

# Ein stochastisches Modell für EEG-Microstates

Masterarbeit

vorgelegt am Institut für Mathematik, Schwerpunkt Stochastik  
Fachbereich Informatik und Mathematik  
Johann Wolfgang Goethe-Universität Frankfurt am Main

von

**Matthias Gärtner**

Betreuerin:

Dr. Gaby Schneider

Zweitgutachter:

Prof. Dr. Anton Wakolbinger

Tag der Abgabe: 28. März 2013



## **Erklärung zur Masterarbeit**

Ich erkläre hiermit, dass ich die Arbeit selbstständig verfasst, keine anderen als die angegebenen Hilfsmittel verwendet und die Stellen, die anderen Werken im Wortlaut oder dem Sinne nach entnommen sind, mit Quellenangaben kenntlich gemacht habe. Dies gilt auch für Zeichnungen, Skizzen sowie bildliche Darstellungen.

Frankfurt, den 28. März 2013

Matthias Gärtner



## Danksagung

Bei Frau Dr. Gaby Schneider bedanke ich mich herzlich für die Unterstützung zur Erstellung dieser Masterarbeit, für Anregungen zu Inhalt und Struktur sowie für das kontinuierliche hilfreiche Feedback.

Herrn Dr. Brooks Ferebee danke ich sehr für die fachliche Begleitung bei der Untersuchung des stochastischen Modells. Er hat sich immer wieder Zeit für meine Fragen genommen und stand mir mit hilfreichen Ratschlägen zur Seite.

Für die Bereitstellung der Daten, die dieser Arbeit zugrunde liegen, sowie für Hintergrundinformationen zu EEG-Microstates und den Daten, danke ich Herrn Dr. Helmut Laufs und Frau Dr. Verena Brodbeck.



# Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b>	<b>9</b>
1.1	Neurophysiologische Motivation . . . . .	9
1.2	Mathematische Problemstellung . . . . .	9
1.3	Aufbau der Arbeit . . . . .	11
<b>2</b>	<b>Die untersuchten Daten</b>	<b>13</b>
2.1	Beschreibung der Daten . . . . .	13
2.2	Auftrittshäufigkeiten von Zuständen in den Sequenzen . . . . .	16
<b>3</b>	<b>Das SMI-Modell</b>	<b>18</b>
3.1	Motivation . . . . .	19
3.2	Verallgemeinerung des SMI-Modells und mögliches Hintergrundmodell . . . . .	25
3.2.1	Die Fälle $\pi = 1$ und $\pi = 0$ . . . . .	33
3.2.2	Anzahl der Parameter . . . . .	33
<b>4</b>	<b>Anpassung des SMI-Modells</b>	<b>35</b>
4.1	Schätzer für $p_r$ . . . . .	36
4.1.1	Vergleich mit ML-Schätzer auf der Basis von Hintergrundmodell-Variablen . . . . .	37
4.1.2	Überprüfung von $\hat{p}_r$ durch Simulation . . . . .	39
4.2	Schätzer für $\pi$ . . . . .	41
4.2.1	Vergleich mit ML-Schätzer auf der Basis von Hintergrundmodell-Variablen . . . . .	43
4.2.2	Überprüfung von $\hat{\pi}^{(u)}$ durch Simulation . . . . .	44
4.3	Vergleich der Schätzer $\hat{p}_r$ und $\hat{\pi}^{(u)}$ mit den ML-Schätzern auf der Basis von $(X_0, \dots, X_n)$ . . . . .	44
4.3.1	Lagrange-Ansatz . . . . .	47
4.3.2	Numerischer Ansatz . . . . .	50
4.4	Schätzer für $p_{rs}$ . . . . .	53
4.5	Die aus den Daten berechneten Schätzer . . . . .	53
<b>5</b>	<b>Varianzen der Schätzer</b>	<b>57</b>
5.1	Varianz von $\hat{p}_r$ . . . . .	57
5.1.1	Intuition . . . . .	57
5.1.2	Mathematische Herleitung . . . . .	59
5.1.3	Einfluss der Parameter des SMI-Modells auf $\text{Var}(\hat{p}_r)$ . . . . .	61
5.2	Kovarianzen von Auftritts- und Übergangshäufigkeiten in Markov-Ketten . . . . .	63
5.2.1	Formel für $\text{Cov}(n_r, n_s)$ . . . . .	63
5.2.2	Formel für $\text{Cov}(n_{rs}, n_{tu})$ . . . . .	70
5.3	Varianz von $\hat{\pi}^{(u)}$ . . . . .	76
5.3.1	Herleitung . . . . .	77
5.3.2	Einfluss der Parameter des SMI-Modells auf die Varianz von $\hat{\pi}^{(u)}$ . . . . .	87

<b>6</b>	<b>Überprüfung des SMI-Modells</b>	<b>97</b>
6.1	Ein $\chi^2$ -Anpassungstest für das SMI-Modell mithilfe von Simulation . . . .	98
6.1.1	Mathematischer Hintergrund . . . . .	99
6.1.2	Beschreibung der Durchführung des Anpassungstests . . . . .	101
6.1.3	Ergebnisse des Anpassungstests bei Anwendung auf die gegebenen Daten . . . . .	103
6.1.4	Überprüfung des Fehlers 1. Art für die Durchführung des Anpas- sungstests . . . . .	108
6.1.5	Gründe für die Ablehnung des SMI-Modells . . . . .	110
6.2	Identifizierung der Abweichungen des SMI-Modells von den Daten . . . .	110
6.2.1	Vorüberlegungen . . . . .	112
6.2.2	Durchführung der Ermittlung des 95%-Konfidenzintervalls durch Simulationen . . . . .	114
6.2.3	Ergebnisse . . . . .	115
6.3	Das Erweiterte SMI-Modell . . . . .	118
6.3.1	Idee und Definition . . . . .	119
6.3.2	Anpassung des Erweiterten SMI-Modells . . . . .	122
6.3.3	Ein $\chi^2$ -Anpassungstest für das Erweiterte SMI-Modell . . . . .	124
<b>7</b>	<b>Schlussbetrachtung</b>	<b>127</b>
7.1	Zusammenfassung . . . . .	127
7.2	Ausblick . . . . .	127
<b>A</b>	<b>Ermittlung von EEG-Microstate-Sequenzen</b>	<b>129</b>
<b>B</b>	<b>Die Delta-Methode</b>	<b>132</b>
	<b>Literaturverzeichnis</b>	<b>134</b>
	<b>Abbildungsverzeichnis</b>	<b>136</b>



# 1 Einleitung

## 1.1 Neurophysiologische Motivation

Mit der Methode der Elektroenzephalografie ist es möglich, das elektrische Feld an der Kopfoberfläche zu erfassen, das von den Neuronen im Cortex hervorgerufen wird. Dafür werden meist mehr als 20 Elektroden an der Kopfoberfläche platziert, die in kurzen Abständen (im Millisekundenbereich) das elektrische Feld messen (siehe z. B. Nunez und Srinivasan, 2007).

Ein Ansatz zur Analyse der dadurch erhobenen hochdimensionalen Zeitreihen ist die Ermittlung von sogenannten „EEG-Microstates“<sup>1</sup> (Brodbeck et al., 2012).

Lehmann et al. beobachteten 1987, dass die hochdimensionale Ausprägung (Topografie genannt) des durch die Elektroenzephalografie gemessenen elektrischen Feldes über gewisse Zeiträume annähernd stabil ist (dabei werden Polarität und Stärke der Ausprägung vernachlässigt). Weiterhin stellten Lehmann et al. fest, dass der Wechsel von einer annähernd stabilen Topografie zur nächsten diskontinuierlich erfolgt (Lehmann et al., 1987). Die Zeiträume mit annähernd stabiler Topografie werden als EEG-Microstates bezeichnet (Brodbeck et al., 2012) und dauern typischerweise jeweils etwa 100 Millisekunden an (Lehmann und Michel, 2011).

Die Topografien der EEG-Microstates, die bei der Durchführung einer Elektroenzephalografie zu den Zeitpunkten aufgetreten sind, zu denen die Stärke der Ausprägung des elektrischen Feldes ein lokales Maximum aufweist<sup>2</sup>, werden mithilfe einer Clusteranalyse in Cluster von ähnlichen Topografien eingeteilt. Sind sich die Topografien zweier EEG-Microstates ähnlich, deutet dies darauf hin, dass die Gehirnzustände während der Zeiträume der beiden EEG-Microstates ähnliche Funktionen erfüllen könnten (Lehmann et al., 1987). Aus diesem Grund wurden EEG-Microstates von Lehmann et al. (1998) als mögliche „atoms of thought“ vorgeschlagen, aus denen sich die mentale Aktivität zusammensetzen könnte.

In den meisten Studien wurden stets vier Cluster gefunden (Lehmann und Michel, 2011) und prototypische Topografien bestimmt, sogenannte Karten, die die Cluster repräsentieren und üblicherweise mit den Buchstaben *A*, *B*, *C* und *D* bezeichnet werden.

## 1.2 Mathematische Problemstellung

Jede Topografie eines EEG-Microstates, die zu den oben beschriebenen Zeitpunkten der lokalen Maxima der Ausprägungsstärke des elektrischen Feldes beobachtet wird, lässt sich einem Cluster zuordnen, sodass man eine Sequenz  $x = (x_0, \dots, x_n)$  von Karten mit

---

<sup>1</sup>In dieser Arbeit wird „EEG“ als Abkürzung für Elektroenzephalogramm verwendet.

<sup>2</sup>Die Stärke der Ausprägung des elektrischen Feldes zu einem festen Zeitpunkt wird gemessen durch die Standardabweichung über die Messwerte aller Elektroden zu diesem Zeitpunkt, die auch „Global Field Power“ (GFP) genannt wird (siehe z. B. Brodbeck et al., 2012).

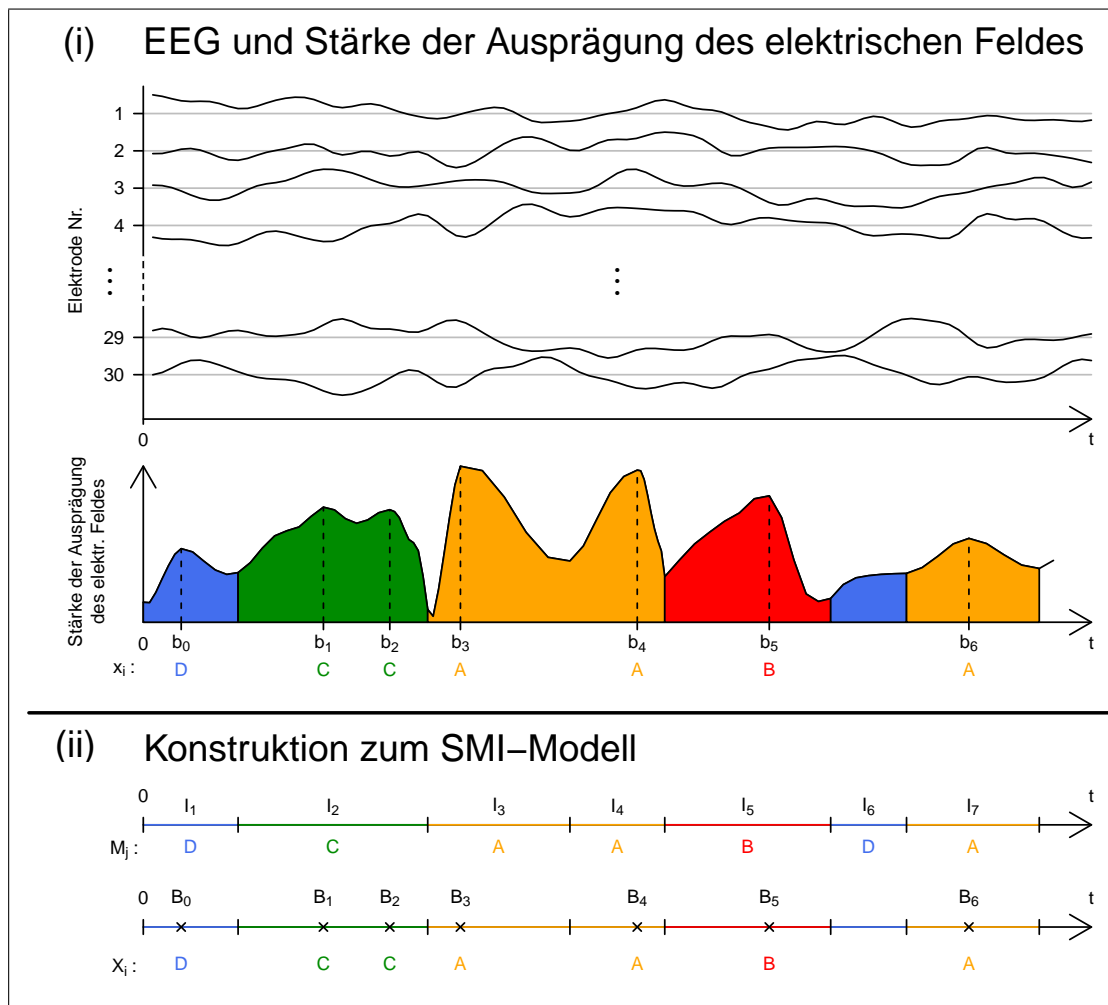


Abbildung 1.1: (i): Exemplarische Darstellung der Ermittlung einer Sequenz  $x$  durch die Zuordnung der Topografien (als EEG dargestellt) zu den Karten  $A, B, C, D$  in den Zeitpunkten  $b_i$  der lokalen Maxima der Ausprägungsstärke des elektrischen Feldes.  
(ii): Darstellung der Ermittlung der  $x$  entsprechenden Markov-Kette  $X$  durch die zu (i) parallele Konstruktion zum SMI-Modell.

Werten in  $\{A, B, C, D\}$  erhält<sup>3</sup>, die die beobachteten EEG-Microstates in ihrer zeitlichen Reihenfolge darstellt (siehe Abbildung 1.1 (i)).

Auf diese Weise gewonnene Sequenzen von Karten werden in dieser Arbeit modelliert als diskrete homogene Markov-Kette 1. Ordnung. Dieser Modellierungsansatz wurde in der Vergangenheit für mehrere Studien verfolgt, indem die Übergangswahrscheinlichkeiten für eine Sequenz  $x$  geschätzt wurden durch die Anzahl der Übergänge in  $x$  von einer

<sup>3</sup>Die Bestimmung der Sequenz  $x$  wird in Anhang A ausführlich erläutert.

Karte  $r$  zu einer Karte  $s$  relativ zur Auftrittshäufigkeit des Ausgangszustandes  $r$  (siehe z. B. Wackermann et al., 1993; Lehmann et al., 2005; Schlegel et al., 2012; Brodbeck et al., 2012).

Ziel der vorliegenden Arbeit ist es, die Struktur der auf diese Weise geschätzten Übergangswahrscheinlichkeiten zu beschreiben und dabei die Anzahl der Parameter zu reduzieren.

Zu diesem Zweck wird zur Modellierung von Sequenzen von Karten ein Modell eingeführt, das durch die spezielle Wahl der Übergangswahrscheinlichkeiten einer homogenen Markov-Kette 1. Ordnung charakterisiert ist. Das Modell wird in dieser Arbeit „Sampled Marked Intervals-Modell“ (SMI-Modell) genannt nach der mathematischen Konstruktion, durch die sich die speziellen Übergangswahrscheinlichkeiten herleiten lassen. Bei dieser Konstruktion wird eine homogene Markov-Kette  $(X_i)_{i \in \mathbb{N}_0}$  1. Ordnung dadurch erzeugt, dass ein zufälliger markierter Intervallprozess  $((I_j, M_j))_{j \in \mathbb{N}}$  zu zufälligen Zeitpunkten  $(B_i)_{i \in \mathbb{N}_0}$  abgetastet wird, wobei die technischen Annahmen getroffen werden, dass die zufälligen Intervalle  $(I_j)_{j \in \mathbb{N}}$  und Abtastungszeitpunkte  $(B_i)_{i \in \mathbb{N}_0}$  jeweils durch unabhängige Poisson'sche Punktprozesse gegeben seien. Für die Markierung  $M_j$  des Intervalls  $I_j$  gelte  $\mathbb{P}(M_j = r) = p_r$  für alle  $r \in \{A, B, C, D\}$  und  $j \in \mathbb{N}$  unabhängig voneinander und unabhängig von  $(I_j)_{j \in \mathbb{N}}$  und  $(B_i)_{i \in \mathbb{N}_0}$  (siehe Abbildung 1.1 (ii)).

