabe zum Testen eines Würfels*** wurde von etwa 80% der Befragten richtig gelöst, wobei die Gruppe der Fragestundenbesucher signifikant schlechter abschnitt als die Gruppe derer, die die „R Aufgaben“ selten oder nie bearbeitetet hatten. Bezüglich der Sicherheiten wurde in keiner Gruppe ein Unterschied zwischen den Angaben zu richtigen und denen zu falschen Lösungen beobachtet. Der einzige

Unterschied zwischen den Gruppen war, dass bei richtiger Antwort die Gruppe der Fragestundenbesucher ihre Sicherheit signifikant niedriger einstufte als es in den anderen beiden Gruppen der Fall war.

- Die ***Post-Aufgabe*** lösten insgesamt etwa 79% der Befragten richtig, wobei die Gruppe der Fragestundenbesucher signifikant schwächer war als die anderen beiden Gruppen. Bezüglich der angegebenen Sicherheiten war nichts Auffälliges zu beobachten, außer dass bei der Gruppe der Fragestundenbesucher im Mittel für falsche Lösungen signifikant niedrigere Sicherheiten angegeben wurden als für richtige. Da es in den anderen beiden Gruppen kaum falsche Lösungen gab, wurde der Vergleich aber auch nur in dieser Gruppe durchgeführt.

Die Ergebnisse bezüglich der Aufgaben, die einen Vergleich von Grundvorstellungen zu Beginn und am Ende des Semesters ermöglichen sollten, sind wie folgt zusammenzufassen:

- Bei der **Krankenhaus- bzw. Grundschule-Aufgabe** gab es keine Veränderung innerhalb der Gruppen, wenn man die Anzahlen richtiger und falscher Antworten zu Beginn und am Ende des Semesters vergleicht. Die Anteile richtiger Lösungen lagen insgesamt zwischen 25% und knapp 80%. Die Gruppe der Fragestundenbesucher war zu Beginn und am Ende signifikant bzw. auffällig schwächer als die anderen beiden Gruppen. Untersucht man das individuelle Antwortverhalten, gab es in dieser Gruppe aber als einziger eine signifikante Verbesserung. Bezüglich der angegebenen Sicherheiten ließ sich bei den richtigen Antworten beobachten, dass die Fragestunden-Besucher zu Beginn des Semesters ihre Sicherheit im Mittel auffällig niedriger bezifferten als die Gruppe derer, die die „R Aufgaben" selten oder nie bearbeitetet hatten. Am Ende gaben sie bei richtiger Antwort sogar eine signifikant niedrigere Einschätzung ab als die Befragten der beiden anderen Gruppen. Ansonsten gab es keine auffälligen Unterschiede.
- Bei der ***Aufgabe zur Erkennung möglicherweise gefälschter Münzwürfe*** war zu Beginn des Semesters kein Unterschied zwischen den Gruppen zu beobachten, am Ende hatte die beiden Gruppen, die die „R Aufgaben" immer oder häufig eigenständig oder mit Hilfe der Fragestunde gelöst hatten, die 3. Reihe signifikant häufiger angekreuzt als die Gruppe derer, die die „R Aufgaben" selten oder nie bearbeitet hatten, was eine Verbesserung darstellt. Bei der 2. Reihe ließ sich zwar auch am Ende

des Semesters kein Unterschied zwischen den Gruppen feststellen, ein individueller Vergleich innerhalb der Gruppen ergab aber, dass die Reihe bei allen signifikant mehr angekreuzt wurde. Bezüglich der Sicherheiten gaben die Fragestundenbesucher am Ende des Semesters bei falscher Lösung im Mittel einen höheren Wert an als zu Beginn, der Wert war aber auch am Ende nicht unterscheidbar von den mittleren Werten der anderen Gruppen. Ansonsten gab es keine Auffälligkeiten.

Bezüglich der Begründungen im Freitextfeld konnte beobachtet werden, dass zu Beginn des Semesters überwiegend die Anzahlen der Nullen und Einsen als Kriterium herangezogen wurden (nur eine Person ging auf die Gleichmäßigkeit ihrer Verteilung als Kriterium ein), am Ende des Semesters wurden dagegen beide Kriterien – das der Anzahlen und das der Gleichmäßigkeit (Anzahl der Vorzeichenwechsel) – herangezogen, wobei besonders die Gruppe derer, die die „R Aufgaben" oft oder immer eigenständig gelöst hatten, auf die Vorzeichenwechsel einging und die Fragestundenbesucher sich hauptsächlich auf die Anzahlen der Nullen und Einsen bezogen. Bei der Gruppe derer, die die „R Aufgaben" selten oder nie bearbeitet hatten, gaben auffällig wenige überhaupt eine (richtige) Begründung an.

Die Aufgabe zu einer möglichen Erweiterung von Grundvorstellungen im Laufe des Semesters führte zu folgendem Ergebnis:

- Die ***Multiple-Choice-Test-Aufgabe*** wurde insgesamt von 55% der Befragten falsch und nur von 43% richtig beantwortet, wobei die Gruppe der Fragestundenbesucher signifikant schlechter abschnitt als die Gruppe derer, die die „R Aufgaben" immer oder häufig selbstständig gelöst hatten, nicht aber als die Gruppe derer, die die „R Aufgaben" selten oder nie bearbeitet hatten. Die angegebenen Sicherheiten bei falscher Lösung waren in der Gruppe der Fragestundenbesucher signifikant niedriger als bei richtiger Lösung, ansonsten gab es keine Auffälligkeiten.

Insgesamt bewegten sich die angegebenen Sicherheiten im Mittel sowohl bei richtigen als auch bei falschen Antworten bei allen Aufgaben im mittleren oder oberen Bereich der Skala.

Bezüglich der **Selbsteinschätzung der Befragten zu ihren Kompetenzen** im Fach Stochastik ließ sich Folgendes beobachten:

- Die Fragestundenbesucher schätzten am Anfang und am Ende ihr Wissen (E1) signifikant niedriger ein als die Gruppe derer, die die „R Aufgaben" immer oder häufig selbstständig gelöst hatten. Im Vergleich zu denen, die die „R Aufgaben" selten oder nie bearbeitet hatten, war die Einschätzung nur zu Beginn signifikant niedriger, am Ende war kein Unterschied mehr feststellbar.
- Am Ende des Semesters war die Gruppe derer, die die „R Aufgaben" selten oder nie bearbeitet hatten, auffällig stärker davon überzeugt, dass sich die Ergebnisse von Aufgaben im Fach Stochastik mit den Vermutungen deckten (E2) als es in der Gruppe derer, die die „R Aufgaben" oft oder immer eigenständig gelöst hatten, der Fall war.
- Die Fragestundenbesucher schätzten am Anfang des Semesters ihre Intuition (E3) signifikant bzw. auffällig schlechter ein als die anderen beiden Gruppen. Am Ende gab es nur noch einen signifikanten Unterschied zur Gruppe derer, die die „R Aufgaben" selten oder nie bearbeitet hatten. Zusätzlich schätzte die Gruppe derer, die die „R Aufgaben" selten oder nie bearbeitet hatten, die eigene Intuition am Ende des Semesters aber auch auffälliger besser ein als diejenigen, die die „R Aufgaben" oft oder immer selbstständig gelöst hatten.
- Die Fragestundenbesucher schätzten ihr Vorstellungsvermögen von Zufall (E4) zu Beginn und am Ende signifikant oder zumindest auffällig schlechter ein als die Befragten der anderen beiden Gruppen. Insgesamt gaben aber alle Gruppen an, sich Zufall eher gut vorstellen zu können.
- Ihre Schwierigkeiten beim Lösen von Stochastik-Aufgaben im Vergleich zu Aufgaben aus anderen Gebieten (E6) schätzten die Fragestundenbesucher am Anfang des Semesters signifikant größer ein als die Befragten der anderen beiden Gruppen, am Ende war nur noch ein auffälliger Wert im Vergleich zu denen, die die „R Aufgaben" oft oder immer selbstständig bearbeitet hatten, zu beobachten. Bei den Vergleichen innerhalb der Gruppen konnte festgestellt werden, dass in der Gruppe derer, die die „R Aufgaben" selten oder nie bearbeitet hatten, am Ende auffällig größere Schwierigkeiten bei Stochastik-Aufgaben im Vergleich zu anderen Aufgaben gesehen wurden als am Anfang des Semesters.
- Zu Beginn des Semesters gaben die Fragestundenbesucher ein signifikant niedrigeres Vertrauen in die eigenen Lösungen (E7) an als die anderen beiden Gruppen. Am Ende

des Semesters war der Wert nur noch im Vergleich zu denen, die die „R Aufgaben“ immer oder häufig selbstständig bearbeitet hatten, auffällig. Bei den individuellen Vergleichen innerhalb der Gruppen gab es bei den Fragestundenbesuchern ebenfalls eine auffällige Steigerung.

Die Ergebnisse bezogen auf die **Einstellungen zur Stochastik** sind wie folgt zusammenzufassen:

- Der Aussage „Ich mag Stochastik.“ (E5) stimmten die Fragestundenbesucher zu Beginn des Semesters signifikant weniger zu als die Gruppe derer, die die „R Aufgaben“ selten oder nie bearbeitet hatten. Am Ende war kein Unterschied mehr feststellbar. Hinzu kommt, dass im individuellen Vergleich innerhalb der Gruppen in der zuletzt genannten Gruppe ein signifikanter Rückgang der Zustimmung zu beobachten war.
- Bezüglich der Einstellung zur Nützlichkeit von Stochastik (E8) gab es keine bemerkenswerten Unterschiede zwischen den Gruppen, bei den individuellen Vergleichen innerhalb der Gruppen ließ sich aber feststellen, dass die Zustimmung bei beiden Gruppen, die die Fragestunde nicht besucht hatten, am Ende des Semesters signifikant oder zumindest auffällig gesunken war. Insgesamt wurde die Nützlichkeit des Faches allerdings von allen Gruppen zu beiden Zeitpunkten als sehr hoch eingeschätzt.
- Während es am Anfang keinen Unterschied zwischen den Gruppen gab, hielt die Gruppe derer, die die „R Aufgaben“ selten oder nie bearbeitet hatten, Stochastik am Ende des Semesters eher für ein typisch mathematisches Fach als die anderen beiden Gruppen (signifikanter bzw. zumindest auffälliger Unterschied). Im individuellen Vergleich innerhalb der Gruppen spiegelt sich dieser Effekt auch wider: Diejenigen, die die „R Aufgaben“ oft oder immer selbstständig oder mit Hilfe der Fragestunde bearbeitet hatten, hielten die Stochastik am Ende des Semesters für weniger typisch mathematisch als zu Beginn.

  Bei ihren Begründungen der Einschätzung von Stochastik als typisch oder nicht typisch mathematisch nahmen die wenigen Studierenden, die sich äußerten, v.a. auf Formalisierungen und das Satz-Beweis-Schema Bezug, genannt wurde aber auch, dass sich Zufall nicht mit Formeln erfassen lasse.

Zu den Schwierigkeiten, auf die sie bei ihrer Beschäftigung mit der Stochastik im Laufe des Semesters gestoßen waren, äußerten sich 38 Studierende, wobei auffällig wenige Kommentare von Personen aus der Gruppe derer, die die „R Aufgaben" selten oder nie bearbeitet hatten, formuliert wurden. Der Großteil der Antworten bezog sich nicht speziell auf das Fach Stochastik, sondern allgemein auf Anpassungsschwierigkeiten an die Hochschulmathematik. Als Hauptschwierigkeiten bei der Beschäftigung mit Stochastik wurden genannt, konkrete Anschauungen und Vorstellungen zu den stochastischen Konzepten, Verteilungsmerkmalen und Sätzen zu entwickeln. Auf der methodischen Ebene benannten zwei Studierende den Umgang mit R als Problem.

II.C.4.d. Evaluation des Modellprojekts „R Aufgaben" durch die Studierenden im Rahmen des Fragebogens am Ende des Semesters

Zur Evaluation des Modellprojekts sollten die Studierenden den Einsatz von Simulationen in den Vorlesungen und der „R Aufgaben" evaluieren. Dazu wurde ihre Zustimmung zu sieben Items R1 bis R7 abgefragt.

R1: „Die Simulationen in der Vorlesung waren nützlich für ein vertieftes Verständnis der Inhalte."

R2: „Die Simulationen in den Übungen waren nützlich für ein vertieftes Verständnis der Inhalte."

Beide Items beziehen sich auf die Nützlichkeit der Simulationen im Hinblick auf ein vertieftes Verständnis der Kursinhalte, wobei mit R1 die Simulationen, die in den Vorlesungen gezeigt wurden, evaluiert werden sollten und mit R2 die „R Aufgaben".

Bezüglich R1 ließ sich beobachten, dass, während die Befragten, die die „R Aufgaben" regelmäßig bearbeitet hatten (unabhängig davon, ob sie dies eigenständig oder in der Fragestunde getan hatten), die Simulationen in der Vorlesung im Mittel als eher nützlich einschätzten, die Befragten, die die „R Aufgaben" selten oder nie bearbeitet hatten, den Einsatz im Mittel signifikant kritischer bewerteten. Bezüglich R2 zeigte sich ein ähnliches Bild, wobei hier nur zwischen denen, die die „R Aufgaben" nie oder selten bearbeitet hatten und denen, die oft oder immer die Fragestunde besucht hatten, ein signifikanter Unterschied besteht. Die anderen Unterschiede waren zwar auffällig, jedoch nicht signifikant. (Vgl. Abb.II.C.38 und Tab.II.C.52.)

Vergleicht man die individuellen Einschätzungen der Nützlichkeit von Simulationen in der Vorlesung mit denen der Nützlichkeit von Simulationen in den Übungen innerhalb der Gruppen, lassen sich im Mittel keine signifikanten Unterschiede feststellen. Bei der Gruppe derjenigen, die die „R Aufgaben" oft oder immer eigenständig bearbeitet hatten, war der p-Wert des gepaarten t-Tests auffällig, möglicherweise schätzten sie die Nützlichkeit in der Vorlesung demnach etwas höher ein. (Vgl. Tab.II.C.52).

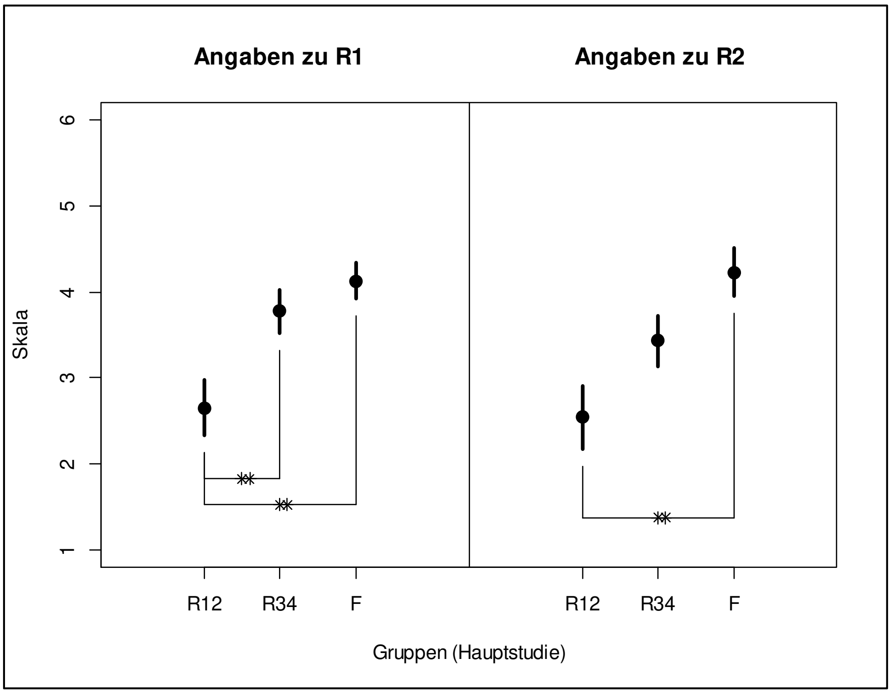


Abb.II.C.38. In der Graphik sind die Angaben der Befragten zu R1 (Evaluation der Nützlichkeit der Simulationen in der Vorlesung für ein vertieftes Verständnis) und R2 (Evaluation der Nützlichkeit der Simulationen in den Übungen für ein vertieftes Verständnis) dargestellt, wobei für jede Experimentalgruppe jeweils der Mittelwert mit Standardfehler abgebildet ist. Signifikante Unterschiede sind mit einem Stern markiert (wobei zwei Sterne auf einen p-Wert kleiner als 1% hinweisen).

Angaben zu R1			
	R12	R34	Frage
MW	2.65	3.77	4.13
Stdf.	0.33	0.25	0.21
Anz.	20	35	32

Angaben zu R2			
	R12	R34	Frage
MW	2.54	3.43	4.23
Stdf.	0.37	0.30	0.28
Anz.	13	35	31

Vergleich der Angaben zu R1:	p-Werte
R12 – R34	0.009717
R12 – Frage	0.000565
R34 – Frage	0.2857

Vergleich der Angaben zu R2:	p-Werte
R12 – R34	0.0712
R12 – Frage	0.00117
R34 – Frage	0.05652

Vergleich der Angaben zu R1 und R2 innerhalb der Gruppen:	p-Werte
R12: R1 – R2	0.613
R34: R1 – R2	0.09075
Frage: R1 – R2	0.6076

Tab.II.C.52: Mittelwerte (MW) und Standardfehler (Stdf.), gerundet auf die zweite Dezimale, sowie Anzahlen der Befragten (Anz.) für die drei Gruppen bezüglich R1 und R2, die Angaben gemacht haben. Außerdem Vergleiche der Mittelwerte zwischen den Gruppen bezüglich R1 und R2: p-Werte der t-Tests (ungepaart); sowie Vergleiche der Angaben zu R1 und R2 innerhalb der Gruppen sofern zu beiden Einschätzungen Angaben gemacht wurden: p-Werte der t-Tests (gepaart).

R3: „Ich habe den Eindruck, dass sich durch die Simulationen meine stochastische Intuition verbessert hat.“

Bezüglich der Einschätzung R3 ließen sich signifikante Unterschiede zwischen der Gruppe der Personen, die die „R Aufgaben“ nie oder selten bearbeitet hatten (also im Wesentlichen nur in der Vorlesung mit Simulationen in Kontakt gekommen waren) und denen, die die „R Aufgaben“ regelmäßig eigenständig oder mit Unterstützung in der Fragestunde bearbeitet hatten, feststellen. Insgesamt schätzten die Befragten die Nützlichkeit von Simulationen für eine Verbesserung der stochastischen Intuition im Mittel als eher gering (R34, Frage) oder sogar sehr gering (R12) ein. (Vgl. Abb.II.C.39 und Tab.II.C.53.)

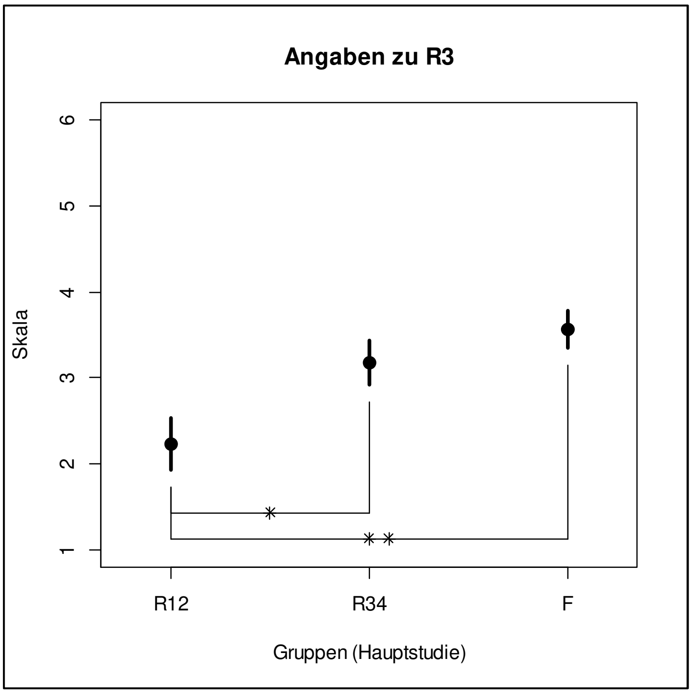


Abb.II.C.39. In der Graphik sind die Angaben der Befragten zu R3 (Evaluation des Eindrucks, dass sich durch die Simulationen die stochastische Intuition verbessert hat) dargestellt, wobei für jede Experimentalgruppe jeweils der Mittelwert mit Standardfehler abgebildet ist. Signifikante Unterschiede sind mit einem Stern markiert (wobei ein Stern einen p-Wert kleiner als 5% anzeigt und zwei Sterne auf einen p-Wert kleiner als 1% hinweisen).

	R12	R34	Frage
MW	2.23	3.17	3.56
Stdf.	0.30	0.26	0.22
Anz.	13	35	32

	p-Werte
R12 – R34	0.02462
R12 – Frage	0.001442
R34 – Frage	0.2482

Tab.II.C.53: Mittelwerte (MW) und Standardfehler (Stdf.), gerundet auf die zweite Dezimale, sowie Anzahlen der Befragten (Anz.) für die drei Gruppen bezüglich R3, die Angaben gemacht haben. Außerdem Vergleiche der Mittelwerte zwischen den Gruppen bezüglich R3: p-Werte der t-Tests (ungepaart).

R4: „Selbst Programme für Simulationen zu schreiben ist besser als nur Simulations-ergebnisse gezeigt zu bekommen.“

Dass es besser ist, selbst zu programmieren als nur mit Ergebnissen von Simulationen konfrontiert zu werden, gaben erwartungsgemäß eher Befragte an, die die „R Aufgaben“ regelmäßig bearbeitet hatten, wobei sich deren Zustimmungen zu der Aussage R4 durchschnittlich im oberen mittleren Bereich bewegte. Lediglich der Unterschied zwischen dem Mittelwert der Gruppe derer, die „R Aufgaben“ nie oder selten bearbeitet hatten, und derer, die die „R Aufgaben“ oft oder immer bearbeitet hatten, war signifikant. Der Unterschied zwischen den Befragten aus der ersten Gruppe und denen, die oft oder immer in der Fragestunde gewesen sind, war aber zumindest auffällig. (Vgl. Abb.II.C.40 und Tab.II.C.54)

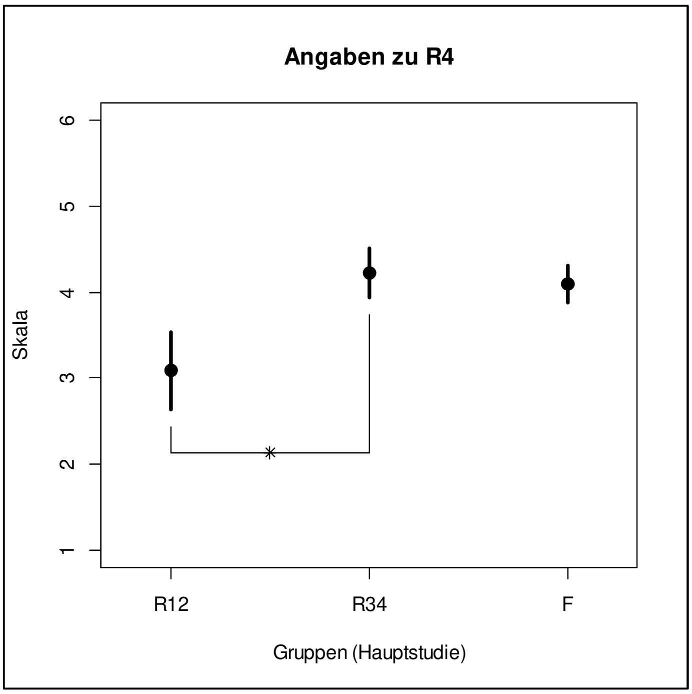


Abb.II.C.40. In der Graphik sind die Angaben der Befragten zu R4 (Evaluation des Eindrucks, dass selbst zu programmieren besser ist als nur Simulationsergebnisse präsentiert zu bekommen) dargestellt, wobei für jede Experimentalgruppe jeweils der Mittelwert mit Standardfehler abgebildet ist. Signifikante Unterschiede sind mit einem Stern markiert.

	R12	R34	Frage
MW	3.08	4.22	4.10
Stdf.	0.45	0.28	0.22
Anz.	12	36	31

	p-Werte
R12 – R34	0.04519
R12 – Frage	0.06009
R34 – Frage	0.7283

Tab.II.C.54: Mittelwerte (MW) und Standardfehler (Stdf.), gerundet auf die zweite Dezimale, sowie Anzahlen der Befragten (Anz.) für die drei Gruppen bezüglich R4, die Angaben gemacht haben. Außerdem Vergleiche der Mittelwerte zwischen den Gruppen bezüglich R4: p-Werte der t-Tests (ungepaart).

R5: „Ohne die R-Fragestunde hätte ich die Aufgaben nicht lösen können."

Erwartungsgemäß war die Zustimmung zu der Aussage, dass die „R Aufgaben" ohne die R-Fragestunde nicht lösbar gewesen wären, bei der Gruppe, die die Fragestunde regelmäßig besucht hatte, am höchsten und bei der Gruppe, die die „R Aufgaben" zwar regelmäßig bearbeitet, die Fragestunde aber nicht besucht hatte, am niedrigsten. Nichtsdestotrotz gab es auch in der zweiten Gruppe wenige Personen, die der Aussage eher zugestimmt haben, was auf einen Austausch mit Fragestunden-Besuchern hindeutet. Aber auch die Gruppe derer, die die „R Aufgaben" nie oder selten bearbeitet hatten, zeigte im Mittel eine hohe Zustimmung. Beide Unterschiede zu der zweiten Gruppe waren signifikant. (Vgl. Abb.II.C.41 und Tab.II.C.55)

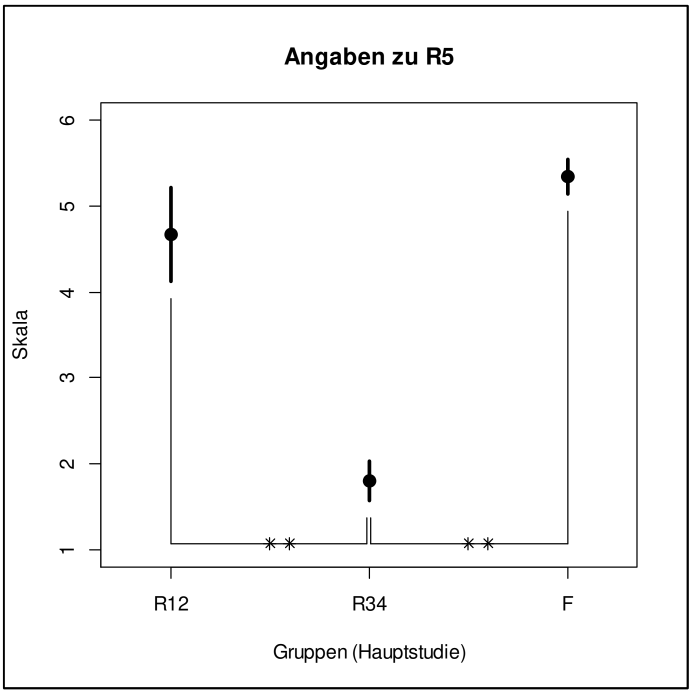


Abb.II.C.41. In der Graphik sind die Angaben der Befragten zu R5 (Evaluation der Notwendigkeit des Besuchs der Fragestunde, um die „R Aufgaben" lösen zu können) dargestellt, wobei für jede Experimentalgruppe jeweils der Mittelwert mit Standardfehler abgebildet ist. Signifikante Unterschiede sind mit einem Stern markiert (wobei zwei Sterne auf einen p-Wert kleiner als 1% hinweisen).

	R12	R34	Frage
MW	4.67	1.8	5.34
Stdf.	0.54	0.23	0.20
Anz.	12	35	32

	p-Werte
R12 – R34	0.0001956
R12 – Frage	0.2598
R34 – Frage	< 2.2e-16

Tab.II.C.55: Mittelwerte (MW) und Standardfehler (Stdf.), gerundet auf die zweite Dezimale, sowie Anzahlen der Befragten (Anz.) für die drei Gruppen bezüglich R5, die Angaben gemacht haben. Außerdem Vergleiche der Mittelwerte zwischen den Gruppen bezüglich R5: p-Werte der t-Tests (ungepaart).

R6: „Die R-Aufgaben waren interessant."

Die Interessantheit der „R Aufgaben" wurde von allen Gruppen unterschiedlich bewertet, wobei die Gruppe derer, die die „R Aufgaben" nie oder selten bearbeitet hatten, die Aufgaben durchschnittlich eher weniger interessant fand. Diejenigen, die die „R Aufgaben" oft oder immer eigenständig bearbeitet hatten, fanden sie im Mittel auch interessanter. Die höchste Zustimmung zu der Aussage R6 war bei der Gruppe derer zu finden, die regelmäßig die Fragestunde besucht hatten. Die Werte der zuletzt genannten Gruppe lagen durchschnittlich im oberen Bereich der Zustimmungsskala, die der zweiten Gruppe im oberen mittleren und die der ersten Gruppe zwischen dem mittleren und unteren Bereich der Skala. (Vgl. Abb.II.C.42 und Tab.II.C.56)

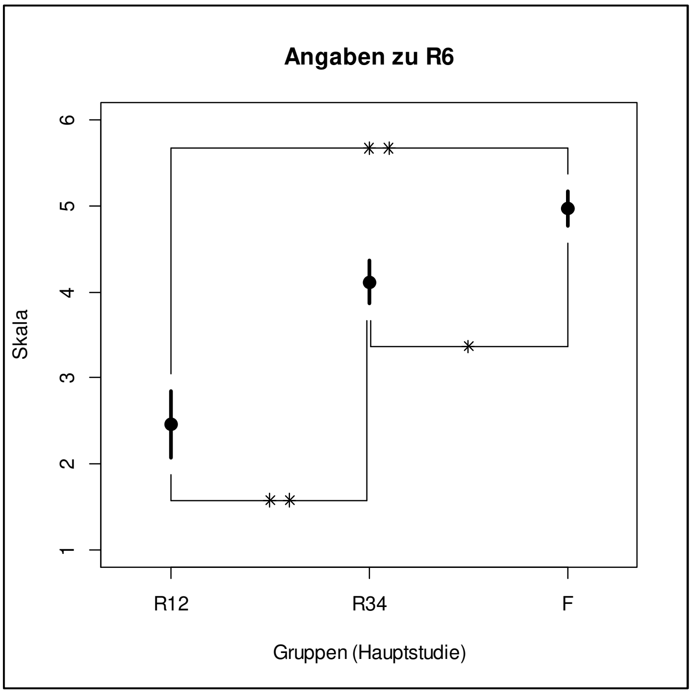


Abb.II.C.42. In der Graphik sind die Angaben der Befragten zu R6 (Evaluation der Interessantheit der „R Aufgaben") dargestellt, wobei für jede Experimentalgruppe jeweils der Mittelwert mit Standardfehler abgebildet ist. Signifikante Unterschiede sind mit einem Stern markiert (wobei ein Stern einen p-Wert kleiner als 5% anzeigt und zwei Sterne auf einen p-Wert kleiner als 1% hinweisen).

	R12	R34	Frage
MW	2.46	4.11	4.97
Stdf.	0.39	0.25	0.20
Anz.	13	35	31

	p-Werte
R12 – R34	0.001536
R12 – Frage	1.544e-05
R34 – Frage	0.01016

Tab.II.C.56: Mittelwerte (MW) und Standardfehler (Stdf.), gerundet auf die zweite Dezimale, sowie Anzahlen der Befragten (Anz.) für die drei Gruppen bezüglich R6, die Angaben gemacht haben. Außerdem Vergleiche der Mittelwerte zwischen den Gruppen bezüglich R6: p-Werte der t-Tests (ungepaart).

R7: „Ich würde empfehlen, auch weiterhin Simulationen in den Übungen zur Elementaren Stochastik zu verwenden.“

Der Grad an Zustimmung zu der Empfehlung, auch weiterhin Simulationen in den Übungen zur Elementaren Stochastik einzusetzen, unterschied sich erwartungsgemäß signifikant zwischen denen, die die „R Aufgaben“ regelmäßig bearbeitet hatten und denen, die sie nie oder nur selten bearbeitet hatten. Die Stärke der Zustimmung lag bei denen, die die „R Aufgaben“ oft oder immer entweder selbstständig oder mit Unterstützung in der Fragestunde bearbeitet hatten, durchschnittlich im oberen mittleren Bereich der Skala, sie stimmten also im Mittel „eher zu“. Bei denen, die die „R Aufgaben“ nie oder selten bearbeitet hatten, lag der Mittelwert dagegen zwischen dem mittleren und dem unteren Bereich der Zustimmungsskala, d.h. sie empfahlen den weiteren Einsatz „eher nicht“. (Vgl. Abb.II.C.43 und Tab.II.C.57)

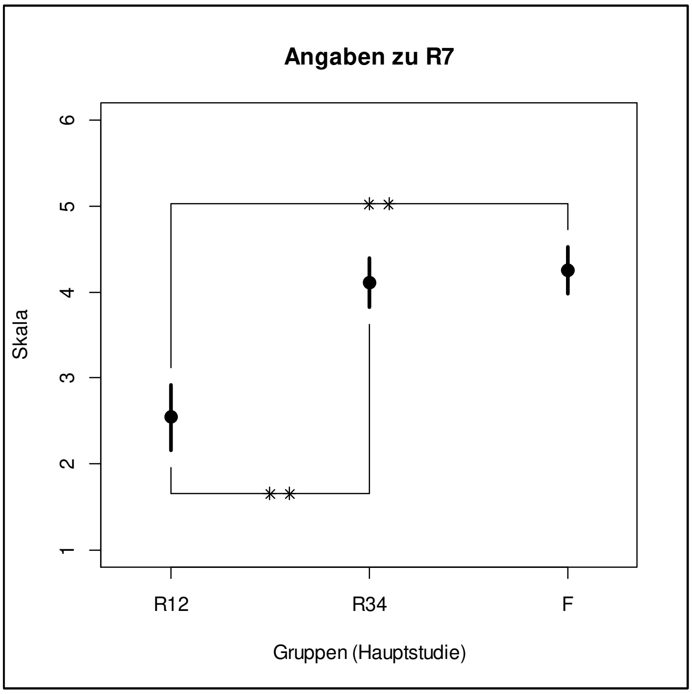


Abb.II.C.43. In der Graphik sind die Angaben der Befragten zu R7 (Empfehlung, auch weiterhin Simulationen in den Übungen zur Elementaren Stochastik einzusetzen) dargestellt, wobei für jede Experimentalgruppe jeweils der Mittelwert mit Standardfehler abgebildet ist. Signifikante Unterschiede sind mit einem Stern markiert (wobei zwei Sterne auf einen p-Wert kleiner als 1% hinweisen).

	R12	R34	Frage
MW	2.54	4.11	4.25
Stdf.	0.39	0.29	0.27
Anz.	13	36	32

	p-Werte
R12 – R34	0.003061
R12 – Frage	0.001326
R34 – Frage	0.7282

Tab.II.C.57: Mittelwerte (MW) und Standardfehler (Stdf.), gerundet auf die zweite Dezimale, sowie Anzahlen der Befragten (Anz.) für die drei Gruppen bezüglich R7, die Angaben gemacht haben. Außerdem Vergleiche der Mittelwerte zwischen den Gruppen bezüglich R7: p-Werte der t-Tests (ungepaart).

Zusätzlich zu den Einschätzungen auf den Skalen hatten die Studierenden noch die Möglichkeit, „**sonstige Kommentare zu den R-Aufgaben/ Simulationen**“ aufzuschreiben. Insgesamt 22 Befragte machten davon Gebrauch (R12: 6, R34: 9, F: 7). Die Äußerungen werden im Folgenden in komprimierter Form wiedergegeben, wobei Mehrfachnennungen vorkamen.

Am häufigsten kommentierten die Studierenden den organisatorischen Rahmen. Zur Fragestunde äußerten sich insgesamt 8 Personen, dabei wurde sechsmal die Notwendigkeit einer Fragestunde, die auch als hilfreich empfunden wurde, hervorgehoben (R12: 1, R34: 1, F: 4), zweimal jedoch ein mangelnder Lerneffekt durch die Fragestunde kritisiert (R12, R34). Drei Befragte betonten, dass es für die Planung wichtig sei, die Termine für die Fragestunde bereits im Vorlesungsverzeichnis bekannt zu geben (R34: 1, F: 2). Die Templates wurden von zwei Studierenden als hilfreich und gelungen empfunden (R34: 2), zwei andere Befragte hätten sich mehr technische Unterstützung gewünscht (R34, F) und ein Proband schlug vor, die erste Vorlesung für einen gemeinsamen Einstieg in R zu nutzen (R12). Sechs Befragte formulierten, dass die Aufgaben ohne vorherige Programmierkenntnisse eigenständig kaum oder gar nicht lösbar gewesen seien (R12: 3, R34: 1, F:2). Zwei Studierende empfanden den zusätzlichen Aufwand als sehr oder sogar zu hoch (F: 2). Zwei L3-Studierende stellten den Nutzen für ihr Berufsziel in Frage (R34, F), wohingegen ein anderer einen Nutzen auch für die Schule feststellte (F). Ein Bachelor-Studierender wünschte sich für seinen Studiengang lieber einen „abstrakten, axiomatischen Zugang“ statt der „R Aufgaben“ (R34). Ein Befragter empfand R als „unschöne“ Programmiersprache (R12). Zum Inhalt der „R Aufgaben“ bzw. der Simulationen und möglichen Lerneffekten gab es insgesamt sieben Äußerungen, die überwiegend positiv waren: Sie wurden als sinnvoll (F), interessant (F), hilfreich für eine Veranschaulichung von Zusammenhängen (R12) und praxisbezogen (R12, R34) empfunden. Zudem ermöglichten sie auch ohne Vorkenntnisse einen Einblick in R (R34), wobei durch sie kein eigenständiges Programmieren erlernbar sei (R34).

II.C.4.e. Wer profitiert genau von Simulationen in der Vorlesung bzw. von der Bearbeitung der „R Aufgaben“?

Um genauer zu untersuchen, wer von Simulationen in der Vorlesung bzw. der Bearbeitung der „R Aufgaben“ profitieren konnte, wurden zur besseren Vergleichbarkeit der Resultate in den verschiedenen Gruppen ein „Leistungsindex“, ein „Vertrauensindex“ und ein „Einstellungsindex“ erstellt, die jeweils verschiedene Items der Fragebögen zusammenfassen. Im *„Leistungsindex“* spiegeln sich die Lösungen der Aufgaben wider, wobei eine richtige Antwort mit einem Punkt bewertet wurde und eine falsche oder keine Antwort mit 0 Punkten. Zu Beginn des Semesters konnten maximal 5 Punkte erreicht werden, am Ende maximal 3 Punkte.[146] In den *„Vertrauensindex“*, einer Art Likert-Skala, gingen die Einschätzungen E1, E2, E3, E4, E6 und E7 ein, wobei aufgrund ihrer unterschiedlichen Formulierungen die angegebenen Skalenwerte zu E1 und E6 mit positivem Vorzeichen und die restlichen mit negativem Vorzeichen addiert wurden. Dadurch ergibt sich die Interpretation, dass das Vertrauen in die eigenen stochastischen Fertigkeiten umso höher ist, je höher auch der Wert des Indexes ist. Der kleinste Wert, der im „Vertrauensindex“ erreicht werden konnte, war –22, der größte 8. In den *„Einstellungsindex“* gingen lediglich zwei Einschätzungen ein, nämlich E5 und E8, beide mit positivem Vorzeichen. Dieser Index gibt an, wie positiv die Einstellung dem Fach gegenüber ist, und wiederum gilt: Je höher der Wert, desto positiver die Einstellung. Der kleinste Wert, der in diesem Index erreicht werden konnte, ist 2, der höchste 12.

Für Studierende, die zu Beginn oder am Ende des Semesters nicht zu allen Items, die in den jeweiligen Index eingingen, eine Angabe gemacht hatten, wurde der entsprechende Index nicht berechnet.

Während eine eher starke positive Korrelation zwischen dem *Vertrauensindex* und dem *Einstellungsindex* vorlag, waren die Korrelationen zwischen dem *Leistungsindex* und dem *Vertrauens-* bzw. *Einstellungsindex* eher schwach (vgl. Tab.II.C.58). Dabei ist allerdings zu beachten, dass die Aussagekraft dieser Korrelationen aufgrund der wenigen möglichen Ausprägungen im *Leistungsindex* nicht überschätzt werden sollte.

[146] Bei dem Leistungsindex wurde bei der *Aufgabe zur Erkennung möglicherweise gefälschter Münzwürfe* aufgrund der Eindeutigkeit ihrer Zweifelhaftigkeit nur die letzte Reihe („0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0“) in die Wertung einbezogen.

Leistung – Vertrauen	0.26 (vorher), 0.21 (nachher)
Leistung – Einstellung	0.20 (vorher), 0.08 (nachher)
Vertrauen – Einstellung	0.64 (vorher), 0.71 (nachher)

Tab.II.C.58: Spearman-Korrelationen zwischen den berechneten Indizes, gerundet auf die 2. Dezimale.

Im Folgenden werden sowohl die Ausgangslage als auch die Ergebnisse am Ende des Semesters zwischen der Kontrollgruppe und den drei Experimentalgruppen im Hinblick auf die Indizes verglichen.

II.C.4.e.1. Vergleich der Ausgangslage der Befragten in der Vorstudie (SoSe 2011) und der Hauptstudie (SoSe 2012)

Die Befragten aus der Vorstudie unterschieden sich weder bezüglich der Geschlechterverteilung, noch bezüglich der drei Indizes zu Beginn des Semesters signifikant von denen aus der Hauptstudie. Ebenfalls gab es keine auffälligen Unterschiede bezüglich der Anteile von Personen, die einen Mathematik-Leistungskurs oder im laufenden Semester zusätzlich das Proseminar zur Stochastik belegt hatten. Wie zu erwarten, hatten die Befragten in der Hauptstudie aber im Mittel ein Jahr später Abitur gemacht als die Befragten aus der Vorstudie. (Vgl. Tab. II.C.59).

	Kontrollgruppe (SoSe 2011)	Experimentalgruppe (SoSe 2012)	
Geschlechterverteilung	m: 16 – w: 28	m: 46 – w: 45	Fisher-Test, p=0.1425
„Leistungsindex“	Mittelwert: 3.41	Mittelwert: 3.56	t-Test, p= 0.4679
„Vertrauensindex“	Mittelwert: -6.28	Mittelwert: -5.94	t-Test, p= 0.7415
„Einstellungsindex“	Mittelwert: 8.49	Mittelwert: 8.39	t-Test, p= 0.8025
Mathematik-LK	belegt: 68%	belegt: 73%	
Proseminar	belegt: 11%	belegt: 13%	

Tab.II.C.59: Vergleich der Kontroll- und Experimentalgruppe zu Beginn des Semesters (Ausgangslagen).

Die Kontrollgruppe konnte also aufgrund der praktisch nicht unterscheidbaren Ausgangslage sinnvoll mit der Experimentalgruppe verglichen werden, um Aussagen über Bedingungen zu treffen, die zu einem Gewinn durch die Auseinandersetzung mit Simulationen in der Vorlesung bzw. der Bearbeitung der „R Aufgaben“ führen könnten.

II.C.4.e.2. Profit bezüglich der Lösung der Aufgaben

Zur näheren Analyse, wer bezüglich des Leistungsindexes von Simulationen profitieren konnte, wurde innerhalb der Gruppen „Kontrolle“, „R12“, „R34“ und „Frage“ noch zwischen anfangs schwachen und anfangs starken Studierenden unterschieden. Der Mittelwert lag zu Beginn des Semesters bei etwa 3.51, wobei die Werte aller Probanden aus beiden Jahrgängen zugrunde gelegt wurden. Daher wurden alle mit einem Index von weniger als 4 zu den „Schwachen“ und alle mit einem Index von 4 oder 5 zu den „Starken“ gezählt. In der Kontrollgruppe gab es dann 21 „Schwache“ und 23 „Starke“, in der Experimentalgruppe 44 „Schwache“ und 47 „Starke“.

In der Experimentalgruppe besuchten von den „leistungsschwachen“ Studierenden die meisten regelmäßig die Fragestunde (n=27), 12 lösten die „R Aufgaben“ oft oder immer selbstständig und nur 5 beschäftigten sich selten oder nie mit den „R Aufgaben“. Die „leistungsstarken“ Studierenden lösten die „R Aufgaben“ dagegen am häufigsten oft oder immer selbstständig (n=24), 18 von ihnen bearbeiteten sie gar nicht oder selten und nur 5 besuchten die Fragestunde regelmäßig.

Zu Beginn des Semesters ließen sich bezüglich der Mittelwerte der Leistungsindizes bei den „leistungsschwachen“ Studierenden keine Unterschiede feststellen. Es kann also davon ausgegangen werden, dass zu Beginn des Semesters alle vier Gruppen vergleichbar schwach waren (vgl. Tab. II.C.60.).

	Kontrolle	R12	R34	Frage
MW	2.43	2.8	2.42	2.63
Stdf.	0.04	0.09	0.08	0.03
Anz.	21	5	12	27

	p-Werte
K – R12	0.1799
K – R34	0.9717
K – Frage	0.3436

	p-Werte
R12 – R34	0.2914
R12 – Frage	0.4976
R34 – Frage	0.5108

Tab.II.C.60: Leistungsschwache Studierende: Mittelwerte (MW) und Standardfehler (Stdf.), gerundet auf die zweite Dezimale, sowie Anzahlen der Befragten (Anz.) für die vier Gruppen bezüglich des Leistungsindexes zu Beginn des Semesters. Außerdem Vergleiche der Mittelwerte zwischen den Gruppen: p-Werte der t-Tests (ungepaart).

Die Leistungsindizes der „leistungsschwachen“ Studierenden am Ende des Semesters sind in der Tabelle Tab.II.C.61 aufgelistet. Da am Ende des Semesters ja nur drei Aufgaben gestellt wurden, bewegten sich alle Mittelwerte auf einem niedrigeren Niveau. Das heißt aber nicht, dass die Studierenden am Ende des Semesters insgesamt schlechter waren, sondern der Referenzrahmen ist ein anderer als zu Beginn des Semesters. Innerhalb der Experimentalgruppe gab es keine auffälligen oder signifikanten Unterschiede, wohl aber zwischen der Kontrollgruppe und denen, die die „R Aufgaben“ entweder regelmäßig selbstständig oder in der Fragestunde bearbeitet hatten.

	Kontrolle	R12	R34	Frage
MW	1.19	1.60	2.00	1.63
Stdf.	0.04	0.18	0.06	0.03
Anz.	21	5	12	27

	p-Werte
K – R12	0.3874
K – R34	0.007359
K – Frage	0.0544

	p-Werte
R12 – R34	0.4094
R12 – Frage	0.9467
R34 – Frage	0.1556

Tab.II.C.61: Leistungsschwache Studierende: Mittelwerte (MW) und Standardfehler (Stdf.), gerundet auf die zweite Dezimale, sowie Anzahlen der Befragten (Anz.) für die vier Gruppen bezüglich des Leistungsindexes am Ende des Semesters. Außerdem Vergleiche der Mittelwerte zwischen den Gruppen: p-Werte der t-Tests (ungepaart).

Bei den „leistungsstarken“ Studierenden war das Bild ein etwas anderes: Die Gruppe der „Leistungsstarken“, die die Fragestunde regelmäßig besucht hatten, war zu Beginn des Semesters signifikant schwächer als die anderen „leistungsstarken“ Gruppen (vgl. Tab.II.C.62). Ansonsten ließen sich aber keine auffälligen Unterschiede finden. Die Gruppe der regelmäßigen Fragestunden-Besucher unter den „Leistungsstarken“ war aber auch sehr klein (n=5) und ohne jegliche Variabilität, so dass ihre Ergebnisse ohnehin nicht aussagekräftig waren.

	Kontrolle	R12	R34	Frage
MW	4.30	4.5	4.54	4
Stdf.	0.02	0.03	0.02	0
Anz.	23	18	24	5

	p-Werte
K – R12	0.218
K – R34	0.1037
K – Frage	0.005196

	p-Werte
R12 – R34	0.7956
R12 – Frage	0.0007103
R34 – Frage	2.75e-05

Tab.II.C.62: Leistungsstarke Studierende: Mittelwerte (MW) und Standardfehler (Stdf.), gerundet auf die zweite Dezimale, sowie Anzahlen der Befragten (Anz.) für die vier Gruppen bezüglich des Leistungsindexes zu Beginn des Semesters. Außerdem Vergleiche der Mittelwerte zwischen den Gruppen: p-Werte der t-Tests (ungepaart).

Die Ergebnisse der „leistungsstarken“ Studierenden am Ende des Semesters finden sich in der Tabelle Tab.II.C.63. Bei dieser Gruppe ließ sich nur ein signifikanter Unterschied zwischen der Kontrollgruppe und jenen feststellen, die die „R Aufgaben“ oft oder immer selbstständig bearbeitet hatten. Innerhalb der Experimentalgruppen gab es keine auffälligen Unterschiede, auch nicht mehr zu den regelmäßigen Fragestunden-Besuchern.

	Kontrolle	R12	R34	Frage
MW	1.78	2.00	2.33	1.80
Stdf.	0.04	0.05	0.03	0.17
Anz.	23	18	24	5

	p-Werte
K – R12	0.4608
K – R34	0.03448
K – Frage	0.9685

	p-Werte
R12 – R34	0.2163
R12 – Frage	0.657
R34 – Frage	0.2405

Tab.II.C.63: Starke Studierende: Mittelwerte (MW) und Standardfehler (Stdf.), gerundet auf die zweite Dezimale, sowie Anzahlen der Befragten (Anz.) für die vier Gruppen bezüglich des Leistungsindexes am Ende des Semesters. Außerdem Vergleiche der Mittelwerte zwischen den Gruppen: p-Werte der t-Tests (ungepaart).

Für einen aussagekräftigen individuellen Vergleich ist die Anzahl der Aufgaben, die sowohl zu Beginn als auch am Ende des Semesters in analoger Weise gestellt wurden, zu klein (es waren lediglich zwei Aufgaben, nämlich die *Krankenhaus-/ Grundschule-Aufgabe* und die *Aufgabe zur Erkennung möglicherweise gefälschter Münzwürfe*). Um mögliche Hinweise auf eine Veränderung zu sichtbar zu machen, wurden dennoch innerhalb der oben beschriebenen Gruppen der Leistungsschwachen und Leistungsstarken die Antworten zu diesen beiden Aufgaben auf individueller Ebene verglichen und dazu ein weiterer Leistungsindex erstellt. Die Ergebnisse finden sich in den folgenden Tabellen Tab.II.C.64 und Tab.II.C.65. Bei den Leistungsschwachen gibt es sowohl bei denen, die die „R Aufgaben“ oft oder immer eigenständig gelöst haben, als auch bei den Fragestundenbesuchern eine signifikante Verbesserung. Die Gruppe derer, die die „R Aufgaben“ selten oder nie bearbeitet hat, zeigt einen auffälligen p-Wert, der jedoch aufgrund der sehr kleinen Gruppengröße kaum aussagekräftig ist. Bei den Leistungsstarken lässt sich dagegen keine Entwicklung beobachten.

Kontrolle	vorher	nachher	R12	Vorher	nachher	R34	vorher	nachher	Frage	vorher	nachher
MW	0.81	0.86	MW	0.6	1.2	MW	0.58	1.5	MW	0.96	1.37
Stdf.	0.51	0.57	Stdf.	0.55	0.45	Stdf.	0.67	0.52	Stdf.	0.59	0.49
Anz.	21		Anz.	5		Anz.	12		Anz.	27	
t-Test	p= 0.7152		t-Test	p= 0.07048		t-Test	p= 0.00474		t-Test	p= 0.008794	

Tab.II.C.64: Leistungsschwache Studierende: Mittelwerte (MW) und Standardfehler (Stdf.) eines weiteren Leistungsindexes nur für die *Krankenhaus-/Grundschule-Aufgabe* und die *Aufgabe zur Erkennung möglicherweise gefälschter Münzwürfe* zu Beginn und am Ende des Semesters, gerundet auf die zweite Dezimale, sowie Anzahlen der Befragten (Anz.) für die vier Gruppen. Außerdem: Individuelle Vergleiche innerhalb der Gruppen: p-Werte der t-Tests (gepaart).

Kontrolle	vorher	nachher	R12	Vorher	nachher	R34	vorher	nachher	Frage	vorher	nachher
MW	1.43	1.43	MW	1.56	1.61	MW	1.71	1.71	MW	1.2	1.4
Stdf.	0.51	0.66	Stdf.	0.51	0.61	Stdf.	0.46	0.46	Stdf.	0.45	0.55
Anz.	23		Anz.	18		Anz.	24		Anz.	5	
t-Test	p=1		t-Test	p= 0.6676		t-Test	p=1		t-Test	p=0.3739	

Tab.II.C.65: Leistungsstarke Studierende: Mittelwerte (MW) und Standardfehler (Stdf.) eines weiteren Leistungsindexes nur für die *Krankenhaus-/Grundschule-Aufgabe* und der *Aufgabe zur Erkennung möglicherweise gefälschter Münzwürfe* zu Beginn und am Ende des Semesters, gerundet auf die zweite Dezimale, sowie Anzahlen der Befragten (Anz.) für die vier Gruppen. Außerdem: Individuelle Vergleiche innerhalb der Gruppen: p-Werte der t-Tests (gepaart).

II.C.4.e.3. Profit bezüglich des Vertrauens in die eigenen stochastischen Fertigkeiten und bezüglich der Einstellungen zur Stochastik

Die Aufteilung in bezüglich des „Vertrauensindex" anfangs „Starke" und „Schwache" wurde ebenfalls anhand des Mittelwerts aller Probanden vorgenommen. Er lag bei etwa -6.1. Alle Studierenden mit einem „Vertrauensindex" von maximal -6 zu Beginn des Semesters wurden als „vertrauensschwach", die anderen als „vertrauensstark" klassifiziert.

Da das Vertrauen in die eigenen stochastischen Fähigkeiten zu Beginn und am Ende des Semesters mit den identischen Items abgefragt wurde, können die Ergebnisse bezüglich des „Vertrauensindex" auf individuelle Veränderungen untersucht werden. Die Ergebnisse im Hinblick auf den „Vertrauensindex" sind für die anfangs „Schwachen" in der Tab.II.C.66 abgebildet. Auffällig ist, dass in den Gruppen derer, die die „R Aufgaben" häufig oder immer selbstständig oder in der Fragestunde bearbeitet hatten, das Vertrauen in die eigenen stochastischen Fähigkeiten signifikant stieg, während es bei den Gruppen derer, die keine „R Aufgaben" bearbeitet hatten, keine auffällige Veränderung gab.

Kontrolle	vorher	nachher	R12	Vorher	nachher	R34	vorher	nachher	Frage	vorher	nachher
MW	-9.62	-9	MW	-8.43	-7.86	MW	-10	-7.5	MW	-11.27	-9.32
Stdf.	0.17	0.28	Stdf.	0.39	0.73	Stdf.	0.24	0.27	Stdf.	0.12	0.18
Anz.	21		Anz.	7		Anz.	14		Anz.	22	
t-Test	p= 0.665		t-Test	p= 0.6112		t-Test	p= 0.01986		t-Test	p= 0.03257	

Tab.II.C.66: Vertrauensschwache Studierende: Mittelwerte (MW) und Standardfehler (Stdf.) des Vertrauensindexes zu Beginn und am Ende des Semesters, gerundet auf die zweite Dezimale, sowie Anzahlen der Befragten (Anz.) für die vier Gruppen. Außerdem: Individuelle Vergleiche innerhalb der Gruppen: p-Werte der t-Tests (gepaart).

Innerhalb der Gruppe derer, die zu Beginn des Semesters ein vergleichsweise großes Vertrauen in ihre stochastischen Fähigkeiten hatten, ließ sich beobachten, dass dieses bei der Gruppe derer, die die „R Aufgaben“ entweder gar nicht oder aber selbstständig bearbeitet hatten, signifikant sank. Am Ende des Semesters hatten alle Gruppen eine vergleichbare Einschätzung, die Werte glichen sich also im Laufe des Semesters an. (Vgl. Tab.II.C.67.)

Kontrolle	vorher	nachher	R12	Vorher	nachher	R34	vorher	nachher	Frage	vorher	nachher
MW	-2.58	-3.63	MW	-1.57	-3.71	MW	-0.53	-2.58	MW	-3.11	-3
Stdf.	0.12	0.20	Stdf.	0.17	0.19	Stdf.	0.16	0.24	Stdf.	0.30	0.44
Anz.	19		Anz.	14		Anz.	19		Anz.	9	
t-Test	p= 0.2008		t-Test	p=0.008666		t-Test	p= 0.02983		t-Test	p= 0.9068	

Tab.II.C.67: Vertrauensstarke Studierende: Mittelwerte (MW) und Standardfehler (Stdf.) des Vertrauensindexes zu Beginn und am Ende des Semesters, gerundet auf die zweite Dezimale, sowie Anzahlen der Befragten (Anz.) für die vier Gruppen. Außerdem: Individuelle Vergleiche innerhalb der Gruppen: p-Werte der t-Tests (gepaart).

Inwiefern ein Ansteigen des Vertrauens im Laufe des Semesters gerechtfertigt ist, lässt sich im Abgleich mit den gezeigten Leistungen beurteilen. Da ein individueller Vergleich der Leistungen zu Beginn und am Ende des Semesters nur bezüglich zwei Aufgaben möglich ist (vgl. Abschnitt II.C.4.e.2), können die folgenden Untersuchungen lediglich Hinweise liefern. Sowohl für die Vertrauensstarken als auch die Vertrauensschwachen wurden die Leistungen bezüglich der *Krankenhaus-/ Grundschule-Aufgabe* und der *Aufgabe zur Erkennung möglicherweise gefälschter Münzwürfe* zu Beginn und am Ende des Semesters auf individueller Ebene verglichen. Dabei konnte nur bei den vertrauensschwachen Fragestundenbesuchern eine signifikante Verbesserung der individuellen Leistungen beobachtet werden (vgl. Tab.II.C.68). Bei den vertrauensstarken Studierenden, die die „R

Aufgaben“ oft oder immer eigenständig bearbeitet haben, gab es zudem eine auffällige Verbesserung (vgl. Tab.II.C.69).

Kontrolle	vorher	nachher	**R12**	Vorher	nachher	**R34**	vorher	nachher	**Frage**	vorher	nachher
MW	1.05	1.24	MW	1.14	1.43	MW	1.21	1.57	MW	1.05	1.45
Stdf.	0.59	0.62	Stdf.	0.69	0.53	Stdf.	0.89	0.51	Stdf.	0.58	0.51
Anz.	21		Anz.	7		Anz.	14		Anz.	22	
t-Test	p= 0.1623		t-Test	p= 0.1723		t-Test	p= 0.2079		t-Test	p= 0.01622	

Tab.II.C.68: Vertrauensschwache Studierende: Mittelwerte (MW) und Standardfehler (Stdf.) des Leistungsindexes, der nur die *Krankenhaus-/ Grundschule-Aufgabe* und die *Aufgabe zur Erkennung möglicherweise gefälschter Münzwürfe* einbezieht, zu Beginn und am Ende des Semesters, gerundet auf die zweite Dezimale, sowie Anzahlen der Befragten (Anz.) für die vier Gruppen. Außerdem: Individuelle Vergleiche innerhalb der Gruppen: p-Werte der t-Tests (gepaart).

Kontrolle	vorher	nachher	**R12**	Vorher	nachher	**R34**	vorher	nachher	**Frage**	vorher	nachher
MW	1.21	1.56	MW	1.36	1.57	MW	1.42	1.74	MW	1	1.22
Stdf.	0.63	0.69	Stdf.	0.63	0.65	Stdf.	0.69	0.45	Stdf.	0.5	0.44
Anz.	19		Anz.	14		Anz.	19		Anz.	9	
t-Test	p= 0.7162		t-Test	p= 0.1894		t-Test	p= 0.08276		t-Test	p= 0.3466	

Tab.II.C.69: Vertrauensstarke Studierende: Mittelwerte (MW) und Standardfehler (Stdf.) des Leistungsindexes, der nur die die *Krankenhaus-/ Grundschule-Aufgabe* und die *Aufgabe zur Erkennung möglicherweise gefälschter Münzwürfe* einbezieht, zu Beginn und am Ende des Semesters, gerundet auf die zweite Dezimale, sowie Anzahlen der Befragten (Anz.) für die vier Gruppen. Außerdem: Individuelle Vergleiche innerhalb der Gruppen: p-Werte der t-Tests (gepaart).

Der „Einstellungsindex“ gibt an, wie positiv die Studierenden dem Fach Stochastik gegenüber stehen. Der Mittelwert lag zu Beginn des Semesters bei etwa 8.42, wobei über alle Studierenden gemittelt wurde. Daher wurden alle Probanden mit einem „Einstellungsindex von maximal 8 zu den „Schwachen“ und alle anderen zu den „Starken“ gezählt.

Da die Einstellung zum Fach Stochastik ebenso wie das Vertrauen in die eigenen stochastischen Fähigkeiten zu Beginn und am Ende des Semesters mit den identischen Items abgefragt wurde, konnten auch hier die Ergebnisse auf individuelle Veränderungen untersucht werden. Die Ergebnisse im Hinblick auf den „Einstellungsindex“ sind für die anfangs „Schwachen“ in der Tab.II.C.70 abgebildet. In keiner Gruppe gab es eine auffällige Veränderung im Laufe des Semesters.

Kontrolle	vorher	nachher	**R12**	Vorher	nachher	**R34**	vorher	nachher	**Frage**	vorher	nachher
MW	6.79	6.47	MW	6.89	6.44	MW	6.83	6.72	MW	6.39	6.83
Stdf.	0.08	0.10	Stdf.	0.19	0.23	Stdf.	0.09	0.11	Stdf.	0.09	0.12
Anz.	19		Anz.	9		Anz.	18		Anz.	18	
t-Test	p= 0.3306		t-Test	p= 0.426		t-Test	p= 0.8012		t-Test	p= 0.3234	

Tab.II.C.70: Einstellungsschwache Studierende: Mittelwerte (MW) und Standardfehler (Stdf.) des Vertrauensindexes zu Beginn und am Ende des Semesters, gerundet auf die zweite Dezimale, sowie Anzahlen der Befragten (Anz.) für die vier Gruppen. Außerdem: Individuelle Vergleiche innerhalb der Gruppen: p-Werte der t-Tests (gepaart).

Bei den Studierenden, die anfangs eine vergleichsweise positive Einstellung dem Fach Stochastik gegenüber hatten, verschlechterten sich die Werte in den Gruppen derer, die die „R Aufgaben" nicht oder aber oft bzw. immer selbstständig bearbeitet hatten. In der Kontrollgruppe und in der Gruppe der regelmäßigen Fragestunde-Besucher gab es keine auffälligen Änderungen, der Mittelwert der Gruppe R34 hatte sich am Ende des Semesters stark diesen beiden Gruppen angeglichen. (Vgl. Tab.II.C.71.)

Kontrolle	vorher	nachher	**R12**	Vorher	nachher	**R34**	vorher	nachher	**Frage**	vorher	nachher
MW	9.95	9.86	MW	10.38	8.92	MW	10.29	9.59	MW	9.85	9.54
Stdf.	0.05	0.07	Stdf.	0.09	0.15	Stdf.	0.05	0.10	Stdf.	0.09	0.13
Anz.	22		Anz.	13		Anz.	17		Anz.	13	
t-Test	p= 0.7473		t-Test	p= 0.002588		t-Test	p= 0.04131		t-Test	p= 0.4549	

Tab.II.C.71: Einstellungsstarke Studierende: Mittelwerte (MW) und Standardfehler (Stdf.) des Vertrauensindexes zu Beginn und am Ende des Semesters, gerundet auf die zweite Dezimale, sowie Anzahlen der Befragten (Anz.) für die vier Gruppen. Außerdem: Individuelle Vergleiche innerhalb der Gruppen: p-Werte der t-Tests (gepaart).

II.C.5. Hauptstudie: Ergebnisse aus den halbstrukturierten Interviews

Insgesamt wurden 15 Interviews zu den „R Aufgaben" geführt. Alle Studierenden, die sich zu einem Interview bereit erklärt hatten, wurden befragt. Die Bitte, sich für ein Interview zur Verfügung zu stellen, wurde auf der Homepage veröffentlicht, um möglichst viele Studierende zu erreichen. Alle Interviews wurden aufgenommen und transkribiert (vgl. Anhang). Die Interviews wurden halbstrukturiert organisiert, wobei jeweils Fragenkomplexe vorformuliert wurden, die allerdings nicht unbedingt immer in derselben Reihenfolge gestellt wurden, um den Gesprächsfluss nicht zu unterbrechen. Der Leitfaden findet sich im Anhang. Je nachdem, ob die Studierenden angaben, die „R Aufgaben" zumindest selten oder aber nie bearbeitet zu haben, wurden sie teilweise unterschiedlich befragt (Variante A bzw. B).
Die Befragten bildeten zufälligerweise annähernd die Verteilung der Studiengänge ab, die auch bei den Fragebögen aufgetreten war, und auch das Verhältnis von Studienanfängern zu Studierenden in höheren Semestern entsprach etwa dem derer, die die Fragebögen ausgefüllt hatten. Bezüglich des Bearbeitungsmodus der „R-Aufgaben" war im Vergleich zu den Fragebögen die Gruppe der Fragestundenbesucher etwas unterrepräsentiert.

Nr.	Studiengang	Gruppe	Interview-Variante	Programmier-erfahrung	Fach-semester	Note in der Klausur
Int1	L3	R12*	A	Fathom, Excel, Geogebra	12	08 P.
Int2	L3	R12*	B (überwiegend)	Excel	4	kA
Int3	Bachelor (M)	R34*	A	Java	2	kA
Int4	L3	Frage	A	Keine	4	kA
Int5	L3	R34	A	Keine	2	13 P.
Int6	Meteorologie	R34	A	Keine	4	1.0
Int7	Bachelor (M)	R12*	A	Matlab, Excel	7	kA
Int8	Bachelor (M)	R34*	A	Maple, VBA, Matlab	6	1.0
Int9	Informatik	R34*	A	Java, Python, Haskell	4	kA
Int10	Bachelor (M)	Frage	A	Python, HTML	2	kA
Int11	Bachelor (M)	R34	A	Python u.a.	2	1.0
Int12	Bachelor (M)	R12*	B	Maple, Matlab, Latex	9	nicht best.
Int13	Bachelor (M)	R34	A	Keine	2	3.7
Int14	L3	Frage	A	Keine	3	1.0
Int15	L3	Frage	A	Keine	3	2.0 (11 P.)

Tab.II.C.72. Übersicht der befragten Studierenden und ihrer Interviewnummer. Die Interviews wurden in der Reihenfolge, in der sie durchgeführt wurden, nummeriert. Bei * wurde die Einstufung in die Experimentalgruppe aufgrund der Angaben im Interview vorgenommen, bei den anderen Befragten konnten die Angaben aus dem Fragebogen übernommen werden. Bemerkung zu Int14 und Int15: Die beiden Studierenden kamen regelmäßig zu einer persönlichen Beratung zu den „R Aufgaben" in die Sprechstunde. Sie selbst haben das im Fragebogen als „R Fragestunde" aufgefasst, daher sind sie der Gruppe „Frage" zugeordnet, obwohl sie die reguläre „R Fragestunde" nicht besucht haben.

	Vertrauensindex am Ende des Semesters	Einstellungsindex am Ende des Semesters	Antwort auf die KH-/GS-Aufgabe am Ende des Semesters
Int1	-11 (schwach)	10 (stark)	-
Int2	-13 (schwach)	7 (schwach)	-
Int3	-10 (schwach)	8 (schwach)	-
Int4	-15 (schwach)	10 (stark)	Richtig
Int5	-6 (stark)	12 (stark)	Falsch
Int6	4 (stark)	10 (stark)	Richtig
Int7	-9 (schwach)	10 (stark)	-
Int8	1 (stark)	12 (stark)	Falsch
Int9	-9 (schwach)	11 (stark)	Richtig
Int10	-1 (stark)	11 (stark)	Falsch
Int11	-9 (schwach)	6 (schwach)	Richtig
Int12	-11 (schwach)	8 (schwach)	-
Int13	-1 (stark)	9 (stark)	Falsch
Int14	-17 (schwach)	4 (schwach)	Richtig
Int15	-1 (stark)	11 (stark)	Richtig

Tab.II.C.73. Die Einordnung der Skalenwerte als „stark“ oder „schwach“ orientiert sich an der Vorgehensweise bei der Auswertung der Fragebögen. Falls im Laufe des Interviews ein Lösungsvorschlag zur *Krankenhaus-Aufgabe* gemacht wurde, wurde dieser in die Tabelle eingetragen. In den anderen Fällen wurden die Lösungen der analogen *Grundschule-Aufgabe* aus den entsprechenden Fragebögen verwendet, sofern sie vorlagen.

Auffällig ist, dass alle befragten Personen aus der Gruppe derer, die die „R Aufgaben“ selten oder nie bearbeitet hatten (R12), bezüglich des Vertrauensindexes als schwach einzuordnen sind. Es waren auch insgesamt unter den Befragten mehr Personen, die ihr Vertrauen in die eigenen Fähigkeiten im Fach Stochastik als relativ gering einstuften, und alle, die dem Fach eher skeptisch gegenüberstanden, hatten auch ein geringes Vertrauen. Die Mehrheit der Befragten zeigte aber eine positive Einstellung dem Fach gegenüber. Bemerkenswert ist an dieser Stelle noch, dass weder die Note in der Klausur noch das Antwortverhalten bezüglich der Krankenhaus- bzw. Grundschule-Aufgabe am Ende des Semesters mit der Einschätzung auf der Vertrauens- oder Einstellungsskala zusammenzuhängen scheinen. (Vgl. Tab II.C.72. und Tab II.C.73.)

Die weitere Auswertung der Interviews erfolgt getrennt nach den Fragekomplexen des Leitfadens, wobei hinter der Nummer der Frage jeweils die Interview-Variante durch den Buchstaben A oder B vermerkt ist. Zur besseren Nachvollziehbarkeit werden bei der Auswertung jeweils die Interviewnummer und dahinter die Zeilennummern genannt, auf die bei der jeweiligen Analyse Bezug genommen wird.

Fragenkomplex 1A/B: Was war Deine erste Empfindung, als Du von den „R-Aufgaben“ gehört hast? Hat sich an Deiner Einstellung bis heute etwas geändert?

Insgesamt schilderten neun der Befragten eine eher negative Empfindung bzw. ablehnende Haltung als erste Reaktion auf die „R Aufgaben“ (Int1/3f., Int3/4ff., Int4/3ff., Int5/3ff., Int6/3, Int7/13, Int9/5f., Int14/3ff., Int15/3f.), drei eine neutrale bzw. nicht eindeutige (Int10/3f., Int12/3ff., Int13/3ff.,) und drei eine positive bzw. eher aufgeschlossene Haltung (Int2/3ff., Int8/3ff., Int11/3ff.).

Am häufigsten wurde als Grund für eine negative Empfindung die Erwartungshaltung angegeben, dass der Umgang mit dem Programm bzw. Programmieren im Allgemeinen eine große Schwierigkeit darstellen würde. Diese Erwartungshaltung resultierte teilweise aus negativen Erfahrungen mit anderen Programmen in der Vergangenheit (Int1/3f., Int4/3f., Int13/3f.), teilweise auch daraus, dass die Studierenden gar keine Vorkenntnisse im Programmieren hatten und daher Angst bekamen, dass die Aufgaben nicht oder nur mit immensem Zeitaufwand zu bewältigen sein würden (Int6/3f., Int14/4f., Int15/4) bzw. sie sich ohne Vorkurs überfordert fühlten (Int7/5ff./14ff.). Drei Befragte gaben an, dass sich ihnen der Sinn des Aufgabenformats nicht erschloss bzw. sie sich nicht auf ein solches eingestellt hätten (Int3/4ff., Int5/4f., Int12/3ff.), wobei zwei von ihnen daraus eine negative erste Empfindung ableiteten und einer eine neutrale.

Drei Befragte werteten ihre erste Empfindung als positiv. Als Gründe gaben sie an, dass R in der Praxis ein häufig verwendetes Paket sei (Int8/5ff.), sie eine aufgeschlossene Haltung zum Einsatz von Programmen bzw. dem Programmieren hätten (Int2/7, Int11/4ff.) und sie einen Nutzen für die Schule, insbesondere in der Oberstufe, erkennen könnten (Int2/5ff.).

Die nachfolgende Frage, ob sich an ihrer ersten Empfindung im Laufe der Zeit etwas geändert hat, bejahten acht der Befragten explizit oder implizit deutlich (Int2/8ff., Int4/4ff., Int5/6ff., Int6/10, Int9/6f., Int10/4ff., Int14/5ff., Int15/5ff.). Nur in einem Fall wurde eine zunächst positive Einstellung im Laufe der Zeit relativiert (Int2/8ff.), wobei als Begründung eine Überforderung aufgrund der Menge an neuen Inhalten angegeben wurde. Als Gründe für eine Verbesserung der Empfindung wurden die Machbarkeit der Aufgaben bzw. hilfreiche Unterstützungsangebote genannt, die den Umgang mit dem Programm erleichterten (Int9/6f., Int14/6ff., Int15/5ff.), die Einsicht in den Nutzen des Programms zur Überprüfung von Lösungen (Anwendungsfeld A1), aber auch in außeruniversitären Kontexten (Int5/7ff.), die Inhalte der Aufgaben (Int10/8f.) sowie der Eindruck, dass sie zu einem besseren Verständnis führten (Int4/5ff.). Zwei Befragte, bei denen keine (deutliche) Veränderung abzulesen war, äußerten, dass sie den mit den „R Aufgaben“ verfolgten Ansatz zwar als sinnvoll bzw.

wichtig für das spätere Berufsleben erachteten, ihnen der tatsächlich eingetretene Nutzen aber fraglich erscheine (Int1/7ff., Int7/16ff.).

Bemerkenswert ist an dieser Stelle, dass unter den Befragten diejenigen, die sich zunächst ablehnend positioniert hatten, die „R Aufgaben“ mindestens selten bearbeitet haben, während die beiden interviewten Personen, die die „R Aufgaben“ nie bearbeitet hatten, eine positive erste Haltung ihnen gegenüber schilderte. Insofern lässt die erste Empfindung keinen eindeutigen Rückschluss auf die tatsächliche Bearbeitungshäufigkeit zu.

Fragenkomplex 2A/B: Wie häufig hast Du die „R Aufgaben“ bearbeitet? Wie häufig warst Du in der Fragestunde? Ggf: Warum hast Du die Fragestunde nicht besucht?
Fragenkomplex 3B: Warum hast Du die „R Aufgaben“ nicht bearbeitet?