Zwischen der Ermittlung der Sequenzen  $x$  von Karten und der Konstruktion der Markov-Ketten  $(X_i)_{i \in \mathbb{N}}$  des SMI-Modells lassen sich Parallelen erkennen, die in Abbildung 1.1 dargestellt werden. Die zufälligen markierten Intervalle  $((I_j, M_j))_{j \in \mathbb{N}}$  lassen sich als Zeiträume  $(I_j)_{j \in \mathbb{N}}$  interpretieren, für die die Topografien des elektrischen Feldes jeweils annähernd stabil sind und die größte Ähnlichkeit zu den jeweiligen Karten  $(M_j)_{j \in \mathbb{N}}$  aufweisen. Die Abtastungszeitpunkte lassen sich vergleichen mit den Zeitpunkten, in denen die Stärke der Ausprägung des elektrischen Feldes ein lokales Maximum aufweist. Zu jedem dieser Zeitpunkte wird die jeweils vorherrschende Topografie dem Cluster zugeordnet, zu dessen Karte sie die größte Ähnlichkeit besitzt, und dadurch die Karte des aktuellen EEG-Microstates identifiziert. Entsprechend wird zu einem Abtastungszeitpunkt  $B_i$ ,  $i \in \mathbb{N}_0$ , die aktuelle Markierung  $M_{j_i}$  des Intervalls  $I_{j_i}$  mit  $B_i \in I_{j_i}$  abgetastet.

### 1.3 Aufbau der Arbeit

Für diese Arbeit liegen Daten zu Sequenzen von Karten für 32 Probanden in vier verschiedenen Schlafstadien vor. Im zweiten Kapitel werden diese Daten vorgestellt und die Notation zur Beschreibung der Daten eingeführt.

Im dritten Kapitel wird das SMI-Modell definiert und die Wahl des Modells anhand der gegebenen Daten motiviert. Zudem wird die oben dargestellte mögliche Konstruktion des SMI-Modells aus zufälligen markierten Intervallen und zufälligen Abtastungszeitpunkten in Form eines „Hintergrundmodells“ ausführlich beschrieben und es wird bewiesen, dass sich aus diesem Hintergrundmodell das SMI-Modell herleiten lässt. Das Hintergrundmodell ermöglicht es dann, die Parameter des SMI-Modells anschaulich zu interpretieren.

Mithilfe der vom Hintergrundmodell gelieferten Interpretationen werden im vierten Kapitel Schätzer für die Parameter des SMI-Modells hergeleitet, die in dieser Arbeit für die Anpassung des SMI-Modells an die gegebenen Daten genutzt werden und SMI-Modell-Schätzer genannt werden.

Mit der Variabilität der SMI-Modell-Schätzer beschäftigt sich das fünfte Kapitel. Es werden dabei zunächst Formeln für die asymptotischen Varianzen der SMI-Modell-Schätzer hergeleitet. Danach wird untersucht, wie sich die asymptotischen Varianzen in Abhängigkeit von den wahren Werten der Parameter des SMI-Modells verhalten.

Das SMI-Modell wird im sechsten Kapitel getestet mithilfe eines Anpassungstests in Anlehnung an Pearsons  $\chi^2$ -Anpassungstest. Anschließend wird untersucht, wie sich die Anpassungsgüte des SMI-Modells verbessern lässt, und eine Abwandlung des SMI-Modells vorgestellt und getestet.

Im siebten Kapitel werden die vorliegende Arbeit zusammengefasst und als Ausblick weitere Forschungsmöglichkeiten präsentiert.

Die computergestützte Auswertung der Daten und sämtliche Simulationen wurden mit R 2.14.1 durchgeführt.

## 2 Die untersuchten Daten

In diesem Kapitel werden die Daten beschrieben, die für diese Arbeit vorliegen. Gleichzeitig werden die im weiteren Verlauf der Arbeit für die Daten verwendeten Notationen eingeführt.

Die Vorstellung der Daten erfolgt in Kapitel 2.1. Dabei wird auch kurz darauf eingegangen, wie die Daten erhoben wurden.<sup>4</sup>

Bei den Daten handelt es sich um Übergangshäufigkeiten von einem Zustand<sup>5</sup>  $r$  zu einem Zustand  $s$  in einer endlichen Sequenz von Zuständen. Die Sequenzen selbst lassen sich aus den Übergangshäufigkeiten nicht rekonstruieren. Das Kapitel 2.2 befasst sich mit der Bestimmung der Auftrittshäufigkeit eines Zustandes  $r$  in einer dieser unbekannt Sequenzen. Da die Bestimmung dieser Auftrittshäufigkeit im Allgemeinen nicht exakt möglich ist, wird in Kapitel 2.2 außerdem angegeben, auf welche Weise die Auftrittshäufigkeit eines Zustandes  $r$  im Rahmen dieser Arbeit approximiert wird.

### 2.1 Beschreibung der Daten

Die für diese Arbeit vorliegenden Daten wurden zur Verfügung gestellt von der Forschungsgruppe REST am Zentrum der Neurologie und Neurochirurgie des Klinikums der Goethe-Universität Frankfurt am Main, namentlich von Herrn Dr. med. Helmut Laufs und Frau Dr. med. Verena Brodbeck.

Für die Erhebung der Daten wurden Elektroenzephalogramme (EEGs) von 32 gesunden Probanden aufgezeichnet. Die Menge aller Probanden sei

$$\mathcal{J} := \{1, 2, \dots, 32\}.$$

Die Aufzeichnungen erfolgten für jeden Probanden zum einen im wachen, entspannten Zustand (W) und zum anderen in verschiedenen Schlafstadien. Dabei wurden nur Schlafstadien des Non-REM-Schlafs<sup>6</sup> berücksichtigt, die mit N1, N2 und N3 bezeichnet werden. Es steht N1 für den Übergang zwischen dem wachen Zustand und dem Schlafen, N2 steht für stabilen Schlaf und N3 für Tiefschlaf (Weeß und Landwehr, 2009).

Der wache Zustand W wird im Folgenden ebenfalls als Schlafstadium bezeichnet, sodass in dieser Arbeit vier Schlafstadien unterschieden werden: W, N1, N2, N3. Die Menge aller Schlafstadien sei

$$\mathcal{Z} := \{W, N1, N2, N3\}.$$

Für alle Probanden gibt es für die Schlafstadien W, N1 und N2 über einen ausreichend langen Zeitraum<sup>7</sup> Messungen (aufgezeichnete EEGs), aus denen die vorliegenden Daten generiert wurden. Das Schlafstadium N3 haben nur 19 von 32 Probanden während

---

<sup>4</sup>Für eine ausführliche Beschreibung der Methoden zur Erhebung der Daten siehe Anhang A.

<sup>5</sup>Zustände entsprechen den Karten aus Kapitel 1.

<sup>6</sup>REM ist die Abkürzung für Rapid Eye Movement.

<sup>7</sup>Ein Schlafstadium muss insgesamt mindestens 105 Sekunden andauern, um in den Messungen berücksichtigt zu werden (Brodbeck et al., 2012).

der Messungen ausreichend lang erreicht, sodass nur für diese Probanden Daten zum Schlafstadium N3 vorliegen.

Für ein Schlafstadium  $Z \in \mathcal{Z}$  bezeichne  $\mathcal{J}_Z$  die Menge aller Probanden, für die es Daten aus dem Schlafstadium  $Z$  gibt. Es gilt

$$\mathcal{J}_W = \mathcal{J}_{N1} = \mathcal{J}_{N2} = \mathcal{J},$$

aber

$$\mathcal{J}_{N3} = \mathcal{J} \setminus \{1, 2, 3, 7, 9, 10, 17, 18, 19, 20, 22, 29, 31\}.$$

Umgekehrt sei  $\mathcal{Z}_J$  die Menge aller Schlafstadien, für die es zu einem Probanden  $J \in \mathcal{J}$  Daten gibt. Beispielsweise ist

$$\mathcal{Z}_3 = \{W, N1, N2\} \quad \text{und} \quad \mathcal{Z}_4 = \mathcal{Z}.$$

Für den Ausdruck „für jeden Probanden  $J \in \mathcal{J}$  und jedes Schlafstadium  $Z \in \mathcal{Z}_J$ “ wird im weiteren Verlauf der Arbeit nur kurz „für alle Probanden  $J$  und alle Schlafstadien  $Z$ “ geschrieben (und analog für ähnliche Formulierungen).

Wenn sich aus dem Kontext nicht ergibt, dass eine Variable zu einem Probanden  $J$  in einem Schlafstadium  $Z$  gehört, oder diese Zugehörigkeit hervorgehoben werden soll, wird an die Variable der Index  $(J, Z)$  geschrieben. Aus einer Variable  $v$  wird dann also  $v^{(J,Z)}$  bzw.  $v_{(J,Z)}$ .

Für jeden Probanden  $J$  und jedes Schlafstadium  $Z$  wurde bei der Datenerhebung aus den entsprechenden EEG-Abschnitten eine Sequenz

$$x = x^{(J,Z)} = \left( x_0^{(J,Z)}, \dots, x_{n^{(J,Z)}}^{(J,Z)} \right)$$

mit

$$x_i^{(J,Z)} \in S := \{A, B, C, D\} \quad \forall i \in \{0, \dots, n^{(J,Z)}\}$$

hergeleitet, wobei  $A, B, C, D$  Zustände und  $S$  Zustandsraum genannt werden. Ist das EEG für einen Probanden  $J$  im Schlafstadium  $Z$  nicht zeitlich zusammenhängend, sondern  $(k - 1)$ -mal unterbrochen, werden für die  $k$  zeitlich zusammenhängenden EEG-Abschnitte Teilsequenzen

$$x^{(j)} = x^{(j),(J,Z)} = \left( x_0^{(j),(J,Z)}, \dots, x_{n^{(j),(J,Z)}}^{(j),(J,Z)} \right) \quad , j \in \{1, \dots, k\},$$

ermittelt mit  $k = k^{(J,Z)} > 1$ , aus denen sich  $x$  zusammensetzt durch

$$x = \left( x_0^{(1)}, \dots, x_{n^{(1)}}^{(1)}, x_0^{(2)}, \dots, x_{n^{(2)}}^{(2)}, \dots, x_0^{(k)}, \dots, x_{n^{(k)}}^{(k)} \right).$$

Die Länge  $n^{(j)} + 1 = n^{(j),(J,Z)} + 1$  der  $j$ -ten Teilsequenz  $x^{(j)}$  hängt dabei sowohl von  $(J, Z)$  als auch von  $j$  ab. Für die Länge  $n + 1 = n^{(J,Z)} + 1$  von  $x$  ergibt sich

$$n + 1 = \sum_{j=1}^k \left( n^{(j)} + 1 \right) = k + \sum_{j=1}^k n^{(j)}.$$

**Bemerkung 2.1.** Die Länge  $n^{(J,Z)} + 1$  der Sequenz  $x^{(J,Z)}$  für Proband  $J$  im Schlafstadium  $Z$  wird im weiteren Verlauf dieser Arbeit als konstant angesehen und kann für jedes  $(J, Z)$  verschieden sein. Sie hängt im Wesentlichen von der durchschnittlichen Anzahl von Auswertungszeitpunkten<sup>8</sup> pro Zeiteinheit ab und zum anderen von der Länge der betrachteten EEG-Abschnitte.

Für diese Arbeit liegen die erhobenen Sequenzen  $x^{(J,Z)}$  nicht vor. Stattdessen sind pro Proband  $J$  und Schlafstadium  $Z$  für jeden Übergang  $(r, s)$  die Summen  $n_{rs} = n_{rs}^{(J,Z)}$  der Übergangshäufigkeiten  $n_{rs}^{(j)} = n_{rs}^{(j),(J,Z)}$  in den Teilsequenzen  $x^{(j),(J,Z)}$ , aus denen sich  $x = x^{(J,Z)}$  zusammensetzt, gegeben. Da die jeweiligen Teilsequenzen zu zeitlich nicht miteinander verbundenen EEG-Abschnitten gehören, werden also für die Ermittlung von  $n_{rs}$  die Übergänge

$$\left( x_{n^{(j)}, x_0}^{(j)}, x_0^{(j+1)} \right)$$

nicht berücksichtigt, die in der Sequenz  $x$  zwischen dem letzten Zustand der  $j$ -ten und dem ersten Zustand der  $(j+1)$ -ten Teilsequenz für ein  $j \in \{1, \dots, k-1\}$  auftreten.

Formal wird  $n_{rs}^{(j)}$  für alle  $r, s \in S$  und  $j \in \{1, \dots, k\}$  definiert durch

$$n_{rs}^{(j)} := \sum_{i=0}^{n^{(j)}-1} \mathbf{1}_{\{x_i=r, x_{i+1}=s\}}.$$

Der Wertebereich eines  $n_{rs}^{(j)}$  ist somit  $\{0, \dots, n^{(j)}\}$  und es gilt

$$\sum_{r,s \in S} n_{rs}^{(j)} = n^{(j)}.$$

Wie oben beschrieben, wird die Anzahl  $n_{rs}$  der Übergänge von  $r$  zu  $s$ , die in der Sequenz  $x$  berücksichtigt werden, definiert durch

$$n_{rs} := \sum_{j=1}^k n_{rs}^{(j)}.$$

Für die Anzahl aller in  $x$  berücksichtigten Übergänge ergibt sich

$$\sum_{r,s \in S} n_{rs} = \sum_{r,s \in S} \sum_{j=1}^k n_{rs}^{(j)} = \sum_{j=1}^k n^{(j)} = n - (k-1).$$

---

<sup>8</sup>Das EEG wird in den Zeitpunkten ausgewertet, in denen die Stärke der Ausprägung des gemessenen elektrischen Feldes lokal maximal ist (siehe Kapitel 1 und Anhang A).

**Bemerkung 2.2.** (i) Für die weiteren Betrachtungen der Sequenzen  $x^{(J,Z)}$  in dieser Arbeit wird zur Vereinfachung stets angenommen, dass sich jedes  $x^{(J,Z)}$  nur aus einer Teilsequenz  $x^{(1),(J,Z)}$  zusammensetzt mit Übergangshäufigkeiten

$$n_{rs}^{(1),(J,Z)} = n_{rs}^{(J,Z)} \quad \forall r, s \in S.$$

Diese Annahme wird auch dann getroffen, wenn keine Sequenz mit in den Daten gegebenen Übergangshäufigkeiten  $n_{rs}^{(J,Z)}$  existiert.

(ii) Unter der Annahme, dass  $x^{(J,Z)} = x^{(1),(J,Z)}$  gilt, unterscheiden sich die Länge  $n^{(J,Z)} + 1$  und die Anzahl  $\sum_{r,s \in S} n_{rs}^{(J,Z)}$  aller Übergänge, die in der Sequenz  $x^{(J,Z)}$  berücksichtigt werden, nur um 1:

$$\sum_{r,s \in S} n_{rs}^{(J,Z)} = n^{(J,Z)}.$$

## 2.2 Auftrittshäufigkeiten von Zuständen in den Sequenzen

Dadurch, dass nicht die Sequenzen  $x^{(J,Z)}$  selbst für diese Arbeit verfügbar sind, sondern lediglich die Übergangshäufigkeiten  $n_{rs}^{(J,Z)}$ , ist es im Allgemeinen nicht möglich, die exakte Auftrittshäufigkeit eines Zustandes  $r$  in einer der Sequenzen  $x^{(J,Z)}$  zu bestimmen. Die Ursache für diesen Umstand wird nachstehend erläutert und es wird angegeben, wie in dieser Arbeit die Auftrittshäufigkeit eines Zustandes  $r$  approximiert wird.