Mit dieser Frage sollte zunächst eine Eingruppierung der Studierenden erreicht werden, anhand derer sich der Fortgang der Befragung entschied. Nur einer der Befragten wurde aufgrund seiner Antwort der Kategorie *„nicht bearbeitet“* (Interview-Variante B) zugeteilt (Int12). Einer weiteren Person, die angab, die ersten drei angeschaut bzw. probiert zu haben (Int2/32), sie dann aber nicht weiter bearbeitet zu haben, wurden ebenfalls überwiegend die Fragen aus dieser Kategorie gestellt, um Gründe für den Abbruch der Bearbeitung herauszufinden. Alle anderen Personen wurden der Kategorie *„bearbeitet“* (Interview-Variante A) zugeteilt und entsprechend interviewt.

Vier der Befragten gaben an, die Aufgaben mit Unterstützung bearbeitet zu haben, wobei zwei davon die Unterstützung in der regulären „R Fragestunde“ (Int4/19, Int10/14) erfuhren und zwei eine regelmäßige persönliche Beratung in Anspruch nahmen, die der „R Fragestunde“ sehr ähnlich war (Int14/13ff., Int15/6ff.). Die meisten Befragten besuchten die „R Fragestunde“ nicht, wobei am häufigsten zeitliche Gründe dagegen sprachen (Int1/20, Int2/34ff., Int3/17ff., Int5/25f., Int6/16, Int9/13ff., Int13/24ff., Int15/17f.). Drei Befragte zogen eine individuelle Beratung vor (nach Bedarf: Int6/18, regelmäßig: Int14/14ff./18ff./24, Int15), zwei arbeiteten lieber in selbstorganisierten Lerngruppen mit Kommilitonen zusammen (Int1/4f./20f, Int5/19ff.). Vier Befragte gaben an, dass sie keinen Bedarf an Unterstützung hatten, weil sie die Aufgaben auch alleine bearbeiten konnten oder generell lieber alleine arbeiteten (Int5/20ff., Int7/37, Int8/15ff., Int11/17/20), wobei eine befragte Person diese Entscheidung aufgrund des fehlenden Feedbacks und weniger Unterstützung bereute (Int7/39ff./44ff.). Eine Person gab an, kein Interesse daran gehabt zu haben, weil sie vermutete, die Klausur auch ohne die „R Aufgaben“ bestehen zu können, und erklärte auf (als suggestiv zu wertende) Nachfrage, dass sie zudem Zeit hätte sparen wollen (Int12/25f./31/39).

Fragenkomplex 3A: Fandest Du die „R Aufgaben“ im Vergleich zu den anderen Aufgaben eher schwer oder leicht? Was war schwerer, was war leichter?

Die „R Aufgaben“ wurden von acht Befragten als eher leichter im Vergleich zu den anderen Aufgaben bewertet (Int1/25ff., Int2/42ff., Int4/26ff., Int5/29, Int7/52f., Int9/18, Int10/17f., Int14/27f.). Drei Befragte schätzten sie als schwieriger ein (Int3/27ff., Int11/23ff., Int15/21/31ff.), die anderen kamen zu keinem eindeutigen Ergebnis bzw. stuften sie als gleichwertig schwierig ein (Int6/21ff., Int8/25ff., Int13/31/45f.).

Als eine Hauptschwierigkeit wurde das algorithmische Denken formuliert, das zusätzlich zu den stochastischen Kompetenzen für die Lösungen der Aufgaben erforderlich war (Int4/24ff./35ff./42, Int7/53ff./57ff., Int10/20ff., Int15/21ff.). Passend dazu empfand eine befragte Person, die sich als vertraut mit Programmieren charakterisierte, die „R Aufgaben“ als leichter als die anderen (Int9/20ff.). Zwei der Befragten gaben an, dass es für sie teilweise schwierig war zu verstehen, was sie überhaupt tun sollten, bzw. dass eine mögliche Herangehensweise erst durch die Templates deutlich geworden sei (Int3/29ff., Int11/23ff./34f.). Dem gegenüber formulierte eine andere befragte Person, dass ihr bei den „R Aufgaben“ klarer gewesen sei, was zu tun sei, als bei den anderen Aufgaben, da ihr die Herangehensweise einleuchtender gewesen sei (Int14/31ff./36f.). Eine weitere Person gab an, dass sie durch die Hilfestellung im Template die „R Aufgaben“ als einfacher empfunden habe (Int5/31ff.). Eine Person führte aus, dass der Zusammenhang zu anderen Aufgaben die „R Aufgaben“ für sie erleichtert habe (Int2/42ff.). Zwei der Befragten hoben hervor, dass der Praxis- bzw. Anwendungsbezug motivierend gewesen sei (Int1/29ff., Int4/48ff.), und einer, dass die Inhalte greifbarer und die Beschäftigung mit den „R Aufgaben“ als nützlich empfunden worden sei (Int4/48ff./53ff.).

Fragenkomplex 4A: Was fandest Du an den „R Aufgaben“ gut? An welche „R Aufgaben“ erinnerst Du Dich noch und warum? Ggf als Nachfrage: Gibt es etwas, was Dir durch eine „R Aufgabe“ klar geworden ist, was Du vorher nicht verstanden hattest? Wenn ja: was?

Fragenkomplex 5B: Hattest Du den Eindruck, dass Du dadurch, dass Du die Aufgaben nicht bearbeitet hast, Nachteile hattest? Wenn ja: welche? // Fragenkomplex 7B: Gibt es etwas, was man mit Simulationsaufgaben Deiner Meinung nach besonders gut lernen kann?

Insgesamt sieben der 13 Personen, denen diese Fragen gestellt wurden, sahen in den „R Aufgaben“ einen Gewinn auf der methodisch-technischen Ebene des Programmierens, da die Aufgaben in den Umgang mit dem Programm R einführten (Int4/83f., Int5/42ff., Int6/29ff., eingeschränkt Int7/83ff., Int10/30ff., Int11/52f., Int15/37ff.), was als nützlich empfunden wurde. Zwei Befragte betonten dabei die große Bedeutung des Programmierens für die moderne Mathematik im Allgemeinen (Int4/109ff., Int10/31ff.). Zwei andere Befragte hoben in diesem Zusammenhang hervor, dass für das Erlernen des Programmierens mit R besonders die Abmilderung technischer Schwierigkeiten durch das Template wichtig gewesen seien (Int6/29ff., Int8/30ff.), während ein Befragter zu viele Vorgaben im Template bemängelte (Int11/53f.) und ein anderer ausführte, dass durch die Vorgaben das Programmieren selbst nicht richtig gelernt wurde, dies aber durchaus für angemessen hielt (Int5/86ff.).

Dass die für die Programmierung notwendige Strukturierung der stochastischen Zusammenhänge auch für das inhaltliche Verständnis hilfreich sein könnte, wurde von einer weiteren Person ausgeführt (Int15/81ff./89f.).

Das Anwendungsfeld (A1), Simulationen als Hilfsmittel zur Überprüfung von theoretischen Ergebnissen oder Verfahrensweisen zu nutzen, wurde von acht Befragten umschrieben, die das Positive bei den „R Aufgaben" in der Möglichkeit sahen, Ergebnisse zu überprüfen oder erste Ideen zu testen (Int5/43f., Int11/66ff./73ff., Int13/79ff.) und theoretische Erkenntnisse zu veranschaulichen (Int3/54ff., Int6/44ff./50, Int9/26ff./56ff./64ff./69f., Int15/73ff.) bzw. zu vertiefen (Int4/84ff./130ff.), wobei eine befragte Person explizit formulierte, dass dies dann auch ihr Vertrauen in die theoretischen Ergebnisse erhöhe (Int15/77f.). Eine andere Befragte merkte an, dass sie durch die Simulationen ein „Gefühl" für Verteilungen bekommen habe, das sie durch Berechnungen nicht hätte erlangen können, da sie nie sicher gewesen sei, ob diese richtig waren (Int9/64ff./69f.). Daran anschließend stand für vier Befragte die Visualisierung bzw. Übersetzung in Graphiken, die eine Interpretation erleichterten, im Vordergrund (Int1/38, Int4/84ff./93, Int13/63ff./68, Int15/41ff./73f.), worin anklingt, dass die „R Aufgaben" eine Hilfe zur Versprachlichung sein könnten. Eine Person formulierte explizit, dass es für sie wichtig war, dass sie die simulierten Ergebnisse interpretieren konnte (Int14/74ff.). Ein Befragter gab an, dass er insbesondere eine Visualisierung von Grenzwertsätzen als lohnenswert empfunden habe, da diese formalisiert schwer vorstellbar seien, anhand von Simulationen aber einleuchtend würden (Int8/49ff.). Eine andere Person wertete als positiv, dass in den „R Aufgaben" zuvor behandelte Themen wiederholt worden seien (Int11/46ff.).
Außerdem beschrieben sechs der Befragten Aspekte der „R Aufgaben" als positiv, die in Zusammenhang mit dem Anwendungsfeld von Simulationen als Hilfsmittel zur Exploration und Generierung neuer Hypothesen (A2) gebracht werden können, nämlich, dass man durch sie die Möglichkeit erhalten habe, etwas auszuprobieren bzw. Experimente mit vielen Wiederholungen durchzuführen, bei denen die Zufallsergebnisse sichtbar gemacht worden

seien (Int1/34ff., Int3/44ff., Int7/64ff., Int10/51ff./69ff., Int14/57ff., Int15/44f.), selbst wenn man die dahinter liegenden Zusammenhänge noch nicht vollständig durchdringen konnte (Int1/44ff.). Ein Befragter formulierte, dass es durch die Simulationen möglich sei, ein „Gefühl" für stochastische Zusammenhänge zu entwickeln (Int10/71f.). Eine andere befragte Person wertete die Möglichkeit des Experimentierens als „Realitätsbezug" (Int14/45ff.) und wieder andere beschrieben, dass die Inhalte greifbar geworden seien (Int1/47/207ff., Int4/48/86ff.). Zwei Befragte betonten das darin enthaltende spielerische Moment (Int5/45f., Int7/66ff.), das Freude bereitet habe. Eine Person sah in der Experimentiermöglichkeit auch einen Bezug zur Schule (Int1/34f/211ff..).

Das Anwendungsfeld (A3), Simulationen als Hilfsmittel zur Annäherung nicht oder nur sehr schwer berechenbarer Größen zu verwenden, deutete nur eine befragte Person an, indem sie sagte, dass durch das Programm auch Ergebnisse für Aufgaben produziert werden könnten, bei denen sie keinen Ansatz für eine rechnerische Lösung hätte (Int1/202ff.), was sie auch als „Anwendungsbezug" umschrieb.

Die beiden Personen, die die „R Aufgaben" nicht oder nur äußerst selten bearbeitet hatten, bezogen sich bei der Frage, ob sie dadurch Nachteile hätten erkennen können, auf keines der drei benannten Anwendungsfelder von Simulationen, sondern antworteten rein strukturell auf die Veranstaltung bezogen, dass sie dadurch weniger Punkte auf die Übungsblätter bzw. in der Klausur hätten erreichen können (Int2/67ff., Int12/46/71f.). Allerdings gingen sie bei der Frage, ob man an den Aufgaben etwas besonders gut hätte lernen können, auf das Anwendungsfeld (A1) der Überprüfung von Lösungen und (A3) der Entwicklung neuer Fragestellungen ein (Int2/92ff./97ff.) bzw. benannten, dass Simulationsverfahren auch eine Verknüpfungsmöglichkeit mit anderen Fächern bieten würden (Int12/78f./81f.).

Auf die Frage an diejenigen, die die „R Aufgaben" zumindest selten bearbeitet hatten, an welche sie sich denn noch erinnerten, wurden am häufigsten Aufgaben zu Monte Carlo Simulationen genannt (Int5/54/67, Int6/33ff., Int8/42, Int9/33, Int10/37, Int11/56f., Int15/46ff./53f.), wobei diesbezüglich fünf Mal auf die Aufgabe 8 zum Rejection Sampling und der anschließenden Monte Carlo Simulation zu Bestimmung des relativen Flächenanteils eines Kugeldreiecks Bezug genommen wurde (Int5/54/67, Int9/33/35, Int10/37, Int11/56f., Int15/46ff./53f.). Außerdem erinnerten sich viele an die Aufgaben 4 bzw. 5 zur Größenverzerrung, illustriert an Belegzeiten von Betten in einem Krankenhaus (Int4/107, Int7/71f., Int9/56, Int10/39, implizit Int13/40ff./50ff., Int14/40). Aber auch alle anderen Aufgaben wurden von mindestens einer befragten Person erwähnt.

Als Gründe für die Erinnerungen wurden genannt:

- eine neue Erkenntnis, nämlich wie durch mehrere Testgruppen eine Verfälschung statistischer Ergebnisse verhindert werden kann (Aufgabe 7 zum Senderwechsel, Int3/38ff./50f. und Int14/66ff.), bzw. ein überraschendes oder unerwartetes Ergebnis (Aufgabe 6 zu den gefälschten Münzwürfen, Int9/39ff./43; Aufgabe 11 zu den bedingten Wahrscheinlichkeiten, Int4/62/65ff.),
- ein Alltagsbezug (Aufgabe 11 zu den bedingten Wahrscheinlichkeiten, Int4/76ff.),
- eine Empfindung von Nützlichkeit auch für andere Kontexte (Aufgabe 3 zur Monte Carlo Simulation, Int6/34ff.),
- eine Verbindung zur Vorlesung (Aufgaben 4 und 5 zu Belegzeiten von Betten im Krankenhaus, Int9/56ff.; Aufgabe 10 zu den Konfidenzintervallen, Int1/39f.) oder zu einer theoretischen Aufgabe (Aufgabe 2 zu Crabs, Int5/53/76ff., Int8/40f.), wobei eine befragte Person angab, durch das Programmieren das Spiel Crabs erst richtig verstanden zu haben,
- technische Schwierigkeiten bei der Lösung (Aufgaben 4 bzw. 5 zur Größenverzerrung, Int13/50ff.) bzw. Unklarheiten bei der Umsetzung im Template (Int5/67f.)
- und inhaltliche Verständnisschwierigkeiten (Aufgabe 2 zu Crabs, Int4/60f.; Aufgabe 8 zum Rejection Sampling, Int9/32, Int10/37f.).

Fragenkomplex 5A: Was hat Dir bei den „R Aufgaben" nicht gefallen?

Auf die Frage, was ihnen bei den „R Aufgaben" nicht gefallen hat, gaben insgesamt zehn der 13 Befragten eine Antwort, die sich auf die technische Umsetzung bezog: Fünf Personen empfanden das Programmieren bzw. die Fehlersuche als mühsam (Int4/135ff., Int5/98ff., Int14/80ff./85f., Int15/93ff.) oder sogar abschreckend (Int1/49ff.), zwei Befragte gaben an, dass der Umgang mit R schwierig war, weil es für sie ein neues Programm war, in das sie noch nicht eingearbeitet waren (Int10/74ff., Int11/78ff.) und eine Person formulierte, dass ihr Programmieren generell keinen Spaß mache (Int13/95ff.). Zwei weitere Befragte empfanden die Aufgaben als zu schwer, um sie ohne Hilfe lösen zu können (Int3/61ff./70ff., Int7/87ff.) und bewerteten auch das Template und die „R Fragestunde" als kritisch, weil sie beides aufgrund zu wenig oder zu viel Hilfestellung nicht als günstig empfanden, um das Programmieren zu erlernen (besonders Int7/91ff./97f./100ff.). Ein Befragter hätte sich bei besserer Einarbeitung gewünscht, dass man noch mehr selbst programmieren hätte müssen

(Int11/80ff.) und bemängelte zudem die unzureichende Integration in das reguläre Tutorium (Int11/82ff.).

Außerdem empfanden zwei der Befragten manche Aufgaben als zu umfangreich (Int8/56ff.) oder für sie schwer verständlich (Int9/72f./75f.) und eine Person bemängelte, dass bei den „R Aufgaben“ besonders viel abgeschrieben werden konnte (Int13/105ff./135/137/139/141). Eine befragte Person konnte nichts Negatives benennen (Int6/58).

Fragenkomplex 6A: Würdest Du aus Studentensicht empfehlen, weiterhin „R Aufgaben“ zu verwenden? Würdest Du etwas verändern?

Fragenkomplex 4B: Was hätte man anbieten können, damit Du die „R Aufgaben“ bearbeitet hättest? // Frage 6B: Findest Du es prinzipiell eine gute Idee, in Stochastik Simulationsaufgaben zu stellen?

Die Befragten positionierten sich eindeutig zur Frage der Empfehlung von „R Aufgaben“, wobei alle bis auf eine Person (Int3/76), diese bejahten (Int1/59, Int2/77, Int4/175, Int5/109/128, Int6/60, Int7/106, Int8/67, Int9/78, Int10/84, Int11/88, Int12/75, Int13/144, Int14/89/103f., Int15/107), insbesondere auch die beiden Studierenden, die diese gar nicht oder nur äußerst selten bearbeitet hatten. Zwei Studierende begründeten ihre Empfehlung mit der Möglichkeit, dadurch stochastische Konzepte sichtbar machen bzw. veranschaulichen zu können (Int2/77ff., Int8/77). Drei andere Befragte argumentierten damit, dass die Fertigkeit, Simulationen einzusetzen bzw. generell programmieren zu können, auch hilfreich für den weiteren Fortgang des Studiums (Int11/92ff., Int13/145) bzw. das spätere Berufsleben (Int7/106ff.) sei. Eine Person nahm auf das Anwendungsfeld (A1) der Überprüfung von Lösungen Bezug (Int5/109ff.) und eine andere empfand die „R Aufgabe“ in der Klausur als angenehmes Format (Int11/88/90) und empfahl den Einsatz aus diesem Grund weiter.

Als Veränderungsvorschläge wurden genannt:

- Es könnte einen erleichterten Einstieg bzw. einen Vorkurs geben (Int2/53ff., Int7/118ff., Int11/80ff.), wobei eine Person diesen sogar verpflichtend machen würde (Int7/122ff.), und, wie bereits bei der Analyse des vorherigen Fragenkomplexes genannt, eine andere Person darauf basierend mehr eigenständige Programmierung wünschenswert fände. Zudem wäre für drei Befragte eine bessere Integration in das reguläre Tutorium wünschenswert (Int1/169ff., Int8/84ff., Int11/82ff.).

- Eventuell könnten die „R Aufgaben“ auch nur für Lehramtsstudierende angeboten werden, da die Verwendung von Simulationen in der Oberstufe oder auch als Angebote für einzelne Schüler gewinnbringend sein könnte (Int1/59ff.).
- Eine Absprache mit anderen Veranstaltungen bezüglich des Programms könnte hilfreich sein, damit weniger unterschiedliche Programme erlernt werden müssen (Int4/146ff.).
- Eine bessere Erläuterung bzw. Reflexion des Nutzens von Computereinsätzen im Studium bzw. der Mathematik allgemein könnte zu mehr Akzeptanz führen (Int4/154ff.).
- Häufiger den Quellcode kommentieren zu lassen statt eigene Programmierungen zu fordern, könnte Frustrationen durch formale Fehler wie falsche Klammern abmildern (Int5/133ff.).
- Simulationen sollten auch in andere Vorlesungen, z.B. Stochastische Prozesse, integriert werden (Int8/67ff./71ff.).
- Häufiger eine Verknüpfung zum theoretischen Hintergrund oder anderen Aufgaben herzustellen, könnte gewinnbringend sein (Int9/84ff., Int11/112f./114f./121).
- Für Studierende, die mehr üben möchten, könnten zusätzlich noch freiwillige Aufgaben angeboten werden (Int10/86ff.).
- Der zeitliche Aufwand, der mit der Bearbeitung verbunden ist, könnte stärker honoriert werden, z.B. durch weitere Bonuspunkte (Int14/92ff./96ff.).

Fragenkomplex 7A/8B: Wurde in deinem Tutorium manchmal auf R Aufgaben Bezug genommen, als eine andere Aufgabe an der Tafel erklärt wurde? Falls ja: Erinnerst Du Dich an Beispiele?

Insgesamt zehn der Befragten verneinten diese Fragen oder gaben an, dass es ihnen zumindest nicht mehr im Bewusstsein geblieben sei und nannten im Folgenden auch kein Beispiel (Int1/92ff./100ff., Int2/104, Int4/185ff., Int6/68f./71, Int7/132ff., Int9/91, Int12/87, Int13/168, Int14/109, Int15/119ff.). Drei weitere Studierende verneinten die Frage ebenfalls, benannten dann aber doch jeweils ein Beispiel für eine solche Situation, nämlich die Aufgabe zu den Fehlständen (Int3/81ff.), Crabs (Int5/143ff.) und die letzte „R Aufgabe“ zu den bedingten Wahrscheinlichkeiten, anhand derer auch noch einmal formale Schreibweisen geklärt worden seien (Int11/102f./105ff.).

<u>Fragenkomplex 8A/9B: Welches Objekt oder welcher Begriff fällt Dir als erstes ein, wenn Du an Stochastik denkst? Erkläre, was das ist!</u>

Als Antwort auf diese Frage wurden neben Gegenständen wie Münzen und Würfeln folgende Begriffe und Objekte genannt, die das Spektrum der Vorlesung überwiegend abdeckten. In der nachfolgenden Tabelle sind die genannten Begriffe nach Gruppenzugehörigkeit bzgl. des Bearbeitungsmodus der „R-Aufgaben" aufgelistet:

- Wahrscheinlichkeit (Int1/110, Int5/154, Int9/96) und Wahrscheinlichkeitsrechnung (Int2/124, Int13/171)
- unkorrelliert (Int1/110)
- unabhängig/Unabhängigkeit bzw. stochastische Abhängigkeit (Int1/111, Int4/218, Int6/80)
- Kombinatorik (Int2/118, Int11/134)
- Münzwurf und fairer Münzwurf (Int3/89, Int5/156)
- Einschluss-Ausschluss-Formel (Int3/97)
- statistisches Beispiel aus der Verlesung (Int 4/198) bzw. allgemein Statistiken (Int13/171)
- Zwei-Stufen-Modelle/ zweistufige Zufallsexperimente (Int4/218, Int15/133)
- Zufallsexperiment (Int5/173)
- Zufallsvariable (Int5/173), Simulationen von Zufallszahlen (Int7/140), Zufall (Int10/124)
- Allgemein Verteilungen (Int6/87) und speziell die Binomialverteilung (Int13/208) bzw. die Bernoulli-Verteilung (Int10/141)
- Erwartungswert (Int7/146, Int14/136)
- Varianz (Int7/146) und Kovarianz (Int14/136)
- Brownsche Bewegung (Int8/97)
- Indikatorvariable (Int8/100)
- Dreisatz (Int12/96)
- Markowketten (Int12/100, Int15/135)
- Approximationen (Int13/175)
- Polya-Urne (Int13/191)
- Hypothesentests (Int14/127)

Gruppe	genannte Begriffe
R12	Wahrscheinlichkeit und Wahrscheinlichkeitsrechnung, unkorrelliert, unabhängig/Unabhängigkeit bzw. stochastische Abhängigkeit, Kombinatorik. Simulationen von Zufallszahlen, Erwartungswert, Varianz, Dreisatz, Markowketten
R34	Wahrscheinlichkeit und Wahrscheinlichkeitsrechnung, unabhängig/Unabhängigkeit bzw. stochastische Abhängigkeit, Kombinatorik, Münzwurf und fairer Münzwurf, Einschluss-Ausschluss-Formel, Statistiken, Zufallsexperiment, Zufallsvariable, Verteilungen bzw. Binomialverteilung, Brownsche Bewegung, Indikatorvariable, Approximationen, Polya-Urne
Frage	unabhängig/Unabhängigkeit bzw. stochastische Abhängigkeit, statistisches Beispiel aus der Vorlesung, Zwei-Stufen-Modelle/ zweistufige Zufallsexperimente, Zufall, Bernoulli, Erwartungswert, Varianz und Kovarianz, Markowketten, Hypothesentests

Tab.II.C.74. Begriffe aus der Stochastik, die von den Befragten im Interview genannt wurden, aufgeschlüsselt nach Experimentalgruppe.

Die in den Interviews formulierten Erklärungsversuche werden im Folgenden nach Begriffen sortiert zusammengefasst. Bei den Interviews entnommenen Zitaten ist an der Seite der von der befragten Person gewählte Bearbeitungsmodus der „R Aufgaben“ vermerkt.

- Wahrscheinlichkeit und Wahrscheinlichkeitsrechnung:
  Eine inhaltlich ausgefüllte Erklärung der Begriffe (zumindest bezüglich Laplace-Experimenten) gelang nur einer Person, die Wahrscheinlichkeiten als Anteile von „Günstigen“ durch „Mögliche“ beschrieb (ohne diese Begriffe zu nennen). Außerdem wurde versucht, die Begriffe an Beispielen von Zufallsexperimenten begreiflich zu machen. Eine andere Herangehensweise war, das Zahlenformat anzugeben. Die Idee, ein Synonym für das Wort „wahrscheinlich“ zu finden, führte nicht zum Erfolg.
    - „Wahrscheinlichkeit gibt an, hm, dass irgendein Fall, also wenn ich jetzt Lotto spiele, ähm, wie wahrscheinlich es ist, dass eben eine Zahl von diesen 49 gezogen wird, oder dann auch dass man unterscheidet, ob irgendwelche Dinge gleichwahrscheinlich sind oder eben ob es da noch einen Unterschied gibt“ (Int1/120ff.)) R12
      „Es ist so was Prozentuales“ (Int1/125))
    - „[lacht] Das ist schwierig, das Wort zu beschreiben. Also, ich denk jetzt, dass ich weiß, was es ist, aber ich glaub, ich kann es nicht beschreiben. […]“ (Int5/159f.) „Das ist halt, ich versuch grad das Wort Wahrscheinlichkeit anders auszudrücken. […]“ (Int5/164f.) „Also, es geht ja äh im Prinzip um […] die Wahrscheinlichkeit von Ereignissen [lacht]. Ja, dass halt irgendwas eintritt oder nicht eintritt, aber […] mir fällt kein besseres Wort für Wahrscheinlichkeiten ein.“ (Int5/167ff.) R34
    - „[…] also ich würd das so übersetzen, dass Stochastik Wahrscheinlichkeitsrechnung ist. Dass ich also Wahrscheinlichkeiten für verschiedene Ereignisse und so weiter ausrechne dann.“ R12

Auf Nachfrage, was eine Wahrscheinlichkeit ist: „Hm, wie wahrscheinlich etwas ist, dass ein Ereignis eintritt. Also, ich glaub ich bin einfach noch sehr am Anfang irgendwie.“ (Int2/124ff.)

- „Wahrscheinlichkeit, ich denke das äh von so viele, zum Beispiel Münzwürfe, von so viel Mal, wenn man eine Münz werfen, ok, wie oft dann kriegt man Kopf oder Zahl. Ja, das ist die klassische Art, wie man erstmalig von Stochastik lernt.“ (Int9/98ff.) R34

- unkorreliert:

  Diese Eigenschaft wurde in Relation zur Unabhängigkeit erklärt:

  „Wenn das erfüllt ist, dann ist es so“ (Int1/116), R12

  was allerdings nicht zutreffend ist, da aus Unkorreliertheit im Allgemeinen keine Unabhängigkeit gefolgert werden kann (außer etwa bei normalverteilten Zufallsvariablen, bei denen die beiden Eigenschaften äquivalent sind). Umgekehrt wäre es zwar richtig, aber auch noch keine inhaltlich erhellende Beschreibung der Eigenschaften.

- unabhängig/ Unabhängigkeit und stochastische Abhängigkeit:

  Den beiden Studierenden, die die Begriffe versuchten zu erläutern, gelang es nicht, den Fachbegriff der (stochastischen) Unabhängigkeit von dem umgangssprachlichen Wort abzugrenzen. Ein Studierender versuchte, den Begriff (im Fall von zwei Zufallsvariablen) durch eine Formel zu erklären, konnte sie aber nicht vollständig angeben. Zudem ist eine Konfusion der Begriffe Unabhängigkeit und Korrelation zu beobachten.

  - „unabhängig, wenn die Ergebnisse von zwei Versuchen voneinander unabhängig sind, logischerweise, wenn die hintereinander ausgeführt werden und dann der erste den zweiten beeinflusst, dann ist es abhängig, ja.“ (Int6/83ff.) R34
  - „Ok, Unabhängigkeit drückt einfach nur aus, ob zwei Ergebnisse unabhängig voneinander auftreten oder nicht. Das berechnet sich dadurch, dass ich einfach gucke, wie nenn ich das eigentlich, also P von X_1 Element A_1 Komma X_2 Element A_2 ist gleich P von X_1 Ele, ähm, und P von X_2. Ja, also so sehr unscheinbar, aber wie gesagt es kann sehr, es ist ein sehr mächtiges Instrument, weil es einfach extreme Auswirkungen hat. Wenn ich wieder zurückgehen würde auf die Statistik zum Beispiel, also ich glaube das statistische Pendent dazu ist die Korrelation, also zum Beispiel, wenn ich zwei, hab ich was sehr lustiges drüber gelesen, wenn ich zwei verschiedene Entwicklungen hab, müsst ich eigentlich immer prüfen, ob die abhängig oder nicht abhängig sind, also zum Beispiel wenn die Politik sagt, seit 2002 gibt es irgendwie 500 000 weniger Arbeitslose, also klappt Hartz 4 irgendwie, muss aber nicht stimmen. Also ich hab da was gelesen von einem Herrn ich glaub Henze war das oder so, der hat geschrieben, wenn naja, also vor 20 Jahren gabs von mir aus weniger Kinder auf der Welt, also weniger, weniger Menschen und vor 20 Jahren gabs von mir aus auch weniger Störche auf der Welt und heute gibt’s mehr Störche und mehr Kinder, das ist dann der Beweis dafür, dass die Störche die Kinder bringen.“ (Int4/222ff.) „Stimmt halt nicht, ne. Und deswegen müsste man halt bei so was, also da bräuchte man A wieder ein Grundwissen in Stochastik, damit einfach so was anzweifelt und man müsste halt auch einfach immer bei so Daten auch einfach mal die Mathematik dahinter mit angeben. Wie groß ist denn die Abhängigkeit oder gibt’s überhaupt ne Abhängigkeit, ist das vollkommen abhängig, ist das unabhängig? Das ist so was, das sind sehr wichtige Fragen, die man eigentlich dazu beantworten sollte.“(Int4/239ff.) Frage

Auf eine Nachfrage zu der von ihm genannten Formel formulierte der Studierende dann aber doch noch eine Aussage, die auf eine richtige Vorstellung des Konzepts hindeutet.

Aus seiner Ausführung lässt sich derweil auch schließen, dass sich sein Begriff von Zufall noch sehr an dem in der Alltagsverwendung orientierte:

> „Achso. Ähm. Wir stellen uns eine Urne vor und wir ziehen zwei Kugeln raus. Irgendwelche, von mir aus die Kugel mit der Nummer 7 und die Kugel mit der Nummer 8. Wir haben die 7 und die Nummer 8 gezogen. Wenn […] ich […]diese beiden Kugeln gezogen habe, und die sind unabhängig davon, heißt das […] ist ne gute Frage […] Was heißt das denn eigentlich? […]Es gibt keine, keine Abhängigkeiten, mh, ich hab darüber noch nie nachgedacht, das ist ne gute Idee. Ähm […] Ich weiß es gar nicht, glaub ich. Also ich mein, es gibt keine Abhängigkeiten und es macht keinen Unterschied, ob du die quasi […] nein, das Ereignis, dass du quasi beide ziehst, ist, wie soll ich sagen, es hängt von nichts ab, es ist quasi purer Zufall. Also wenn du sagen würdest, etwas ist nicht unabhängig, also es ist abhängig, hieße es ja, wenn ich die 7 ziehe, würde sich irgendwie die Chance erhöhen, dass ich die 8 bekomme. Wenn sie unabhängig sind, ist das eben nicht so. So würd ich das für mich, glaub ich, grad stehenlassen. (Int4/247ff.)

Frage

- Kombinatorik:

  Den beiden Personen, die diesen Begriff genannt hatten, gelang eine inhaltlich zutreffende Beschreibung, wobei in einem Fall keine ganz klare Trennung der Begriffe Kombinatorik und (Laplace-) Wahrscheinlichkeiten vorgenommen wurde.

  - „Ähm, dass ich, ja, mit Kombinatorik versuch ich ja auszurechnen, wieviele Möglichkeiten es gibt und dann Günstige durch Mögliche, also in jedem Falle wieviele Möglichkeiten gibt es für verschiedene Ereignisse. “ (Int2/120ff.) R12
  - „Ähm, im Grunde ist das ja ein Zählen, also so […] ja irgendwie Methoden, um das zu vereinfachen, dass man eben nicht alles zählen muss, sondern Methoden dafür hat.“ (Int11/140f.) R34

- Fairer Münzwurf:

  Bei diesem Begriff ließ sich aus der Erklärung trotz sprachlicher Ungenauigkeit die richtige Beschreibung ableiten.

  > „Ja, also das ist halt mit der Wahrscheinlichkeit 0,5 dass beide Parteien, also dass es gleich, dass es gleich wahrscheinlich fällt.“ (Int3/91f.) R34

- Einschluss-Ausschluss-Formel:

  Die Formel wurde von dem Studierenden nicht genannt (auch nicht für einen Spezialfall), sondern anhand einer Visualisierung die Grundidee beschrieben. Allerdings konnte kein Bezug zur Stochastik hergestellt werden, da dem Studierenden die Mengenvorstellung von Ereignissen nicht präsent war.

  > „Also der NN. (Mitarbeiter des Lernzentrums, F.) hat mir mal die Einschluss-Ausschluss-Formel erklärt, das fand ich eigentlich, da hat er hier glaub ich sechs Kreise oder was gemalt, weil ich es einfach nicht verstanden hab und hat mir hier alle wirklichen Einzelheiten rausgeschrieben und das war dann am Ende sooo ein Term und, ja also daran kann ich mich noch ganz gut erinnern, weil ich das wirklich dann auch verstanden hab.“ (Int3/96ff.) R34
  >
  > Auf Nachfrage, wozu man sie braucht: „Puh, [lacht], das ist ne gute Frage, äh, also jetzt zu Stochastik, ähm, also da ging‘s ja um, um, dass man äh sozusagen mehrere Mengen vereinigt und äh die

Schnittmengen halt nicht doppelt zählt, also doppelt und dreifach, irgendwie x-fach zählt, und dass man die dann halt sozusagen rausnimmt um dann insgesamt die gesamte Menge zu haben, aber was das jetzt genau in Stochastik war, das weiß ich nicht mehr.“ (Int3/104ff.)

- statistisches Beispiel aus der Vorlesung:

  „[…] ich find, das ist interessant, dass du aus einem Satz Daten versuchst, Information zu ziehen […]“ (Int4/202f.) und „[…]weil man da eben versucht Information aus den Daten zu gewinnen und die irgendwie aufzubereiten und darzustellen für die, die eben nicht die extreme Leuchte in der Statistik sind.“ (Int4/210ff.) Frage

  Die Grundidee der Statistik, Daten auszuwerten, um aus ihnen (verallgemeinerbare) Schlüsse zu ziehen, war dem Studierenden präsent, wenngleich unklar ist, wie der letzte Teil seiner Ausführung gemeint ist.

- zweistufige Zufallsexperimente:

  Die zweistufigen Zufallsexperimente erläuterte der Studierende an einem Beispiel aus der Vorlesung, wobei im Fokus der Ausführungen bedingte Wahrscheinlichkeiten standen, die in diesem Kontext zutreffend beschrieben wurden. Ob dem Studierenden klar war, dass das Ergebnis von 50% von der Prävalenz der Krankheit abhängt, kann dabei allerdings nicht beurteilt werden.

  - „zweistufiges Zufallexperiment ist ähm, da geht’s auch um bedingte Wahrscheinlichkeiten und da ist immer die Frage, dass ähm zum Beispiel ähm bei Krankheiten beispielsweise, man macht einen Test und man will wissen, ob jemand diese Krankheit hat. Und dann ähm ist immer so die Frage, ja, wie hoch ist die Wahrscheinlichkeit, dass der Test äh so ausfällt, dass dieser Mensch krank ist, er aber tatsächlich gesund ist.“ (Int15/139ff.) Frage
    „So und diese zweistufige also Zufallexperimente, dass man sagt, ok, für den Fall, er ist gesund und der Test ist positiv oder negativ, oder er ist krank und positiv, negativ, das ist so. Das fand ich auch ziemlich, da fand ichs ziemlich cool auch zu erfahren, dass ähm man nur ne 50%ige Wahrscheinlichkeit hat, dass wenn der Test so ausfällt, dass er sagt, man sei krank, dass man im Endeffekt mit 50%iger Wahrscheinlichkeit trotzdem gesund sein kann, das fand ich ziemlich cool.“ (Int15/146ff.)

- Zufallsvariable:

  Eine (nicht maßtheoretische) Vorstellung, was eine Zufallsvariable ausmacht, wurde auf Nachfrage nach einer Umschreibung statt einer Definition von dem Studierenden formuliert, wobei sich die Vorstellung möglicherweise nur auf diskrete Zufallsvariablen, nämlich die Summe von Indikatorvariablen (Zählen), beschränken könnte.

  „[…] Also, mir fällt auch keine wirklich gute Definition für ne Zufallsvariable ein.“ (Int5/178f.) Frage

Auf Nachfrage: „Also, man zählt immer, halt ne Variable, die halt aus nem bestimmten, aus nem bestimmten Bereich irgendwelche Werte annimmt mit ner gewissen Wahrscheinlichkeit oder halt auch nicht.“ (Int5/181ff.)

- Simulationen von Zufallszahlen:

  Der Studierende, der als erstes an Simulationen von Zufallszahlen dachte, bezog sich bei seiner Erläuterung auf technische Aspekte. Aus seiner Ausführung lässt sich eine differenzierte Beschäftigung mit dem Thema Pseudozufallszahlen schließen:

  „Dann denk ich sofort an Simulationen, ja, also an Simulationen von Zufallszahlen, ja, und auch die Frage, was ist guter Zufall, was ist stabiler Zufall, ja.“ (Int7/140f.) „Ja, da ist ja jede Maschine auch beschränkt. Ja, das geht ja schon fast auch in ne philosophische Frage dann auch rein.“ (Int7/143f.) R12

- Zufall:

  Der Begriff Zufall wird durch seine Verwendung im Kontext von Zufallsexperimenten erläutert:

  „Ähm, ja, ah ja, gut, ne formale Definition, also man hat ähm, man hat Ereignisse und jedes Ereignis tritt eben mit irgend ner bestimmten Wahrscheinlichkeit ein, die man vorher kennen kann, aber nicht muss, und ähm das Zufallsexperiment ist dann eben, dass man äh ja, so, so nen ähm Durchlauf macht, wo alle Ereignisse vorkommen können und dann guckt man eben, ja, der Zufall bestimmt dann eben, welches Ereignis halt rauskommt.“ (Int10/130ff.) Frage

- Verteilungen und Binomialverteilung bzw. Bernoulli-Verteilung:

  Beide Studierende, die verschiedene Verteilungen bzw. nur die Binomialverteilung genannt hatten, versuchten eine Erklärung anhand von Beispielen, wobei der eine lediglich die Namen verschiedener Verteilungen aufzählte und dabei eine Begriffsverwirrung in Form des Wortes „Dichteverteilung“ zeigte, der andere die Binomialverteilung mit dem Spezialfall der Bernoulli-Verteilung verwechselte.

  - „Sehr wichtig, Normalverteilung, Exponentialverteilung, vor allem die Dichteverteilung. Die diskreten Verteilungen, was hatten wir denn da Schönes, geometrisch, hypergeometrisch, uniform verteilt, ja. Vor allem die kontinuierlichen, die sind, glaub ich, auch für mich dann nochmal wichtig, die kann ich noch gebrauchen.“ (Int6/88ff.) R34
  - Binomialverteilung: „Also so was wie Münzwurf oder ja. Das ist, aber das hat dann auch viel, hat dann nen größeren Bezug zur Realität, muss ich sagen. Also, man kann eigentlich damit ganz gut erklären, warum, wenn man ne Münze wirft, wie die Wahrscheinlichkeit liegt, dass Kopf oder Zahl geworfen wird. Also ich hab das jetzt zum Beispiel, vor allem wenn man die zinkt, dann ist das ganz interessant.“ (Int13/212ff.) R34

    „Mhm, naja, also die hat ja, die Binomialverteilung hat ja nur den Ausgang entweder, also wenn man das jetzt auf den Münzwurf bezieht, ähm, entweder Richtig oder Falsch. Und ähm, im Prinzip, wenns fair ist, 50-50 Chance, ähm, ja, bei der Münze, bei der Münze ist es halt Kopf oder Zahl.“ (Int13/221ff.)

    „Aber außer solchen Dingen, Ereignissen, ob Erfolg oder Misserfolg eintritt, kann man eigentlich nichts damit modellieren, also nur solche Sachen.“ (Int13/226f.)

    „[…] Ja, also […] ja, Null oder Eins, kann sie annehmen.“ (Int13/233) „Oh, beziehungsweise, das hängt davon ab, wie man sich das wählt, also, es kann auch, glaub ich, -1 und 1 sein. Aber 0 oder 1 finde ich sinnvoller." (Int13/235f.)

Der Studierende, der Bernoulli genannt hatte, versteht unter dem Begriff Bernoulli nicht die Person, sondern die Bernoulli-Verteilung, die er – im Rückgriff auf Schulwissen – zutreffend beschreiben kann, wobei er in der Erklärung auch Bernoulli-Ketten andeutet.

> „Ja. Bernoulli ist äh, das hat man auch schon in der Schule, also ein Experiment, wo ähm immer dieselbe Wahrscheinlichkeit auftritt und eigentlich nur zwei Möglichkeiten, ähm ja, Treffer oder Nicht-Treffer waren das immer, und ja, es gibt eben eine Trefferwahrscheinlichkeit und dann konnt man damit mehrstufige Experimente machen oder einstufige, wie auch immer." (Int10/143ff.) Frage

- Erwartungswert:

Der Begriff des Erwartungswertes wurde zweimal genannt. Im ersten Fall deutete der Studierende zunächst die formale Definition als Integral an, wobei er diese nicht näher ausführte. Bei dem Versuch der Übersetzung des Ansatzes in Alltagssprache beschränkte er sich anhand eines Beispiels auf den Erwartungswert der Binomialverteilung, wobei die Erläuterung entweder eine sprachliche Ungenauigkeit aufweist oder nicht korrekt ist (die Anzahl der Versuche wird mit „Vorkommnissen" bezeichnet). Zudem wird die Aussage des Gesetzes der großen Zahlen auf missverständliche Weise einbezogen.

> „Erwartungswert ist definiert als Integral, ja, was mich auch immer noch [?] irritiert, ist, warum die Stochastik über die Maßtheorie definiert ist, das kann ich mir nach wie vor nicht erklären, ja, warum Wahrscheinlichkeiten als Maße betrachtet werden." (Int7/146ff.) R12
>
> Auf Nachfrage, wie man das jemandem erklären könnte, der nichts mit Mathe zu tun hat: „Das ist ne gute Frage. Also beim Erwartungswert geht's ja schon los, ja, also das kann man dann natürlich nur plastisch erklären, anhand von Beispielen, am Würfelwurf oder so was, ja, und dann würd ich halt auch irgendwie sagen, also du nimmst jetzt nen Würfel und du würfelst den zweimal und wenn du Glück hast, kommt zweimal Sechs, ja, aber du kannst jetzt nicht davon ausgehen, dass immer Sechs kommt, weil wenn du eben ganz oft würfeln würdest, das ist dann das Gesetz der großen Zahlen, würdest du die Sechs so oft bekommen, wie im Wahrscheinlichkeitsmodell, ja, manifestiert ist, nämlich von der Anzahl ein Sechstel mal der Vorkommnisse. Und das würd ich dann probieren, so plastisch zu erklären, zumindest den Erwartungswert. Bei der Varianz, schwierig." (Int7/156ff)

Der zweite Studierende leitet die Bedeutung aus der Alltagsverwendung ab (was keineswegs bei allen Verteilungen zutreffend ist) und bezieht sich dann nur auf binomialverteilte Zufallsvariablen, indem er die Formel für den Erwartungswert bei dieser Verteilung ohne weitere Erläuterung nennt:

> „Ja, das sagt ja der Name schon. Ich erwarte einen bestimmten Wert. Das ist n mal p." (Int14/146) Frage

- Varianz und Kovarianz:

  Bei den Begriffen Varianz und Kovarianz lässt sich bei keinem der beiden Studierenden eine inhaltlich gefüllte Vorstellung der Bedeutung feststellen. Beide hatten eine Grundidee der Formel für die Varianz, konnten sie aber nicht vollständig formulieren. Bei der Kovarianz nannte der Studierende den Begriff der Abhängigkeit, konnte diesen aber nicht richtig einordnen.

  - „Ich würd sagen, der Erwartungswert minus irgendwas und dann die Summe darüber, ja, das ist schwierig.“ (Int7/167) R12
  - „Ähm, Varianz ist die quadratische Abweichung, Sigma Quadrat, […] boah ich erinner mich nicht mehr an Stochastik, irgendwas mit Abweichung und quadriert, quadratische Abweichung, ähm, Kovarianz, ist […] ich hab mich häufig gefragt, was Kovarianz eigentlich ist, Wikipedia hat keinen Eintrag dazu, ähm, heißt nur, glaub ich, ob irgendwelche Varianzen miteinander kollidieren, wie weit sie miteinander kollidieren […] irgendwas mit Abhängigkeiten war da, wenn man nur eine Zufallsvariable hat, dann ist Kovarianz gleich Varianz, hm, ja, so was.“ (Int14/138ff.) Frage

- Indikatorvariable:

  Bei der Erklärung der Indikatorvariablen wurde die Grundvorstellung einer Zufallsvariablen als Funktion deutlich, die in der Vorlesung eigentlich weniger im Vordergrund stand:

  „Ja, eine Funktion, die die Werte Null oder Eins annimmt im Prinzip, ja, und die ist immer Null für das Komplementärereignis und Eins für das Ereignis selber, ja.“ (Int8/105f.) R34

- Markowketten:

  Das hinter Markowketten stehende Konzept konnte von keinem der beiden Studierenden, die den Begriff genannt hatten, explizit formuliert werden. Während dem einen Studierenden keine Erklärung möglich war, nannte der andere zwar Grundideen, konnte diese aber nicht in einer inhaltlich korrekten und nachvollziehbaren Weise sortieren:

  - keine Erklärung möglich (Int12) R12
  - „Ähm, ja, Markovketten, da haben wir das Coole mit den Pfeilen gemacht und mit diesen Bereichen, ne?“ (Int15/153f.) Frage

    „Ähm […] das war immer ähm entweder kann man drin bleiben oder man kann mit ner Wahrscheinlichkeit in die anderen Bereiche gehen und dann startet man, ah genau, dann startet man immer wieder von dem Bereich, in dem man dann ist,startet man von neu. Also, man geht nicht zurück, sondern […] Kann ichs wirklich erklären, warscheinlich nicht [lacht]. Oh, ich hab dieses Bild vor Augen mit diesen ganz vielen Pfeilen, wo wir dann überlegt haben, ok, wie kam man jetzt dahin.“ (Int15/156ff.)

    „Ja, und man startet jedenfalls, man hat immer dieses neue Wahrscheinlichkeits, wie sagt man denn, ähm, Wahrscheinlichkeitsbereich oder so was, das hat nen Namen, ich weiß nicht, hab ich vergessen, ja.“ (Int15/163ff.)

> „Hm, ist das, also man kann natürlich wieder zurück fallen, je nachdem, was danach kommt, welches Ereignis, aber man startet, das ist ja ne Kette, man startet von dort, wo man hingelangt ist, also man hat am Anfang weiß ich nicht, wenn wir irgendwie äh macht das mit dem Würfeln Sinn, weiß ich gar nicht, also wenn wir würfeln und ich starte von einer Wahrscheinlichkeit und es gibt beispielsweise jetzt ganz einfach gerade und ungerade Zahl, und dann hab ich ne, und ich will beispielsweise äh grade, grade ungrade haben oder so und dann würfel ich das erste Mal und es ist grade, dann komm ich aufs nächste Feld, ok, grade und dann gehe ich von dort aus weiter. " (Int15/167ff.)
> „Und wenn ich dann ungrade würfel, geh ich dann aufs Ungrade oder wenn ich grade und so weiter, ja." (Int15/176f.)

- Approximationen:

  Bei den Ausführungen zum Begriff der Approximationen nannte der Studierende die Stirling-Approximation, ohne sie allerdings zu formulieren, und zeigte dann eine Begriffsverwirrung bezüglich Approximation und Modellierung:

  > „Stirling oder, also, ich fand interessant einfach zu sehen, dass nicht jede Aufgabe oder jede Thematik mit EINER Wahrscheinlichkeit abgearbeitet wird, so im Sinne von es gibt nicht nur DIE Wahrscheinlichkeit, sondern man hat verschiedene Modelle einfach. Ähm, dass es doch so komplex ist." (Int13/180ff.) R34

- Polya-Urne:

  Das Grundprinzip eines Polya-Urnen-Modells wurde zutreffend beschrieben, wobei der Studierende betonte, keine Anwendung dafür nennen zu können:

  > „Oh, ich versuchs, ich versuchs. Also, […] das ist eines dieser berühmten Modelle zum Berechnen von Wahrscheinlichkeiten. Mh, ich hätt mir jetzt ein einfacheres Beispiel aussuchen sollen [lacht]. Ähm, ist eigentlich peinlich, wenn ichs jetzt nicht wüsste, weil ich habs behandelt, ähm, also ich weiß zumindest, wie die funktioniert." (Int13/193ff.) „Wenn man, also man hat ja eine gewisse Anzahl von Kugeln in dieser Polya-Urne, die man von vornherein festlegt. Äh und ähm dann zieht man immer eine aus dieser Urne und je nachdem, was für eine Farbe es ist, legt man eine zu, eine Kugel, Extra-Kugel der gleichen Farbe wieder dazu und so äh ja, so kann man halt feststellen, ähm die Wahrscheinlichkeit, wie oft welche Kugel gezogen wird, würd ich jetzt sagen." (Int13/198ff.) „Ähm, aber wofür mans jetzt verwendet, wir hatten da so ein paar schöne Bildchen, aber […] ich hätte einfach was anderes nehmen sollen. [lacht]" (Int13/204f.) R34

- Hypothesentests:

  Bei der Erklärung von Hypothesentests wurde die Bedeutung aus dem Wort selbst abgeleitet, was zu einer unzutreffenden Erläuterung des Signifikanzniveaus führte und ausblendete, dass Hypothesentests auf der Berechnung bedingter Wahrscheinlichkeiten beruhen. Es wurde auch nicht erläutert, auf welcher Basis die „Überprüfung" der Hypothese erfolgt.

  > Man stellt eine Hypothese auf, ja, und will mit irgendeiner Wahrscheinlichkeit überprüfen, ob die stimmt und dann prüft man das und dann kann man mit einer gewissen Wahrscheinlichkeit, Frage

je nachdem, was man halt vorher gesagt hat, 95% oder was weiß ich, kann man sagen, ob die Hypothese wahrscheinlich stimmt oder nicht. (Int14/129ff.)

Fragenkomplex 9A/10B: Was verstehst Du unter „stochastischer Intuition"? Gibt es einen Unterschied zu „Verständnis von Stochastik"?

Bemerkenswert ist, dass insgesamt acht der Befragten im Laufe des Interviews entweder betonten, selbst keine gute stochastische Intuition zu haben (Int1/131, Int2/132, Int4/265f., Int7/181ff., Int11/126, Int14/114) oder dass sie vielen Menschen entweder fehle bzw. viele zumindest anzweifelten, eine gute zu haben (Int6/97ff., Int15/182).

Der Begriff **Intuition** wurde von 12 Studierenden mit einem inneren Gefühl bzw. einer unmittelbaren Erwartung in Verbindung gebracht, was die Lösung einer Aufgabe ist, wobei teilweise auch eine mögliche Korrekturfunktion der Intuition angedeutet wurde:

- „Also, stochastische Intuition, haben wir ja auch in diesem einen Stochastik-Kurs mal drüber gesprochen, ähm, ist ja eigentlich so dieses innere Gefühl, was da irgendwie bei rauskommt." (Int1/129ff.)
- „[…] Also dass ich jetzt intuitiv ne Idee hätte, wie sich was verhält […]. Und dass ich jetzt intuitiv sagen könnte, das ist eher wahrscheinlich oder eher unwahrscheinlich." (Int2/132ff.)
- „[…] und intuitiv wär dann für mich, dass ich sage, ich würde aber erwarten, dass irgendwie so was rauskommen würde, dass die Wahrscheinlichkeit null ist, dass es überhaupt nicht eintritt, oder so." (Int2/148ff.)
- „[…] Intuition versteh ich halt darin, dass man ähm sich das halt anguckt und sagt, ok, es ist jetzt nur 5 Mal Kopf und 15 Mal Zahl gefallen, würd ich mal sagen, ja, grenzwertig, also, kann mal passieren, aber ich weiß nicht, ob das wirklich so oft passiert. Aber wenn ich jetzt z.B. 11 und 9 sehe oder 12 und 8, dann würd ich schon sagen, ja, kann gut sein, also das ist so Intuition, dass ich sag, könnt hinkommen." (Int3/114ff.)
- „[…] was man sich so denkt, was rauskommen müsste […]" (Int4/266)
- „Also, ich versteh darunter, ähm, also wenn ich jetzt irgendein Experiment hab, ob ich mir dann intuitiv vorstellen kann, wie groß die Wahrscheinlichkeiten für irgendwelche Ausgänge sind, ob das wahrscheinlich ist, ob das so funktioniert." (Int5/187ff.)
- „Ähm, stochastische Intuition. Ob man, wenn man irgendne Aufgabe sieht, meistens sind so stochastische Aufgaben, denk ich, so, dass man nicht das tippen würde, was am Ende dann raus kommt. Und mit Intuition ist dann vielleicht, wenn mans doch richtig tippt." (Int6/97ff.)
- „[…] stochastische Intuition ist für mich ein Gefühl dafür, Wahrscheinlichkeiten zu verstehen und auch richtig einzuschätzen […]" (Int7/173f.)
- „Ja, ok. Also äh stochastische Intui, also stochastische Intuition ähm, also wenn man jetzt irgendne Aufgabe bekommt, die halt äh nen stochastischen Inhalt hat, und man soll jetzt spontan ohne zu rechnen irgendwie sagen ähm, was wäre dann da so denkbar, was die Lösung wohl, also in welche Richtung geht die Lösung vielleicht. Und stochastische Intuition wär dann eben irgendwo zu wissen, ok, welches Modell trifft darauf zu und ähm ja, dann kann man eben, es ist ja auch so ein bißchen, für stochastische Intuition muss man ja nicht unbedingt ähm also zumindest für einfachere Sachen muss man ja nicht unbedingt jetzt Mathematik studiert haben. […] Stochastische Intuition ist für mich auch irgendwie so ein bißchen Logik mit drin, ja." (Int10/153ff.)
- „Ja im Grunde ist das, das richtige Modell zu finden für äh ja, irgendne Problemstellung." (Int11/145f.) und: „[…] ich denke, man kann auch das richtige Modell auch finden, ohne es hundertprozentig verstanden zu haben." (Int11/150f.)
- „Ja, das ist […] ähm, also jetzt allein von dieser Wortkombination her würd ich sagen, jemand, der sich damit zum Teil beschäftigt hat, vor allem mit diesen Themen, und dann einfach ohne irgendwelche Nachrechnungen oder ähm Aufschriebe ungefähr sagen kann, was für ein Ergebnis

rauskommt.“ (Int13/239ff.)
„Oder was für eine Wahrscheinlichkeit dieses Modell oder dieser Versuch hat. Würd ich jetzt sagen, das ist stochastische Intuition.“ (Int13/244f.)

- „[…]dass man intuitiv die richtige Antwort hat.“ (Int14/156)
- „[…] würd ich zum Beispiel bezeichnen ähm mit der Fähigkeit etwas einschätzen zu können, obs wahrscheinlich ist oder nicht, also dass wenn irgendjemand zu mir kommt und sagt ähm weiß ich nicht, wenn Sie jetzt Lotto spielen, dann äh ist es wesentlich wahrscheinlicher, dass Sie gewinnen, weil blablabla, dass ich dann die Fähigkeit habe, ja, intuitiv zu sagen, ja ok, es könnte sein oder irgendwie zu sagen ne, das glaub ich nicht.“ (Int15/183ff.)
- „Oder dass beispielsweise auch so Erwartungswerte oder so was, dass irgendjemand sagt, ja der Erwartungswert, man erwartet das, dass ich mir überlegen kann, hm, stimmt das jetzt oder stimmt das nicht. Oder dass ich mir selbst überlege, was erwarte ich denn eigentlich, dass es kommt. Dass man so ein bißchen ein Gefühl dafür hat, wie wahrscheinlich oder wie oft Dinge vorkommen oder nicht vorkommen.“ (Int15/191ff.)

Außerdem wurde die Intuition auch als unmittelbare Idee zur Herangehensweise an eine Problemstellung beschrieben bzw. weitergehend als innere Einsicht in die Zusammenhänge:

- „[…] die Art und Weise, wie du an eine Aufgabe ran gehst und wie schnell du da auf ne Lösung kommst oder auf den richtigen Weg kommst […]“ (Int8/120f.)
- „Ja, also, das Rechnen an sich hilft noch nicht dafür, also es ist schon so, also es ist dann so, ich hab das gerechnet, ich hab das richtig gemacht und die Rechenregel ist mir auch klar, aber es bleibt immer noch die Verwunderung, krass, warum ist das so, also das Warum ist nicht beantwortet […]“ (Int1/152ff.)

Zu Möglichkeiten, die Intuition zu verändern bzw. zu verbessern, äußerten sich vier der Befragten. Sie betonten, dass Erfahrung bzw. Ausprobieren entscheidend sei, nicht aber bloßes Nachrechnen:

- „[…] und so diese innere Intuition, die man hat, die lässt sich nicht davon wegbringen, obwohl ichs gerechnet hab und weiß es geht so. Also, es bleibt so ein fahles Gefühl […]“ (Int1/142ff.)
- „[…] und wenns dann irgendwie durch Ausprobieren so ist, dann werd ich bestätigt, aber andererseits lernt man ja auch in anderen Situationen, so paarmal ausprobieren bestätigt mathematisch noch gar nichts. Klar, wenn man das mit den Simulationen macht, klar, wenn man bei 1000, 10000, immer strebt es gegen irgendeinen Wert, dann macht es schon Sinn, dass es so ist, was ja auch beim Würfel eigentlich so ist, ja, also ist ja gar nicht die Frage. Aber trotzdem, ich weiß nicht, was man für die Intuition, also Intuition ist auf jeden Fall nochmal durch dieses Ausprobieren, das stärkt sie.“ (Int1/156ff.)
- „[…] ich würd am Anfang niemandem raten, sich zu sehr darauf verlassen, wenn er nicht grad wirklich extrem, ja extreme Erfahrung in Stochastik hat.“ (Int4/308ff.)
- „[…] je mehr man hört, desto einfacher fällt es einem, ja wenn man so eine Intuition entwickelt, wie man das machen kann.“ (Int8/112f.)
- „[…]anstatt das so alles in formales Theorie zu lernen, dann lernt man so, ich denke von, von das Leben oder was, von experience, ne. Ja, man sagt, von experience ich habe das schon einmal erlebt und, ja.“ (Int9/104ff.)

Eine weitere Eigenschaft, die der Intuition zugeordnet wurde, war, dass sie ohne logische Begründungen auskomme, und somit aber auch – im Gegensatz zum Verständnis – fehleranfällig sei:

- „[…] ja, aber das ist SO fehleranfällig, da kann ich mich nicht drauf verlassen. Und unbegründet.“ (Int2/142f.)

- „[…] das ist ja im Endeffekt äh sag ich ja damit nichts, ja. Es ist ja für mich persönlich die Intuition, andere sehen das vielleicht anders, die sagen, ach das könnt doch so und so sein, ja.“ (Int3/123ff.)
- „[…] diese Intuition in der Stochastik ist sehr, sehr irreführend und klappt sehr, sehr wenig.“ (Int4/270f.)
- „Ich denk mal die, naja, normalerweise es kommt ja auch öfters vor, dass man sich was vorstellt, was rauskommen könnte und tatsächlich kommt aber was ganz anderes raus. Und ich denke mal, wenn man das Verständnis von Stochastik hat, dann kommt man eher zum richtigen Ergebnis und mit der Intuition könnt man auch mal falsch liegen.“ (Int5/197ff.)
- Auch die folgende Beschreibung könnte darauf hindeuten: „Überdimensionales Denken.“ (Int12/110)

In der Frage, ob eine gute Intuition das Verständnis der stochastischen Zusammenhänge voraussetzt, waren die Befragten dagegen unterschiedlicher Meinung:

- „Ja, also ich glaub, die Intuition braucht das Verständnis. Ich glaub, je mehr Verständnis man hat, desto besser wird die Intuition.“ (Int4/332f.)
- „Also, ich find Verständnis an sich ist, glaub ich, das grundlegende Prinzip und das muss man erstmal hinkriegen, um überhaupt irgendwas oberhalb von diesem Begriff äh zu machen […].“ (Int13/247ff.)
- „Intuition hat ja nichts mit Verständnis zu tun. Intuition kann ja auch kommen, wenn man gar kein Verständnis für irgendwas hat. Man kann ja intuitiv auf was tippen.“ (Int14/158ff.)

Der Begriff **Verständnis** wurde am häufigsten, nämlich von insgesamt sechs Befragten, mit Nachvollziehbarkeit und logischen Schlussfolgerungen bzw. der Fähigkeit zur Versprachlichung in Verbindung gebracht:

- „Ähm, ich, würd glaub ich, also bei mir sind das irgendwie zwei verschiedene Ebenen, einmal diese Formelsprache, also dass ich irgendwie sage, also wenn das so ist dann ist das so, wenn ich dann mal irgendwie durchgeblickt hab […]“ (Int2/146ff.)
- „Ne, ja, also, dass ichs nachvollziehen kann […]“ (Int2/154)
- „Und äh das ist natürlich ein großer Unterschied, ob ich jetzt wirklich Fakten vor mir habe oder irgendwie von mir aus denke, ach, so oder so.“ (Int3/125ff.)
- „Also dass ich diesen Beweis eben auch nachvollziehen kann, sag ich mal. Das ist für mich eigentlich das Verständnis.“ (Int3/131f.)
- „Weil Verständnis würd ich eher so mathematisch, formelmäßig verstehen, so.“ (Int6/103f.)
- „Äh, die ganzen Sätze und Herleitungen, das versteh ich unter, wenn mans tatsächlich verstanden hat.“ (Int6/106f.)
- „Wenn mans selber erklären kann.“ (Int10/175)
- „[…] das Verständnis ist, denk ich, wenn man das wirklich äh alles auch korrekt beweisen kann […]“ (Int11/167f.)
- „[…] dass man versteht, woher das kommt.“ (Int15/218)
- „Oder auch argumentativ verwenden kann […]“ (Int15/231)

Ein weiterer Bezug, der von drei Studierenden formuliert wurde, war das korrekte Anwenden von Rechenregeln, wobei einer auch die Einsicht in die Zusammenhänge, die hinter den Rechnungen stehen, für erforderlich hielt:

- „[…] also die Rechenregeln sind mir klar, also würde ich sagen, ich habe ein Verständnis dafür, wie ichs rechne, äh, mir macht es aber trotzdem nicht bewusst, warum bei diesem Verständnis, das ich ja eigentlich hab, so etwas rauskommt […]“ (Int1/138ff.)
- Und diese Erkenntnis, ich würde sagen, dann hat man Stochastik verstanden oder so die Grundzüge von der Stochastik verstanden, wenn man nicht mehr denkt, dass alles einfach irgendwie passiert, sondern dass wirklich, dass man das wirklich teilweise berechnen kann.“ (Int15/226ff.)
- „Also, ich kann ja quasi Beweise oder Rechenschritte, Herleitungen einfach verstehen, mathematisch, ohne jetzt sag ich mal den so großen Sinn dahinter zu verstehen. […] stochastisches Verständnis heißt eben für mich auch, dass man diesen Aha-Effekt dahinter mitbekommt […]“ (Int4/315ff./322f.)

Insgesamt drei Befragte verbanden den Begriff Verständnis mit der Fähigkeit, Übungsaufgaben lösen zu können. Für sie bedeutete Verständnis also explizit auch das Anwendenkönnen theoretischer Erkenntnisse auf konkrete Sachverhalte:

- „Ähm, also vom Stoff her geht’s immer relativ schnell, aber ich glaub so wirklich dieses tiefgehende Verständnis, das kommt erst mit den Übungsaufgaben […]“ (Int8/126ff.)
- „Wenn ich mir das im Skript durchlese, wenn ich die Übung dazu mache, dann hab ich das verstanden.“ (Int12/120f.)
- „Wahrscheinlich wenn man beides besitzt, wenn man sowohl die Aufgabe versteht, als auch intuitiv, also wenn man intuitiv schon mal weiß, wo ungefähr, wo's lang gehen muss, und dann das Verständnis dafür hat, um die Aufgabe lösen zu können.“ (Int14/162ff.)
  „Hm […] vielleicht bräuchte man da noch ein bißchen Wissen dazu.“ (Int14/166)

Drei Befragte äußerten sich dazu, wann man denn ein Verständnis entwickelt habe. Einer von ihnen brachte Verständnis in diesem Kontext – wie es auch mehrere beim Begriff der Intuition getan hatten – mit Erfahrung in Zusammenhang. Zwei formulierten, dass Verständnis nie umfassend werden könne:

- „Das ist Erfahrungssache.“ (Int8/134)
- Ähm, […] ich persönlich würde sagen, man hats eigentlich, also bei den einfachen Dingen ist es relativ schnell, hat mans verstanden, aber bei so komplexeren würd ich sagen vor allem, wenn man das dann ja auf die Allgemeinheit bezieht, hat mans eigentlich nie ganz verstanden.“ (Int13/253ff.)
- Antwort auf die Frage, wann man in Stochastik etwas richtig gut verstanden hat: „Ich glaube, nie.“ (Int7/207)

Für einen Befragten stand im Vordergrund, dass stochastisches Verständnis immer abstrakt bleibe, es nicht die tatsächliche Welt abbilde:

- „Und was auch super wichtig ist, was ich auch immer merk, ist irgendwie, die Stochastik ist wirklich ne reine Kopfsache, ja, weil wenn ich so nen Würfel werfe, dann hab ich ne Modell-Denke, alles passiert, also jede Augenzahl erscheint mir ein Sechstel Wahrscheinlichkeit, aber dann lass ich ja total die physikalischen Gegebenheiten raus […]“ (Int7/193ff.)

Eine explizite Nachfrage, ob Simulationen bei der Entwicklung stochastischen Verständnisses helfen könnten, wurde bejaht. Ansonsten nahm keiner der Befragten Bezug zu Simulationen.

- „Hm, ja, wäre ein Baustein. Also ich möcht jetzt nicht sagen, hätt ich die Simulationen gehabt, hätt ichs sofort verstanden, aber ich denk es wär ein hilfreicher, oder es IST ein hilfreicher Baustein, ja.“ (Int2/163ff.)

Fragenkomplex 10A/11B: Möchtest Du noch etwas anmerken?

Die Möglichkeit, weitere Anmerkungen zu tätigen, nutzten insgesamt zehn Befragte, wobei sich drei Befragte zum Programmieren an sich äußerten, die alle eine positive Einstellung diesbezüglich hatten (Int8/185f., Int10/229, Int5/287f.). Zwei betonten aber auch, dass es nötig sei, sich darauf einzulassen und auch ausreichend mitzuarbeiten, da man sonst schnell den Anschluss verliere (Int5/285ff., Int8/186ff.). Ein Befragter führte aus, dass diesbezüglich auch die Verwendung desselben Programms in anderen Veranstaltungen, etwa in einem Proseminar, eine hilfreiche Ergänzung gewesen sei (Int10/230ff.).

Drei Befragte schätzten die „R Aufgaben" ausdrücklich dafür, dass sie einen Einstieg in das Programm geboten hätten (Int5/281ff., Int9/159ff., Int15/340ff./347ff.), wobei zwei von ihnen angaben, diesen nun selbstständig auszubauen bzw. für eigene Projekte zu nutzen (Int5/281ff., Int9/159ff./165f./168/170f.). Eine andere befragte Person formulierte, dass sie Simulationen an sich auch für Lehramts-Studierende mit Blick auf ihren späteren Beruf als sinnvoll erachte, das Programm R für dieses Anwendungsfeld jedoch zu schwierig sei (Int14/218ff./224ff./244ff.).

Zwei Studierende äußerten sich zu den Unterstützungsangeboten, wobei deutlich wurde, dass entweder die Angebote von Fragestunden unbedingt beibehalten werden und für individuelle Fragen sogar noch mehr Ansprechpartner zur Verfügung stehen sollten (Int3/160ff., Int10/233f.), oder eventuell die Aufgaben aber auch einfacher gestaltet werden könnten, so dass weniger Hilfe nötig würde (Int3/157ff.).

Drei Personen formulierten allgemeine Kommentare zum Aufbau und Struktur der Vorlesung sowie zur Klausur (Int11/242/246ff./260ff./264f./269ff., Int13/395ff./406, Int14/249/251/257ff./287ff.), wobei ein Befragter explizit den Einsatz von Simulationen auch in der Vorlesung lobte (Int11/252). In diesem Zusammenhang ist noch der Kommentar einer anderen befragten Person bemerkenswert, dass sie Stochastik für ein spannendes, aber unmathematisches Fach halte, weil es weniger Schemata gebe als in anderen Fächern, dafür ein spielerisches Moment im Vordergrund stehe, man stärker mitdenken müsse und es viel stärker um Modellierungen gehe als in anderen Fächern (Int7/303ff.). Auch eine weitere Person beschrieb das Fach in einem anderen Kontext als unmathematisch und brachte eine ähnliche Argumentation vor, nämlich dass die Vorgehensweise weniger schematisch als in anderen Gebieten sei und für sie auch wesentlich abstrakter und unzugänglicher (Int2/107ff.).

Gespräch über Simulationsplan für Krankenhausaufgabe

Am Ende des Interviews wurde den Studierenden, die angegeben hatten, sich mit den „R Aufgaben" häufig beschäftigt zu haben[147], die Krankenhausaufgabe aus den Fragebögen vorgelegt mit der Bitte, für den darin beschriebenen Sachverhalt einen Simulationsplan zu formulieren. Ziel war es insbesondere herauszufinden, ob bei den Studierenden die Einsicht, dass für die Auswertung stochastischer Simulationen erst eine häufige Wiederholung zu belastbaren Ergebnissen führt, verankert werden konnte.

Die Struktur der Simulation des Krankenhaus-Sachverhalts für einen Tag (Ziehen von 45 bzw. 15 Bernoulli-verteilten Zufallsvariablen mit Trefferwahrscheinlichkeit 0,5 bzw. uniform auf {0;1} verteilten Zufallsvariablen, Berechnen des Prozentsatzes an Treffern, am einfachsten durch Summieren der Zufallsvariablen (sofern der Wert 1 ein Mädchen abbildet) und Dividieren durch die Anzahl 45 bzw. 15, und Überprüfung, ob ein größerer Wert als 0,6 erreicht wurde) beschrieben zwei der 12 dazu befragten Studierenden sehr detailliert (Int4/369ff./384ff./388/393ff. und Int15/255ff./266ff.) und sieben weitere gröber, aber so, dass die Schritte abgeleitet werden konnten (Int5/225f./228f./231, Int6/123ff., Int8/152ff./156, Int10/205ff., Int11/193ff./199f./202f./205, Int14/197ff.; mit Hilfe: Int13/275f./313ff./323ff./ 327/329/331f.). Die restlichen drei Befragten konnten das nicht leisten oder lieferten eine Beschreibung, die die Schritte nicht erkennen ließ (Int3/141ff./145ff., Int7/243ff./274ff., Int9/133ff./141ff.).

Keine der neun Personen, die einen Simulationsplan aufstellen konnten, schlugen eine Simulation von jeweils nur einem Tag pro Krankenhaus vor. Um ein Jahr zu modellieren, wäre eine 365fache Wiederholung des oben beschriebenen Vorgehens nötig, bei denen jeweils gezählt werden sollte, wie oft ein Prozentsatz an Treffern, der größer als 60% ist, erreicht wurde. Das beschrieben sieben der 12 Studierenden (Int4/375f./378ff./390ff./397ff., Int5/235f./240, Int8/156f., Int10/207ff.[148], Int11/211f./214, Int13/323ff., Int15/280ff./287ff./298ff./306ff./311ff.). Zwei von ihnen würden die Simulation auch bei einem Jahr belassen. Dass für eine zuverlässige Antwort auf die Frage eine sehr viel größere Anzahl von Wiederholungen erforderlich ist, formulierten insgesamt sieben der 12 Befragten, allerdings teilweise nur auf Nachfrage. Zwei Befragte würden direkt eine größere Anzahl als 365 Wiederholungen durchführen, was natürlich ebenfalls zum richtigen Ergebnis führen würde (eigenständig: Int6/131/133ff./140/142, Int10/220ff., Int14/201ff.; auf Nachfrage: Int4/413ff/424ff., Int11/229f./232/234/236ff., Int13/339f./342f./347ff., Int15/330f./333ff.).

[147] Int7 wurde trotz der Zuordnung zu R12 hinzugenommen, da die Person angegeben hatte, zumindest die ersten vier Aufgaben selbst bearbeitet zu haben.

[148] Die richtige Grundidee des Zählens war vorhanden, die Umsetzung dann aber falsch beschrieben.