Gemäß Bemerkung 2.2 (i) wird angenommen, dass sich die Sequenzen  $x^{(J,Z)}$  jeweils nur aus einer Teilsequenz zusammensetzen. Somit ist  $n_{rs} = n_{rs}^{(1)}$  die Anzahl der Übergänge von einem Zustand  $r$  zu einem Zustand  $s$  in einer Sequenz  $x = x^{(J,Z)}$  der Länge  $n + 1$ .

Es seien

$$n_{r\cdot} := \sum_{s \in S} n_{rs} = \sum_{i=0}^{n-1} \mathbf{1}_{\{x_i=r\}} \quad \text{und} \quad n_{\cdot r} := \sum_{t \in S} n_{tr} = \sum_{i=1}^n \mathbf{1}_{\{x_i=r\}} \quad \forall r \in S.$$

Für die Auftrittshäufigkeit  $n'_r$  von  $r$  in der Sequenz  $x$  folgt dann

$$n'_r = \sum_{i=0}^n \mathbf{1}_{\{x_i=r\}} = \mathbf{1}_{\{x_0=r\}} + n_{r\cdot} = n_{r\cdot} + \mathbf{1}_{\{x_n=r\}}. \quad (2.1)$$

Ist also  $n_{\cdot r} \neq n_{r\cdot}$  für ein  $r \in S$ , ergibt sich direkt

$$n'_r = \max\{n_{\cdot r}, n_{r\cdot}\}.$$

Gilt jedoch  $n_{\cdot r} = n_{r\cdot}$  für alle  $r \in S$ , dann ist es im Allgemeinen nicht möglich,  $n'_r$  exakt zu bestimmen.



So kann  $n'_A$  beispielsweise nicht exakt ermittelt werden, wenn für den Zustandsraum  $\{A, B\}$  gegeben ist, dass

$$n_{AB}^{(J,Z)} = n_{BA}^{(J,Z)} = 1 \quad \text{und} \quad n_{AA}^{(J,Z)} = n_{BB}^{(J,Z)} = 0$$

gelten. Denn dann ist sowohl die Sequenz  $(A, B, A)$  als auch die Sequenz  $(B, A, B)$  möglich. In einem Fall tritt  $A$  zweimal auf, im anderen einmal.

Wenn die Auftrittshäufigkeit des Zustandes  $r$  lediglich in den ersten  $n$  Zuständen

$$(x_0, \dots, x_{n-1})$$

anstatt in den ersten  $n + 1$  Zuständen  $(x_0, \dots, x_n)$  bestimmt wird, unterscheidet sich die so ermittelte Auftrittshäufigkeit  $n_r$  maximal um 1 von  $n'_r$ . Der dadurch begangene Fehler ist für große  $n$  vernachlässigbar.

Daher wird in dieser Arbeit die Auftrittshäufigkeit  $n'_r$  des Zustandes  $r$  in  $x$  stets durch die Auftrittshäufigkeit

$$n_r := n_r = \sum_{s \in S} n_{rs} = \sum_{i=0}^{n-1} \mathbf{1}_{\{x_i=r\}}$$

des Zustandes  $r$  in den ersten  $n$  Zuständen  $(x_0, \dots, x_{n-1})$  von  $x$  approximiert.<sup>9</sup>

**Bemerkung 2.3.** Setzt sich  $x$  entgegen der für diese Arbeit getroffenen Annahme aus  $k$  Teilsequenzen  $x^{(1)}, \dots, x^{(k)}$  zusammen, dann werden durch die Definition

$$n_r := \sum_{s \in S} n_{rs}$$

die Zustände  $x_{n^{(j)}}^{(j)}$  für alle  $j \in \{1, \dots, k\}$  nicht berücksichtigt beim Bestimmen von Auftrittshäufigkeiten von Zuständen. Dadurch unterscheidet sich  $n_r$  maximal um  $k$  von der tatsächlichen Auftrittshäufigkeit von  $r$  in  $x$ . Da  $k^{(J,Z)}$  für alle Probanden  $J$  und alle Schlafstadien  $Z$  im Vergleich zur Länge der Sequenzen  $x^{(J,Z)}$  klein ist, ist der Fehler zu vernachlässigen, sodass die Annahme aus Bemerkung 2.2 (i) keine Probleme bereitet und weiterhin für die folgenden Kapitel dieser Arbeit getroffen wird.

---

<sup>9</sup>Sofern nicht explizit auf  $n'_r$  verwiesen wird.

### 3 Das SMI-Modell

Die in Kapitel 2.1 beschriebenen Sequenzen  $x^{(J,Z)} = (x_0^{(J,Z)}, \dots, x_n^{(J,Z)})$  werden für jeden Probanden  $J$  und jedes Schlafstadium  $Z$  in dieser Arbeit als homogene Markov-Kette 1. Ordnung auf folgende Weise modelliert:

**Modell 1** (SMI-Modell, mit Zustandsraum  $S = \{A, B, C, D\}$ ). Die Sequenz

$$x = x^{(J,Z)} = \left( x_0^{(J,Z)}, \dots, x_n^{(J,Z)} \right)$$

werde aufgefasst als die ersten  $n + 1 = n^{(J,Z)} + 1$  Zustände

$$\left( X_0^{(J,Z)}, \dots, X_n^{(J,Z)} \right)$$

einer Markov-Kette  $(X_i^{(J,Z)})_{i \in \mathbb{N}_0}$  mit Zustandsraum  $S := \{A, B, C, D\}$  und der Startverteilung  $(p_A, p_B, p_C, p_D)$  sowie der Übergangsmatrix  $P = (p_{rs})_{r,s \in S}$ , gegeben durch

$$p_{rs} := \begin{cases} \pi p_s & , \text{ falls } r \neq s, \\ \pi p_s + 1 - \pi & , \text{ falls } r = s, \end{cases}$$

wobei

$$\pi \in (0, 1]$$

und

$$p_A, p_B, p_C, p_D \in (0, 1) \text{ mit } p_A + p_B + p_C + p_D = 1.$$

Dieses Modell und im Speziellen die Struktur der Übergangswahrscheinlichkeiten  $p_{rs}$  werden in Kapitel 3.1 motiviert und in Kapitel 3.2 durch Modell 2 für beliebige endliche Zustandsräume verallgemeinert. Im Anschluss daran wird ebenfalls in Kapitel 3.2 ein Hintergrundmodell (Modell 3) beschrieben, aus dem sich das SMI-Modell mathematisch herleiten lässt. Zudem wird dieses Hintergrundmodell es erlauben, die Parameter  $p_A, p_B, p_C, p_D$  und insbesondere  $\pi$  zu interpretieren.

Das Hintergrundmodell wird mithilfe von stochastischen Prozessen definiert. Dabei wird die Zeitachse durch einen Prozess in zufällige Intervalle geteilt, denen jeweils ein zufälliger Zustand zugeordnet wird (zufällige markierte Intervalle). Ein zweiter Prozess beschreibt die Abtastungszeitpunkte, zu denen die Zustände des aktuellen Intervalls bestimmt werden. Diese Zustände ergeben in zeitlich aufsteigender Reihenfolge eine Folge  $(X_i)_{i \in \mathbb{N}_0}$  von zufälligen Zuständen (siehe Abbildung 3.1 für ein Beispiel).

Für die Prozesse und deren Zusammenwirken werden im Hintergrundmodell bestimmte technische Annahmen getroffen, sodass sich die Formeln für die Übergangswahrscheinlichkeiten  $p_{rs}$  des SMI-Modells mathematisch aus dem Hintergrundmodell herleiten lassen. Aufgrund der Konstruktion des Hintergrundmodells werden sowohl die Modelle 1 und 2 als auch das Hintergrundmodell selbst in dieser Arbeit „Sampled Marked Intervals-Modell“ (SMI-Modell) genannt.

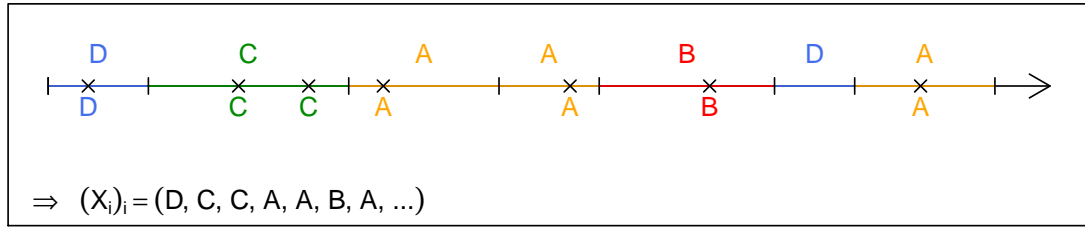


Abbildung 3.1: Exemplarische Darstellung des Hintergrundmodells

Abschließend werden in Kapitel 3.2 Eigenschaften des SMI-Modells bzw. des Hintergrundmodells untersucht, speziell Eigenschaften der Markov-Kette  $(X_i)_{i \in \mathbb{N}_0}$  sowie der Parameter  $p_A, p_B, p_C, p_D$  und  $\pi$ .

### 3.1 Motivation

Die Sequenz  $x$  im oben angegebenen SMI-Modell (Modell 1) ist durch eine homogene Markov-Kette 1. Ordnung  $X$  modelliert. Brodbeck et al. (2012) betrachteten bei ihren Untersuchungen der gegebenen Daten ebenfalls  $x$  als Markov-Kette, trafen jedoch im Unterschied zum SMI-Modell keine besonderen Annahmen an die Struktur der Übergangswahrscheinlichkeiten. Stattdessen wurden die Übergangswahrscheinlichkeiten

$$\hat{p}_{rs}^{(v)} := \frac{n_{rs}}{n_r}$$

für das volle Modell geschätzt. Unter der Voraussetzung, dass  $X$  eine homogene Markov-Kette 1. Ordnung ist, stellen diese geschätzten Übergangswahrscheinlichkeiten  $\hat{p}_{rs}^{(v)}$  des vollen Modells durch die Verwendung aller möglichen 12 Parameter das genaueste Abbild der Realität dar.

In Abbildung 3.2 sind in Anlehnung an Brodbeck et al. (2012) die aus den gegebenen Daten geschätzten Übergangswahrscheinlichkeiten  $\hat{p}_{rs}^{(v),(J,Z)}$  getrennt nach Schlafstadium  $Z$  für jeden Probanden  $J$  eingetragen. Jeder Proband wird dabei in einer anderen Farbe dargestellt.

Bei der Betrachtung der Übergangswahrscheinlichkeiten  $\hat{p}_{rs}^{(v),(J,Z)}$  in Abbildung 3.2 lässt sich insbesondere für  $Z \in \{N1, N2, N3\}$  vermuten, dass diese Übergangswahrscheinlichkeiten keine rein zufälligen Werte annehmen, da ein Großteil der eingezeichneten, verbindenden Linien der verschiedenen Probanden pro Schlafstadium einen ähnlichen, gegebenenfalls parallel verschobenen Verlauf aufweist.

Ziel ist es nun, die Struktur der geschätzten Übergangswahrscheinlichkeiten  $\hat{p}_{rs}^{(v)}$  des vollen Modells zu charakterisieren. Dazu werden anhand von Abbildung 3.2 zwei Beobachtungen angestellt, die anschließend mithilfe der gegebenen Daten diskutiert werden.

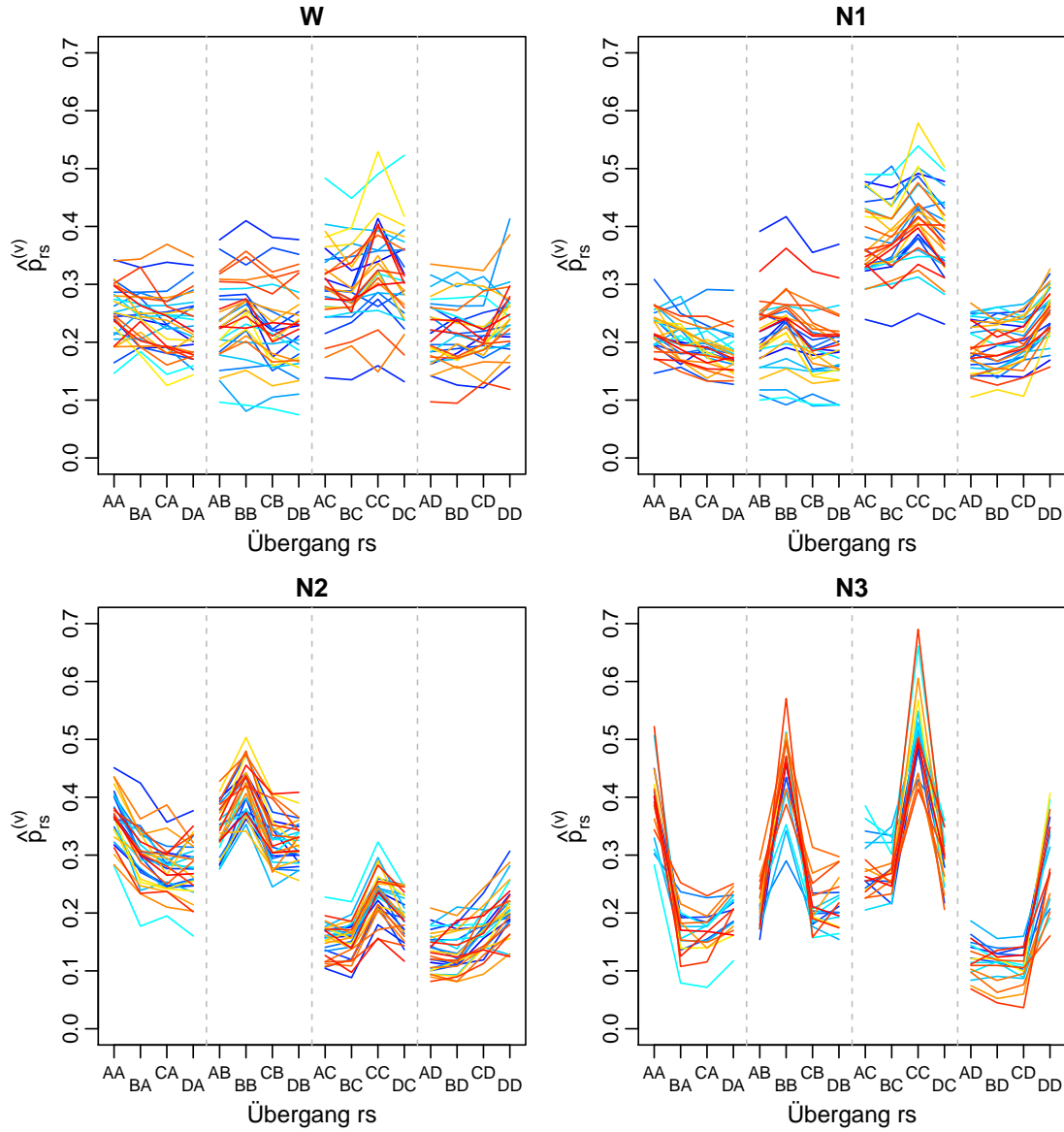


Abbildung 3.2: Übergangswahrscheinlichkeiten aller Probanden getrennt nach Schlafstadium. Die Übergangswahrscheinlichkeiten des vollen Modells sind pro Proband und Schlafstadium geschätzt durch  $\hat{p}_{rs}^{(v)} = n_{rs}/n_r$ . Die Linien verbinden für jeden Probanden, jedes Schlafstadium und jeden Zustand  $s$  jeweils die vier geschätzten Übergangswahrscheinlichkeiten  $\hat{p}_{As}^{(v)}$ ,  $\hat{p}_{Bs}^{(v)}$ ,  $\hat{p}_{Cs}^{(v)}$ ,  $\hat{p}_{Ds}^{(v)}$ .