III. Diskussion der Ergebnisse und Ansätze für weiterführende Forschung

III.A. Diskussion der Ergebnisse bezüglich der stochastischen Fertigkeiten und Wahl des Bearbeitungsmodus von Simulationsaufgaben

In der **Vorstudie** wurden die in der Literatur beschriebenen Beobachtungen zu hartnäckigen Fehlvorstellungen überwiegend bestätigt. Die Fehlvorstellung, dass die Variabilität des Mittelwerts nicht von der Stichprobengröße abhängig wäre (F2(b)), konnte durch die theoretische Beschäftigung in der Vorlesung und den Übungsgruppen trotz intuitiven Zugangs über die Zufallsvariablen nicht ausgeräumt werden, wie sich in der *Krankenhaus- bzw. Grundschulaufgabe* zeigte. Wie erwartet, wählte die Mehrheit die zu der Fehlvorstellung F2 passende Antwort c), nur jeweils etwa ein Drittel der Befragten gab die richtige Antwort. Die Tatsache, dass am Ende des Semesters bei falscher Lösung im Mittel eine niedrigere Sicherheit angegeben wurde als bei richtiger, könnte darauf hindeuten, dass zumindest Zweifel an der Fehlvorstellung gestreut werden konnten, die aber noch nicht zu einer Korrektur geführt haben. Musste der Einfluss der Stichprobengröße auf die Genauigkeit einer Schätzung dagegen im Kontext von Häufigkeitsverteilungen erkannt werden (F2(a)), lieferte ein Großteil der Befragten das richtige Ergebnis (*Post-Aufgabe*). Darin kann eine Bestätigung der von Sedlmeier und Gigerenzer empfohlenen Unterscheidung zwischen den Situationen der Kennwerte- und Häufigkeitsverteilungen gesehen werden. Die beschriebene Fehlvorstellung bezieht sich bei einer großen Mehrheit der Personen lediglich auf den ersten Fall. Allerdings scheint die richtige Grundvorstellung, dass bei einer größeren Stichprobe eine genauere Schätzung erwartet werden kann, nicht ohne Weiteres auf komplexere Situationen übertragbar zu sein, wie sich an den Ergebnissen der *Multiple-Choice-Aufgabe* zeigt, die deutlich schlechter sind als bei der *Post-Aufgabe*. Es wurden bei falschen Ergebnissen zwar niedrigere Sicherheiten angegeben, dennoch löste nur etwa ein Drittel die Aufgabe richtig. Es bieten sich zwei mögliche Interpretationen für diesen Sachverhalt an: Entweder könnte die Grundvorstellung für einen Großteil der Befragten nicht mehr abrufbar sein, sobald der Zusammenhang weniger explizit in der Aufgabenstellung adressiert wird. Oder die Schwierigkeit bestand darin, dass der Inhalt der Aufgabe selbst von einem nicht zu vernachlässigenden Anteil der Studierenden nicht richtig erfasst wurde. Für die zweite Möglichkeit spricht, dass von den falschen Antworten nur etwa die Hälfte der Fehlvorstellung F2(a) zuzuordnen ist. Der Anteil derer, die dieser Fehlvorstellung unterlagen, unterscheidet sich damit gar nicht wesentlich vom entsprechenden Anteil bei der *Post-Aufgabe*.

Bezüglich der Fehlvorstellung der lokalen Repräsentativität (F1) zeigt sich im Kontext statistischer Fragestellungen ein ähnliches Bild: Dass es keinen auffälligen Unterschied in der Beantwortung der *Aufgabe zum Erkennen möglicherweise gefälschter Münzwürfe* zwischen dem Beginn und dem Ende des Semesters gab, deutet auch hier darauf hin, dass die Fehlvorstellung weiterhin vorhanden war. Das einzige Kriterium, das den Angaben in den Freitextfeldern zur Begründung nach für ein Anzweifeln der Zufälligkeit herangezogen wurde, war das numerische Verhältnis von Nullen und Einsen. Viele Wechsel, also kurze Runs, gaben dagegen keinen Anlass dafür, von einer Fälschung auszugehen, wie sich an der zweiten Münzwurfreihe in der Aufgabe zeigte. Wurde die Fehlvorstellung jedoch noch expliziter adressiert, indem in einem Vergleich zweier Reihen danach gefragt wurde, welche Abfolge wahrscheinlicher wäre (*Aufgabe zur Wahrscheinlichkeit von Münzwurfergebnissen*), zeigte sich die beschriebene Fehlvorstellung schon zu Beginn des Semesters bei den Probanden nicht. Ebenso bestätigte sich die Beobachtung von Biehler und Maxara, dass bei den Studierenden die Vorstellung, stochastische Fragestellungen mit Hilfe von Experimenten bzw. Simulationen zu untersuchen, fehle, ebenfalls nicht (*Aufgabe zum Testen eines Würfels*). Ob diese Resultate darin begründet sein könnten, dass es sich um Mathematik-Studierende mit dem Ziel Bachelor bzw. Gymnasiallehramt handelte und sie bereits umfassendere stochastische Vorkenntnisse mitbrachten als die Studierenden, die von Biehler und Maxara befragt wurden, lässt sich nicht abschließend beurteilen.

Die Einschätzung der Studierenden bezüglich ihrer Sicherheiten, dass die Aufgaben richtig gelöst wurden, war zu Beginn des Semesters bei fast allen Aufgaben im Mittel unzuverlässig (mit Ausnahme der *Aufgabe zum Testen eines Würfels* konnte kein auffälliger Unterschied zwischen den Angaben bei richtiger und falscher Lösung beobachtet werden). Am Ende des Semesters waren in zwei von drei Aufgaben die angegebenen Sicherheiten für richtige Lösungen im Mittel höher als für falsche, was für eine Verbesserung der Selbsteinschätzung spricht.

In der **Hauptstudie** zeigte sich zu Beginn des Semesters ein vergleichbares Bild, so dass beide Jahrgänge sinnvoll verglichen werden können. Auch hier ließen sich zu Beginn des Semesters kaum Unterschiede zwischen den Sicherheiten für falsche und für richtige Antworten beobachten, was darauf hindeutet, dass die Befragten ihre Fertigkeiten eher schlecht einschätzen konnten. Anders als es sich in der Vorstudie abzeichnete, konnte in der Hauptstudie keine deutliche Verbesserung dieser Selbsteinschätzung im Laufe des Semesters beobachtet werden, bei der *Grundschule-Aufgabe* gab es im Mittel in keiner Gruppe

Unterschiede bei den angegebenen Sicherheiten zwischen richtiger und falscher Lösung, bei der *Multiple-Choice-Test-Aufgabe* nur bei den Fragestundenbesuchern.

Die Wahl des Bearbeitungsmodus der „R Aufgaben" scheint nicht zufällig gewesen zu sein: Auffällig ist, dass die Gruppe der Fragestundenbesucher zu Beginn des Semesters bei allen Aufgaben signifikant schlechtere Ergebnisse zeigte als mindestens eine der beiden anderen Gruppen. Im Fall der *Aufgabe zum Testen eines Würfels* und der *Krankenhausaufgabe* zeichneten sich die zukünftigen Fragestundenbesucher zusätzlich dadurch aus, dass sie ihren richtigen Lösungen im Mittel weniger trauten als die anderen Studierenden. Daraus lässt sich schließen, dass das Angebot der Fragestunde hauptsächlich für die leistungsschwächeren und tendenziell unsichereren Studierenden attraktiv war. Ob ein Studierender die „R Aufgaben" oft oder immer eigenständig (R34) oder selten bzw. gar nicht bearbeitete (R12), scheint dagegen nicht mit den stochastischen Fertigkeiten zusammenzuhängen.

Untersucht man die Entwicklungen innerhalb des Semesters, dann zeigt die Gruppe der Fragestundenbesucher die meisten Veränderungen. Bezüglich der Fehlvorstellung, dass die Variabilität des Mittelwertes nicht von der Stichprobengröße abhänge (F2(b)), lässt sich eine Verbesserung feststellen, wenn man das individuelle Antwortverhalten im Fall der *Krankenhaus-Aufgabe* mit dem im Fall der *Grundschule-Aufgabe* vergleicht. Hierbei muss allerdings bemerkt werden, dass in dieser Gruppe das Verbesserungspotenzial auch am größten war, da zu Beginn des Semesters signifikant weniger richtige Antworten zur *Krankenhaus-Aufgabe* gegeben wurden als in den anderen Gruppen. Trotz der nachweisbaren Verbesserung auf individueller Ebene ist die Fehlvorstellung in dieser Gruppe aber auch am Ende des Semesters noch signifikant häufiger vertreten als in den anderen beiden Gruppen. Außerdem hatten die Fragestundenbesucher im Mittel auch ein signifikant niedrigeres Zutrauen bei richtiger Lösung als es in den anderen beiden Gruppen der Fall war. In der Gruppe derer, die die „R Aufgaben" oft oder immer eigenständig gelöst hatten (R34), ließ sich genauso wie in der Gruppe, die die „R Aufgaben" selten oder nie bearbeitet hatten (R12), keine Veränderung beobachten. Insgesamt führte also die Beschäftigung mit den Simulationsaufgaben trotz der stetigen impliziten Adressierung der Fehlvorstellung F2 nicht zu einer deutlichen Verminderung dieser Fehlvorstellung.

Bemerkenswert ist an dieser Stelle das Ergebnis des Befragungsteils aus den Interviews, in dem die Studierenden aufgefordert wurden, einen Simulationsplan für die *Krankenhaus-Aufgabe* zu entwerfen (vgl. den letzten Abschnitt von Kap. II.C.5.). Die Fragestundenbesucher zeigten hier auffallend gute Leistungen. Ob die in der Fragestunde trainierte Fähigkeit, die Struktur von Aufgaben zu versprachlichen und in Programmiercode zu übersetzen, diesen Erfolg erklärt, oder es aufgrund der geringen Anzahl an Befragten aus dieser Gruppe ein zufälliges Ergebnis ist, kann an dieser Stelle nicht abschließend beurteilt werden. Hier wären weitere Untersuchungen nötig. Bezüglich der in der Aufgabe adressierten Fehlvorstellung lässt sich folgendes festhalten: Alle Fragestundenbesucher und mehr als die Hälfte derer, die die „R Aufgaben" oft oder immer selbstständig bearbeitet hatten, konnten einen korrekten Simulationsplan aufstellen, hatten also die Struktur der Aufgabe verstanden. Lediglich zwei von diesen Personen schlugen nur die Simulation eines Jahres vor, allen anderen war klar, dass eine häufige Wiederholung zu einem genaueren Ergebnis der Simulation führte – unabhängig davon, ob sie für die Aufgabe die richtige oder eine falsche Lösung angegeben hatten. Der Transfer dieser Erkenntnis für Simulationen auf die vorgelegte Aufgabe war aber offensichtlich so schwierig, dass er nicht allen gelang. Die Variabilität von Schätzungen sollte daher vermutlich in den „R Aufgaben" noch expliziter thematisiert werden. Allein eine Beschäftigung mit dieser Thematik in Form des Experimentierens mit unterschiedlichen Wiederholungsanzahlen von Simulationen scheint nicht auszureichen, um die Fehlvorstellung, Kennwerteverteilungen seien nicht von den Stichprobengrößen abhängig, abzubauen. Wie genau diese explizitere Thematisierung umzusetzen wäre, könnte ein Ansatzpunkt für weiterführende Forschung sein.

Bei der *Multiple-Choice-Aufgabe* konnte kein Unterschied zwischen dem Antwortverhalten der Fragestundenbesucher und denen, die die „R Aufgaben" selten oder nie bearbeitet hatten (R12), festgestellt werden. Die Fragestundenbesucher scheinen hier also eine Angleichung erreicht zu haben, da sie zu Beginn des Semesters ja bei allen Aufgaben signifikant schlechter abgeschnitten hatten als die anderen beiden Gruppen. Dies könnte auf eine Verbesserung ihrer Fertigkeiten hindeuten. Eine Angleichung an die Gruppe derer, die die „R Aufgaben" immer oder oft selbstständig bearbeitet hatte (R34), gelang dagegen nicht. Ein möglicher Interpretationsansatz könnte sein, dass die Bearbeitung der „R Aufgaben" eine gewisse Verbesserung der stochastischen Fertigkeiten, die sich auf die Fehlvorstellung F2(a) beziehen, ermöglicht – wovon die Gruppe derer, die die Simulationsaufgaben selten oder nie bearbeitet hatte, folgerichtig nicht profitieren konnten. Allerdings ist bei dieser Interpretation insofern Vorsicht geboten, als dass auf der anderen Seite kein signifikanter Unterschied zwischen den Lösungen der Gruppen R12 und R34 gefunden werden konnte – falls es einen Verbesserungseffekt geben sollte, kann er also nicht allzu groß sein. Die Ergebnisse der Multiple-Choice-Aufgabe sind aber auch in der Hauptstudie ohnehin schwierig zu interpretieren, da auch hier – ähnlich wie in der Vorstudie – ein nicht unwesentlicher Teil der

Befragten den Inhalt der Aufgabe nicht vollständig erfasst haben könnte. Hierauf deutet der nicht zu vernachlässigende Anteil der falschen Antwort b) hin, der nicht zu der Fehlvorstellung F2(a) passt.

Bei der *Aufgabe zur Erkennung möglicherweise gefälschter Münzwürfe* zeigte sich in der Gruppe der Fragestundenbesucher dagegen eine deutliche Verbesserung, am auffälligsten allerdings bei der Reihe, die aufgrund der Anzahlen von Nullen und Einsen verdächtig erscheint (Reihe 3). Bei der zweiten Reihe, die auf die Fehlvorstellung der lokalen Repräsentativität zielt (F1), zeigte sich bei allen Gruppen eine Verbesserung im individuellen Antwortverhalten, also auch bei den Fragestundenbesuchern. Bei den Begründungen in den Freitextfeldern fand sich in dieser Gruppe am Ende des Semesters auch die Anzahl der Runs als Kriterium, hauptsächlich blieb es aber bei dem der Anzahlen von Nullen und Einsen. Am meisten profitierte bezüglich der Korrektur dieser Fehlvorstellung jedoch die Gruppe derer, die die „R Aufgaben" oft oder immer eigenständig gelöst hatte (R34). Das ist insofern bemerkenswert, als dass diese Fehlvorstellung explizit sowohl in der Vorlesung als auch in theoretischen Übungen wie auch in einer „R Aufgabe" adressiert wurde. Betrachtet man insbesondere die Begründungen, so scheint das experimentelle Hantieren mit Münzwurfreihen in Form von Simulationen tatsächlich zu einer Verbesserung der Vorstellung zu führen. Ob die Studierenden der Gruppe R34 deshalb noch mehr davon profitierten als die Fragestundenbesucher, weil sie insgesamt auch auf einem solidieren Fachwissen aufbauen zu können schienen oder ob die Eigenständigkeit der Bearbeitung entscheidend ist, lässt sich an dieser Stelle nicht abschließend beurteilen. Sollte die Eigenständigkeit der Bearbeitung der wesentliche Erfolgsfaktor sein, ist möglicherweise die Einführung eines „Vorkurses" hilfreich, durch den mehr Studierende befähigt werden könnten, die Simulationsaufgaben mit weniger technischer Unterstützung zu lösen.

Auf das Zutrauen in die eigenen Lösungen scheint die Beschäftigung mit den „R Aufgaben" keinen Einfluss zu haben. In den meisten Fällen gab es gar keine Unterschiede zwischen den Untersuchungszeitpunkten bzw. den Gruppen. Bei den Fragestundenbesuchern gab es zwar Veränderungen, diese waren aber uneinheitlich.

Es schließt sich die Frage an, welche Studierende von den Simulationsaufgaben besonders profitieren. Der in Kapitel II.C.4.e.2 beschriebene Leistungsindex könnte hierzu Hinweise liefern. Bemerkenswert ist, dass bei den „Leistungsschwachen" zwar am Ende des Semesters innerhalb der Experimentalgruppe keine signifikanten Unterschiede gefunden werden konnten, es aber zwischen denen, die die „R Aufgaben" regelmäßig bearbeitet hatten (unabhängig davon, ob sie das eigenständig oder mit Unterstützung in der Fragestunde getan hatten) und jenen aus der Kontrollgruppe, die also mit gar keinen Simulationen in Kontakt gekommen waren, auffällige Unterschiede gab. Dass innerhalb der Experimentalgruppe kein Unterschied zwischen jenen gefunden werden konnte, die die „R Aufgaben“ selten oder nie bearbeitet hatten (R12), und jenen, die sie regelmäßig gelöst hatten (R34 und Frage), könnte damit zusammenhängen, dass die Stichprobengröße von R12 deutlich geringer ist als die der anderen Gruppen. Betrachtet man die individuellen Leistungen bei den beiden Aufgaben, die in analoger Weise zu Beginn und am Ende des Semesters gestellt wurden, bestätigt sich der Eindruck: Bei denen, die die „R Aufgaben“ regelmäßig bearbeitet haben, verbessern sich diese Leistungen. Insgesamt könnte das Ergebnis also darauf hinweisen, dass eine regelmäßige Beschäftigung mit den „R Aufgaben“ (kombiniert mit dem Einsatz von Simulationen in der Vorlesung) bei den „leistungsschwachen“ Studierenden zu einer Verbesserung der Grundvorstellungen führt.

Bei den „leistungsstarken“ Studierenden ist das Ergebnis weniger eindeutig. Hier war nur zwischen der Kontrollgruppe und jenen, die die „R Aufgaben“ oft oder immer selbstständig bearbeitet hatten, ein signifikanter Unterschied in der Leistung am Ende des Semesters zu verzeichnen. Allerdings besuchten von den „leistungsstarken“ Studierenden auch nur äußerst wenige die Fragestunde, so dass die Tatsache, dass bei ihnen kein Unterschied zur Kontrollgruppe festzustellen war, auch auf die geringe Stichprobengröße zurückzuführen sein könnte. Zudem waren sie zu Beginn des Semesters auffällig schwächer als die „Leistungsstarken“ der Kontrollgruppe, so dass ein Vergleich ohnehin schwierig ist. Innerhalb der Experimentalgruppen schienen sich die Leistungen am Ende des Semesters anzugleichen, die „leistungsstarken“ Fragestundenbesucher unterschieden sich am Ende nicht mehr signifikant von den „leistungsstarken“ anderen Gruppen, was auf eine Verbesserung ihrer Fertigkeiten hindeuten könnte. Obwohl bei den „Leistungsstarken“ die Gruppe derer, die die „R Aufgaben“ selten oder nie bearbeitet hatten, nicht wesentlich kleiner war als die derer, die „R Aufgaben“ oft oder immer selbstständig gelöst hatten, ließen sich zwischen diesen Gruppen keine Unterschiede in der Entwicklung feststellen. Ob also schon der Kontakt mit Simulationen in der Vorlesung zu einer Verbesserung der Grundvorstellungen geführt hat oder es sich bei dem Ergebnis um ein statistisches Artefakt handelt, bleibt offen. Bei den „leistungsstarken“ Studierenden gibt es folglich auch Hinweise darauf, dass eine Beschäftigung mit den „R Aufgaben“ einen Nutzen für die Verbesserung der Grundvorstellungen bringt, allerdings ist er deutlich schwächer einzustufen als bei den „leistungsschwachen“ Studierenden.

Da die Zuordnung zu den Gruppen sowohl aus ethischen als auch praktischen Gründen auf einer individuellen Wahl und nicht einem randomisierten Verfahren beruhen konnte, ist es schwierig, aus den dargestellten Ergebnissen allgemeingültige Schlüsse zu ziehen. Zusammenfassend lässt sich aber dennoch festhalten, dass die Bearbeitung der Simulationsaufgaben ein Potenzial zu haben scheint, die beschriebenen Fehlvorstellungen zu vermindern. Bezieht man die Erkenntnisse von Kahneman und Tversky mit ein, so scheint es aber eine Grenze zu geben, die nicht überschritten werden kann: Eine vollständige Auflösung der Fehlvorstellungen scheint (zumindest mit dem hier beschriebenen Ansatz) nicht möglich zu sein. Am meisten von Simulationsaufgaben zu profitieren scheinen „leistungsschwächere" Studierende, die das volle Maß an angebotener Unterstützung (in Form der zusätzlichen Fragestunde) in Anspruch nehmen. Um die leistungsstärkeren Studierenden bezüglich ihrer stochastischen Fertigkeiten durch Simulationsaufgaben noch besser zu fördern, müsste das Angebot evtl. angepasst werden. Denkbar wären beispielsweise Simulationsaufgaben unterschiedlicher Schwierigkeitsgrade. Eine weitere Schlussfolgerung, die aus der Untersuchung gezogen werden kann, ist, dass eine explizite Adressierung von Fehlvorstellungen in Simulationsaufgaben, wie bei F1 erfolgt, hilfreicher zu sein scheint als eine wiederholte implizite, wie es bei F2 versucht wurde. Eine Empfehlung könnte diesbezüglich sein, sowohl in der Vorlesung als auch in den Übungsaufgaben die Studierenden zu einer expliziten Auseinandersetzung mit verbreiteten Fehlvorstellungen aus dem Gebiet der Stochastik zu animieren.

Die bisher diskutierten Ergebnisse beziehen sich auf die stochastischen Fertigkeiten bezüglich der Lösung von Aufgaben. Ein anderer Bereich, der bei dieser Arbeit zwar nicht im Fokus stand, dennoch am Rande mit untersucht wurde, ist die Frage, wie sich Simulationen auf die Fähigkeit zur Versprachlichung stochastischer Konzepte auswirken könnten. Hierzu lassen sich in den Interviews Hinweise finden. Bei den Erklärungsversuchen der von den Befragten genannten Begriffe fällt auf, dass viele Studierende große Schwierigkeiten hatten, die hinter den Objekten bzw. Begriffen stehenden Konzepte in Worte zu fassen. Selbst grundlegende Begriffe konnten oft nicht treffend erklärt werden, wobei der Erfolg oder Misserfolg des Erklärungsversuchs nicht eindeutig mit dem Bearbeitungsmodus der „R Aufgaben" oder der Selbsteinschätzung bezüglich des Vertrauens- oder Einstellungsindexes zusammenzuhängen schien. Die Auswahl der Begriffe (vgl. Tab II.C.74.) unterschied sich dagegen zwischen den Gruppen: Diejenigen, die die „R Aufgaben" selten oder nie bearbeitet hatten (R12), wählten auffällig grundlegendere, eher aus der Schule bekannte und weniger komplexe Begriffe als die Befragten der anderen Gruppen, besonders derjenigen, die die „R Aufgaben“ oft oder immer selbstständig bearbeitet hatten (R34). Bei letzteren deutete die Begriffswahl auf eine größere Sachkenntnis hin. Die genannten Begriffe bezogen sich nicht nur auch auf komplexere Konzepte, sondern waren teilweise auch wesentlich konkreter auf fortgeschrittene Vorlesungsinhalte zu beziehen.

Bei den Erklärungsversuchen wurden im Wesentlichen drei Strategien verfolgt: Am häufigsten wurde probiert, eine innermathematische Beschreibung zu liefern, teilweise unter Verwendung formal-technischer Begriffe, teilweise in Alltagssprache. Viele dieser Beschreibungen waren jedoch entweder unzureichend oder sogar inhaltlich falsch. Zwei weitere, etwa gleich häufig angewandte Strategien bestanden darin, eine Erläuterung anhand von Beispielen zu formulieren oder aber Ableitungen aus der Wortbedeutung in der Alltagssprache zu versuchen. Während die Beispiele in den meisten Fällen passend gewählt, wenn auch nicht unbedingt richtig ausgeführt waren, führte der Versuch einer Erklärung über die Wortbedeutung in der Regel in die Irre, da der Gebrauch in der Alltagssprache nicht mit dem in der Fachsprache übereinstimmte. Keiner der Studierenden nahm explizit Bezug auf eine „R Aufgabe“ oder eines der Anwendungsfelder von Simulationen im Allgemeinen, und auch die in den „R Aufgaben“ behandelten Themen (vgl. Tab.II.B.1.) scheinen für die Befragten nicht in besonderer Weise im Vordergrund zu stehen. Eventuell könnte hier ein expliziteres Angebot von Reflexionsanlässen zum Nutzen der Aufgaben Abhilfe schaffen.[149]

Sicherlich ist zu berücksichtigen, dass die Studierenden nicht darauf vorbereitet waren, einen Fachbegriff erklären zu müssen, und es ist auch keine allzu leichte Aufgabe, dies spontan zu leisten (vgl. hierzu auch die von Zieffler et al. zusammengefassten Studienergebnisse, ausgeführt in Kapitel I.A.1.). Dennoch ist es nicht befriedigend, dass selbst grundlegende Konzepte am Ende des Semesters sogar bei teilweise hervorragenden Klausurergebnissen nicht richtig in Worte gefasst werden können. Insbesondere für Lehramtsstudierende, aber auch für Studierende mit anderem Abschlussziel ist diese Fähigkeit für spätere berufliche Kontexte immens wichtig – vom Beitrag der sprachlichen Dimension zum Verständnis dieser Konzepte einmal ganz abgesehen. Auch wenn fast alle „R Aufgaben“ mindestens einen Aufgabenteil enthielten, bei denen explizit ein Auftrag zur sprachlichen Ausformulierung oder Interpretation der Ergebnisse enthalten war (vgl. Tab.II.B.1.), könnte der Fokus noch stärker auf die Versprachlichung stochastischer Konzepte gelegt werden. Denkbar wäre

[149] Vgl. hierzu bspw. Regina Bruder, Timo Leuders & Andreas Büchter, Mathematikunterricht entwickeln. Bausteine für kompetenzorientiertes Unterrichten, Berlin 2012, 47ff.

beispielsweise, den Anteil an Aufgaben zu erhöhen, in denen Ergebnisse schriftlich ausformuliert und interpretiert oder aber stochastische Konzepte in Worte gefasst werden müssen. Die „R Aufgaben" böten sich dafür insofern an, als dass sie konkrete und anschauliche Beispiele für solche Konzepte beinhalten, mit denen im Sinne des Anwendungsfeldes A2 auch experimentiert werden kann. Wie genau ein solch erweiterter Einsatz von Simulationsaufgaben aussehen könnte und ob sich dadurch die Fähigkeit der Studierenden zur Versprachlichung verbessern ließe, ist ein möglicher Ansatzpunkt für weitergehende Forschung auf diesem Gebiet.

III.B. Diskussion der Ergebnisse bezüglich der Selbsteinschätzung der Befragten

Entsprechend der Erwartung aufgrund der in der Literatur beschriebenen Ergebnisse ließen sich in der Vorstudie kaum Veränderungen in den Einstellungen der Studierenden bezüglich des Fachs Stochastik beobachten. Es gab keine auffälligen Unterschiede zwischen den Zeitpunkten am Anfang und am Ende des Semesters. Dem Fach wird eine große Nützlichkeit zuerkannt (E8), die Zuneigung dem Fach gegenüber ist im mittleren Bereich angesiedelt (E5) und es wird nicht als besonders typisch, aber auch nicht besonders untypisch für eine mathematische Disziplin betrachtet (E9). Bemerkenswert ist an dieser Stelle, dass Praxis- und Anwendungsorientierung sowie Anschaulichkeit als eher untypisch für die Mathematik angesehen werden. Daraus könnte sich schließen lassen, dass diese Aspekte des Faches in den sonstigen theoretischen Veranstaltungen des Studiums kaum abgebildet werden und z.B. durch den Einsatz von Simulationen stärker in den Mittelpunkt gerückt werden könnten. Auch bei den benannten Schwierigkeiten lassen sich Anknüpfungspunkte für Simulationen finden. Beispielsweise könnte das Finden einer geeigneten Herangehensweise für die Lösung von Aufgaben bzw. das Generieren hierfür brauchbarer Hypothesen zumindest in einigen Fällen durch Simulationen unterstützt werden (Anwendungsfeld A2).

Auch bezüglich ihrer Selbsteinschätzungen der Kompetenzen im Fach Stochastik gab es im Laufe des Semesters kaum Unterschiede. Weder auf die stochastische Intuition (durch E2 und E3 abgefragt) noch die Vorstellung von Zufall hatte die Veranstaltung in der Einschätzung der Studierenden einen merklichen Einfluss. Der auffällige Unterschied bezüglich des Vertrauens in die eigenen Lösungen (E7) könnte darauf hindeuten, dass es eher zu einer Verunsicherung gekommen ist. Dass diese Verunsicherung durch die Möglichkeit der Kontrolle etwa durch Simulationen vermindert werden kann, deutet sich aus einem Ergebnis der **Hauptstudie** an: Bei den Besuchern der Fragestunde kam es diesbezüglich zu einer auffälligen Verbesserung.

Allerdings hatten sie zu Beginn des Semesters ein signifikant niedrigeres Vertrauen in die eigenen Lösungen als die anderen beiden Gruppen, d.h., die Bearbeitung der „R Aufgaben" scheint das Potenzial einer Angleichung zu bieten.

Die späteren Fragestundenbesucher schätzten ihre Kompetenzen insgesamt hinsichtlich fast aller Items schlechter ein als diejenigen, die sich später einer der anderen Gruppen zuordneten, zeigten sich also unsicherer. Die deutlichen Unterschiede relativierten sich teilweise im Laufe des Semesters: Bezüglich der Einschätzung des eigenen Wissens (E1) und der Schwierigkeiten beim Lösen von Stochastik-Aufgaben im Vergleich zu anderen Aufgaben (E6) kam es zu einer Angleichung mit denen, die die „R Aufgaben" selten oder nie bearbeitet hatten. Gerade sich selbst als unsicher einschätzenden Studierenden könnte also der Besuch der Fragestunde empfohlen werden.

Bemerkenswert ist eine Entwicklung bei der Gruppe derer, die die „R Aufgaben" selten oder nie bearbeitet hatten: Bezüglich der Items, die sich auf die Intuition beziehen (E2 und E3), unterschieden sich ihre Angaben am Ende des Semesters anders als am Anfang insbesondere von denen, die die „R Aufgaben" oft oder immer selbstständig bearbeitet hatten, in der Art, dass sie ihre Intuition als besser einschätzten. Auf der anderen Seite ergab ein Vergleich der Angaben auf individueller Ebene, dass sie die Schwierigkeiten bei der Lösung von Stochastik-Aufgaben im Vergleich zu Aufgaben aus anderen Gebieten (E6) am Ende größer einschätzten als am Anfang. Möglicherweise überschätzten diejenigen, die aufgrund fehlenden Zugangs zu Simulationen weniger Kontrollmöglichkeiten ihrer Vermutungen zur Verfügung hatten, ihre Intuition. Möchte man die Bearbeitung der „R Aufgaben" auch weiterhin als fakultativ beibehalten, wäre zu überlegen, ob sich für Studierende, die diese nicht bearbeiten wollen oder können, andere Feedbackmechanismen zu intuitiven Vorstellungen einrichten lassen könnten. Denkbar wäre beispielsweise, auch in theoretischen Aufgaben mehr Wert auf einen expliziten Abgleich von intuitiven mit berechneten Lösungen zu legen, etwa indem zunächst Hypothesen aufgestellt werden müssen, die anschließend überprüft werden sollen. Inwiefern sich ein solcher Zugang positiv auf eine realistische Einschätzung der eigenen Intuition auswirken könnte, müsste empirisch untersucht werden.

Anders als in der Vorstudie gab es hinsichtlich der Einstellungen zur Stochastik einige Entwicklungen im Laufe des Semesters. Auffällig ist insbesondere, dass sich bei denjenigen, die die „R Aufgaben" selten oder nie bearbeitet hatten, die Einstellungen verschlechterten (signifikanter Rückgang bezüglich E5 und E8 innerhalb der Gruppe). Hinsichtlich der Nützlichkeit von Stochastik (E8) zeigten nur die Fragestundenbesucher keinen Rückgang ihrer Einschätzungen, was darauf schließen lässt, dass der Austausch über die Simulationsaufgaben einer aufkeimenden leichten (die Zustimmungswerte waren ja trotzdem noch vergleichsweise hoch) Skepsis dem Fach gegenüber Einhalt gebieten konnte.

Bemerkenswert ist ebenso wie in der Vorstudie das Bild von Mathematik, das sich in den Ergebnissen bezüglich des Items E9 zeigt. Diejenigen, die die „R Aufgaben" oft oder immer bearbeitet hatten – unabhängig davon, ob sie dies eigenständig oder mit Unterstützung in der Fragestunde taten – empfanden das Fach Stochastik am Ende des Semesters als weniger typisch mathematisch als am Anfang. Ein experimenteller Zugang scheint für viele Studierende kein typischer Bestandteil mathematischer Arbeit zu sein – was im Hinblick darauf, welch große Bedeutung Simulationen in der modernen Mathematik zukommt, als Fehleinschätzung zu werten ist. Hierzu passen auch die Ausführungen in den Freitextfeldern und entsprechende Passagen in den Interviews (vgl. z.B. Antworten zum Fragenkomplex 10A/11B, Kap. II.C.5.), die darauf hindeuten, dass sich das Bild von Mathematik rein auf formalisiertes Arbeiten beschränkt.

Wie auch in der Vorstudie wurden am Ende des Semesters (neben allgemeinen Anpassungsschwierigkeiten an die Hochschulmathematik) vor allem Schwierigkeiten genannt, die möglicherweise durch Simulationen adressiert werden könnten. In diesem Fall standen Schwierigkeiten, konkrete und anschauliche Vorstellungen zu stochastischen Konzepten zu entwickeln, im Mittelpunkt. Dass diese Hürden trotz der Beschäftigung mit Simulationsaufgaben bestehen blieben, deutet darauf hin, dass das Potenzial des Zugangs nicht voll ausgeschöpft wurde. Auch diesbezüglich könnte eine stärkere Fokussierung auf Aufgaben zur Versprachlichung stochastischer Konzepte, aufbauend auf Simulationen, Verbesserungsmöglichkeiten bieten.

Ebenso wie hinsichtlich der bei den Aufgaben gezeigten Leistungen soll auch bezüglich der Selbsteinschätzungen versucht werden, herauszufinden, wer genau von der Beschäftigung mit den „R Aufgaben" profitieren konnte. Dazu wurde in Anlehnung an den *Leistungsindex* zusätzlich ein *Vertrauensindex* und ein *Einstellungsindex* erstellt (vgl. Kapitel II.C.4.e.3). Bemerkenswert ist, dass unabhängig vom Modus der Bearbeitung (selbstständig oder mit Unterstützung in der Fragestunde) bei Studierenden mit zunächst geringem Vertrauen in die eigenen stochastischen Fähigkeiten ein signifikanter Anstieg des *Vertrauensindex* im Laufe des Semesters zu verzeichnen war. Die Beschäftigung mit den „R Aufgaben" brachte also zunächst unsicheren Studierenden einen Gewinn an Vertrauen in die eigenen Fähigkeiten. An dieser Stelle schließt sich die Frage an, ob dieser Vertrauensgewinn auch gerechtfertigt ist, also mit den tatsächlichen Fertigkeiten im Fach Stochastik korreliert. Mit den hier erhobenen Daten lässt sich diese Frage nicht abschließend beantworten, durch einen Vergleich mit den individuellen Leistungen kann aber nach Hinweisen gesucht werden. So zeigte sich bei den vertrauensschwachen Besuchern der Fragestunde bezüglich der beiden zu Beginn und am Ende des Semesters analog gestellten Aufgaben eine signifikante Verbesserung des individuellen Antwortverhaltens. Bei den zunächst vertrauensschwachen Studierenden, die die „R Aufgaben" häufig oder immer selbstständig bearbeitet hatten, rührte die Steigerung des Vertrauens dagegen nicht von einer nachweisbaren Steigerung der Leistung bezüglich dieser beiden Aufgaben her (die Leistungen waren aber auch vorher schon gut). Vermutlich profitieren diese Studierenden insbesondere vom Anwendungsfeld A1 (Simulationen als Hilfsmittel zur Überprüfung von theoretischen Ergebnissen oder Verfahrensweisen), da sie eigene Überlegungen mit den Simulationsresultaten abgleichen können und so bei richtiger Herangehensweise unmittelbare Bestätigung erfahren.

Für diejenigen, die schon zu Beginn des Semesters ein eher großes Vertrauen in ihre Fähigkeiten hatten, brachte die Beschäftigung mit den Aufgaben dagegen keinen Gewinn – auch nicht im Vergleich zu denen, die keinen oder wenig Kontakt zu Simulationen hatten. Bei diesen Studierenden kam es unabhängig davon, ob die „R Aufgaben" bearbeitet wurden oder nicht, zu einer Angleichung der Vertrauenswerte zwischen den Gruppen. Bezüglich der individuellen Leistungen ließen sich bei den vertrauensstarken Studierenden kaum Entwicklungen beobachten, lediglich bei denen, die die „R Aufgaben" oft oder immer selbstständig bearbeitet hatten, ergab sich ein auffälliger p-Wert. Diese mögliche Verbesserung führte aber nicht zu einer weiteren Steigerung des Vertrauens.

Hinsichtlich des *Einstellungsindexes* konnte bei keiner Gruppe ein Profit durch die „R Aufgaben" beobachtet werden. Auch hier kam es bei den anfangs „Stärkeren" zu einer Angleichung der Einschätzungen – unabhängig davon, ob die „R Aufgaben" bearbeitet wurden oder nicht. Dies deckt sich mit den Studienergebnissen, die von Zieffler et al. zusammengefasst wurden (vgl. Kapitel I.A.1.).

Wie die Selbsteinschätzungen bezüglich der unterschiedlichen Kategorien zusammenhängen, kann durch die Berechnung von Korrelationen untersucht werden. Die Ergebnisse sind aber aufgrund der wenigen möglichen Ausprägungen im *Leistungsindex* nur hinsichtlich „Vertrauen" und „Einstellung" aussagekräftig. Wie zu erwarten war, besteht eine eher starke Korrelation zwischen dem *Vertrauensindex* und dem *Einstellungsindex*: Je stärker die

Studierenden ihren eigenen Kompetenzen im Fach Stochastik vertrauten, desto positiver standen sie auch dem Fach gegenüber.

III.C. Diskussion der Ergebnisse bezüglich der Evaluation der „R Aufgaben“

Die in den Items R1/R2 und R3 verwendeten Begriffe „Verständnis“ und „Intuition“ wurden in den Fragebögen nicht definiert, sondern es wurde angenommen, dass die Alltagsverwendung der Begriffe sich im Wesentlichen mit derjenigen dieser Arbeit deckt (vgl. Kap. I.A.3.). Um diese These zu überprüfen, wurde in den Interviews nach dem Verständnis der beiden Begriffe gefragt (Fragenkomplex 9A/10B). Bezüglich beider Begriffe kann in so weit von einer Deckung ausgegangen werden, dass die Items sinnvoll interpretierbar sind. Der Intuition wurden insbesondere die Eigenschaften der Unmittelbarkeit und des Unbewussten zugeschrieben. Dabei schien den Befragten auch die Fehleranfälligkeit von Intuitionen bewusst zu sein, einige äußerten sich auch zu handlungsorientierten Möglichkeiten einer Verbesserung. Mit „Verständnis“ brachten die Studierenden einerseits Nachvollziehbarkeit und logische Schlussfolgerungen, andererseits aber auch die bewusste Anwendung theoretischer Erkenntnisse und Rechenregeln auf konkrete Aufgaben in Verbindung.

Zunächst kann beobachtet werden, dass Simulationen von den Personen, die die „R Aufgaben“ häufig oder immer bearbeitet haben, tendenziell als nützlicher für das Verständnis angesehen werden als von denen, die sie nicht bearbeitet haben – und zwar sowohl bezüglich eines Einsatzes in der Vorlesung als auch einer Verwendung in Form von Übungsaufgaben. Die Fragestundenbesucher scheinen die Nützlichkeit in den Übungen am höchsten einzuschätzen. Insgesamt bewegen sich die Werte aber bei allen Gruppen im mittleren Bereich, d.h., den Simulationen wird ein eher geringer Einfluss auf das Verständnis zuerkannt. Dabei wird nur ein geringer Unterschied zwischen Demonstration und Experiment gesehen: Die Nützlichkeit für das Verständnis wird in der Vorlesung und den Übungen vergleichbar eingeschätzt (R1 und R2). Lediglich bei jenen, die die „R Aufgaben“ oft oder immer eigenständig gelöst haben, gibt es einen Hinweis darauf, dass sie den Einsatz in der Vorlesung für etwas nützlicher halten als den in den Übungen. Allerdings scheinen trotzdem diejenigen, die die „R Aufgaben“ häufig oder immer bearbeitet haben, es tendenziell besser zu finden, selbst Programme zu schreiben, als nur Simulationen gezeigt zu bekommen, was sich an recht hohen Zustimmungswerten zu Item R4 zeigt. Dieses inkonsistente Ergebnis könnte darauf hindeuten, dass die Studierenden noch einen weiteren Gewinn in der Bearbeitung sehen, als nur eine Verbesserung ihres Verständnisses. Diejenigen, die die „R Aufgaben“ selten oder nie bearbeitet haben, scheinen diesen zusätzlichen Gewinn niedriger einzuschätzen (R4). Die Nützlichkeit für einen Ausbau der stochastischen Intuition (R3) wird insgesamt noch etwas geringer eingeschätzt als die für das Verständnis, insbesondere von jenen, die nur in der Vorlesung in Kontakt mit Simulationen gekommen sind.

Worin ein Gewinn unabhängig von einer möglichen Verbesserung des stochastischen Verständnisses oder der stochastischen Intuition bestehen könnte, lässt sich aus den Fragebögen nicht erschließen, es finden sich aber Hinweise in den Interviews: Für knapp die Hälfte der Befragten stand der Gewinn an technischen Fertigkeiten im Vordergrund. Die Möglichkeit, anhand der „R Aufgaben“ einen Einstieg in die Nutzung des Programms R zu erhalten und auch grundlegende Programmierfähigkeiten zu erwerben, scheint für die Studierenden unabhängig vom Studiengang einen großen Wert zu haben (vgl. Fragenkomplexe 4A/5B, 6A/4B/6B sowie 10A/11B). Dass ein Großteil derer, die die „R Aufgaben“ oft oder immer bearbeitet hatten, in den Interviews einen Simulationsplan für die *Krankenhaus-Aufgabe* entwerfen konnte (vgl. den letzten Abschnitt von Kap. II.C.5.), deutet darauf hin, dass entsprechende Fähigkeiten im algorithmischen Denken bei diesen Studierenden vorhanden sind. Ob die „R Aufgaben“ hierfür maßgeblich waren, lässt sich an diesem Ergebnis allerdings nicht ablesen. Hierzu wären weitere Untersuchungen mit entsprechenden Kontrollgruppen nötig.

Die Notwendigkeit, weitere Unterstützung in Form der Fragestunde anzunehmen, um die „R Aufgaben“ lösen zu können (R5), wurde erwartungsgemäß von denen, die die Aufgaben selbständig bearbeitet haben, als sehr gering beziffert und von denen, die die Fragestunde besucht haben, als sehr hoch. Bemerkenswert ist das Ergebnis derer, die die „R Aufgaben“ selten oder nie bearbeitet haben. Auch sie schätzten die Notwendigkeit sehr hoch ein. Das deutet darauf hin, dass technische Schwierigkeiten ein Hemmnis für die Bearbeitung sein könnten. Diese Schwierigkeiten lassen sich möglicherweise durch die Einrichtung eines Vorkurses, in dem grundlegende Fertigkeiten des Programmierens in R vermittelt werden, abfedern.

Bezüglich des Items R6, bei dem es darum ging, wie interessant die Studierenden die „R Aufgaben“ fanden, lässt sich Folgendes beobachten: Die Simulationsaufgaben wurden als umso interessanter eingeschätzt, je mehr Kontakt mit ihnen bestand. Am interessantesten fanden sie die Besucher der Fragestunde, möglicherweise steigert also der Austausch über die

Aufgaben ihre Attraktivität. Diejenigen, die sie nie oder selten bearbeitet haben, fanden die Aufgaben tendenziell eher uninteressant. Dieses Ergebnis könnte einerseits bedeuten, dass die „R Aufgaben" nicht auf den ersten Blick interessant sind, sondern ihr Potenzial erst durch die Beschäftigung mit ihnen entfalten, andererseits aber auch, dass es auf die Entscheidung, ob man die „R Aufgaben" bearbeiten möchte, auch einen Einfluss hat, ob man die Aufgaben interessant findet. Dass einige Studierende die Simulationsaufgaben nicht bearbeitet haben, könnte also daran liegen, dass sie ihnen einerseits (möglicherweise auf der Ebene der technischen Umsetzung[150]) zu schwer waren, andererseits aber nicht interessant genug erschienen, um den zusätzlichen Aufwand des Besuchs der Fragestunde in Kauf zu nehmen, um sie bearbeiten zu können. Allein die extrinsische Motivation, durch sie Punkte für die Klausur zu sammeln, reichte offenbar nicht aus. Eventuell wäre eine explizitere Offenlegung der Anwendungsfelder von Simulationen, vielleicht in Form eines Einführungsvortrags und der Zuordnung einzelner Aufgaben zu den Feldern, eine Möglichkeit, den inhaltlichen Gewinn, der sich durch die Bearbeitung der „R Aufgaben" bietet, transparenter zu machen und so auch das Interesse an den Aufgaben zu steigern. Dass den Studierenden dieser Aspekt der Aufgaben nicht so stark im Bewusstsein ist wie ein möglicher Gewinn auf formal-technischer Ebene, zeigen besonders die Antworten auf den Fragenkomplex 4A/5B in den Interviews. Denkbar wäre auch, die „R Aufgaben" stärker in die regulären Tutorien oder auch die Vorlesung einzubeziehen. Beispielsweise könnten dort Hintergründe der Aufgaben erörtert oder Ergebnisse diskutiert werden und so mögliche Gewinne durch eine Auseinandersetzung mit ihnen auf inhaltlicher Ebene verdeutlicht werden.

Einer Empfehlung, „R Aufgaben" auch weiterhin einzusetzen (R7), stimmten erwartungsgemäß diejenigen, die die „R Aufgaben häufig oder immer bearbeitet hatten, mit signifikant höheren Werten zu als diejenigen, die sie selten oder nie bearbeitet hatten. Alle mittleren Werte waren allerdings im mittleren Bereich der Skala, d.h. keine Gruppe stand einer weiteren Verwendung sehr skeptisch oder auch sehr überzeugt gegenüber. Aus den Freitextantworten lässt sich schließen, dass die technische Unterstützung für die Akzeptanz der Simulationsaufgaben eine große Rolle zu spielen scheint. Zum Inhalt der „R Aufgaben" gab es überwiegend positive Äußerungen in den Freitextfeldern, auch von Studierenden, die sie nicht selbst bearbeitet haben. Diesbezüglich scheint also kein großes Verbesserungspotenzial gesehen zu werden. Will man noch einen größeren Anteil der

[150] Die Antworten der in den Interviews befragten Studierenden zum Fragenkomplex 3A deuten darauf hin, dass die Schwierigkeiten der „R Aufgaben" eher weniger in ihrem stochastischen Inhalt als vielmehr in der algorithmischen Umsetzung gesehen wurden.

Studierenden für die Simulationsaufgaben gewinnen, scheint folglich eine noch stärkere technische Unterstützung, insbesondere auch ein erleichterter Einstieg in die Programmiersprache, ein geeigneter Ansatzpunkt zu sein. Darauf hatten ja auch die Antworten auf Item R5 hingedeutet und auch in den Interviews finden sich entsprechende Hinweise (vgl. Fragenkomplexe 1A/B, 5A und 10A/11B). Möglich wäre beispielsweise die Einrichtung eines Vorkurses, um die technische Einführung von der inhaltlichen Beschäftigung mit stochastischen Fragestellungen zu trennen. Inwieweit ein solcher Vorkurs auf Akzeptanz bei den Studierenden stoßen würde, ist allerdings schwer einzuschätzen.

Fazit

Der Einsatz von Simulationen in einem Einführungskurs zur Stochastik für Bachelor- und Gymnasiallehramtsstudierende kann als lohnenswert betrachtet werden und wird von den Studierenden mehrheitlich befürwortet. Insbesondere für leistungsschwächere und für tendenziell unsichere Studierende, die zusätzliche Unterstützung durch Teilnahme an einer begleitenden „Fragestunde“ suchen, scheint die Bearbeitung von Simulationsaufgaben gewinnbringend zu sein. Die Studierenden, die die Aufgaben selbstständig bearbeiten, scheinen tendenziell bereits vor der Veranstaltung als fachlich kompetenter gelten zu dürfen als diejenigen, die zusätzliche (methodische) Unterstützung benötigen. Auch sie können aber noch bezüglich des Abbaus von Fehlvorstellungen profitieren. Außerdem zeigen beide Gruppen, die die Simulationsaufgaben regelmäßig bearbeitet hatten, in den Interviews die Fähigkeit, ein stochastisches Problem so zu analysieren, dass es mit Hilfe einer Simulation bearbeitet werden kann. Daraus kann – auf Grundlage dessen, dass ein großer Teil der Befragten, die die Aufgaben regelmäßig bearbeitet hatten, zuvor keinen Kontakt zum Gebiet des Programmierens hatte – geschlossen werden, dass die „R Aufgaben“ auch auf der methodischen Ebene des algorithmischen Denkens gewinnbringend sind.

Bei zukünftigen Einsätzen von Simulationsaufgaben könnte es nützlich sein, noch mehr Wert auf die Versprachlichung entsprechender Ergebnisse zu legen, um Fehlvorstellungen besser entgegen zu wirken. In diesem Zusammenhang sollten Fehlvorstellungen auch explizit thematisiert werden bzw. eine unmittelbare Auseinandersetzung mit ihnen angeregt werden. Außerdem könnte für die Akzeptanz der Aufgaben und den Lernerfolg der Studierenden eine größere Transparenz bezüglich der Anwendungsfelder von Simulationen und ihres Beitrags zum inhaltlichen Verständnis stochastischer Konzepte geboten sein.

Die Wahl des Programms R ist in dem untersuchten Setting als günstig zu erachten, da es einerseits frei zugänglich ist und in der Praxis Verwendung findet, andererseits von den Studierenden als Softwarelösung akzeptiert wird, da die technischen Fähigkeiten, die durch die Bearbeitung der Simulationsaufgaben erworben werden können, geschätzt werden. Ein Angebot an technischer Unterstützung wird dabei aber als essentiell und ausbaufähig angesehen. Insbesondere ist in diesem Zusammenhang die Einrichtung eines Vorkurses, in dem wesentliche technische Fertigkeiten vermittelt werden, zu prüfen. Dieser könnte einerseits die anfängliche Hemmschwelle senken, andererseits auch eine breitere Eigenständigkeit bei der Bearbeitung der Simulationsaufgaben ermöglichen.

Als Ansätze für weiterführende Forschung auf dem Gebiet sind die genauere Analyse des Unterschieds von expliziter und impliziter Thematisierung von Fehlvorstellungen sowie der Einfluss von Versprachlichung auf den Abbau von Fehlvorstellungen zu nennen. Außerdem könnte bei Einführung eines Vorkurses interessant sein, inwiefern die Befähigung zu mehr Eigenständigkeit bei der Bearbeitung der Simulationsaufgaben insbesondere auch von leistungsschwächeren und tendenziell unsicheren Studierenden zu einer Verbesserung ihrer stochastischen Fertigkeiten und Selbsteinschätzungen führen könnte.

LITERATURVERZEICHNIS

Rolf Biehler (1997), Software for Learning and for Doing Statistics, International Statistical Review 65/2, 167-189.

Rolf Biehler, Working styles and obstacles: Computer-supported collaborative learning in statistics, ICOTS-7 (2006), 1-6. Abrufbar unter: https://iase-web.org/Conference_Proceedings.php?p=ICOTS_7_2006 (zuletzt aufgerufen am 04.01.2020).

Rolf Biehler & Joachim Engel, Stochastik: Leitidee Daten und Zufall, in: Handbuch der Mathematikdidaktik, hrsg. von Regina Bruder, Lisa Hefendehl-Hebeker, Barbara Schmidt-Thieme & Hans-Georg Weigand, Berlin, Heidelberg 2015.

Rolf Biehler & Carmen Maxara, Eingangstest Stochastik – Vorkenntnisse von Lehramtsstudierenden, Beiträge zum Mathematikunterricht 2005, Hildesheim, 91-94. Abrufbar unter: https://eldorado.tu-dortmund.de/handle/2003/30643 (zuletzt aufgerufen am 04.01.2020).

Rolf Biehler & Andreas Prömmel, Developing students‘ computer-supported simulation and modelling competencies by means of carefully designed working environments, Proceedings of ICOTS 8 (2010), 1-6. Abrufbar unter: https://www.researchgate.net/publication/257927054_Developing_Students'_Computer-Supported_Simulation_and_Modelling_Competencies_by_Means_of_Carefully_Designed_Working_Enviroments (zuletzt aufgerufen am 04.01.2020)

Jürgen Bortz, Statistik für Human- und Sozialwissenschaftler, 6. Aufl., Heidelberg 2005.

Regina Bruder, Timo Leuders & Andreas Büchter, Mathematikunterricht entwickeln. Bausteine für kompetenzorientiertes Unterrichten, 2. Aufl., Berlin 2012.

Markus Bühner, Einführung in die Test- und Fragebogenkonstruktion, 3. Aufl., München 2011.

Gail Burrill, Simulation as a tool to develop statistical understanding, ICOTS-6 (2002),1-6. Abrufbar unter: https://iase-web.org/Conference_Proceedings.php?p=ICOTS_6_2002 (zuletzt aufgerufen am 05.01.2020)

Beth Chance, Robert delMas & Joan Garfield, Reasoning About Sampling Distributions, in: The Challenge of Developing Statistical Literacy, Reasoning and Thinking, hrsg. von Dani Ben-Zvi & Joan Garfield, Dordrecht, The Netherlands 2004, 295-323 [zitiert nach: Zieffler et al. (2008)].

Emanuel Czuber, Wahrscheinlichkeitsrechnung und ihre Anwendung auf Fehlerausgleichung, Statistik und Lebensversicherung, 1. Bd., Leipzig und Berlin 1908.

Joseph W. Dauben, Abraham Robinson and Nonstandard Analysis – History, Philosophy, and Foundations of Mathematics, in: History and Philosophy of Modern Mathematics, hrsg. von William Aspray & Philip Kitcher, Univ. of Minnesota 1988, 177–200. Als pdf abrufbar unter: http://www.mcps.umn.edu/philosophy/11_7Dauben.pdf (zuletzt aufgerufen am 15.04.2020).

Robert delMas & Yan Liu, Exploring Students' Conceptions of the Standard Deviation, Statistics Education Research Journal [Online], 4/1(2005), 55-82 [zitiert nach: Zieffler et al. (2008)].

Hermann Dinges, Variables, in particular random variables, in: Activity and Sign. Grounding Mathematics Education, hrsg. von Michael H.G. Hoffmann, Johannes Lenhard und Falk Seeger, New York 2005, 305-311.

Dudenredaktion (Bibliographisches Institut), Duden online, Eintrag zum Stichwort „Intuition", https://www.duden.de/rechtschreibung/Intuition (zuletzt aufgerufen am 15.04.2020).

Joachim Engel & Rudolf Grübel, Bootstrap – oder die Kunst, sich selbst aus dem Sumpf zu ziehen, Mathematische Semesterberichte 55 (2008), 1-19. Preprint abrufbar unter: https://www.researchgate.net/publication/225158275_Bootstrap__oder_die_Kunst_sich_selbst_aus_dem_Sumpf_zu_ziehen/link/0912f5125f57c4da6c000000/download (zuletzt aufgerufen am 02.01.2020)

William Feller, An Introduction to Probability Theory and Its Applications, Vol. I, Third Edition, New York, London, Sydney 1968.

Petra Hauer-Typpelt, Angemessene Grundvorstellung zu Wahrscheinlichkeit und Zufall entwickeln – Vorschläge für den Stochastikunterricht, in: Didaktikhefte der Österreichischen Mathematischen Gesellschaft, Heft 43 (2010), 75-87. Abrufbar unter: https://www.oemg.ac.at/DK/Didaktikhefte (zuletzt aufgerufen am 26.01.20)

Christopher J. Jones & Paul L. Harris, Insight into the law of large numbers: A comparison of Piagetian and judgement theory, Quarterly Journal of Experimental Psychology 34A (1982), 479-488 [zitiert nach: Sedlmeier/ Gigerenzer (1997)].

Daniel Kahneman & Amos Tversky, Subjective Probability: A Judgement of Representativeness, in: Cognitive Psychology 3/3 (1972), 430-454.

Götz Kersting, G, Random Variables – without Basic Space, in: Trends in Stochastic Analysis, hrsg. von Jochen Blath, Peter Mörters, Michael Scheutzow, Cambridge 2009, 13-34.

Götz Kersting & Anton Wakolbinger, Elementare Stochastik, in Reihe: Mathematik kompakt, hrsg von Martin Brokate et. al., Basel 2008.

Götz Kersting & Anton Wakolbinger, Stochastische Prozesse, in Reihe: Mathematik kompakt, hrsg. von Martin Brokate et. al, Basel 2014.

Herbert Kütting & Martin J. Sauer, Elementare Stochastik. Mathematische Grundlagen und didaktische Konzepte, in Reihe: Mathematik Primarstufe und Sekundarstufe I + II, hrsg. von Friedhelm Padberg, 3. Aufl., korrigierter Nachdruck, Berlin, Heidelberg 2014.

Johannes Lenhard, Mit allem rechnen – zur Philosophie der Computersimulation, in Reihe: Ideen&Argumente hrsg. von Wilfried Hinsch & Lutz Wingert, Berlin/Boston 2015.

Thierry Martin, Wahrscheinlichkeit – ein mehrdeutiger Begriff, in: Zufall und Chaos, Spektrum der Wissenschaft Spezial 1 (2010), 12-16.

Carmen Maxara, Simulationskompetenzen und stochastische Kompetenzen – Ergebnisse einer explorativen Fallstudie, Vortrag auf der 43. Jahrestagung der GDM in Oldenburg (2009), 1-4. Abrufbar unter: https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=2ahUKEwjwvK01unmAhXLzoUKHe5iCmEQFjAAegQIAxAC&url=https%3A%2F%2Fcore.ac.uk%2Fdownload%2Fpdf%2F46913768.pdf&usg=AOvVaw3Oj_jQX1ETG7qctFu6uYb- (zuletzt aufgerufen am 04.01.2020).

Carmen Maxara & Rolf Biehler, Constructing stochastic simulations with a computer tool – students' competencies and difficulties, Proceedings of the Fifth Congress of the European Society for Research in Mathematics Education (2007), 762-771. Abrufbar unter: https://www.researchgate.net/publication/257927049_Constructing_stochastic_simulations_with_a_computer_tool_-_students'_competencies_and_difficulties (zuletzt aufgerufen am 04.01.2020).

Carmen Maxara & Rolf Biehler, Students' probabilistic simulation and modeling competence after a computer-intensive elementary course in statistics and probability, ICOTS-7 (2006), 1-6. Abrufbar unter: https://iase-web.org/Conference_Proceedings.php?p=ICOTS_7_2006 (zuletzt aufgerufen am 04.01.2020).

Thorsten Meyfarth, Die Konzeption, Durchführung und Analyse eines simulationsintensiven Einstiegs in das Kurshalbjahr Stochastik der gymnasialen Oberstufe. Eine explorative Entwicklungsstudie, Dissertation, Universität Kassel 2008, erschienen in: Kasseler Online-Schriften zur Didaktik der Stochastik, Bd. 6.

Jamie D. Mills, Using Computer Simulation Methods to Teach Statistics: A Review of the Literature, Journal of Statistics Education 10/1 (2002), 1-20. DOI (pdf) abrufbar unter: https://doi.org/10.1080/10691898.2002.11910548

Leonard Mlodinow, The Drunkard's Walk. How randomness rules our lives, New York 2008.

David Mumford, The Dawning of the Age of Stochasticity, in: Mathematics. Frontiers and Perspectives, hrsg. von Vladimir I. Arnold, Michael Atiyah, Peter D. Lax & Barry Mazur, AMS 2000,197-218.

Edward Nelson, Radically elementary probability theory, Annals of Mathematics Studies 117, Princeton 1987.

Jean Piaget & Bärbel Inhelder, The Origin of the Idea of Chance in Children (L. Leake et al. trans.), New York 1975 [zit. nach Sedlmeier/ Gigerenzer (1997)].

Kristina Reiss & Christoph Hammer, Grundlagen der Mathematikdidaktik. Eine Einführung für den Unterricht in der Sekundarstufe, in Reihe: Mathematik kompakt, Basel 2013.

Sheldon M. Ross, Statistik für Ingenieure und Naturwissenschaftler (übersetzt von Carsten Heinisch), 3. Aufl., München 2006.

Rainer Schlittgen, Das Statistiklabor. R leicht gemacht, 2. akt. Aufl., Berlin Heidelberg 2009.

Peter Sedlmeier & Gerd Gigerenzer, Intuitions About Sample Size: The Empirical Law of Large Numbers, in: Journal of Behavioral Decision Making 10 (1997), 33-51.

Terence Tao, Topics in random matrix theory, in Reihe: Graduate Studies in Mathematics, Vol. 132, AMS 2012. Zitiert nach Preprint (pdf), abrufbar unter: https://terrytao.wordpress.com/category/teaching/254a-random-matrices (zuletzt aufgerufen am 10.04.2020)

G. S. Tune, Response preferences: A review of some relevant literature, Psychological Bulletin 61 (1964), 286-302 [zit. nach Kahneman/ Tversky (1972)].

Amos Tversky & Daniel Kahneman, Belief in the Law of Small Numbers, Psychological Bulletin 76/2 (1971), 105-110.

Zalman Usiskin, What does it mean to understand some mathematics, in: Selected Regular Lectures from the 12th International Congress on Mathematical Education, hrsg. von Sung Je Cho, Springer Switzerland 2015, 821-841.

Rudolf vom Hofe, Grundvorstellungen mathematischer Inhalte, Heidelberg 1995.

Rudolf vom Hofe, Grundvorstellungen mathematischer Inhalte als didaktisches Modell, Journal für Mathematik-Didaktik 13/4 (1992), 345–364.

Willam A. Wagenaar, Subjective randomness and the capacity to generate information, in: Acta Pychologica 33 (1970), 233-242 [zit. nach Kahneman/ Tversky (1972)].

Arnold D. Well, Alexander Pollatsek & Susan J. Boyce, Understanding the effects of sample size on the variability of the mean, in: Organizational Behavior and Human Decision Processes 47/2 (1990), 289-312 [zit, nach Sedlmeier/ Gigerenzer (1997)].

Hans Wolpers, Didaktik der Stochstik, in Reihe: Mathematikunterricht in der Sekundarstufe II, hrsg. von Uwe-Peter Tietze, Manfred Klika, Hans Wolpers, Bd. 3, Braunschweig/ Wiesbaden 2002.

Michael Wood, The Role of Simulation Approaches in Statistics, Journal of Statistics Education 13/3 (2005), 1-11. DOI (pdf) abrufbar unter: https://doi.org/10.1080/10691898.2005.11910562

Andrew Zieffler & Joan B. Garfield, Studying the Role of Simulation in Developing Students' Statistical Reasoning, Bulletin of the International Statistical Institute 56th Session (2007), 1-5.

Andrew Zieffler, Joan Garfield, Shirley Alt, Danielle Dupuis, Kristine Holleque & Beng Chang, What Does Research Suggest About the Teaching and Learning of Introductory Statistics at the College Level? A Review of the Literature, Journal of Statistics Education 16/2 (2008), 1-25. DOI (pdf) abrufbar unter https://doi.org/10.1080/10691898.2008.11889566

ANHANG

1. Templates für die „R Aufgaben”
2. Lösungen der „R Aufgaben“
3. Beispiel für einen Fragebogen (hier SoSe 2012)
4. Ergebnisse der Kolmogorov-Smirnov- sowie der Wilcoxon- bzw. U-Tests
5. Leitfaden für die halbstrukturierten Interviews
6. Transkripte der Interviews mit farblichen Markierungen für die Zuordnungen zu den Fragenkomplexen