**1. Beobachtung:** Für einen festen Zustand  $s$  liegen die drei Übergangswahrscheinlichkeiten  $\hat{p}_{rs}^{(v)}$  mit  $r \neq s$  eines Probanden  $J$  im Schlafstadium  $Z$  ungefähr auf einer Höhe,

also beispielsweise für  $s = D$  die Übergangswahrscheinlichkeiten  $\hat{p}_{AD}^{(v)}, \hat{p}_{BD}^{(v)}, \hat{p}_{CD}^{(v)}$ .

**2. Beobachtung:** In den Schlafstadien N1, N2 und N3 sind die Übergangswahrscheinlichkeiten  $\hat{p}_{ss}^{(v)}$  des vollen Modells von einem festen Zustand  $s$  zu sich selbst für den Großteil der Probanden erhöht im Vergleich zu den  $\hat{p}_{rs}^{(v)}$  mit  $r \neq s$ . Diese Erhöhung für einen festen Probanden  $J$  in einem festen Schlafstadium  $Z$  ist für jedes  $s \in S$  ungefähr gleich groß. Im Schlafstadium W lässt sich diese Beobachtung teilweise ebenfalls anstellen.

Diese beiden Beobachtungen werden nun quantifiziert und aus ihnen jeweils eine formale Schlussfolgerung für die Beschreibung der ungefähren Struktur der Übergangswahrscheinlichkeiten des vollen Modells gezogen.

### Zur 1. Beobachtung

Sei ein Proband  $J \in \mathcal{J}$  und ein Schlafstadium  $Z \in \mathcal{Z}_J$  gegeben. Für ein festes  $s \in S$  ist der maximale Abstand zwischen den drei Übergangswahrscheinlichkeiten  $\hat{p}_{rs}^{(v)} = \hat{p}_{rs}^{(v),(J,Z)}$  mit  $r \neq s$  gegeben durch

$$h_s^{(J,Z)} := \max_{r,t \neq s} |\hat{p}_{rs}^{(v)} - \hat{p}_{ts}^{(v)}|.$$

Für die Menge  $\{h_s^{(J,Z)} : s \in S, J \in \mathcal{J}, Z \in \mathcal{Z}_J\}$  ergeben sich die in der Tabelle angegebenen Quantile.

Minimum	1. Quartil	Median	3. Quartil	90%-Quantil	Maximum
0.002	0.018	0.030	0.042	0.058	0.101

Die Hälfte aller  $h_s^{(J,Z)}$  ist also kleiner oder gleich 0.03 und 90% aller  $h_s^{(J,Z)}$  sind kleiner als 0.06.

Wählt man für jeden Probanden  $J$  und jedes Schlafstadium  $Z$  aus allen  $s \in S$  den schlechtesten Fall aus, also

$$h^{(J,Z)} := \max_s h_s^{(J,Z)},$$

sind 50% von diesen  $h^{(J,Z)}$  kleiner als 0.05 und 90% kleiner als 0.08. Die in den Daten beobachteten Werte  $h^{(J,Z)}$  sind in Abbildung 3.3 für jeden Probanden getrennt nach Schlafstadium eingezeichnet.

Insgesamt bestätigt sich also die 1. Beobachtung, dass für einen Probanden  $J$  im Schlafstadium  $Z$  für festes  $s$  die  $\hat{p}_{rs}^{(v),(J,Z)}$  mit  $r \neq s$  sich in den meisten Fällen nur wenig voneinander unterscheiden. Dies führt zu der ersten Bedingung an die Struktur der Übergangswahrscheinlichkeiten  $p_{rs}$ , die vom SMI-Modell erfüllt werden soll:

Für einen festen Zustand  $s$  seien alle drei Übergangswahrscheinlichkeiten  $p_{rs}$  mit  $r \neq s$  gleich groß.

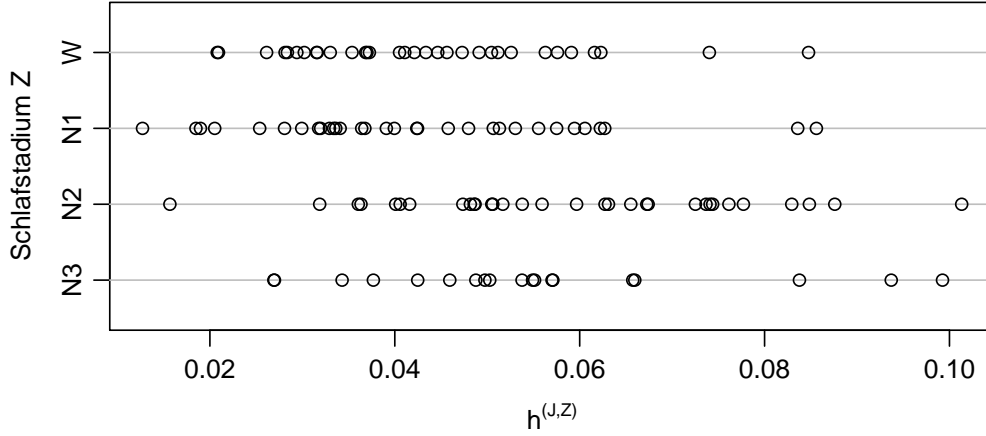


Abbildung 3.3: Maximale Abstände  $h^{(J,Z)}$  pro Proband getrennt nach Schlafstadium. Jeder Punkt entspricht  $h^{(J,Z)} = \max_s \left( \max_{r,t \neq s} |\hat{p}_{rs}^{(v)} - \hat{p}_{ts}^{(v)}| \right)$  für einen festen Probanden  $J$  und das jeweilige Schlafstadium  $Z$ .

Formal bedeutet dies

$$\forall s \in S : \quad p_{rs} = p_{ts} \quad \forall r, t \neq s. \quad (3.1)$$

### Zur 2. Beobachtung

Um die beobachteten Erhöhungen der geschätzten Übergangswahrscheinlichkeiten  $\hat{p}_{ss}^{(v)}$  im vollen Modell von einem Zustand  $s$  zu sich selbst im Vergleich zu den jeweils anderen drei  $\hat{p}_{rs}^{(v)}$  mit  $r \neq s$  abschätzen zu können, ist es notwendig, einen ungefähren gemeinsamen Wert  $\hat{p}_s^{(1)}$  für die drei  $\hat{p}_{rs}^{(v)}$  zu ermitteln. Hier bietet sich der Mittelwert an, sodass zunächst als Schätzer

$$\hat{p}_s^{(1)} := \frac{1}{3} \sum_{t \neq s} \hat{p}_{ts}^{(v)}$$

betrachtet wird.

Es fällt auf, dass die vier Differenzen  $d_A^{(J,Z)}, d_B^{(J,Z)}, d_C^{(J,Z)}, d_D^{(J,Z)}$ , die durch

$$d_s^{(J,Z)} := \hat{p}_{ss}^{(v),(J,Z)} - \hat{p}_s^{(1),(J,Z)} \quad \forall s \in S$$

gegeben sind, pro Proband  $J$  und Schlafstadium  $Z$  in vielen Fällen dicht beieinander liegen. Abbildung 3.4 zeigt die maximale Abweichung

$$\max_{s,u} |d_s^{(J,Z)} - d_u^{(J,Z)}|$$

zwischen den vier Differenzen  $d_A^{(J,Z)}, d_B^{(J,Z)}, d_C^{(J,Z)}, d_D^{(J,Z)}$  jedes Probanden.

Konkret unterscheiden sich die Differenzen  $d_A^{(J,Z)}, d_B^{(J,Z)}, d_C^{(J,Z)}, d_D^{(J,Z)}$  eines festen Probanden  $J$  in den Schlafstadien W, N1 und N2 in der Hälfte aller Fälle um ungefähr

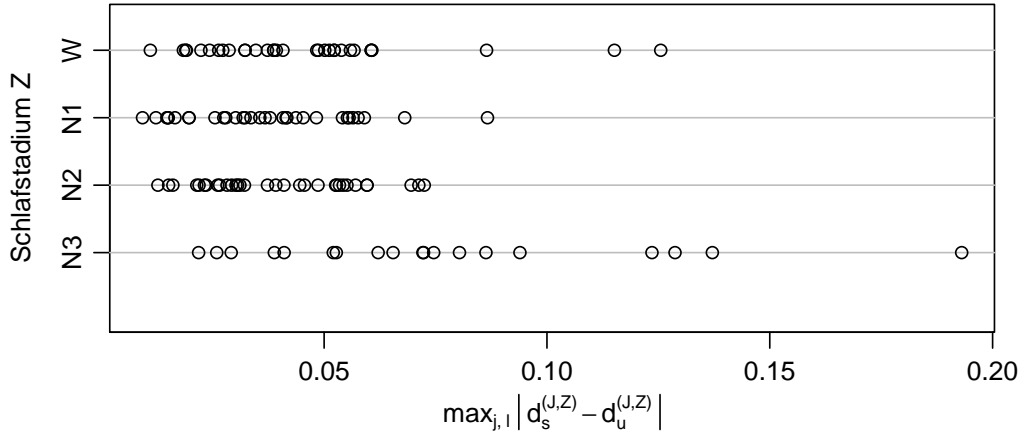


Abbildung 3.4: Maximale Abweichung der Differenzen  $d_A, d_B, d_C, d_D$  voneinander pro Proband getrennt nach Schlafstadium. Jeder Punkt entspricht  $\max_{s,u} |d_s - d_u|$  für einen festen Probanden und das jeweilige Schlafstadium.

höchstens 0.04 und in 90% aller Fälle jeweils um ca. höchstens 0.06. In N3 lassen sich vermehrt größere maximale Abweichungen  $\max_{s,u} |d_s^{(J,Z)} - d_u^{(J,Z)}|$  zwischen den vier Differenzen eines Probanden  $J$  finden, in der Höhe von bis zu 0.2. Der Median liegt für N3 bei ca. 0.07.

Insbesondere für die Schlafstadien W, N1 und N2 liegen  $d_A^{(J,Z)}, d_B^{(J,Z)}, d_C^{(J,Z)}, d_D^{(J,Z)}$  in vielen Fällen also relativ dicht beieinander. Daraus ergibt sich die zweite Forderung an das SMI-Modell:

Für jeden Probanden  $J$  und jedes Schlafstadium  $Z$  gelte

$$d = d^{(J,Z)} := d_A^{(J,Z)} = d_B^{(J,Z)} = d_C^{(J,Z)} = d_D^{(J,Z)}.$$

D.h. an das SMI-Modell wird für die Übergangswahrscheinlichkeiten  $p_{ss}$  von einem Zustand  $s$  zu sich selbst die Bedingung

$$\exists d \in [0, 1] : \quad p_{ss} = p_{rs} + d \quad \forall r, s \in S, r \neq s, \quad (3.2)$$

gestellt, wobei nach Bedingung (3.1) für festes  $s$  gilt, dass  $p_{rs} = p_{ts}$  ist für alle  $r, t \neq s$ .

**Beispiel 3.1.** Zur Verdeutlichung der beiden Beobachtungen sind in Abbildung 3.5 exemplarisch für die Probanden 2 und 5 im Schlafstadium N2 die Werte der oben erwähnten  $\hat{p}_{rs}^{(v)}, \hat{p}_{rs}^{(1)}$  und  $d_s$  eingetragen.

Proband 2 dient dabei als Beispiel für einen Probanden, bei dem die Bedingungen (3.1) und (3.2) annähernd erfüllt werden. Mit Proband 5 wird ein Beispiel gezeigt, bei dem die beiden Bedingungen nur grob erfüllt werden.

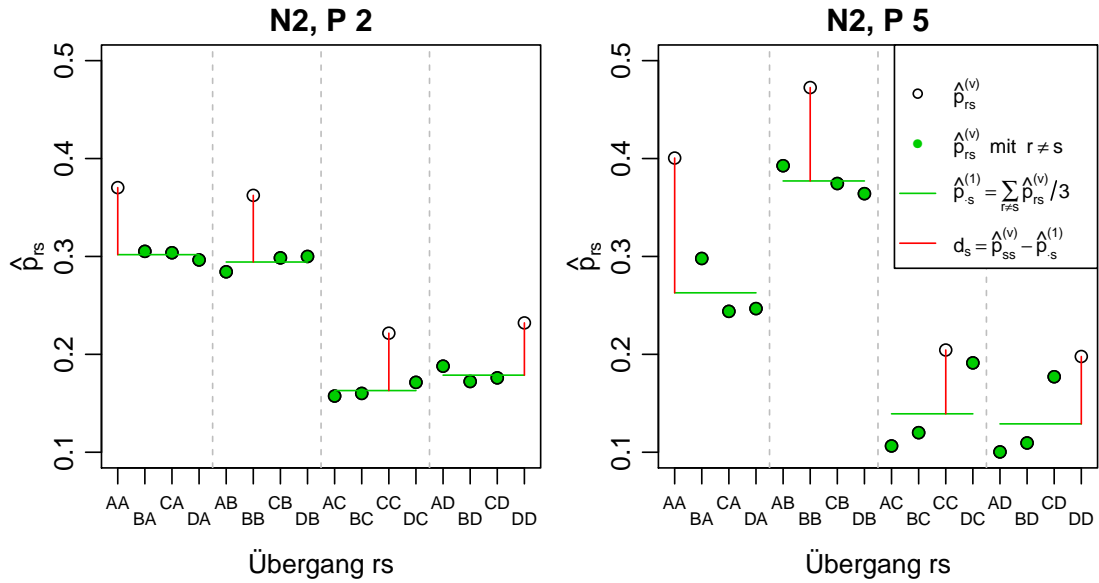


Abbildung 3.5: Geschätzte Übergangswahrscheinlichkeiten im Schlafstadium N2 für Proband 2 (links) und Proband 5 (rechts). Die grünen Linien geben die Höhe von  $\hat{p}_{\cdot s}$  an, die Längen der roten Linien entsprechen  $d_s$ . Die Kreise markieren  $\hat{p}_{rs}^{(v)}$ , die Kreuze  $\hat{p}_{rs} = \hat{p}_{\cdot s} + \delta_{r,s} d$  (1. Modell).

Die Schätzer  $\hat{p}_{rs}^{(v)}$  werden in Form von Kreisen in Abbildung 3.5 dargestellt. Für jedes feste  $s \in S$  sind die vier Übergänge  $(r, s)$  mit  $r \in S$  durch gestrichelte vertikale Linien gruppiert. Zudem sind die Kreise der drei  $\hat{p}_{rs}^{(v)}$  mit  $r \neq s$  einer jeden Gruppe grün ausgefüllt. Für jedes  $s$  werden sie zur Berechnung von  $\hat{p}_{\cdot s}^{(1)}$  (Mittelwertbildung) herangezogen. Die berechneten  $\hat{p}_{\cdot s}^{(1)}$  werden jeweils durch eine grüne horizontale Linie markiert.

Bei Proband 2 sind für jedes feste  $s \in S$  die drei  $\hat{p}_{rs}^{(v)}$  mit  $r \neq s$  ungefähr gleich groß, sodass sie nur schwach um ihren Mittelwert streuen. Hier wird die Bedingung (3.1) also annähernd erfüllt. Dagegen weichen für Proband 5 die entsprechenden Werte der  $\hat{p}_{rs}^{(v)}$  relativ stark voneinander ab mit der Konsequenz, dass sie stärker um ihren Mittelwert streuen. Beispielsweise unterscheiden sich  $\hat{p}_{AC}^{(v),(5,N2)}$  und  $\hat{p}_{DC}^{(v),(5,N2)}$  um etwa 0.085.