1. Templates für die „R Aufgaben”

```
# Kopf oder Zahl (1. Aufgabe)

############################################################
#############
# Allgemeines zu R:

# Kommentare werden mit "#" eingeleitet.

# R kann man als Taschenrechner benutzen:
3+4
1/2
2^3
exp(10)

# Zuordnungen macht man mit "<-"
# Vektoren: Um mehrere Zahlen zu einem Vektor
zusammenzufuegen gibt es den Befehl c():
a<-c(5,6,7)
# den Vektor a anschauen:
a

# Die Laenge von a (Anzahl der Eintraege):
length(a)

# Ein Vektor, der aus einer aufsteigenden Folge 1 bis 3
besteht:
b<-c(1:3)
b

# Rechnungen mit Vektoren werden komponentenweise
ausgefuehrt:
a+b
3*a

# Logische Abfragen (koennen wahr oder falsch sein):
3 < 5
4 >= 10
3 == 7 # doppeltes "=" !
7 == 7

# Auf Elemente eines Vektors zugreifen:
# Der 3. Eintrag von a:
a[3]
# Alle Eintraege von a, die groesser oder gleich 6 sind:
a[a>=6]

# Matrizen:
mat<-matrix(c(3,4,5,6,7,8),3,2) # 3 Zeilen und 2 Spalten
mat
# auf ein Element zugreifen:
mat[1,2] # Element in 1. Zeile und 2. Spalte

############################################################
#############
############################################################
#############

# Alle vorherigen Zuordnungen loeschen:
rm(list=ls())

############################################################
#############
# Wir simulieren einen Muenzwurf:

# Die Muenze wird 40 Mal geworfen.
n<-40

# "sample" waehlt aus einer Menge eine bestimmte Anzahl aus
(mit oder ohne Zuruecklegen). Wenn man sich die Hilfe zu
dieser Funktion anschauen moechte, gibt man ein:
?sample

# Wir waehlen aus der Menge {"Kopf","Zahl"} n Werte aus. Die
Ergebnisse werden in "spiel" gespeichert.
spiel<-sample(c("Kopf","Zahl"),n,replace=T)

# Rufe "spiel" auf, um sich die Eintraege anzuschauen:
spiel

# Gewinne und Verluste von Peter: Bei "Kopf" gewinnt Peter
einen Euro (+1), bei Zahl verliert er einen Euro (-1). Wir
erstellen einen neuen Vektor namens "peter" und benutzen
dafuer eine "for"-Scheife und eine "if"-Abfrage.

peter<-rep(0,n) # neuen Vektor vorbereiten

for(i in 1:n)
     {
      if(spiel[i]=="Kopf"){peter[i]<-1}else{peter[i]<--1}
     }


# Wieviel Geld gewinnt (oder verliert) Peter insgsamt? Wir
addieren alle Eintraege von "peter" und speichern das
Ergebnis in "gewinnpeter".
gewinnpeter<-sum(peter)
gewinnpeter

############################################################
#############
# Eine eigene Funktion definieren: Man kann in R recht
einfach eigene Funktionen definieren. In unserem Beispiel ist
es nicht unbedingt noetig, wir zeigen aber trotzdem mal, wie
das geht. Nach dem Befehl "function" gibt man in runden
Klammern die Input-Variablen an. Falls man keine Input-
Variablen verlangen möchte, bleiben die runden Klammern
einfach leer.

# Unsere Funktion soll das Spiel "Kopf oder Zahl" simulieren
und den Spielverlauf sowie die Gewinne von Peter in der
Konsole ausgeben.
```

```
kopfoderzahl<-function(n)
{
 readline("neues Spiel")
 # Um zu starten, muss man die Konsole aktivieren (darauf
klicken) und "Enter" druecken. Der "readline"-Befehl ist oft
nuetzlich, deshalb wird er hier gezeigt.
 n<<-n
 spiel<<-sample(c("Kopf","Zahl"),n,replace=T)

 # Zuordnung innerhalb einer Funktion mit "<<-", wenn man
auch spaeter noch darauf zurueckgreifen koennen moechte.
 peter<<-rep(0,n)
 for(i in 1:n)
     {
      if(spiel[i]=="Kopf"){peter[i]<<-1}else{peter[i]<<--1}
     }
 gewinnpeter<<-sum(peter)
 print(spiel)
 print(paste("Gewinn von Peter:",gewinnpeter))
}

# Funktionsaufruf für n=10: Hier müssen Sie die Anzahl
aendern!
kopfoderzahl(10)

# Gewinne und Verluste von Peter in einem Spielverlauf
gewinnverlauf<-function()
{
verlauf<<-rep(0,n)

for(i in 1:n)
     {
      verlauf[i]<<-sum(peter[1:i])
      # Die Eintraege 1 bis i von "peter" werden aufsummiert
und an die i-te Stelle des neuen Vektors "verlauf"
geschrieben.
     }
}

gewinnverlauf()
verlauf

##############################################################
#############
# Nun sollen Sie das Spiel mit 40 Wuerfen 10 Mal wiederholen.
Tipp: Am einfachsten geht es mit der Funktion kopfoderzahl,
die wir gerade definiert haben, und einer "for"-Schleife.


n<-40
w<-10
ergebnissepeter<-rep(0,w)

# Hier sollten Sie eine for-Schleife definieren

ergebnissepeter

##############################################################
#############
# Graphische Darstellung der Gewinne und Verluste von Peter
in einem Spielverlauf


plot(c(1:n),verlauf,type="l",main="Gewinne und Verluste von
Peter im
Spielverlauf",ylim=c(min(verlauf),max(verlauf)),xlab="Wurfnum
mer",ylab="Gewinn/Verlust")

abline(h=0)


##############################################################
#############
# Graphische Darstellung der Gewinne von Peter am Ende der
Spiele
# Hier tricksen wir ein wenig, um ein Stabdiagramm zu
erhalten:

# Zunaechst erzeugen wir w Wiederholungen des Spiels. Wenn w
groß ist, ist es sinnvoll, den "readline"-Befehl und die
"print"-Befehle in der Funktion kopfoderzahl()
auszukommentieren. Das koennen Sie hier machen:

kopfoderzahl<-function(n)
{
 readline("neues Spiel")
 n<<-n
 spiel<<-sample(c("Kopf","Zahl"),n,replace=T)
 peter<<-rep(0,n)
 for(i in 1:n)
     {
      if(spiel[i]=="Kopf"){peter[i]<<-1}else{peter[i]<<--1}
     }
 gewinnpeter<<-sum(peter)
 print(spiel)
 print(paste("Gewinn von Peter:",gewinnpeter))
}


# Jetzt wiederholen wir das Spiel w Mal und speichern die
Ergebnisse in ergebnissepeter:

w<-100   # Hier können Sie die Anzahl der Wiederholungen
festlegen.
ergebnissepeter<-rep(0,w)

for(i in 1:w)
{
kopfoderzahl(40)
ergebnissepeter[i]<-gewinnpeter
}
```

```
breaks<-seq(from=-n,to=n,by=0.1) # gibt die Breite der Stäbe
an
hist(ergebnissepeter,breaks=breaks,main=paste("Peters
Gewinnsummen bei", w, "Wiederholungen"),xlab="Peters
Gewinnsummen",ylab="Absolute
Haeufigkeiten",freq=T,col="black")

#############################################################
#############
#############################################################
#############
```

```
# Crabs (2. Aufgabe)

# Wuerfelwurf
w<-sample(1:6,2,replace=T)
w

# Augensumme
x<-sum(w)

########################################################
####################
# a) Einzelnes Spiel zum Anschauen


spiel<-function()
{
ergebnis<-"leer"
ersterwurf<-sample(1:6,2,replace=T)
x<-sum(ersterwurf)
readline("Neues Spiel")
readline(paste("Erster Wurf: X=",x))

# Spielende nach einer Runde
   if(x==7|x==11)
     {
      ergebnis<-"G"
      readline("Spieler gewinnt")
     }

  # Hier muss etwas ergaenzt werden!

# Spielende erst später
while(ergebnis=="leer")
     {
      neuerwurf<-sample(1:6,2,replace=T)
      neuesumme<-sum(neuerwurf)
      readline(paste("Neuer Wurf: Summe:",neuesumme))

      if(neuesumme==x)
          {
           # Hier muss etwas ergaenzt werden
          }

      # Hier muss etwas ergaenzt werden

      }
}


# Ergaenze: Wiederhole 10 Mal (am einfachsten geht das
mit einer for-Schleife):
```

```
########################################################
#######################
# b) Gewinnws aus n Spielen schätzen

n<-10000

ergebnis<-rep("leer",n)

 # Ergaenze: Wiederhole n Mal (am einfachsten geht das
mit einer for-Schleife) und speichere  die Ergebnisse
in Vektor ergebnis ab:



# Wie oft gewinnt der Spieler?
ergebnisnum<-rep(0,n)
for(i in 1:n)
     {
      # Ergaenze: if-Abfrage
     }

sum(ergebnisnum)/n

########################################################
#######################
# c) Vergleich mit fairem Spiel

k<-10000

# Mit der Funktion rbinom() kann man binomialverteilte
Zufallszahlen ziehen, vgl Hilfe zu rbinom(): ?rbinom
# Bsp: Ziehe 50 binomialverteilte Zufallszahlen mit
n=100 und Erfolgswahrscheinlichkeit p=0.6:
rbinom(50,100,0.6)

# Definiere hier zwei Vektoren fair und crabs mithilfe
von rbinom

min<-min(fair,crabs)
max<-max(fair,crabs)



# Faires Spiel: Stabdiagramm zeichnen

breaks<-seq(from=min,to=max,by=0.2)
hist(fair,breaks=breaks,border="red",col="red",main="V
ergleich von Crabs mit fairem Spiel: Anzahl Gewinne
bei 100
Spielrunden",cex.main=0.9,ylab=paste("Haeufigkeit
bei",k,"Simulationen"),xlab="Anzahl Gewinne")
text(min,max(table(fair)),"faires
Spiel",col="red",adj=0)


# Faires Spiel und Crabs vergleichen: Stabdiagramme
```

```
übereinander zeichnen
hist(fair,breaks=breaks,border="red",col="red",main="V
ergleich von Crabs mit fairem Spiel: Anzahl Gewinne
bei 100
Spielrunden",cex.main=0.9,ylab=paste("Haeufigkeit
bei",k,"Simulationen"),xlab="Anzahl Gewinne")

hist(crabs,breaks=breaks,border="blue",add=T)

text(min,max(table(fair)),"faires
Spiel",col="red",adj=0)
text(min,max(table(fair))*0.95,"Crabs",col="blue",adj=
0)
```

```
# Monte Carlo Methode zur Schaetzung von pi (3. Aufgabe)

# Ein Punkt
abstand<-function(x,y)
{
abstand<-sqrt((x-0.5)^2+(y-0.5)^2)
abstand
}

x<-runif(1,min=0,max=1)
y<-runif(1,min=0,max=1)

# im Kreis?
imkreis<-0
if(abstand(x,y)<=0.5){imkreis<-imkreis+1}

#############################################################
##########
# Wiederhole mit w=10000 Punkten: am besten definiert man
hier eine Funktion mit Input w. Wie bekommt man daraus einen
Schaetzer für den Flaecheninhalt? Erechnen Sie damit einen
Schaetzer für pi.

montecarlo<-function(w)
{
# Hier müssen Sie etwas ergaenzen
}


#############################################################
############

# Verfahren 100 Mal wiederholen und Bereich für pi-
Schaetzungen angeben

schaetzerpivektor<-rep(0,100)

for(j in 1:100)
{
# Hier muss etwas ergaenzt werden
schaetzerpivektor[j]<-schaetzerpi
}

# Bereich, in dem Schaetzungen liegen:
min(schaetzerpivektor)
max(schaetzerpivektor)

# oder
summary(schaetzerpivektor)

# Wie genau ist die Schaetzung?

#############################################################
############

# Wieviele Punkte braucht man, um bis auf 2 Nachkommastellen
genau zu schaetzen? Das können Sie mit Ihrer Funktion
montecarlo herausfinden oder theoretisch herleiten!
```

```
# Belegzeiten von Betten (4. Aufgabe)


rm(list=ls())

########################################################
######################
# Belegungszeit von Betten
# exp(0.1)-verteilt => MW=10
# Studie: In einem großen neu eröffnetem KH mit 800
Betten soll nach 100 Tagen untersucht werden, wie
lange Patienten im KH liegen. Dazu wird an diesem Tag
für jedes Bett notiert, welcher Patient darin liegt
und nach der Entlassung aufgeschrieben, wie lange
dieser Patient das Bett belegt hat.

########################################################
#######################

# Matrix vorbereiten
mat<-matrix(rep(0,800*50),nrow=800)

# Simuliere für 800 Betten die Liegezeiten von 100
Patienten

n<-800
a<-1 # Anzahl exp(10)-verteilter ZV, die erzeugt
werden soll: muss man aendern!

for(j in 1:n)
     {
      # Hier muss etwas ergaenzt werden
      # Ausreichende Anzahl a exp(0.1)-verteilter ZV
erzeugen, hier 100 Patienten
      # Befehl: rexp
     }

########################################################
#######################
# Wie lange hat Patient, der zum Zeitpunkt t=100 im
Bett lag, dieses belegt?

liege100<-rep(0,800)

for(j in 1:n)
     {
      index<-0
      z<-0
      for(i in 1:a)
          {
           if(z<100)
               {
                # Hier muss etwas ergaenzt werden
                # Aufsummierien, bis Summe groesser ist
als 100
```

```
                # Index, bei dem 100 ueberschritten
wird (Stoppzeit)
               }
          }
      # Hier muss etwas ergaenzt werden
      # Wert der (Stoppzeit)ten ZV ist gesuchte
Intervall-Laenge
     }

# Mittelwert:
mean(liege100)

########################################################
#######################
# Histogramm der Liegezeiten der Patienten, die zum
Zeitpunkt t=100 im Bett lagen
breaks<-seq(from=0,to=max(liege100)+1,by=1) # Breite
der Säulen im Histogramm
hist(liege100,breaks=breaks,freq=FALSE,col="yellow",yl
im=c(0,0.15),main="Wie lange belegten die Patienten
ihr Bett, die zum Zeitpunkt t=100 darin
lagen?",cex.main=1,ylab="",xlab="Liegezeit der
Patienten")

text(9,0.9,expression(Zeiten~exponential~verteilt~zum~
Parameter~0.1),cex=1.1)

#Zum Vergleich: exp(0.1)-Dichte einzeichnen:
l<-length(breaks)
c<-rep(0,l)
for(i in 1:l){c[i]<-0.1*exp(-0.1*breaks[i])}  #
Funktionswerte der Dichtefunktion
lines(breaks,c,lwd=1.8)

# => Groessenverzerrung bei dieser Schaetzmethode, man
braucht anderes Verfahren.

########################################################
#######################
# Ende
graphics.off()
```

```
# Fortsetzung: Belegzeiten von Betten (5. Aufgabe)


rm(list=ls())


#############################################################
###############
# Belegungszeit von Betten
# exp(0.1)-verteilt => MW=10
# Studie: In einem großen neu eröffnetem KH mit 800 Betten
soll nach 100 Tagen untersucht werden, wie lange Patienten im
KH liegen. Dazu wird an diesem Tag für jedes Bett notiert,
welcher Patient darin liegt und nach der Entlassung
aufgeschrieben, wie lange dieser Patient das Bett belegt hat.

#############################################################
################

# Matrix vorbereiten
mat<-matrix(rep(0,800*50),nrow=800)

# Simuliere für 800 Betten die Liegezeiten von 100 Patienten

n<-800

for(j in 1:n)
     {
      #Ausreichende Anzahl exp(10)-verteilter ZV erzeugen,
hier 100 Patienten
      x<-rexp(50,0.1)
      mat[j,]<-x
     }

#############################################################
################
# Wie lange hat Patient, der zum Zeitpunkt t=100 im Bett lag,
dieses belegt?

liege100<-rep(0,800)
indexvek<-rep(0,800)

for(j in 1:n)
     {
      z<-0
      for(i in 1:50)
          {
           if(z<100)
               {
                z<-z+mat[j,i] # Aufsummierien, bis Summe
groesser ist als 100
                indexvek[j]<-i # Index, bei dem 100
ueberschritten wird (Stoppzeit)
                }
           }
      # Wert der (Stoppzeit)ten ZV ist gesuchte Intervall-
Laenge
      liege100[j]<-mat[j,indexvek[j]]
     }

# Index
indexvek

#############################################################
################
# Ab hier kommt etwas Neues: Frage: Wie kommt es zu der
Verzerrung? Erklaren Sie die Verzerrung mit Hilfe der
folgenden Graphik!


plot(c(0,120),c(0,10),type="n",main="Belegzeiten
verschiedener Betten", ylab="",xlab="Tage",axes=F)
axis(1)
box()
col<-rainbow(max(indexvek)+1,alpha=0.5)

for(j in 1:10)
     {
      polygon(c(0,mat[j,1],mat[j,1],0),c(j-0.3,j-
0.3,j+0.3,j+0.3),col=col[1])
      for(i in 1:indexvek[j])
          {

polygon(c(sum(mat[j,1:i]),sum(mat[j,1:(i+1)]),sum(mat[j,1:(i+
1)]),sum(mat[j,1:i])),c(j-0.3,j-
0.3,j+0.3,j+0.3),col=col[i+1])
          }
     }

abline(v=100,col="red",lwd=3)

#############################################################
################
# Ueberlegen Sie sich eine andere Methode, die die
Liegezeiten unverzerrt schaetzt! Es gibt verschiedene
Moeglichkeiten. Sie koennen wieder die Matrix mat verwenden.


liege<-rep(0,800)

# Hier muessen Sie etwas ergaenzen

mean(liege)

#############################################################
##########

# Histogramm der Belegzeiten Betten
breaks<-seq(from=0,to=max(liege)+1,by=1) # Breite der Säulen
im Histogramm
hist(liege,breaks=breaks,freq=FALSE,col="yellow",ylim=c(0,0.1
```

```
5),main="Belegzeiten der Betten",xlab="",ylab="")

#Zum Vergleich: exp(0.1)-Dichte einzeichnen:
l<-length(breaks)
c<-rep(0,l)
for(i in 1:l){c[i]<-0.1*exp(-0.1*breaks[i])}  #
Funktionswerte der Dichtefunktion
lines(breaks,c,lwd=1.8)


#############################################################
################
```

```
# Vorzeichenwechsel beim Münzwurf (6. Aufgabe)

############################################################
############
# Alle vorherigen Zuordnungen loeschen:
rm(list=ls())

############################################################
############

# Wir untersuchen zunaechst einen fairen Muenzwurf

n<-20
p<-0.5

# Zeichne den Verlauf von 3 fairen Muenzwuerfen:

wuerfegraphik<-function(n,p)
{
par(mfrow=c(3,1))

for(j in 1:3)
     {


     #Hier muss etwas ergaenzt werden: In einem Vektor
"wurfnum" sollen die Ergebnisse eines Muenzwurfs
abgespeichert werden als 1 fuer Kopf und 0 fuer Zahl


     readline(paste("Muenzwurf Nr",j,"p=",p))

     plot(wurfnum,type="p",pch=16,main=paste("Muenzwurf",j,"(
p=",p,")"),axes=F,ylim=c(-4,3), xlim=c(-
2/n,n),ylab="",xlab="")

     abline(h=0)
     text(0,1,"Kopf")
     text(0,-1,"Zahl")
     text(n/2,-3,paste("Anzahl Vorzeichenwechsel:",
vorzwechsel))
     }
}

wuerfegraphik(n,p)

############################################################
############
# Wiederhole haeufig und zaehle Vorzeichenwechsel:

wurffunktion<-function(p)
{

# Hier muss etwas ergaenzt werden. Die Anzahl der
Vorzeichenwechsel soll in einer Variablen gespeichert werden.

}

w<-10000
vorzwechselvektor<-rep(0,w) # In diesem Vektor sollen die
Anzahlen der Vorzeichenwechsel fuer w Muenzwuerfe gespeichert
werden

for(i in 1:w)
     {
      # Hier muss etwas ergaenzt werden
     }


# Zeichne Stabdiagramm:

par(mfrow=c(1,1))
breaks<-seq(from=0,to=max(vorzwechselvektor),by=0.01)
histvorzwechsel<-
hist(vorzwechselvektor,breaks=breaks,main=paste("Anzahl
Vorzeichenwechsel bei",n,"p=", p, "-
Muenzwuerfen"),ylab="Absolute Hauefigkeit",xlab="Anzahl
Vorzeichenwechsel",xlim=c(1,max(vorzwechselvektor)))

# Erwartete Anzahl Vorzeichenwechsel
erwartung<-mean(vorzwechselvektor)

abline(v=erwartung,lwd=2,col="green")

text(1,max(histvorzwechsel$counts)-
0.1*max(histvorzwechsel$counts),"erwartete Anzahl
Vorzeichenwechsel",col="green",cex=0.8,adj=0)

############################################################
############
# Vergleich mit Vorzeichenwechsel aus gefaelschtem Muenzwurf:

#Anzahl Vorzeichenwechsel aus gefaelschtem Munezwurf
eintragen, Bsp: gefaelscht=10
gefaelscht <-10

par(mfrow=c(1,1))
breaks<-seq(from=0,to=max(vorzwechselvektor),by=0.01)
histvorzwechsel<-
hist(vorzwechselvektor,breaks=breaks,main=paste("Anzahl
Vorzeichenwechsel bei",n,"p=", p, "-
Muenzwuerfen"),ylab="Absolute Hauefigkeit",xlab="Anzahl
Vorzeichenwechsel",xlim=c(1,max(vorzwechselvektor)))
abline(v=erwartung,lwd=2,col="green")

text(1,max(histvorzwechsel$counts)-
0.1*max(histvorzwechsel$counts),"erwartete Anzahl
Vorzeichenwechsel",col="green",cex=0.8,adj=0)


text(1,max(histvorzwechsel$counts)-
0.1*max(histvorzwechsel$counts),paste("Vorzeichenwechsel bei
```

```
gefaelschtem Muenzwurf",
gefaelscht),col="blue",cex=0.8,adj=0)

points(gefaelscht,0,pch=16,col="blue",cex=2)


###############################################################
############
# Wir wollen untersuchen, wie hauefig eine so grosse
Abweichung vom (geschaetzten) Erwartungswert bei fairem
Muenzwurf vorkommt.

# Abweichung vom Erwartungswert
abw<-abs(gefaelscht-erwartung)

par(mfrow=c(1,1))
breaks<-seq(from=0,to=max(vorzwechselvektor),by=0.01)
histvorzwechsel<-
hist(vorzwechselvektor,breaks=breaks,main=paste("Anzahl
Vorzeichenwechsel bei",n,"p=", p, "-
Muenzwuerfen"),ylab="Absolute Hauefigkeit",xlab="Anzahl
Vorzeichenwechsel",xlim=c(1,max(vorzwechselvektor)))

abline(v=erwartung,lwd=2,col="green")

text(1,max(histvorzwechsel$counts)-
0.1*max(histvorzwechsel$counts),"erwartete Anzahl
Vorzeichenwechsel",col="green",cex=0.8,adj=0)

text(1,max(histvorzwechsel$counts)-
0.2*max(histvorzwechsel$counts),paste("Vorzeichenwechsel bei
gefaelschtem Muenzwurf",
gefaelscht),col="blue",cex=0.8,adj=0)

points(gefaelscht,0,pch=16,col="blue",cex=2)


text(1,max(histvorzwechsel$counts)-
0.3*max(histvorzwechsel$counts),"mind. so große Abweichung
vom Erwartungswert",col="red",cex=0.8,adj=0)

hist(vorzwechselvektor[vorzwechselvektor>=erwartung+abw],brea
ks=breaks,add=T,col="red",border="red")

hist(vorzwechselvektor[vorzwechselvektor<=erwartung-
abw],breaks=breaks,add=T,col="red",border="red")

abline(h=0,col="black",lwd=1.5)

#######

#Bsp: Wie oft gibt es so grosse Abweichung vom (geschaetzten)
Erwartungswert?
step<-seq(from=0,to=length(histvorzwechsel$counts),by=100)
abshauefigkeitenvorzwechsel<-
c(histvorzwechsel$counts[1],histvorzwechsel$counts[step])

ah<-
sum(abshauefigkeitenvorzwechsel[(floor(erwartung+abw)+1):leng
th(abshauefigkeitenvorzwechsel)])+sum(abshauefigkeitenvorzwec
hsel[1:ceiling(erwartung-abw)])


# Schaetzer für Ws, dass es mehr als 10 Vorzeichenwechsel
gibt: Hier muessen Sie etwas ergaenzen.



###############################################################
################
###############################################################
################

# Vorbereitung: Erwartungswert beim p- Muenzwurf
p<-0.1 # kann man veraendern
w<-10000
vorzwechselvektor<-rep(0,w)

for(i in 1:w)
     {
      #Hier muessen Sie etwas ergaenzen
     }

mean(vorzwechselvektor)


# Nun die eigentliche Aufgabe: Abhaengigkeit von p

w<-10000
vorzwechselmatrix<-matrix(rep(0,w*11),nrow=11)
pvektor<-seq(from=0,to=1,by=0.1)

for(j in 1:11)
     {
      for(i in 1:w)
         {
          # Hier muessen Sie etwas ergaenzen
         }
     }

# Erwartungswerte in Abhaengigkeit von p
erwartungsvektor<-rep(0,11)

for(j in 1:11)
     {
      # Hier muessen Sie etwas ergaenzen
     }


# Plot Erwartungswerte
plot(pvektor,erwartungsvektor,type="p",pch=16,main="Erwartete
Anzahl Vorzeichenwechsel als Funktion von p",ylab="Erwartete
```

```
Anzahl Vorzeichenwechsel",xlab="p",ylim=c(0,10))

text(pvektor[1],max(erwartungsvektor),"simulierte
Erwartungswerte",adj=0,cex=0.8)
```

```
# Regression zur Mitte (7. Aufgabe)
rm(list=ls())

##############################################################
###############
# Teil i
# Um die Daten (X,Y) zu simulieren, gibt es zwei
Moeglichkeiten. Am einfachsten geht es mit der Funktion
mvrnorm aus dem Paket MASS. Falls Sie diese Funktion nicht
zur Verfuegung haben, waeheln Sie bitte Variante 2.

############
# Variante 1 mit mvrnorm
library(MASS)
n<-10 # muss veraendert werden!
kappa<-0.7
varcov<-matrix(c(64,kappa*64,kappa*64,64),nrow=2,byrow=T)
daten<-mvrnorm(n,c(82,82),varcov)
gewicht1<-daten[,1]
gewicht2<-daten[,2]


# Variante 2
n<-1 # muss veraendert werden!
kappa<-0.7

x1<-rnorm(n)
x2<-rnorm(n)

gewicht1<-x1*8+82
gewicht2<-(kappa*x1+sqrt(1-kappa^2)*x2)*8+82

############

# Vergewissern Sie sich, dass die Daten richtig erzeugt
wurden! Kontrollieren Sie die Mittelwerte (mit dem Befehl
mean), die Standardabweichungen (mit dem Befehl sd) und die
Korrelation (mit dem Befehl cor)

# Hier muss etwas ergaenzt werden

############
# Zeichnen Sie ein Streudiagramm. Sie brauchen den Graphik-
Befehl plot.

# Hier muss etwas ergaenzt werden


##############################################################
###############
# Teil ii
# Wir speichern die Indizes, bei denen in gewicht1 Zahlen
kleiner als 100 stehen, in einem Vektor index ab. Machen Sie
sich klar, was durch index<-c(index,i) in der for-Schleife
passiert.

index<-0

for(i in 1:n)
{
 if(# hier muss etwas ergaenzt werden){index<-c(index,i)}
}

gewicht1vonSchweren<-gewicht1[-index]
gewicht2vonSchweren<-gewicht2[-index]

# Was sind die Mittelwerte der schweren Männer vorher und
nachher?
# hier muss etwas ergaenzt werden

# Hier werden die Histogramme erzeugt

ming<-min(gewicht1vonSchweren,gewicht2vonSchweren)
maxg<-max(gewicht1vonSchweren,gewicht2vonSchweren)

breaks<-seq(from=min(75,ming-5),to=maxg+5,by=5)

hist(gewicht1vonSchweren,breaks=breaks,col="#ff000099",main="
Gewichte der Schweren vorher und nachher (rot
vorher)",xlab="")
hist(gewicht2vonSchweren,breaks=breaks,xlab="",col="#0000ff44
",add=T)
abline(v=82,lwd=2,col="green")
```

```
# Rejection Sampling (8. Aufgabe)

# Alle vorherigen Zuordnungen loeschen:
rm(list=ls())

#############################################################
#############s

# Umrechnung in kartesische Koordinaten (%*% bildet
Skalarprodukt)
# Frankfurt
ffm<-c(cos(0.14)*cos(0.87),cos(0.14)*sin(0.87),sin(0.14))
einffm<-ffm/sqrt(ffm%*%ffm)


# Kapstadt
kap<-c(cos(0.31)*cos(-0.6),cos(0.31)*sin(-0.6),sin(0.31))
einkap<-kap/sqrt(kap%*%kap)


# Peking
pek<-c(cos(2.02)*cos(0.7),cos(2.02)*sin(0.7),sin(2.02))
einpek<-pek/sqrt(pek%*%pek)

#############################################################
#############


# Uniforme Punkte auf Kugelsspäre:

# 1. uniforme Punkte in [-1,1]^3 erzeugen
# Erzeuge 3 Vektoren: einen Vektor "x" fuer die x-
Koordinaten, einen Vektor "y" fuer die y-Koordinaten und
einen Vektor "z" fuer die z-Koordinaten

n<-10000

# Hier muss etwas ergaenzt werden

# Pruefe fuer jeden Punkt: Ist er in der Einheitskugel?
# Falls nein: speichere den Index im Vektor "Index"
# Zusatz: in der Variablen "inkugel" soll gezaehlt werden,
wieviele Punkte in der Kugel lagen

inkugel<-0
index<-rep(0,n)

# Hier muss etwas ergaenzt werden

# Nur Punkte, die in Einheitskugel liegen:
xkug<-x[-index]
ykug<-y[-index]
zkug<-z[-index]

l<-length(xkug)

# 2. Berechne U_i

norm<-rep(0,l)

for(i in 1:l)
{
help<-c(xkug[i],ykug[i],zkug[i])
norm[i]<-sqrt(help%*%help)
}

xproj<-rep(0,l)
yproj<-rep(0,l)
zproj<-rep(0,l)

for(i in 1:l)
{
# Hier muss etwas ergaenzt werden
}


# Naehere Flaecheninhalt des Kugeldreiecks an

imkugeldreieck<-0

for(i in 1:l)
     {
      help<-c(xproj[i],yproj[i],zproj[i])

      #Hier muss etwas ergaenzt werden

      if(det(mat1)<=0 & det(mat2)<=0 & det(mat3)<=0)
           {
            #Hier muss etwas ergaenzt werden
           }
     }


#############################################################
#############

# Schaetzer fuer relativen Anteil der Flaeche:
#Hier muss etwas ergaenzt werden
```

```
# Fehlstaende rein zufaelliger Permutationen (9. Aufgabe)


################################################################
################
rm(list=ls())

################################################################
################
# Teil i

# rein zufaellige Permutation der Laenge 5
n<-5
# Hier muss etwas ergaenzt werden: speichern Sie die
Permutation in einem Vektor p


# Fehlstaendevektor y
# y2 = y[1] = alle Fehlstaende, an denen die 2 zusammen mit
kleinerem Partner beteiligt ist usw

y<-rep(0,n-1)

for(j in 2:n)
     {
      for(i in 1:(j-1))
           {
            # Hier muss etwas ergaenzt werden
           }
     }

# alle Fehlstände
sum(y)


# w Wiederholungen

w<-10000

ymat<-matrix(rep(0,((n-1)*w)),ncol=(n-1))

for(m in 1:w)
{
# Hier muss etwas ergaenzt werden

ymat[m,]<-y
}

# Histogramme

par(mfrow=c(2,2))

for(i in 1:(n-1))
{
hist(ymat[,i],breaks=seq(from=0,to=i,by=0.01),
main=paste("Fehlstaende, an denen die", i+1, "mit kleinerem
Partner beteiligt"),cex.main=0.8,xlab="Anzahl
Fehlstaende",ylab=paste("Haeufigkeit
bei",w,"Wdh"),cex.lab=0.8)
}


################################################################
################

# Histogramm der Gesamtanzahlen der Fehlstaende bei w
Permutationen fuer n=5, n=10 und n=15

n<-15 # hier kann man den Wert von n veraendern!
w<-10000

ymat<-matrix(rep(0,((n-1)*w)),ncol=(n-1))

for(m in 1:w)
{

# Hier muss etwas ergaenzt werden (siehe oben)

ymat[m,]<-y
}


gesamt<-rep(0,w) # hier sollen die Gesamtanzahlen gespeichert
werden

for(i in 1:w){# Hier muss etwas ergaenzt werden}

breaks<-seq(from=0,to=max(gesamt)+10,by=1)
hist(gesamt,breaks=breaks,ylim=c(0,0.25),freq=F)


# Vergleich mit Normalverteilung

mu<-# Hier muss etwas ergaenzt werden
sigma<-# Hier muss etwas ergaenzt werden

x<-seq(0,max(gesamt)+10,0.1)
l<-length(x)
c<-rep(0,l)
for(i in 1:l)
{
c[i]<-1/(sqrt(2*pi)*sigma)*exp(-((x[i]-mu)^2/(2*sigma^2)))
}

lines(x,c,lwd=2,col="red")
```

```
# Konfidenzintervalle (10. Aufgabe)

######################################################
######################
rm(list=ls())

######################################################
######################
# Teil i: Simulation einer Umfrage mit p=1/3 und n=100

# Hier muss etwas ergaenzt werden. Speichern Sie das
geschätzte p in einer Variablen pdach.

# Konfidenzintervall berechnen. Berechnen Sie die
untere und die obere Grenze getrennt und speichern Sie
die Ergebnisse in zwei Variablen unteregrenze und
oberegrenze

######################################################
######################
# Teil ii

# Wiederhole w=1000 Mal

w<-1000

# Hier muss etwas ergaenzt werden. Speichern Sie das
geschätzte p in einem Vektor pdachvektor.

# Konfidenzintervalle berechnen
unteregrenzevektor<-rep(0,w)

# Hier muss etwas ergaenzt werden

oberegrenzevektor<-rep(0,w)

# Hier muss etwas ergaenzt werden


# Stelle graphisch dar
# zunaechst nur 100 Konfidenzintervalle, damit man es
besser erkennen kann:

plot(c(min(unteregrenzevektor),max(oberegrenzevektor))
,c(0,100),type="n",xlab="Konfidenzintervalle",ylab="",
main="Überdeckungen der Konfidenzintervalle")

for(i in 1:100)
     {

lines(c(unteregrenzevektor[i],oberegrenzevektor[i]),c(
i,i),lwd=1)
     }

abline(v=1/3,lwd=2,col="green")

# jetzt alle:
plot(c(min(unteregrenzevektor),max(oberegrenzevektor))
,c(0,w),type="n",xlab="Konfidenzintervalle",ylab="",ma
in="Überdeckungen der Konfidenzintervalle")

for(i in 1:w)
     {

lines(c(unteregrenzevektor[i],oberegrenzevektor[i]),c(
i,i),lwd=0.2)
     }

abline(v=1/3,lwd=2,col="green")


# relativer Anteil, der wahres p ueberdeckt:

ueberdeckt<-0

# Hier muss etwas ergaenzt werden

# relativer Anteil:
# Hier muss etwas ergaenzt werden

# Mittlere Laenge der Konfidentintervalle

# Hier muss etwas ergaenzt werden. Speichern Sie die
Längen in einem vektor laenge.

######################################################
######################
# Teil iii

N<-seq(from=10,to=10000,by=10)
l<-length(N)

laengeN<-rep(0,l) # Hier sollen die mittleren Laengen
gespeichert werden

for(j in 1:l)
     {

     # Hier muss etwas ergaenzt werden

     }

plot(
# Hier muss etwas ergaenzt werden
,main="Log-log plot: Länge der Konfidenzintervalle bei
verschiedenen N")

# Zum Vergleich: Gerade mit Steigung m (in der Aufgabe
nicht gefordert!)
m<-0.3
```

```
y<-rep(0,l)
for(i in 1:l)
{
y[i]<-m*log(N[i])-0.5
}
lines(log(N),y)
```

2. Lösungen der „R Aufgaben“

```
# Kopf oder Zahl (1. Aufgabe)

#############################################################
#############
# Allgemeines zu R:

# Kommentare werden mit "#" eingeleitet.

# R kann man als Taschenrechner benutzen:
3+4
1/2
2^3
exp(10)

# Zuordnungen macht man mit "<-"
# Vektoren: Um mehrere Zahlen zu einem Vektor
zusammenzufuegen gibt es den Befehl c():
a<-c(5,6,7)
# den Vektor a anschauen:
a

# Die Laenge von a (Anzahl der Eintraege):
length(a)

# Ein Vektor, der aus einer aufsteigenden Folge 1 bis 3
besteht:
b<-c(1:3)
b

# Rechnungen mit Vektoren werden komponentenweise
ausgefuehrt:
a+b
3*a

# Logische Abfragen (koennen wahr oder falsch sein):
3 < 5
4 >= 10
3 == 7 # doppeltes "=" !
7 == 7

# Auf Elemente eines Vektors zugreifen:
# Der 3. Eintrag von a:
a[3]
# Alle Eintraege von a, die groesser oder gleich 6 sind:
a[a>=6]

# Matrizen:
mat<-matrix(c(3,4,5,6,7,8),3,2) # 3 Zeilen und 2 Spalten
mat
# auf ein Element zugreifen:
mat[1,2] # Element in 1. Zeile und 2. Spalte

#############################################################
#############
#############################################################
#############
```

```
# Alle vorherigen Zuordnungen loeschen:
rm(list=ls())

#############################################################
#############
# Wir simulieren einen Muenzwurf:

# Die Muenze wird 40 Mal geworfen.
n<-40

# "sample" waehlt aus einer Menge eine bestimmte Anzahl aus
(mit oder ohne Zuruecklegen). Wenn man sich die Hilfe zu
dieser Funktion anschauen moechte, gibt man ein:
?sample

# Wir waehlen aus der Menge {"Kopf","Zahl"} n Werte aus. Die
Ergebnisse werden in "spiel" gespeichert.
spiel<-sample(c("Kopf","Zahl"),n,replace=T)

# Rufe "spiel" auf, um sich die Eintraege anzuschauen:
spiel

# Gewinne und Verluste von Peter: Bei "Kopf" gewinnt Peter
einen Euro (+1), bei Zahl verliert er einen Euro (-1). Wir
erstellen einen neuen Vektor namens "peter" und benutzen
dafuer eine "for"-Scheife und eine "if"-Abfrage.

peter<-rep(0,n) # neuen Vektor vorbereiten

for(i in 1:n)
     {
      if(spiel[i]=="Kopf"){peter[i]<-1}else{peter[i]<--1}
     }


# Wieviel Geld gewinnt (oder verliert) Peter insgsamt? Wir
addieren alle Eintraege von "peter" und speichern das
Ergebnis in "gewinnpeter".
gewinnpeter<-sum(peter)
gewinnpeter

#############################################################
#############
# Eine eigene Funktion definieren: Man kann in R recht
einfach eigene Funktionen definieren. In unserem Beispiel ist
es nicht unbedingt noetig, wir zeigen aber trotzdem mal, wie
das geht. Nach dem Befehl "function" gibt man in runden
Klammern die Input-Variablen an. Falls man keine Input-
Variablen verlangen möchte, bleiben die runden Klammern
einfach leer.

# Unsere Funktion soll das Spiel "Kopf oder Zahl" simulieren
und den Spielverlauf sowie die Gewinne von Peter in der
Konsole ausgeben.
```

```
kopfoderzahl<-function(n)
{
 readline("neues Spiel")
 # Um zu starten, muss man die Konsole aktivieren (darauf
klicken) und "Enter" druecken. Der "readline"-Befehl ist oft
nuetzlich, deshalb wird er hier gezeigt.
 n<<-n
 spiel<<-sample(c("Kopf","Zahl"),n,replace=T)

 # Zuordnung innerhalb einer Funktion mit "<<-", wenn man
auch spaeter noch darauf zurueckgreifen koennen moechte.
 peter<<-rep(0,n)
 for(i in 1:n)
     {
      if(spiel[i]=="Kopf"){peter[i]<<-1}else{peter[i]<<--1}
     }
 gewinnpeter<<-sum(peter)
 print(spiel)
 print(paste("Gewinn von Peter:",gewinnpeter))
}

# Funktionsaufruf für n=40:
kopfoderzahl(40)

# Gewinne und Verluste von Peter in einem Spielverlauf
gewinnverlauf<-function()
{
verlauf<<-rep(0,n)

for(i in 1:n)
     {
      verlauf[i]<<-sum(peter[1:i])
      # Die Eintraege 1 bis i von "peter" werden aufsummiert
und an die i-te Stelle des neuen Vektors "verlauf"
geschrieben.
     }
}

gewinnverlauf()
verlauf

##############################################################
#############
# Nun sollen Sie das Spiel mit 40 Wuerfen 10 Mal wiederholen.
Tipp: Am einfachsten geht es mit der Funktion kopfoderzahl,
die wir gerade definiert haben, und einer "for"-Schleife.


n<-40
w<-10
ergebnissepeter<-rep(0,w)

for(i in 1:w)
{
kopfoderzahl(n)
ergebnissepeter[i]<-gewinnpeter
}

ergebnissepeter

##############################################################
#############
# Graphische Darstellung der Gewinne und Verluste von Peter
in einem Spielverlauf


plot(c(1:n),verlauf,type="l",main="Gewinne und Verluste von
Peter im
Spielverlauf",ylim=c(min(verlauf),max(verlauf)),xlab="Wurfnum
mer",ylab="Gewinn/Verlust")

abline(h=0)


##############################################################
#############
# Graphische Darstellung der Gewinne von Peter am Ende der
Spiele
# Hier tricksen wir ein wenig, um ein Stabdiagramm zu
erhalten:

# Zunaechst erzeugen wir w Wiederholungen des Spiels. Wenn w
groß ist, ist es sinnvoll, den "readline"-Befehl und die
"print"-Befehle in der Funktion kopfoderzahl()
auszukommentieren. Das koennen Sie hier machen:

kopfoderzahl<-function(n)
{
 #readline("neues Spiel")
 n<<-n
 spiel<<-sample(c("Kopf","Zahl"),n,replace=T)
 peter<<-rep(0,n)
 for(i in 1:n)
     {
      if(spiel[i]=="Kopf"){peter[i]<<-1}else{peter[i]<<--1}
     }
 gewinnpeter<<-sum(peter)
 #print(spiel)
 #print(paste("Gewinn von Peter:",gewinnpeter))
}


# Jetzt wiederholen wir das Spiel w Mal und speichern die
Ergebnisse in ergebnissepeter:

w<-10000
ergebnissepeter<-rep(0,w)

for(i in 1:w)
{
kopfoderzahl(40)
ergebnissepeter[i]<-gewinnpeter
```

```
}

breaks<-seq(from=-n,to=n,by=0.1) # gibt die Breite der Stäbe
an
hist(ergebnissepeter,breaks=breaks,main=paste("Peters
Gewinnsummen bei", w, "Wiederholungen"),xlab="Peters
Gewinnsummen",ylab="Absolute
Haeufigkeiten",freq=T,col="black")

############################################################
#############
```

```
# Crabs (2. Aufgabe)

# Würfelwurf
w<-sample(1:6,2,replace=T)
w

# Augensumme
x<-sum(w)

######################################################
####################
# a) Einzelne Spiel zum Anschauen: hier wurden
doppelte Pfeilspitzen ergaenzt!!


spiel<-function()
{
ergebnis<<-"leer"
ersterwurf<-sample(1:6,2,replace=T)
x<-sum(ersterwurf)
readline("Neues Spiel")
readline(paste("Erster Wurf: X=",x))

# Spielende nach einer Runde
if(x==7|x==11)
     {
      ergebnis<<-"G"
      readline("Spieler gewinnt")
     }
if(x==2|x==3|x==12)
     {
      ergebnis<<-"V"
      readline("Spieler verliert")
     }

# Spielende erst später
while(ergebnis=="leer")
     {
      neuerwurf<-sample(1:6,2,replace=T)
      neuesumme<-sum(neuerwurf)
      readline(paste("Neuer Wurf: Summe:",neuesumme))

      if(neuesumme==x)
          {
           ergebnis<<-"G"
           readline("Spieler gewinnt")
          }
      if(neuesumme==7)
          {
           ergebnis<<-"V"
           readline("Spieler verliert")
          }
      }
}

for(i in 1:10)
     {
      spiel()
     }

######################################################
#######################
# b) Gewinnws aus n Spielen schätzen

n<-10000

ergebnisvektor<-rep("leer",n)
#Name "ergebnisvektor" besser, damit es keine
Ueberschneidung mit der Variablen "ergebnis" in der
Funktion "spiel" gibt

#Moeglichkeit 1 mit Funktion "spiel", aber ohne
readline-Befehle:

spiel<-function()
{
ergebnis<<-"leer"
ersterwurf<-sample(1:6,2,replace=T)
x<-sum(ersterwurf)

# Spielende nach einer Runde
if(x==7|x==11)
     {
      ergebnis<<-"G"
     }
if(x==2|x==3|x==12)
     {
      ergebnis<<-"V"
     }

# Spielende erst später
while(ergebnis=="leer")
     {
      neuerwurf<-sample(1:6,2,replace=T)
      neuesumme<-sum(neuerwurf)

      if(neuesumme==x)
          {
           ergebnis<<-"G"
          }
      if(neuesumme==7)
          {
           ergebnis<<-"V"
          }
      }
}

for(i in 1:n)
{
spiel()
```

```
ergebnisvektor[i]<-ergebnis
}

# Moeglichkeit 2 ohne Funktion "spiel"

ergebnisvektor<-rep("leer",n)
for (i in 1:n)
     {
      ersterwurf<-sample(1:6,2,replace=T)
      x<-sum(ersterwurf)

      # Spielende nach einer Runde
      if(x==7|x==11){ergebnisvektor[i]<-"G"}
      if(x==2|x==3|x==12){ergebnisvektor[i]<-"V"}

      # Spielende erst später
      while(ergebnisvektor[i]=="leer")
           {
            neuerwurf<-sample(1:6,2,replace=T)
            neuesumme<-sum(neuerwurf)

            if(neuesumme==x){ergebnisvektor[i]<-"G"}
            if(neuesumme==7){ergebnisvektor[i]<-"V"}
           }
     }



# Wie oft gewinnt der Spieler?
ergebnisnum<-rep(0,n)
for(i in 1:n)
     {
      if(ergebnisvektor[i]=="G"){ergebnisnum[i]<-
1}else{ergebnisnum[i]<-0}
     }

sum(ergebnisnum)/n

######################################################
#######################
# c) Vergleich mit fairem Spiel

k<-10000
fair<-rbinom(k,100,0.5)
crabs<-rbinom(k,100,0.493)

min<-min(fair,crabs)
max<-max(fair,crabs)

breaks<-seq(from=min,to=max,by=0.2)

# Faires Spiel

hist(fair,breaks=breaks,border="red",col="red",main="V
ergleich von Crabs mit fairem Spiel: Anzahl Gewinne
bei 100
Spielrunden",cex.main=0.9,ylab=paste("Häufigkeit
bei",k,"Simulationen"),xlab="Anzahl Gewinne")
text(min,max(table(fair)),"faires
Spiel",col="red",adj=0)


# Faires Spiel und Crabs vergleichen: kann man
praktisch nicht unterscheiden
hist(fair,breaks=breaks,border="red",col="red",main="V
ergleich von Crabs mit fairem Spiel: Anzahl Gewinne
bei 100
Spielrunden",cex.main=0.9,ylab=paste("Häufigkeit
bei",k,"Simulationen"),xlab="Anzahl Gewinne")

hist(crabs,breaks=breaks,border="blue",add=T)

text(min,max(table(fair)),"faires
Spiel",col="red",adj=0)
text(min,max(table(fair))*0.95,"Crabs",col="blue",adj=
0)
```

```
# Monte Carlo (3. Aufgabe)

# Ein Punkt
abstand<-function(x,y)
{
abstand<-sqrt((x-0.5)^2+(y-0.5)^2)
abstand
}

x<-runif(1,min=0,max=1)
y<-runif(1,min=0,max=1)

# im Kreis?
imkreis<-0
if(abstand(x,y)<=0.5){imkreis<-imkreis+1}

#####
# Wiederhole w Mal
montecarlo<-function(w,print)
{
imkreis<-rep(0,w)
x<-runif(w,min=0,max=1)
y<-runif(w,min=0,max=1)
for(i in 1:w)
     {
      if(abstand(x[i],y[i])<=0.5){imkreis[i]<-
imkreis[i]+1}
     }

# Schaetzung für Flaecheninhalt
schaetzerflaeche<<-sum(imkreis)/w
if(print==1){print(paste("Schaetzer fuer den
Flaecheninhalt:", schaetzerflaeche))}

# Schaetzung für pi: exakter Flaecheninhalt ist pi/4
schaetzerpi<<-schaetzerflaeche*4
if(print==1){print(paste("Schaetzer fuer pi:",
schaetzerpi))}
if(print==1){print(paste("Wahrer Wert von pi:", pi))}
}

montecarlo(10000,1)

#####################################################
###################

# Verfahren 100 Mal wiederholen: Bereich für pi-
Schätzungen

schaetzerpivektor<-rep(0,100)

for(j in 1:100)
{
montecarlo(10000,2)
schaetzerpivektor[j]<-schaetzerpi
}

# Bereich, in dem Schätzungen liegen:
min(schaetzerpivektor)
max(schaetzerpivektor)

# oder
summary(schaetzerpivektor)

# Also: Schaetzung bis auf eine Nachkommastelle genau
(bis auf wenige Ausnahmen, die gerundet werden
muessen)

###################################

# Wieviele Punkte braucht man, um bis auf 2
Nachkommastellen genau zu schaetzen?

# Wiederhole w Mal
montecarlo(1000000,1)
```

```
 # Belegzeiten von Betten (4. Aufgabe)


rm(list=ls())

#######################################################
#######################
# Belegungszeit von Betten
# exp(0.1)-verteilt => MW=10
# Studie: In einem großen neu eröffnetem KH mit 800
Betten soll nach 100 Tagen untersucht werden, wie
lange Patienten im KH liegen. Dazu wird an diesem Tag
für jedes Bett notiert, welcher Patient darin liegt
und nach der Entlassung aufgeschrieben, wie lange
dieser Patient das Bett belegt hat.

#######################################################
########################

# Matrix vorbereiten
mat<-matrix(rep(0,800*50),nrow=800)

# Simuliere für 800 Betten die Liegezeiten von 50
Patienten

n<-800
a<-50
for(j in 1:n)
     {
      #Ausreichende Anzahl exp(0.1)-verteilter ZV
erzeugen, hier 50 Patienten
      x<-rexp(a,0.1)
      mat[j,]<-x
     }

#######################################################
########################
# Wie lange hat Patient, der zum Zeitpunkt t=100 im
Bett lag, dieses belegt?

liege100<-rep(0,800)

for(j in 1:n)
     {
      index<-0
      z<-0
      for(i in 1:50)
          {
           if(z<100)
               {
                z<-z+mat[j,i] # Aufsummierien, bis Summe
groesser ist als 100
                index<-i # Index, bei dem 100
ueberschritten wird (Stoppzeit)
               }
          }
      # Wert der (Stoppzeit)ten ZV ist gesuchte
Intervall-Laenge
      liege100[j]<-mat[j,index]
     }

# Mittelwert:
mean(liege100) # Die Schaetzmethode überschaetzt die
mittlere Liegedauer!

#######################################################
########################
# Histogramm der Liegezeiten der Patienten, die zum
Zeitpunkt t=100 im Bett lagen

breaks<-seq(from=0,to=max(liege100)+1,by=1) # Breite
der Säulen im Histogramm
hist(liege100,breaks=breaks,freq=FALSE,col="yellow",yl
im=c(0,0.15),main="Wie lange belegten die Patienten
ihr Bett, die zum Zeitpunkt t=100 darin
lagen?",cex.main=1,ylab="",xlab="Liegezeit der
Patienten")

text(9,0.9,expression(Urspruengliche~Zeiten~exponentia
l~verteilt~zum~Parameter~0.1),cex=1.1)

#Zum Vergleich: exp(0.1)-Dichte einzeichnen:
l<-length(breaks)
c<-rep(0,l)
for(i in 1:l){c[i]<-0.1*exp(-0.1*breaks[i])}  #
Funktionswerte der Dichtefunktion
lines(breaks,c,lwd=1.8)

# => Groessenverzerrung bei dieser Schaetzmethode, man
braucht anderes Verfahren.

#######################################################
########################
# Ende
graphics.off()
```

```
# Fortsetzung: Belegzeiten von Betten (5. Aufgabe)


rm(list=ls())

######################################################
######################
# Belegungszeit von Betten
# exp(0.1)-verteilt => MW=10
# Studie: In einem großen neu eröffnetem KH mit 800
Betten soll nach 100 Tagen untersucht werden, wie
lange Patienten im KH liegen. Dazu wird an diesem Tag
für jedes Bett notiert, welcher Patient darin liegt
und nach der Entlassung aufgeschrieben, wie lange
dieser Patient das Bett belegt hat.

######################################################
#######################

# Matrix vorbereiten
mat<-matrix(rep(0,800*50),nrow=800)

# Simuliere für 800 Betten die Liegezeiten von 50
Patienten

n<-800

for(j in 1:n)
     {
      #Ausreichende Anzahl exp(10)-verteilter ZV
erzeugen, hier 50 Patienten
      x<-rexp(50,0.1)
      mat[j,]<-x
     }

######################################################
#######################
# Wie lange hat Patient, der zum Zeitpunkt t=100 im
Bett lag, dieses belegt?

liege100<-rep(0,800)
indexvek<-rep(0,800)

for(j in 1:n)
     {
      z<-0
      for(i in 1:50)
          {
           if(z<100)
               {
                z<-z+mat[j,i] # Aufsummierien, bis Summe
groesser ist als 100
                indexvek[j]<-i # Index, bei dem 100
ueberschritten wird (Stoppzeit)
                }
          }
      # Wert der (Stoppzeit)ten ZV ist gesuchte
Intervall-Laenge
      liege100[j]<-mat[j,indexvek[j]]
     }

# Index
indexvek

######################################################
#######################
# Wie kommt es zu der Verzerrung? Längere Liegezeiten
haben höhere Ws getroffen zu werden:

plot(c(0,120),c(0,10),type="n",main="Belegzeiten
verschiedener Betten", ylab="",xlab="Tage",axes=F)
axis(1)
box()
col<-rainbow(max(indexvek)+1,alpha=0.5)

for(j in 1:10)
     {
      polygon(c(0,mat[j,1],mat[j,1],0),c(j-0.3,j-
0.3,j+0.3,j+0.3),col=col[1])
      for(i in 1:indexvek[j])
          {

polygon(c(sum(mat[j,1:i]),sum(mat[j,1:(i+1)]),sum(mat[
j,1:(i+1)]),sum(mat[j,1:i])),c(j-0.3,j-
0.3,j+0.3,j+0.3),col=col[i+1])
          }
     }

abline(v=100,col="red",lwd=3)

######################################################
#######################
# andere Methoden

# Wähle zufällig Patienten aus, Bsp: notiere vom
nächsten Patienten (der also nach Entlassung dessen,
der zum Zeitpunkt t=100 das Bett belegt hat) die
Liegezeit

liege<-rep(0,800)
for(i in 1:800)
     {
      liege[i]<-mat[i,indexvek[i]+1]
     }

mean(liege)


#ODER je Bett, wähle rein zufällig Patient aus (dann
müssten aber von allen Patienten bis t=100 die
Liegezeiten notiert worden sein) oder
```

```
(aequivalent)wähle einfach rein zufällig 800 Patienten
aus


liege<-rep(0,800)
for(i in 1:800)
     {
      s<-sample(1:indexvek[i],1)
      liege[i]<-mat[i,s]
     }

mean(liege)

######################################################
#################

# Histogramm der Belegzeiten Betten
breaks<-seq(from=0,to=max(liege)+1,by=1) # Breite der
Säulen im Histogramm
hist(liege,breaks=breaks,freq=FALSE,col="yellow",ylim=
c(0,0.15),main="Belegzeiten der
Betten",xlab="Liegezeiten von 800 zufälligen
Patienten",ylab="")

#Zum Vergleich: exp(0.1)-Dichte einzeichnen:
l<-length(breaks)
c<-rep(0,l)
for(i in 1:l){c[i]<-0.1*exp(-0.1*breaks[i])}  #
Funktionswerte der Dichtefunktion
lines(breaks,c,lwd=1.8)


######################################################
#######################

# Ende
graphics.off()
```

```
# Vorzeichenwechsel beim Münzwurf (6. Aufgabe)

############################################################
############
# Alle vorherigen Zuordnungen loeschen:
rm(list=ls())

############################################################
############

# Wir untersuchen zunaechst einen fairen Muenzwurf

n<-20
p<-0.5

# Zeichne den Verlauf von 3 fairen Muenzwuerfen:

wuerfegraphik<-function(n,p)
{
par(mfrow=c(3,1))

for(j in 1:3)
     {
      wurf<<-sample(c("K","Z"),n,replace=T,prob=c(p,1-p))

      vorzwechsel<<-0
      for(i in 1:(n-1))
          {
           if(wurf[i]!=wurf[i+1]){vorzwechsel<<-vorzwechsel+1}
          }

     wurfnum<<-rep(0,n)
     for(i in 1:n){if(wurf[i]=="K"){wurfnum[i]<<-
1}else{wurfnum[i]<<--1}}

     readline(paste("Muenzwurf Nr",j,"p=",p))

     plot(wurfnum,type="p",pch=16,main=paste("Muenzwurf",j,"(
p=",p,")"),axes=F,ylim=c(-4,3), xlim=c(-
2/n,n),ylab="",xlab="")

     abline(h=0)
     text(0,1,"Kopf")
     text(0,-1,"Zahl")
     text(n/2,-3,paste("Anzahl Vorzeichenwechsel:",
vorzwechsel))
     }
}

wuerfegraphik(n,p)

############################################################
############
# Wiederhole haeufig und zaehle Vorzeichenwechsel:

wurffunktion<-function(p)
```

```
{
      wurf<<-sample(c("K","Z"),n,replace=T,prob=c(p,1-p))

      vorzwechsel<<-0
      for(i in 1:(n-1))
          {
           if(wurf[i]!=wurf[i+1]){vorzwechsel<<-vorzwechsel+1}
          }
}

w<-10000
vorzwechselvektor<-rep(0,w)

for(i in 1:w)
     {
      wurffunktion(p)
      vorzwechselvektor[i]<-vorzwechsel
     }


# Zeichne Stabdiagramm:

par(mfrow=c(1,1))
breaks<-seq(from=0,to=max(vorzwechselvektor),by=0.01)
histvorzwechsel<-
hist(vorzwechselvektor,breaks=breaks,main=paste("Anzahl
Vorzeichenwechsel bei",n,"p=", p, "-
Muenzwuerfen"),ylab="Absolute Hauefigkeit",xlab="Anzahl
Vorzeichenwechsel",xlim=c(1,max(vorzwechselvektor)))

# Erwartete Anzahl Vorzeichenwechsel
erwartung<-mean(vorzwechselvektor)

abline(v=erwartung,lwd=2,col="green")

text(1,max(histvorzwechsel$counts)-
0.1*max(histvorzwechsel$counts),"erwartete Anzahl
Vorzeichenwechsel",col="green",cex=0.8,adj=0)

############################################################
############
# Vergleich mit Vorzeichenwechsel aus gefaelschtem Muenzwurf:

#Anzahl Vorzeichenwechsel aus gefaelschtem Munezwurf
eintragen, Bsp: gefaelscht=10
gefaelscht <-10

par(mfrow=c(1,1))
breaks<-seq(from=0,to=max(vorzwechselvektor),by=0.01)
histvorzwechsel<-
hist(vorzwechselvektor,breaks=breaks,main=paste("Anzahl
Vorzeichenwechsel bei",n,"p=", p, "-
Muenzwuerfen"),ylab="Absolute Hauefigkeit",xlab="Anzahl
Vorzeichenwechsel",xlim=c(1,max(vorzwechselvektor)))
abline(v=erwartung,lwd=2,col="green")
```

```
text(1,max(histvorzwechsel$counts)-
0.1*max(histvorzwechsel$counts),"erwartete Anzahl
Vorzeichenwechsel",col="green",cex=0.8,adj=0)


text(1,max(histvorzwechsel$counts)-
0.2*max(histvorzwechsel$counts),paste("Vorzeichenwechsel bei
gefaelschtem Muenzwurf",
gefaelscht),col="blue",cex=0.8,adj=0)

points(gefaelscht,0,pch=16,col="blue",cex=2)


##############################################################
############
# Wir wollen untersuchen, wie hauefig eine so grosse
Abweichung vom (geschaetzten) Erwartungswert bei fairem
Muenzwurf vorkommt.

# Abweichung vom Erwartungswert
abw<-abs(gefaelscht-erwartung)

par(mfrow=c(1,1))
breaks<-seq(from=0,to=max(vorzwechselvektor),by=0.01)
histvorzwechsel<-
hist(vorzwechselvektor,breaks=breaks,main=paste("Anzahl
Vorzeichenwechsel bei",n,"p=", p, "-
Muenzwuerfen"),ylab="Absolute Hauefigkeit",xlab="Anzahl
Vorzeichenwechsel",xlim=c(1,max(vorzwechselvektor)))

abline(v=erwartung,lwd=2,col="green")

text(1,max(histvorzwechsel$counts)-
0.1*max(histvorzwechsel$counts),"erwartete Anzahl
Vorzeichenwechsel",col="green",cex=0.8,adj=0)

text(1,max(histvorzwechsel$counts)-
0.2*max(histvorzwechsel$counts),paste("Vorzeichenwechsel bei
gefaelschtem Muenzwurf",
gefaelscht),col="blue",cex=0.8,adj=0)

points(gefaelscht,0,pch=16,col="blue",cex=2)


text(1,max(histvorzwechsel$counts)-
0.3*max(histvorzwechsel$counts),"mind. so große Abweichung
vom Erwartungswert",col="red",cex=0.8,adj=0)

hist(vorzwechselvektor[vorzwechselvektor>=erwartung+abw],brea
ks=breaks,add=T,col="red",border="red")

hist(vorzwechselvektor[vorzwechselvektor<=erwartung-
abw],breaks=breaks,add=T,col="red",border="red")

abline(h=0,col="black",lwd=1.5)
```

```
#######

#Bsp: Wie oft gibt es so grosse Abweichung vom (geschaetzten)
Erwartungswert?
step<-seq(from=0,to=length(histvorzwechsel$counts),by=100)
abshauefigkeitenvorzwechsel<-
c(histvorzwechsel$counts[1],histvorzwechsel$counts[step])

ah<-
sum(abshauefigkeitenvorzwechsel[(floor(erwartung+abw)+1):leng
th(abshauefigkeitenvorzwechsel)])+sum(abshauefigkeitenvorzwec
hsel[1:ceiling(erwartung-abw)])


# Schaetzer für Ws, dass es mehr als 10 Vorzeichenwechsel
gibt
ah/w


##############################################################
################
##############################################################
################

# Erwartungswert beim p- Muenzwurf
p<-0.1
w<-10000
vorzwechselvektor<-rep(0,w)

for(i in 1:w)
     {
      wurffunktion(p)
      vorzwechselvektor[i]<-vorzwechsel
     }

mean(vorzwechselvektor)


# Abhaengigkeit von p

w<-10000
vorzwechselmatrix<-matrix(rep(0,w*11),nrow=11)
pvektor<-seq(from=0,to=1,by=0.1)

for(j in 1:11)
     {
      for(i in 1:w)
           {
            wurffunktion(pvektor[j])
            vorzwechselmatrix[j,i]<-vorzwechsel
           }
     }

# Erwartungswerte in Abhaengigkeit von p
erwartungsvektor<-rep(0,11)
```

```
for(j in 1:11)
     {
      erwartungsvektor[j]<-mean(vorzwechselmatrix[j,])
     }


# Plot Erwartungswerte
plot(pvektor,erwartungsvektor,type="p",pch=16,main="Erwartete
Anzahl Vorzeichenwechsel als Funktion von p",ylab="Erwartete
Anzahl Vorzeichenwechsel",xlab="p",ylim=c(0,10))

text(pvektor[1],max(erwartungsvektor),"simulierte
Erwartungswerte",adj=0,cex=0.8)

formelerw<-function(p,n)
{
erw<-p*(1-p)*(n-1)*2
erw
}

x<-seq(from=0,to=1,by=0.01)
l<-length(x)
y<-rep(0,l)

for(i in 1:l){y[i]<-formelerw(x[i],n)}

lines(x,y,col="red") # tatsächliche Erwartungswerte
```

```
# Regression zur Mitte (7. Aufgabe)
rm(list=ls())

##############################################################
###############
# Teil i
# Um die Daten (X,Y) zu simulieren, gibt es zwei
Moeglichkeiten. Am einfachsten geht es mit der Funktion
mvrnorm aus dem Paket MASS. Falls Sie diese Funktion nicht
zur Verfuegung haben, waeheln Sie bitte Variante 2.

############
# Variante 1 mit mvrnorm
library(MASS)
n<-10000
kappa<-0.7
varcov<-matrix(c(64,kappa*64,kappa*64,64),nrow=2,byrow=T)
daten<-mvrnorm(n,c(82,82),varcov)
gewicht1<-daten[,1]
gewicht2<-daten[,2]


# Variante 2
n<-10000
kappa<-0.7

x1<-rnorm(n)
x2<-rnorm(n)

gewicht1<-x1*8+82
gewicht2<-(kappa*x1+sqrt(1-kappa^2)*x2)*8+82

############

# Vergewissern Sie sich, dass die Daten richtig erzeugt
wurden! Kontrollieren Sie die Mittelwerte (mit dem Befehl
mean), die Standardabweichungen (mit dem Befehl sd) und die
Korrelation (mit dem Befehl cor)

mean(gewicht1)
mean(gewicht2)
sd(gewicht1)
sd(gewicht2)
cor(gewicht1,gewicht2)

############
# Zeichnen Sie ein Streudiagramm. Sie brauchen den Graphik-
Befehl plot.

plot(gewicht1,gewicht2,main="Gewichte vorher und
nachher",xlab="Gewicht vorher",ylab="weG")


##############################################################
###############
# Teil ii
# Wir speichern die Indizes, bei denen in gewicht1 Zahlen
kleiner als 100 stehen, in einem Vektor index ab. Machen Sie
sich klar, was durch index<-c(index,i) in der for-Schleife
passiert.

index<-0

for(i in 1:n)
{
 if(gewicht1[i]<=100){index<-c(index,i)}
}

gewicht1vonSchweren<-gewicht1[-index]
gewicht2vonSchweren<-gewicht2[-index]

# Was ist der Mittelwert des zweiten Gewichts der vorher
schweren Maenner?
mean(gewicht1vonSchweren)
mean(gewicht2vonSchweren)

# Hier werden die Histogramme erzeugt

ming<-min(gewicht1vonSchweren,gewicht2vonSchweren)
maxg<-max(gewicht1vonSchweren,gewicht2vonSchweren)

breaks<-seq(from=min(75,ming-5),to=maxg+5,by=5)

hist(gewicht1vonSchweren,breaks=breaks,col="#ff000099",main="
Gewichte der Schweren vorher und nachher (rot
vorher)",xlab="")
hist(gewicht2vonSchweren,breaks=breaks,xlab="",col="#0000ff44
",add=T)
abline(v=82,lwd=2,col="green")


##############################################################
###############
# Teil iii
# entweder zusätzlich eine Kontrollgruppe, die Radiosender
beibehält, oder schwere und leichte Personen bei
Radiosenderwechsel beobachten: Mittelwert vorher und nachher
gleich
```

```
# Rejection Sampling (8. Aufgabe)


# Alle vorherigen Zuordnungen loeschen:
rm(list=ls())

# Umrechnung in kartesische Koordinaten (%*% bildet
Skalarprodukt)
# Frankfurt
ffm<-c(cos(0.14)*cos(0.87),cos(0.14)*sin(0.87),sin(0.14))
einffm<-ffm/sqrt(ffm%*%ffm)


# Kapstadt
kap<-c(cos(0.31)*cos(-0.6),cos(0.31)*sin(-0.6),sin(0.31))
einkap<-kap/sqrt(kap%*%kap)


# Peking
pek<-c(cos(2.02)*cos(0.7),cos(2.02)*sin(0.7),sin(2.02))
einpek<-pek/sqrt(pek%*%pek)


# Uniforme Punkte auf Kugelsspäre:
# 1. uniforme Punkte in Einheitswuerfel

n<-10000

x<-runif(n,-1,1)
y<-runif(n,-1,1)
z<-runif(n,-1,1)

# in Einheitskugel? Wenn nein, verwerfe Punkt
inkugel<-0
index<-rep(0,n)

for(i in 1:n)
     {
      if(sqrt((x[i]-0)^2+(y[i]-0)^2+(z[i]-0)^2)<=1)
          {
           inkugel<-inkugel+1
          }
      else
          {
           index[i]<-i
          }
     }


# Nur Punkte, die in Einheitskugel liegen:
xkug<-x[-index]
ykug<-y[-index]
zkug<-z[-index]

l<-length(xkug)
```

```
# 2. Projektion auf Kugeloberfläche
norm<-rep(0,l)

for(i in 1:l)
{
help<-c(xkug[i],ykug[i],zkug[i])
norm[i]<-sqrt(help%*%help)
}

xproj<-rep(0,l)
yproj<-rep(0,l)
zproj<-rep(0,l)

for(i in 1:l)
{
xproj[i]<-xkug[i]/norm[i]
yproj[i]<-ykug[i]/norm[i]
zproj[i]<-zkug[i]/norm[i]
}


# Naehere Flaecheninhalt des Kugeldreiecks an

imkugeldreieck<-0

for(i in 1:l)
{
help<-c(xproj[i],yproj[i],zproj[i])
mat1<-matrix(c(einffm,einkap,help),nrow=3)
mat2<-matrix(c(einffm,help,einpek),nrow=3)
mat3<-matrix(c(help,einkap,einpek),nrow=3)
if(det(mat1)<=0 & det(mat2)<=0 &
det(mat3)<=0){imkugeldreieck<-imkugeldreieck+1}
}

# Schaetzer fuer relativen Anteil der Flaeche
imkugeldreieck/l
```

```
# Fehlstaende rein zufaelliger Permutationen (9. Aufgabe)

##############################################################
################
rm(list=ls())

##############################################################
################
# Teil i

# rein zufaellige Permutation der Laenge 5
n<-5
p<-sample(1:n,replace=F)
p

# Fehlstaendevektor y

# y2 = y[1] = alle Fehlstaende, an denen die 2 zusammen mit
kleinerem Partner beteiligt ist usw

y<-rep(0,n-1)

for(j in 2:n)
     {
      for(i in 1:(j-1))
            {
             if(p[i]>p[j]){y[j-1]<-y[j-1]+1}
            }
     }

sum(y)

# Beh: yj(X) ist uniform auf {0,...,j-1} verteilt

#

w<-10000

ymat<-matrix(rep(0,((n-1)*w)),ncol=(n-1))

for(m in 1:w)
{
p<-sample(1:n,replace=F)
y<-rep(0,n-1)

for(j in 2:n)
     {
      for(i in 1:(j-1))
            {
             if(p[i]>p[j]){y[j-1]<-y[j-1]+1}
            }
     }
ymat[m,]<-y
}
```

```
# Histogramme

par(mfrow=c(2,2))

for(i in 1:(n-1))
{
hist(ymat[,i],breaks=seq(from=0,to=i,by=0.01),border="green",
main=paste("Fehlstaende, an denen die", i+1, "mit kleinerem
Partner beteiligt"),cex.main=0.8,xlab="Anzahl
Fehlstaende",ylab=paste("Haeufigkeit
bei",w,"Wdh"),cex.lab=0.8)
}

# Bezug zur Aufgabe 35: jeweils annaehernd uniform auf {0,j-
1}

##############################################################
################

# Histogramm der Gesamtanzahlen der Fehlstaende bei w
Permutationen fuer n=5, n=10 und n=15

n<-15 # hier kann man den Wert von n veraendern!
w<-10000

ymat<-matrix(rep(0,((n-1)*w)),ncol=(n-1))

for(m in 1:w)
{
p<-sample(1:n,replace=F)
y<-rep(0,n-1)

for(j in 2:n)
     {
      for(i in 1:(j-1))
            {
             if(p[i]>p[j]){y[j-1]<-y[j-1]+1}
            }
     }
ymat[m,]<-y
}


gesamt<-rep(0,w)

for(i in 1:w){gesamt[i]<-sum(ymat[i,])}

breaks<-seq(from=0,to=max(gesamt)+10,by=1)
hist(gesamt,breaks=breaks,ylim=c(0,0.25),freq=F)

mu<-n*(n-1)/4
sigma<-sqrt((2*n^3+3*n^2-5*n)/72)


x<-seq(0,max(gesamt)+10,0.1)
```

```
l<-length(x)
c<-rep(0,l)
for(i in 1:l)
{
c[i]<-1/(sqrt(2*pi)*sigma)*exp(-((x[i]-mu)^2/(2*sigma^2)))
}

lines(x,c,lwd=2,col="red")

# Man scheint hier eine Art Zentralen Grenzwertsatz zu haben.
Wir haben eine Summe von ZV, die zwar unabhängig sind, aber
nicht identisch verteilt (das koennte einen wundern)!
```

```
# Konfidenzintervalle (10. Aufgabe)

############################################################
###############
rm(list=ls())

############################################################
###############
# Teil i: Simulation einer Umfrage mit p=1/3 und n=100

p<-1/3
n<-100

pdach<-rbinom(1,n,p)/n

# Konfidenzintervall berechnen:
unteregrenze<-pdach-2*sqrt(pdach*(1-pdach)/n)
oberegrenze<-pdach+2*sqrt(pdach*(1-pdach)/n)

############################################################
###############
# Teil ii

# Wiederhole w=1000 Mal

w<-1000
pdachvektor<-rbinom(w,n,p)/n

# Konfidenzintervalle berechnen
unteregrenzevektor<-rep(0,w)

for(i in 1:w)
     {
      unteregrenzevektor[i]<-pdachvektor[i]-
2*sqrt(pdachvektor[i]*(1-pdachvektor[i])/n)
     }

oberegrenzevektor<-rep(0,w)

for(i in 1:w)
     {
      oberegrenzevektor[i]<-
pdachvektor[i]+2*sqrt(pdachvektor[i]*(1-pdachvektor[i])/n)
     }


# Stelle graphisch dar
# zunaechst nur 100 Konfidenzintervalle, damit man es besser
erkennen kann:

plot(c(min(unteregrenzevektor),max(oberegrenzevektor)),c(0,10
0),type="n",xlab="Konfidenzintervalle",ylab="",main="Überdeck
ungen der Konfidenzintervalle")

for(i in 1:100)
     {

lines(c(unteregrenzevektor[i],oberegrenzevektor[i]),c(i,i),lw
d=1)
     }

abline(v=1/3,lwd=2,col="green")

# jetzt alle:
plot(c(min(unteregrenzevektor),max(oberegrenzevektor)),c(0,w)
,type="n",xlab="Konfidenzintervalle",ylab="",main="Überdeckun
gen der Konfidenzintervalle")

for(i in 1:w)
     {

lines(c(unteregrenzevektor[i],oberegrenzevektor[i]),c(i,i),lw
d=0.2)
     }

abline(v=1/3,lwd=2,col="green")


# relativer Anteil, der wahres p ueberdeckt:

ueberdeckt<-0
for(i in 1:w)
     {
      if(unteregrenzevektor[i]<=p &
oberegrenzevektor[i]>=p){ueberdeckt<-ueberdeckt+1}
     }

# relativer Anteil:
ueberdeckt/w

# Mittlere Laenge der Konfidentintervalle

laenge<-oberegrenzevektor-unteregrenzevektor
mean(laenge)

############################################################
###############
# Teil iii

N<-seq(from=10,to=10000,by=10)
l<-length(N)

laengeN<-rep(0,l)

for(j in 1:l)
     {
      pdachvektor<-rbinom(w,N[j],1/3)/N[j]

      unteregrenzevektor<-rep(0,w)

      for(i in 1:w)
          {
```

```
           unteregrenzevektor[i]<-pdachvektor[i]-
2*sqrt(pdachvektor[i]*(1-        pdachvektor[i])/N[j])
          }

      oberegrenzevektor<-rep(0,w)

      for(i in 1:w)
          {
           oberegrenzevektor[i]<-
pdachvektor[i]+2*sqrt(pdachvektor[i]*(1-
pdachvektor[i])/N[j])
          }


      laengevektor<-oberegrenzevektor-unteregrenzevektor
      laengeN[j]<-mean(laengevektor)
}

plot(log(N),log(laengeN),main="Log-log plot: Länge der
Konfidenzintervalle bei verschiedenen N")

# Im log-log-plot erhält man eine Gerade, wenn der
Zusammenhang zwischen n und laenge(n) von der Form
laenge(n)=n^a*c ist, denn:
log(laenge(n))=log(n^a*c)=a*log(n)+log(c).

# Zum Vergleich: Gerade mit Steigung -1/2 (in der Aufgabe
nicht gefordert!)
y<-rep(0,l)
for(i in 1:l)
{
y[i]<--1/2*log(N[i])-0.5
}
lines(log(N),y)

# Schlussfolgerung: die Länge/Breite faellt mit der
Größenordnung Wurzel(n), denn es gilt: laenge(n)=c*n^(-1/2)
(-1/2 ist die Steigung der Geraden im log-log-plot).
Erklärung: Das Konfidenzintervall ist gerade so konstruiert:
pdach+/-sqrt(pdach*(1-pdach)/n). Die Breite ist also
2*sqrt(pdach*(1-pdach)/n) und damit gilt: c=2*sqrt(pdach*(1-
pdach)) und a=-1/2.
```

3. Beispiel für einen Fragebogen (hier SoSe 2012)

Liebe Studentin, lieber Student, wie auch schon zu Beginn des Semesters bitten wir Sie einen Fragebogen auszufüllen. Da wir insbesondere gerne vergleichen möchten, ob sich bestimmte Einschätzungen und Einstellungen von Ihnen geändert haben, ist es wichtig, dass Sie auch die statistischen Informationen zu Ihrer Person angeben, damit wir Sie richtig zuordnen können. VIELEN DANK!

Die folgenden Angaben dienen statististischen Zwecken:

Studiengang: ______________________________________

Zur Zuordnung benötigen wir noch folgende Angaben:

Den dritten Buchstaben Ihres Vornamens	
Die ersten beiden Buchstaben des Vornamens Ihrer Mutter	
Die ersten beiden Buchstaben des Vornamens Ihres Vaters	
Der erste Buchstabe Ihres Straßennamens	
Die letzte Ziffer Ihrer Hausnummer	

Geschlecht: ☐ männlich ☐ weiblich

Ich habe das Proseminar Stochastik besucht: ☐ ja ☐ nein

TEIL I: AUFGABEN

Bitte lösen Sie die folgenden Aufgaben und kreuzen Sie anschließend an, wie sicher Sie sind, dass Ihr Ergebnis richtig ist.

1) In einer Stadt gibt es zwei Grundschulen. In der größeren werden jedes Jahr etwa 100 Kinder eingeschult, in der kleineren etwa 50 Kinder. Wir gehen davon aus, dass etwa 50% aller eingeschulten Kinder Mädchen sind. Der exakte Prozentsatz der Mädchen in den beiden Grundschulen variiert aber natürlich von Jahr zu Jahr.
Für den Zeitraum von 20 Jahren wurden in beiden Grundschulen jeweils die Jahre gezählt, an denen mehr als 60% der eingeschulten Kinder weiblich waren. In welcher Grundschule erwarten Sie mehr solche Jahre?

☐ **in der der kleineren Grundschule**
☐ **in der größeren Grundschule**
☐ **in beiden sind etwa gleich viele zu erwarten**

Wie sicher sind Sie, dass Ihr Ergebnis richtig ist?

absolut unsicher	ziemlich unsicher	eher unsicher	eher sicher	ziemlich sicher	absolut sicher
○	○	○	○	○	○

2) Eine Person A behauptet, Sie habe 20 mal eine faire Münze[1] geworfen und „Kopf“ als 0 und „Zahl“ als 1 notiert. Sind bei den folgenden 01-Folgen welche dabei, bei denen Sie daran zweifeln, dass sie auf diese Weise zustande gekommen sind? Falls ja, kreuzen Sie diese bitte an.

- ☐ **1 1 1 0 1 0 0 0 1 1 0 1 1 1 0 0 0 0 1 0**
- ☐ **0 1 0 0 1 0 1 0 0 1 1 0 1 0 0 1 1 0 1 0**
- ☐ **1 1 0 0 0 0 0 0 0 1 0 0 0 1 0 0 0 0 0 0**
- ☐ **1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1**

Ggf. Begründung?

Wie sicher sind Sie, dass Ihre Einschätzung richtig ist?

absolut unsicher	ziemlich unsicher	eher unsicher	eher sicher	ziemlich sicher	absolut sicher
○	○	○	○	○	○

3) Am Ende einer Vorlesung soll ein Multiple-Choice Test geschrieben werden. Test A besteht aus 10 Fragen, Test B aus 20 Fragen. Pro Frage gibt es zwei Antwortmöglichkeiten, von denen genau eine richtig ist. Beide Tests gelten als bestanden, wenn mindestens 60% der Fragen richtig beantwortet sind. Bei welchem der beiden Tests hat ein bei jeder Frage rein zufällig eine Antwort ankreuzt?

- ☐ **Bei Test A ist die Wahrscheinlichkeit größer.**
- ☐ **Bei Test B ist die Wahrscheinlichkeit größer.**
- ☐ **Bei beiden Tests ist die Wahrscheinlichkeit gleich groß.**

Wie sicher sind Sie, dass Ihr Ergebnis richtig ist?

absolut unsicher	ziemlich unsicher	eher unsicher	eher sicher	ziemlich sicher	absolut sicher
○	○	○	○	○	○

[1] Bei einer „fairen“ Münze ist die Wahrscheinlichkeit, dass sie bei einem Wurf „Kopf“ zeigt, genauso groß wie die Wahrscheinlichkeit, dass sie „Zahl“ zeigt, nämlich jeweils 1/2.

TEIL II: FRAGEN ZU R-SIMULATIONEN UND R-AUFGABEN

Wie häufig haben Sie die R- Aufgaben bearbeitet?

nie	selten	häufig	immer
○	○	○	○

Wie häufig haben Sie die R-Fragestunden besucht?

nie	selten	häufig	immer
○	○	○	○

Stimmen Sie folgenden Aussagen zu?

„Die Simulationen in der Vorlesung waren hilfreich für ein vertieftes Verständnis der Inhalte."

trifft überhaupt nicht zu	trifft weitgehend nicht zu	trifft eher nicht zu	trifft eher zu	trifft weitgehend zu	trifft vollständig zu
○	○	○	○	○	○

„Die Simulationen in den Übungen waren hilfreich für ein vertieftes Verständnis der Inhalte."

trifft überhaupt nicht zu	trifft weitgehend nicht zu	trifft eher nicht zu	trifft eher zu	trifft weitgehend zu	trifft vollständig zu
○	○	○	○	○	○

„Ich habe den Eindruck, dass sich durch die Simulationen meine stochastische Intuition verbessert hat ."

trifft überhaupt nicht zu	trifft weitgehend nicht zu	trifft eher nicht zu	trifft eher zu	trifft weitgehend zu	trifft vollständig zu
○	○	○	○	○	○

„Selbst Programme für Simulationen zu schreiben ist besser als nur Simulationsergebnisse gezeigt zu bekommen."

trifft überhaupt nicht zu	trifft weitgehend nicht zu	trifft eher nicht zu	trifft eher zu	trifft weitgehend zu	trifft vollständig zu
○	○	○	○	○	○

„Ohne die R-Fragestunde hätte ich die Aufgaben nicht lösen können."

trifft überhaupt nicht zu	trifft weitgehend nicht zu	trifft eher nicht zu	trifft eher zu	trifft weitgehend zu	trifft vollständig zu
○	○	○	○	○	○

„Die R-Aufgaben waren interessant."

trifft überhaupt nicht zu	trifft weitgehend nicht zu	trifft eher nicht zu	trifft eher zu	trifft weitgehend zu	trifft vollständig zu
○	○	○	○	○	○

„Ich würde empfehlen auch weiterhin Simulationen in den Übungen zur Elementaren Stochastik zu verwenden."

trifft überhaupt nicht zu	trifft weitgehend nicht zu	trifft eher nicht zu	trifft eher zu	trifft weitgehend zu	trifft vollständig zu
○	○	○	○	○	○

Sonstige Kommentare zu den R-Aufgaben/ Simulationen:

TEIL III: EINSTELLUNGEN ZU STOCHASTIK, EINSCHÄTZUNG DES WISSENS

„Mein Wissen in Stochastik schätze ich als sehr gut ein.“

trifft überhaupt nicht zu	trifft weitgehend nicht zu	trifft eher nicht zu	trifft eher zu	trifft weitgehend zu	trifft vollständig zu
◯	◯	◯	◯	◯	◯

„Bei Stochastik-Aufgaben kommt meistens etwas anderes raus, als man vermutet.“

trifft überhaupt nicht zu	trifft weitgehend nicht zu	trifft eher nicht zu	trifft eher zu	trifft weitgehend zu	trifft vollständig zu
◯	◯	◯	◯	◯	◯

„Bei Stochastik versagt meine Intuition.“

trifft überhaupt nicht zu	trifft weitgehend nicht zu	trifft eher nicht zu	trifft eher zu	trifft weitgehend zu	trifft vollständig zu
◯	◯	◯	◯	◯	◯

„Ich kann mir nur sehr schwer vorstellen, was Zufall bedeuten soll.“

trifft überhaupt nicht zu	trifft weitgehend nicht zu	trifft eher nicht zu	trifft eher zu	trifft weitgehend zu	trifft vollständig zu
◯	◯	◯	◯	◯	◯

„Ich mag Stochastik.“

trifft überhaupt nicht zu	trifft weitgehend nicht zu	trifft eher nicht zu	trifft eher zu	trifft weitgehend zu	trifft vollständig zu
◯	◯	◯	◯	◯	◯

„Das Lösen von Stochastik-Aufgaben fällt mir im Allgemeinen nicht schwerer als Aufgaben aus anderen mathematischen Bereichen.“

trifft überhaupt nicht zu	trifft weitgehend nicht zu	trifft eher nicht zu	trifft eher zu	trifft weitgehend zu	trifft vollständig zu
◯	◯	◯	◯	◯	◯

„Ich traue meinen Lösungen von Stochastik-Aufgaben im Allgemeinen nicht.“

trifft überhaupt nicht zu	trifft weitgehend nicht zu	trifft eher nicht zu	trifft eher zu	trifft weitgehend zu	trifft vollständig zu
◯	◯	◯	◯	◯	◯

„Stochastik ist nützlich.“

trifft überhaupt nicht zu	trifft weitgehend nicht zu	trifft eher nicht zu	trifft eher zu	trifft weitgehend zu	trifft vollständig zu
◯	◯	◯	◯	◯	◯

„Stochastik ist kein typisch mathematisches Fach.“

trifft überhaupt nicht zu	trifft weitgehend nicht zu	trifft eher nicht zu	trifft eher zu	trifft weitgehend zu	trifft vollständig zu
◯	◯	◯	◯	◯	◯

Falls noch Zeit ist:

Beim Lernen von Stochastik fand ich folgendes schwer:

4. Ergebnisse der Kolmogorov-Smirnov- sowie der Wilcoxon- bzw. U-Tests

Ergebnisse von Kolmogorov-Smirnov-Tests auf Normalverteilung

Alle Kolmogorov-Smirnov-Tests sowie auch die folgenden Wilcoxon-Tests wurden mit dem Programm R durchgeführt. Hochsignifikante Ergebnisse ($p<1\%$) bei den Kolmogorov-Smirnov-Tests sind gelb markiert.