Die roten Linien in Abbildung 3.5 entsprechen ihrer Länge nach den Differenzen

$$d_s^{(J,Z)} = \hat{p}_{ss}^{(v),(J,Z)} - \hat{p}_{\cdot s}^{(1),(J,Z)}.$$

Im Idealfall, der für die Bedingung (3.2) sprechen würde, sind alle roten Linien gleich lang. Dies ist für Proband 2 näherungsweise der Fall, denn es ist

$$\max_{s,u \in S} |d_s^{(2,N2)} - d_u^{(2,N2)}| \approx 0.015.$$

Für Proband 5 unterscheiden sich die Längen der roten Linien um bis zu

$$\max_{s,u \in S} |d_s^{(5,N2)} - d_u^{(5,N2)}| \approx 0.072.$$



Die Bedingung (3.2) ist also für Proband 5 nur grob erfüllt, wobei jedoch das Ergebnis für eine andere Definition von  $\hat{p}_s^{(1)}$  besser ausfallen kann.

### 3.2 Verallgemeinerung des SMI-Modells und mögliches Hintergrundmodell

Das SMI-Modell (Modell 1) für die Sequenzen  $x^{(J,Z)}$  von Zuständen wird nun auf beliebige endliche Zustandsräume  $S$  verallgemeinert formuliert und es wird gezeigt, dass das (verallgemeinerte) SMI-Modell per Definition die Bedingungen (3.1) und (3.2) erfüllt.

Danach wird der Frage nachgegangen, wie eine Markov-Kette  $X$  mit den durch das SMI-Modell beschriebenen Eigenschaften zustande kommen könnte. Dafür wird ein mögliches Hintergrundmodell für  $X$  angegeben und bewiesen, dass eine durch das Hintergrundmodell konstruierte Markov-Kette  $X$  die geforderten Eigenschaften des SMI-Modells besitzt. Mithilfe des Hintergrundmodells lassen sich die Parameter  $p_r$ ,  $r \in S$ , und  $\pi$  des SMI-Modells anschaulich interpretieren, sodass Hinweise für sinnvolle Schätzer geliefert werden.

Den Abschluss bilden Ergebnisse zur stationären Verteilung und zur Irreduzibilität der Markov-Kette  $X$ , zu den Erwartungswerten  $\mathbb{E}[n_r]$  und  $\mathbb{E}[n_{rs}]$  sowie Bemerkungen zu den Parametern.

Für die Verallgemeinerung des SMI-Modells werden die Zustände der Markov-Kette  $X$  nicht mehr mit den Buchstaben  $A, B, C, D$  bezeichnet, sondern zur einfacheren Handhabung durchnummeriert, sodass

$$S := \{1, 2, \dots, a\}$$

der verallgemeinerte endliche Zustandsraum ist, wobei die Anzahl  $a \in \mathbb{N}$  der paarweise verschiedenen Zustände fest gegeben sei. Durch die Zuordnung

$$A \leftrightarrow 1, B \leftrightarrow 2, C \leftrightarrow 3, D \leftrightarrow 4$$

lässt sich das verallgemeinerte SMI-Modell auf die gegebenen Daten anwenden.

**Modell 2** (SMI-Modell, mit Zustandsraum  $S = \{1, \dots, a\}$ ). Die Sequenz

$$\left( x_0^{(J,Z)}, \dots, x_n^{(J,Z)} \right)$$

werde aufgefasst als die ersten  $n + 1 = n^{(J,Z)} + 1$  Zustände

$$(X_0, \dots, X_n)$$

einer Markov-Kette  $(X_i)_{i \in \mathbb{N}_0}$  mit Zustandsraum

$$S := \{1, \dots, a\}$$

und der Startverteilung

$$p^\top = (p_1, \dots, p_a)$$

sowie der Übergangsmatrix  $P = (p_{rs})_{r,s \in S}$ , gegeben durch

$$p_{rs} := \begin{cases} \pi p_s & , \text{ falls } r \neq s, \\ \pi p_s + 1 - \pi & , \text{ falls } r = s, \end{cases}$$

wobei

$$\pi \in (0, 1]$$

und

$$p_1, \dots, p_a \in (0, 1) \text{ mit } \sum_{r=1}^a p_r = 1.$$

**Bemerkung 3.2.** Die Übergangswahrscheinlichkeiten  $p_{rs}$ ,  $r, s \in S$ , erfüllen die Bedingung (3.1), denn für  $s \in S$  hängen die  $p_{rs}$  mit  $r \neq s$  nur von  $s$  ab, sodass gilt

$$\forall s \in S: \quad p_{rs} = \pi p_s = p_{ts} \quad \forall r, t \in S, r, t \neq s.$$

Weiterhin erfüllen die Übergangswahrscheinlichkeiten  $p_{rs}$  auch Bedingung (3.2) wegen

$$p_{ss} = \pi p_s + 1 - \pi = p_{rs} + d \quad \forall r, s \in S, r \neq s,$$

mit  $d := 1 - \pi \in [0, 1)$ .

Im nächsten Schritt wird ein mögliches Hintergrundmodell konstruiert, das eine Erklärung für das Zustandekommen des SMI-Modells mit seiner speziellen Struktur der Übergangswahrscheinlichkeiten bietet.

Dieses Hintergrundmodell setzt sich zusammen aus vier Konstruktionsschritten:

1. Der Einteilung der Zeitachse in zufällige Intervalle,
2. der zufälligen Zuordnung von Zuständen zu den Intervallen (Markierung der Intervalle),
3. der Bestimmung von zufälligen Abtastungszeitpunkten und
4. dem Ablesen des Zustandes des jeweils aktuellen Intervalls zu den Abtastungszeitpunkten

(siehe Abbildung 3.1 auf Seite 19 für ein Beispiel).

Damit sich aus dem Hintergrundmodell das SMI-Modell ableiten lässt, müssen technische Annahmen an die zufälligen Elemente der Konstruktion gestellt werden:

1. Die zufälligen Intervalle seien durch einen homogenen Poisson'schen Punktprozess auf  $\mathbb{R}^+$  mit Intensität  $\lambda_1 > 0$  gegeben,
2. die zufälligen Abtastungszeitpunkte seien durch einen homogenen Poisson'schen Punktprozess auf  $\mathbb{R}^+$  mit Intensität  $\lambda_2 > 0$  gegeben,

3. beide Prozesse seien unabhängig voneinander und
4. für jedes Intervall erfolge die Zuordnung des Zustandes  $r \in S$  mit konstanter Wahrscheinlichkeit  $p_r$  unabhängig voneinander und unabhängig von beiden Prozessen.

Für die folgende formale Beschreibung des Hintergrundmodells wird die Tatsache ausgenutzt, dass für die durch einen homogenen Poisson'schen Punktprozess  $Z = \sum_i \delta_{P_i}$  auf  $\mathbb{R}^+$  mit Intensität  $\lambda > 0$  gegebenen Punkte  $P_1, P_2, \dots$  gilt, dass die Wartezeiten  $P_1$  und  $P_i - P_{i-1}$  unabhängig  $\exp(\lambda)$ -verteilt sind,  $i \geq 2$  (siehe z. B. Meintrup und Schäffler, 2005).

**Modell 3** (SMI-Modell, Hintergrundmodell). Sei  $a \in \mathbb{N}$  die Anzahl der paarweise verschiedenen Zustände und  $S := \{1, \dots, a\}$  die Menge dieser Zustände. Sei  $(\tau_j)_{j \in \mathbb{N}}$  eine Folge unabhängiger  $\exp(\lambda_1)$ -verteilter Zufallsvariablen. Seien  $T_0 := 0$  und  $T_j := T_{j-1} + \tau_j$  für alle  $j \in \mathbb{N}$ . Die zufälligen Intervalle  $I_1, I_2, \dots$  seien definiert durch

$$I_j := [T_{j-1}, T_j) \quad \forall j \in \mathbb{N}.$$

Es seien Wahrscheinlichkeiten  $p_1, \dots, p_a \in (0, 1)$  gegeben mit  $\sum_{r \in S} p_r = 1$  und dazu eine Folge  $(M_j)_{j \in \mathbb{N}}$  unabhängiger identisch verteilter Zufallsvariablen mit

$$\mathbb{P}(M_j = r) = p_r \quad \forall r \in S, j \in \mathbb{N}.$$

Die Zufallsvariable  $M_j$  steht für den Zustand, der dem Intervall  $I_j$  zugeordnet wird (Markierung des Intervalls  $I_j$ ).

Sei weiterhin  $(\beta_i)_{i \in \mathbb{N}_0}$  eine Folge unabhängiger  $\exp(\lambda_2)$ -verteilter Zufallsvariablen und  $B_0 := \beta_0$  sowie  $B_i := B_{i-1} + \beta_i$  für alle  $i \in \mathbb{N}$ . Dadurch werden die zufälligen Abtastungszeitpunkte  $B_0, B_1, \dots$  definiert.

Sämtliche Zufallsvariablen  $(\tau_j)_{j \in \mathbb{N}}$ ,  $(M_j)_{j \in \mathbb{N}}$  und  $(\beta_i)_{i \in \mathbb{N}_0}$  seien unabhängig voneinander. Ist  $j \in \mathbb{N}$  der Index des Intervalls mit  $B_i \in I_j$  für ein  $i \in \mathbb{N}_0$ , wird der zugehörige Zustand  $M_j$  abgelesen. Die Folge  $(X_i)_{i \in \mathbb{N}_0}$  der abgelesenen Zustände ist somit für alle  $i \in \mathbb{N}_0$  definiert durch

$$X_i := \sum_{j \in \mathbb{N}} \mathbf{1}_{\{B_i \in I_j\}} M_j.$$

Die Intervalle  $I_j$  und die zugehörigen Zustände  $M_j$  seien für alle  $j \in \mathbb{N}$  nicht beobachtbar. Die Abtastungszeitpunkte  $B_i$  und die Zustände  $X_i$  zu diesen Abtastungszeitpunkten seien in diesem Modell beobachtbar für alle  $i \in \mathbb{N}_0$ .

Das Beispiel zum Hintergrundmodell aus Abbildung 3.1 wird im folgenden Beispiel aufgegriffen und erweitert.

**Beispiel 3.3.** In Abbildung 3.6 wird das Beispiel zum Hintergrundmodell aus Abbildung 3.1 ergänzt um die Bezeichnungen der nun eingeführten Variablen  $I_j$  (Intervalle),  $M_j$  (Markierungen der jeweiligen  $I_j$ ) und  $B_i$  (Abtastungszeitpunkte). Zur besseren Übersicht wurden dabei die Intervalle und die Abtastungszeitpunkte jeweils auf einer eigenen Zeitachse dargestellt.

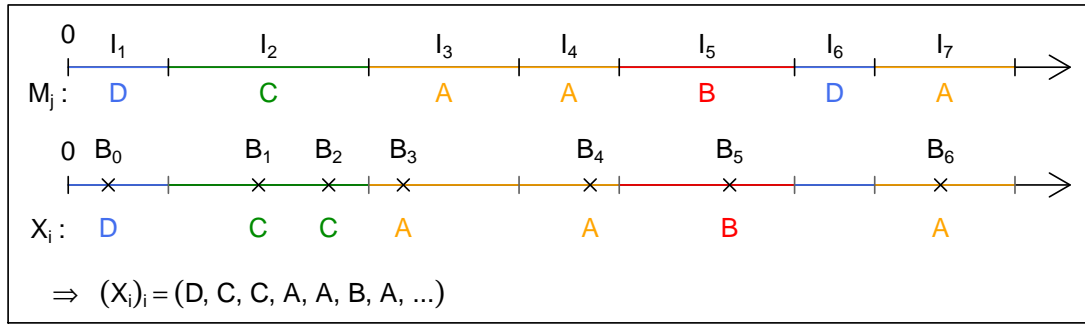


Abbildung 3.6: Exemplarische Darstellung des Hintergrundmodells mit Variablenbezeichnungen

In Abbildung 3.6 liegt  $B_0$  in  $I_1$  und somit erhält  $X_0$  den Wert  $M_1 = D$ . Dementsprechend ist  $X_1 = C$ , da  $B_1 \in I_2$  gilt und  $M_2 = C$  ist. Weil  $B_2$  im selben Intervall wie  $B_1$  liegt, muss  $X_2 = X_1 = C$  gelten.

Es ist auch möglich, dass zwei hintereinander folgende  $X_i$  und  $X_{i+1}$  gleich sind, wenn zwei verschiedenen Intervallen  $I_{j_i}$  und  $I_{j_{i+1}}$ , in denen  $B_i$  bzw.  $B_{i+1}$  liegen, der gleiche Zustand zugeordnet wurde, d.h. wenn  $M_{j_i} = M_{j_{i+1}}$  erfüllt ist. Dies ist in diesem Beispiel der Fall für  $X_3$  und  $X_4$ .

Zudem kann es Intervalle  $I_j$  geben, in denen keines der  $B_i$  für  $i \in \mathbb{N}_0$  liegt, wie im Fall von Intervall  $I_6$ . Der Zustand  $M_j$  des Intervalls  $I_j$  wird dann überhaupt nicht abgetastet.

**Bemerkung 3.4.** Nach den Voraussetzungen des Hintergrundmodells sind zwar  $(B_i)_{i \in \mathbb{N}_0}$  und  $(X_i)_{i \in \mathbb{N}_0}$  beobachtbar, sie liegen für diese Arbeit jedoch nicht als gegebene Daten tatsächlich vor. Die einzigen gegebenen Daten sind für alle  $r, s \in S$  die absoluten Häufigkeiten der Übergänge

$$n_{rs} := \sum_{i=0}^n \mathbf{1}_{\{X_i=r \wedge X_{i+1}=s\}}$$

in  $(X_0, \dots, X_n)$  von  $r$  nach  $s$  für ein  $n \in \mathbb{N}$ .

Es wird nun gezeigt, dass die durch das Hintergrundmodell gewonnene Folge  $(X_i)_{i \in \mathbb{N}_0}$  gerade die Eigenschaften des SMI-Modells erfüllt.

**Proposition 3.5.** Die durch das Hintergrundmodell 3 beschriebene Folge  $(X_i)_{i \in \mathbb{N}_0}$  ist eine Markov-Kette mit Zustandsraum  $S = \{1, \dots, a\}$ , Startverteilung  $(p_1, \dots, p_a)$  und Übergangsmatrix  $P = (p_{rs})_{r,s \in S}$ , gegeben durch

$$p_{rs} = \begin{cases} \pi p_s & , \text{ falls } r \neq s, \\ \pi p_s + 1 - \pi & , \text{ falls } r = s, \end{cases}$$

wobei

$$\pi := \mathbb{P}(B_{i+1} \in I_k \text{ für ein } k > j \mid B_i \in I_j)$$

konstant ist für alle  $i \in \mathbb{N}_0, j \in \mathbb{N}$ . Insbesondere gilt

$$\pi = \frac{\lambda_1}{\lambda_1 + \lambda_2} \in (0, 1).$$

Bevor die Proposition bewiesen wird, wird in der nachstehenden Bemerkung der Parameter  $\pi$  genauer betrachtet.

**Bemerkung 3.6.** (i) Der Parameter  $\pi$  lässt sich interpretieren als die Wahrscheinlichkeit dafür, dass zwischen den Abtastungszeitpunkten  $B_i$  und  $B_{i+1}$  für festes  $i \in \mathbb{N}_0$  mindestens ein Intervallwechsel stattfindet, d.h.

$$\pi = \mathbb{P}(\text{Intervallwechsel zwischen } B_i \text{ und } B_{i+1}).$$

Entsprechend gibt  $1 - \pi = \mathbb{P}(B_{i+1} \in I_j \mid B_i \in I_j)$  die Wahrscheinlichkeit dafür an, dass beide Abtastungszeitpunkte  $B_i$  und  $B_{i+1}$  im selben Intervall liegen, gegeben  $B_i \in I_j$ . In diesem Fall gilt  $X_i = X_{i+1}$ .