Untersuchung der angegebenen Sicherheiten zu den Aufgaben, bei denen ein Vergleich (Anfang vers. Ende des Semesters) angestrebt wird, in der **Kontrollgruppe**:

Item	Gruppe	p-Wert
KH richtig	K	0.7952
GS richtig	K	0.02713
KH falsch	K	0.2252
GS falsch	K	0.09905
WS Münzwurferg. richtig	K	0.09982
WS Münzwurferg. falsch	K	0.7727
Erkennung mögl. gefälschter Münzwürfe falsch vorher	K	0.1093
Erkennung mögl. gefälschter Münzwürfe falsch nachher	K	0.01124
Testen eines Würfels richtig	K	0.2276
Testen eines Würfels falsch	K	0.8103
Post richtig	K	0.1544
Post falsch	K	0.9083
Multiple-Choice-Test richtig	K	0.5489
Multiple-Choice-Test falsch	K	0.008865

Items zu den Einstellungen (E1 bis E9) in der **Kontrollgruppe**:

Item	Gruppe	p-Wert	Item	Gruppe	p-Wert
E1 (vorher)	K	0.006677	E5 (nachher)	K	0.03279
E1 (nachher)	K	0.1425	E6 (vorher)	K	0.1688
E2 (vorher)	K	0.04586	E6 (nachher)	K	0.1507
E2 (nachher)	K	0.1701	E7 (vorher)	K	0.04692
E3 (vorher)	K	0.003752	E7 (nachher)	K	0.07862
E3 (nachher)	K	0.167	E8 (vorher)	K	0.01924
E4 (vorher)	K	0.04613	E8 (nachher)	K	0.01001
E4 (nachher)	K	0.06566	E9 (vorher)	K	0.2755
E5 (vorher)	K	0.04078	E9 (nachher)	K	0.3642

Angegebene Sicherheiten zu den Aufgaben, bei denen ein Vergleich (Anfang vers. Ende des Semesters oder Experimentalgruppen untereinander) angestrebt wird, in der **Experimental-gruppe**:

Item	Gruppe	p-Wert
KH richtig	R12	0.1062
	R34	0.39
	Frage	0.7936
GS richtig	R12	0.3125
	R34	0.02508

	Frage	0.172
KH falsch	R12	0.1074
	R34	0.5595
	Frage	0.1437
GS falsch	R12	0.214
	R34	0.4035
	Frage	0.296
WS Münzwurferg. richtig	R12	0.1121
	R34	0.03369
	Frage	0.2097
WS Münzwurferg. falsch	R12	0.9992
	R34	0.6986
	Frage	0.5152
Erkennung mögl. gefälschter Münzwürfe richtig nachher	R12	0.8327
	R34	0.3338
	Frage	0.577
Erkennung mögl. gefälschter Münzwürfe falsch vorher	R12	0.07229
	R34	0.05215
	Frage	0.1071
Erkennung mögl. gefälschter Münzwürfe falsch nachher	R12	0.429
	R34	0.02354
	Frage	0.1118
Testen eines Würfels richtig	R12	0.2318
	R34	0.1242
	Frage	0.2392
Testen eines Würfels falsch	R12	-
	R34	0.6807
	Frage	0.7127
Post richtig	R12	0.04379
	R34	0.01948
	Frage	0.2294
Post falsch	R12	0.9992
	R34	0.9052
	Frage	0.2454
Multiple-Choice-Test richtig	R12	0.8257
	R34	0.3574
	Frage	0.5205
Multiple-Choice-Test falsch	R12	0.8838
	R34	0.532
	Frage	0.02485

Items zu den Einstellungen (E1 bis E9) und zur Evaluation der „R Aufgaben“ (R1 bis R7) in der **Experimentalgruppe**:

Item	**Gruppe**	**p-Wert**	**Item**	**Gruppe**	**p-Wert**
E1 (vorher)	R12 R34 Frage	0.1039 0.2956 0.2969	E7 (nachher)	R12 R34 Frage	0.3621 0.1171 0.06336
E1 (nachher)	R12 R34 Frage	0.3162 0.05569 0.05475	E8 (vorher)	R12 R34 Frage	0.09606 0.01667 0.01056
E2 (vorher)	R12 R34 Frage	0.1669 0.1399 0.00368	E8 (nachher)	R12 R34 Frage	0.3412 0.04827 0.286
E2 (nachher)	R12 R34 Frage	0.1502 0.09899 0.04701	E9 (vorher)	R12 R34 Frage	0.08822 0.179 0.0678
E3 (vorher)	R12 R34 Frage	0.07792 0.07319 0.06587	E9 (nachher)	R12 R34 Frage	0.04359 0.2053 0.4974
E3 (nachher)	R12 R34 Frage	0.1412 0.124 0.2033			
E4 (vorher)	R12 R34 Frage	0.1651 0.05007 0.1414	R1	R12 R34 Frage	0.2764 0.02853 0.1255
E4 (nachher)	R12 R34 Frage	0.02325 0.04033 0.04835	R2	R12 R34 Frage	0.2889 0.2512 0.006856
E5 (vorher)	R12 R34 Frage	0.4032 0.09527 0.1632	R3	R12 R34 Frage	0.6815 0.1499 0.1433
E5 (nachher)	R12 R34 Frage	0.2186 0.137 0.3775	R4	R12 R34 Frage	0.9236 0.1021 0.2874
E6 (vorher)	R12 R34 Frage	0.5052 0.0608 0.3427	R5	R12 R34 Frage	0.3847 1.904e-05 0.0009746
E6 (nachher)	R12 R34 Frage	0.3188 0.02194 0.173	R6	R12 R34 Frage	0.4538 0.1886 0.01278
E7 (vorher)	R12 R34 Frage	0.3651 0.08821 0.07574	R7	R12 R34 Frage	0.7414 0.1233 0.2166

<u>Wilcoxon-Tests (für gepaarte bzw. ungepaarte Stichproben)</u>

Folgende Vergleiche wurden mit entsprechenden Wilcoxon-Tests wiederholt:

- alle Vergleiche, bei denen mind. eine Gruppe beim Kolmogorov-Smirnov-Test einen p-Wert von weniger als 1% aufweist („starke Verletzung" der Normalverteilungs-annahme)

und

- Vergleiche, bei denen mind. eine Gruppe beim Kolmogorov-Smirnov-Test einen p-Wert von weniger als 5% aufweist und zusätzlich der im Ergebnisteil dokumentierte t-Test einen signifikanten Unterschied angezeigt hat.

Kontrollgruppe:

Item	**Vergleich**	**p-Wert des Wilcoxon- Tests**	**Beachtenswerter Unterschied zum t-Test Ergebnis?**
GS Si	richtig-falsch	0.03901	Nein
MC Si	richtig-falsch	0.02864	Nein

Item	**Vergleich**	**p-Wert des Wilcoxon- Tests**	**Beachtenswerter Unterschied zum t-Test Ergebnis?**
E1	vorher-nachher	0.8635	Nein
E3	vorher-nachher	0.8602	Nein

Experimentalgruppe:

Item	**Vergleich**	**p-Wert des Wilcoxon- Tests**	**Beachtenswerter Unterschied zum t-Test Ergebnis?**
GS Si	R34-Frage (richtige Lsg.)	0.0005955	Nein
MC Si	Frage: richtig-falsch	0.02166	Nein

Item	**Vergleich**	**p-Wert des Wilcoxon- Tests**	**Bedeutsamer Unterschied zum t-Test Ergebnis?**
E2	vorher-nachher bei Frage	0.3014	Nein
E2	vorher: R12-Frage R34-Frage	0.111 0.3177	Nein Nein
E4	nachher: R12-Frage	0.02707	Nein
E8	Vorher-nachher bei R34	0.04982	Nein
R1	R12-R34	0.009133	Nein
R2	R12-Frage R34-Frage	0.001809 0.05009	Nein Nein
R2	Vgl. mit R1 bei Frage	0.4165	Nein

R5	R12-R34 R12-Frage R34-Frage	5.322e-05 0.3044 6.03e-11	Nein Nein Nein
R6	R12-Frage R34-Frage	1.12e-05 0.009262	Nein Nein

5. Leitfaden für die halbstrukturierten Interviews

Leitfadeninterview: Fragen

1) Was war deine erste Empfindung, als du von den R-Aufgaben gehört hast? Hat sich an deiner Einstellung bis heute etwas geändert?
2) Wie häufig hast du die R Aufgaben bearbeitet? Wie häufig warst du in der Fragestunde? Warum hast du die Fragestunde besucht/nicht besucht?

mind. selten bearbeitet

3) Fandest du die R Aufgaben im Vergleich zu den anderen Aufgaben eher schwer oder leicht? Was war schwerer, was leichter?
4) Was fandest du an den R Aufgaben gut? An welche erinnerst du dich noch und warum? Ggf. als Nachfrage: Gibt es etwas, was dir durch eine R Aufgabe klar geworden ist, was du vorher nicht verstanden hattest? Wenn ja: was?
5) Was hat dir bei den R Aufgaben nicht gefallen?
6) Würdest du aus Studentensicht empfehlen, weiterhin R Aufgaben zu verwenden? Würdest du etwas verändern?
7) Wurde in deinem Tutorium manchmal auf R Aufgaben Bezug genommen, als eine andere Aufgabe an der Tafel erklärt wurde? Falls ja: Erinnerst du dich an Beispiele?
8) Welches Objekt fällt dir als erstes ein, wenn du an Stochastik denkst? Erkläre, was es ist.
9) Was verstehst du unter „stochastischer Intuition“? Gibt es einen Unterschied zu „Verständnis von Stochastik“?
10) Möchtest du noch etwas anmerken?
11) Angaben zur Person (separat)

nie bearbeitet

3) Warum hast du die R Aufgaben nicht bearbeitet?
4) Was hätte man anbieten müssen, damit du sie bearbeitet hättest?
5) Hattest du den Eindruck, dass du dadurch, dass du die Aufgaben nicht bearbeitet hast, Nachteile hattest? Wenn ja: welche?
6) Findest du es prinzipiell eine gute Idee, in Stochastik Simulationsaufgaben zu stellen?
7) Gibt es etwas, was man damit deiner Meinung nach besonders gut lernen kann?
8) Wurde in deinem Tutorium manchmal auf R Aufgaben Bezug genommen, als eine andere Aufgabe an der Tafel erklärt wurde? Falls ja: Erinnerst du dich an Beispiele?
9) Welches Objekt fällt dir als erstes ein, wenn du an Stochastik denkst? Erkläre, was es ist.
10) Was verstehst du unter „stochastischer Intuition“? Gibt es einen Unterschied zu „Verständnis von Stochastik“?
11) Möchtest du noch etwas anmerken?
12) Angaben zur Person (separat)

6. Transkripte der Interviews mit farblichen Markierungen für die Zuordnungen zu den Fragenkomplexen

Fragenkomplexe	Farbe
1 A/B	gelb
2 A/B / 3 B	hellgrün
3 A	türkis
4 A / 5 B / 7 B	rosa
5 A	blau
6 A / 4 B / 6 B	rot
7 A / 8 B	dunkelblau
8 A / 9 B	blaugrün
9 A / 10 B	grün
10 A / 11 B	violett
Simulationsplan	dunkelrot

Int1

F: Also die erste Frage ist, was war deine erste Empfindung, als du von den R Aufgaben gehört hast und hat sich an deiner Einstellung bis heute etwas geändert?

A: Also, der erste Gedanke, der mir gekommen ist, war Oh mein Gott, ich kann mit so Programmen in der Regel nicht umgehen. Und der zweite Gedanke war dann, ich hab nämlich einen Kommilitonen, mit dem ich zusammenarbeite und der das sehr gerne macht, solche Programme, weil der parallel auch noch Informatik studiert als drittes Fach sozusagen und dann hab ich gedacht, dann wird das schon irgendwie gehen und jetzt so im Nachhinein ist es so, dass ich es schon sinnvoll finde, aber mich selber damit noch nicht so vertraut fühle oder [ähm] viel Lust habe mich damit zu beschäftigen.

F: Ok. Wie häufig hast du die Aufgaben denn bearbeitet?

A: Also, kommt drauf an, wie man jetzt Bearbeitung definiert, ich hab sie immer durchgelesen und thematisch so Gedanken gemacht, was man da machen könnte, aber bearbeitet in dem Sinne eigentlich nur der andere Kommilitone und im Nachhinein auf die Klausurvorbereitung bin ich ein paar wichtige Sachen durchgegangen, so grundlegende Dinge.

F: Ok. Und warst du in der Fragstunde?

A: Nein.

F: Warum nicht?

A: [lacht] Hm, ja, zu zeitaufwändig dann, weil ich sie ja nicht so in dem Umfang bearbeiten wollte und die Fragen konnte ich dann an meinen Kommilitonen stellen.

F: Ok, das heißt, dann fällst du hier in meine Kategorie mind. selten bearbeitet. Da ist die erste Frage: Fandest du die R Aufgabe im Vergleich zu den anderen Aufgaben eher schwer oder eher leicht und was war schwerer und was leichter?

A: Ähm, also tendenziell eher leichter, aber, ich weiß nicht, welche Übung es jetzt genau war, also einmal, da wars relativ schwer, weil ich da nur so ne Idee hatte, und mein Kommilitone gesagt hat, oh, er kriegts auch nicht hin und wenn ers nicht hinkriegt, dann hab ich mir gedacht, hm, dann muss es ein bißchen schwieriger sein, aber tendenziell auf jeden Fall leichter. Und, aufgrund von [ähm] ja Praxisbezug oder Anwendung vielleicht auch motivierender als jetzt irgendwie ne Stochastikaufgabe, die nach rein Schema F geht und man eigentlich gar nicht weiß, was da gemacht wird.

F: Ähm, also es ist jetzt schon so bißchen mitbeantwortet, aber trotzdem nochmal: Was fandest du an den R Aufgaben gut und an welche erinnerst du dich noch und warum?

A: Also, allgemein gut, dass man eben sagt, man untersucht was, mit Experimenten, was man mit Schülern ja auch macht, irgendwie, und dann auch sagt, ok, wir, wir persönlich würden das jetzt 1, 2, 10, vielleicht auch noch 100 mal ausprobieren, aber ab 10000, 1000 hörts auf, so oft machen wir das nicht, dass man das so macht, und was ich sonst noch gut fand, war dass man in Diagramme übersetzt, sozusagen. Ähm, ja, mit diesen Konfidenzintervallen, das ist mir noch so im Kopf, weil du ja da auch

die Vorlesung drüber gehalten hast, aber so richtig die Klausuraufgabe konnt ich dann nicht lösen, vielleicht auch weil ich mich dann nicht genau mehr damit beschäftigt hab. So grob, was es ist oder was man damit machen kann, ja, aber, ich überleg grad, ob mir noch ne R Aufgabe [Pause] war da nicht auch was mit dem Münzwurf, so ganz am Anfang, also ich fands immer praktisch, dass es so Praxisbezug hat, und man wusste, dass man da jetzt was simuliert, und innerhalb, wenn dann irgendwelche Regeln, die man nicht kennt aufgetaucht sind, war das auch erstmal egal, aber dass es einfach um was geht und man das dann untersuchen kann.

F: Ok. Was hat dir denn bei den R Aufgaben nicht gefallen?

A: Hm, also das müsst ich eher global beantworten, so nach dem Motto generell so Programme wie auch aus dem PC-Einsatz, so Fathom, und was wir da alles gemacht haben, dass ich einfach generell nicht motiviert bin und so Angst nicht, aber Respekt vor diesen Aufgaben habe und schon so mit normalen Computerprogrammen, wenn ich Word oder so bediene, überfordert ist übertrieben, aber ich mich so rantasten muss und dann sage, ach, wenns anders geht, verwende ich das. Und ich das auch selber kritisch sehe, weil ich denke, dass es nicht gut ist, ich das machen müsste, aber irgendwie noch so Respekt davor hab.

F: Würdest du denn aus Studentensicht empfehlen, weiterhin R Aufgaben zu verwenden und sollte man irgendwas ändern, wenn man es weiter verwendet?

A: Auf jeden Fall weiter verwenden, die Frage ist, ich würde dann vielleicht ne Unterscheidung zwischen L3 und anderen machen oder bzw. das hinterfragen, weil für L3 ist das mit Sicherheit sinnvoll, weil man das spätestens mit Oberstufenschülern machen könnte oder, naja, Schülern, die SEHR motiviert sind, denen könnte man sowas mal in die Hand geben und sagen, probier das einfach mal aus und mach dein Ding. Ähm, ich weiß nicht wie Bachelor oder diese anderen Studiengänge das sehen. Ähm, wie war nochmal der zweite Teil?

F: Ob man etwas verändern sollte.

A: Ja, verändern. Ich überleg, ich hab schon die ganze Zeit im Vorhinein überlegt, was die Veränderung hätte sein müssen, damit ich die Aufgaben mehr gemacht hätte. [Ähm] und da ich dann aber von vornherein eigentlich wusste, dass die R Aufgabe sagen wir mal nur ein Zehntel der Klausur ist und irgendwie die anderen Aufgaben, die ja drei Viertel waren, irgendwie mehr gewichtet sind, weil die ja einfach mehr drankommen, hab ich da versucht den Schwerpunkt darauf zu legen und hab mir gesagt, die R Aufgabe krieg ich vielleicht auch irgendwie so hin. Aber es macht ja auch keinen Sinn zu sagen, die Hälfte der Klausur ist R Code, also, ähm, aber vielleicht mehr Punkte dafür, aber dann wäre es ja nur so, wenn jemand sagen würde, dass er den R Code nicht sinnvoll findet, der wird dann sagen, ja warum gibts da so viel Punkte, habt ihr sie noch alle, das ist ja eigentlich gar nicht so wichtig. Also, ich könnt mir vorstellen, dass Bachelor-Studenten das trivial finden, oder in Richtung trivial und Lehramtsstudenten eher froh um so ne Aufgabe sind.

F: Ich hatte eher den Eindruck, dass gerade L3 Studenten skeptisch waren.

A: Das überrascht mich. Vielleicht hat das aber was damit zu tun, also ich hab jetzt Stochastik ganz am Ende meines Studiums gemacht, was überhaupt nicht Sinn der Sache ist, aber das hat sich durch mehrere Dinge, durch Uniwechsel und so ergeben,

und ich hatte vor dieser Veranstaltung PC-Einsatz gehört. Und hier machen das Leute, glaube ich im ersten oder zweiten Semester und haben dann PC-Einsatz noch nicht gehört und vielleicht ist deshalb auch die Verknüpfung noch nicht da. Also, durch dieses PC-Einsatz wird einem nämlich schon nahe gelegt, dass man auch Computerprogramme, die irgendwas simulieren, nicht verzichten sollte oder DARF, also vielleicht ist das die Erklärung. Weil das finde ich persönlich komisch.

F: Ok, das ist interessant. Die nächste Frage ist: Wurde in deinem Tutorium manchmal auf R Aufgaben Bezug genommen, wenn eine andere Aufgabe erklärt wurde?

A: Also spontan würd ich sagen nie, aber so darauf eingehen, also höchstens bei Tipps für die Aufgaben, da hieß es manchmal, das müsst ihr auch im R Code machen oder so, aber ich war jetzt bis auf zweimal glaub ich immer im Tutorium und da kann ich eigentlich sagen, dass wir da NICHT drauf eingegangen sind und dass ich immer das Gefühl hatte, dass die Tutorin hinter den R Aufgaben auch nicht steht, oder die nicht so bearbeitet.

F: Aber es wurde jetzt nicht irgendwas gesagt, dass man die R Aufgabe nicht machen sollte, oder …

A: Nein. Es hat einfach keine Rolle gespielt. Genau, also die R Aufgabe hat so irgendwie keine Rolle gespielt und wenn, es gab vereinzelt mal, aber die haben das dann auch schnell wieder gelassen, so Kommilitonen, die nachgefragt haben, wie muss ich das denn machen oder so und dann wurd eben auf die R Fragestunde hingewiesen, was ja auch in Ordnung ist, weil die Zeit dann mit den anderen drei Aufgaben eigentlich auch gefüllt war. Aber es war nicht so „Macht die nicht!“, auf gar keinen Fall.

F: Ok, ja, dann jetzt noch etwas Inhaltliches: Welcher Begriff fällt dir als erstes ein, wenn du an Stochastik denkst?

A: Allgemein Stochastik?

F: Ja.

A: Äh, spontan, äh, Wahrscheinlichkeit, äh, unkorreliert, ich weiß zwar nicht warum, aber, unabhängig. Nur einen?

F: Drei sind auch gut. Ok. Und kannst du die Begriffe erklären, also was das jeweils ist?

A: [Pause]

F: Also, muss jetzt keine Definition sein.

A: Ok, also die drei, die ich jetzt genannt habe, ich glaube, dass ich die letzten zwei, also unkorreliert und unabhängig eher so Wenn das erfüllt ist, dann ist es so, also mit diesen Formeln und Wahrscheinlichkeit könnt ich eher so mehr oder weniger umgangssprachlich formulieren.

F: Ja, mach das ruhig.

A: Ich würd sagen, Wahrscheinlichkeit gibt an, hm, dass irgendein Fall, also wenn ich jetzt Lotto spiele, ähm, wie wahrscheinlich es ist, dass eben eine Zahl von diesen 49 gezogen wird, oder dann auch dass man unterscheidet, ob irgendwelche Dinge

gleichwahrscheinlich sind oder eben ob es da noch einen Unterschied gibt. Also, es ist halt so, so eine rein mathematische Definition von Wahrscheinlichkeit könnte ich eher nicht geben. Es ist so was Prozentuales.

F: Ok. Ähm, was verstehst du unter stochastischer Intuition? Das war so ein Begriff, der in den Fragebögen vorkam. Und gibt es einen Unterschied zu Verständnis von Stochastik?

A: Also, stochastische Intuition, haben wir ja auch in diesem einen Stochastik-Kurs mal drüber gesprochen, ähm, ist ja eigentlich so dieses innere Gefühl, was da irgendwie bei rauskommt. Und MEINE stochastische Intuition ist anscheinend nicht so gut, weil ich irgendwie immer denke, das kann nicht sein, warum ist das so, das kann nicht sein, also so beim Roulette, wo das irgendwie nach Schema F gerechnet wird, versteh ich das auch, und da ist die Intuition, deckt sich vielleicht auch, also z.B. beim Würfeln, dass es ½ ist, dass eine gerade Zahl kommt, also, das ist klar, da deckt sich vielleicht Intuition auch mit Verständnis, ABER bei so was wie dem Geburtstagsproblem geht es überhaupt nicht bei mir, da ist die Intuition, ach ist eher unwahrscheinlich, und ähm, ja, die Intuition ist da nicht gut, wenn ich es aber rechne, also die Rechenregeln sind mir klar, also würde ich sagen, ich habe ein Verständnis dafür, wie ichs rechne, äh, mir macht es aber trotzdem nicht bewusst, warum bei diesem Verständnis, das ich ja eigentlich hab, so etwas rauskommt, weil es mit der Intuition ja immer noch nicht klar geht und so diese innere Intuition, die man hat, die lässt sich nicht davon wegbringen, obwohl ichs gerechnet hab und weiß es geht so. Also, es bleibt so ein fahles Gefühl und dann denk ich, ich habs ja gerechnet, also müsste es eigentlich so gehen, aber es ist so, bis heut noch nicht, das Geburtstagsproblem, warum das quasi doch so wahrscheinlich ist, dass man an einem doppelten Tag Geburtstag hat.

F: Wenn du sagst, Verständnis deckt sich mit Rechnen können ...

A: ja

F: dann ist der Sprung zur Intuition also noch groß, also um den zu überbrücken müsste man etwas anderes machen als rechnen

A: Ja, also, das Rechnen an sich hilft noch nicht dafür, also es ist schon so, also es ist dann so, ich hab das gerechnet, ich hab das richtig gemacht und die Rechenregel ist mir auch klar, aber es bleibt immer noch die Verwunderung, krass, warum ist das so, also das Warum ist nicht beantwortet, was dann natürlich bei mir so ist, also ich probier dann immer aus, und wenns dann irgendwie durch Ausprobieren so ist, dann werd ich bestätigt, aber andererseits lernt man ja auch in anderen Situationen, so paarmal ausprobieren bestätigt mathematisch noch gar nichts. Klar, wenn man das mit den Simulationen macht, klar, wenn man bei 1000, 10000, immer strebt es gegen irgendeinen Wert, dann macht es schon Sinn, dass es so ist, was ja auch beim Würfel eigentlich so ist, ja, also ist ja gar nicht die Frage. Aber trotzdem, ich weiß nicht, was man für die Intuition, also Intuition ist auf jeden Fall nochmal durch dieses Ausprobieren, das stärkt sie.

[Beispiel aus einer Didaktik-Veranstaltung: Spiel mit Zauberer, Krieger und Bär]

F: Ok. Möchtest du noch irgendetwas anmerken?

A: Hm [lacht]. [Pause] Ich muss erstmal überlegen. Hm, ich weiß nicht.

F: Was vielleicht nochmal interessant wäre: Gibt es irgendwas, was man hätte anbieten können, damit du die R Aufgaben öfter selber gelöst hättest?

A: Also ich glaube, wenn sie im Tutorium besprochen worden wären. Also, ja, weil das für mich immer so ein, also es war klar, dass es diese R Fragestunden gibt, und dass die wichtig sind oder wie auch immer, aber es war halt mehr so ein Angebot und kein Muss, das Tutorium ist zwar auch nur ein Angebot, aber doch eher als Muss zu verstehen für Studenten, und alles, was da nicht besprochen wird, ist irgendwie zweitrangig. Ja, also das im Tutorium hätte bestimmt geholfen, ja, also dass es vielleicht auch jemand präsentiert hätte und allein schon, wenn ich mir das in Anführungsstrichen anhören MUSS, wie derjenige, der Student das jetzt gelöst hat oder wie ers zeigt, was er jetzt gemacht hat, dann würd ich glaub ich unbewusst schon so Begriffe im R Code mitkriegen, also replace oder was da war, das fänd ich schwer, das zu verstehen oder irgendwas, aber das kriegt man erst, wenn mans paar Mal hört, rein. Oder es wiederholt sich ja ständig, gewisse Dinge. Also, ich glaub, das hätte geholfen. Ja gut, in der Vorlesung wars ja schon ein paar Mal, obwohl glaub ich der Herr [xxx] irgendwann mal gesagt hat, er wills nochmal machen, dann hat ers doch nicht gemacht. Ja, das sind dann immer so technische Sachen oder was auch immer, ähm, aber ich weiß nicht, obs viel gebracht hätte, wenn man da mehr gemacht hätte, und ich glaub auch, ein Student geht nicht unbedingt in die Vorlesung, also würden es eh nicht alle mitkriegen. [Pause] Ja, durch MEHR Aufgaben, also, auf keinen Fall MEHR Aufgaben, dass es dann 5 oder 6 sind, würd ich sagen, weil ähm, das wäre dann einfach zu viel für die Woche und dann wärs aber auch nicht klar, ist das jetzt hier grad ein PROGRAMMIER äh Vorlesung, Kurs, oder ist es Stochastik. Deshalb fand ich das mit dem ein Viertel drei Viertel schon gut, aber klar, umso mehr man das integriert umso mehr müssen die Leute das machen, also wär ja auch so wenn man zwei andere Aufgaben und zwei R Code Aufgaben hätte, wär ja auch Zeit da, die R Aufgaben zu besprechen. Also, ich hatte nämlich das Stochastik schon mal belegt und da waren vier normale Aufgaben, die haben wir ja auch durchgesprochen bekommen, deshalb hab ich mich sowieso gefragt, wieso man quasi den R Code nicht mit rein nimmt.
So, aber, noch ne Anmerkung, hm [Pause] Ne, also eigentlich wars gut aufgebaut, grade mit den Fragestunden, also das Angebot reicht, aber es ist halt kein Muss und dadurch ist es für Studenten immer, ich mach nur das Minimale oder, ja.

F: Und hast du vielleicht ein Argument für R, irgendwas, was man an den R Aufgaben lernen kann, was einem sonst vielleicht entgeht?

A: Also, es ist halt anwendungsbezogen, also mir ist es irgendwie so in Erinnerung, das ist irgendwie ein Sachverhalt, die Aufgabe ist klar, ich wüsste aber überhaupt nicht, wie ich die Aufgabe jetzt mathematisch lösen sollte, und geh dann übers Programm und krieg irgendwie ein Ergebnis. Das ist irgendwie für mich ein ganz gutes Argument, warum einfach, ja, warum man das einfach einsetzt. Wobei man die reine Mathematik dann nicht weglassen sollte. Und es sind immer Aufgaben gewesen, die man, sag ich jetzt mal VERSTANDEN hat, also das und das ist der Sachverhalt, und bei anderen reinen Stochastikaufgaben denk ich manchmal auch, also erstens weiß ich nicht, worum es geht, zweitens weiß ich auch gar nicht, was das soll und was kann ich damit anwenden und erklären, also und dann frag ich mich halt ständig, wie könnt ich das einem Oberstufenschüler, also es ist dann ja doch nur was für einen Oberstufenschüler, wie könnt ich das dem jetzt verklickern oder wo, bei welchem Thema kann ich das mit reinbringen. Und DIESE Aufgaben, die könnt ich ja ständig bringen und die reinen Stochastikaufgaben, frag ich mich, also ich will halt immer über den Tellerrand hinaus Aufgaben haben und ich denk aber schon manchmal, das sind jetzt aber schon Aufgaben auf nem SEHR hohen Niveau, die ich nem Schüler nicht mehr zutrauen will, KANN, nicht KANN, weil ichs selbst nicht verstanden hab auf dem Niveau und deshalb ist es dieser Anwendungsbezug auf jeden Fall, das finde ich das beste Argument für die R Aufgaben.

F: Wunderbar, fällt dir sonst noch etwas ein?

A: Ne, ne, ich glaub das war alles.

F: Vielen Dank, das war wirklich sehr hilfreich.

Int2

F: Also, die erste Frage ist: Was war deine erste Empfindung, als du von den R Aufgaben gehört hast und hat sich an deiner Einstellung was geändert?

A: Ähm, also eigentlich erstmal kannt ich das wirklich überhaupt nicht, dann hab ich mir überlegt, dass ich das eigentlich ne gute Sache finde, dass man sich mit so was beschäftigt und das könnte man in der Oberstufe, also ich studier ja auf Lehramt, dass man den Schülern das zeigen kann und dass man auch noch ein bißchen mehr kann, als die Schüler, fand ich eigentlich ganz gut und bin auch recht aufgeschlossen gegenüber solchen Programmen. Ähm, also eigentlich erstmal neutral bis positiv, ja, ähm, und mittlerweile ja, hat sich das halt zu so nem riesen Berg angehäuft, wo ich keine Chance sehe, das irgendwie zu bewältigen und ähm, ja, also hab so ein bißchen versucht das so zu verdrängen, vor mir herzuschieben, und weiß noch nicht so recht, ob ich das noch so hinkriege.

F: Ok, und der riesen Berg kommt jetzt durch die Befehle zustande, oder durch was genau?

A: Ähm, also ich hab das am Anfang noch probiert, hab mich da drangesetzt und hab versucht, das da reinzukopieren in diese, hab mir das runtergeladen, mit dem, also die wichtigsten Befehle und hab auch noch im Internet, bevor ich diese Liste gesehen hab, geguckt, ob ich irgendwas finde, so R Aufgaben und den Befehl, was man so gerne hätte, und hab auch zum Teil was gefunden, das hat aber nicht funktioniert [lacht], also Schrott, und zum Teil hat er auch gar nichts gerechnet, hab so, also das ganze Prinzip einfach noch nicht verstanden. Und hab versucht, dann auch später nochmal, als die Lösungen online waren, das nachzuvollziehen, und selbst da hab ichs zum Teil nicht hinbekommen, dass er das wenigstens, konnt noch nicht mal das rüberkopieren, und dann hab ichs auch irgendwann sein gelassen, ich hatte ja auch noch anderes zu tun. Und ähm, deswegen dadurch weiß ich halt, wieviel Arbeit das jetzt ist, wenn ich das jetzt richtig noch machen will für die Klausur.

F: Ok, also es ist eher das Technische, was jetzt so Schwierigkeiten gemacht hat, nicht so das Inhaltliche?

A: Ja, also eigentlich kann ichs noch nicht so beurteilen.

F: Ok, gut, ähm, oder schade. Wie häufig hast du denn die R Aufgaben bearbeitet, also angeschaut oder sogar gelöst?

A: Also, die ersten drei Blätter hab ich noch probiert.

F: Und warst du mal in der Fragestunde?

A: Das hab ich zeitlich irgendwie, also das hätte mir den ganzen Tag zerrupft und dann gabs ein paar Donnerstage, da ist es dann eh ausgefallen (Feiertage, F.) und ein paar Donnerstage hab ich mir dann irgendwie jobmäßig noch zugeballert und dann hab ichs leider nicht geschafft, obwohl ichs gerne wahrgenommen hätte.

F: Für mich wäre trotzdem, obwohl du nur drei angeguckt hast, eine Frage zu den R Aufgaben interessant: Und zwar, wenn du die vergleichst mit den anderen Aufgaben, also unabhängig jetzt von den technischen Schwierigkeiten, fandest du, dass die eher

leicht oder eher schwer waren? Und was war schwer und was leicht?

A: Also, was ich ganz angenehm, oder tendenziell eher leicht fand, war, dass die ja zum Teil mit den Aufgaben davor in Zusammenhang standen, also man hat irgendwas ausgerechnet und dann hätte man es so irgendwie, ja, nachvollziehen können. Ähm, ja, ich hab mir generell mit den Aufgaben schwer getan, also, Stochastik ist nichts, was mir so irgendwie zufliegt, und als ich dann, also ich hab dann versucht, wenigstens DIE Aufgaben zu machen, bin da aber auch gescheitert, von daher, ist das alles für mich in unerreichbare Weite gerückt.

F: Hätte man denn irgendwas anbieten können, damit du die R Aufgaben bearbeitet hättest, oder dass es irgendwie leichter gefallen wäre?

A: Also, ich mein, das Angebot der Fragestunde gabs ja, da hätt ich halt irgendwie meinen Zeitplan über Bord werfen müssen, ähm, [Pause], ich weiß nicht, vielleicht, ich hätte mich, glaub ich, gefreut, wenn einem der Einstieg leichter gemacht worden wäre, also, statt dass es sofort R Aufgaben gibt, sagen, es gibt so ne Art Tutorium oder Crashkurs, für die, die das noch nie gehört haben, vier Stunden, und dann hat man ne Idee, wie das Programm funktioniert, ähm, ja, wär wahrscheinlich auch, also, wär wahrscheinlich schwierig gewesen, das für alle Studenten anzubieten, aber, so was, so ein Einstieg, dass ich mal ne Idee kriege, wie funktioniert das, was ist die Grundidee von dem Programm.

F: Ok. Und das hätte man auch im Tutorium machen können?

A: Äh, ja, vom Prinzip her schon, wobei ich, zumal ich ja für den Tutorientermin, also da hatt ich ja frei, aber, wobei ich nicht weiß, ob die Tutoren das hingekriegt hätten.

F: Ja aber das ist sicher ein wichtiger Punkt, so eine Starthilfe hat gefehlt.

A: Ja, genau, irgendwie schon.

F: Hattest du denn den Eindruck, dass du dadurch, dass du die R Aufgaben jetzt nicht bearbeitet hast, oder selten bearbeitet hast, irgendwelche Nachteile hattest?

A: Ähm, also der Druck, dass ich wusste, das kommt in der Klausur dran und ich muss das irgendwie hinkriegen und ähm, gleichzeitig auch, war die Frustrationsgrenze einfach geringer, weil ich schon wusste, da sind vier Aufgaben, eine kann ich sowieso nicht [lacht], dann sinds schon nur noch drei, und wenn ich von den drei dann noch eine nicht konnte, oder ähm, mich irgendwie schwer getan hab, ja dann. [Pause] Also dann hab ich das relativ schnell abgeschrieben.

F: Also das Ganze abgeschrieben? Das ganze Fach?

A: Ja.

F: Ok. Ähm, findest du es denn prinzipiell ne gute Idee, in Stochastik Simulationsaufgaben …

A: Ja, also ich find das genial, dass man ne Idee hat, was da passiert, dass man das sieht, das ist ja zum Teil also nicht intuitiv oder selbst wenns intuitiv ist, dann kann man sichs nicht wirklich vorstellen, und das so auch mit den mit den Diagrammen dann zu

verbildlichen, find ich absolut genial und wär genau das Richtige für mich eigentlich, aber ähm, also zum Vergleich vielleicht, ich mach das ganz gerne ganz primitiv in Excel, dass ich in Analysis ne Formel eingebe, mir ein paar Werte reinziehe, das rüber kopiere und dann einfach mal so ne Idee hab, was das überhaupt is, wie der Graph jetzt aussieht. Und das hätt ich auch gerne machen können für mich, dass ich sagen kann ok, ich hab jetzt die und die Verteilung und wenn ich jetzt sage, was für nen Wert kann ich denn erwarten, bin ich dann, also wenn das Programm jetzt natürlich nicht DEN Erwartungswert ausrechnet, sondern vielleicht natürlich durch Experimente, aber trotzdem sollte sich das dem ja irgendwie annähern, also ich sollte dann in meinem Diagramm irgendwie sehen können, DER Wert ist am häufigsten. Also das hätt ich schon, also auch als Werkzeug für mich, hätt ich das gerne gehabt. Ja.

F: Gibt’s denn irgendwas, das man mit den Simulationen besonders gut lernen kann?

A: Also, sich selbst zu überprüfen. Also, was man so intuitiv irgendwie denkt, wie das a aussehen könnte, das zu überprüfen. Und dadurch, denk ich, würde man auch Sicherheit gewinnen. Also dass man, ich mein, man fängt an, irgendne Aufgabe zu rechnen, das kann ja völlig falsch sein, und ich rechne dann mit dem Wert weiter oder verrenn mich total und dann könnt ich gleich überprüfen, ob das irgendwie Sinn macht. Oder auch, wenn ich das Diagramm seh, dass ich vielleicht auf ne Idee komme, oder dass sich ganz neue Fragestellungen entwickeln überhaupt, wie sieht das so aus, ja, hat schon Potenzial denk ich.

F: Ok. Wurde denn in deinem Tutorium manchmal auf R Aufgaben Bezug genommen, wenn..

A: Nie.

F: Auch nicht, wenn irgendeine andere Aufgabe erklärt wurde?

A: Ne. [Pause]

F: Ok. Dann nochmal so eine bisschen allgemeinere Frage: Woran denkst du denn als erstes, wenn du an Stochastik denkst?

A: An das unmathematischste Fach, weil es äh, also zumindest wars in der Oberstufe schon so und das hat sich auch in gewisser Weise bestätigt, für mich nicht logisch ist. Also, es ist nicht so konsequent mathematisch, als dass ich nur stur irgendwelchen, ja, diese Formelsprache ist ja in gewisser Weise noch abstrakter, mit viel mehr Worten verknüpft, denke ich, als rechnerisch, die Rechengesetze, die man sonst so kennt, benutz ich ja überhaupt nicht, auch das, was wir so in der Vorlesung gesehen haben, diese Schlussfolgerungen und Rechenketten, also wenn da mal ein plus vorkam, stand es bestimmt nicht für Addition [lacht], also Stochastik generell, ich weiß, wie wichtig das ist fürs Leben und für die Realität, aber auch Kombinatorik allein schon ist nichts, was mir zufliegt oder was sich für mich so erschließt.

F: Ok. Und an was für ein Objekt oder was für einen Begriff denkst du?

A: Bei Stochastik? Oh Gott, ja, Kombinatorik, ähm, [Pause].

F: Ok, und kannst du erklären, was das ist?

A: Ähm, dass ich, ja, mit Kombinatorik versuch ich ja auszurechnen, wieviele Möglichkeiten es gibt und dann Günstige durch Mögliche, also in jedem Falle wieviele Möglichkeiten gibt es für verschiedene Ereignisse.

F: Ok, und fällt dir noch ein anderer Begriff ein?

A: Ähm, Wahrscheinlichkeitsrechnung insgesamt, also ich würd das so übersetzen, dass Stochastik Wahrscheinlichkeitsrechnung ist. Dass ich also Wahrscheinlichkeiten für verschiedene Ereignisse und so weiter ausrechne dann.

F: Und was ist ne Wahrscheinlichkeit?

A: Hm, wie wahrscheinlich etwas ist, dass ein Ereignis eintritt. Also, ich glaub ich bin einfach noch sehr am Anfang irgendwie. Aber, ja, ja.

F: Was verstehst du denn unter stochastischer Intuition? Das ist ein Begriff, der in den Fragebögen vorkam.

A: Also auf jeden Fall etwas, was ich nicht hab. [lacht] Also dass ich jetzt intuitiv ne Idee hätte, wie sich was verhält, ähm, wie sich Wahrscheinlichkeiten zueinander verhalten oder ja, wie wahrscheinlich was ist bei der Verkettung von äh Ereignissen, ja oder halt [Pause] Werden Ereignisse verkettet? Ja, also [Pause] bedingte Wahrscheinlichkeiten, also dass ich irgendwas hab, das abhängig ist und wie wahrscheinlich ist dann, dass das eintritt. Und dass ich jetzt intuitiv sagen könnte, das ist eher wahrscheinlich oder eher unwahrscheinlich. Und das Ganze ja allein schon in, also symbolischer, in Formelsprache, aber, [Pause], das ist, dann hab ich manchmal so ne Intuition und denk sofort, bestimmt falsch [lacht], also da verlass ich mich ÜBERHAUPT nicht drauf und dann bin ich im Gegenteil sogar überrascht, wenn ich dann in der Vorlesung mal recht gehabt hätte, ja, aber das ist SO fehleranfällig, da kann ich mich nicht drauf verlassen. Und unbegründet.

F: Und gibt es einen Unterschied zu Verständnis von Stochastik?

A: Ähm, [Pause], ähm, das ist halt für mich echt schwierig, weil ich ja nicht wirklich was verstanden hab. [lacht] Ähm, ich, würd glaub ich, also bei mir sind das irgendwie zwei verschiedene Ebenen, einmal diese Formelsprache, also dass ich irgendwie sage, also wenn das so ist dann ist das so, wenn ich dann mal irgendwie durchgeblickt hab, und intuitiv wär dann für mich, dass ich sage, ich würde aber erwarten, dass irgendwie so was rauskommen würde, dass die Wahrscheinlichkeit null ist, dass es überhaupt nicht eintritt, oder so. Ähm, also Verständis, natürlich sollte man dann auch verstehen, was man sich da intuitiv irgendwie gedacht hat, aber das ist, tu ich in den seltensten Fällen.

F: Ok, aber Verständnis ist dann für dich schon eher mit Formeln kombiniert?

A: Ne, ja, also, dass ichs nachvollziehen kann, also ich brauch jetzt nicht immer per se so diese lange Kette an Schlussfolgerungen, aber, ja, also ich mein Formelsprache ist ja im Endeffekt auch nur ne Abkürzung für das, was es eigentlich ist, und, das zu verstehen, und es vielleicht auch gar nicht so kurz zu machen, oder wiederum aber so prägnant, dass man es versteht ohne dass es so abgekürzt ist, dass man die fünf Begriffe erstmal nachschlagen muss, also, das fällt mir auch nicht leicht, also so Formelsprache übersetzen geht noch, aber sofort die Erkenntnisse zu ziehen und das so zu begreifen, das dauert lange. Und das ist da genauso.

F: Könnten denn an der Stelle Simulationen helfen?

A: Hm, ja, wäre ein Baustein. Also ich möcht jetzt nicht sagen, hätt ich die Simulationen gehabt, hätt ichs sofort verstanden, aber ich denk es wär ein hilfreicher, oder es IST ein hilfreicher Baustein, ja.

F: Ok. Möchtest du denn noch irgendwas anmerken?

A: Ne, fällt mir jetzt nichts ein. [Pause]

F: Gut, dann wars das schon. Vielen Dank!

Int3

F: Gut, also dann kanns losgehen. Die erste Frage ist: Was war deine erste Empfindung, als du von den R Aufgaben gehört hast, und hat sich an deiner Einstellung bis heute etwas geändert?

A: Meine erste Empfindung, ja, also ich hatt, wie soll ich das erklären, also ich hab jetzt bei Stochastik eher damit gerechnet, dass man da nur Aufgaben zum Rechnen bzw. Beweisen hat und äh mit den R Aufgaben muss ich jetzt ehrlich sagen konnt ich auch nicht so viel anfangen und das hat sich bis heute nicht so wirklich geändert, also ich kann schon ein bisschen damit umgehen mittlerweile aber äh ich reiß mich nicht drum die zu machen, sag ich mal.

F: Ok. Und wie häufig hast du die R Aufgaben bearbeitet?

A: Ja, also gut, also ich hab mir natürlich auch öfter mal von meinen Kollegen unter die Arme greifen lassen, die in dem Tutorium waren (R Fragestunde, F.), muss ich klar sagen. Äh, wirklich selber richtig bearbeitet hab ich vielleicht die Hälfte, so.

F: Ok. Und warst du mal in der Fragestunde?

A: Nein, gar nicht.

F: Warum nicht?

A: [Schmunzelt] Weil ich mit anderen Aufgaben gut beschäftigt war und das war immer blöd, wenn man da mittendrin rausgerissen wird, wenn ich dann hier um 12, ich weiß gar nicht, die Fragestunde war ja immer relativ zwischendrin, ich bin halt immer früh im Lernzentrum gewesen, hab dann unter anderem auch die Stochastikaufgaben gemacht, halt nicht die R Aufgabe, oder andere Aufgaben und dann halt dann da rausgerissen zu werden und dann nochmal zu kommen, das hab ich dann irgendwie, das war für mich dann nicht vereinbar, sag ich mal.

F: Alles klar, das heißt du bist hier in der Kategorie „mindestens selten bearbeitet“. Fandest du die R Aufgaben im Vergleich zu den anderen Aufgaben eher schwer oder leicht und was war schwerer und was leichter?

A: Ich fands sie unverständlicher zum Teil. Ich mein die Aufgaben selber, die nicht R Aufgaben, die konnt man sich eigentlich schon relativ häufig ganz gut erklären oder erklären lassen. Bei den R Aufgaben, ich fand das so ein bißchen, gut ich war jetzt auch nicht was die Codes betrifft so richtig sicher, aber ich fand das so ein bißchen undurchsichtig und man wurde ja am Anfang mehr oder weniger direkt ins kalte Wasser geschmissen, wobei, es waren natürlich schon auch Vorbereitungsaufgaben dabei, aber ich hatte manchmal das Gefühl, woher soll ich wissen, wie jetzt dieser Code lautet, prinzipiell, also ganz alleine hätt ich die wahrscheinlich nicht hinbekommen.

F: Und gabs irgendetwas, was du an den R Aufgaben gut fandest? Und an welche erinnerst du dich noch?

A: Äh, ich kann mich noch, äh, an die äh mit dem Senderwechsel bei dem Radio erinnern, die fand ich eigentlich relativ, die fand ich eigentlich interessant, also dass man noch

ne Vergleichsgruppe gestalten muss, sonst kommt da irgend ein Blödsinn raus, dann kann man irgendwas mehr oder weniger reininterpretieren, was gar nicht da ist, das fand ich eigentlich ganz interessant.

F: Mmh. Fällt dir sonst noch eine ein?

A: Hm, gut, die einfachen mit den Münzwürfen, das war ja so der Einstieg, das fand ich eigentlich ganz interessant zu sehen, ach ich hab irgendwie ein Programm, joa, das sagt mir so und so oft ist das und das rausgekommen, das fand ich so ganz, ganz interessant. Aber sonst jetzt Weiteres [Pause] fällt mir erstmal nicht ein, nein.

F: Ok, und gibt's irgendetwas, was dir durch eine R Aufgabe klar geworden ist, was dir vorher nicht klar war?

A: Ja, also das mit den Testgruppen eben, dass man da so ne Testgruppe machen muss, damit die Aussagen nicht verfälscht werden. Das ist mir klar geworden dadurch.

F: Und gibt es irgendwas, was man deiner Meinung nach besonders gut lernen kann an den R Aufgaben?

A: Also, ich fand die in der Hinsicht eigentlich ganz cool, wenn man ja sozusagen die Sachen, die man vorher bewiesen hat, äh, dann nochmal dastehen hat, natürlich nicht so gut und so genau, aber sehr nah dran. Also das fand ich an den R Aufgaben richtig gut.

F: Also dass sie veranschaulichen?

A: Genau, ja.

F: Ok. Was hat dir bei den R Aufgaben nicht gefallen?

A: Ja, also, ich fand die, also als Laie sag ich mal, fand ich die schon relativ kompliziert, so dass man da jetzt extra auch noch, also die Lösungen hat man ja im Stützkurs bzw. im R Kurs mehr oder weniger, so wie ich das jetzt von den anderen mitbekommen habe, nachgeschmissen bekommen und ich hätt mir da gewünscht, die Aufgaben vielleicht bißchen einfacher zu machen, vielleicht auch paar mehr Tipps, aber dass man dann wirklich hingehen muss und sagen muss, ok, ich programmier das jetzt mal selber, ich brauch keinen Stützkurs, ja, weil ich hatt manchmal das Gefühl, wenn ich nicht in den Stützkurs geh, dann krieg ich davon irgendwie gar nichts gebacken.

F: Ok. Also auch nicht mit diesem Template, das im Internet war?

A: Ich weiß nicht, also, ich fand immer diese Zeile mit „Hier muss noch was eingefügt werden“, fand ich immer so ein bißchen verwirrend, sag ich mal, ich wusst jetzt nicht zu 100% was da jetzt reinkam. Auch wenn oben, natürlich stand da auch einiges dabei, das muss ich schon sagen, aber irgendwie bin ich damit nicht zurecht gekommen.

F: Ok. Würdest du denn empfehlen, weiterhin R Aufgaben zu verwenden, vielleicht auch in veränderter Form?

A: Nein [lacht], also mir wärs lieber, wenn wirklich nur Stochastik, also keine R Aufgaben drin vorkommen würden. Aber ok, ich bin da auch ein bisschen engstirnig,

glaub ich [lacht].

F: Nein, nein, ist ja gut zu wissen. Wurde denn in deinem Tutorium manchmal Bezug auf R Aufgaben genommen, als eine andere erklärt wurde?

A: Puh, also, ich war eigentlich immer im Tutorium, hm, also so spontan würde mir eigentlich nichts einfallen, ich mein wir hatten, glaub ich, mal eine Aufgabe, wo wir, das war mit den Fehlständen, da wurde, glaub ich, mal Bezug drauf genommen, aber mehr wüsst ich äh auch nicht. Ne, spontan würd mir da nichts einfallen.

F: Ok. Dann jetzt nochmal was Allgemeineres zu Stochastik: Woran denkst du als erstes, wenn du an Stochastik denkst?

A: [lacht] An Münzen, oder Zocken [lacht]

F: Ah, ok, und an was für einen mathematischen Begriff?

A: Mathematischen Begriff? Äh, puh, ja ein fairer Münzwurf oder so was [Pause]

F: Und kannst du erklären, was das ist?

A: Ja, also das ist halt mit der Wahrscheinlichkeit 0.5 dass beide Parteien, also dass es gleich, dass es gleich wahrscheinlich fällt.

F: Mmh. Und fällt dir noch etwas anderes ein?

A: [lacht] Also, ich muss ehrlich zugeben, ich war jetzt auch nicht in der Klausur und ich hab mich auch erstmal nicht weiter damit beschäftigt, weil ich, also das kommt jetzt erst für die Nachklausur. Also der NN. (Mitarbeiter des Lernzentrums, F.) hat mir mal die Einschluss-Ausschluss-Formel erklärt, das fand ich eigentlich, da hat er hier glaub ich sechs Kreise oder was gemalt, weil ich es einfach nicht verstanden hab und hat mir hier alle wirklichen Einzelheiten rausgeschrieben und das war dann am Ende sooo ein Term und, ja also daran kann ich mich noch ganz gut erinnern, weil ich das wirklich dann auch verstanden hab.

F: Mmh. Und kannst du noch sagen, wofür man die Einschluss-Ausschluss-Formel braucht?

A: Puh, [lacht], das ist ne gute Frage, äh, also jetzt zu Stochastik, ähm, also da ging's ja um, um, dass man äh sozusagen mehrere Mengen vereinigt und äh die Schnittmengen halt nicht doppelt zählt, also doppelt und dreifach, irgendwie x-fach zählt, und dass man die dann halt sozusagen rausnimmt um dann insgesamt die gesamte Menge zu haben, aber was das jetzt genau in Stochastik war, das weiß ich nicht mehr.

F: Kein Problem. Dann kommen wir zu was anderem: Was verstehst du unter dem Begriff „stochastische Intuition“? Das kam in den Fragebögen vor.

A: Achso, ja, also wenn man, äh, also sowas wie, dass man halt hergeht und man kriegt halt wirklich ne Frage, ja, also da war ja im Fragebogen ein Beispiel mit Wir werfen die Münze 10 Mal oder 20 Mal und dann waren ja verschiedene Verläufe davon gezeigt, also wieviel Mal Kopf und wieviel Mal Zahl gefallen ist und Intuition versteh ich halt darin, dass man ähm sich das halt anguckt und sagt, ok, es ist jetzt nur 5 Mal

Kopf und 15 Mal Zahl gefallen, würd ich mal sagen, ja, grenzwertig, also, kann mal passieren, aber ich weiß nicht, ob das wirklich so oft passiert. Aber wenn ich jetzt z.B. 11 und 9 sehe oder 12 und 8, dann würd ich schon sagen, ja, kann gut sein, also das ist so Intuition, dass ich sag, könnt hinkommen.

F: Ok. Und gibt es einen Unterschied zu Verständnis von Stochastik?

A: Puh, ja, ich find schon, also ich mein, wenn ich dann wirklich nen Beweis oder irgendwas versteh, dann ist das natürlich ein Unterschied, als wenn ich dann sag vorher, ach, es könnt ja so und so sein, das ist ja im Endeffekt äh sag ich ja damit nichts, ja. Es ist ja für mich persönlich die Intuition, andere sehen das vielleicht anders, die sagen, ach das könnt doch so und so sein, ja. Und äh das ist natürlich ein großer Unterschied, ob ich jetzt wirklich Fakten vor mir habe oder irgendwie von mir aus denke, ach, so oder so.

F: Also das heißt Verständnis ist für dich, dass man irgendwie einen Beweis hat.

A: Ja, klar.

F: Ok.

A: Also dass ich diesen Beweis eben auch nachvollziehen kann, sag ich mal. Das ist für mich eigentlich das Verständnis.

F: Mmh. So jetzt hab ich noch eine Aufgabe, und zwar kam ja in den Fragebögen diese Krankenhausaufgabe vor, ich weiß nicht, ob du dich noch dran erinnerst, ich hab sie hier ausgedruckt. [Pause] Und wenn du jetzt die Aufgabe hättest, dieses Problem zu simulieren, wie würdest du vorgehen? Also du brauchst keinen R Code sagen, nur die Schritte.

A: Ok, ich schau mirs mal an. [Pause]
Also wie ich das simulieren würde? Also ohne Programm, nur die Schritte?

F: Genau.

A: [lacht] Puh, also wie ich es simuliere, äh, puh, also um ehrlich zu sein, ich glaub ich würd da, wenn ich jetzt was dazu sagen würde, etwas ins Schwafeln kommen, ich wüsst es jetzt nicht hundertprozentig genau.

F: Ist nicht schlimm, sag einfach, was du für Ideen hast. Muss kein kompletter Plan sein.

A: Äh, also ich weiß jetzt nicht, was für Zahlen vorkommen, aber ich würd da erstmal hergehen und gucken, um welche Anzahlen es da überhaupt geht, um was für Anzahlen von Kindern. Und dann steht da, wir gehen davon aus, also äh, also irgendwie dann vielleicht ne Wahrscheinlichkeit sich errechnen, indem ich ne bestimmte Anzahl, also die Gesamtanzahl durch die Anzahl der Jungen oder Mädchen teile und ja und [lacht] würd ich versuchen, mir das zusammenzubandeln. Also wenn jetzt jemand neben mir stehen würde und sagen, du MUSST das jetzt machen, dann würd ich mir natürlich was einfallen lassen, aber SO ist das jetzt, ich weiß es jetzt einfach nicht genau, nein.

F: Ok. Kein Problem. Möchtest du denn noch irgendetwas anmerken?

A: An sich nicht, nee. Also, was ich so an Kritikpunkten hatte, haben Sie ja hier drauf. Sonst weiteres, wüsst ich jetzt nicht. Also wie gesagt mit dem Stützkurs, das fand ich bisschen, also natürlich ist das ne Sache, die Sie nicht machen müssen, ja, ist ja klar, aber ich hätt mir halt, es wär angebrachter gewesen, es bisschen einfacher zu machen und dafür die Leute bisschen selber versuchen zu lassen. Was, was auch ganz cool gewesen wäre, wär halt, wenn auch im Lernzentrum, gut, kann man vielleicht nicht erwarten, aber wenn da ein paar Leute da gewesen wären, die einem so ein bisschen unter die Arme können. Sie waren ja so ab und an mal da, aber ich weiß nicht, aber wär schon cool gewesen, wenn sich da einer gut ausgekannt hätte, also richtig. Weil ja, also die NN (Mitarbeiterin des Lernzentrums, F.) kannte sich so ein bisschen damit aus, aber ja, es wär schon cool gewesen, noch nen Ansprechpartner so außerhalb der Fragestunde zu haben. Ja. Aber sonst nichts.

F: Ok, dann vielen Dank!

Int4

F: Gut, also die erste Frage ist: Was war deine erste Empfindung, als du von den R Aufgaben gehört hast, und hat sich an deiner Einstellung etwas geändert bis heute?

P: Hm, also meine allererste Empfindung war: Das ist totaler Scheiß. Weil mit PCs […] konnt ich überhaupt nicht. Ich glaub nicht, dass ich jetzt mit PCs wesentlich besser kann, aber ich hab so ne Grundidee davon, wie PCs funktionieren. Also, am Anfang extrem negativ und jetzt positiv. Also, ich glaub nicht, dass ich jetzt alles, diese R Aufgaben, jetzt partout hinbekommen würde, aber ich hab dafür jetzt viel größeres Verständnis, z.B. also wo ich nachfragen könnte und was ungefähr so die Idee dahinter ist.

F: Mmh.

P: Also das ganz eindeutig.

F: Mmh, perfekt. Freut mich natürlich. Ähm, wie häufig hast du die R Aufgaben bearbeitet? Ist die nächste Frage.

P: Oft. Moment […] Häufig im Sinne von Wie lange?

F: Ne, ne, einfach häufig im Sinne von wie oft, also welchen Anteil davon du gemacht hast. Und du warst ja auch eigentlich immer in der Fragestunde.

P: Ja, schon viel.

F: Viel, genau. Und du warst auch oft in der Fragestunde?

P: Ja.

F: Gut, dann fällst du hier in diese Kategorie. Das sind jetzt erstmal inhaltliche Fragen zu den R-Aufgaben bzw. Fragen zu den R-Aufgaben selbst. Ähm, fandest du die R-Aufgaben im Vergleich zu den anderen Aufgaben eher schwer oder leicht und was war schwerer und was war leichter?

P: Ich fands sie eher schwieriger, weil man eben zusätzlich zu Stochastik noch sich in eine ganz andere PC-Gedankenwelt eindenken musste und das war aber auch schon das eigentliche Schwierige daran, also die Aufgaben an sich waren gar nicht schwer, du bist ne ziemlich Nette und du hast sie eher leicht gestellt im Vergleich zu den anderen. Ich glaub, das einzig schwierige war wirklich, dass man einfach, ja, versteht, wie ein PC denkt. Eben, obwohl man die Mathematik eben rafft und genau weiß, worum's eigentlich geht, man das noch dem PC klar machen muss. Also so eine Art doppeltes Übersetzen. So die Aufgabe in Mathe und dann eben von Mathe PC. Mm. Ich find, dieser Schritt von Mathe in PC war immer das, was ich nicht hinbekommen hab.

F: Mmh. Ok.

P: Und dann dieses ganz genaue Arbeiten, dass der PC schon nicht weiß, was er machen soll, wenn du ein Komma zu viel hast oder eins zu wenig oder ne Klammer ist rund statt eckig, sowas. Ja das war ziemlich gut, das hat mich viele Nerven gekostet,

einfach nur mich mein Skript angucken und es heißt Fehler und für mich heißt Fehler eigentlich, da ist irgendwas falsch, aber eigentlich heißt es, da ist nur ne Klammer falsch oder ein Leerzeichen zu viel, so was.

F: Mmh.

P: So formale Sachen. Fand ich sehr anstrengend.

F: Mmh. Ähm, und kannst du sagen, was war leichter an den Aufgaben?

P: Was mir die Aufgaben leichter gemacht haben oder was an der Aufgabenstellung an sich leichter war?

F: Was an der Aufgabenstellung an sich leichter war. Weil du eben meintest, dass es an und für sich doch leichter war, außer jetzt diese Übersetzung.

P: Es war immer sehr, sehr greifbar. Ich fand, die waren immer sehr anwendungsorientiert, man wusste halt auch, warum man sich damit beschäftigt, was hatte man davon, das war echt schnell geklärt. Und ich fand, die waren einfach so… ja, greifbarer, orientierter. Eben nicht so was wie, ich berechne mir jetzt etwas, einfach nur damit ich sagen kann in der 7. Dimension bei Vollmond, wenn eine Jungfrau eine Rose ohne Handschuhe pflückt, dann […] ist da ne Nullstelle. Sondern es war wirklich so was, wo ich gedacht hab, hey, das kann man vielleicht wirklich mal irgendwo gebrauchen. Das fand ich wirklich gut.

F: Ok, wunderbar. Also, du hast es schon ein bißchen mit beantwortet, also, was fandest du an den R Aufgaben gut? Und an welche erinnerst du dich noch und warum?

P: Mmh, also das ist blöd, weil wie gesagt, ich hab vor 2, 3 Wochen angefangen, Stochastik zu lernen und ich bin noch die R Aufgaben nicht durchgegangen, also ich erinnere mich an Crabs, erinnere ich mich nach wie vor heute, weil die Aufgabe hab ich immer noch nicht ganz verstanden, ich erinnere mich an das Würfeln, irgendwas mit Abhängigkeiten, ich erinnere mich an die letzte Aufgabe, wo wir einfach erklärt haben, worum's ging.

F: Weißt du noch, worum's ging?

P: Bei der allerletzten? Oh ja, da hab ich viel drüber nachgedacht, wir hatten bedingte Wahrscheinlichkeiten, wir hatten quasi ein zweistufiges Experiment und haben dann rausgefunden, also auf der einen Seite, dass man der Intuition irgendwie nie trauen darf bei der Statistik [lacht] oder Stochastik, und auf der anderen Seite eben, wenn man ne Menge hat, wo nehmen wir mal an, da sind rote und schwarze Kugeln drin und wir haben eigentlich nur rote und nur ganz wenige schwarze, muss das nicht so ein großes Gewicht haben, wie es auf den ersten Blick scheint. Und das hab ich da mitgenommen. Weil es eben diese zwei verschiedenen Wege gibt. Genau, wir hatten da auch so ne Aufgabe aus der Schule mit Wie gut Schwangerschaftsteste funktionieren, genau das gleiche, das hab ich mir dann extra nochmal angeguckt. Und das ist krass, von wegen 99,9 % ist fürn Arsch, also da kommt knapp ein bißchen mehr als 50% raus. Aber da ist es wieder so ne Sache, wo man sagt, hey, das kann man wirklich gebrauchen, das ist mal sinnvolles Wissen, dass ich auch jemandem erzählen würde, der Mathe nicht mag und ja, ich fand das einfach wichtig. Ich find, so was ist was, wo ich sagen würde, deshalb sollte jeder Stochastik ein bißchen können, weil so

was, so was gibt es eben, ne. Das fand ich echt gut.

F: Ok, ähm, gibt es etwas, was man an den R-Aufgaben deiner Meinung nach besonders gut lernen kann?

P: Ähm, den Umgang mit, äh, was ist R, einem Rechenprogramm, einem Darstellungsprogramm, den auf alle Fälle. Hm, durch die Graphiken, die einem, die man von R bekommt, kann man auch sehr viel Sachen verdeutlichen, die man sonst nicht verdeutlichen oder die man sonst nicht so gut verdeutlichen könnte. Z.B. kann man mit R ja Graphiken machen, die auf 10 Milliarden Versuche oder 10000 Versuchen basieren. Das kann man natürlich real nicht durchwürfeln. Das kann man. Auch Sachen abschätzen, mit extrem hohen Wahrscheinlichkeiten, oder allein schon, dass ich Sachen ausrechnen kann, also ich glaub, ein normaler Taschenrechner [?] schon ab 15 Fakultät und R bekommt das noch locker hin. Problem daran ist halt, was ich bei mir und auch bei vielen andern gesehen hab, dass man zwar mit der Graphik, dass man die Graphik sehr gut interpretieren konnte und auch immer wusste, worum es ging in der Graphik, aber R dann dazu zu bringen, dass es mir diese Graphik ausspuckt, das ist dann nochmal ein anderes. Und da hätt ich mir eigentlich, also da hätt ich den Vorschlag von Herrn [xxx] fast gemacht, dass er vielleicht beides macht. Dass er sowohl sagt, also hier ist ein R-Programm, das geben Sie einfach ein und das spuckt ihnen ne Graphik aus und Sie interpretieren diese Graphik einfach mal, weil ich glaub, Stochastiker und Statistiker machen das ganz viel, dass sie einfach ne Graphik haben und die versuchen zu interpretieren. Und das man eben gleichzeitig dieses R auch selbst programmiert, quasi zwei R- Aufgaben hätte.

F: Mmh, mmh.

P: Ja, weil ich glaub, die Interpretation von Graphen ist in der Stochastik wichtig. Sehr wichtig.

F: Ja.

P: Finde, die sollte auch nicht zu kurz kommen, also ich mein, ja, da hat man wieder diese Sachen, da kann man sehr viele Sachen mit ausdrücken, z.B. die Sache mit den Krankenbetten und so und diese Abhängigkeit und Unabhängigkeit, also wann ist es unabhängig und wann ist es nicht unabhängig. Das ist nicht wirklich [?] , so was. Und ja, das Programmieren. Ich glaub, das Programmieren ist einfach extrem wichtig für die Mathematik und wird auch immer wichtiger für die Mathematik, weil das so … viele der modernen Probleme haben ne ganz andere Dimension, wenn man die mit Rechnerleistung angeht.

F: Mmh.

P: Und ich mein, es gibt einen Mathematiker, der hat sein Leben lang die Nachkommastellen von pi berechnet, irgendwie, kam auf 200 richtige, hatte einen Fehler gemacht und der Rest des Lebens ist falsch, hat natürlich, wenn man zurückdenkt, einen ganz anderen Stellenwert als wenn man heute sagt, hier sind 200 Stellen von pi, das hat irgendwie ein Hochleistungs-PC innerhalb von einem Tag gemacht, ja. So ein ganz anderes Gewicht.

F: Mmh

P: Und deswegen sollte man auch einfach PCs integrieren, so früh, wie's geht. Aus gutem Grund hat man, glaub ich, im 1. Semester so ne Vorlesung dafür, wie man mit PCs umgeht und ich glaub, das ist einfach extrem wichtig und wird immer wichtiger werden.

F: Mmh, ok. Und gibt es irgendwas inhaltliches, was du jetzt irgendwie durch eine R Aufgabe gelernt hast oder was dir klargeworden ist, was dir vorher nicht klar war?

P: Diese zweistufigen Tests auf alle Fälle und ähm mehr oder weniger Abhängigkeiten und Unabhängigkeiten. Verstanden hab ich die zwar irgendwie immer, aber was für ein Gewicht das haben kann, wurd mir erst durch R klar. Ja, weil ich mein, die Formel ist jetzt nicht allzu schwer, aber was diese Formel, was diese eine Zeile aussagen kann, das ist nochmal ein anderes [?] und das wurd mir durch R viel klarer durch diese Aufgabe mit diesem Krankenhaus.

F: Mmh, ok. Schön. Ähm, was hat dir bei den R Aufgaben nicht gefallen?

P: Dass R so rumpienst [lacht] , also die Genauigkeit, die man wirklich extrem an den Tag legen muss, noch viel mehr als normal bei Mathematik und ähm auch so ein bißchen das hilflosige, das Hilflosigkeitsgefühl, dass ich zwar die Aufgabe verstanden hab und in der R Fragestunde war und mir das eigentlich alles mitgeschrieben hab und ich den Fehler trotzdem nicht gefunden hab, weil's dann doch irgendwo ein Komma ist, an irgendeiner Stelle, wo es 5 Kommas sind und ich [?]. So was. Ansonsten, ich fand's das echt ziemlich gut, äh erfrischend auch mal, eine neue Dimension, grad so als Lehrämtler hat man von PC-Kram nicht allzu viel, das, ich find das wichtig, das ist so ne neue Ebene, die man betreten sollte.

F: Ähm, würdest du denn aus Studentensicht empfehlen, weiterhin R-Aufgaben zu verwenden und würdest du irgendwas ändern?

P: Hm, ich hab mir mal überlegt, vielleicht könnte man mal für's ganze Studium ein Programm haben, ich weiß nicht, ob so was geht, aber ich find das Schwierigste ist ja meistens, dass man die Befehle kennt und dass man sich ungefähr in dieses Programm eindenkt und das ist natürlich immer wieder Arbeit, grad wenn man 5 Programme auf einmal hat. Und da hätt ich mir oder ich hab mir überlegt, es wär natürlich am didaktisch sinnvollsten, wenn's ein Programm gibt, das alles kann. Ne, es muss halt jeder haben und ich wette, so ein Programm ist bestimmt mega schweineteuer und das kann sich die Uni gar nicht leisten, geschweige denn der Student. Aber das fände ich sinnvoll. Und ähm ansonsten würd ich versuchen, viel mehr zu betonen, warum es denn diese PC-Aufgaben eigentlich gibt. Also in der Stochastik stellt sich diese Frage eigentlich nicht so sehr, weil man einfach den PC so der so braucht, um diese 10000 Versuche zu simulieren, aber ich mein, wenn wir erstmal die Numerik weglassen, die ja quasi nur mit PCs umgeht, wozu braucht man den eigentlich PCs in der Mathematik? Ich mein, alles klappt mehr oder weniger, ich mein, die Algebra klappt mehr oder weniger ohne PC, Geometrie kann man vielleicht am PC veranschaulichen, kann aber auch ohne klappen. Und ich find, vielleicht könnt man einfach versuchen, den PC so ins Mathestudium zu integrieren, dass es eben nicht so abgehackt ist, so wir benutzen jetzt den Pc, wenn wir was am PC machen müssen, sondern dass man die Studenten wirklich mal an den Punkt hier ranführt, wo man den PC einfach braucht. So dass man z.B. pi sehr genau berechnen möchte oder man approximiert die Nullstelle oder auch in der Algebra, dass man z.B. nen Eigenwert berechnen möchte oder, was weiß ich, irgendwas, was zwar ne Formel hat, aber die Formel würd total lang dauern und ist dann natürlich total fehleranfällig, [?] dass man dafür PCs braucht und dass die dann auch anerkannt sind, weil es sonst immer heißt, man soll alleine rechnen und ich find, da könnt man betonen, wofür man die PCs hat und wo die einem helfen und wo die einem nicht helfen, es ist nämlich oft so, dass die Studenten wissen, welches Programm sie haben.

F: Mmh, aber du würdest schon aus Studentensicht empfehlen, R Aufgaben weiter zu verwenden?

P: Ja, definitiv.

F: Ähm, wurde denn in deinem Tutorium manchmal auf R-Aufgaben Bezug genommen, als eine andere Aufgabe erklärt wurde und erinnerst du dich, wenn ja, an Beispiele?

P: Mmh, also mein Tutor war jetzt nicht besonders motiviert darin und hat auch überhaupt keine Ahnung von R gehabt und also das Tutorium bei meinem Tutor war eh ganz, ganz komisch, weil er zwar irgendwie, ja wie soll ich sagen. Er hat sehr ausführlich Tipps gegeben, nach dem Motto, ok, ich rechne euch das mal komplett vor mit dem Faktor 4 und ihr macht's halt für den Faktor 5, aber ich finde dann hat man, also man hat zwar die Aufgabe irgendwie gerafft, ne, weil man sie nachvollziehen konnte, aber man hat nicht viel dabei gelernt, weil man hat halt nur nachvollzogen, was er gerechnet hat und wirklich […] also vielleicht hat er mal gesagt, ja, das ist so ähnlich wie die eine R Aufgabe, aber wirklich, wo das wirklich geholfen hat, wo es irgendwie eine ganz andere Dimension geöffnet hat, das kam leider nicht.

F: Mmh, mmh. Ähm, so jetzt nochmal was allgemeineres zu Stochastik. Ähm, woran denkst du als erstes, wenn du an Stochastik denkst?

P: Hm, soll ich das wirklich sagen?

F: Mmh.

P: Ich denk immer an den NN [Mitarbeiter des Lernzentrums].

F: Ok [lacht]. Und an was für einen mathematischen Begriff oder was für ein Objekt?

P: Hm, also ich weiß nicht. Ich denke hm, also ich denke, wenn ich an.. Also ich denke natürlich nicht nur an NN [lacht], NN ist nur so, dass hat sich irgendwie so in meinen Kopf eingebrannt, wie er daran verzweifelt ist, mir Stochastik zu erklären und ich hab festgestellt, ich interessier mich anscheinend sehr für Statistik, und deswegen denk ich auch immer viel an diese Krabbenpopulation, die wir mal hatten.

F: An die Krabbenpopulation?