- (ii) Das  $\pi$  aus dem Hintergrundmodell liegt in  $(0, 1)$ , während in den Modellen 1 und 2 auch  $\pi = 1$  zugelassen ist. Der Fall  $\pi = 1$  lässt sich jedoch als Grenzfall des Hintergrundmodells interpretieren mit  $\lambda_1 \rightarrow \infty$  und endlichem  $\lambda_2$ .

Da beim Hintergrundmodell (Modell 3) zufällige markierte Intervalle  $((I_j, M_j))_{j \in \mathbb{N}}$  abgetastet werden zu Abtastungszeitpunkten  $(B_i)_{i \in \mathbb{N}_0}$ , werden die nach Proposition 3.5 (bis auf den Fall  $\pi = 1$  bzw. bis auf die Mächtigkeit von  $S$ ) äquivalenten Modelle 1, 2 und 3 in dieser Arbeit alle „Sampled Marked Intervals-Modell“ (SMI-Modell) genannt. Wenn auf die spezielle Konstruktion des Modells 3 verwiesen werden soll, wird im weiteren Verlauf der Arbeit jedoch weiterhin die Bezeichnung „Hintergrundmodell“ für das Modell 3 verwendet.

*Beweis (Proposition 3.5).* Wegen der Unabhängigkeit von  $(\tau_i)_{i \in \mathbb{N}}$ ,  $(\beta_i)_{i \in \mathbb{N}_0}$  und  $(M_i)_{i \in \mathbb{N}}$  voneinander, ist es nicht von Bedeutung für die Startverteilung von  $(X_i)_{i \in \mathbb{N}_0}$ , in welchem Intervall  $B_0$  liegt. Bedingt darauf, dass  $B_0 \in I_j$  gilt, folgt  $X_0 = M_j$  und somit

$$\mathbb{P}(X_0 = r \mid B_0 \in I_j) = \mathbb{P}(M_j = r) = p_r \quad \forall r \in S.$$

Mit dem Satz von der totalen Wahrscheinlichkeit ergibt sich die Behauptung für die Startverteilung.

Als Nächstes wird gezeigt, dass  $1 - \pi = \mathbb{P}(B_{i+1} \in I_j \mid B_i \in I_j)$  für alle  $i \in \mathbb{N}_0, j \in \mathbb{N}$  konstant ist.

Seien  $i \in \mathbb{N}_0, j \in \mathbb{N}$ . Ist  $B_i \in I_j = [T_{j-1}, T_{j-1} + \tau_j)$  gegeben, so ist bekannt, dass

$$\tau_j > B_i - T_{j-1}$$

erfüllt wird. Bedingt auf dieses Ereignis und darauf, dass  $B_i$  und  $T_{j-1}$  bekannt sind, ist  $\tau_j - (B_i - T_{j-1}) \exp(\lambda_1)$ -verteilt (wie  $\tau_j$ ) wegen der Gedächtnislosigkeit der Exponential-Verteilung. Bildlich gesprochen lässt sich also eine  $\exp(\lambda_1)$ -verteilte Zufallsvariable  $\tau'_j$  ab dem Zeitpunkt  $B_i$  starten, um die rechte Grenze von  $I_j$  zu ermitteln. Zum selben Zeitpunkt startet ebenfalls die  $\exp(\lambda_2)$ -verteilte Zufallsvariable  $\beta_{i+1}$ . Das Ereignis  $\{B_{i+1} \in I_j\}$  tritt nun genau dann ein, wenn  $\beta_{i+1} < \tau'_j$  ist.

Formal bedeutet das

$$\begin{aligned} 1 - \pi &= \mathbb{P}(B_{i+1} \in I_j \mid B_i \in I_j) \\ &= \mathbb{P}(B_{i+1} < T_j \mid T_{j-1} \leq B_i < T_j) \\ &= \mathbb{P}(B_i + \beta_{i+1} < T_{j-1} + \tau_j \mid T_{j-1} \leq B_i < T_{j-1} + \tau_j) \\ &= \mathbb{P}(B_i - T_{j-1} + \beta_{i+1} < \tau_j \mid 0 \leq B_i - T_{j-1} < \tau_j) \end{aligned}$$

und mit dem Satz von der totalen Wahrscheinlichkeit

$$\begin{aligned} 1 - \pi &= \mathbb{P}(B_i - T_{j-1} + \beta_{i+1} < \tau_j \mid 0 \leq B_i - T_{j-1} < \tau_j) \\ &= \int_0^\infty \int_0^\infty \mathbb{P}(c + b < \tau_j \mid c < \tau_j, c = B_i - T_{j-1}, \underbrace{b = \beta_{i+1}}_{\substack{\text{unabh.} \\ \text{von } \tau_j}}) \mathbb{P}(\beta_{i+1} \in db) \mathbb{P}(B_i - T_{j-1} \in dc) \\ &= \int_0^\infty \int_0^\infty \mathbb{P}(b < \tau_j \mid c = B_i - T_{j-1}) \mathbb{P}(\beta_{i+1} \in db) \mathbb{P}(B_i - T_{j-1} \in dc) \\ &= \int_0^\infty \mathbb{P}(b < \tau_j \mid \underbrace{0 \leq B_i - T_{j-1}}_{\text{unabh. von } \tau_j}) \mathbb{P}(\beta_{i+1} \in db) \\ &= \int_0^\infty \mathbb{P}(b < \tau_j) \mathbb{P}(\beta_{i+1} \in db), \end{aligned}$$

wobei die Gedächtnislosigkeit der Exponential-Verteilung in der dritten Zeile eingeht.

Da  $\tau_j \exp(\lambda_1)$ -verteilt ist für alle  $i \in \mathbb{N}_0$  und  $\beta_{i+1} \exp(\lambda_2)$ -verteilt ist für alle  $j \in \mathbb{N}$ , hängt der letzte Ausdruck und damit  $1 - \pi$  nicht von  $i$  oder  $j$  ab. Genauer hängt  $1 - \pi$  und somit auch  $\pi$  nur von den Parametern  $\lambda_1$  und  $\lambda_2$  ab, wie die folgende Rechnung zeigt:

$$\begin{aligned} 1 - \pi &= \int_0^\infty \mathbb{P}(b < \tau_j) \mathbb{P}(\beta_{i+1} \in db) \\ &= \int_0^\infty e^{-\lambda_1 b} \lambda_2 e^{-\lambda_2 b} db \\ &= \frac{\lambda_2}{\lambda_1 + \lambda_2} \int_0^\infty (\lambda_1 + \lambda_2) e^{-(\lambda_1 + \lambda_2)b} db \\ &= \frac{\lambda_2}{\lambda_1 + \lambda_2}. \end{aligned}$$

Daraus folgt direkt

$$\pi = \frac{\lambda_1}{\lambda_1 + \lambda_2}.$$

Da  $\lambda_1, \lambda_2 > 0$  sind, gilt außerdem  $0 < \pi < 1$ .

Nun sind noch die Markov-Eigenschaft zu zeigen und die Übergangswahrscheinlichkeiten  $p_{rs}$  zu bestimmen.

Mit Wahrscheinlichkeit  $\pi$  ist  $B_i$  mit  $i > 0$  der erste Abtastungszeitpunkt in einem Intervall  $I_j$  für ein  $j \in \mathbb{N}$ , d.h.  $B_i \in I_j$ , aber  $B_{i-1} \notin I_j$ . Dann nimmt  $M_j$  und somit auch  $X_i$  unabhängig von der Vergangenheit den Wert  $r$  mit Wahrscheinlichkeit  $p_r$  an, also unabhängig von  $B_k, X_k, I_l, M_l$  für  $k = 0, \dots, i-1$  und  $l = 1, \dots, j-1$ .

Mit Wahrscheinlichkeit  $1 - \pi$  findet zwischen  $B_{i-1}$  und  $B_i$  kein Intervallwechsel statt. Dann gilt  $X_{i-1} = X_i$ . Ist zudem gegeben, dass auch ein  $B_k$  mit  $k < i-1$  im selben Intervall liegt wie  $B_{i-1}$  und  $B_i$ , liefert dies keine zusätzliche Information für  $X_i$ . Von allen  $X_k$ , für die die zugehörigen  $B_k$  mit  $k < i-1$  in einem vorangegangenen Intervall liegen, ist  $X_i$  unabhängig.

Insgesamt gilt also

$$\mathbb{P}(X_i = m_i | X_{i-1} = m_{i-1}, \dots, X_0 = m_0) = \mathbb{P}(X_i = m_i | X_{i-1} = m_{i-1})$$

für alle  $i \in \mathbb{N}$  und  $m_0, \dots, m_i \in S$ . Somit ist  $(X_i)_{i \in \mathbb{N}_0}$  eine Markov-Kette 1. Ordnung.

Zum Abschluss des Beweises werden nun die Übergangswahrscheinlichkeiten bestimmt. Es sei das Ereignis

$$A_i := \{\text{Intervallwechsel zwischen } B_{i-1} \text{ und } B_i\} \quad \forall i \in \mathbb{N}$$

definiert und  $A_i^c$  sein Gegenereignis. Die Ereignisse haben die Wahrscheinlichkeit

$$\mathbb{P}(A_i) = \pi \quad \text{bzw.} \quad \mathbb{P}(A_i^c) = 1 - \pi.$$

Wie oben näher ausgeführt, sind  $X_{i-1}$  und  $X_i$  unabhängig voneinander, bedingt darauf, dass zwischen  $B_{i-1}$  und  $B_i$  mindestens ein Intervallwechsel stattfindet. Außerdem ist  $(M_j)_{j \in \mathbb{N}}$  unabhängig von  $(A_i)_{i \in \mathbb{N}_0}$ . Somit gilt

$$\mathbb{P}(X_i = s | X_{i-1} = r, A_i) = \mathbb{P}(X_i = s | A_i) = \mathbb{P}(X_i = s).$$

Findet zwischen  $B_{i-1}$  und  $B_i$  dagegen kein Intervallwechsel statt, gilt  $X_{i-1} = X_i$ . Das führt zu

$$\mathbb{P}(X_i = s | X_{i-1} = r, A_i^c) = \mathbb{P}(X_{i-1} = s | X_{i-1} = r).$$

Daraus ergibt sich für die Übergangswahrscheinlichkeiten

$$\begin{aligned} p_{rs} &= \mathbb{P}(X_i = s | X_{i-1} = r) \\ &= \mathbb{P}(A_i) \mathbb{P}(X_i = s | X_{i-1} = r, A_i) + \mathbb{P}(A_i^c) \mathbb{P}(X_i = s | X_{i-1} = r, A_i^c) \\ &= \pi \mathbb{P}(X_i = s) + (1 - \pi) \mathbb{P}(X_{i-1} = s | X_{i-1} = r) \\ &= \begin{cases} \pi p_s & , \text{ falls } r \neq s, \\ \pi p_s + 1 - \pi & , \text{ falls } r = s. \end{cases} \end{aligned}$$

□

Aus der Konstruktion des Hintergrundmodells folgt direkt, dass

$$\mathbb{P}(X_i = r) = \mathbb{P}(M_{j_i} = r) = p_r$$

gilt, wobei  $j_i \in \mathbb{N}$  so gewählt sei, dass  $B_i \in I_{j_i}$  ist. Dies impliziert die folgende Proposition.

**Proposition 3.7.** *Für die Markov-Kette  $(X_i)_{i \in \mathbb{N}_0}$  des SMI-Modells ist die Startverteilung  $(p_1, \dots, p_a)$  die stationäre Verteilung.*

*Rechnerischer Beweis.* Auch ohne das Hintergrundmodell lässt sich die Stationarität nachweisen. Dafür ist zu zeigen, dass  $p = (p_1, \dots, p_a)^\top$  ein linker Eigenvektor der Übergangsmatrix  $P = (p_{rs})_{r,s \in S}$  zum Eigenwert 1 ist.

Für die Einträge von  $P$  gilt

$$p_{rs} = \pi p_s + \delta_{r,s}(1 - \pi),$$

wobei  $\delta_{r,s}$  das Kronecker-Delta sei. Damit lässt sich die Übergangsmatrix darstellen als  $P = \pi e p^\top + (1 - \pi)I$ , wobei  $e = (1, \dots, 1)^\top \in \mathbb{R}^a$  und  $I$  die  $(a \times a)$ -Einheitsmatrix sei. Wegen  $p^\top e = p_1 + \dots + p_a = 1$  folgt direkt

$$p^\top P = \pi p^\top e p^\top + (1 - \pi)p^\top = \pi p^\top + (1 - \pi)p^\top = p^\top.$$

□

**Korollar 3.8.** *Seien  $n \in \mathbb{N}$  und  $Z_{i,r} := \mathbf{1}_{\{X_i=r\}}$  für alle  $i \in \{0, \dots, n\}$  und  $r \in S$ . Für die in Kapitel 2 definierten Auftritts- bzw. Übergangshäufigkeiten*

$$n_r = \sum_{s \in S} n_{rs} = \sum_{i=0}^{n-1} Z_{i,r} \quad \text{und} \quad n_{rs} = \sum_{j=0}^{n-1} Z_{j,r} Z_{j+1,s}$$

gelten dann

$$\mathbb{E}[n_r] = n p_r \quad \text{und} \quad \mathbb{E}[n_{rs}] = n p_r p_{rs} \quad \forall r, s \in S.$$

*Beweis.* Seien  $r, s \in S$ . Da nach Proposition 3.7 der Vektor  $(p_1, \dots, p_a)$  die stationäre Verteilung von  $(X_i)_{i \in \mathbb{N}_0}$  ist, gilt

$$\mathbb{P}(X_i = r) = p_r \quad \forall i \in \mathbb{N}_0.$$

Somit folgen direkt

$$\mathbb{E}[n_r] = \sum_{i=0}^{n-1} \mathbb{E}[Z_{i,r}] = n \mathbb{P}(X_0 = r) = n p_r$$

und

$$\mathbb{E}[n_{rs}] = \sum_{j=0}^{n-1} \mathbb{E}[Z_{j,r} Z_{j+1,s}] = \sum_{j=0}^{n-1} \mathbb{P}(X_j = r, X_{j+1} = s) = n p_r p_{rs}.$$

□

**Proposition 3.9.** *Die Markov-Kette  $(X_i)_{i \in \mathbb{N}_0}$  des SMI-Modells ist irreduzibel.*

*Beweis.* Da  $\pi > 0$  und  $p_s > 0$  für alle  $s \in S$  gelten, sind alle  $p_{rs} = \pi p_s + \delta_{r,s}(1 - \pi) > 0$  und die Markov-Kette irreduzibel. □



### 3.2.1 Die Fälle $\pi = 1$ und $\pi = 0$

Angenommen, man betrachtet die Markov-Kette  $(X_i)_{i \in \mathbb{N}_0}$  mit Übergangsmatrix

$$P = \pi e p^\top + (1 - \pi)I$$

des SMI-Modells, dann ist  $\pi \in (0, 1]$ .

Der Fall  $\pi = 1$  führt zu  $P = e p^\top$ , also zu  $p_{rs} = p_s$  für alle  $r, s \in S$ . Eine Unterscheidung, ob  $r = s$  oder  $r \neq s$  gilt, muss nicht mehr durchgeführt werden. Daher sind die Übergangswahrscheinlichkeiten  $p_{rs}$  unabhängig von  $r$  und alle  $X_i$  sind unabhängig voneinander. Durch die Unabhängigkeit der  $X_i$  ist  $(n_1, \dots, n_a)$  multinomial( $n, p_1, \dots, p_a$ )-verteilt.