P: Hast du mal ne Vorlesung gemacht. Da wurden 53 Krabben gezogen, da sind z.B: 23 weiblich und was weiß ich der Rest männlich. Was kann man daraus für [?] schließen, ich find, das ist interessant, dass du aus einem Satz Daten versuchst, Information zu ziehen, ich mein, du kannst diese Krabbenaufgabe eins zu eins übersetzen auf wie sinnvoll ist ein Medikament zum Beispiel oder was weiß ich für Versicherungen könnt man das auch machen, glaub ich, wenn man so und die Leute hat und von denen haben die und die die Verletzung, lohnt es sich, das zu versichern. Ok, ist vielleicht ein bisschen assi, aber ich glaub, das könnt man auch machen. Aber das, das find ich sehr

interessant und ich glaub, da spielt auch R wieder eine ganz andere Rolle, weil man R vielmehr braucht und so weit ich weiß, wird R auch verlangt in den Statistikvorlesungen aus gutem Grund, weil man da eben versucht Information aus den Daten zu gewinnen und die irgendwie aufzubereiten und darzustellen für die, die eben nicht die extreme Leuchte in der Statistik sind. Ich glaub das ist sehr wichtig und diese, diese Population und die statistischen Gedanken im weitesten Sinn ist das, woran ich als erstes denke und was ich gern vertiefen würde und ich würd auch gern in die Statistik-Vorlesung gehen, wenn ich Zeit hab. Also muss ich so oder so machen und deshalb werd ich's auch tun.

F: Mmh. Und fällt dir noch irgendwie so ein stochastischer Begriff ein?

P: Zwei-Stufen-Modelle und Unabhängigkeit.

F: Mmh. Und kannst du erklären, was das ist?

P: Jetzt mathematisch oder in meinen Worten?

F: Wie du möchtest.

P: Ok, Unabhängigkeit drückt einfach nur aus, ob zwei Ergebnisse unabhängig voneinander auftreten oder nicht. Das berechnet sich dadurch, dass ich einfach gucke, wie nenn ich das eigentlich, also P von X_1 Element A_1 Komma X_2 Element A_2 ist gleich P von X_1 Ele, ähm, und P von X_2. Ja, also so sehr unscheinbar, aber wie gesagt es kann sehr, es ist ein sehr mächtiges Instrument, weil es sehr extreme Auswirkungen hat. Wenn ich wieder zurückgehen würde auf die Statistik zum Beispiel, also ich glaube das statistische Pendent dazu ist die Korrelation, also zum Beispiel, wenn ich zwei, hab ich was sehr lustiges drüber gelesen, wenn ich zwei verschiedene Entwicklungen hab, müsst ich eigentlich immer prüfen, ob die abhängig oder nicht abhängig sind, also zum Beispiel wenn die Politik sagt, seit 2002 gibt es irgendwie 500 000 weniger Arbeitslose, also klappt Hartz 4 irgendwie, muss aber nicht stimmen. Also ich hab da was gelesen von einem Herrn ich glaub Henze war das oder so, der hat geschrieben, wenn naja, also vor 20 Jahren gabs von mir aus weniger Kinder auf der Welt, also weniger, weniger Menschen und vor 20 Jahren gabs von mir aus auch weniger Störche auf der Welt und heute gibt's mehr Störche und mehr Kinder, das ist dann der Beweis dafür, dass die Störche die Kinder bringen.

F. Mmh

P: Stimmt halt nicht, ne. Und deswegen müsste man halt bei so was, also da bräuchte man A wieder ein Grundwissen in Stochastik, damit einfach so was anzweifelt und man müsste halt auch einfach immer bei so Daten auch einfach mal die Mathematik dahinter mit angeben. Wie groß ist denn die Abhängigkeit oder gibt's überhaupt ne Abhängigkeit, ist das vollkommen abhängig, ist das unabhängig? Das ist so was, das sind sehr wichtige Fragen, die man eigentlich dazu beantworten sollte.

F: Mmh. Und bei dieser Formel, die du gerade genannt hast, kannst du da sozusagen in eigenen Worten sagen, was das bedeutet? Also, wenn du hast, P von ...

P: Achso. Ähm. Wir stellen uns eine Urne vor und wir ziehen zwei Kugeln raus. Irgendwelche, von mir aus die Kugel mit der Nummer 7 und die Kugel mit der Nummer 8. Wir haben die 7 und die Nummer 8 gezogen. Wenn […] ich […]diese

beiden Kugeln gezogen habe, und die sind unabhängig davon, heißt das […] ist ne gute Frage […] Was heißt das denn eigentlich? […]Es gibt keine, keine Abhängigkeiten, mh, ich hab darüber noch nie nachgedacht, das ist ne gute Idee. Ähm […] Ich weiß es gar nicht, glaub ich. Also ich mein, es gibt keine Abhängigkeiten und es macht keinen Unterschied, ob du die quasi […] nein, das Ereignis, dass du quasi beide ziehst, ist, wie soll ich sagen, es hängt von nichts ab, es ist quasi purer Zufall. Also wenn du sagen würdest, etwas ist nicht unabhängig, also es ist abhängig, hieße es ja, wenn ich die 7 ziehe, würde sich irgendwie die Chance erhöhen, dass ich die 8 bekomme. Wenn sie unabhängig sind, ist das eben nicht so. So würd ich das für mich, glaub ich, grad stehenlassen.

F: Mmh. Ähm, was verstehst du unter stochastischer Intuition? Das ist so ein Begriff aus den Fragebögen.

P: Ähm ja, traue auf keinem Fall deiner stochastischen Intuition! [lacht]

F: Ok.

P: Ähm, ja, ich weiß es nicht. Ich weiß nicht, ob die sich auch im Laufe der Ausbildung schult, so viel Stochastik hab ich noch nicht gemacht, aber bis jetzt wurd fast immer meine stochastische Intuition, also was man sich so denkt, was rauskommen müsste, kam irgendwie nie raus. Ja, oder wenn, nur ganz selten. Und wenn es rauskam, kam noch ein richtig dicker Klotz mit bei, ABER, aber es gibt da noch so ein Hintertürchen, was man hätte aufmachen können. Also, ich weiß nicht, es ist ganz komisch, irgendwie, diese Intuition in der Stochastik ist sehr, sehr irreführend und klappt sehr, sehr wenig. Vielleicht liegt es daran, dass man Stochastik auch generell recht wenig hatte in der Schule und ähm, ich kann mir gut vorstellen, dass dieses Gefühl für Zahlen, für Wurzeln, für Quadratzahlen einfach länger und besser ist, weil man einfach in der Schule viel mehr damit gerechnet hat als in der Stochastik. Und ich glaub einfach, diese, dieser Begriff des Zufalls, das ist sehr schwer zu begreifen, also allein schon, ich mein, wenn ich jetzt nen Würfel würfele und ich hab jetzt, ich hätte jetzt 20 Mal nur ne 1 gewürfelt und ich würd jetzt sagen, eigentlich ist doch die Chance, dass jetzt nochmal 1 kommt, gleich Null, aber das stimmt zum Beispiel gar nicht. Allein, allein das ist schon so ein Punkt, da setzt die Intuition sehr aus, weil man dann eben nicht die gesamte Kette betrachten muss, sondern einfach nur den nächsten Schritt. Und das ist, das ist schwierig, ich glaub, das ist sehr viel verzweigt, was da wie eben abhängt und was da nicht voneinander abhängt. Das, das ist schwierig. Von daher würd ich meiner Intuition in der Stochastik überhaupt nicht trauen, was auch für mich viele Aufgaben extrem schwierig macht, weil ich […] Wenn man sich nicht komplett einfach auf die Mathematik verlassen kann, sich irgendwas herleiten muss, das ist, das find ich extrem schwierig. Also wenn es nicht grad so ein kombinatorisches Problem ist, wo man ein bißchen abzählt. Aber, ja, das ist einer der Punkte, die Stochastik für mich extrem schwierig und auch extrem spannend machen, weil man da eben immer was Neues hat, man hat immer was Überraschendes, aber man hat auch einfach […], also ich brauch für Stochastik, glaub ich, immer am allermeisten, um überhaupt irgendwas hinzubekommen, irgendwas selbst zu rechnen, weil ich ganz, ganz lange, einfach die Masse der Zeit einfach auf dem Schlauch stehe und erst mal nen Ansatz brauche. Das hab ich nirgendwo so stark wie in Stochastik, das hatt ich nicht mal in Analysis 1 und ich war da sehr, sehr schwach Schrägstrich schlecht am Anfang vom Mathestudium, aber sogar da hatt ich irgendwie ein besseres Gefühl dafür, was rauskam, worum es geht, als jetzt in der Stochastik.

F: Mmh.

P: Dafür, wie gesagt, hat die Stochastik immer so einen Aha-Effekt.

F: Mmh.

P: Lohnt sich auch, also es ist so ein zweischneidiges Schwert. Es ist sehr anstrengend für mich, es ist sehr komplex, und ich hab oft genug Bock, den ganzen Kram hinzuschmeißen, und zu sagen, ach, studier ich eben nicht nebenbei noch Mathe, werd ich eben nur Lehrer, ist ok, aber eigentlich will ich das nicht und ich will das schon noch raffen.

F: Mmh.

P: Also, stochastische Intuition, schön, wenn man die hat, schön, wenn die was bringt, aber ich hab sie nicht und deswegen, ich würd am Anfang niemandem raten, sich zu sehr darauf verlassen, wenn er nicht grad wirklich extrem, ja extreme Erfahrung in Stochastik hat. Also ich könnte mir gut vorstellen, dass die Intuition vom Herrn Kersting wesentlich besser ist als meine, weil er einfach mehr Erfahrung hat, ich hab sie nicht und deswegen kann ich mich auf keinen Fall darauf verlassen.

F: Mmh. Und gibt es einen Unterschied zu Verständnis von Stochastik?

P: Ja, auf alle Fälle. Ähm, Verständnis von Stochastik kann ja auch heißen, dass ich einfach nur die Mathematik verstehe. Also, ich kann ja quasi Beweise oder Rechenschritte, Herleitungen einfach verstehen, mathematisch, ohne jetzt sag ich mal den so großen Sinn dahinter zu verstehen. Also, auch wieder so, beim zweistufigen Experiment kann ich verstehen, was man da rechnet, der Rechenarbeit ist sehr gering, aber der eigentliche, die eigentliche Aussagekraft ist ja jetzt nicht, dass man diese zwei Zahlen irgendwie miteinander malnehmen kann, oder durcheinander teilt, das ist nicht so wichtig, die Aussage dahinter, die ist nicht immer zu verstehen. Und ähm, stochastisches Verständnis heißt eben für mich auch, dass man diesen Aha-Effekt dahinter mitbekommt, dass man eben jetzt weiß, ok, ich brauch das wirklich weil... Nicht nur, damit ich irgendwie zwei Zahlen miteinander multipliziere, dass ich das auch mal wieder mache, sondern damit ich auch wirklich sagen kann, warum brauch ich das denn, was steckt denn dahinter. Wenn das wirklich jeder Mathestudent im ersten, äh im zweiten Semester mitbekommen soll, muss das ja irgendeinen Grund haben. Das gehört für mich dazu.

F: Mmh.

P: Äh, was war die Frage nochmal?

F: Was der Unterschied zu Intuition ist. Also...

P: Ja, also ich glaub, die Intuition braucht das Verständnis. Ich glaub, je mehr Verständnis man hat, desto besser wird die Intuition.

F: Mmh.

P: Und desto mehr lernt man glaub ich auch mit diesen Zweifeln an der Intuition umzugehen. Ich glaub, das ist ein sehr wichtiger Punkt.

F: Mmh. Ok,schön. Jetzt hab ich noch eine Aufgabe für dich. Und zwar gabs in den Fragebögen diese Krankenhausaufgabe, ich hab sie hier nochmal ausgedruckt, wo man

P: Ah, ja

F: wo man zwei Krankenhäuser hat, ähm, ein großes und ein kleines, und ähm, du kannst es auch gern nochmal in Ruhe durchlesen, und die Frage wäre: Wenn du jetzt die Aufgabe hättest, diese Krankenhausaufgabe zu simulieren, welche Schritte würdest du machen? Also jetzt nicht den genauen R Code, sondern nur, welche Schritte würdest du machen? Kannst dir auch gern ein bißchen Zeit nehmen, um es in Ruhe durchzulesen.

P: [liest] Für wen möchte ich das denn simulieren?

F: Also du möchtest wissen, was ist das Ergebnis davon und weißt nicht, wie du es ausrechnen sollst.

P: Ah, ok, also für mich quasi. Ok.

F: Mmh.

P: Ähm, also spontan, hab ich schon beim ersten Mal dran gedacht, es gibt natürlich diese absolute und die relative Wahrscheinlichkeit. Ne, also, ähm, weiß nicht, stell dir vor, du hast drei Kinder, da kann ich sagen, ok, auf eine Mutter kommen drei Kinder, das muss aber im Bundesdurchschnitt nicht heißen, dass auf eine Mutter drei Kinder kommen und so was kann man da zum Beispiel denken. Meine erster Gedanke war auch, dass wahrscheinlich im ersten, im kleinen Krankenhaus leichter ist, über diese 60 Prozent Hürde zu kommen, weil es da ja einfach weniger Mütter gibt. Auf der anderen Seite gibt es ja auch weniger Mütter, die potenziell Kinder bekommen könnten, also vielleicht ist es auch einfach die […], obwohl es leichter ist, über die 60 Prozent Hürde zu kommen, die Anzahl der Tage geringer, wo man über die 60 Prozent Hürde kommt. Und ähm..

F: Mmh. Und wenn du das jetzt aber mit Hilfe des Computers rausfinden willst, was die Antwort ist

P: Ah, ok.

F: Genau, also wenn du jetzt ein Programm schreiben solltest, was wären deine Schritte oder wie würdest du vorgehen?

P: Ich darf darauf rum malen? [meint das Blatt]

F: ja, sicher

P: wunderbar, ok, was haben wir denn da? Also wir haben das K 1, K1 groß, 45 Kinder [schreibt] ok, ja, und [schreibt] Also, ich hab mal so angefangen. Jetzt würd ich mich erstmal fragen, welche Verteilung würd ich R dafür anbieten. Also, ich würde versuchen, etwas zu […], ach ne, ist ja viel leichter, als ich gedacht hab […] Ok, also ich muss irgendwie dafür sorgen, dass jeden Tag also oder jede Einheit aus einer Masse von uniform verteilten Kugeln 15 oder 45 rausgezogen werden und ich zähle dann einfach die roten, weil die für die Weibchen stehen. Und dann später, also das

mach ich dann 365 Mal, wenn wir nicht im Schaltjahr sind

F: mmh

P: und lass dann quasi, definier mir dann noch ne for- Schleife, ne, ich definier mir keine for-Schleife, ich brauch nen Parameter, der bei Null anfängt und zu dem immer +1 addiert wird, wenn ich eben 60% hab oder mehr

F: mmh

P: ok, das heißt, ich muss erstmal random ziehen, wie hieß der Befehl nochmal...

F: Befehl ist nicht wichtig, das ist schon wunderbar

P: mmh, also ich muss R dazu bringen, dass es zufällig aus einem Haufen zieht, in dem nur Mädchen und Jungs drin sind und uniform natürlich, also ich brauch ne Menge mit 2 Elementen, aus dem R uniform zieht

F: mmh

P: eben entweder 15 Mal oder 45 Mal.

F: mmh

P: Dann brauch ich so etwas wie eine Variable 60 plus, die ordne ich der Null zu und dann brauch ich ne for-Schleife […] mit der if-Abfrage, dass meine Variable quasi eins dazu bekommt, jedes Mal, wenn an meinem Tag x irgendwie die Anzahl der Kinder 60 Prozent übersteigt, das heißt, ich brauch auch noch irgendwo etwas, was mir den Anteil von Mädchen und Jungs darstellt. Das kann ich durcheinander teilen, dann hätt ich das schon. […]

F: mmh, und wenn du das gemacht hättest, was hättest du dann simuliert?

P: […] dann hätt ich quasi simuliert, dass ich aus einer Urne mit zwei Elementen drin so und so oft rausziehe und einfach gucke, wie oft ist es äh ein Junge oder ein Mädchen und dann könnt ich mir die Wahrscheinlichkeit von einander ja äh angeben, klar, ich muss noch die Anzahlen irgendwo speichern, also die jeweiligen Prozentzahlen müsste ich noch irgendwo speichern und dann, wenn ich quasi […] mit meiner if-Abfrage Schleife, also wenn ich es richtig gemacht hab und meine Zählvariable auch schön fleißig zählt, könnt ich dann nachher einfach die […] die Mächtigkeit dieser Variable oder die Zahl dieser Variable, die sie zugeordnet ist, einfach benutzen und dann hätt ich die Lösung auf die Frage und dann könnt ich sagen in der und der, also in der ersten Variable ist es von mir aus nur 17 Mal und in der zweiten 35 Mal, dann hätt ich dann schon mal ne grobe Antwort darauf.

F: mmh, also das heißt, du hast das dann 365 Mal wiederholt, das heißt du hättest ein Jahr simuliert.

P: genau

F: mmh, und würdest du es dann so machen, also würdest du dann bei einem Jahr stehen bleiben?

P: Es käm dann drauf an, was ich dann am Ende aussagen möchte. Also, dann könnt ich sagen, also für ein Jahr kann es passieren, dass es so ist. Wenn ich es jetzt aber so generell betrachten würde, würd ich versuchen, diesen Vorgang öfters zu wiederholen. Logischerweise. Also einfach noch ne for-Schleife, also dann könnt man in der Funktion ne Funktion definieren, die einfach dafür sorgt, dass ich diesen Vorgang x Mal wiederhole. Je nachdem, für wen ich spreche. Ich glaub, wenn ich jetzt einfach nur für mich gesprochen hätte, würd ich das vielleicht 20, 30 Mal machen. Wenn ich jetzt etwas machen möchte, was wirklich Gewicht hat, wo ich wirklich auf die Aussagekraft äh mich berufen möchte, würd ich es schon so […] ich weiß nicht, ob R das schafft, 10 000 Mal anpeilen, 10 000 ist ein guter Wert.

F: Mmh

P: Man darf da das Wurzel n Gesetz nicht vergessen, das ist sehr scheiße, aber es ist halt so [lacht] Also, ich mein, je nachdem, was dahinter steckt, also wenn ich einfach nur für mich wissen möchte, ist das irgendwie zufällig, dass das so passiert ist, ich glaub, für mich würden ein, zweimal reichen, also, es steckt nichts dahinter, es gibt keinen Nach- oder Vorteil, wenn ich mich irren würde, sag ich mal. Wenn das aber zum Beispiel irgendwas ist, wo wirklich was von abhängt, zum Beispiel, ob Medikamente als gut oder schlecht geprüft werden, dann würd ich wiederholen und dann wär vielleicht sogar 10 000 noch zu wenig. Dann würd ich auf ne Million Schrägstrich 10 Millionen hochgehen. Ich weiß nicht, ob mein PC das schafft, aber wie gesagt […] Aber da hängt auch mehr von ab, ne. Ich glaube aber für mich persönlich würde sogar ein Durchgang reichen und dann würd ich mir erstmal angucken, wie sind denn so die Wahrscheinlichkeiten, also wenn es einen extrem großen Unterschied gibt, könnte man sich natürlich fragen, ist das denn eigentlich äh ja jetzt Zufall gewesen oder nicht, aber ich glaub, ich würds bei einem Mal lassen, wenn nichts davon abhängt. Also, mein Wissensdurst wär gestillt und dann wärs erstmal ok.

F: Wunderbar. Ja, möchtest du noch irgendwas anmerken?

P: Zu was?

F: Allgemein, zu R Aufgaben oder was auch immer, liegt dir noch irgendwas auf dem Herzen, was ich jetzt nicht gefragt hab?

P: Nö.

F: Wunderbar, dann wars das. Ganz herzlichen Dank.

Int5

F: Also, die erste Frage ist, was war deine erste Empfindung, als du von den R Aufgaben gehört hast und hat sich an deiner Einstellung bis heute etwas geändert?

A: Also die erste Empfindung war äußerst negativ. Ich hab überhaupt nichts mit Programmieren zu tun und ich hab auch nicht wirklich nen Sinn darin gesehen, dass ich im Mathe-Studium jetzt unbedingt ein Simulationsprogramm programmieren können muss, aber ich mein im Laufe der Zeit hat sich das deutlich geändert, weil ähm ich mein für die Aufgaben wars ganz nützlich, nicht nur die, die man damit gemacht hat, wenn man das R Programm dann halbwegs konnte, konnt man halt auch andere Aufgaben überprüfen, ob das auch richtig ist und äh auch außerhalb vom Mathe-Studium organisier ich halt so nen Tippspiel und so weiter und so fort und da muss ich halt immer Gruppeneinteilungen machen und da ist es vielleicht auch manchmal ganz günstig dieses R Programm zu nutzen anstatt äh das alles per Hand mit so Zettel zu machen.

F: Ok. Ok. Und wie häufig hast du die R Aufgaben bearbeitet?

A: Äh, immer.

F: Immer. Und warst du mal in der Fragestunde?

A: Einmal. [lacht]

F: Einmal, ok. Ähm, warum bist du dann nicht mehr gekommen? Oder hast dus nicht gebraucht, oder?

A: Also, ich sag mal, ich hab halt, wir waren ja in der ganzen Gruppe zusammen, und da war halt der ein oder andere, der kannte sich gut damit aus, und im Prinzip wars auch nicht, ist es auch nicht so schwer gewesen, dadurch dass ja das meiste schon vorgegeben war und man irgendwann mal die Befehle drauf hatte, musste man die halt einfach nur einsetzen und im Zweifel hab ich halt den Kollegen gefragt.

F: Mh, ok.

A: Und dann war, war dann halt auch, an dem Tag hatt ich halt leider auch immer frei, das war dann auch ein Grund.

F: Ok, ne, umso besser, wenns ohne geht. Ähm, fandest du denn die Aufgaben im Vergleich zu den anderen Aufgaben auf dem Blatt eher schwer oder eher leicht?

A: Ich find, das waren immer die einfachsten Aufgaben. [lacht]

F: Immer die einfachsten, ok. Und was war so leicht?

A: Also, ich meine, es war ja im Prinzip äh also man musste halt einmal verstanden haben, was diese ganzen Befehle bedeuten, und dann musst man ja im Prinzip, weil es ja schon vorgegeben war, das war ja der Punkt, dann wars halt relativ einfach die letzten paar Zeilen einzufügen. Ich denk, wenn man das komplett hätte selbst machen müssen, dann wärs ein bißchen schwierig gewesen. Aber so wars eigentlich wirklich einfach.

F: Mh.

A: Wobei natürlich die letzten Aufgaben und auch die in der Klausur nochmal nen Ticken einfacher waren, weil man halt nur noch erklären musste, was passiert.

F: Ok. Ok. Ähm, was fandest du an den R Aufgaben gut? Und an welche erinnerst du dich noch?

A: Hm, [lacht], also hm, schwierige Frage. Also, gut fand ich auf jeden Fall wie gesagt, dass äh dass man, also nicht mal wirklich die Aufgaben selbst, sondern mehr, dass man dadurch auch das R Programm für andere Aufgaben nutzen konnte, um zu überprüfen, ob das halt passt, die Ergebnisse zu testen, oder auch dass mans auch im Freizeitbereich immer mal wieder machen kann. Oder was halt wirklich lustig ist, zum Beispiel bei Crabs, ähm diese Spiele zu simulieren, das macht schon Spaß.

F: Mh.

A: Weil selbst zu spielen, das dauert natürlich ewig. Und manche Sachen sind schon sehr interessant so sich anzugucken.

F: Mh.

A: Ich mein, soll ich jetzt jede Aufgabe aufzählen, die ich noch weiß?

F: Ja, wenn du... Gerne.

A: Also, Crabs war die erste. Ähm, was hatten wir dann noch? Dann hatten wir zweimal das mit den Betten, ähm dann die mit der Kugel mit den Koordinaten und den Städten. Ähm, was hatten wir noch? Die letzte war die Polya-Urne, das war die Klausuraufgabe [lacht].

F: [lacht] Das ist auch gut.

A: Ähm, was war davor? Der ähm wie heißt das, Vorzeichenwechsel, hatten wir auch mal gehabt, das war aber glaub ich die Wirtschaftspädagogik-Dings, das war gar nicht unsere [lacht]

F: Ne, kam auch mal bei euch..

A: Doch, es kam einmal vor, doch stimmt. Da mussten wir, da sollten wir doch vorher überlegen.

F: Mh, genau, genau.

A: Ist alles noch da [lacht].

F: Perfekt. Ist nicht so selbstverständlich, deswegen äh, freu ich mich.

A: Nö, also, ich mein, was wir die, da war noch irgendne Aufgabe, ne, die hatten wir ja schon, die mit den Koordinaten, die hab ich nicht, die hatt mir nicht so gefallen. Die war etwas, weiß nicht. Da war der Code ein bißchen komisch. Ähm, ja, fällt mir noch eine ein? Was haben wir noch gemacht? Jetzt wird's schwierig. [...]

F: Ja, ist auch gut, kein Problem.

A: Ok, ja. Wieviele warens denn?

F: Eins, zwei, drei, vier, fünf, sechs, ja, sechs. Das ist die Hälfte.
A: Ist doch ganz gut. [lacht]

F: [lacht] Gibt es denn irgendwas, was dir durch eine R Aufgabe klar geworden ist, was du vorher nicht verstanden hattest?

A: Also, auf jeden Fall das Crabs-Spiel […] Das ist dadurch, ich mein, man musste sich ja damit nochmal beschäftigen, und dadurch, ich wusste am Anfang überhaupt nicht, wie das Spiel überhaupt funktioniert, das war schon das erste Problem. Dann auch das mit dieser Verzerrung, das war ja auch bei den Betten, ich mein, die Aufgabe war ja doppelt, und man hatte ja gesehen, dass das beim ersten Mal nicht gepasst hat, wie hieß das, Größenverzerrung, oder?

F: Mh, genau.

A: Das ist alles schon abgelegt.

F: Ok. Und gibt's irgendwas, was man deiner Meinung nach an den R Aufgaben besonders gut lernen kann?

A: Hm, also das Programmieren auf jeden Fall nicht [lacht], weil äh, ich mein, man musste ja wirklich nur einsetzen. Was aber, denk ich mal, für den Fall sinnvoll war. Weil, ich mein, so das ganze zu lernen, ein ganzes Programm, wär vielleicht ein bißchen viel gewesen.

F: Mh.

A: Ähm, was hat man denn noch gut gelernt? […] Schwierig, also, ich halts wirklich für sehr sinnvoll, das Ganze, aber mir fällt jetzt grad nichts Konkretes ein, was man daraus gelernt hat. [lacht]

F: Ist kein Problem. Wenn dir noch was einfällt...

A: sag ichs, genau.

F: Ähm, was hat dir denn bei den R Aufgaben nicht gefallen?

A: […] Hm, […]

F: Oder gabs nichts?

A: Doch, es gab schon hin und wieder, also, ich sag mal, vor allem in der Anfangsphase, da wars halt schon, also als ich die Befehle noch nicht so richtig drauf hatte, war das schon manchmal frustrierend, weil es dann an einer Klammer gehängen hat, gehongen hat, oder an irgendeinem Komma, was man nicht gesetzt hat, oder vielleicht einem kleinen Wort, was man vergessen hat, dann war das ganze Ding falsch, und dann hat man, also bei den ersten Aufgaben schon einige Stunden dran gesessen. Und das hat dann schon manchmal ein bißchen Verzweiflung gebracht oder Wut auf das System [lacht].

F: Ok. [lacht] Ok.

A: [lacht] Aber ich mein, das hatte sich mit Zeit gelegt.

F: Ok. Ähm, würdest du denn aus Studentensicht empfehlen, weiterhin R Aufgaben zu verwenden und sollte man irgendetwas ändern, falls mans weiter benutzt?

A: Also, ich bin der Meinung, dass man das auf jeden Fall weiter machen soll. Es ist ne gute Hilfe auch für, um Aufgaben sich klarer zu machen oder zu testen, obs Ergebnis stimmt. Weil, zumindest bei den Aufgaben, bei denen man Zahlen benutzt, kann man ja auf jeden Fall herausfinden, wenn das nämlich in etwas ähnlich ist, dann stimmt das auch, und wenns nicht ähnlich ist... Ach, jetzt fällt mir auch noch ne Aufgabe ein.

F: Ja.

A: Einen Schätzer haben wir doch auch einmal, zweimal gemacht. Einmal für den ach, was haben wir denn, was haben wir denn geschätzt, […] [lacht]

F: Aber es stimmt...

A: Wir haben auf jeden Fall noch nen Schätzer gemacht. Hatte damit auch irgendwas zu tun, mit den Wahrscheinlichkeiten, mit dem p Dach, wie heißt das, Konfidenzintervall oder so war das. [lacht] Ja, ich weiß nicht, irgendwie legt man das nach der Klausur erstmal so ein bißchen ab.

F: Ja, ist doch klar.

A: Ist halt die Frage, ob mans nochmal alles so braucht.

F: Mh.

A: Also, wie war die Frage eben?

F: Ob man irgendwas ändern sollte, wenn mans nochmal benutzt. Also, und ob du empfiehlst, es weiter zu benutzen.

A: Also, ich empfehls auf jeden Fall weiter zu machen und ich find auch, also so wie mans gemacht hat mit dem Code vorschreiben halt ich auch für sinnvoll. Ich würd vielleicht das ein oder andere Mal öfter die, die letzte Variante wählen, dass mans nur erklären muss, weil das ist nämlich auch interessant. Also, ich fand das deutlich cooler.

F: Ok.

A: Weil man halt dabei weniger den Fehler macht, dass man irgendwelche komischen eckigen Klammern vergisst, sondern einfach, da kann man nämlich auch testen, ob mans verstanden hat. Weil, wenn mans erklären kann, dann sollte man normalerweise auch die ganzen Befehle verstanden haben.

F: Mh, ok.

A: Und ich sag mal, für die Klausur ist es natürlich auch ganz günstig, weil die Aufgabe war ja wirklich eine von den leichteren. Und, ich mein, das tut den Studenten ja auch gut, denk ich mal.

F: Mh, ja. Ich hoffs [lacht]. Wurde denn in deinem Tutorium manchmal auf R Aufgaben Bezug genommen, als eine andere Aufgabe erklärt wurde?

A: […] Hm, kann ich meinen Telefon-Joker nutzen? [lacht] Also, ich glaube nicht, dass wir das gemacht haben. Höchstens auf, bei Crabs sind wir, glaub ich, weil die war ja, das war ja so ne Doppelaufgabe, einmal wo wir rechnen sollten und einmal mit der R Aufgabe und da wurde glaub ich, so weit ich mich erinnern kann, das wurde noch kurz erklärt oder erwähnt. Ansonsten haben wir halt eigentlich bei R immer nur drüber gesprochen, wenn halt irgendwelche Probleme waren.

F: Mh.

A: Und das war aber relativ selten, weil die, entweder habens die Leute nicht gemacht oder sie konntens [lacht].

F: Ok. Aha, ok. Ähm, dann jetzt nochmal was ein bißchen Allgemeineres: Woran denkst du als erstes, wenn du an Stochastik denkst?

A: [lacht] Das erste Wort wär jetzt Wahrscheinlichkeit, was mir da einfällt.

F: Aha.

A: Sonst, Münzwürfe [lacht]

F: Mh. Und ähm gut bei Münzwürfen jetzt vielleicht nicht, aber bei Wahrscheinlichkeiten: Kannst du erklären, was das ist?

A: [lacht] Das ist schwierig, das Wort zu beschreiben. Also, ich denk jetzt, dass ich weiß, was es ist, aber ich glaub, ich kann es nicht beschreiben. […]

F: Ok.

A: Wie macht man das denn am besten? [lacht]

F: Also, muss nicht irgendwie mathematisch sein, kannst auch...

A: Ja, ne, aber […] Das ist halt, ich versuch grad das Wort Wahrscheinlichkeit anders auszudrücken. […]

F: Ok.

A: Also, es geht ja äh im Prinzip um […] die Wahrscheinlichkeit von Ereignissen [lacht]. Ja, dass halt irgendwas eintritt oder nicht eintritt, aber […] mir fällt kein besseres Wort für Wahrscheinlichkeiten ein.

F: Ok. Macht nichts. Fällt dir noch ein anderer Begriff ein?

A: Für Wahrscheinlichkeit?

F: Ne, generell, wenn du jetzt an Stochastik denkst?

A: Hm, Zufallsexperimente, Zufallsvariablen.

F: Mh, und kannst du da...

A: Aber da wird's auch schon wieder, deshalb wollt ich's nicht sagen [lacht]

F: [lacht] Aber trotzdem: Kannst du was dazu sagen? Also zu Zufallsexperiment oder Zufallsvariable?

A: [lacht] Das sind tolle Wörter. Also, mir fällt auch keine wirklich gute Definition für ne Zufallsvariable ein.

F: Ok. Und irgendwie so ne Umschreibung?

A: Also, man zählt immer, halt ne Variable, die halt aus nem bestimmten, aus nem bestimmten Bereich irgendwelche Werte annimmt mit ner gewissen Wahrscheinlichkeit oder halt auch nicht.

F: Mh. Ok. Ist doch gut. Ähm, genau, jetzt noch zwei Begriffe, die in den Fragebögen vorkommen, und zwar war einmal stochastische Intuition ein Begriff, der dort vorkam. Und die Frage wär, was du darunter verstehst.

A: Also, ich versteh darunter, ähm, also wenn ich jetzt irgendein Experiment hab, ob ich mir dann intuitiv vorstellen kann, wie groß die Wahrscheinlichkeiten für irgendwelche Ausgänge sind, ob das wahrscheinlich ist, ob das so funktioniert.

F: Mh.

A: Also beispielsweise wenn ich nen fairen Münzwurf hab, ob ich mir dann vorstellen kann, dass ich bei 100 Würfen in etwa 50 Mal Kopf werfe oder so was.

F: Ok. Und gibt es nen Unterschied zu Verständnis von Stochastik?

A: […] Hm, vermutlich gibt's einen, sonst würd man die Frage nicht stellen, [lacht].

F: Ne, ist eher so gemeint, ob du einen siehst. Vielleicht gibt's auch keinen, ich weiß es nicht so genau.

A: Ich denk mal die, naja, normalerweise es kommt ja auch öfters vor, dass man sich was vorstellt, was rauskommen könnte und tatsächlich kommt aber was ganz anderes raus. Und ich denke mal, wenn man das Verständnis von Stochastik hat, dann kommt man eher zum richtigen Ergebnis und mit der Intuition könnt man auch mal falsch liegen.

F: Ok.

A: Das ist so mein Verständnis davon.

F: Aha, gut. Jetzt hab ich noch eine, sozusagen, Aufgabe...

A: Oh Gott [lacht]

F: Und zwar, also nichts Schlimmes [lacht]. Aber in den Fragebögen kam diese

Krankenhausaufgabe vor, da gabs ein großes und ein kleines Krankenhaus und ähm im großen wurden jeden Tag ungefähr 45 Kinder geboren, im kleinen ungefähr 15. Und jetzt sagt man, ok, der Mädchenanteil ist ungefähr bei 50% oder die Wahrscheinlichkeit für ein Mädchen. Ja, und jetzt wurde für den Zeitraum von einem Jahr jeweils die Tage gezählt, an denen mindestens 60% der neugeborenen Kinder Mädchen waren und in welchem Krankenhaus erwartet man mehr solcher Tage? […] Ich will jetzt keine Antwort auf die Frage, sondern: Angenommen, du müsstest das simulieren, dieses Problem: Ohne Code, ja, ohne irgendwelche Befehle, aber wir würdest du vorgehen?

A: Hm.

F: Kannst auch gerne nochmal in Ruhe durchlesen.

A: Die Frage kam ziemlich oft in den Fragebögen. […] Hm, das wird bestimmt wieder so was Hässliches mit Matrizen sein. [lacht]

F: Also, du musst jetzt nicht irgendwie sagen, was das für, also wie du jetzt den Code aufbauen würdest, sondern einfach nur, welche Schritte du machen würdest.

A: Also ich vermute mal, dass es ähnlich wie die andere mit den Bettlegezeiten vermutlich sein könnte. Dass man es vielleicht so machen könnte, wie wirs auch da gemacht haben. Aber da müsst ich, glaub ich, ein bißchen überlegen, wie das funktioniert.

F: Mh. Ok, aber wenn du versuchst, so Schritte zu sagen, wie man da vorgehen könnte? Also jetzt wirklich sehr allgemein, nicht irgendwie konkrete...

A: Also, ich muss ja im Prinzip, muss ich ja die Betten simulieren, also nicht Betten, also vermutlich dann bei dem einen 15 und bei dem andern 45.

F: Mh, oder Kinder in dem Fall.

A: Ja, ok. Und äh die muss ich halt, stimmt, sind ja jetzt Kinder [lacht]. Ja, und die muss ich ja auffüllen mit nem gewissen Prozentsatz.

F: Ok.

A: Ja, ok, der Prozentsatz ist ein halb.

F: Mh, ok.

A: […]

F: Und dann machst du das einmal 45 Mal und einmal 15 Mal. Und dann?

A: Und dann, was ist das, ein Jahr. Und das ist ja täglich, ne. Und dann müsst ich das ja theoretisch 365 wiederholen.

F: Mh.

A: Für jeden Tag.

F: Mh, genau.

A: Und dann kann ich die Tage zählen, wo 60% Mädchen sind.

F: Jawohl, wunderbar. Und ähm, das wärs?

A: [lacht] Vermutlich nicht.

F: Also, [lacht], wunderbar. Dann hast du ja für ein Jahr zwei Werte, also wie oft kam im größeren es vor, dass über 60% Mädchen waren, wie oft im kleineren. Und die Frage ist ja irgendwie, wie viel, also in welchem erwartet man mehr.

A: Stimmt, da war noch was. [lacht]

F: Vielleicht ist es auch zu simpel, als dass man es jetzt sagt.

A: Ne, also, vielleicht [lacht]. Also, ich hab jetzt, wenn ich die Tage gezählt hab einfach noch auf das Jahr beziehen, also auf die kompletten 365 Tage.

F: Mh.

A: […]

F: Ok. Und dann guckst du einfach, wo mehr sind und dann ist gut?

A: Ja.

F: Ok, wunderbar. Passt.

A: Krieg ich noch ne Antwort auf die Frage?

F: Ja, also, auf was ich hinaus wollte, war, dass man vielleicht das noch öfter wiederholen könnte.

A: Ok.

F: Aber ist nicht so wichtig. Ähm, aber die Antwort ist...

A: Also für mich, ich dachte ja eigentlich immer, dass es gleich viel wären, aber.

F: Aha.

A: Was ist die richtige Antwort?

F: Die richtige Antwort ist, dass es im kleineren wahrscheinlicher ist, weil ähm, wenn du in dem kleineren Krankenhaus, also im Grunde schätzt du ja diese 50%, ja. Und wenn du weniger Beobachtungen hast, weil du ja jeweils nur 15 Kinder am Tag hast, ist der Schätzer ungenauer.

A: Ok.

F: Oder halt, wenn mans mathematischer ausdrücken will, dann ist der Standardfehler größer, weil man da ja durch n teilt oder durch Wurzel n und dann ist das n ja einmal 15 und einmal 45, das heißt in dem kleinen ist das n größer, äh, das n kleiner, damit der Standardfehler größer und das heißt, da erwartet man mehr Tage, also der Schätzer ist ungenauer in dem kleinen Krankenhaus.

A: Ok.

F: Aber da vertut man sich gerne, ja. Deswegen ist die Frage auch drin, weil die...

A: [lacht]

F: wurde auch schon viel untersucht, weil sich da sehr viele Leute vertun. Aber das ist so ein Intuitionsding, wo man normalerweise daneben liegt.

A: [lacht] Das ist cool.

F: Ok, aber vielen Dank. Ja, möchtest du sonst noch irgendwas anmerken zu den R Aufgaben oder hast du irgendwelche Vorschläge?

A: Ne, ich find eigentlich, dass das gut gemacht ist so. Ich mein, jetzt hab ich zu Hause ein Simulationsprogramm, wo ich ein bisschen rumspielen kann, ist auch, wenn ich halt mir irgendwelche Spiele halt ausdenk, das kann man halt auch gut testen, also das ist schon lustig.

F: Mh.

A: Also, ich sag mal, am Anfang, wie gesagt, ich denke mal die Leute, die halt nichts mit Informatik zu tun haben, die werden am Anfang eher unglücklich sein, aber ich denk, wenn die das halt wirklich machen und sich drauf einlassen, dann ist das schon ne gute Sache. Gibt’s nichts Negatives zu sagen.

F: Ok, perfekt. Vielen Dank.

F: Ok, also die erste Frage ist: Was war deine erste Empfindung, als du von den R Aufgaben gehört hast, und hat sich an deiner Einstellung bis heute etwas geändert?

A: Ja, ich habe erstmal gedacht: Hilfe, was soll das? Hab ich noch nie was, ich hab noch nie programmiert oder so.

F: Ok.

A: Und da hab ich gedacht, naja, toll, ja, ich hab immer gern schön viele Punkte in den Übungsaufgaben, hier hats sichs ja auch mal gelohnt [lacht], aber mittlerweile, das war ja so aufgebaut, dass man auf jeden Fall mitkommen konnte, find ich.

F: Ok.

A: Ich fands dann doch in Ordnung im Nachhinein.

F: Ok. Und wie häufig hast du die R Aufgaben bearbeitet?

A: Relativ oft. Ich glaub die ersten zwei oder drei hab ich nicht gemacht. Aber ich mirs immer angeguckt und dann hab ich sie immer bearbeitet, so viel ich halt hingekriegt hab.

F: Mh. Und warst du mal in der Fragestunde?

A: Ne, da hatt ich keine Zeit.

F: Ok. Also aus Zeitgründen.

A: Ja, aber ich hab dir ja immer mal emails geschrieben.

F: Genau, genau. Ok, ähm, ja, fandest du denn die R Aufgaben im Vergleich zu den anderen Stochastik-Aufgaben eher schwer oder eher leicht?

A: Hm. Das kann man gar nicht so generell sagen, find ich. Also manche von den Stochastik-Aufgaben waren recht einfach und manche ziemlich schwierig, find ich.

F: Mh.

A: Die R-Aufgaben, die waren eigentlich so ähnlich vom Schwierigkeitsgrad her. Manchmal halt schon zeitintensiver. Weil man halt, ich weiß nicht, ob das jetzt für jeden so ist, ich habs halt noch nie gemacht Programmieren, da hats schon ein bißchen länger gedauert. Aber so vom Schwierigkeitsgrad find ichs ok.

F: Mh. Und ähm gabs irgendwas, was du an den R Aufgaben besonders gut fandst?

A: Dass es die als Lückentext quasi schon gab und dass man dadurch das sich selber beibringen konnte ein bißchen, weil ich hab das vorher noch nie gemacht, wie gesagt, und trotzdem hab ichs einigermaßen hingekriegt irgendwann, das fand ich gut.

F: Mh. Und an welche erinnerst du dich noch?

A: Ähm, besonders gut fand ich die erste mit der Monte Carlo Methode, das mit dem, was war das, mit dem Kreis, genau, das fand ich lustig, vor allem weil das auch so, ich bin ja gar keine Mathe-Studentin und das war dann eher so was, was ich mir vorstellen kann, was ich noch gebrauchen kann, irgendwie. Ja.

F: Mh.

A: Fand ich lustig.

F: Ok. Gibt's denn irgendwas, was dir durch eine R Aufgabe klar geworden ist, was du vorher nicht verstanden hattest? Oder wo du vorher unsicher warst?

A: Oh, fällt mir jetzt nichts ein spontan,ne.

F: Ok. Und gibt's irgendwas, was man deiner Meinung nach besonders gut lernen kann an den R Aufgaben? Also auch inhaltlich.

A: Ja, ich finds eigentlich ganz schön, weil man da wirklich mal stochastische Ergebnisse sieht, weil sonst rechnen wir es aus, aber da, weil da ja schon zufällig irgendwelche Zahlen gewürfelt wurden, aber trotzdem kommt das raus, was man sich denkt, das ist schön.

F: Mh. Und hat dir das irgendwie was gebracht so für die anderen Aufgaben oder generell für dein Verständnis?

A: Ja, ich denk schon, ein bißchen besser anschaulich sich das vorzustellen. Bei den normalen Aufgaben waren dann halt oft auch so theoretische Herleitungsaufgaben, das liegt mir dann nicht so, aber die Anwendungs- oder R Aufgaben waren dann schon schöner.

F: Ok. Ähm würdest du denn, achso ne. Was hat dir bei den R Aufgaben nicht gefallen? Die Frage vergess ich gern [lacht].

A: Hm, […], da fällt mir jetzt spontan auch nicht wirklich was ein. Also manchmal hab ich auch da gesessen und gehongen und keine Ahnung gehabt, aber im Großen und Ganzen ist jetzt nichts Negatives hängen geblieben bis jetzt.

F: Mh, ok. Und würdest du empfehlen auch weiterhin R Aufgaben zu verwenden?

A: Ja, ich fands gut. Ich weiß jetzt nicht, wie groß so der Anteil von Studenten war, die es genutzt haben, aber ja.

F: Schon recht viele, ja.

A: Ok.

F: Und würdest du irgendwas verändern an dem Aufbau oder...?

A: Ne, das fand ich so ok. Lückentext, ich konnt nachfragen, ja.

F: Ok. Ähm, wurde denn manchmal in deinem Tutorium auf R Aufgaben Bezug genommen, wenn eine andere Aufgabe erklärt wurde?

A: Hm, eigentlich nicht. Vielleicht mal irgend, in einem halben Satz gesagt, das habt ihr ja in der R Aufgabe gemacht, aber eigentlich sonst nicht.

F: Und die R Aufgabe wurde auch nicht besprochen weiter?

A: Ne.

F: Ähm, gut, jetzt nochmal was Inhaltliches zu Stochastik.

A: Oh je.

F: Woran denkst du als erstes, wenn du an Stochastik denkst?

A: An Würfel.

F: An Würfel, ok. Und...

A: und Münzen und erstmal die ganzen Gegenstände, was war das, Roulettespiel, weil wir das in der Schule mit Stochastik immer ja, Lotto.

F: Mh. Und fällt dir irgendein mathematischer Begriff ein?

A: Äh, stochastische Abhängigkeit, Unabhängigkeit, was gibt's noch schönes? Das ist einfach schon zu lang her [lacht].

F: Ne, kein Problem. Kannst du denn erklären, was abhängig und unabhängig heißt?

A: Ähm, ja, unabhängig, wenn die Ergebnisse von zwei Versuchen voneinander unabhängig sind, logischerweise, wenn die hintereinander ausgeführt werden und dann der erste den zweiten beeinflusst, dann ist es abhängig, ja.

F: Mh, ok. Noch ein anderer Begriff? Wenn nicht, ist auch nich schlimm.

A: Ja, was hatten wir denn noch so Schönes? Die ganzen verschiedenen Verteilungen. Sehr wichtig, Normalverteilung, Exponentialverteilung, vor allem die Dichteverteilung. Die diskreten Verteilungen, was hatten wir denn da Schönes, geometrisch, hypergeometrisch, uniform verteilt, ja. Vor allem die kontinuierlichen, die sind, glaub ich, auch für mich dann nochmal wichtig, die kann ich noch gebrauchen.

F: Mh, ok. Ähm, ok. Dann hab ich noch so einen Begriff, der in dem Fragebogen vorkam, und zwar wurde da mal von stochastischer Intuition gesprochen.

A: Mh.

F: Und mich würd interessieren, was du darunter verstehst.

A: Ähm, stochastische Intuition. Ob man, wenn man irgendne Aufgabe sieht, meistens sind so stochastische Aufgaben, denk ich, so, dass man nicht das tippen würde, was am Ende dann raus kommt. Und mit Intuition ist dann vielleicht, wenn mans doch richtig tippt.

F: Mh. Und gibt es einen Unterschied zu Verständnis von Stochastik?

A: […] Ich denk schon, weil Intuition ist eher so […], auch wenn man sichs dann richtig vorstellen kann, heißt ja nicht, dass man es dann mathematisch verstanden hat. Weil Verständnis würd ich eher so mathematisch, formelmäßig verstehen, so. […]

F: Also formelmäßig. Wie genau meinst du das?

A: Äh, die ganzen Sätze und Herleitungen, das versteh ich unter, wenn mans tatsächlich verstanden hat.

F: Mh. Ok. Ähm, dann hab ich noch ne kleine Aufgabe.

A: Nein [lacht].

F: [lacht] Nichts Schlimmes. Also in den, auch in dem einen Fragebogen kam die vor und zwar geht's um diese Krankenhausaufgabe.

A: Ja.

F: Und ähm, also kannst auch gern nochmal in Ruhe durchlesen, wie sie genau war.

A: Ja. […] Im kleinen.

F: Ja. Und angenommen, du müsstest jetzt ähm diese Aufgabe simulieren, also jetzt ohne den Code, ja, also nur so von dem, wie würdest du rangehen?

A: Ähm, jetzt muss ich aufpassen, weil ich mach grad ein Programmier-Praktikum bei den anderen...

F: Ah.

A: in einer anderen Programmiersprache, jetzt muss ich erstmal überlegen, aber das geht ja ähnlich. Da bin ich jetzt mehr drin. Ähm […]

F: Also, es reichen die Schritte.

A: Ja. Wir können ja zufällig würfeln lassen, was weiß ich, Mädchen sind 1 und Jungen sind 0. Einfach würfeln lassen und dann 15 Mal und 45 Mal und das dann keine Ahnung wie oft wiederholen. Und ihn dann zählen lassen, wo […] mehr Mädchen, also anteilsmäßig mehr Mädchen geboren wurden, ja.

F: Mh.

A: Ja, jetzt fällt mir nicht mehr ein, wie mans weiter machen könnte.

F: Das ist gut. Und wenn du sagst, irgendwie so und so oft wiederholen, wie oft würdest du es wiederholen?

A: Ich glaub, wir hatten immer so 10000 Mal. [lacht]

F: Ok, mh.

A: Ist zwar jetzt da eher nicht so realistisch, weil ich das ne lange Zeit ist in Tagen, aber wir habens in den Simulationen immer so gemacht, damit man stochastisch sinnvolle Ergebnisse hat.

F: Mh. Und das heißt ähm du würdest jetzt nicht irgendwie ein Jahr simulieren, weil es heißt für den Zeitraum von einem Jahr...

A: Achso, das hab ich nicht genau genug gelesen.

F: Ok.

A: […] Ja, ein Jahr ist jetzt […] würd ich sagen noch ein bißchen kurz.

F: Ok, mh.

A: Ja, um das zu schätzen.

F: Mh. Perfekt, wunderbar. Perfekte Lösung.

A: [lacht] Jetzt weiß ich auch mal, dass es richtig war, was ich da immer angekreuzt hab.

F: Ja, es ist tatsächlich richtig. Ähm, ja, möchtest du sonst noch irgendwas anmerken? […] Gibt es noch irgendwas, was du loswerden möchtest?

A: Insgesamt, ja wie gesagt, ich mach seit letzter Woche Fortran-Programmierpraktikum und ich merk total, dass ich das mit den R Aufgaben gemacht hab, weil ich konnt von Anfang an das viel schneller mir merken so, also die Befehle sind natürlich anders, aber so wie man anfängt erstmal, das konnt ich dann schon einigermaßen, wie man an so ne Aufgabe ran geht. Also geholfen hats mir auf jeden Fall, ja.

F: Das ist schön, ok. Sonst noch irgendwas, was du sagen möchtest?

A: Nö, eigentlich nicht.

F: Ok, perfekt. Dann vielen Dank.

Int7

A : Ich hab mich halt gar nicht auf die R-Sachen fokussiert, gell, ich hab das so ein bisschen ausgelassen, in der Anfangszeit hab ich ein bisschen probiert da so rumzutüfteln,

F : ja

A : und hab gemerkt, das ist schon gut, das macht schon Spaß auch, aber dann dacht ich mir uh, das ist ja schon, dann muss ich ja alles lernen, die ganzen Schleifen, die ganzen Befehle sozusagen, die ganze Logik dahinter, und dann dacht ich mir, hm, wär schon gut, wenn man dann da irgendwie so nen Vorkurs dafür gehabt hätte, so wie in der Numerik jetzt mit Matlab, ja. Das man das irgendwie so ein bisschen übt, ja, weil ich fand das so ein bisschen ähm, ja, das war dann halt einfach so da.

F : Ja, das haben schon mehrere Leute gesagt. Ich denke auch, dass man darauf achten sollte, wenn man das nochmal macht.

- Eigentlicher Beginn Interview -

F : Also, die erste Frage ist: Was war deine erste Empfindung, als du von den R Aufgaben gehört hast, und hat sich an deiner Einstellung etwas geändert?

A : Also die erste Empfindung, ich war überrascht, aber nicht positiv überrascht, ich hab mir gedacht, gut, jetzt müssen wir nochmal programmieren im Mathestudium, das nimmt ja eh irgendwie zu und das auch noch ohne Vorkurs, ja, also das war dann einfach so da. Ähm, im Endeffekt denk ich mir ähm, das ist wichtig, ja, weil gerade auch in über die Stochastik in der Anwendung später, wenn man ins Berufsleben geht, muss man das ja irgendwie implementieren, und es ist schon gut, das zu können, die Frage ist nur, ob das jetzt auch wirklich jeder dadurch jetzt kann.

F : Mhm.

A : Ja, weil viele blenden das auch aus und ähm sagen dann, ok, ich mach die R Aufgaben jetzt nicht, und ich konzentrier mich auf die reine Stochastik und ja, lernen es dann sozusagen auch nicht.

F : Mhm.

A : Weil die das einfach so umgehen.

F : Ok. Mhm. Und wie häufig hast du die R Aufgaben selbst bearbeitet?

A : Mhm, die ersten vier Übungsblätter, hab ich mich dran probiert.

F : Mhm.

A : Die ersten gingen auch noch recht gut. Hat auch Spaß gemacht, weil sie vielleicht auch einfach waren. Ich glaub, das war so sample Funktion und so was.

F : Mhm.

A : Und dann, als das dann mit if und den ganzen Sachen losging, und for-Schleife und irgendwie Abfrage, dann war mir das so ein bißchen zu viel Arbeit.

F : Mhm.

A : Ja, dann dacht ich mir, gut, ich lass das jetzt lieber.

F : Ok. Und warst du mal in der Fragestunde?

A : Ne, da bin ich leider nicht hingegangen. Weil ich bin so ähm ein Autodidakt.

F : Ok .

A : Ja, ich mach in der Regel alles alleine. Aber wär jetzt im Nachhinein, glaub ich, schlauer gewesen, und ähm, ich glaub, die Anwesenheit in dieser Fragestunde hätte, glaube ich, die Einstellung zu dem ganzen R, also zu dem ganzen R äh Tool irgendwie geändert.

F : Ok .

A : Ja, weil man ja auch Feedback bekommen hätte, und das hab ich dann sozusagen nicht genutzt, ja, aber was sehr interessant ist, dass ich auch wiederum gemerkt habe, dass wenn man das so alleine machen will, das so in den Anfängen recht gut ist, also da hat es recht gut funktioniert, aber wenn man dann Sachen komplizierterer Natur dann irgendwie darstellen möchte, dann sind die Skripte, die man hat, allein nicht so gut, dann braucht man gute Informatik-Vorkenntnisse, würd ich sagen.

F : Mhm, ok. Und fandest du denn die R Aufgaben im Vergleich zu den anderen Aufgaben eher schwer oder eher leicht?

A : Ähm, ich würde sagen, qualitativ waren sie einfacher als die meisten Stochastik-Aufgaben. Das Schwierige lag darin, dann wiederum, dass ähm ja, dieses Programmieren-Denken, also dieses algorithmische Denken, das macht dann so ne Herausforderung.

F : Mhm.

A : Also, man muss dann sozusagen, sich in so in so nen Rechner eindenken, und auch ähm wenn man es, also auch wenn das Programmieren gar nicht geht so, sozusagen einfach mal überlegen, worum geht's, ja, ich hab jetzt irgendwie 718 Gäste, da fallen einige mit so ner Wahrscheinlichkeit irgendwie aus, ja, also das heißt, ich muss die irgendwie da drinne haben im System, dann muss ich jeden einzelnen eventuell bewerten, und das fiel mir schwer, so ein Maschinen-Denken zu bekommen.

F : Ja, ok. Und gab es etwas, was du an den R Aufgaben gut fandst?

A : Ja, ich fand diese äh diese sample Funktion, die fand ich sehr gut. Dass man direkt Zufalls ähm einfache Zufallssimulationen direkt einfach ausführen kann und das fand ich gut, das hat so ein spielerisches Moment und das macht auch Spaß, dann denkt man, gut, ich hab das jetzt simuliert und jetzt kann ich vielleicht noch das simulieren, und das macht dann auch Lust, eventuell weiter zu machen.

F : Mhm. Ähm, erinnerst du dich denn noch an einzelne R Aufgaben?

A : Oh, das ist ne gute Frage, ist ja lang her, hm, also ich weiß, dass es irgendwas mit so ner Bettenbelegung im Krankenhaus gab und dazu musste man ein Histogramm ausspucken, ja und dann kann ich auch noch an die sample-Funktion erinnern, dann an eine zufällige Anordnung von Zahlen, genau.

F : Mhm, wunderbar. Gibt es den irgendwas, was dir durch eine R Aufgabe klar geworden ist, was du vorher nicht verstanden hattest?

A : Hm, […] eher nicht.

F : Mhm.

A : Also, alles, was ich vorher kannte, ich sag mal so, es gibt jetzt nichts, was ich durch R jetzt mathematisch neu gelernt hätte. Ja, also ich wusste z.B., dass es Zufallsgeneratoren gibt, dass man die am Rechner implementiert, und ähm, dass sie dir auch irgendwelche Permutationen von Zahlen geben, wenn du einfach Zahlen hast, also die umordnen, von daher hab ich mathematisch gesehen nichts Neues gelernt. Was ich, wenn überhaupt, gelernt hab, ist so ein bisschen diese ersten formalen Schritte, diese Prozeduren, auch Schleifen, so was, so weit ich mich da auch reingearbeitet hab.

F : Aha, ok. Was hat dir bei den R Aufgaben nicht gefallen?

A : Gute Frage. […] Ich fand manche, also dass, ich fand, bei manchen Aufgaben, die waren mir einfach ähm von der Anforderung her zu weit voraus. Also, wie halt so ne Belegung mit Betten halt auch, ja, da musste man ziemlich viel eingeben, daran kann ich mich noch irgendwie erinnern, bei der Hotel äh Krankenhaus-Bettbelegungsaufgabe, und ziemlich viel selbst erarbeiten. Gut, der Vorteil war halt, dass da die Lösungen ja schon fast im Netz waren, aber da hab ich mich auch wiederum gefragt, hm, bringt mir das was. Wär es nicht besser, sie so ein bisschen einfacher zu machen, und dann wirklich selbst dran zu tüfteln, also ich fand sie manchmal zu schwer.

F : Ok, und dann war das Template im Internet auch nicht so richtig hilfreich?

A : Hm, ja, also ich konnte damit schon so die Lösung so ja wiederherstellen, oder selbst aufbauen, aber ich habs dadurch, glaub ich, nicht selbst begriffen, ja.

F : Ok.

A : Da bin ich dann immer so ein bisschen umgangen, ich hab dann gedacht, oh, das steht ja dann da, das muss ich jetzt an den richtigen Stellen einsetzen, an den richtigen Knotenpunkten muss man sich die richtigen Gedanken machen, aber im Gesamtüberblick würd ich sagen, eher kontraproduktiv.

F : Ok. Würdest du denn aus Studentensicht empfehlen, weiterhin R Aufgaben zu verwenden?

A : Mhm, würde ich definitiv machen. Ähm, zumal ich auch jetzt merke, also, ich bin in der Arbeitswelt und auch in so einem Bereich, wo man Simulationen braucht, also weil ich da nen Nebenjob grade mach, ja, und da merk ich, dass dieses Programmieren unglaublich wichtig ist, ja. Und letzten Endes muss man als Mathematiker, wenn man zumindest nicht Lehrer wird, sondern in die Arbeitswelt geht, irgendwo auch programmieren.

F : Mhm.

A : Im mathematischen Sinn. Das ist schon super erforderlich.

F : Mhm.

A : Aber ich würde, ja, also die Fragestunde gabs, das darf ich natürlich auch nicht vergessen und die hab ich nicht genutzt, vergess ich dann immer, gut, wär wahrscheinlich anders ausgefallen die Bewertung davon, wenn ich in die Fragestunde gegangen wär, aber ich weiß in der Numerik gibts diesen Vorkurs und der ist super, ja, also in den Semesterferien in der Woche davor, da hat man dann ein erstes Kennenlernen mit dem Tool, ja, aber ich würd das auf jeden Fall empfehlen, ja.

F : Mhm. Aber dann mit so einem Vorkurs.

A : Ja, ich würd den vorne dran setzen und auch verpflichtend, ähm, eventuell auch als Credit Point Zahl, ja, wie in der Numerik, da gibt's, glaub ich, 3 Credit Points für den Kurs unbenotet, also man kriegt Punkte drauf, aber es gibt nur bestehen und nicht bestehen.

F : Aha.

A : Genau, und dann ja, würd ich das sozusagen outsourcen, in den Vorkurs.

F : Ok.

[kurzes Gespräch über den Nebenjob]

F : Wurde denn in deinem Tutorium manchmal auf R Aufgaben Bezug genommen, als eine andere Aufgabe erklärt wurde? Erinnerst du dich da an was?

A : Ähm, weitestgehend nicht, also es wurden die stochastischen Aufgaben einfach durchgegangen, auch ziemlich gut, aber R, ja, da war, glaub ich, mein Tutor auch nicht so gut drin.

F : Mhm.

A : Der war nicht gut in R und hat auch gesagt, die R Aufgaben korrigiert jemand anders, ja, also da wurd kaum drauf Bezug genommen im Tutorium.

F : Mhm, ok. Ähm, jetzt nochmal was Inhaltliches zu Stochastik: Woran denkst du als erstes, wenn du an Stochastik denkst?

A : Dann denk ich sofort an Simulationen, ja, also an Simulationen von Zufallszahlen, ja, und auch die Frage, was ist guter Zufall, was ist stabiler Zufall, ja.

F : Mhm.

A : Ja, da ist ja jede Maschine auch beschränkt. Ja, das geht ja schon fast auch in ne philosophische Frage dann auch rein.

F : Mhm, ok. Und denkst du auch an einen mathematischen Begriff?

A : Klar, natürlich auch an den Erwartungswert, die Varianz, ja, Erwartungswert ist definiert als Integral, ja, was mich auch immer noch [?] irritiert, ist, warum die Stochastik über die Maßtheorie definiert ist, das kann ich mir nach wie vor nicht erklären, ja, warum Wahrscheinlichkeiten als Maße betrachtet werden.

F : Mhm.

A : [?] Das ist für mich immer noch offen, also, ich war total überrascht, als gesagt wurde, die Stochastik ist in allem ein Spezialfall der Analysis, ja, also jetzt ganz theoretisch betrachtet. Und ich frag mich, was haben die zwei eigentlich miteinander zu tun, ja.

F : Mhm. Und wenn du jetzt jemanden, der nichts mit Mathe zu tun hat, diese Begriffe erklären müsstest, also Erwartungswert und Varianz, wie würdest du das machen?

A : Das ist ne gute Frage. Also beim Erwartungswert geht's ja schon los, ja, also das kann man dann natürlich nur plastisch erklären, anhand von Beispielen, am Würfelwurf oder so was, ja, und dann würd ich halt auch irgendwie sagen, also du nimmst jetzt nen Würfel und du würfelst den zweimal und wenn du Glück hast, kommt zweimal Sechs, ja, aber du kannst jetzt nicht davon ausgehen, dass immer Sechs kommt, weil wenn du eben ganz oft würfeln würdest, das ist dann das Gesetz der großen Zahlen, würdest du die Sechs so oft bekommen, wie im Wahrscheinlichkeitsmodell, ja, manifestiert ist, nämlich von der Anzahl ein Sechstel mal der Vorkommnisse. Und das würd ich dann probieren, so plastisch zu erklären, zumindest den Erwartungswert. Bei der Varianz, schwierig.

F : Mhm.

A : Ich würd sagen, der Erwartungswert minus irgendwas und dann die Summe darüber, ja, das ist schwierig.

F : Ok. Ähm, in den Fragebögen, ich weiß nicht, ob du die ausgefüllt hast, da kam mal so ein Begriff vor – stochastische Intuition.

A : Mhm.

F : Und mich würde interessieren, was du darunter verstehst.

A : Super, also, eigentlich, boah, das ist schon […] stochastische Intuition ist für mich ein Gefühl dafür, Wahrscheinlichkeiten zu verstehen und auch richtig einzuschätzen, ja ähm, also wenn man zum Beispiel irgendwie ne Aussage bekommt, die Ausfallrate von irgendwas ist mit 10% gesichert übers Jahr, dann denken auch gleich alle falsch, ja, ok, und denken an irgendeinem Zeitpunkt, an einem einzigen wird irgendetwas passieren, ja in einem kleinen Ausmaß, was dann 10% sind. Aber dass es eventuell auf alle Zeitpunkte in noch kleinerem Ausmaß irgendwie passiert, also sozusagen geglättet ist, ähm, das widerstrebt dann vielen, also ich hoffe, ich hab, ich hoff, ich hab das einigermaßen plastisch erklärt, aber da merk ich auch, wie schwierig das ist, stochastische Intuition richtig, ja, also die auch richtig für sich zu haben, muss ich mir noch erarbeiten.

F : Mhm. Und gibt es einen Unterschied zu Verständnis von Stochastik?

A : Ja, würd ich schon sagen, also Stochastik ist sozusagen ähm die Einschätzung von heuristischen Wahrscheinlichkeiten, von Wahrscheinlichkeiten, die man überhaupt messen kann. Und das als Wissenschaft zu begreifen, das ist schon schwierig, weil viele denken sich ja gut, Stochastik, das ist einfach, ja, die reduzieren das auf Wahrscheinlichkeit, ja, aber die wissen nicht, dass es darum geht, dass es da um viel mehr kreist, ja, und für die ist es banal ne Wahrscheinlichkeit, die glauben, für alles gibt's ne Wahrscheinlichkeit, ja, und die ist auch konsistent, die ändert sich nicht, was ja Quatsch ist, also zumindest nicht in jedem Modell ist es so, in festen Modellen schon, ja, und das ist schon ne Schwierigkeit auch, ja. Und was auch super wichtig ist, was ich auch immer merk, ist irgendwie, die Stochastik ist wirklich ne reine Kopfsache, ja, weil wenn ich so nen Würfel werfe, dann hab ich ne Modell-Denke, alles passiert, also jede Augenzahl erscheint mir ein Sechstel Wahrscheinlichkeit, aber dann lass ich ja total die physikalischen Gegebenheiten raus, ja, ich hab ja nen Tisch, auf den der Würfel fällt, und da gibt's gewisse Reibungskräfte und wenn ich jetzt in ne Kule werfe, ja, die ich zum Beispiel nicht sehe, dann manipuliert es ja das ganze Ergebnis und den Vorgang, ja.

F : Mhm.

A : Und ich frage mich dann auch immer, zumindest beim Würfelwurf, ja, da muss ja auch irgendwo noch ne physikalische Schnittstelle sein, ja, aber das wird dann wahrscheinlich auch viel zu kompliziert.

F : Mhm. Und wann würdest du sagen, hat man etwas in Stochastik richtig gut verstanden?

A : Ich glaube, nie.

F : [lacht] Ok.

A : [lacht] Ich glaube, Stochastik und der gesunde Menschenverstand, das sind zwei Dinge, die gehen nicht grundsätzlich d'accord, ja, sondern bei Wahrscheinlichkeiten oder Einschätzungen von Wahrscheinlichkeiten bis hin zum Erwartungswert, wo man im praktischen, ja, im realen Leben, oft mal daneben liegt, ja. Also das ist schwierig, das mit dem gesunden Menschenverstand in Einklang zu bringen.

F : Mhm.

A : Die grundlegenden Sachen schon, ja, das macht Sinn, ja, ähm, das kann man sich aneignen, auch verstehen, aber wenn du dann höher gehst, dann wird das schwierig.

F : Mhm. Jetzt hab ich noch eine, sozusagen Aufgabe.

A : Ok.

F : Und zwar, äh, die kennst du jetzt nicht, wenn du die Fragebögen nicht gesehen hast, ähm also es geht um äh zwei Krankenhäuser in einer Stadt, das eine ist ein großes und das andere ein kleines. Und man sagt jetzt, also die Wahrscheinlichkeit für ein Mädchen nehmen wir mal als 50% an.

A : Mhm.

F : Und in dem kleineren ähm werden jeden Tag 15 Kinder geboren im Schnitt und im großen ungefähr 45.

A : Ok.

F : Und jetzt ähm beobachtet man das erstmal über ein Jahr und guckt, äh an wie vielen Tagen waren jeweils in diesen beiden Krankenhäusern mehr als 60% der Kinder weiblich.

A : Mhm. Ok.

F : Und zählt diese Tage. Und die Frage ist, in welchem der beiden Krankenhäusern erwartet man mehr solcher Tage.

A : Ok, Das ist ne gute Frage. […]

F : Angenommen, du müsstest das jetzt simulieren [lacht]

A : [lacht] Ich wusste, dass ich [?]

F : weil das auf den ersten Blick vielleicht nicht ganz klar ist, ähm, wie würdest du da vorgehen? Also du musst jetzt keinen Code oder so was sagen, sondern es ist nur interessant, welche Schritte du machen würdest

A : Ok. […]

F : Also, du kannst sehr allgemein beschreiben.

A : Ich guck's mir kurz nochmal ganz kurz durch, ja.

F : Mhm, ja, lass dir ruhig Zeit.

A : […] Ist ne gute Frage. Ähm, wenn ich das simulieren würde, […] also ich würde irgendwie die ähm […] ich würde zwei Simulationen laufen lassen, ja, eine Simulation für das eine Krankenhaus und eine für das andere. Dann würde ich wahrscheinlich äh ähm, das muss ja über den Erwartungswert laufen […], der Erwartungswert, na der Erwartungswert ist ja hier dann die Wahrscheinlichkeit von 50%, würd ich sagen. […] In welchem Krankenhaus erwarten Sie mehr solcher Tage? Und die Mädels […] mit einer Wahrscheinlichkeit von 60% geboren werden. […] Da muss ich echt überlegen. Also eigentlich sind die ja, ich weiß nicht, ob man die eine, das eine Krankenhaus aufs andere einfach transformieren kann oder irgendwie ähm einfach irgendwie ne Simulation für 15 Kinder machen kann und das dann hochrattern auf 45 Kinder, ja. Weil von den, also von der Anfangswahrscheinlichkeit sind die ja Grundannahmen ja gleich, fifty fifty, ja. Und dann schwankt das sozusagen um den, um den Durchschnittswert ab und zu, von Tag zu Tag, ja, stabilisiert sich aber gegen die Wahrscheinlichkeit von 50%. […]

F : Aber, also du zählst quasi an jedem Tag, also du betrachtest dann immer nur einen Tag, also du hast halt sozusagen in dem einen Krankenhaus im Schnitt immer 15 Beobachtungen und im anderen ungefähr 45.

A : ok

F : und dann zählst du einfach immer pro Tag, also du sagst, ok, an dem Tag waren irgendwie wurden 15 Kinder geboren und davon waren, was weiß ich, 11 Mädchen, also mehr als 60%, und bei dem anderen entsprechend .

A : Ok, ok. Gut, dann muss ich einfach überlegen, was ist den, äh, boah, also ich tu mich irgendwie ein bisschen schwer. […]

F : Ist kein Problem, also du musst dich jetzt nicht irgendwie gezwungen fühlen. Aber ich mein, vielleicht nochmal zurück zu deiner ersten Idee, dass du zwei Simulationen machst.

A : Mhm.

F : Also du hast einmal das eine Krankenhaus und das andere.

A : Mhm.

F : Und, wenn du jetzt, also vielleicht kannst du einfach nochmal beschreiben, was du da simulieren würdest, das reicht schon, also muss jetzt gar nicht so detailliert sein.

A : Mhm. Also, ich würde probieren bei dem ersten Krankenhaus ähm, muss ich kurz nochmal durchlesen […] warden 45 Kinder geboren, im kleineren 15, ok. Beim ähm ersten Krankenhaus, um das nochmal zusammenzufassen, werden 15 Kinder pro Tag geboren, ja, und dann würd ich einfach sagen, ähm würd ich ne Prozedur oder irgendwie nen Befehl suchen in R und würd dann sagen ähm in dem ersten Krankenhaus 15 Kinder und das über einen Zeitraum von 365 Tagen ähm spuck mir irgendwie ähm also mit dem Erwartungswert ähm 50% Jungs, 50% Mödels aber mit Schwankung drum herum um den Erwartungswert, spuck mir irgendwie, spuck mir irgendwie nen Histogramm aus, ja. Dann würd ich mir das Histogramm dazu anschauen, ja, oder ne Datenreihe, und dann würd ich probieren äh nochmal diese Ausreißer zusammenzufassen, da gibt's, glaub ich, ne table-Funktion oder so was in R und die listet dann Ausreißer auf. Nicht die bzw. die Anzahl der Ausreißer würd ich notieren, wo es wirklich mehr als 15, mehr als äh 50% sind, mehr als 60%. Entschuldigung. Die dann summieren und mit der zweiten Simulation vergleichen.

F : Ok.

A : Das wär so meine Überlegung.

F : Wunderbar, wunderbar. Mhm. Und wenn du das für die 365 Tage gemacht hast, ähm, dann vergleichst du einfach die Werte und sagst, wo ist es mehr.

A : Genau. Dann würd ich sozusagen ähm die erste Simulation mir anschauen von dem

Krankenhaus, wo nur 15 Kinder geboren werden, und gucken, hat der mehr Ausreißer bezüglich der 60%igen Erwartung von Mädels, und das einfach vergleichen, ja mit der gleichen Simulation angewandt auf das zweite Krankenhaus, ja.

F : Vielen Dank, Ähm, möchtest du sonst noch irgendetwas anmerken?

A : Äh, eigentlich nicht, oder? Muss ich jetzt mal überlegen. Fällt mir noch irgendetwas ein? Ja, keine Ahnung, also ich tu mich generell ein bisschen schwer mit Stochastik einfach und deswegen auch war diese Frage der stochastischen Intuition bei mir so, ja, eher, eher mau ausgedrückt, ja. Und ähm, ich lieg halt immer daneben irgendwie, wenn ich Wahrscheinlichkeiten einschätze, hab auch, ja, und das find ich dann immer so interessant, ich hab auch, wie kriegt man das denn jetzt einigermaßen jetzt hin, ja, oder was sind die Kennzahlen, ja und ähm, ja, also ich find das schon spannend, super spannendes Fach, aber für mich ist das ein sehr unmathematisches Fach, wenn ich das sagen kann, ja. Also wenn ich das mit Ana, mit der Höheren Analysis oder äh mit der Diskreten Mathematik vergleiche oder mit der Numerik, ähm, wo man ganz klare Schemata hat, ja, und die sozusagen durchgeht und dann auch die Gesetzmäßigkeiten dahinter begreift, das ist hier in diesem Stochastik-Schein eher ein spielerisches Moment, ja. Ja, da musst du mitdenken. Und auch selbst tüfteln und auch, das ist mir auch schwierig gefallen, ist, diese ganzen Aufgaben oder Aufgabenstellungen zurückzuführen auf geeignete Modelle. Ja, ist das jetzt ein Urnenmodell mit Zurücklegen, ohne Zurücklegen, ist das jetzt ein Münzwurf, und was auch immer, und da muss man sich erst reinarbeiten. Aber ist auch spannend. Aber ist anders. Ich wars bisher nicht gewohnt [lacht].

F : Ok..

A : War ein bisschen überrascht.

F : Ok. Ja, wunderbar. Vielen Dank.

Int 8

F.: Genau, also, die erste Frage ist: Was war deine erste Empfindung, als du von den R Aufgaben gehört hast, und hat sich an deiner Einstellung etwas geändert bis heute?

A.: Ähm, ich fand es eigentlich ganz gut, wir haben ja auch im ersten Semester diese ECM Aufgaben, ähm, und da benutzen wir ja MAPLE und ähm ich fand es eigentlich ganz gut, dass in der Vorlesung R gemacht wurde, weil, ähm, das Packet halt oft benutzt wird und ist auch Freeware und deswegen kann man es auch selber benutzen und im Studium auch weitergehend machen, weil sonst, also mit R kommt man sonst ja nur im Statistikschwerpunkt in Berührung und ich mache halt Wahrscheinlichkeitstheorie als Spezial, also als Spezialisierung im Bachelor und da war es ganz gut.

F.: Mhm, mhm. Wie häufig hast Du denn die R Aufgaben bearbeitet?

A.: Ähm, ich glaube so 70%.

F.: Ok und warst Du mal in der Fragestunde?

A.: Nee.

F.: Und warum nicht?

A: Also, ich hab, ich bin jetzt schon im 6. Semester und ich habe, die Vorlesung sollte ja eigentlich im zweiten Semester gehört werden und ich habe sie damals geskipped, habe aber schon die Vertiefung bei Herr Neininger gehört, also die Prozesse und die Algorithmenanalyse und dann, also ich habe mich dann in der Lage gefühlt das auch alleine hinzukriegen.

F.: Ok, wunderbar, ich wollte nur wissen, ob es irgendwie aus zeitlichen Gründen nicht ging.

A.: Nee, nee, überhaupt nicht.

F.: Ähm, fandest du die R-Aufgaben im Vergleich zu den anderen Aufgaben eher schwer oder eher leicht?

A.: Ähm. Teilsteils. Also ich fand da waren ein paar dabei, die ich relativ schwierig fand, aber größtenteils eigentlich ok, also ich glaube so ungefähr wie die anderen Übungsaufgaben.

F.: Mh, ok. Und, ähm, gab es an den R Aufgaben irgendetwas, was Du besonders gut fandest?

A.: Ähm, ich fand es gut, dass man die Skripte schon so ein bisschen, so ein Skelett von den Skripten schon so bekommen hat, weil man sich dann nicht so viele Gedanken machen muss, wie beschrifte ich jetzt Sachen und so was, das war immer schon vorgemacht, das fand ich ganz gut also, dass man nur noch die Schleifen einfügen muss und so, weil das kostet auch immer so ewig viel Zeit, diese Achsentitel und was

weiß ich was zu machen, da muss man sich immer stundenlang durchquälen, wenn man da noch nichts gemacht hat mit.

F: Mhm.

A: Also das fand ich gut.

F: Mhm. Und an welche R Aufgaben erinnerst du dich noch?

A: Ähm, an das Crabs-Spiel, ähm, aber daran auch, weil ich die Übungsaufgabe dazu, ich glaub, da hab ich ziemlich lang für gebraucht, irgendwie [lacht], das weiß ich noch, das mit dem Münzwurf weiß ich noch, ähm, die Monte Carlo Simulation, das fällt mir jetzt spontan ein.

F: Mh, ist gut. Ähm, gibt es irgendetwas, was dir durch eine R Aufgabe klar geworden ist, was du vorher nicht verstanden hattest?

A: Hmm, ne, ich glaub nicht.

F: Mhm, und gibt es irgendetwas, was man deiner Meinung nach an den R Aufgaben besonders gut lernen kann, also jetzt bezogen auf Stochastik?

A: Ähm, ich find halt Grenzwertsätze immer ganz schön mit R, weil das immer sehr, sehr formalisiert ist und man kann sich das schwer vorstellen, aber obwohl das relativ in Anführungsstrichen relativ leicht ist und man kann da mit R ein bisschen rumprobieren und Mittelwerte sich da angucken und schauen, wie sich da was verschiebt, ja.

F: Mhm. Gibt es irgendetwas, was du bei den R Aufgaben nicht gut fandst, was dir nicht gefallen hat?

A: Da war eine dabei, das weiß ich noch, die war unheimlich viel, also da war auch mega viel schon vorgegeben an Text, aber ich weiß nicht mehr, welche Aufgabe das war, und ich glaub, die hab ich dann im Endeffekt auch nicht gemacht, weil es mir zu viel war.

F: Waren das die Aufgabe mit den Kugelkoordinaten?

A: Ja, das könnt sogar sein.

F: Wo man so eine Kugeldreiecksfläche berechnen sollte?

A: Ja, genau, ich glaub schon.

F: Ok.

A: Ja, ansonsten, ich weiß nicht, also ich fand sie sonst gut.

F: Mhm, würdest du denn empfehlen, weiterhin R Aufgaben zu verwenden und

A: Ja, auf jeden Fall. Und ich würde es auch in anderen Vorlesungen anwenden, also zum Beispiel jetzt bei den stochastischen Prozessen, da hätte ich es mir auch gewünscht, wenn es da gemacht worden wäre.

F: Mhm.

A: Ja, oder auch in Analysis, jetzt nicht mit R unbedingt, aber ja, es ist immer gut, die Programmieraufgaben zu machen, also das hilft schon, das kann man ja auch weiter verwenden, wenn das mal in ner Vorlesung nicht gemacht wird, man aber mit dem Paket vertraut ist, dann kann man ja auch bei Übungsaufgaben ein bisschen rumprobieren, also mal Kugeln zeichnen oder für Differenzialgleichungen Sachen machen und so was, also das find ich schon ganz cool, und gerade auch bei Stochastik bietet sich das ja an, weil man viel auch sichtbar machen kann.

F: Ja, mhm, und würdest du an dem Konzept etwas verändern?

A: Ich fands eigentlich ganz gut. Ich würde eventuell, ja, das aus der, schwierig, also, ich hatte, ich war immer ein bisschen unter Druck, die zu machen, wegen den Bonuspunkten, aber ich glaube, wenn ich, also wenn es nicht die Bonuspunkte-Regelung gegeben hätte, hätte ich es wahrscheinlich nicht gemacht, von daher, ähm, ja. Ansonsten, ich weiß halt nicht, in welchem Umfang da noch diese Übungen angeboten wurden, weil ich ja nicht da war [R Fragestunde] , ich hätt mir vielleicht gewünscht, dass das in Tutorien, ich weiß nicht, ob das gewünscht war, aber in den Tutorien wurde das ziemlich oberflächlich behandelt nur, also vielleicht mal kurz erklärt, was der Sinn der Aufgabe ist, aber dass da auch ein bisschen was mit R auch richtig gemacht wird, also ein paar Skript-Beispiele mal aufgeschrieben werden oder so, mal mit dem Beamer so was gemacht, ja das sollte man vielleicht machen.

F: Mhm, ok. Wurde denn manchmal im Tutorium auf R Aufgaben Bezug genommen, wenn andere Aufgaben erklärt wurden?

A: Ähm, eher selten.

F: Fällt dir irgendein Beispiel ein?

A: Ähm, ne.

F: Mhm, ok. Jetzt mal ein bißchen was inhaltliches, ähm, woran denkst du als erstes, wenn du an Stochastik denkst?

A: Ähm, Brownsche Bewegung, ehrlich gesagt.

F: Gut, da merkt man, dass du schon weiterführendes gehört hast [lacht]. Und aus der Vorlesung irgendwas?

A: Ich fands schön, dass wir so viel mit Indikatorvariablen gemacht haben, also, das sind so die Sachen, wo ich am meisten mitgenommen habe, weil ich das, ich hab natürlich schon öfter was damit gemacht, aber so nochmal ein paar Beispiele zu sehen, für was man das verwenden kann und so, das fand ich schon ganz schön.

F: Mhm. Und kannst du erklären, was eine Indikatorvariable ist?

A: Ja, eine Funktion, die die Werte Null oder Eins annimmt im Prinzip, ja, und die ist immer Null für das Komplementärereignis und Eins für das Ereignis selber, ja.

F: Mhm. Was verstehst du unter stochastischer Intuition? Das ist ein Begriff, der in den Fragebögen aufgetaucht ist.

A: Hm, das bring ich immer mit Sachen abzählen, also Kombinatorik in Verbindung, das fällt mir da ein, also das mag ich persönlich nicht so gerne, aber es stimmt schon, das ist mir früher auch unheimlich schwer gefallen, so wieviele Kombinationen gibt es, die und die Sachen zu verteilen und so was, aber je mehr man hört, desto einfacher fällt es einem, ja wenn man so eine Intuition entwickelt, wie man das machen kann.

F: Mhm.

A: Also Kombinatorik ist das, was ich darunter verstehe.

F: Mhm, und gibt es einen Unterschied zu Verständnis von Stochastik?

A: Ähm, meinst du jetzt zwischen Intuition und Verständnis von Stochastik?

F: Mhm.

A: Ja, auf jeden Fall, ich glaub schon. Also ich sag mal, den Stoff zu verstehen ist ja was anderes als die Art und Weise, wie du an eine Aufgabe ran gehst und wie schnell du da auf ne Lösung kommst oder auf den richtigen Weg kommst. Das sind auf jeden Fall zwei unterschiedliche Sachen, aber es beschränkt sich auch nicht nur auf Stochastik, würd ich sagen.

F: Mhm. Und kannst du sagen, wann du sagen würdest, man hat etwas wirklich verstanden in Stochastik?

A: Ähm, also vom Stoff her geht's immer relativ schnell, aber ich glaub so wirklich dieses tiefgehende Verständnis, das kommt erst mit den Übungsaufgaben, und oft auch nicht während der Vorlesung, sondern was später auch teilweise, wenn man höhere Vorlesungen hört oder so was, die noch mal ein anderes Licht auf die Sachen werfen und dann fallen die Sachen auf, wo du denkst, ok die hab ich damals gar nicht so verstanden oder dachtest, das wär irgendwie anders, also ich denk, das kommt dann mit der Zeit oder so.

F: Mhm.

A: Das ist Erfahrungssache.

F: Mhm, ok. So, dann hab ich noch eine Aufgabe sozusagen und zwar kam es auch

A: Oh Gott.

F: bei den Fragebögen vor, ist nichts Tragisches, keine Sorge. Und zwar geht es um dieses Krankenhausproblem, also man hat ein großes und ein kleines Krankenhaus

A: ja

F: und im großen werden jeden Tag ungefähr 45 und im kleinen ungefähr 15 Kinder geboren. Und jetzt sagt man, die Wahrscheinlichkeit für ein Mädchen ist 50%

A: ja

F: man beobachtet jetzt für den Zeitraum von einem Jahr diese beiden Krankenhäuser und zählt jeweils die Tage, an denen mehr als 60% der Kinder weiblich waren

A: ja

F: und die Frage ist, in welchem Krankenhaus erwartet man mehr solcher Tage. So, und für den Fall, du möchtest das simulieren. Wie würdest du vorgehen? Also ohne Code jetzt natürlich, nur die Schritte.

A: So was mit Krankenhäusern hatten wir doch auch bei der Größenverzerrungsaufgabe, oder?

F: Ja, wobei es hier wahrscheinlich leichter ist.

A: Ja. […] Also ich mein, die Wahrscheinlichkeit von Junge oder Mädchen ist ja wie beim Münzwurf, ähm, das heißt, ich müsste im Prinzip für das große Krankenhaus 45 Münzwürfe pro Tag simulieren und für das kleine 15.

F: Mhm.

A: [...] Und dann, wenn ich mehr als 60% habe, müsst ich irgendwie ein 1 speichern und es dann aufsummieren über die 365 Tage. [...]

F: Mhm.

A: Ok. Ja, aber die müssten doch gleich sein, oder? [lacht] Ungefähr. Die Anzahl der Tage in beiden Krankenhäusern, oder?

F: Ok, klären wir gleich [lacht]. Nochmal ein Schritt zurück, also du machst das 365 Mal? Wir können gleich nochmal drüber reden.

A: Ähm, ja, jetzt so.

F: Mhm. Und dann schaust du einfach, wo kommt es öfter vor und das nimmst du?

A: Ja, genau.

F: Ok. Und du sagst jetzt, in beiden ist es gleich?

A: Ja, oder?

F: Das ist interessant, das sagen viele Leute. Es stimmt aber nicht, also im kleineren ist die Wahrscheinlichkeit größer. Also, du hast ja, also 50% ist ja, angenommen, die wahre Wahrscheinlichkeit und du schätzt ja im Grunde diese 50%.

A: Ah, ja, ja.

F: Und da wo 60% sind, ist das ja eine schlechte Schätzung.

A: Ja.

F: Und da du im kleinen einen höheren Standardfehler hast, kannst du es dort eher erwarten.

A: Ja, verstehe. Ja, das macht Sinn. Ich, äh, ja, aber rein theoretisch ist es gleich, oder? Also, wenn man jetzt nicht schätzt, sondern es genau ausrechnest, dann ist es doch, also im Mittel zumindest.

F: Also wenn du in beiden eine große Anzahl hast, also z.B. über alle Tage mittelst, dann würde es sich in beiden Fällen der 50% annähern.

A: Ah, ja.

F: Aber du hast ja immer nur 45 oder 15 Beobachtungen pro Schätzung.

A: Ja, stimmt. OK.

F: Gibt es sonst noch etwas, was du anmerken möchtest?

A: Also ich fands durchweg positiv, also mir hat es gut gefallen. Ich weiß, dass viele das nicht so gerne mögen, diese Programmiersachen, oder dass sie, ja, ist halt immer schwierig, weil man das am Anfang immer machen muss und wenn man einmal den Faden verliert, dann kann man es in der Regel nicht mehr machen, weil es dann zu schwierig ist und man zu lange brauchen würde um es zu machen. Aber ich fands super.

F: Mhm, vielen Dank.

[Angaben zur Person.]

F: Also, die erste Frage ist, was war Ihre erste Empfindung, als Sie von den R Aufgaben gehört haben und hat sich an Ihrer Einstellung bis heute etwas geändert?

A: Ähm, ich verstehe die Frage nicht.

F: Achso, also was war so Ihr erstes Gefühl, als Sie von den R Aufgaben gehört haben?

A: Ähm, ja, ich habe gedacht, vielleicht das würde schwierig sein, aber danach, ja am Anfang hatte ich ein bisschen Angst gehabt, ne. Aber nach ich dann die Codes geguckt, und dann sehe ich die Dinge und sag, ok, das war nicht so schlimm [lacht].

F: Ah, ok. Und wie häufig haben Sie die R Aufgaben bearbeitet?

A: Ähm, alle außer eins.

F: Ok. Und waren Sie mal in der Fragestunde?

A: Nein, war ich nicht.

F: Aus Zeitgründen, oder gings so?

A: Äh, nein, aus Zeitgründen. Weil wenn Sie die, die, während diese schedule für die Fragestunde gehabt, muss, musste ich mein eigenes Tutorium leiten, während dieser Übungsstunden, ja.

F: Ah, ok. Fanden Sie denn die R Aufgaben im Vergleich zu den anderen Aufgaben eher leicht oder eher schwer?

A: Leichter.

F: Leichter. Und was war leichter?

A: Ähm, ja, weil das war, es hatte zu tun mit programming, ja, ok, dass man hat diese Vorkenntnis mit programming und das hat zu tun mit wie man eine Funktion schreibt oder wie man diese variable nutzt und so weiter.

F: Mh.

A: Ja.

F: Ok. Und was fanden Sie an den R Aufgaben gut?

A: Ähm, in den R Aufgaben, dass wir haben dann gelernt die Theorie und dann sieht man das in einer Graph. Das sieht gut aus, ne. Jetzt hat man so ein bisschen mehr von Theorie zu etwas man kann das sehen, füllen, so irgendwie.

F: Ah, ok.

A: Ja.

F: Und ähm erinnern Sie sich noch an eine R Aufgabe?

A: Ähm, eine R Aufgabe, ähm was ich habe so schwer zum Beispiel gefunden war diese über diese Kugel, ähm diese ähm so wie Monte Carlo Schätzung.

F: Ja. Ja.

A: Ja, ob diese rein zufällige Punkte in diese Dreieck fallen, ja reinfallen oder nicht.

F: Mhm. Und fallen Ihnen noch weitere ein?

A: Ähm, was ist auch, ah, was hat mich ganz überrascht, war diese Run von Münzwürfen.

F: Aha.

A: Ja. Habe ich immer gedacht, das wird fifty fifty, wenn man da so rein zufälliger Münzwurf hat. Ich habe nie gedacht, bis ich diese Aufgabe gemacht habe, ja, dass wenn man ähm so eine ob das ist ähm […] ähm wie sagt man das […]

F: Sie können das auch gern auf Englisch sagen.

A: Äh, ob das ist ein Betrug oder nicht.

F: Ah, ok.

A: Ja. Ich habe das nicht gewusst, bis ich diese Aufgabe gemacht habe.

F: Ah.

A: Das war interessant, ja.

F: Und gibt es sonst noch irgendetwas, was Ihnen bei einer R Aufgabe klar geworden ist, was Sie vorher nicht verstanden hatten?

A: Äh, in einer […] äh, nochmal?

F: Ähm, ich kanns auch auf Englisch formulieren, wenn das leichter..

A: Ne, ne, einfach nochmal.

F: Ok. Gibt es irgendetwas, was Sie durch eine R Aufgabe verstanden haben, was vorher nicht klar war?

A: Achso. Ähm, was vorher nicht […] ähm, ja, ich glaube so, zum Beispiel wenn man ähm, das war diese Aufgabe mit diese äh Krankenhaus mit Bettliegen und so weiter ja. Und bevor das hat man das in Theorie so alle diese ähm Sätze und Definition gelernt und jetzt mit dieser Aufgabe das, irgendwie das ist alles, es gibt eine Beziehung zwischen alle diese Sätze und Definitionen.

F: Ok.

A: Ja.

F: Ok, schön, das freut mich. Und gibt es irgendwas, was man allgemein an diesen R Aufgaben besonders gut lernen kann? Also jetzt speziell auf Stochastik gesehen?

A: Ähm, was man gut lernen kann, ähm, was besonders, ähm, […] ja, ich glaube, dass ähm vielleicht dass durch diese R Aufgaben kann man dann so viele, zehntausend, zwanzigtausend Fälle simulieren und dann kann man, kann man sehen, ok, wie die Verteilung dann fällt und so weiter, ne.

F: Ok.

A: Und das ist ganz anders, wenn man nur mit einer Papier kalkulieren und dann sagt, ja ok, aber ist das wahr oder nicht, ne. Man hat keine Gefühl, ja.

F: Ok. Und gibt es etwas, was Sie bei den R Aufgaben nicht gut fanden?

A: Ähm bei den R, vielleicht weil ich bin von der Informatiker-Seite, manchmal ich verstehe die Aufgabe nicht.

F: Ok.

A: Ok, ich muss das manchmal so viele, nochmal lesen, bevor ich die Aufgabe verstehe, manchmal habe ich das falsch verstanden.

F: Ok. Würden Sie denn empfehlen, auch weiterhin R Aufgaben zu benutzen?

A: Ja.

F: Mhm. Und würden Sie irgendwas verändern?

A: Ähm, […] vielleicht, was, was hat mir gefallen, ist eine von diesen Aufgaben, wo, ich kann mich nicht erinnern, das war eine Aufgabe hat zu tun mit die Nicht-R-Aufgabe […]

F: Ah ja, das war bei den Fehlständen bei den Permutationen.

A: Genau, bei den Fehlständen der Permutationen. Und das ist gut, wenn man so eine Aufgabe vorher so eine Aufgabe hat und dann direkt danach eine R Aufgabe hat, dann hat man von Theorie zur Praxis eine Beziehung, ne.

F: Ok. Und das sollte öfter vorkommen?

A: Genau.

F: Mhm, ok. Ähm, wurde denn in Ihrem Tutorium manchmal auf die R Aufgaben Bezug genommen, wurden die irgendwie angesprochen?

A: Gar nicht, nein.

F: Gar nicht, ok.

A: Meine Tutorin hatte keine Ahnung von R Aufgaben.

F: Ok. Jetzt nochmal was so ein bißchen allgemeineres: Woran denken Sie als erstes, wenn Sie an Stochastik denken?

A: Ähm, die Wahrscheinlichkeit, ne [lacht].

F: Ok. Und können Sie erklären, was das ist?

A: Wahrscheinlichkeit, ich denke das äh von so viele, zum Beispiel Münzwürfe, von so viel Mal, wenn man eine Münz werfen, ok, wie oft dann kriegt man Kopf oder Zahl. Ja, das ist die klassische Art, wie man erstmalig von Stochastik lernt.

F: Ok, wunderbar. Und was äh verstehen Sie unter stochastischer Intuition? Also, intuition ist das auf...

A: Ja, ich verstehe das. Ja, ich habe das von textbook gelesen, dass Herr [xxx] hat diese Terminologie benutzt, ich glaube, das ist so, dass anstatt das so alles in formales Theorie zu lernen, dann lernt man so, ich denke von, von das Leben oder was, von experience, ne. Ja, man sagt, von experience ich habe das schon einmal erlebt und, ja.

F: Ok. Und gibt es einen Unterschied zu Verständnis?

A: Äh, vielleicht von Intuition es ist leichter, ich denke, es zu lernen anstatt nur von Theorie. Ähm vielleicht liegt das, ich bin auch nicht von der Mathe-Seite, dann für mich, die Schwierigkeit ist, ich habe nicht diese Basis, ja, Mathe-Basis, so eine Mathe-Satz zu lesen und dann zu sagen, ja, ich verstehe das.

F: Ja, ok. Ähm, ich habe jetzt noch eine kleine Aufgabe und zwar das war eine Aufgabe, die auch in den Fragebögen vorkam.

A: Ah, ja.

F: Ähm, und zwar ähm. Also man hat zwei Krankenhäuser. Und in dem großen werden jeden Tag ungefähr 45 Kinder geboren und in dem kleinen ungefähr 15.

A: Mhm.

F: Und man sagt jetzt, man geht davon aus, dass etwa 50% aller Kinder Mädchen sind, also die Wahrscheinlichkeit für Mädchen 50%.

A: Mhm.

F: Und jetzt guckt man sich für einen Zeitraum von einem Jahr jeweils an, an wievielen Tagen in den beiden Krankenhäusern mehr als 60% der Kinder Mädchen waren.

A: Mhm.

F: Und die Frage ist, in welchem Krankenhaus erwartet man mehr von solchen Tagen.

A: Die kleinere. [lacht]

F: Ok, stimmt, genau. Die Frage ist jetzt, wenn Sie das simulieren müssten, dieses Problem, wie würden Sie da vorgehen? Also jetzt nich mit Code, sondern nur was für

Schritte würden Sie machen, wenn Sie das simulieren würden?

A: Ähm, wie meinen Sie das, nicht mit Code?

F: Also, wie, ich meine, Sie müssen jetzt nicht den Code irgendwie aufschreiben...

A: Achso, die Idee?

F: Die Idee, genau.

A: Wie man das simuliert. Ähm, ja, ok, man hat ähm die Wahrscheinlichkeit p ist 1 durch, ist halbe. Ja, ok, dann hat man dann ähm eine ist so 15 hoch halbe, ne halb hoch 15 oder halb hoch 45. Und dann sieht man dann, das Ergebnis, ja, welche ist näher zu 60%.

F: Äh, ok. Aber wenn Sie es jetzt mit dem Computer simulieren müssten?

A: Ja, also wenn ich das jetzt mit R Code, dann so simulieren.

F: Ah, ok, Sie würden das ausrechnen? Sie würden quasi ein halb hoch 15 ausrechnen und ein halb hoch 45.

A: Achso. Ähm, Moment. Ja, aber ich muss das dann mit dieser Verteilung, mit diese Standard, rnorm oder was und dann kriegt man, so man simuliert das 15 Mal und man simuliert das 45 Mal.

F: Ok. Und vergleicht?

A: Ja. Je nachdem, was man kriegt. Ist das die richtige Antwort?

F: Also, die Lösung ist richtig, dass das in dem kleinen Krankenhaus ist.

A: Mhm.

F: Aber die Erklärung stimmt nicht.

A: Ok. [lacht]

F: Weil wenn Sie ein halb hoch 15 rechnen, dann ähm würden Sie quasi ausrechnen die Wahrscheinlichkeit für eine bestimmte Abfolge von Jungen und Mädchen.

A: Mhm.

F: Aber es geht ja darum, wieviele Mädchen insgesamt, also die Reihenfolge ist ja egal.

A: Mhm.

F: Das heißt, eigentlich ist es binomialverteilt.

A: Ah, ok.

F: Ok, aber wunderbar. Es geht ja nur um die Idee, von daher ist alles prima. Ähm,

möchten Sie sonst noch irgendetwas sagen zu den R Aufgaben?

A: Ähm, R Aufgaben, ähm, ich habe viele von diese R Aufgaben dann ähm die äh zum Beispiel die Beispiele kurz äh das aus dem Internet gegoogelt und dann gelesen, zum Beispiel wenn man die Funktion oder was von matrix nutzen möchte, ist das leichter das aus Internet zu finden als in ein Buch zu lesen, aber jetzt hab ich angefangen äh mit diese Buch für die graphs zu lesen, das ist interessant.

F: Ok.

A: Weil damals, während diese Hausaufgabe, ähm ich habe nur äh die Codes von Ihnen für histograms oder was so kopiert, ja, nicht so in die parameters geguckt.

F: Ja.

A: Aber jetzt habe ich jetzt das richtig angefangen, das alles zu lernen..

F: Ach, klasse.

A: Ja, weil ich finde das so, so super, ja wie man dann eine Histogramm oder äh wie diese charts und so weiter zu malen, ja.

F: Mhm. Also Sie benutzen das noch weiter?

A: Ja, genau.

F: Das ist natürlich toll.

A: Ja.

F: Ja, wunderbar, das wars dann schon, vielen Dank.

Int10

F: Was war denn deine erste Empfindung, als du von den R Aufgaben gehört hast, und hat sich deiner Einstellung bis heute etwas geändert?

A: Ok, ganz am Anfang ähm konnt ich damit erstmal nicht so viel anfangen, weil ich ja R gar nicht kannte, ähm, gut ja, die Programmiersprache war eigentlich ganz in Ordnung, also die konnte man dann ganz gut lernen, wir hatten ja auch im ersten Semester ähm Einführung in die computerorientierte Mathematik, da hatten wir Python so ein bißchen gemacht, ähm, ja, das ist natürlich nicht dasselbe, aber ach es war dann ganz in Ordnung. Und jetzt am Ende rückblickend würd ich sagen, die Aufgaben waren immer ganz gut. Also die haben auch Spaß gemacht.

F: Mhm. Ok. Und ähm du hast die Aufgaben immer bearbeitet oder oft?

A: Ich glaube, die allerletzten hab ich nicht mehr bearbeitet, weil da hat ich dann schon genug Punkte, und ja.

F: Mhm. Und du warst auch häufig in der Fragestunde oder sogar immer.

A: In der Fragestunde war ich ziemlich häufig, ja. Zweimal hab ich gefehlt oder so.

F: Mhm. Ok, wunderbar. Ähm, fandest du denn die R Aufgaben im Vergleich zu den anderen Aufgaben eher schwer oder eher leicht?

A: Ähm, meistens ein bisschen leichter, es waren glaub ich zwei oder ja zwei, drei Blätter waren dabei, wo die R Aufgaben auch ein bisschen schwieriger waren, ja.

F: Und was war schwerer oder was war leichter?

A: Ähm, also was schwerer war, oder fang ich mal an, was leichter war bei den anderen Aufgaben, da war's, da hatten wir die Vorlesung daneben und wussten schon, ok, mit welchen Sätzen oder mit äh welchen Schritten wir da drangehen müssen, ähm bei den R Aufgaben, gut das haben wir in der Vorlesung ja nicht so viel gemacht, äh und dadurch dass es eben auch Programmieren war, was wir da in der Vorlesung auch nicht gemacht haben, ähm, war's dahingehend ein bisschen schwieriger. Also die Ansätze waren, meistens konnt man sich die schon überlegen, also ähm, es war dann halt vom Programmieren her ein bisschen was Schwieriges, aber ich mein mit den R Fragestunden hat sich das dann gut ergeben, die haben da dann sehr gut geholfen, ja.

F: Mhm. Ok. Und ähm was fandest du an den R Aufgaben gut?

A: Ähm, ich fand, allgemein find ich's gut, dass ne Programmiersprache mit reingenommen wurde, ich glaub, das ist ganz wichtig, dass man ähm auch programmieren irgendwo mitlernt und dass man auch sieht, weil ähm man kriegt ja später nicht so Aufgaben, die man per Hand immer ausrechnet, sondern so rechnerunterstützte äh ähm Mathematik ist, fand ich, also ich fands einfach sehr hilfreich oder auch für die Zukunft irgendwie ähm interessant zu wissen, ja.

F: Mhm. Und an welche R Aufgaben erinnerst du dich noch?

A: Ähm, also an diese ganz schwierige mit den drei Städten auf der Kugel, das war, glaub ich, die, die ich so am schwierigsten fand, ach, es gibt da ganz viele, wir hatten diese Aufgabe mit den, mit den Tests, ähm und die mit den Betten, mit den Belegzeiten, dann hatten wir so Urnenaufgaben ähm, muss mal überlegen, was hatten wir alles, ähm, wir haben diese ähm Bernoulli-Experimente hatten wir was, mir fallen jetzt noch ein paar Graphiken ein, aber die Aufgaben nicht mehr ganz dazu.

F: Ja, aber das sind ja schon viele.

A: Mhm.

F: Ist nur auch interessant, was als erstes genannt wird, deshalb frag ich.

A: Ok, gut.

F: Ähm, gibt's denn irgendwas, was dir mal durch eine R Aufgabe klar geworden ist, was du vorher nicht verstanden hattest?

A: Ja, da war was. Ähm, da war tatsächlich was, das war äh, was war denn das nochmal? Ich muss mal überlegen, wir hatten in einer R Aufgabe hatten wir ähm oder ja, was heißt klarer geworden, also wir hatten was, was in der Vorlesung dran war, hatten wir in der R Aufgabe irgendwie nochmal aufgenommen und ähm da ist dann, mir ist dann also, ich fands erstmal interessant, weil wir haben ja oft diese ähm nicht gleich nen Rechenweg, wo man ne Lösung hatte, sondern immer so Versuche gemacht, irgendwie so Durchläufe, um mal zu sehen, wie man mathematisch da ran geht. Äh jetzt fällt mir nicht mehr genau ein, was das war, was ich, irgendwas war da, was ich in der Vorlesung nicht ganz verstanden habe und was durch die R Aufgabe dann irgendwie klar geworden ist. Jetzt […]

F: Ich start mal ein paar Versuche, äh Konfidenzintervalle?

A: Ne, die warens nicht.

F: Diese Sache mit den bedingten Wahrscheinlichkeiten?

A: Die warens auch nicht, bedingte Wahrscheinlichkeiten warens nicht.

F: Die äh Fehlstände?

A: Die Fehlstände, ich glaub, die könntens gewesen sein. Ja, ja genau, doch Fehlstände.

F: Ich glaub, das war ne Sache, die auch in der Vorlesung vorkam, ok.

A: Mh. Ja.

F: Ok. Und gibt es allgemein irgendetwas, was man an den R Aufgaben deiner Meinung nach besonders gut lernen kann bezogen auf Stochastik?

A: Ähm naja, wie mans halt rechnerunterstützt äh bearbeitet und auch ähm, was ich gut fand, war, man kann da Ideen ausprobieren. Weil äh man kann eben erstmal versuchen, so ähm na so Proben zu machen sozusagen. Und äh ein Gefühl dafür bekommen, was, was man eigentlich oder was eigentlich hinter ner Aufgabe steckt.

F: Mhm, mhm. Ähm, was hat dir bei den R Aufgaben nicht gefallen?

A: Ähm, ich fand, also was ich schwer fand, war am Anfang ähm dass, dass wir halt überhaupt nichts von R wussten. Äh da gabs dann, wie gesagt die R Fragestunden haben das Problem ja dann gelöst im Endeffekt, ähm es war vielleicht für die Leute, die die R Fragestunden nicht wahrnehmen konnten, könnt ich mir vorstellen, also wenn ich nicht die R Fragestunden hätte wahrnehmen können, wär es für mich ziemlich schwierig gewesen. Also das war vielleicht, aber von den Aufgaben her, also, da war die Unterstützung immer da, da war, und die waren auch in nem angemessenen Schwierigkeitsgrad, also ja, gibt's nichts zu meckern [lacht].

F: Ok, umso besser [lacht]. Ähm, wurde denn, achso ne. Würdest du denn aus Studentensicht empfehlen, weiterhin R Aufgaben zu verwenden...

A: Ja.

F: und würdest du etwas ändern?

A: Ähm, also ich würde es auf jeden Fall weiter empfehlen. Ändern äh was vielleicht nicht schlecht ist, ist wenn man so freiwillige Aufgaben macht, ähm also die waren ja jetzt immer in die Übung eingebunden, so dass man auch irgendwo nen gewissen Druck hatte jetzt, sag ich mal, die auch zu machen, was auch gut ist, also so ist jeder dazu im Prinzip gezwungen, sich auch damit auseinander zu setzen, aber für die Leute, denen das vielleicht nicht reicht oder die ein bisschen mehr Übung haben wollen, wären so freiwillige Übungen vielleicht nicht schlecht.

F: Ok. Mhm. Das ist ein guter Tipp. Ähm wurde in deinem Tutorium denn manchmal auf R Aufgaben Bezug genommen, wenn eine andere Aufgabe erklärt wurde?

A: Ähm, jein, also es äh von der Tutorin wurde das erwähnt, immer wenn irgendwie ne Aufgabe mit der R Aufgabe irgendwie zusammenhing, aber wir haben dann, über die R Aufgaben haben wir sehr wenig eigentlich nur gesprochen. Also, ja, die wurden halt korrigiert und ich weiß nicht, wie das jetzt war, ich hatte nicht so oft, war ich in ner Situation, wo ich ne Aufgabe nicht lösen konnte, eben weil ich dann, wenn ich irgendwas nicht verstanden hatte oder so, bin ich in die R Fragestunde gekommen, ähm, ich weiß jetzt nicht, wie das für diejenigen war, die ähm die Aufgaben nicht verstanden haben, im Tutorium haben wirs halt nicht wirklich so viel besprochen.

F: Ja, ja. Ok. Ähm aber es war jetzt auch nicht so, dass irgendwie Studenten, die ne andere Aufgabe jetzt erklärt haben, die jetzt nicht die R Aufgabe war, mal gesagt haben, ja, das haben wir ja in der R Aufgabe gesehen und hier kommt so was vor?

A: Doch, das kam vor. Das wurde schon mal gesagt.

F: Aber dann eher von den Studenten, nicht von der Tutorin?

A: Doch, doch.

F: Auch von der Tutorin, ok.

A: Ja, die hat auch Hinweise gesagt, wenn da ne Aufgabe war, das hat sie schon gesagt.

F: Ok. Ähm, das ist nämlich interessant, äh, da haben nämlich, ich hab die Tutoren nämlich auch gefragt so einige Sachen und ähm es haben irgendwie mehrere Tutoren geschrieben, dass es häufig vorkam, und du bist jetzt aber der erste von den Studenten, die sagen, dass es tatsächlich so war.

A: Ach ja.

F: Und das hat mich jetzt, also ist interessant, jetzt zumindest einer mal, der das sagt. Weil die anderen haben immer gesagt: Was? Nein, nie. [lacht]

A: Also erwähnt wurd's schon, aber wie gesagt, aber was nicht [?]

F: Ihr habt es nicht besprochen, sozusagen?

A: Mhm, genau, das war nicht der Fall.

F: Mhm. Ähm jetzt nochmal so ein bisschen was Inhaltliches zu Stochastik.

A: Mhm.

F: Ähm, woran denkst du als erstes, wenn du an Stochastik denkst?

A: Ähm, ich glaube, woran ich jetzt immer als erstes denke, ist Zufall, ja. Das war ganz am Anfang schon im Buch wurde dieser Begriff aufgegriffen und das äh ist wohl ziemlich zentral, also ist es ja auch.

F: Aha. Und kannst du erklären, was das bedeutet?

A: Zufall?

F: Mhm.

A: Ähm, ja, ah ja, gut, ne formale Definition, also man hat ähm, man hat Ereignisse und jedes Ereignis tritt eben mit irgend ner bestimmten Wahrscheinlichkeit ein, die man vorher kennen kann, aber nicht muss, und ähm das Zufallsexperiment ist dann eben, dass man äh ja, so, so nen ähm Durchlauf macht, wo alle Ereignisse vorkommen können und dann guckt man eben, ja, der Zufall bestimmt dann eben, welches Ereignis halt rauskommt.

F: Mhm. Mhm. Und fällt dir noch ein anderer vielleicht Begriff ein aus der Vorlesung?

A: Ähm, ja, jede Menge. Also [lacht] ähm, ein zentraler, oder?

F: Mhm.

A: Äh, hm.

F: Also irgendwas, was dir so sofort in den Sinn kommt.

A: Bernoulli.

F: Bernoulli, ok.

A: Ja. Bernoulli ist äh, das hat man auch schon in der Schule, also ein Experiment, wo ähm immer dieselbe Wahrscheinlichkeit auftritt und eigentlich nur zwei Möglichkeiten, ähm ja, Treffer oder Nicht-Treffer waren das immer, und ja, es gibt eben eine Trefferwahrscheinlichkeit und dann konnt man damit mehrstufige Experimente machen oder einstufige, wie auch immer.

F: Mhm. Wunderbar, mhm. Ähm, dann noch ein Begriff, der in den Fragebögen vorkam, ähm was verstehst du unter stochastischer Intuition? Also, es kam in den Fragebögen und im Buch vor.

A: Ah, mhm.

F: Und mich würd interessieren, was du darunter verstehst.

A: Ja, ok. Also äh stochastische Intui, also stochastische Intuition ähm, also wenn man jetzt irgendne Aufgabe bekommt, die halt äh nen stochastischen Inhalt hat, und man soll jetzt spontan ohne zu rechnen irgendwie sagen ähm, was wäre dann da so denkbar, was die Lösung wohl, also in welche Richtung geht die Lösung vielleicht. Und stochastische Intuition wär dann eben irgendwo zu wissen, ok, welches Modell trifft darauf zu und ähm ja, dann kann man eben, es ist ja auch so ein bißchen, für stochastische Intuition muss man ja nicht unbedingt ähm also zumindest für einfachere Sachen muss man ja nicht unbedingt jetzt Mathematik studiert haben. Wenn man irgenwie nen Würfel hat und man wird irgendwie gefragt, wenn ich jetzt zehn Mal würfel, wie oft kommt jetzt da die Sechs vor oder so, da kann man ja auch logisch denken und da, das ist dann irgendwie klar, ok, Wahrscheinlichkeit ist ein Sechstel, ich werf zehn Mal, dann ist das halt zehn Sechstel. So, und das ist, also. Stochastische Intuition ist für mich auch irgendwie so ein bißchen Logik mit drin, ja.

F: Mhm. Und gibt es nen Unterschied zu Verständnis von Stochastik?

A: Ja, auf jeden Fall. Also ähm Verständnis find ich deutlich schwieriger teilweise, also ich glaube, es gibt dann immer, oder was mir aufgefallen ist, es gibt immer so nen Klick-Moment, wo man vorher denkt man sich oft ähm das ist ziemlich schwierig, und dann gibt's irgendwann den Moment, wo mans verstanden hat und dann merkt man eigentlich, dass es gar nicht so schwierig war. Weil, oder es fällt einem dann vielleicht im Nachhinein denkt man sich, eigentlich wars ganz einfach.

F: Aha. Und kannst du irgendwie festmachen, wann du sagen würdest, man hat wirklich was verstanden in Stochastik?

A: Wenn mans selber erklären kann.

F: Wenn mans selber erklären kann, aha. Ok. Mh. Dann hab ich noch eine Aufgabe für dich.

A: Mhm.

F: Und zwar äh, die kam auch in den Fragebögen vor, ist das diese Krankenhausaufgabe.

A: Ah.

F: Ich hab sie mal ausgedruckt.

A: Ok.

F: Ähm, also man hat ein größeres und ein kleineres Krankenhaus und äh im größeren werden 45 Kinder im Schnitt am Tag geboren und im kleineren 15.

A: Mhm.

F: Und man geht jetzt einfach mal davon aus, dass für ein Mädchen 50% unf für einen Jungen 50% Wahrscheinlichkeit ist.

A: Mhm.

F: Und jetzt miss, äh zählt man für den Zeitraum von einem Jahr in beiden Krankenhäusern die Tage, an denen mehr als 60% der Babies Mädchen waren.

A: Mhm.

F: Und die Frage ist, in welchem erwartet man mehr.

A: Mhm.

F: Wenn du jetzt nicht die Antwort sagen sollst, sondern das Ganze simulieren solltest,

A: Mhm.

F: was würdest du für Schritte machen? Also es geht nicht um den Code, sondern sozusagen die Ideen, wie du so ne Simulation aufbauen würdest.

A: Ok. Also, Moment, kurz durchlesen.

F: Mhm. Lass dir ruhig Zeit.

A: […] Ok, also die Simulation soll jetzt enthalten ähm äh ja, die soll mir im Prinzip rausgeben, ähm, wieviele, sagen wir mal wir nehmen jetzt eben das Jahr, ähm an wievielen Tagen in dem Jahr eben in dem großen und in dem Kranken, in dem kleinen Krankenhaus äh 60% der Neugeborenen weiblich waren.

F: Ok.

A: So was, ähm, da könnte man jetzt Folgendes machen, also man müsste erstmal ne Zufallsvariable irgendwie definieren, oder ja, mit Trefferwahrscheinlichkeit 50% und ähm dann müsste man sagen, ok, ein Jahr hat 365 Tage ähm, jetzt muss ich mal überlegen, dann muss man äh, also ich würd dann also ne Matrix oder so nen Vektor nehmen mit 365 Einträgen und äh dann würde ich gucken, in dem großen haben wir 45 Kinder ähm dann muss man eben für jedes der Kinder also 45 Mal muss man dieses Zufallsexperiment durchlaufen lassen ähm Mädchen – Junge, und dann muss eben in dem Vektor eingespeichert werden, wieviele Mädchen das waren. Und dasselbe in nem zweiten Vektor dann für die, für das kleinere mit den 15 Kindern, und dann hat man da eben seinen Vektor mit ähm 365 Einträgen, wieviele Mädchen an jedem Tag geboren wurden, und dann müsste man halt ähm sagen, ok, man nimmt jetzt irgendne

Zählvariable, die durchzählt, wie oft ähm waren jetzt mehr als 60% Mädchen. Also, genau, Null Komma Sechs mal 45 und Null Komma Sechs mal 15. Genau, und dann hätte man das rausgezählt und dann ähm, gut dann hätte man das einmal gemacht. Und äh jetzt kann man natürlich sagen, äh ok, man hat dann ein Ergebnis, sagen wir mal, in dem größeren waren mehr. Um das jetzt zu überprüfen, kann man das eben für noch mehr machen, für noch ein Jahr und für noch ein Jahr. Und dann hätte man da, kann man ne Matrix draus machen oder so. Oder zwei Matrizen, genau.

F: Perfekt. Perfekt. Wunderbar, vielen Dank. Ähm, du bist übrigens der zweite, der darauf kommt, dass man das mehr als einmal machen sollte.

A: [lacht]

F: Alle andern haben gesagt, ich mach das einmal und dann [?]. Ähm, möchtest du sonst noch irgendwas anmerken zu den R Aufgaben, den Fragestunden oder was auch immer?

A: Also ich, ich sag mal so ähm, mir hat das Ganze sehr viel Spaß gemacht ähm also es hatte auch, für mich wars vielleicht äh auch ne besondere äh, also was besonders Tolles, weil ich hatte auch noch als Proseminar eben Den Zufall erleben, das hat sich ganz toll ergänzt irgendwie und ähm ja, also ich find, das kann man so weiter machen, das hat Spaß gemacht. Was wichtig sind, sind diese R Fragestunden, also die sollte man auf jeden Fall beibehalten. Weil die helfen dann doch auch viel, ja.

F: Ok, wunderbar. Ja, vielen Dank.

F: Gut, also die erste Frage ist ähm: Was war deine erste Empfindung, als du von den R Aufgaben gehört hast und hat sich an deiner Einstellung bis heute etwas geändert?

A: Hm, also R Aufgaben fand ich prinzipiell erstmal gut, weil ich Informatik noch studiere und programmieren sowieso schon mal gut finde und ich auch gemerkt hab letztes Semester in ECM äh hatten einige doch Probleme damit und kann man ja schon für Mathe auch gebrauchen.

F: Mhm.

A: Als ich R dann selbst gesehen hab, war ich erstmal von der Syntax überrascht.

F: Ja.

A: Und fands erstmal nicht so gut. Aber man gewöhnt sich halt an die Syntax und dann ist es auch ok.

F: Ok. Und wie häufig hast du die R Aufgaben bearbeitet?

A: Ich hab die immer gemacht.

F: Immer. Und in der Fragestunde warst du aber nicht?

A: Nein, war ich nie.

F: Mhm. Und warum nicht? Keine Zeit?

A: Keine Lust. [lacht]

F: Keine Lust, ok. [lacht] Also jetzt nicht Zeitgründe, sondern es ging auch ohne sozusagen.

A: Genau.

F: Ok. Mhm. Fein. Ähm fandest du denn die R Aufgaben im Vergleich zu den anderen Aufgaben eher schwer oder eher leicht?

A: Hm, ich wusste oft nicht genau, was gemeint war erst, weil wir haben ja den, ja im Grunde den kompletten Quelltext bekommen und ähm den musste ich halt erstmal verstehen, teilweise, und ähm das war dann halt ein bisschen schwer. Außerdem fand ich es war ein bisschen undurchsichtig teilweise, wie die Tutoren sich das angeguckt haben.

F: Ok.

A: Also eine Tutorin, mit der hab ich gesprochen, die hat gesagt, ja, sie guckt sich die Bilder an, wenn die ok aussehen, dann passt es, dann muss sie nicht groß den Quelltext angucken und ähm meine Tutorin hat auch ähm sich sehr stark an die Musterlösung gehalten ähm und hat uns auch gesagt, wir sollen bloß nicht den also man hätte ja auch den Quelltext selbst schreiben können, aber sie selbst konnte damit wohl nichts

anfangen und also man hat sich schon irgendwie an die Vorgaben halten müssen also an die, das hast du ja geschrieben.

F: Ja, genau. Aha. Also das heißt schwerer war eigentlich das Formale daran, also dass man sich an diesen Quelltext halten musste und jetzt gar nicht das Inhaltliche, sondern...

A: Ja, es war eigentlich, fand ich, immer gut erklärt mit Kommentaren und ähm klar gabs bei der einen oder anderen Aufgabe mal ne Schwierigkeit, aber im Allgemeinen fand ichs gut.

F: Ok. Ähm gut, jetzt hast du es schon so ein bisschen mitbeantwortet, aber vielleicht noch ein bisschen allgemeiner. Was fandest du an den R Aufgaben gut, also gibt's irgendwas, was du daran prinzipiell hilfreich fandest?

A: Hm […] teilweise hats nochmal ein bisschen wiederholt, was schon ein bisschen vorher war, also irgendeine der letzten Aufgaben war ja nochmal die Polya-Urne, die war ja schon zwei Monate oder ein Monat vorher irgendwie besprochen worden. Ähm, fand ich einerseits gut, andererseits fand ich halt auch, es hat nicht so ganz reingepasst, teilweise.

F: Aha, aha.

A: Ähm, ja ansonsten einfach, dass man halt selbst ein bisschen programmieren musste, fand ich gut, wobei ich mir gewünscht hätte, mehr halt selbst zu programmieren, weniger Vorgaben.

F: Mhm. An welche R Aufgaben erinnerst du dich denn noch?

A: Ähm, ja, zum einen die Polya-Urne, die mit ähm dem Dreieck auf der, äh mit dem Sphärendreieck, ähm […] sonst weiß ich gar nicht mehr so genau, also ich weiß teilweise, also zum Beispiel bei der ersten Aufgabe mussten wir irgendwie nicht viel machen, da mussten wir nur ne äh for-Schleife oder ne while-Schleife einbauen, also so ein bisschen, was wir machen mussten, [lacht], nicht mehr so genau, was es jetzt im Hintergrund hatte.

F: Ok. Aber Polya-Urne und Kugeldreieck ist doch schon mal gut. Ähm gibt es denn irgendetwas, was dir jetzt inhaltlich durch eine R Aufgabe klar geworden ist, was du vorher nicht verstanden hattest?

A: Hm, bei der […] ja genau, es gab noch die Aufgabe zu den Konfidenzintervallen ähm […] bei der einen Aufgabe musste man, ich weiß gar nicht genau, welche das war, was anwenden, was man in der dritten Aufgabe beweisen und in der vierten musste man das halt programmieren. Und der Zusammenhang wurde dann deutlicher irgendwie.

F: Ah, ok, das waren die Fehlstände, galub ich, was du meinst, ne?

A: Das kann sein. Ja, genau, ja.

F: Ah, ok. Und gibt es irgendwas, was man allgemein an diesen R Aufgaben deiner Meinung nach besonders gut lernen kann?

A: Hm, ja oft ist es halt schon so, dass man was ganz anderes denkt und dann äh macht man ne Simulation und ähm sieht, ok, es ist doch was ganz anderes. Ähm, ja, das find ich schon hilfreich.

F: Mhm. Ok. Was hat dir bei den R Aufgaben nicht gefallen? Also, du hattest schon ein bisschen was gesagt, aber gibt's sonst noch Punkte?

A: Hm […] ja eigentlich nur das, was ich gesagt hab, also das Formale irgendwie und ähm ich weiß, dass es ziemlich viel geworden wär, wenn wir R nochmal von Vornherein irgendwie komplett lernen hätten müssen, aber ich hätts besser gefunden, wenn man vielleicht auch kleinere Programme macht und äh die dann aber komplett selbst geschrieben sind. Und dass die halt auch im Tutorium ein bisschen [lacht] besprochen werden, also im Tutorium hieß es immer, ja, geht zur R Fragestunde ähm aber das ging für mich ja auch ohne.

F: Ja. Ja. Ok. Ähm würdest du denn aus Studentensicht jetzt sozusagen empfehlen, weiterhin R Aufgaben zu verwenden und würdest du etwas verändern. Gut, haben wir auch schon ein bisschen [lacht].

A: [lacht] Ähm, ja, ich fands auch für die Klausur sehr hilfreich. [lacht]

F: Ja, ok.

A: Also, es war halt ne Bonusaufgabe.

F: Ja, mhm.

A: Ähm und […] ja, ich könnt schon einige Studenten verstehen, die sagen, dass sie's nicht gerne machen ähm aber ich finds auf jeden Fall hilfreich, weil man dann halt auch ein bisschen, wenn man Stochastik-Vertiefung macht oder mehr in die Richtung, dann braucht mans ja öfter mal, dann hat man schon nen Einblick gehabt.

F: Mhm. Und würdest du irgendwie, also außer dass man das in die Tutorien einbeziehen sollte und eventuell darüber nachdenken sollte mit dem Quelltext, sonst noch was verändern?

A: Hm […] so spontan fällt mir da nichts ein.

F: Ok. Ähm wurde denn in deinem Tutorium manchmal von Studentenseite auf eine R Aufgabe Bezug genommen, als irgendeine andere Aufgabe an der Tafel erklärt wurde?

A: […] Ich glaube nicht. Also ich weiß die letzte oder eine der letzten, das war mit dem rot äh grün oder...

F: Mhm, diese bedingte Wahrscheinlichkeit.

A: Genau. Ähm die stand auf jeden Fall an der Tafel, die haben wir auch kurz besprochen […] aber es war nicht, dass wir irgendne andere Aufgabe besprochen haben und dann gesagt haben, hier, zu der R Aufgabe, oder da kann man was sehen oder so. Aber man konnte damit nochmal äh alles Formale sehen, wie man so ne bedingte Wahrscheinlichkeit ausrechnet.

F: Ja. Ok.

A: Das fällt mir auch grad ein, das fand ich auch gut, äh dass in den letzten R Aufgaben immer stand: Was ist der theoretische Hintergrund?

F: Ja. Ja.

A: Also, das war in den ersten ja irgendwie gar nichts so, da, das fand ich auch ja teilweise nicht so gut, dass es halt überhaupt keinen Bezug hatte.

F: Mhm.

A: Du hattest dann zwar alles erklärt, auch mit dem Kugeldreieck, aber es war halt einfach mal was komplett anderes [lacht].

F: Ja. Ja. Ok. Also das ist irgendwie schöner, wenn sich das wirklich so einbettet auch in den Rest, dass es jetzt nicht irgendwie...

A: Fand ich, ja.

F: Ok. Ähm gut, jetzt nochmal so ein bisschen was Inhaltliches: Woran denkst du als allererstes, wenn du an Stochastik denkst?

A: Dass ich es überhaupt nicht kann [lacht].

F: Ok. [lacht]

A: Also, das ist übertrieben, aber ich hab irgendwie nicht so ne gute Intuition, scheinbar, und ja, also, […] es ging [lacht], aber es ist nicht was, was ich irgendwie vertiefen möchte.

F: Ok. Und an was für einen mathematischen Begriff zum Beispiel denkst du oder an was für ein Objekt?

A: […] Hm, ja, Objekt, ja im Grunde das, was Stochastik äh schon in der Schule war, irgendwie Würfeln und Urnen und Kugeln ziehen und so.

F: Mhm.

A: Ähm und als Begriff […] Kombinatorik, wobei mir das auch überhaupt nicht liegt.

F: Ok.

A: [lacht]

F: Ok. Und kannst du erklären, was du darunter verstehst?

A: Unter Kombinatorik?

F: Ja.

A: Ähm, im Grunde ist das ja ein Zählen, also so […] ja irgendwie Methoden, um das zu vereinfachen, dass man eben nicht alles zählen muss, sondern Methoden dafür hat.

F: Mhm. Ok. Wunderbar. Jetzt hast du den Begriff schon erwähnt, äh, was verstehst du unter stochastischer Intuition? Also, du hast gesagt, dass du sie vielleicht eher nicht hast, aber was verstehst du drunter?

A: Ja im Grunde ist das, das richtige Modell zu finden für äh ja, irgendne Problemstellung.

F: Mhm. Und gibt es einen Unterschied zu Verständnis von Stochastik?

A: Ähm, ich glaube schon. Aber […] ich kanns nicht genau beschreiben. Also wenn man Stochastik versteht, dann findet man auch das richtige Modell, denk ich mal, aber ähm ich finde, man, oder ich denke, man kann auch das richtige Modell auch finden, ohne es hundertprozentig verstanden zu haben.

F: Aha. […] Und wann, würdest du sagen, hat man in Stochastik was richtig gut verstanden?

A: Ähm […] wenn man nicht mehr ähm […] allzu lange überlegen muss über irgendwas, also wenn man, also bei den Aufgaben auf den Übungsblättern zum Beispiel fand ich auch, hab ich die Lösung irgendwie gefunden, weil ich gedacht hab, das ist grad unser Thema, das muss was damit zu tun haben [lacht].

F: Mhm. Mhm. Mhm.

A: Ähm, aber […] ja Stochastik ist ja auch mehr als nur die Aufgaben, die wir gemacht haben.

F: Mhm. Mhm.

A: Ähm […] deswegen weiß ich jetzt nicht so genau, also ich fand auch zur Intuition in den Vorlesungen haben wir halt viel gesagt, ja das ist halt so und das passt schon irgendwie. [lacht]

F: Mhm.

A: Ja, das hat mir halt nicht so gut gefallen. Ja, deswegen weiß ich nicht, der ähm […] äh das Verständnis ist, denk ich, wenn man das wirklich äh alles auch korrekt beweisen kann. Und das, also ich weiß, dass ich das selbst nicht könnte und wir habens ja auch nicht so richtig gemacht in der Vorlesung. Ja, deswegen sag ich auch, dass mir da das Verständnis fehlt.

F: Mhm. Mhm. Ok. Interessant. Ähm, jetzt hab ich noch sozusagen eine Aufgabe. Und zwar ist das eine, die auch in den Fragebögen vorkam, ich weiß nicht, ob du die ausgefüllt hast.

A: Ja.

F: Ähm und zwar die Krankenhausaufgabe. Also man hat in einer Stadt zwei Krankenhäuser, ein großes und ein kleines. Und man sagt jetzt im großen werden jeden Tag im Schnitt 45 Kinder geboren am Tag und im kleinen im Schnitt 15. Und

wir gehen jetzt einfach davon aus, dass die Wahrscheinlichkeit für ein Mädchen 50% ist. Und ähm jetzt zählt man für den Zeitraum von einem Jahr in beiden Krankenhäusern die Tage, an denen mehr als 60% der Babies Mädchen waren. Und die Frage ist: In welchem Krankenhaus erwartet man mehr solcher Tage.

A: Im kleineren.

F: Ok. Sehr gut. Aber gesetzt den Fall du sollst jetzt nicht die richtige Lösung sagen [lacht], sondern das Ganze simulieren. Wie würdest du da vorgehen? Also nicht, kein Code oder so, sondern mich interessiert nur, was für Ideen du hast, welche Schritte du sozusagen angehen würdest. Also, wie gesagt, kann sehr allgemein sein, ohne Code. Kannst auch gerne nochmal in Ruhe hier durchlesen. [...]

A: Mhm. […] Ja, im Grunde braucht man ne Urne mit unendlichen Kugeln, gleich viel blau und weiß, wie auch immer, Männlein, Weiblein und ähm muss dann halt ziehen, wobei natürlich auch nicht jeden Tag genau 15 und 45 Kinder geboren werden...

F: Ok, das können wir einfach mal annehmen, also wir nehmen vereinfachend an nur 45 und 15.

A: Genau, und dann muss man halt ziehen, also wenn man annimmt, dass es genau 50% sind und ähm durch die Unendlichkeit der Anzahl der Kugeln äh gibt's da auch keine Veränderung, wenn man eine Kugel rauszieht, man kann natürlich auch einfach zwei Kugeln nur nehmen und immer zurücklegen, geht auch. […] Ja, und dann schaut man halt, was rauskommt.

F: Ok. Und also sozusagen, du ziehst jetzt erstmal wie oft?

A: Hm, wenn man davon ausgeht, dass im kleinen Krankenhaus immer 15 Kinder geboren werden, würd ich da 15 Mal ziehen.

F: Einmal, oder? Also sozusagen einmal 15 ziehen, oder?

A: Ja, wenn ich, wenn ich ne Urne hab mit nur 2 Kugeln, dann zieh ich eine, leg die Kugel wieder zurück und zieh dann so noch 14 Mal eine.

F: Ok.

A: Oder halt 45 Mal.

F: Ok. Gut, dann hast du sozusagen zwei Sets, einmal 15 Werte von blau und weiß und einmal 45 Werte von blau und weiß. Und dann?

A: Ähm, dann nehm ich an, dass es die Anzahl Mädchen und Jungen widerspiegelt. [lacht]

F: Mhm. Ok. Und weiter? [lacht] Ist ja schon mal sehr gut, ok.

A: Und dann schau ich ähm, ja schau ich mir die Tage an oder halt die Sets, also ich muss das ganze ja dann für ein Jahr lang machen, also dann halt 365 Mal.

F: Aha.

A: […] Und dann guck ich halt von den 365 Tagen die Sets an, bei denen mehr als 60% blau waren.

F: Ok. Und ähm um die Frage zu beantworten, schließt du dann was?

A: Ähm ja, dass in der Simulation, also dass das der Realität entspricht.

F: Ok.

A: Ähm, ja [lacht], ich weiß nicht genau, worauf du hinaus willst.

F: Also ich mein jetzt, du hast ja dann sozusagen 365 Mal, also ein Jahr, hast du dir angeguckt.

A: Genau.

F: Und zählst jetzt einfach ab, wo waren mehr, im größeren oder im kleineren.

A: Mhm.

F: Und dann nimmst du das Ergebnis und sagst, das ist dein, das ist jetzt sozusagen deine Prognose, so, das erwartest du auch.

A: Genau.

F: Ok. Alles klar.

A: Also, ich mein, das Ganze kann man dann noch öfter machen, also nicht nur ein Jahr simulieren, sondern..

F: Genau [lacht].

A: hundert Jahre oder so, aber ja.

F: Ok. Und so vom Bauchgefühl her, wie oft würdest du es wiederholen?

A: […] Das ist alles nicht so ganz aufwendig. Deswegen […] vielleicht 100000 Mal.

F: Ok.

A: Vielleicht auch ne Million, je nachdem. Also wenn ich bei den R Aufgaben gesehen hab, es ging in einer Sekunde oder weniger, dann hab ich halt noch ne Null hinten dran gehängt oder so, also.

F: Ok. Wunderbar. Also ich wollte auf das nochmal Wiederholen hinaus, zur Aufklärung. Ok, perfekt, eins a. Ähm, möchtest du denn sonst noch irgendwas anmerken? […] Irgendwas, was dir noch auf der Seele brennt?

A: […] Ja, also ich fand die Vorlesung nicht so gut.

F: Ok.

A: Ich war eigentlich immer da, aber ich hab nicht unbedingt immer aufgepasst. [lacht]

F: Ok. Und kannst du sagen, was dir nicht gefallen hat?

A: Ja zum einen, was ich vorhin schon angesprochen hatte, dass es halt öfter mal so war, ja, das ist halt irgendwie so und ich fand auch, man hat mit dem Buch nicht so wirklich viel anfangen können, weil ich find im Buch, was alles schriftlich ist, da kann mans auch richtig genau beweisen, also wenn man in der Vorlesung nicht die Zeit hat, ok, aber ich versteh's halt richtig, also ich versteh's halt am liebsten komplett richtig und das kann ich am besten, wenn ich's halt alles nachvollziehen kann durch richtige Beweise. Ähm […] die Simulationen sonst waren gut. Hm […] ja, hm […] die Klausur fand ich sehr einfach. Also, hab ich nichts dagegen, aber ich fand halt auch, also ich hab mir am Ende das dann angeguckt, es gab irgendwie von 82 Mitgeschriebenen 27 1.0, ähm, das fand ich schon ein bisschen [lacht], bisschen viel vielleicht, aber ich mein, das lag daran, dass er ja gesagt hat von Vornherein, dass man mit 90% 1.0 bekommt und es ja eigentlich nen doppelten Bonus gab. Ich möcht mich da auch nicht beschweren [lacht].

F: Ja.

A: Weil es hat mir ja auch geholfen, ähm, aber ich kann's jetzt auch einfach sagen, ich hatte in der Klausur 81 Punkte plus 10 hab ich dann 91, hab ich 1.0, aber ich fand die Note halt nicht gerechtfertigt.

F: Ok, aha.

A: Also, weil ich hab halt eigentlich 81 von 112,5 Punkten erreicht, das find ich schon nicht so viel [lacht].

F: Ok, mhm.

A: Ähm, ja […]

F: Ok.

A: Und ansonsten […] ja, teilweise konnt ich nicht so gut folgen, weil ich fand, dass wir auch immer mal hin und her gesprungen sind im Stoff. […] Aber sonst […] ne, ich weiß auch nicht, wir haben auch nicht so viel gemacht, also im Vergleich zu was Herr [xxx] gemacht hatte ähm ja, letztes Jahr oder so, also da hatte ich mit ein paar Kommilitonen geredet, ja, die hatten um einiges mehr gemacht. Ja, hm […], also ich fand mich teilweise ein bisschen wie ein Schüler, der halt irgendwas macht, was er nicht gerne macht, was er machen muss, und natürlich froh ist ne gute Note zu bekommen, aber ich fands nicht so befriedigend. Aber kann auch gut sein, dass ich da äh zu ner Minderheit gehöre.

F: Ne, ist auf jeden Fall gut, dass du das sagst, also, das ist interessant.

A: Ich hab dann auch [lacht], ich hab dann auch beschlossen, dass ich keine Stochastik-Vertiefung mache oder so was, ja.

F: Mhm. […] Ja, also vielen Dank. Das ist wirklich hilfreich.

F: Also, die erste Frage ist: Was war deine erste Empfindung, als du von den R Aufgaben gehört hast und hat sich an deiner Einstellung bis heute etwas geändert?

A: Was war meine erste Empfindung, naja, empfunden hab ich nichts, ich fands halt nur komisch, weil äh die letzten Semester in Stochastik wurde ja so was nicht angeboten, deswegen war ich halt ein bisschen verwundert, dass es jetzt auf einmal auftaucht, aber großartig gewundert hab ich mich eigentlich nicht.

F: Mhm. Ok, und ähm, ok. Also, du fandst es komisch und hat sich daran was geändert oder fandst du es auch noch am Schluss irgendwie ne seltsame Sache?

A: [lacht] Ja, die Sache ist, die R Aufgaben hab ich zum Teil gar nicht bearbeitet.

F: Ja.

A: Und äh beschäftigt hab ich mich jetzt auch nicht viel damit.

F: Ok.

A: Aber ich glaub, das war ja das, weshalb ich gekommen bin.

F: Ja, das ist gut.

A: [lacht] Ja, ok.

F: Also, das heißt, wie häufig hast du die R Aufgaben bearbeitet? Nie oder selten?

A: Ich glaub, nur einmal.

F: Mhm, einmal. War das die erste?

A: Ja, ich glaub das war das erste Blatt, ja.

F: Mhm. Und warst du irgendwann mal in der Fragestunde?

A: Einmal.

F: Einmal.

A: [lacht]

F: Und warum hast du dann die R Aufgaben nicht mehr bearbeitet?

A: Weil ich mir gedacht hab, dass, wenn ich die andern Aufgaben mache, das mir das reicht für die Klausur. [schmunzelt]

F: Ok. Und, also fandst dus irgendwie, also wenn du jetzt mehr Zeit gehabt hättest, oder so..

A: Ich hätts auf jeden Fall gemacht.

F: Also es war in erster Linie ein Zeitproblem?

A: Genau.

F: Ok.

A: Ich mein, ich werd das bestimmt mal brauchen, und dann weiß ichs nicht. Also wär auf jeden Fall sinnvoll, das zu lernen.

F: Mhm.

A: Schade, ganz ehrlich.

F: Also, das heißt, warum hast du die R Aufgaben nicht bearbeitet, ist schon beantwortet. Also in erster Linie weil du keine Zeit hattest.

A: Genau.

F: Und hätte man irgendetwas anbieten können, damit du sie bearbeitet hättest?

A: […] Ne.

F: Ok.

A: Also mir fällt da jetzt nichts ein, spontan.

F: Ok. Hattest du den Eindruck, dass du irgendwelche Nachteile dadurch hattest, dass du die R Aufgaben nicht bearbeitet hast?

A: Äh, ja, in der Klausur.

F: Ok. Und inwiefern in der Klausur?

A: […]

F: Also, weil du die R Aufgabe nicht bearbeiten konntest oder auch auf die anderen Aufgaben bezogen?

A: Ne, allgemein für die Stochastik-Klausur. Diesmal, ich hab so gut wie nichts gemacht.

F: Ok.

A: Weil ich mit den anderen Fächern beschäftigt war.

F: Ja.

A: Aber ich wollt das halt unbedingt, ich wollt unbedingt in die erste Klausur, damit ich das dieses Semester abschließe. Weil, falls dann irgendwie, falls ich in die Nachklausur gehe und dann durchfalle, muss ich das nächstes Jahr machen. Und ich will nächstes Semester fertig werden, deswegen.

F: Ja.

A: Damit ich halt, wenn alles schief laufen sollte, ich noch in die dritte Prüfung bei Herrn [xxx] reingehen könnte.

F: Ok, ja. Und wenn du jetzt sagst, dass du dadurch, dass du die R Aufgaben nicht gemacht hast, dass du da Nachteile hattest in der Klausur: Woran hast du das gemerkt?

A: Nachdem ich mein Klausurergebnis hatte [lacht].

F: Ok.

A: Aber ich bin zufrieden damit, also. Ich bin durchgefallen, aber ich hab auch einen Tag davor angefangen, mir das so richtig anzugucken. War blöd von mir, aber wie gesagt, in der Nachklausur mach ichs besser.

F: Aber das heißt, es hat jetzt nicht nur daran gelegen, dass du die R Aufgaben nicht gemacht hast, sondern generell.

A: Naja, in dem Sinne vielleicht, hätt ich die Aufgabe bearbeitet, hätt ich vielleicht bestanden, weil mir haben nur 3 Punkte oder so gefehlt.

F: Ah, ok. Findest du es denn generell eine gute Idee in Stochastik so Simulationsaufgaben zu stellen?

A: […] Ja.

F: Und gibt es irgendwas, was man daran deiner Meinung nach besonders gut lernen kann?

A: […] Was man daran besonders gut lernen kann. […] Ja, vielleicht Verknüpfung mit anderen Fächern.

F: Ok. Fällt dir irgendein Beispiel ein?

A: Ja, so allgemeine Simulationsverfahren, keine Ahnung, Monte Carlo Verfahren, gibt's genug.

F: Mhm. Ok. Wurde denn in deinem Tutorium manchmal auf R Aufgaben Bezug genommen? Ich weiß nicht, warst du im Tutorium?

A: Ich war ab und zu im Tutorium, ja.

F: Und wurde da irgendwie über die R Aufgaben gesprochen?

A: Ähm, also, als ich da war nicht wirklich.

F: Mhm. Ok. Jetzt nochmal so ein bißchen was Inhaltliches zu Stochastik: Woran denkst du als erstes, wenn du an Stochastik denkst?

A: […] Wenn ich an Stochastik denke […] Prozentrechnung [lacht].

F: Prozentrechnung, ok.

A: Das war halt immer meine Vorstellung. [lacht]

F: Ok, ok. Und an was für einen Begriff oder was für ein Objekt denkst du?

A: Ganz banal?

F: Kann ganz banal sein.

A: Dreisatz.

F: Dreisatz, ok. Das ist interessant. Noch irgendwas?

A: Ähm, ansonsten, […] fallen mir halt die Inhalte aus der Vorlesung ein.

F: Ja, sag ruhig.

A: Keine Ahnung, Markovketten.

F: Und kannst du erklären, was eine Markovkette ist?

A: Leider nicht. [lacht]

F: [lacht] Kein Problem. Ist sehr interessant, die Wörter sind bisher nicht gefallen. Ähm, was verstehtst du unter stochastischer Intuition? Das ist so ein Begriff, der in den Fragebögen vorkam. Ich weiß nicht, ob du die ausgefüllt hast, wenn du nicht so oft im Tutorium warst.

A: Nein.

F: Ah, ok. Das war einer am Anfang des Semesters und einer am Ende. Aber unabhängig davon, wenn du das Wort stochastische Intuition hörst, an was denkst du da?

A: Puh. Überdimensionales Denken. […]

F: Und gibt es einen Unterschied zu Verständnis von Stochastik? Also zwischen Intuition und Verständnis?

A: […] Ja, wenn mans versteht, dann […] Ich weiß es nicht. Das ist halt, viel kann ich da jetzt nicht sagen, weil Stochastik ist ein Fach, was mir so nicht wirklich liegt.

F: Ok. Ja, macht ja nichts. Aber wann würdest du sagen, hätte man in Stochastik etwas richtig gut verstanden?

A: […] In erster Linie, wenn man sich damit beschäftigt hat.

F: Ok. […] Und wenn du jetzt zum Beispiel diese Markovketten nimmst, wann würdest du sagen, hast du verstanden, was ne Markovkette ist?

A: Wenn ich mir das im Skript durchlese, wenn ich die Übung dazu mache, dann hab ich das verstanden.

F: Ok, gut. Möchtest du noch irgendwas anmerken?

123 A: Nein. [lacht]

124 F: Ok. Dann wars das schon, vielen Dank.

F: Gut, also die erste Frage ist: Was war deine erste Empfindung, als du von den R Aufgaben gehört hast, und hat sich an deiner Einstellung bis heute etwas geändert?

A: Naja, am Anfang, ich bin nicht so wirklich computerverbunden, das war schon beim Herrn [xxx] so, dass solche Aufgaben nicht so toll waren, aber an sich, also, ich fands interessant, so was auszuwerten.

F: Ok.

A: Also, jetzt abgesehen davon, dass man diese R Stunden dann hatte, wo man dann zusammen im Prinzip den fertigen Code serviert bekommt, aber dann die Auswertungen an sich musste man dann trotzdem selber machen.

F: Mhm.

A: Und ich fand das eigentlich, so Statistiken fand ich schon interessant.

F: Ok. Ähm, wie häufig hast du denn die R Aufgaben bearbeitet?

A: Also, so genau weiß ichs nicht mehr, aber ich würde sagen maximal 50% der ganzen Aufgaben.

F: Ok. Und wie häufig warst du in der Fragestunde?

A: Selten.

F: Selten.

A: Ja, schreib mal selten hin.

F: Gut. Und warum hast du die Fragestunde besucht oder nicht besucht?

A: Also, besucht, kann ich sagen, ähm, am Anfang wusst ich nicht so genau, was ich davon halten soll, man ist da so relativ kritisch, was man so da macht und dann über einen Freund bin ich dann da drauf gekommen.

F: Ok.

A: Dass ich da jetzt auch mal hingehen soll und das hat mir eigentlich ganz gut gefallen, nur es hat in das Gesamtzeitbild nicht so immer reingepasst, also, deshalb war ich auch häufig nicht anwesend.

F: Ok.

A: Ähm, ja. An sich, ich fands ne gute Sache.

F: Ok. Ähm, fandest du denn die R Aufgaben im Vergleich zu den anderen Aufgaben auf dem Übungsblatt eher schwer oder eher leicht und was war schwerer und was leichter?

A: Also, teils, teils.

F: Mhm.

A: Also manche Aufgaben waren relativ einfach, da konnte man ohne großes Know-how oder was auch immer die bearbeiten bzw. die zumindest mal selber für sich durchdenken, ohne irgendwas mit dem Code zu tun, zu arbeiten. Ähm, manche Aufgaben waren schon heftig, muss ich selber zugestehen, also ich weiß gar nicht welche das waren, also so hinten raus gabs ein paar Aufgaben, die waren ziemlich heftig.

F: Ok.

A: So was mit Matrizen zum Beispiel. Das war, da wo man in eine Matrix alles einspeichern musste, das waren solche Aufgaben, die sind schon, also, keine Ahnung, bestimmt hat ein Großteil das so lösen können, aber ich würde behaupten, der größere Teil der Studenten, die konnten es nicht.

F: Ok. Und die waren auch schwerer als die anderen Aufgaben auf dem Blatt?

A: Also im Verhältnis manchmal ja. Also, so auf den letzten Blättern waren ein paar Aufgaben schon dabei, wo ich dachte: Nein, das weiß ich nicht, wie das geht.

F: Ok. Ähm, was fandest du an den R Aufgaben gut und an welche erinnerst du dich noch? […] Also Matrizen sind schon gefallen, ich glaube, du meinst diese Krankenhausaufgaben.