Im Hintergrundmodell ist  $\pi = 1$  nicht möglich (siehe Proposition 3.5 bzw. Bemerkung 3.6.(ii)). Übertragen auf die Bilder des Hintergrundmodells (als Grenzfall mit  $\lambda_1 \rightarrow \infty$  und endlichem  $\lambda_2$ ) bedeutet  $\pi = 1$  jedoch, dass je zwei verschiedene Abtastungszeitpunkte  $B_i$  und  $B_{i'}$  fast sicher in verschiedene Intervalle  $I_{j_i}$  und  $I_{j_{i'}}$  fallen, also mit  $j_i \neq j_{i'}$  für alle  $i, i' \in \mathbb{N}_0$ , denn mit Wahrscheinlichkeit  $\pi = 1$  findet zwischen  $B_i$  und  $B_{i+1}$  mindestens ein Intervallwechsel statt.

Der Fall  $\pi = 0$  ist weder im SMI-Modell noch im Hintergrundmodell möglich, denn es würde der entartete Fall  $P = I$  und somit  $X_i = X_j$  fast sicher für alle  $i, j \in \mathbb{N}_0$  auftreten. In Bildern des SMI-Modells, würde dies bedeuten, dass es nur ein Intervall  $I_1 = \mathbb{R}^+$  gibt, in das alle Abtastungszeitpunkte  $B_i$  fallen.

Für  $\pi = 0$  und  $\pi = 1$  ergeben sich also Markov-Ketten  $(X_i)_{i \in \mathbb{N}_0}$ , die sich auf besondere und einfache Weise verhalten. Aus diesem Grund werden im weiteren Verlauf dieser Arbeit bei Untersuchungen zum SMI-Modell die Fälle  $\pi = 0$  (bzw.  $\pi \rightarrow 0$ ) und  $\pi = 1$  besonders hervorgehoben, wenn sie zur besseren Anschauung beitragen.

### 3.2.2 Anzahl der Parameter

Als Ziel für diese Arbeit wurde in Kapitel 1 formuliert, die Struktur der Übergangswahrscheinlichkeit im vollen Modell zu beschreiben und dabei die Anzahl der Parameter ausgehend vom vollen Modell zu reduzieren. Die Übergangswahrscheinlichkeiten einer Markov-Kette mit 4 möglichen Zuständen weisen im vollen Modell  $4^2 - 4 = 12$  Parameter auf, denn insgesamt gibt es  $4^2 = 16$  Übergangswahrscheinlichkeiten  $p_{rs}$  und für jedes  $r \in \{A, B, C, D\}$  wird ein Parameter abgezogen wegen  $\sum_s p_{rs} = 1$ .

Dagegen lassen sich die Übergangswahrscheinlichkeiten des SMI-Modells für den Zustandsraum  $S = \{A, B, C, D\}$  durch 4 Parameter beschreiben, nämlich durch den Parameter  $\pi$  sowie o.B.d.A. durch die Parameter  $p_A, p_B, p_C$ , wobei  $p_D = 1 - (p_A + p_B + p_C)$  ist. Das SMI-Modell lässt sich also durch  $1 + 4 - 1 = 4$  Parameter beschreiben, wenn  $S$  die Mächtigkeit  $|S| = 4$  hat.

Für einen Zustandsraum  $S$  mit  $|S| = a \in \mathbb{N}$  ergeben sich analog  $a^2 - a$  Parameter für das volle Modell und  $a$  Parameter für das SMI-Modell. Durch das SMI-Modell wird also eine Reduktion um  $a^2 - 2a$  Parameter erreicht.

**Bemerkung 3.10.** Ist  $S = \{1, \dots, a\}$  der Zustandsraum, werden im weiteren Verlauf dieser Arbeit zur Beschreibung des SMI-Modells stets die Parameter  $p_1, \dots, p_a$  und  $\pi$  angegeben, obwohl o.B.d.A.  $p_a$  ersetzt werden kann durch

$$p_a = 1 - \sum_{r \neq a} p_r$$

und das SMI-Modell durch die Parameter  $p_1, \dots, p_{a-1}$  und  $\pi$  schon ausreichend beschrieben wird.

Um das SMI-Modell auf die gegebenen Daten anwenden zu können, müssen  $p_1, \dots, p_a$  und  $\pi$  geschätzt werden. Mit der Schätzung beschäftigt sich das folgende Kapitel.

## 4 Anpassung des SMI-Modells

In diesem Kapitel werden Schätzer für  $p_1, \dots, p_a$  und  $\pi$  für die Anpassung des SMI-Modells an die Daten eines Probanden  $J$  in einem Schlafstadium  $Z$  hergeleitet und Eigenschaften der Schätzer diskutiert.

Bei den Schätzern handelt es sich für  $p_r$  um

$$\hat{p}_r := \frac{n_r}{n} \quad \forall r \in S,$$

wobei  $n_r$  in Kapitel 2.2 definiert wurde durch

$$n_r = \sum_{s \in S} n_{rs} = \sum_{i=0}^{n-1} \mathbf{1}_{\{x_i=r\}} \quad \forall r \in S,$$

und für  $\pi$  um

$$\hat{\pi} := \hat{\pi}^{(g)} := \frac{n - \sum_r n_{rr}}{n \left( 1 - \sum_t \hat{p}_t^2 \right)} \wedge 1.$$

Die Sequenzen  $(x_0^{(J,Z)}, \dots, x_{n^{(J,Z)}}^{(J,Z)})$ , die den gegebenen Daten  $n_{rs} = n_{rs}^{(J,Z)}$  zugrunde liegen, werden gemäß dem SMI-Modell als erste  $n + 1 = n^{(J,Z)} + 1$  Zustände einer Markov-Kette

$$X^{(J,Z)} = X = (X_i)_{i \in \mathbb{N}_0}$$

aufgefasst. Die Überlegungen in diesem Kapitel zu den Schätzern werden sich stets auf diese Markov-Kette  $X$  bzw. auf die weiteren Zufallsvariablen des Hintergrundmodells (Modell 3),  $(I_j)_{j \in \mathbb{N}}$ ,  $(M_j)_{j \in \mathbb{N}}$  und  $(B_i)_{i \in \mathbb{N}_0}$ <sup>10</sup>, beziehen.

Beim Schätzen der Parameter ist zu beachten, dass lediglich die Übergangshäufigkeiten  $n_{rs}$  in den ersten  $n + 1$  Zuständen der Markov-Kette  $(X_i)_{i \in \mathbb{N}_0}$  gegeben sind, nicht jedoch  $(I_j)_{j \in \mathbb{N}}$ ,  $(M_j)_{j \in \mathbb{N}}$ ,  $(B_i)_{i \in \mathbb{N}_0}$  oder  $(X_i)_{i \in \mathbb{N}_0}$  selbst. Dadurch sind Informationen nicht verfügbar, die für das Schätzen hilfreich wären, etwa ob zwischen zwei hintereinander folgenden Abtastungszeitpunkten ein Intervallwechsel stattfindet oder nicht. Auf diese Einschränkung wird in den entsprechenden Unterkapiteln näher eingegangen.

In Kapitel 4.1 wird der Schätzer  $\hat{p}_r$  für  $p_r$  behandelt und in Kapitel 4.2 der Schätzer  $\hat{\pi}$  für  $\pi$ . In beiden Unterkapiteln werden zunächst die Ideen vorgestellt, die zu den jeweiligen Schätzern geführt haben.

Anschließend werden in den Kapiteln 4.1.1 bzw. 4.2.1 die Maximum-Likelihood-Schätzer (ML-Schätzer) für  $p_r$  bzw. für  $\pi$  bestimmt auf der Basis von Zufallsvariablen, die die für das Schätzen von  $p_r$  bzw.  $\pi$  relevanten Informationen aus dem Hintergrundmodell

<sup>10</sup> $(I_j)_{j \in \mathbb{N}}$ : zufällige Intervalle;  $(M_j)_{j \in \mathbb{N}}$ : den Intervallen zugeordnete zufällige Zustände;  $(B_i)_{i \in \mathbb{N}_0}$ : zufällige Abtastungszeitpunkte

enthalten. Konkret werden die ML-Schätzer ermittelt auf der Basis von Teilfolgen von  $(M_j)_{j \in \mathbb{N}}$  bzw. auf der Basis von  $(A_i)_{i \in \mathbb{N}}$  mit

$$A_i := \{\text{Intervallwechsel zwischen } B_{i-1} \text{ und } B_i\},$$

wenn die realisierten Werte der jeweiligen Zufallsvariablen gegeben wären. Die ML-Schätzer werden dann mit den oben definierten Schätzern  $\hat{p}_r$  bzw.  $\hat{\pi}$  verglichen, um festzustellen, wie sich die Tatsache bei den Schätzern  $\hat{p}_r$  und  $\hat{\pi}$  äußert, dass die relevanten Informationen aus dem Hintergrundmodell nicht verfügbar sind.

Zudem werden die Schätzer  $\hat{p}_r$  und  $\hat{\pi}$  in den Kapiteln 4.1.2 bzw. 4.2.2 für simulierte Datensätze  $(n_{rs})_{r,s \in S}$  berechnet, um die Abweichung der Schätzer von den wahren Werten  $p_r$  und  $\pi$  für endliche Datensätze zu untersuchen. Die Simulationen erfolgten dabei systematisch für verschiedene wahre Werte der Parameter  $p_1, \dots, p_a$  und  $\pi$ .

Das Kapitel 4.3 beschäftigt sich mit einer weiteren Möglichkeit für die Anpassung der Daten an das SMI-Modell, bei der die ML-Schätzer diesmal auf der Basis von  $(X_0, \dots, X_n)$  berechnet werden. Es wird sich zeigen, dass für ihre Berechnung nur die Übergangshäufigkeiten  $n_{rs}$  sowie der realisierte Wert  $x_0$  von  $X_0$  benötigt werden und wie mit dem unbekanntem  $x_0$  verfahren werden kann.

Wie aus den Schätzern  $\hat{p}_1, \dots, \hat{p}_a$  und  $\hat{\pi}$  die Schätzer  $\hat{p}_{rs}$  für  $p_{rs}$  in dieser Arbeit im Kontext des SMI-Modells ermittelt werden, wird in Kapitel 4.4 erläutert.

In Kapitel 4.5 werden zum Abschluss die Werte der Schätzer präsentiert, die für die Anpassung des SMI-Modells an die gegebenen Daten berechnet wurden.

#### 4.1 Schätzer für $p_r$

Für alle  $r \in S$  wird in dieser Arbeit für die Anpassung des SMI-Modells

$$\hat{p}_r := \frac{n_r}{n} = \frac{\sum_s n_{rs}}{n} = \frac{\sum_{i=0}^{n-1} \mathbf{1}_{\{X_i=r\}}}{n} \quad (4.1)$$

als Schätzer für  $p_r$  verwendet.

Dies ist ein naheliegender Schätzer, da  $(p_1, \dots, p_a)$  nach Proposition 3.7 die stationäre Verteilung von  $(X_i)_{i \in \mathbb{N}_0}$  ist und somit

$$\mathbb{P}(X_i = r) = p_r$$

gilt für alle  $i \in \mathbb{N}_0$ . Daraus folgt direkt, dass  $\hat{p}_r$  erwartungstreu ist.

**Korollar 4.1.** *Für alle  $r \in S$  ist der Schätzer  $\hat{p}_r = n_r/n$  ein erwartungstreuer Schätzer für  $p_r$ .*

*Beweis.* Sei  $r \in S$ . Nach Korollar 3.8 gilt  $\mathbb{E}[n_r] = np_r$  und damit

$$\mathbb{E}[\hat{p}_r] = \mathbb{E}[n_r/n] = p_r.$$

□

**Bemerkung 4.2.** 1. Für  $\hat{p}_r$  wird  $X_n$  nicht berücksichtigt, sondern nur  $X_0, \dots, X_{n-1}$ . Dies liegt an dem in Kapitel 2.2 beschriebenen Umstand, dass aus den gegebenen  $(n_{rs})_{r,s \in S}$  im Allgemeinen lediglich

$$n_r = n_{r\cdot} = \sum_{i=0}^{n-1} \mathbf{1}_{\{X_i=r\}} \quad \text{bzw.} \quad n_{\cdot r} = \sum_{i=1}^n \mathbf{1}_{\{X_i=r\}}$$

ermittelt werden kann, aber nicht die Auftrittshäufigkeit von  $r$  in  $(X_0, \dots, X_n)$ , die in Kapitel 2.2 bezeichnet wurde mit

$$n'_r = \sum_{i=0}^n \mathbf{1}_{\{X_i=r\}}.$$

2. Mit Ausnahme des Spezialfalls  $\pi = 1$  ist der Schätzer  $\hat{p}_r$  kein ML-Schätzer auf der Basis von  $(X_0, \dots, X_{n-1})$ , da die Zufallsvariablen  $X_0, \dots, X_{n-1}$  für  $\pi < 1$  abhängig voneinander sind. Auf den ML-Schätzer auf der Basis von  $(X_0, \dots, X_n)$  wird in Kapitel 4.3 näher eingegangen.<sup>11</sup>

#### 4.1.1 Vergleich mit ML-Schätzer auf der Basis von Hintergrundmodell-Variablen

Es wird nun ein ML-Schätzer für  $p_r$  hergeleitet, der sich ermitteln lässt, wenn Informationen aus dem Hintergrundmodell gegeben wären, die für das Schätzen von  $p_r$  relevant sind. Konkret wird der ML-Schätzer  $\hat{p}_r^{\text{ML,H}}$  bestimmt auf der Basis aller  $M_j$ , für die es ein  $B_i$  gibt, das in dem zu  $M_j$  gehörenden Intervall  $I_j$  liegt, d.h. auf der Basis von

$$\{M_j : j \in \mathbb{N}, \exists i \in \{0, \dots, n\} \text{ mit } B_i \in I_j\}.$$

Der ML-Schätzer  $\hat{p}_r^{\text{ML,H}}$  wird dann verglichen mit dem Schätzer  $\hat{p}_r$ , sodass deutlich wird, wie sich die Nicht-Verfügbarkeit der relevanten Informationen aus dem Hintergrundmodell auf den Schätzer  $\hat{p}_r$  auswirkt.

Die Wahrscheinlichkeit  $p_r$  wurde im Hintergrundmodell eingeführt als Wahrscheinlichkeit

$$p_r = \mathbb{P}(M_j = r),$$

dass dem Intervall  $I_j$  der Zustand  $r$  zugeordnet wird. Im Gegensatz zum ML-Schätzer auf der Basis von  $(X_0, \dots, X_n)$  lassen sich ML-Schätzer für  $p_r$  auf der Basis von gewissen Teilsequenzen  $(M_{j_k})_{k \in \{1, \dots, l\}}$  von  $(M_j)_{j \in \mathbb{N}}$  der Länge  $l \in \mathbb{N}$  auf einfache Weise durch die relative Häufigkeit

$$\frac{\sum_{k=1}^l \mathbf{1}_{\{M_{j_k}=r\}}}{l}$$

<sup>11</sup>Bei diesem ML-Schätzer wird  $X_n$  wieder mitberücksichtigt.

berechnen, da die  $M_j$  unabhängig voneinander und identisch verteilt sind und somit die absolute Auftrittshäufigkeit

$$\sum_{k=1}^l \mathbf{1}_{\{M_{j_k}=r\}}$$

des Zustandes  $r$  in der Teilsequenz gerade binomial( $l, p_r$ )-verteilt ist (siehe z. B. Kersting und Wakolbinger, 2008).