A: Genau. Die waren sicher ganz interessant und cool, aber da auf diese Matrix zu kommen, dass man das da irgendwie überall reinspeichern soll und R arbeitet dann mit diesen speziellen Werten dann weiter, das war dann, naja, von der Umsetzungen, von der Denkweise ganz klar, aber von der Umsetzung her nicht so klar, also.

F: Mhm.

A: Ähm, an was für Aufgaben erinnere ich mich denn noch? […] Ja, es gab schon ein paar, das ist schwierig. Also, ich erinnere mich mehr an Aufgaben vom Seminar als […] als Aufgaben von, aber ich hab jetzt schon Ewigkeiten nichts mehr für Stochastik gemacht. Also, was heißt Ewigkeiten, seitdem das Semester zu Ende ist. Aber so, wenn ich jetzt Statistiken sehen würde von diesen Aufgaben, würde ich ganz genau wissen, was wir da jetzt gemacht haben. Also so, sie waren ok. Was war nochmal der zweite Teil?

F: Was fandest du an den R Aufgaben gut?

A: Ähm. […] Ja, die Visualisierung einfach an sich. Ähm, es gab nämlich so Aufgaben, die waren von der Denkweise her relativ schwierig zu verstehen, aber dann, wenn man das Bild vor Augen hatte, dann wusste man eigentlich ganz genau, ganz klar, was damit gemeint ist.

F: Mhm.

A: Mit Bildern funktionierts einfach besser.

F: Ja, ok. Und gibt es irgendwas, was dir durch eine R Aufgabe mal klar geworden ist, was du vorher nicht verstanden hattest?

A: Wüsst ich jetzt spontan nichts.

F: Ok. Und gibt es irgendwas, was man deiner Meinung nach besonders gut lernen kann mit diesen R Aufgaben?

A: Naja, also wenn ich jetzt auch das Seminar dazu zähle, dann muss ich sagen, da sind so ein paar Dinge in der Wahrscheinlichkeitsrechnung, in diesen Theorien, die sind für den ein oder anderen selber klar, aber da kommt dann ein ganz anderes Bild raus, eine ganz andere Statistik wie man sich das eigentlich überlegt hat.

F: Mhm.

A: Also das ist nicht immer so einfach, von daher find ich das eigentlich, kann man mit so etwas ganz gut überprüfen, ob diverse Dinge, wo man sich eigentlich sicher ist, das muss man dazu sagen, ich bin nämlich auch schon in ein paar Fallen reingetappt, wo ich dachte, die Verteilung wär ganz anders als sie es in Wirklichkeit ist.

F: Mhm. Also, dass man irgendwie nachprüfen kann, ob das, was man sich überlegt hat, auch stimmt.

A: Ja, zum Beispiel.

F: Mhm. Gibts sonst noch irgendwas, was dir einfällt?

A: Wenn man so lange Zeitintervalle hat mit vielen verschiedenen Daten, also ich für mich persönlich fände es schwierig, sie per Hand auszuwerten.

F: Mhm.

A: Ja, mit einer Software ist natürlich alles viel einfacher.

F: Mhm. Ok. Ähm, was hat dir bei den R Aufgaben nicht gefallen?

A: Naja, jetzt, gute Frage. Ich hab sie bis zur Hälfte gemacht, äh. Also, jetzt abgesehen davon, dass es ja schon zum Teil so war, dass sie, das sind geschenkte Punkte gewesen, ähm, fand ichs an manchen Stellen vielleicht der […] ist schwierig, das ist schwierig, die Frage, was einem nicht gefallen hat. Also, ich würde jetzt sagen bei manchen das Programmieren macht manchen keinen Spaß, also mir machts eigentlich auch keinen Spaß, aber ähm es hängt natürlich vom Kontext ab, also manche Aufgaben, die mich dann interessiert hatten, was kommt da raus, oder wie sieht diese Statistik aus, da sitzt man sich ganz anders dahinter und macht sie als bei so langweiligen Themen, aber da fallen mir jetzt keine spontan ein.

F: Mhm. Und du hast am Anfang gesagt, abgesehen davon, dass es geschenkte Punkte waren, wie meinst du das?

A: Also […]

F: Ganz ehrlich.

A: Also, wenn man ehrlich ist, muss ich sagen, also die Fragestunde war gut, also die, die da drin saßen, denen hats gut getan, ähm, nun ist es halt so, es ist eigentlich wie in der Schule, aus einem Pool werden auf einmal 50 Pools, also es ist dann so kurz vor knapp, so, ist schade, aber manche, im Prinzip kriegen dann die ein oder anderen es einfach so rüber, ich mein, sie werten das dann aus und raten manchmal einfach, was da eigentlich rauskommen sollte, und das ist eigentlich schade. Also, ich hätt, ja, aber, an diesem Punkt kann man eigentlich so gut nichts ändern [lacht] und es wird immer so bleiben. Also, die, die es interessiert, die machens eben und beschäftigen sich damit und die, die es eben nicht interessiert, na gut.

F: Mhm.

A: Denen wird's halt vielleicht später dann schwerer fallen. Ich weiß nicht.

F: Mhm. Aber das war bei den R Aufgaben, also es geht ja darum, dass sie einfach den Code kopiert haben, zum Beispiel.

A: Ja, zum Beispiel.

F: Ähm, das war bei den R Aufgaben stärker als bei den anderen Aufgaben? Weil ich mein, abschreiben sonst...

A: Also, ich weiß hier [Lernzentrum], also ich hab meine Aufgaben in der Regel immer hier gemacht, und, klar, wenn man dann zu fünft oder sechst am Tisch sitzt, ähm, und dann die Aufgaben macht, dann macht mans meistens so, dass die einen die erste Aufgabe machen, die andern die zweite, ähm, also ich weiß nicht, NN, sagt dir vielleicht noch was, also auf alle Fälle, der war sehr gut, und wir haben dann immer so, der hats dann, er hats dann zumindest als einer der wenigen auch erklärt, warum das gerade so ist, das ist dann der Sinn und Zweck der Sache eigentlich, dann kann man von mir aus diese Aufgabe zu 80% oder was zu übernehmen, wenn mans dann verstanden hat, aber so einfach abschreiben, das ist dann halt wie in der Schule.

F: Mhm.

A: Ja, ohne Wissen schreibt mans hin, ähm, gut. Wenn mans dann nicht vorrechnen muss, gut, wenn mans vorrechnen sollte, was manchmal passiert ist, sieht man ganz genau, ob ers kann oder nicht.

F: Aber es war jetzt bei den R Aufgaben irgendwie häufiger so, dass man einfach …

A: Also, es ging einfach am schnellsten, muss man sagen.

F: Aha.

A: Es war klick, kopiert man hier oder kopiert man da.

F: Ok.

A: Meistens per mail oder sonstiges und dann ist es erledigt.

F: Mhm. Ok.

A: Zwangsweise ist es so.

F: Mhm. Würdest du denn aus Studentensicht empfehlen, weiterhin R Aufgaben zu verwenden und würdest du etwas verändern?

A: Also, ich wär voll und ganz damit einverstanden, dass man es weiter verwendet. Jetzt, mal abgesehen davon, dass auch diverse Seminare damit zu tun haben, ähm, fänd ichs eigentlich schade wenn mans nicht schon von Anfang an macht.

F: Mhm.

A: Das ist im Prinzip so wie mit Englisch-Lernen. Desto später man damit anfängt, desto schwieriger wird es. An sich, ob man das vielleicht, […], ne, da fällt mir nichts ein.

F: Also nichts verändern, einfach so.

A: Ne.

F: Und die Fragestunde auch beibehalten?

A: Die Fragestunde beibehalten.

F: Mhm. Aber vielleicht in ner anderen Form?

A: Also, […] ich fand eigentlich dieses zusammen erarbeiten eigentlich ganz gut, also ich mein, man hats zusammen erarbeitet und konnt am Ende es im Prinzip, also der- oder diejenige, die es nicht so ganz verstanden hat, konnt es dann für sich selber nochmal reflektieren, ähm, fand ich eigentlich ganz gut. Also ich fände, ich wüsste da jetzt nichts, was man verändern könnte, damit es besser wird oder ja, mehr Leute kommen. Es ist halt natürlich, terminlich gesehen, ist es natürlich immer schwierig. Also, das muss man dazu sagen, weil, man hats ja für die Klausur gesehen, das war natürlich nicht so toll, Herr [xxx] hätte von Anfang an einen festen Termin machen sollen. Den einen hats so nicht gepasst, den andern hats so nicht gepasst, äh, gut, Samstag war jetzt nicht schlimm, für die meisten nicht schlimm, aber nächstes Mal, manche schreiben sicherlich zwei Klausuren an einem Tag, ich mein, darauf kommts jetzt nicht drauf an.

F: Ja, ok. Ähm, wurde in deinem Tutorium manchmal auf R Aufgaben Bezug genommen, wenn eine andere Aufgabe an der Tafel erklärt wurde? Gabs so was mal?

A: […] Da muss ich ganz ehrlich sagen, ich weiß es nicht mehr.

F: Mhm. Ok. Jetzt nochmal etwas Inhaltliches. Ähm, woran denkst du als erstes, wenn du an Stochastik denkst?

A: Wahrscheinlichkeitsrechnung, ähm, Statistiken, würde ich jetzt so in Anführungsstriche setzen, ähm, ja eigentlich alles, was in dieses Gebiet reinfällt, ich mein, dieses Gebiet ist riesig, aber so spontan jetzt […]

F: Vielleicht irgendein Begriff?

A: Ja, also diese ganzen Approximationen, die man so kennt, also die man jetzt kennt,

muss ich schon dazu sagen.

F: Mhm.

A: Ähm.

F: Zum Beispiel?

A: Stirling oder, also, ich fand interessant einfach zu sehen, dass nicht jede Aufgabe oder jede Thematik mit EINER Wahrscheinlichkeit abgearbeitet wird, so im Sinne von es gibt nicht nur DIE Wahrscheinlichkeit, sondern man hat verschiedene Modelle einfach. Ähm, dass es doch so komplex ist.

F: Mhm. Ok.

A: Ähm, das weiß man jetzt. Davor wär wirklich nur Wahrscheinlichkeit, also ich hatte es auch in der Schule nur, ich hatte es im Abitur zum Beispiel gar nicht, deshalb kann ich da nicht so viel, also es ist schon viel länger her als bei manch anderen, sagen wir es mal so.

F: Mhm. Und ähm wenn du nochmal an so ein zentrales Objekt oder so was denkst, was in der Vorlesung vorkam?

A: Polya-Urne.

F: Die Polya-Urne, ok. Und kannst du erklären, was das ist?

A: Oh, ich versuchs, ich versuchs. Also, […] das ist eines dieser berühmten Modelle zum Berechnen von Wahrscheinlichkeiten. Mh, ich hätt mir jetzt ein einfacheres Beispiel aussuchen sollen [lacht]. Ähm, ist eigentlich peinlich, wenn ichs jetzt nicht wüsste, weil ich habs behandelt, ähm, also ich weiß zumindest, wie die funktioniert.

F: Mhm. Erklär mal.

A: Wenn man, also man hat ja eine gewisse Anzahl von Kugeln in dieser Polya-Urne, die man von vornherein festlegt. Äh und ähm dann zieht man immer eine aus dieser Urne und je nachdem, was für eine Farbe es ist, legt man eine zu, eine Kugel, Extra-Kugel der gleichen Farbe wieder dazu und so äh ja, so kann man halt feststellen, ähm die Wahrscheinlichkeit, wie oft welche Kugel gezogen wird, würd ich jetzt sagen.

F: Ok.

A: Ähm, aber wofür mans jetzt verwendet, wir hatten da so ein paar schöne Bildchen, aber […] ich hätte einfach was anderes nehmen sollen. [lacht]

F: Kein Problem. Fällt dir denn noch ein anderer Begriff ein?

A: Mir fällt noch ein […] ähm die ganz einfache Verteilung, warte, wie heißt sie denn [...] Binomialverteilung.

F: Mhm.

A: Das ist zum Beispiel was einfacheres, muss ich ehrlich sagen.

F: Mhm.

A: Also so was wie Münzwurf oder ja. Das ist, aber das hat dann auch viel, hat dann nen größeren Bezug zur Realität, muss ich sagen. Also, man kann eigentlich damit ganz gut erklären, warum, wenn man ne Münze wirft, wie die Wahrscheinlichkeit liegt, dass Kopf oder Zahl geworfen wird. Also ich hab das jetzt zum Beispiel, vor allem wenn man die zinkt, dann ist das ganz interessant.

F: Mhm. Und bei der Binomialverteilung, kannst du erklären, was das für eine Verteilung ist?

A: […]

F: Also, was man damit modelliert?

A: Mhm, naja, also die hat ja, die Binomialverteilung hat ja nur den Ausgang entweder, also wenn man das jetzt auf den Münzwurf bezieht, ähm, entweder Richtig oder Falsch. Und ähm, im Prinzip, wenns fair ist, 50-50 Chance, ähm, ja, bei der Münze, bei der Münze ist es halt Kopf oder Zahl.

F: Mhm.

A: Aber außer solchen Dingen, Ereignissen, ob Erfolg oder Misserfolg eintritt, kann man eigentlich nichts damit modellieren, also nur solche Sachen.

F: Ok.

A: Also, mir fällt jetzt grad nichts ein, was man damit noch modellieren könnte.

F: Und, also, weißt du noch was, also sozusagen, ähm, also wenn du sagst, Kopf oder Zahl oder Richtig oder Falsch, also was sind die Werte, die so ne binomialverteilte Zufallsvariable annehmen kann?

A: […] Ja, also […] ja., Null oder Eins, kann sie annehmen.

F: Ok.

A: Oh, beziehungsweise, das hängt davon ab, wie man sich das wählt, also, es kann auch, glaub ich, -1 und 1 sein. Aber 0 oder 1 finde ich sinnvoller.

F: Ok. Ähm, dann jetzt noch ein Begriff, der auch in den Fragebögen vorkam. Was verstehst du unter stochastischer Intuition?

A: Ja, das ist […] ähm, also jetzt allein von dieser Wortkombination her würd ich sagen, jemand, der sich damit zum Teil beschäftigt hat, vor allem mit diesen Themen, und dann einfach ohne irgendwelche Nachrechnungen oder ähm Aufschriebe ungefähr sagen kann, was für ein Ergebnis rauskommt.

F: Mhm.

A: Oder was für eine Wahrscheinlichkeit dieses Modell oder dieser Versuch hat. Würd ich jetzt sagen, das ist stochastische Intuition.

F: Mhm. Und gibt es einen Unterschied zu Verständnis von Stochastik?

A: Ja, ich denk schon. Also, ich find Verständnis an sich ist, glaub ich, das grundlegende Prinzip und das muss man erstmal hinkriegen, um überhaupt irgendwas oberhalb von diesem Begriff äh zu machen, weil wenn ich das Verständnis nicht habe von den Verteilungen, von wie, was passiert wann, wenn, ich wie was mache, äh, dann, ja ich find das elementar, also es ist schon ein Unterschied.

F: Mhm. Und wann würdest du sagen, hat man wirklich etwas verstanden in Stochastik?

A: [lacht] Ähm, […] ich persönlich würde sagen, man hats eigentlich, also bei den einfachen Dingen ist es relativ schnell, hat mans verstanden, aber bei so komplexeren würd ich sagen vor allem, wenn man das dann ja auf die Allgemeinheit bezieht, hat mans eigentlich nie ganz verstanden.

F: Mhm.

A: Also, es gibt bestimmt immer irgendwo so einen kleinen Kniff, der anders ist, ich finde, das ist so ne offene Frage.

F: Ok.

A: Ich könnte mir durchaus vorstellen, dass so diverse Themen da rein fallen, wo es eigentlich nicht so ganz klar ist, was immer das Ergebnis ist oder der Ausgang. Obwohl es immer der gleiche Startpunkt ist.

F: Ok. Ähm, jetzt noch eine kleine Aufgabe, dann sind wir fast fertig. Und zwar gabs in den Fragebögen diese Krankenhausaufgabe, ich hab sie hier nochmal ausgedruckt.

A: Ach ja.

F: Also, man hat ein großes und ein kleines Krankenhaus, im kleinen werden jeden Tag 15 Kinder geboren im Schnitt und im großen ungefähr 45. Und man sagt, die Wahrscheinlichkeit für Mädchen und Junge ist gleich, 50%. Und man zählt jetzt für den Zeitraum von einem Jahr in beiden Krankenhäusern die Tage, an denen mehr als 60% der Neugeborenen Mädchen waren. Die Frage ist, in welchem Krankenhaus erwartet man mehr solcher Tage. […] Ich möchte jetzt nicht die Antwort wissen, sondern möcht gern wissen, wenn du das jetzt simulieren müsstest, wie würdest du da vorgehen? Also, ohne Code, nur die Ideen. Was müsstest du simulieren?

A: […] Also, ok, im Prinzip müsste man erstmal die zwei Krankenhäuser mal getrennt betrachten.

F: Mhm.

A: Ähm, ja gut, mit den Werten, die da jetzt angegeben sind.

F: Mhm.

A: Ähm, gut, und dann […] Hm, also, eigentlich müsste dann das Programm das, man müsste dann irgendwie so eine Schleife basteln, äh, die dann, die dann das für diesen Zeitraum von einem Jahr oder was auch immer, ich glaub es ist ein Jahr...

F: Mhm, ein Jahr.

A: das simuliert und dann müsste man das irgendwie so machen, dass das Programm sich ähm […] diesen Wert, den exakten, ne nicht den exakten, also den Wert rauspickt, ähm ab dem Zeitpunkt, wo mehr als 60% weiblich sind.

F: Also für jeden Tag, ne, man guckt ja immer..

A: Äh, ja, ja. Man muss jeden Tag sich schon ansehen und dann muss man sich den Wert rauspicken, ab dem zum ersten Mal die 60% erreicht sind. Find ich, ja, aber […] Das wüsst ich wirklich nicht, wie man das schreibt [schmunzelt].

F: Ähm, vielleicht nochmal einen Schritt zurück. Der Anfang war gut, du hast zwei Krankenhäuser, und ähm machst quasi für ein Jahr immer in dem einen 15 Kinder und in dem anderen 45 und zählst, jetzt hab ichs verraten, naja gut, und überlegst, wie kannst herausfinden, wie oft mehr als 60% der Kinder weiblich waren.

A: Ja, genau.

F: Jetzt hast du gesagt, „ab welchem Zeitpunkt die 60% erreicht sind".

A: Achso, ja, jetzt machts auch Sinn. Ich mein, das werden ja pro Tag 45 Kinder oder jetzt n Kinder, wenn ich es aufs Allgemeine beziehen würde, geboren und von diesen n Kindern müsste man halt erstmal sehen, wieviel sind davon weiblich und wieviel männlich und ja, klar, dann müsste man im Prinzip jeden Tag dieses, den weiblichen Teil zählen …

F: Ja.

A: Ja, und wenn er dann diese 60% erreicht, was auch immer 60% sind, ich weiß halt nicht, was der Gesamtwert ist […], dann wär, dann müsste man sich diesen Tag merken, irgendwie. […] Ja, so was mein ich, das ist so ne Aufgabe [lacht].

F: Ich bin nicht sicher, ob du die Aufgabe richtig verstanden hast, weil du es, glaub ich, komplizierter machst, als es ist.

A: Wahrscheinlich.

F: Also du hast die beiden Krankenhäuser und du willst pro Krankenhaus die Anzahl der Tage zählen, an denen mehr als 60% der Kinder weiblich waren.

A: Ja, genau.

F: Das heißt, du hast eigentlich ja die Grundgesamtheit, oder?

A: Achso, ja. Ups, gut. Ja, ok, dann, ja, ok, stimmt. Dann zählt von den 45 Kindern in dem großen Krankenhaus muss man einfach sehen, wann, ja, VON diesen 45, wenn 60% davon weiblich sind.

F: Genau, genau.

A: Ja, jetzt hab ichs auch.

F: Genau, und dann würdest du, also machst du das einmal für das kleine und einmal für das große Krankenhaus, sagst du.

A: Also ich würde es getrennt machen, aber man kanns bestimmt auch zusammen machen.

F: Nö, machs ruhig getrennt. Ok.

A: Ja gut, also bei dem großen ist es dann so, wenn dann 60% von diesen 45 weiblich waren, dann, ja, dann soll er sich diesen Tag merken ja, und das natürlich auf das Jahr bezogen. Ähm, ja, dann gibt's irgendeine Zahl, wie viele von diesen Tagen eintreffen.

F: Mhm.

A: Und das auch bei dem kleinen Krankenhaus, funktioniert genauso.

F: Mhm.

A: Ja, und dann würd man ja am Ende sehen, welche dieser Zahlen größer ist.

F: Mhm.

A: Ja, und je nachdem, ja entweder im größeren oder im kleineren werden da mehr weibliche Neugeborene erwartet, ich würd sagen, ja, das kann man so machen.

F: Ok. Das heißt, du lässt es sozusagen für ein Jahr durchlaufen und zählst einfach, wieviele Tage waren da und wieviele waren da...

A: Genau.

F: und den größeren Wert nimmst du.

A: […] Eigentlich, eigentlich schon, ja.

F: Mhm.

A: Aber, das kann ja aber auch, wenn man das dann nochmal laufen lässt, ganz anders rauskommen.

F: Aha. Das heißt, was könnte man machen, um sich sicherer zu sein?

A: Naja, man könnte das vielleicht von einem Jahr auf, ja, auf alle Fälle diese Zahl extrem erhöhen.

F: Mhm, mhm.

A: Aber […]

F: Ist doch gut.

A: Aber ich würd die Zahl erhöhen, ich weiß jetzt nicht, auf wieviel, aber, also was Sinn macht, aber ich denke, ein Jahr ist zu klein. Also, ich würd sagen, mindestens würd ich mal 10 nehmen, ich würd fast 100 nehmen. 100 Jahre.

F: Mhm.

A: Auch wenn das so ein bißchen, ja, ein bißchen an der Realität abweicht, aber man kanns ja dokumentieren heutzutage.

F: Mhm, wunderbar. So kann mans machen.

A: Aber ne Antwort geb ich trotzdem nicht, weil ich bin mir da immer noch ziemlich unsicher.

F: [lacht] Doch, gib mal eine. Nicht fürs Protokoll, nur so.

A: Also, rein von der Logik her, müsste, also die meisten würden jetzt sagen, im größeren, aber wenn man das ausrechnet, würd ich sagen, also wie vorher gesagt, ist es eigentlich gar nicht so eindeutig zu sagen, ob es im größeren oder im kleineren mehr kommt, weil es kann ja sein, […] also wir können ja mal so sagen, wenn von 45 Neugeborenen 5 nur weiblich sind und 40 männlich und in dem anderen sind, müssen ja nur mindestens 60 weiblich sein, also es können auch weitaus mehr weiblich sein, 15 weiblich, also wenn man das hochspielt auf ein paar Jahre, dann sind im kleineren definitiv mehr weibliche. Können mehr weibliche Neugeborene sein als im anderen. Also ich würde sagen, gabs da nicht drei Antwortmöglichkeiten?

F: Ja, also entweder im größeren oder im kleineren oder in beiden gleich viele.

A: Ich glaub, ich hab damals angekreuzt, in beiden gleich viele [lacht].

F: Ja, also es stimmt nicht, es ist tatsächlich im kleineren. Und zwar im kleineren deshalb, weil wenn man, also im Grunde ist es ja die Anzahl der Mädchen geteilt durch die Gesamtanzahl der Kinder an dem Tag, das ist ja eigentlich ein Schätzer für die Wahrscheinlichkeit, dass man ein Mädchen bekommt.

A: Ja.

F: Und äh, das heißt, man erwartet eigentlich 50%.

A: Ja, man erwartet, ja.

F: So, und wenn man ne kleine Stichprobe hat, also man hat ja einmal ne kleine Stichprobe, aus der man das schätzt, nämlich nur 15 Kinder am Tag, und einmal ne große mit 45 am Tag. Und wo wird dann eine Schätzung genauer, wenn du mehr oder wenn du weniger nimmst?

A: Ja, wenn man weniger nimmt.

F: Wird dann die Schätzung genauer?

A: Äh, ne, wenn man mehr nimmt, wird sie genauer. Ja, ich glaub, ich hab damals das mittlere angekreuzt, weil ich wirklich nicht genau sagen konnte, obs im großen oder im kleinen jetzt war. Also, die meisten würden tatsächlich das größere ankreuzen, würd ich jetzt behaupten.

F: Die meisten haben angekreuzt, es macht keinen Unterschied.

A: Oder keinen Unterschied. Also ich hab keinen Unterschied angekreuzt. Ja gut, aber das ist dann.

F: Aber wenn mans simuliert, sieht mans sogar sehr schön.

A: Ja.

F: Vielleicht sollte mans noch reinstellen.

A: [lacht] Ja, so aus dem Stehgreif ist es schwierig.

F: Ne, ne, aber es ist auch ne Frage, die erforscht ist, dass die sehr viele Schwierigkeiten macht. Deswegen war die auch da drin als schwierige Aufgabe. Ok. Möchtest du noch irgendwas anmerken?

A: […] Also die Vorlesung war nicht geordnet, also man muss sich ja nicht unbedingt nach dem Buch halten, also zumindest Herr [xxx] so ein kleiner Leitfaden in der Vorlesung wär cool gewesen [lacht], weil ähm, ok, nicht mal der muss drin sein, aber ne Überschrift wär schön gewesen auf der Tafel, weil wenn man dann drei verschiedene Modelle irgendwo an der Tafel stehen hat, alle drei aus verschiedenen Themen, ohne Überschrift, da kommt man dann durcheinander, ja, der Aufschrieb wird dann nicht so toll. Also, ich weiß nicht, ich hab seltener mitgeschrieben, muss ich so sagen, aus dem Grund. Ich hab mehr das Buch durchgelesen, ich hab sein altes Skript mir ausgedruckt, muss ich sagen, weil im Prinzip, auch wenn er sich da dran auch nicht gehalten hat, das alte Skript nah an den Tafelaufschrieb ranging.

F: Mhm.

A: Aber sonst. […] War gut, war fair. War sehr fair.

F: Ok. Wunderbar.

F: Also, die erste Frage ist: Was war deine erste Empfindung, als du von den R Aufgaben gehört hast und hat sich an deiner Einstellung bis heute etwas geändert?

A: Also, die erste Empfindung war reiner Schock, weil ich mich mit Programmieren überhaupt nicht auskenne und den Zeitaufwand ziemlich krass fand, der da beansprucht werden sollte, wenn man das Programm noch nicht kennt. Das hat sich dann während des Semesters etwas gelegt, weil es super Unterstützung gab und jetzt find ichs gut, dass wir nen Einblick bekommen haben in diese Programmiersprache, weil man, wenn man das noch nie gemacht hat, gar kein Verständnis hat, was das, was man da machen muss.

F: Mhm. Wie häufig hast du die R Aufgaben bearbeitet?

A: Jedes Mal.

F: Immer, mhm. Und du warst nicht in der Fragestunde?

A: Ne, wir waren …

F: Wir haben das ja extra gemacht. [Sie und zwei weitere Lehramtsstudentinnen kamen regelmäßig in die Sprechstunde, um Fragen zum Programm zu klären.]

A: Genau.

F: Warum hast du denn die Fragestunde nicht besucht?

A: Weil wir vorgearbeitet haben. Wir wollten, ja doch, die Aufgaben immer so schnell wie möglich lösen, um zu gucken, ob wir auch Probleme haben, die länger als ne Woche dauern und da es ja dann so lang gedauert hat, bis die ersten Aufgaben besprochen wurden, wollten wir nicht noch länger warten, sondern sofort an die neuen Aufgaben gehen, uns auch Gedanken machen.

F: Mhm.

A: Weil wir hatten ja auch durchaus Ansätze und wollten das so machen.

F: Ähm, ok. Fandest du die R Aufgaben im Vergleich zu den anderen Aufgaben eher schwer oder eher leicht? Und was war schwerer und was war leichter?

A: Hm, also da Stochastik überhaupt nicht mein Fall ist und ich immer nur Bahnhof verstanden hab, fand ich die R Aufgaben leichter als die Stochastik-Aufgaben, aber ganz ohne Hilfe hätt ichs nicht geschafft.

F: Und was war leichter an den R Aufgaben?

A: Dieses direkte Logische. Also, wenn da stand, man soll das und das machen, dann sollte, dann musste man auch genau das machen und bei Stochastik, wenn man da ne Fragestellung hat, muss man erstmal herausfinden, was man überhaupt tun soll und wie und, das war bei R viel klarer.

F: Mhm.

A: Auch wenn man dann nicht unbedingt immer wusste, wie mans umsetzen soll, aber den Weg kannte man ungefähr, die Richtung zumindest.

F: Mhm. […] Was fandest du an den R Aufgaben gut? An welche erinnerst du dich noch und warum?

A: Ich erinner mich noch an Größenverzerrung, ähm, an […] ich weiß nicht, ob wir das in dem Kontext besprochen hatten, als wir irgendwann mal über die, über diesen Placebo-Effekt und diese Testergebnisse besprochen haben, ich glaub Brustkrebs oder so, also […] ich fand, was ich gut an denen fand, war die Frage, richtig?

F: Mhm. Und an welche du dich erinnerst.

A: Ich fands gut, wenn man diesen Realitätsbezug hatte, und wenn man daran gesehen hat, also häufig ist es so, dass man irgendwas berechnet in Mathe und man sieht keine Anwendung und da wars so, dass man ne Rechnung hatte und ne Anwendung hatte und durch das Programm, durch die Simulation gesehen hat, welche Auswirkung falsche Messergebnisse, falsche Mess-Herangehensweisen haben können. Oder auch wie man das einfach nur falsch interpretiert, was da rauskommen kann.

F: Mhm.

A: Das fand ich sehr, sehr gut. Und auch, war es das mit den Schildkröten? In der Vorlesung? Oder Fische oder irgendwelche Tiere?

F: Ah. Ja, das war in der Vorlesung.

A: Mhm, da fand ich es auch gut, dass man...

F: Krebse.

A: Ach, Krebse, genau [lacht]. Dass man durch Simulation einfach verschiedene ähm reale Situationen sim, ja nachstellen kann und das Verhalten daran beobachten kann. Das hat mir ganz gut gefallen. Was man normalerweise, wenn man ne normale Stichprobe macht, nimmt man ja immer so nen kleineren Bereich und dann können da, gibt's viele Ausfälle, und da sind wir ja auf ne Million manchmal gegangen und dann war das so ganz schön zu erkennen. Das fand ich sehr, sehr gut daran.

F: Mhm, mhm. Gibt es irgendetwas, was dir durch eine R Aufgabe klar geworden ist, was du vorher nicht verstanden hast?

A: Ja, also vor allem das mit der Größenverzerrung hab ich besser verstanden ähm und diese, ach ja genau, das war das mit dem Radio, mit dem Radiosenderwechsel, also mir ist viel bewusster geworden, wie man mit Statistik oder mit, mit ähm Messungen umzugehen hat, dass man wirklich ganz genau sein muss, um das zu testen, was man testen möchte.

F: Mhm. Ähm, gibt es denn allgemein etwas, was man an den R Aufgaben deiner Meinung nach besonders gut lernen kann?

A: Besonders gut lernen kann?

F: Mhm.

A: Naja, wenn man was simuliert, sieht man daran, kann man ja Erkenntnisse gewinnen, das ist so das Wichtigste. Also, man kann natürlich verschiedene Programme schreiben, die dann für irgendwas nützlich sind, aber damit werd ich nie was zu tun haben, wichtiger fand ich, dass man irgendwas simulieren kann und dadurch dann so ein Ergebnis hat und das interpretieren kann.

F: Mhm. Was hat dir an den R Aufgaben nicht gefallen?

A: Manchmal die zu krasse Logik [lacht]. Also, wenn man halt, wenn man kein Informatik studiert und noch nie was damit zu tun gehabt hat, dann ist es schwer, auf einmal, zu akzeptieren, dass der Computer dumm ist. Dass man ihn wirklich mit allem füttern muss.

F: Aha.

A: Und dann verzweifelt man sehr schnell alleine. Das, also wie gesagt, wenn keine Unterstützung da gewesen wäre, hätt ich die Aufgaben nicht lösen können.

F: Mhm. Würdest du aus Studentensicht empfehlen, weiterhin R Aufgaben zu verwenden und würdest du etwas verändern?

A: […] Ich fands nicht schlecht, das zu machen. Ich würde aber auf jeden Fall immer sagen, dass ne Unterstützung da sein sollte durch ne Fragestunde, durch, also irgendwas, wo sich Studenten hin wenden können, weil es sonst nicht machbar ist, und dass vielleicht irgendwo berücksichtigt werden sollte, wenn jemand alle diese R Aufgaben macht, dass man vielleicht irgendwie nochmal einen Extra-Punkt kriegt oder sonst irgendwas, weil es ein enormer zusätzlicher Aufwand ist.

F: Mhm.

A: Also, allein sich selbst damit auseinander zu setzen, und dann nochmal zur Fragestunde zu gehen und dann nochmal zu Hause alles durchlaufen zu lassen und die Fragen zu beantworten, das dauert fast so lange wie die ganzen anderen drei Aufgaben.

F: Mhm.

A: Man muss es ja immer in Relation sehen zu dem, was man noch so zu tun hat.

F: Mhm, mhm. Aber würdest du es dann mit diesen Änderungen weiterhin empfehlen oder würdest du sagen, man sollte es lieber wieder sein lassen?

A: […] Ne, ich fands eigentlich gut, dass man Simulationen hatte, dass man die Möglichkeit hatte, was zu simulieren.

F: Mhm.

A: Weil anders würde mans nicht machen, weil man kein Wissen darüber hätte.

F: Mhm, gut. Wurde in deinem Tutorium manchmal auf R Aufgaben Bezug genommen, wenn eine andere Aufgabe an der Tafel erklärt wurde?

A: Nein. Nein, meine Tutorin konnte das Programm nicht.

F: Ok. So, jetzt nochmal kurz was Inhaltliches zu Stochastik: Woran denkst du als erstes, wenn du an Stochastik denkst?

A: Kann ich nicht, daran denke ich.

F: Gut.

A: Ich denke vor allem daran, dass ich keine stochastische Intuition besitze.

F: Ok, darauf kommen wir gleich noch. Ähm, und an was für einen Begriff oder was für ein Objekt denkst du?

A: […] Ähm, […] an was für ein Objekt denk ich, […] irgendwas Grausames, irgendwas Schlimmes [lacht], […] weiß ich nicht, an welches Objekt ich denke.

F: Und ein Begriff?

A: Verzweiflung und Hoffnungslosigkeit.

F: [lacht] Und ein stochastischer Begriff?

A: Stochastischer Begriff?

F: Mhm.

A: Einer, der alles beschreiben soll?

F: Nö, einfach irgendwas, was dir als erstes in den Sinn kommt, wenn du an Stochastik denkst.

A: Achso, hm, Hypothesentests.

F: Ok. Und kannst du erklären, was das ist?

A: Man stellt eine Hypothese auf, ja, und will mit irgendeiner Wahrscheinlichkeit überprüfen, ob die stimmt und dann prüft man das und dann kann man mit einer gewissen Wahrscheinlichkeit, je nachdem, was man halt vorher gesagt hat, 95% oder was weiß ich, kann man sagen, ob die Hypothese wahrscheinlich stimmt oder nicht.

F: Mhm.

A: War das gut erklärt?

F: Ok, können wir gleich noch thematisieren. Ähm, fällt dir noch ein anderer Begriff ein?

A: Ja, es gibt einige. Es gibt Kovarianz, Erwartungswert, so was.

F: Mhm. Und kannst du da erklären, was es ist?

A: Ich hab Semesterferien [lacht]. Ähm, Varianz ist die quadratische Abweichung, Sigma Quadrat, […] boah ich erinner mich nicht mehr an Stochastik, irgendwas mit Abweichung und quadriert, quadratische Abweichung, ähm, Kovarianz, ist […] ich hab mich häufig gefragt, was Kovarianz eigentlich ist, Wikipedia hat keinen Eintrag dazu, ähm, heißt nur, glaub ich, ob irgendwelche Varianzen miteinander kollidieren, wie weit sie miteinander kollidieren […] irgendwas mit Abhängigkeiten war da, wenn man nur eine Zufallsvariable hat, dann ist Kovarianz gleich Varianz, hm, ja, so was.

F: Ok, so wars, und Erwartungswert?

A: Ja, das sagt ja der Name schon. Ich erwarte einen bestimmten Wert. Das ist n mal p.

F: Ok. […] Wunderbar. Ähm, jetzt hast du den Begriff eben schon fallen lassen. Was verstehst du unter stochastischer Intuition?

A: Dass man ein stochastisches Problem vor sich hat und dann intuitiv weiß, meistens sind es ja ähm zum Beispiel so was wie das Geburtstagsproblem, das wird häufig angebracht, wenn man auf stochastische Intuition hinaus will, und dann wird häufig also sagen alle dann, dass es ähm nicht wahrscheinlich ist, dass bei 25 Leuten zwei am selben Tag Geburtstag haben, und dann sagt man, siehst du, das ist die stochastische Intuition, und daran sieht man, dass das fast keiner hat. Denn sonst würde man ja, würden ja ein paar darauf tippen, dass zwei am selben Tag Geburtstag haben. Also würd ich sagen, ist das, dass man intuitiv die richtige Antwort hat.

F: Mhm. Und gibt es nen Unterschied zu Verständnis von Stochastik?

A: […] Bestimmt. Intuition hat ja nichts mit Verständnis zu tun. Intuition kann ja auch kommen, wenn man gar kein Verständnis für irgendwas hat. Man kann ja intuitiv auf was tippen.

F: Mhm. Und wann, würdest du sagen, hat man in Stochastik was richtig gut verstanden?

A: Wahrscheinlich wenn man beides besitzt, wenn man sowohl die Aufgabe versteht, als auch intuitiv, also wenn man intuitiv schon mal weiß, wo ungefähr, wo's lang gehen muss, und dann das Verständnis dafür hat, um die Aufgabe lösen zu können.

F: Mhm. Also das heißt, Verständnis wär sozusagen das Lösen von, von Aufgaben.

A: Hm […] vielleicht bräuchte man da noch ein bißchen Wissen dazu. Aber zumindest schon mal die Aufgabe so einordnen zu können, dass man weiß, worauf man hinaus will. Weil ganz häufig bei stochastischen Fragen ist es ja so, dass man erstmal gar keine Ahnung hat, was man überhaupt machen soll. Außer da steht: Berechne die Kovarianz oder so was. Aber das kommt ja ganz selten nur vor.

F: Mhm, mhm.

A: Also man fühlt sich häufig verloren.

F: Ok. Ähm, ich hab jetzt noch eine Aufgabe für dich. Und zwar kam in den Fragebögen ab und zu oder zweimal besser gesagt diese Krankenhausaufgaben oder

Abwandlungen davon vor.

A: Ja.

F: Ich fass es nochmal zusammen, also man hat zwei Krankenhäuser, ein großes und ein kleines.

A: Ach ja, das war diese Abweichung.

F: Genau, im Großen werden im Schnitt 45 Kinder am Tag geboren und im kleinen im Schnitt 15. Und man geht jetzt davon aus, dass man mit 50%iger Wahrscheinlichkeit ein Mädchen bekommt. Und jetzt wird für den Zeitraum von einem Jahr in beiden Krankenhäusern jeweils die Tage gezählt, an denen mehr als 60% der Mä, äh der Neugeborenen Mädchen waren...

A: Ja.

F: Und man fragt sich, in welchem Krankenhaus erwartet man mehr solche Tage.

A: Im kleinen.

F: Das ist richtig. Wenn du jetzt aber nicht die Antwort sagen solltest, sondern das simulieren würdest...

A: Oh Gott.

F: Wie würdest du vorgehen? Also, es geht nicht um den Code, sondern nur um die Idee, um so Schritte.

A: Ooh [lacht], ich glaub, ich weiß das nicht mehr. Ok. […] Ich soll irgendwas simulieren. [...]

F: Also sozusagen wirklich nur so nen Simulationsplan. Ja, also ich würde erst das machen, dann das, dann das. Keine Details.

A: Ähm, naja, hm, […] naja, ich hab nen n gegeben, ich habe, ich habe nen p gegeben, und dann […] ich weiß nicht mehr genau, wies ging, aber man müsste es simulieren ein paar Mal, also erst einmal simulieren, dass ne Stichprobe praktisch gezogen wird, und dann geguckt, wieviele von den 15 Kindern mit p gleich 0,5 Mädchen sind, und das einmal schreiben und das, den Code, den ich weiß nicht, ziemlich hoch, ziemlich oft wiederholen, ziemlich oft sind ja immer so 1000, 10000 Mal, irgendwie so was, und dann hinterher so ne Art Mittelwert oder so was bestimmen und dann gucken, oder zählen, man kann da ja so ne Zählvariable einführen und dann jedes Mal, also wenn, so ein wenn-dann Befehl schreiben und immer wenn bei meiner Stichprobe rauskam, dass mehr als 60%, mehr als 60% ähm Mädchen waren, ja, dann soll der, was weiß ich, eins drauf rechnen. Also so ne Zählvariable und dann könnte man ja die vergleichen für beide Krankenhäuser …

F: Also du machst das einmal mit n gleich 15 und einmal mit n gleich 45?

A: Ja, würd ich, doch.

F: Mhm, ja, du hattest es nicht gesagt, deswegen...

A: Achso, ja, ja, für beide würd ich das machen. Ja, und dann, was genau ist die Frage? In welchem Krankenhaus erwarten Sie mehr solche Tage. […] Ja, und dann würde mans ja ganz genau sehen an den Krankenhaus. Das es in dem kleinen Krankenhaus ja wahrscheinlich viel, viel häufiger vorgekommen ist, weil die Schwankung viel größer ist, und bei dem größeren nicht.

F: Gut, danke, absolut perfekt. Möchtest du noch irgendwas anmerken?

A: Hm, […] also, für Lehrämtler ist es, finde ich, fürs Studium sehr nett, ähm während Stochastik so simulieren zu können, aber es ist nicht wirklich, also ich fühl mich jetzt nicht in der Lage, mit dem Programm jetzt hinterher im Schuldienst arbeiten zu können, dafür ist es mir noch zu schwierig, und auch für meine Schüler ist es zu, zu schwer, weil die Plattform, es ist, ich weiß nicht, wie man das nennt […]

F: die Oberfläche?

A: Die Oberfläche, ja genau, ist nicht schülerfreundlich, nicht so was wie, ich hab zum Beispiel mit Geogebra und so was gearbeitet, das ist sehr schülerfreundlich und ich fänd es, ich fänds noch irgendwie schön, wenn man vielleicht, ich weiß nicht, man kann ja auch eine andere Oberfläche wahrscheinlich erstellen, aber, ja. Wenn man noch irgendwas machen könnte, was noch ein bisschen mehr lehrerbezogen ist, weil ja Stochastik auch für Lehrämtler eben ist und ich fänds schön, wenn ich das, was ich lerne in der Vorlesung auch hinterher für mich verwenden kann.

F: Mhm. Also ich muss dazu sagen, es gibt ein anderes Programm, das heißt Fathom. Das wird auch bei den Didaktikern...

A: Ja, das hab ich gehört.

F: eingesetzt, genau. Das Problem an dem ist halt, dass man nicht so fortgeschrittene Sachen damit machen kann, also es ist halt wirklich für Schüler.

A: Ja.

F: Ja, gut, also viel mehr als die hypergeometrische Verteilung kommt da nicht vor. Also man könnte da kaum so was wie Größenverzerrung damit simulieren.

A: Ja. Das ist halt rein...

F: Das ist halt für Schüler.

A: Ja, genau, das ist Fathom schon und R ist halt wirklich um den Uni-Stoff zu verarbeiten, was natürlich Lehramtsstudenten mit der Sicht auf später nichts bringt.

F: Ok.

A: Das ähm, ich finds halt immer schön, wenn die Vorlesung an sich irgendeinen Bezug zur Realität haben, zu dem eben, was man später machen wird, aber das ist bei Mathematik sowieso nicht der Fall, von daher ist es vielleicht gut, dass man einfach irgendwas simulieren kann. Was einem das veranschaulicht.

F: Mhm, ok. Hast du sonst noch was auf dem Herzen?

A: Ich fand die R Aufgabe in der Klausur nicht schön. Ganz viele fanden die nicht schön.

F: Ok. Warum?

A: Weil die viel schwerer war als die, die wir in der Probeklausur hatten.

F: Tatsächlich? Also sie ist mit am besten ausgefallen als Aufgabe.

A: Echt? Weißt du viele Punkte ich hab bei der Aufgabe?

F: Ich glaube, du hast volle Punktzahl.

A: Ich war nämlich ganz unsicher.

F: Ich kanns nochmal nachgucken. Also, es war ja die Polya-Urne.

A: Genau. Man konnt nur, also bei der Probeklausur konnte mans theoretisch einfach runter lesen, also man musste nicht wirklich viel interpretieren können, aber da gabs ja auch dieses Graphik-interpretieren, es gab viel mehr Interpretationsspielraum und darauf waren wir so gar nicht vorbereitet, also nicht, dass wirs nicht gekonnt hätten, aber wir dachten, dass es ähm halt auf dem Niveau ist von der Probeklausuraufgabe und die war schon, also die in der Klausur war schon um einiges schwerer.

F: Ok. Also war nicht beabsichtigt. Also ist auch nicht, ist nicht irgendwie schlecht ausgefallen oder so die ähm die Aufgabe. Aber, ja, tut mir leid, wenn das irgendwie, also ich wollte halt eigentlich was nehmen, was halt, was man erkennt, deshalb hab ich auch die Polya-Urne genommen.

A: Ja.

F: Und, ok, diese Graphik unten, die hab ich halt eigentlich deswegen noch reingetan, damit die Leute, die vielleicht oben sich irgendwie, die da Schwierigkeiten haben, vielleicht bei der Graphik noch irgendwas holen können.

A: Ich fand grad die Graphik am schwersten [lacht].

F: Ok.

A: Und ich fand schwer, weil irgendwo, ich glaub es gab ein i und ein j, ich bin nicht mehr ganz sicher, aber eine Laufvariable existierte nicht im Code und da hab ich [NN.] gefragt, wo denn diese Laufvariable ist und er hat nur gesagt, keine Ahnung, ich kenn das Programm auch nicht so genau.

F: Ja, das kommt manchmal vor, wenn man einfach ne Schleife oft wiederholt, dann lässt man ein j laufen...

A: Ja.

F: Aber das j kommt gar nicht drin vor, das heißt dann nur, es wird so und so oft wiederholt.

A: Ja, das hat mich kurz irritiert.

F: Ja, es ist halt echt irgendwie, also es war wirklich nicht mit äh Absicht gemacht, dass die irgendwie schwer war, also sie ist auch wirklich sehr gut ausgefallen und es wurde nett korrigiert, aber es tut mir leid, wenn so ein negativer äh Geschmack da geblieben ist. Das war nicht beabsichtigt.

A: Ne, wir dachten nur so, also wir hatten uns alle auf die R Aufgabe gefreut und dachten, ja, damit können wir auf jeden Fall volle Punktzahl holen und danach waren wir uns alle nicht sicher, ob wir ein paar Punkte überhaupt kriegen. Und dachten so, ah, Mist, jetzt war alles praktisch umsonst, was wir gemacht haben. Weil wir fanden alle, dass halt die aus der Probeklausur um einiges leichter war.

F: Ich weiß gar nicht mehr, was ich bei der Probeklausur, achso, das war das mit diesen, mit diesen Urnen, ok.

A: Das war das mit diesem Ziehen, grün aus blau und keine Ahnung rot aus schwarz, was auch immer das war, aber das war so, man konnte so jeden Schritt ganz leicht verstehen, und, ich weiß die Klausuraufgabe auch nicht mehr auswendig...

F: Da wurde dann in jedem Schritt noch eins dazu gelegt, das war eigentlich das. Und dann wurde der Anteil, also man hatte rote und schwarze Kugeln und dann wurde am Schluss der Anteil der roten Kugeln bestimmt und der ist dann uniform verteilt. Das konnte man da sehen.

A: Also ich fand schon, also ich hab ja alles ausfüllen können, das wars ja jetzt nicht, ja, ähm, wir hatten nur mit was anderem gerechnet.

F: Ok.

A: Also das war von uns, von unsrer Gruppe war das der Eindruck.

F: Ok. Ne, ist ja gut zu wissen.

F: Gut, also die erste Frage ist: Was war deine erste Empfindung, als du von den R Aufgaben gehört hast, und hat sich an deiner Einstellung bis heute etwas geändert?

A: Ja, meine erste Empfindung war nicht sehr positiv [lacht], ja, ich war ein bisschen schockiert auch, weil ich mir dachte, oje, wie soll ich das denn schaffen, ähm, ja, meine Einstellung hat sich aber verändert dadurch, dass wir ja Hilfe auch bekommen haben und Unterstützung bekommen haben und das Ganze nicht alleine machen mussten [kam zusammen mit zwei anderen Lehramtsstudentinnen regelmäßig in die Sprechstunde], so dass es dann doch irgendwo machbar war, also ging, und am Ende fand ichs sogar, hats sogar fast Spaß gemacht, doch ja.

F: Ok. Zumindest fast [lacht].

A: Ja, doch [lacht].

F: Ähm, wie häufig hast du denn die R Aufgaben bearbeitet?

A: Jede, ja.

F: Und in der Fragestunde warst du nicht?

A: Ne.

F: Warum nicht?

A: Zeitliche Gründe. Ich wär gerne gegangen, aber ich hab gearbeitet, glaub ich, an dem Tag.

F: Mhm. Ok. Fandest du die R Aufgaben im Vergleich zu den anderen Aufgaben eher schwer oder eher leicht? Und was war schwerer, was war leichter?

A: Also, ich fand sie eher schwer, weil ähm, die anderen Aufgaben irgendwie einfach vertrauter waren, von dem, was man aus der Schulzeit kennt, wenn man das ja, oder beziehungsweise dann aus anderen Bereichen, dass man einfach sich hinsetzt, ne Aufgabe lösen muss und was dazu lesen muss, und, das kennt man mehr, und R Aufgaben, die sehen so kryptisch aus am Anfang, dass man das Gefühl hat, oh mein Gott, kann ich nicht. Und dass diese, dieser Schock oder diese Barriere erstmal da ist, die dann erstmal überwunden werden muss, bis man dann sieht, ok, naja, also irgendwie geht das dann doch.

F: Ok. Und vom Inhaltlichen her, kannst du da sagen, ob die schwerer oder leichter waren? Also, das versteh ich jetzt so, dass eher das Formale schwer war.

A: Ja, also inhaltlich waren sie sicherlich am Angang leichter, weil ähm sie ja eben in das Programm erstmal eingeführt haben, so dass sie von Inhalt ziemlich leicht waren, […] im Nachhinein, am Ende fand ich sie nicht unbedingt leichter, da sind mir andere Dinge dann leichter gefallen.

F: Mhm. Was fandest du an den R Aufgaben gut? Und an welche erinnerst du dich noch und warum?

A: [lacht] Oh je. Ähm, ja, also was ich gut finde, ist, dass man ähm ja ne andere Darstellungsweise mal kennenlernt, kennenlernt wie so nen ähm Programm arbeitet, wies aufgebaut ist und wie man Befehle formulieren muss, dass das Programm irgendwas weiß, weil das hat ja kein Gehirn oder keine stochastischen, stochastisches Wissen, sondern es ist nur, ist ja reine Programmierung und Eingabe, ähm und am Ende die Graphiken fand ich immer sehr schön, die dann rausgekommen ist, so dass man schön vergleichen konnte und sehen konnte, ja, was das Ergebnis der Überlegung ist, ähm, das Gute ist halt auch, dass man da wirklich auch mal simulieren kann, dass man irgendwie tausend Würfe macht oder so was, das kann man ja in echt nicht. Ähm, woran ich mich erinner, ja, ich fands eigentlich ganz, was ich, die eine Aufgabe fand ich ganz cool äh wo man das darstellen musste ähm wo wir das projiziert haben auf diese Kugel. Das fand ich ganz cool. Und dann man sich überlegt hat, ok, welche muss ich jetzt abziehen, also es ging ja irgendwie um diesen Würfel und innerhalb hatten wir die Kugel und dann haben wir uns überlegt, ok, warum, die dürfen nicht außerhalb liegen und innerhalb und…

F: Mhm.

A: Wo wir dann überlegt haben mit Flächen dann die Wahrscheinlichkeit da auszurechnen. Das fand ich ganz gut. Das hat mir ganz gut gefallen.

F: Mhm. Erinnerst du dich noch an eine andere?

A: Ähm, gut, die paar am Anfang, waren ja nicht so spektakulär, die waren ja recht leicht, also vom Inhaltlichen her, welche hat mir noch besonders gefallen, hm, […] hm [lacht] der Sommer lässt grüßen.

F: Kein Problem [lacht]. Ist nur, also wenn dir noch eine einfällt, wenn nicht, ist auch nicht schlimm.

A: Ja, wir haben ja immer wieder irgendwas mit Urnen gehabt. Und wir ziehen, gut, das war auch die letzten vor der Klausur, auch wir ziehen von dieser Urne und ähm gucken dann ähm, wo wir dann irgendwie nen Vektor zugeordnet haben ähm immer wenn eine Urne aus einer Kugel gezogen wird, wird dann was anderes, ne Zählerzahl, die Zählervariable kriegt dann eins mehr und solche Sachen. Also, ich glaub auch, ja, das war glaub ich […]. In der Klausur die Aufgabe war auch so, aber ich glaub davor die Übungsaufgabe war auch so mit der Zählervariable. Aber die davor, die Übungsaufgabe war wesentlich leichter als die Klausur. Das fand ich schon.

F: Mhm. Ähm, ok. Gibt es etwas, was dir durch eine R Aufgabe klar geworden ist, was du vorher nicht verstanden hast? Und wenn ja, was?

A: Hm […], das ist ne gute Frage. Ähm […], jetzt aufs Stochastische ja bezogen.

F: Ja.

A: Hm, also ich fand das, wie gesagt, mit diesen Graphiken und Simulationen am Ende gut, was, was man sich einfach nicht so vorstellen kann, wenn man, ja, das berechnet oder wenn einem da gesagt wird, ok, das ist die und die Verteilung. Ok, gut, dann ist es halt so, aber mit dem Programm konnt man halt auch wirklich sehen, ok, ist es tatsächlich so, macht das tatsächlich keinen Unterschied und das ist, dann glaubt man

das alles auch mehr.

F: Mhm. Und gibt es sonst noch was, was man deiner Meinung nach besonders gut lernen kann, also jetzt allgemein mit diesen R Aufgaben?

A: Hm […], vielleicht hilft es bei der Strukturierung, dass es einem einfach klarer wird, dadurch dass das Programm ja nichts weiß und man dem Programm so alles stückchenweise ja ordnen muss, sortieren muss und was ordne ich wem zu und was muss ich jetzt haben und brauch ich jetzt nen Vektor oder irgendwie noch ne Zählvariable, dadurch ähm grade beispielsweise auch bei, bei ähm den Aufgaben mit […] jetzt überleg ich grad, am Ende, wie heißen die, ähm Zweifallexperimente, ne, am Ende die […]

F: zweistufige Experimente?

A: Ja, Gott, genau, zweistufige Experimente [lacht]. Dass man da sich halt genau Gedanken machen muss, was gehört jetzt wie und was ist von wem abhängig. Also das, dieser ganze Aufbau hilft einem zu strukturieren.

F: Mhm. Wunderbar. Was hat dir bei den R Aufgaben nicht gefallen?

A: Hm […], also dass es natürlich frustrierend ist, wenn man irgendwo nen kleinen Fehler hat, der manchmal sogar so klein ist, dass es wirklich nur ne Klammer ist oder irgendwas und das Programm funktioniert halt nicht so wie man das gerne hätte. Das ist natürlich frustrierend, wenn man dann auf die Fehlersuche geht und sagt, ok, wo ist jetzt hier der klitzekleine Fehler, der gemacht wurde, das stimmt. Ähm, ja, ich denk, der Einstieg ist halt wahrscheinlich ein bisschen frustrierend ins Programm, weil man ähm einfach nicht vertraut damit ist, vor allem, wenn man so wie ich überhaupt keinen Informatik-Unterricht hatte oder wenn man halt eigentlich mit dem Computer nicht so viel zu tun hat, so dass es am Anfang so ein bisschen, ja, man muss ins Rollen kommen so ein bisschen, um sich damit ein bisschen ja anzufreunden, sag ich mal.

F: Mhm. Mhm. Würdest du aus Studentensicht empfehlen, weiterhin R Aufgaben zu verwenden und würdest du etwas verändern?

A: Hm, weiterhin in dem Einführungsseminar für Stochastik?

F: Mhm.

A: Ähm, ja, ich würd aber die Hilfe weiterhin so lassen. Also ich würd das weiterhin so machen, dass es was ist, was man zusätzlich machen kann, aber nicht muss, dass man auch ohne die R Aufgaben weiterhin seine Punkte erreichen kann, auch in der Klausur […]

F: Also so wie es jetzt war?

A: Genau. So wie es jetzt war. Dass es frei gestellt ist und dass auch die Hilfe da bleibt, weil ich finde, es soll schon irgendwo dann mehr Spaß machen und ja nicht so auch noch mit Druck irgendwo dahinter. Ich glaub, das würde sich dann, dann würden sich noch weniger Leute irgendwie trauen zu sagen, ja ok, ich mach das oder der Aufwand ist mir das irgendwie wert, ja.

F: Mhm. Ähm, wurde in deinem Tutorium manchmal auf R Aufgaben Bezug genommen, als eine andere Aufgabe an der Tafel erklärt wurde?

A: Ne. Ne, weil ich glaub, das Problem war, dass viele Tutoren das R Programm selbst nicht so kennen und von daher das dann so ein bisschen außen vor gelassen haben, glaub ich.

F: Ok. Jetzt noch was Inhaltliches.

A: Hm, oh je [lacht].

F: Woran denkst du als allererstes, wenn du an Stochastik denkst?

A: An Stochastik. Dann denk ich an Würfel, ich denk an Glücksräder und an Urnen und Kugeln [lacht].

F: Ok.

A: War das ne gute Antwort?

F: Das ist ne sehr gute Antwort. Ähm, und an was für ein stochastisches Objekt oder was für nen Begriff denkst du vielleicht noch?

A: An ein Objekt wie ein Gegenstand, irgendwas, was man anfassen kann?

F: Ne, muss nicht sein. Irgendwas aus der Vorlesung, kann auch ein Begriff sein.

A: Ähm, was ich ja sehr mochte, waren eigentlich diese zweistufigen Zufallsexperimente und am Ende fand ich das auch cool mit ähm diesen Möglichkeitsbereichen ähm mit den Markovketten. Das war leider, haben wir ja nicht so, also kam ja nicht in der Klausur drin vor, aber das mit den Markovketten, diese ja mit den verschiedenen Bereichen, die man dann immer hat, das fand ich ziemlich, das hat mir gut gefallen.

F: Mhm. Und kannst du erklären, was ein zweistufiges Experiment ist?

A: Hehe [lacht], ich hätte vom Anfang was nehmen müssen, ne, zweistufiges Zufallexperiment ist ähm, da geht's auch um bedingte Wahrscheinlichkeiten und da ist immer die Frage, dass ähm zum Beispiel ähm bei Krankheiten beispielsweise, man macht einen Test und man will wissen, ob jemand diese Krankheit hat. Und dann ähm ist immer so die Frage, ja, wie hoch ist die Wahrscheinlichkeit, dass der Test äh so ausfällt, dass dieser Mensch krank ist, er aber tatsächlich gesund ist.

F: Mhm.

A: So und diese zweistufige also Zufallexperimente, dass man sagt, ok, für den Fall, er ist gesund und der Test ist positiv oder negativ, oder er ist krank und positiv, negativ, das ist so. Das fand ich auch ziemlich, da fand ichs ziemlich cool auch zu erfahren, dass ähm man nur ne 50%ige Wahrscheinlichkeit hat, dass wenn der Test so ausfällt, dass er sagt, man sei krank, dass man im Endeffekt mit 50%iger Wahrscheinlichkeit trotzdem gesund sein kann, das fand ich ziemlich cool.

F: Mhm, und ähm, kannst du erklären, was eine Markovkette ist?

A: Ähm, ja, Markovketten, da haben wir das Coole mit den Pfeilen gemacht und mit diesen Bereichen, ne?

F: Mhm.

A: Ähm […] das war immer ähm entweder kann man drin bleiben oder man kann mit ner Wahrscheinlichkeit in die anderen Bereiche gehen und dann startet man, ah genau, dann startet man immer wieder von dem Bereich, in dem man dann ist,startet man von neu. Also, man geht nicht zurück, sondern […] Kann ichs wirklich erklären, warscheinlich nicht [lacht]. Oh, ich hab dieses Bild vor Augen mit diesen ganz vielen Pfeilen, wo wir dann überlegt haben, ok, wie kam man jetzt dahin.

F: Ok.

A: Ja, und man startet jedenfalls, man hat immer dieses neue Wahrscheinlichkeits, wie sagt man denn, ähm, Wahrscheinlichkeitsbereich oder so was, das hat nen Namen, ich weiß nicht, hab ich vergessen, ja.

F: Ok. Und was meinst du mit man geht nicht zurück?

A: Hm, ist das, also man kann natürlich wieder zurück fallen, je nachdem, was danach kommt, welches Ereignis, aber man startet, das ist ja ne Kette, man startet von dort, wo man hingelangt ist, also man hat am Anfang weiß ich nicht, wenn wir irgendwie äh macht das mit dem Würfeln Sinn, weiß ich gar nicht, also wenn wir würfeln und ich starte von einer Wahrscheinlichkeit und es gibt beispielsweise jetzt ganz einfach gerade und ungerade Zahl, und dann hab ich ne, und ich will beispielsweise äh grade, grade ungrade haben oder so und dann würfel ich das erste Mal und es ist grade, dann komm ich aufs nächste Feld, ok, grade und dann gehe ich von dort aus weiter.

F: Ok.

A: Und wenn ich dann ungrade würfel, geh ich dann aufs Ungrade oder wenn ich grade und so weiter, ja.

F: Ok, man hat eine Idee. [lacht]

A: [lacht]

F: Ähm, was verstehst du unter stochastischer Intuition? Das ist so ein Begriff, der in den Fragebögen aufgetaucht ist.

A: [lacht] Ja, stochastische Intuition ist das, was man anzweifelt zu haben [lacht] Ähm, ne, stochastische Intuition, ja gut, ist ja immer so die Frage ähm […] würd ich zum Beispiel bezeichnen ähm mit der Fähigkeit etwas einschätzen zu können, obs wahrscheinlich ist oder nicht, also dass wenn irgendjemand zu mir kommt und sagt ähm weiß ich nicht, wenn Sie jetzt Lotto spielen, dann äh ist es wesentlich wahrscheinlicher, dass Sie gewinnen, weil blablabla, dass ich dann die Fähigkeit habe, ja, intuitiv zu sagen, ja ok, es könnte sein oder irgendwie zu sagen ne, das glaub ich nicht.

F: Mhm.

A: Oder dass beispielsweise auch so Erwartungswerte oder so was, dass irgendjemand sagt, ja der Erwartungswert, man erwartet das, dass ich mir überlegen kann, hm, stimmt das jetzt oder stimmt das nicht. Oder dass ich mir selbst überlege, was erwarte ich denn eigentlich, dass es kommt. Dass man so ein bißchen ein Gefühl dafür hat, wie wahrscheinlich oder wie oft Dinge vorkommen oder nicht vorkommen.

F: Mhm. Mhm. Und gibt es einen Unterschied zu Verständnis von Stochastik?

A: Von der Intuition her?

F: Mhm. Also Unterschied zwischen Intuition und Verständnis.

A: Ja, ich denk schon. Also, ähm, weil man kann ja intuitiv beispielsweise sagen, dass, ja, es kann ja Schüler eigentlich schon sagen, dass, wenn man sie fragt, ja wie, wenn du ganz oft würfelst, wie oft erwartest du eine bestimmte Zahl, dann können Schüler ja schon intuitiv sagen, naja, also ich schätze so und so häufig kommt die Zahl, je nachdem wie oft ich würfel. Und dass sie schon auch intuitiv sagen können, naja, ich erwarte, dass jede Zahl gleich oft vorkommt beispielsweise.

F: Mhm.

A: Oder wenn ich jetzt ähm einmal würfel, ich erwarte ne 3,5 im Prunzip zu würfeln, ne. Ähm aber ich denke, dass ein Verständnis dahinter nochmal was anderes ist, weil Schüler nicht unbedingt wissen, ok, warum kommt denn jetzt jede Zahl gleich oft vor oder wenn es nicht so ist, ist es dann ein Beweis dafür, dass der Würfel gezinkt ist beispielsweise. Also das wüssten sie auch nicht, wenn sie ein Verständnis nicht dafür haben, also ich denk schon, dass da ein, ja, Unterschied ist.

F: Mhm. Und kannst du irgendwie ausdrücken, wann man was richtig gut verstanden hat in Stochastik?

A: Ist ne gute Frage, ähm, was richtig gut verstanden, ähm wenn die Aufgabe richtig gelöst ist. Nein [lacht]. Ja dann natürlich, dann hat mans verstanden. Ähm […] ja, ich denke mir aber auch, dass man versteht beispielsweise jetzt auch mit dieser Krankheit und dem Test, dass man sagt, gut mit 50%iger Wahrscheinlichkeit kann man trotzdem noch gesund sein, dass man versteht, woher das kommt.

F: Aha.

A: Also wir hatten das auch mit Beispiel aus diesen Urnen und verschiedenen Anzahl von denen, also dass man sich einmal das veranschaulicht hat und dann auch versteht, woher diese 50% kommen und woher das Ganze, dass es auch wirklich im Prinzip ja fest ist, fest steht, dass die Wahrscheinlichkeit 50% ist und dass es nicht, weil viele denken ja, Wahrscheinlichkeit ist ja trotzdem noch irgendwie ja Zufall. Aber trotzdem sind ja manche Sachen berechnbar durch Wahrscheinlichkeiten. Dass eben doch nicht immer alles so zufällig ist wie man, wie man denkt. Und diese Erkenntnis, ich würde sagen, dann hat man Stochastik verstanden oder so die Grundzüge von der Stochastik verstanden, wenn man nicht mehr denkt, dass alles einfach irgendwie passiert, sondern dass wirklich, dass man das wirklich teilweise berechnen kann.

F: Mhm.

A: Oder auch argumentativ verwenden kann wie beispielsweise mit der Aufgabe mit der vergewaltigten Frau, die da ähm, wo's darum geht, ob das der Mann war oder ob sie irgendwo vergewaltigt wurde.

F: Mhm. Ok. Dann hab ich noch eine Aufgabe für dich. Und zwar kam ja in den Fragebögen diese Krankenhausaufgabe vor.

A: Hmm, schön, hmm.

F: Ich fass sie nochmal zusammen. Also man hat ein großes und ein kleines Krankenhaus. In dem großen werden im Schnitt 45 Kinder am Tag geboren und im kleinen im Schnitt 15.

A: Mhm.

F: Und wir gehen jetzt davon aus, dass die Wahrscheinlichkeit für ein Mädchen 50% ist.

A: Mhm.

F: Und jetzt möchte man für den Zeitraum von einem Jahr in beiden Krankenhäusern jeweils die Tage zählen, an denen mehr als 60% der Neugeborenen weiblich, also weiblich waren, Mädchen waren. Und die Frage ist, in welchem der beiden Krankenhäuser erwartet man mehr solcher Tage. Ich möchte jetzt keine Antwort hören oder du kannst auch gerne eine sagen, aber damit bist du noch nicht erlöst [lacht]...

A: Ok [lacht].

F: sondern, ähm, mich interessiert, wie du das simulieren würdest. Also du brauchst keinen Code zu nennen, sondern es geht sozusagen nur um die Schritte, also was für Schritte du machen würdest, wenn du das simulieren würdest.

A: Mhm. Also die Krankenhäuser, hm, ok, Moment, jetzt muss ich erstmal selber überlegen.

F: Mhm.

A: Hm […], jeden Tag werden 45 Kinder geboren […], ok, also […] ich denke gerade spontan an Urnen ähm, dass ich das mit Urnen und Kugeln simuliere, aber ich überlege […] hm […] also dass ich beispielsweise, ja gut, aber das müsste vorher bestimmt sein, wieviel Urnen da sind. Hm […] das geht nicht, ne. […] Ja, wobei doch, es geht. Also, ich würde vielleicht eine große Urne nehmen, in der, sagen wir mal, blaue und rosa Kugeln drin sind [lacht]. Die rosa Kugeln sind die Weibchen, die blauen Kugeln sind die Männchen, ganz klassisch, ähm und dann, diese Urne würd ich nur nutzen als, ja, Gebärmaschine, sag ich jetzt mal. Also, dass da zufällig einfach die Kinder geboren werden, so dass man da halt zufällig ziehen kann. Deswegen ist eigentlich egal, wieviel Kugeln da drin sind.

F: Mhm.

A: Und dann würd ich die Krankenhäuser darstellen, ähm, achso, ich weiß ja fest, dass immer 45 in dem einen und 15 in dem andern geboren werden, ok. Dann könnte man

die entweder als ähm […] ich überleg grad, als Vektor oder ne als Zählvariable, eigentlich könnte man als Zählvariable ähm die Krankenhäuser verschieden, also einmal das große und einmal das kleine Krankenhaus klar, aber dann nochmal unterschieden zwischen Jungen und Mädchen und Mädchen, Junge. Und dann müsste man halt gucken, jedes Mal, wenn da halt, wenn ein, aus der Urne ne Kugel gezogen wird, können wir ja erstmal dem einen Krankenhaus 45 Kugeln zuordnen und dann dem andern 15 Kugeln zuordnen und die immer, je nachdem zugeordnet, dass wenn ich jetzt beispielsweise ne rosane Kugel ziehe und die dem großen Krankenhaus zuordnen möchte, dass dann die Zählervariable für ähm und wenn die Kugel rosa war, also wenn die Kugel rosa war, dass dann die Zählvariable fürs große Krankenhaus einen mehr bekommt.

F: Aha.

A: So dass ich am Ende dann ähm die Anzahl habe. Und das wäre jetzt für einen Tag. Das müsste man dann natürlich wiederholen für ein komplettes Jahr ähm und dann müsste man sich noch überlegen, und dann müsste man noch ähm dass die Tage, mehr als 50%, ne?

F: Mehr als 60.

A: Mehr als 60%.

F: Mhm.

A: Ok, dass wenn diese Zählervariable von der Anzahl her mehr als 60% sind, ähm, dass dann, könnt ich nen neuen Vektor oder so was machen, eine neue Zählvariable oder so was machen, dass dieser neuen Zählvariable dann auch ne Eins zugeordnet wird, so dass ich dann weiß, ok, in welchem Krankenhaus wurden wo, also an wievielen Tagen wurden in diesen Krankenhäusern mehr als ähm 60% weiblich, genau, 60% Mädchen geboren.

F: Mhm. Und das machst du dann für ein Jahr.

A: Genau. Und das würd ich simulieren dann genau wiederholen für ein Jahr.

F: Mhm. Und dann hast du also zwei Werte für dieses eine Jahr.

A: Mhm.

F: Und dann?

A: Ja, jetzt muss ich grade gucken, was eigentlich die Frage ist. Achso, mh, genau, das wär jetzt eine Simulation, so dass ich dann, ja zwei Werte hätte und wüsste, in dem großen Krankenhaus beispielsweise sind, weiß ich nicht, sagen wir mal, waren im großen, ist ja egal jetzt, aber wenn wir irgendein Ergebnis haben vom großen Krankenhaus und dann haben wir ein Ergebnis vom kleinen Krankenhaus kann ich die ja vergleichen, ob die gleich sind die Tage oder annähernd gleich oder ob sie unterschiedlich sind.

F: Mhm.

A: Und wenn sie stark voneinander abweichen und das ne signifikante Abweichung gibt, dann weiß ich ja, gut, dann wird das in dem einen Krankenhaus generell eben häufiger, kommen diese Tage häufiger vor, so dass es häufiger vorkommt, dass mehr ähm Neugeborene weiblich sind.

F: Mhm.

A: Und wenn die Tage annähernd gleich sind, dann kann ich mir denken, gut, dann macht es keinen Unterschied wie groß oder klein das Krankenhaus ist, sondern dass es generell, ja, dass generell halt das die gleiche Wahrscheinlichkeit ist.

F: Mhm. Also, das heißt, du ähm machst pro Krankenhaus äh simulierst du jeweils immer erst die Tage...

A: Mhm.

F: und zählst wieviele Mädchen waren an jedem Tag...

A: Mhm.

F: und das merkst du dir irgendwie pro Tag...

A: Mhm.

F: und dann guckst du, wo und wann waren mehr als 60% Mädchen...

A: Mhm.

F: und dann zählst du jeweils mit ner neuen Zählvariable hoch und guckst einfach, welcher Wert von beiden ist am Schluss größer...

A: Ja.

F: nach 365 Wiederholungen.

A: Genau, so hätt ich das gemacht. Wenn man dem nicht glaubt, kann man das ja ganz leicht nochmal machen per copy und paste.

F: Aha.

A: Kann man ja oft genug wiederholen, so dass man dann ein paar Werte hat für ein paar Jahre, was eigentlich ein bißchen besser ist als nur ein Jahr.

F: Aha.

A: Das geht ja wahnsinnig leicht mit dem Programm, das ist ja der Vorteil, und dann ist es natürlich, hat man ein besseres Ergebnis, ein genaueres Ergebnis, so dass man dann sehen kann, ok, es war nicht nur ein Zufall dieses eine Jahr, sondern in der Tat hat das eine mehr oder weniger solcher Tage.

F: Perfekt. Auf letzteres wollt ich hinaus, dass man das nochmal wiederholen kann.

A: [lacht] Ja, genau.

F: Wunderbar. Möchtest du sonst noch irgendetwas anmerken?

A: Ähm […] ne, eigentlich nicht. Also, was ich halt schön fand, ist, dass ähm ja, durch die Simulation man mehr den Zugang zu äh also ich hatte auch in der Uni noch nicht äh ein Seminar mit nem Computerprogramm für ähm die Mathematik, aber ähm deswegen fand ich das ganz gut auch mal so ein Programm kennenzulernen, wobei ich auch mir denke, ja, es hat mir schon geholfen, aber ob ich's wirklich dann auch, ob ich fähig bin es zu benutzen, sag ich mal, für mich selbst, weiß ich noch nicht.

F: Mhm.

A: Also ich denk für so ein paar einfache Sachen oder ich mein, wir haben ja jetzt Unterlagen darüber, so ein paar Ideen kann man sich behalten, so dass man die eventuell wieder vorzeigen kann, aber von mir aus selbstständig, ja, weiß ich nicht, ob ich das noch in ein paar Jahren kann, aber, ja, das müsste man gucken, denk ich.

F: Mhm. Wunderbar.

A: Ja.

F: Dann vielen Dank.

A: Bitte.