Für die Bestimmung des ML-Schätzers  $\hat{p}_r^{\text{ML,H}}$  wählt man die Teilsequenz  $(M_{j_k})_{k=1,\dots,l}$  so, dass für jedes  $k \in \{1, \dots, l\}$  ein  $i \in \{0, \dots, n\}$  mit  $B_i \in I_{j_k}$  existiert und umgekehrt für jedes  $i \in \{0, \dots, n\}$  ein  $k \in \{1, \dots, l\}$  mit  $B_i \in I_{j_k}$ . Das  $l$  gibt dann die Anzahl der durch  $(B_0, \dots, B_n)$  getroffenen Intervalle an.

Um die Anzahl der getroffenen Intervalle zu bestimmen, ist es nützlich, die Anzahl  $n_{rr}$  der Übergänge von  $r$  zu  $r$  zu zerlegen in

$$n_{rr} = \tilde{n}_{rr} + \bar{n}_{rr},$$

wobei  $\tilde{n}_{rr}$  die Anzahl der Übergänge mit Intervallwechsel und  $\bar{n}_{rr}$  ohne Intervallwechsel bezeichne. Die Anzahl der Intervalle, die von  $B_0, \dots, B_n$  getroffen wurden, ist mit dieser Notation gerade

$$(n+1) - \sum_{t \in S} \bar{n}_{tt} = 1 + \sum_{\substack{t, u \in S \\ t \neq u}} n_{tu} + \sum_{t \in S} \tilde{n}_{tt},$$

denn von  $B_0, \dots, B_n$  können maximal  $n+1$  paarweise verschiedene Intervalle getroffen werden und  $\sum_t \bar{n}_{tt}$  gibt die Anzahl der Paare  $(B_{i-1}, B_i)$  an mit  $i \in \{1, \dots, n\}$ , für die  $B_{i-1}$  und  $B_i$  im selben Intervall liegen.

Bedingt auf

$$(n+1) - \sum_t \bar{n}_{tt} = l$$

für ein  $l \in \mathbb{N}$  ist die Anzahl  $n'_r - \bar{n}_{rr}$  der getroffenen Intervalle, denen der Zustand  $r$  zugeordnet ist, binomial( $l, p_r$ )-verteilt und als ML-Schätzer  $\hat{p}_r^{\text{ML,H}}$  für  $p_r$  auf der Basis von

$$\{M_j : j \in \mathbb{N}, \exists i \in \{0, \dots, n\} \text{ mit } B_i \in I_j\}$$

erhält man

$$\hat{p}_r^{\text{ML,H}} = \frac{n'_r - \bar{n}_{rr}}{n+1 - \sum_t \bar{n}_{tt}}.$$

Wegen

$$\mathbb{E} \left[ \frac{n'_r - \bar{n}_{rr}}{n+1 - \sum_t \bar{n}_{tt}} \mid n+1 - \sum_t \bar{n}_{tt} = l \right] = \frac{lp_r}{l} = p_r$$

für alle  $l \in \mathbb{N}$  ist  $\hat{p}_r^{\text{ML,H}}$  erwartungstreu.

**Bemerkung 4.3.** Für  $\pi = 1$  tritt in Bildern des Hintergrundmodells<sup>12</sup> stets ein Intervallwechsel zwischen zwei Abtastungszeitpunkten auf. Dann ist  $\bar{n}_{tt} = 0$  für alle  $t \in S$  und

$$\hat{p}_r^{\text{ML,H}} = \frac{n'_r}{n+1}$$

ist der ML-Schätzer für  $p_r$  auf der Basis von  $(X_0, \dots, X_n)$ , denn die Anzahl der  $i \in \{0, \dots, n\}$  mit  $X_i = r$  ist in diesem Fall binomial( $n+1, p_r$ )-verteilt. Analog ist

$$\hat{p}_r = \frac{n_r}{n}$$

der ML-Schätzer für  $p_r$  auf der Basis von  $(X_0, \dots, X_{n-1})$ .

### Fazit

Im Gegensatz zum Schätzer  $\hat{p}_r$  werden beim ML-Schätzer  $\hat{p}_r^{\text{ML,H}}$  sowohl im Zähler als auch im Nenner nur die Übergänge mit Intervallwechsel mitgezählt.

Für den Schätzer  $\hat{p}_r$  sind die Informationen bezüglich der Intervallwechsel nicht verfügbar, sodass ein Zustand  $M_j = r$  eines Intervalls  $I_j$  doppelt oder häufiger gezählt werden kann. Wurde ein  $I_j$  bereits durch genau ein  $B_i$  getroffen, wird es mit Wahrscheinlichkeit  $1 - \pi$  auch durch  $B_{i+1}$  getroffen, sodass  $M_j$  ein zweites Mal gezählt wird. Wiederum mit Wahrscheinlichkeit  $1 - \pi$  wird  $I_j$  dann auch von  $B_{i+2}$  getroffen und  $M_j$  ein drittes Mal gezählt usw. Dies geschieht mit der gleichen Wahrscheinlichkeit unabhängig davon, ob  $M_j = r$  ist oder nicht, und unabhängig vom Intervall  $I_j$ . Dadurch erwartet man, dass sich die durch doppeltes und häufigeres Zählen eines Zustandes  $M_j$  verursachten möglichen Abweichungen des Schätzers  $\hat{p}_r$  vom wahren Wert  $p_r$  im Mittel ausgleichen. Dementsprechend ist der Schätzer  $\hat{p}_r$  erwartungstreu, wie in Korollar 4.1 gezeigt wurde. Je kleiner  $\pi$  jedoch ist, desto stärkeren Schwankungen ist  $\hat{p}_r$  intuitiv unterworfen im Vergleich zu  $\hat{p}_r^{\text{ML,H}}$ .

### 4.1.2 Überprüfung von $\hat{p}_r$ durch Simulation

Um das Verhalten der Schätzer  $\hat{p}_r$  zu untersuchen für typische Längen  $n^{(J,Z)}$  der für diese Arbeit gegebenen Sequenzen  $x^{(J,Z)}$ , wurden für  $a = 4$  und

$$p_1, p_2, p_3, p_4 \in \{0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7\}, \pi \in \{0.25, 0.50, 0.75, 1.00\}$$

für jede nicht redundante Kombination  $(p_1, p_2, p_3, p_4, \pi)$  mit  $\sum_r p_r = 1$  jeweils 1000 Markov-Ketten mit  $n = 1000$  Übergängen dem SMI-Modell gemäß simuliert. Für jede dieser simulierten Markov-Ketten wurden anschließend die Übergangshäufigkeiten  $(n_{rs})_{r,s \in S}$  bestimmt und damit  $\hat{p}_1, \hat{p}_2, \hat{p}_3, \hat{p}_4$  berechnet.

In Abbildung 4.1 wird eine Auswahl der ermittelten  $\hat{p}_1$  abgebildet. Dafür wurden für jedes  $p_1 \in \{0.1, 0.3, 0.5, 0.7\}$  und jedes  $\pi \in \{0.25, 0.50, 0.75, 1.00\}$  jeweils die Simulationen einer der oben beschriebenen Parameter-Kombinationen  $(p_1, p_2, p_3, p_4, \pi)$  herangezogen.

<sup>12</sup>Im Hintergrundmodell selbst ist der Fall  $\pi = 1$  nicht möglich (siehe Kapitel 3.2.1).

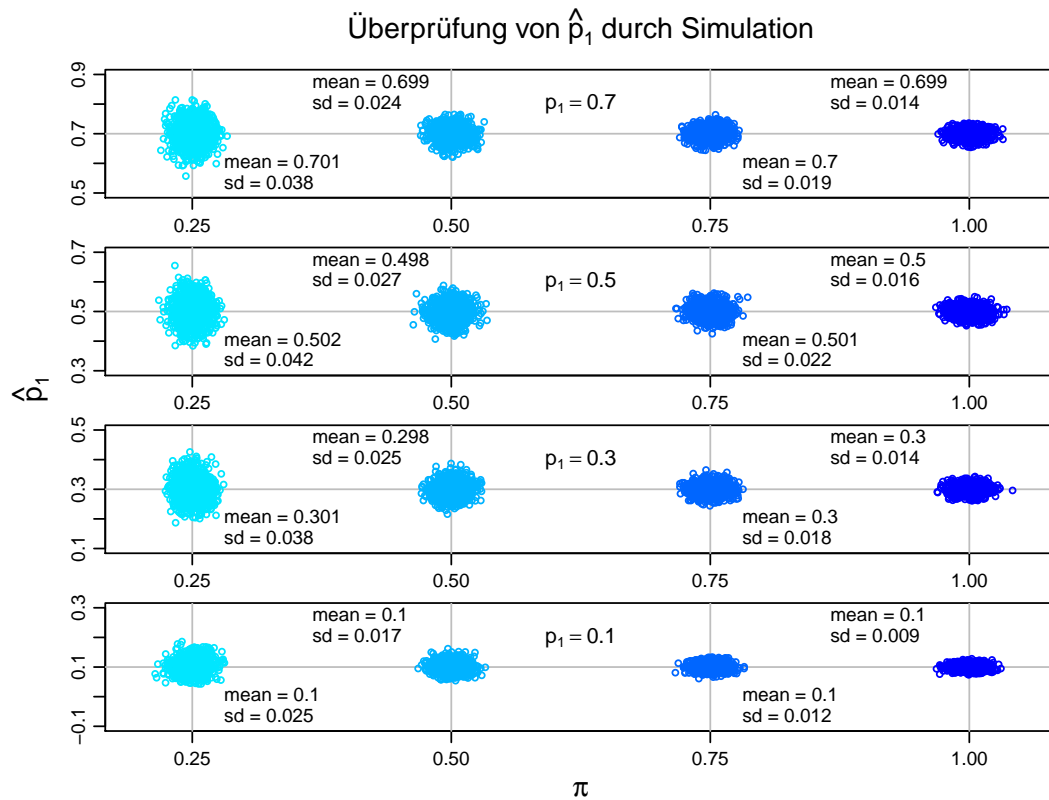


Abbildung 4.1: Schätzer  $\hat{p}_1$  für simulierte Daten.

Jede Punktwolke besteht aus 1000 Punkten, die für die Werte der Schätzer  $\hat{p}_1$  stehen, die aus den für eine ausgewählte Parameterkombination  $(p_1, p_2, p_3, p_4, \pi)$  simulierten Übergangshäufigkeiten  $(n_{rs})_{r,s \in \{1,2,3,4\}}$  berechnet wurden. Die horizontalen bzw. vertikalen grauen Linien, die sich hinter den Punktwolken kreuzen, geben den Wert von  $p_1$  bzw.  $\pi$  an, die für die jeweiligen Simulationen verwendet wurden. Die y-Koordinate eines Punktes gibt  $\hat{p}_1$  für eine Simulation an, die x-Koordinate ist die Summe des wahren Wertes von  $\pi$  und einer  $\mathcal{N}(0, 0.01)$ -verteilten Zufallsvariable (Punkte sind „gejittert“).

Mittelwerte (mean) und Standardabweichungen (sd) beziehen sich auf alle  $\hat{p}_1$  einer Punktwolke und sind auf drei Stellen gerundet.



Es ist zu beobachten, dass der Mittelwert der simulierten  $\hat{p}_1$  höchstens um  $\pm 2/n$  vom wahren Wert  $p_1$  abweicht und die Streuung von  $\hat{p}_1$  für jedes Parameterpaar  $(p_1, \pi)$  umso größer ist, je kleiner  $\pi$  ist und je näher  $p_1$  bei 0.5 liegt. Die Abhängigkeit der Varianz von  $\hat{p}_1$  von  $(p_1, \pi)$  wird in Kapitel 5.1 genauer untersucht, insbesondere im Abschnitt 5.1.3.

## 4.2 Schätzer für $\pi$

Der Schätzer für  $\pi$ , der in der vorliegenden Arbeit zur Anpassung des SMI-Modells verwendet wird, ist

$$\hat{\pi} := \hat{\pi}^{(g)} := \frac{n - \sum_r n_{rr}}{n(1 - \sum_s \hat{p}_s^2)} \wedge 1. \quad (4.2)$$

Es wird nun die Idee hinter diesem Schätzer erläutert.

Wie in Kapitel 3.2 beschrieben wurde, lässt sich  $\pi$  im Kontext des Hintergrundmodells interpretieren als die Wahrscheinlichkeit, mit der bei einem Übergang von  $X_i$  zu  $X_{i+1}$ , also zwischen zwei aufeinanderfolgenden Abtastungspunkten  $B_i$  und  $B_{i+1}$ , mindestens ein Intervallwechsel stattfindet,  $i \in \mathbb{N}_0$ . Diese Interpretation dient als Ansatz für die Herleitung des Schätzers. Ziel ist es dabei, die relative Häufigkeit aller Übergänge mit Intervallwechsel zu ermitteln bzw. zu approximieren.

Bei einem Übergang von einem Zustand  $r$  zu einem Zustand  $s$ , der sich von  $r$  unterscheidet, findet stets ein Intervallwechsel statt (siehe Kapitel 3.2). Es gibt also mindestens  $\sum_{r \neq s} n_{rs}$  Übergänge mit Intervallwechsel.

Bei einem Übergang von einem Zustand  $r$  zum selben Zustand  $r$  lässt sich nicht unterscheiden, ob dabei ein Intervallwechsel auftritt oder nicht. Daher werden diese Übergänge bei diesem Ansatz für das Schätzen von  $\pi$  nicht berücksichtigt.

Die absolute Häufigkeit der Übergänge mit Intervallwechsel wird hier deswegen (nach unten) abgeschätzt durch

$$\sum_{r \neq s} n_{rs} = n - \sum_r n_{rr}.$$

Es folgen drei Schritte bei der Herleitung von  $\hat{\pi}$ :

1. Ermittlung eines sinnvollen Terms, durch den die absolute Häufigkeit  $n - \sum_r n_{rr}$  geteilt wird, um die relative Häufigkeit der Übergänge mit Intervallwechsel zu schätzen. Als Kriterium dient dabei die Erwartungstreue der relativen Häufigkeit. Dadurch erhält man

$$\hat{\pi}^{(1)} := \frac{n - \sum_r n_{rr}}{n \sum_{t \neq u} p_t p_u}.$$

2. Ersetzen von  $p_t$  und  $p_u$  durch ihre Schätzer  $\hat{p}_t$  und  $\hat{p}_u$ . Dadurch erhält man

$$\hat{\pi}^{(u)} := \frac{n - \sum_r n_{rr}}{n \sum_{t \neq u} \hat{p}_t \hat{p}_u}.$$

3. Abschneiden von  $\hat{\pi}^{(u)}$  bei 1. Dadurch erhält man

$$\hat{\pi} := \hat{\pi}^{(g)} := \frac{n - \sum_r n_{rr}}{n \sum_{t \neq u} \hat{p}_t \hat{p}_u} \wedge 1.$$

**1. Schritt:** Um aus der absoluten Häufigkeit

$$n - \sum_r n_{rr} = \sum_{r \neq s} n_{rs}$$

von Übergängen zwischen sich unterscheidenden Zuständen die relative Häufigkeit der Übergänge mit Intervallwechsel abzuschätzen, ist es zweckmäßig, die absolute Häufigkeit nicht durch die Anzahl  $n$  aller Übergänge zu teilen, sondern zu berücksichtigen, dass sämtliche Übergänge von einem Zustand  $r$  zum selben Zustand  $r$  nicht verwendet werden.

Teilt man die absolute Häufigkeit  $n - \sum_r n_{rr}$  durch

$$n \sum_{t \neq u} p_t p_u,$$

erhält man wegen

$$\mathbb{E} \left[ \frac{n - \sum_r n_{rr}}{n \sum_{t \neq u} p_t p_u} \right] = \mathbb{E} \left[ \frac{\sum_{r \neq s} n_{rs}}{n \sum_{t \neq u} p_t p_u} \right] = \frac{\sum_{r \neq s} n p_r \pi p_s}{n \sum_{t \neq u} p_t p_u} = \pi$$

einen erwartungstreuen Schätzer für  $\pi$ :

$$\hat{\pi}^{(1)} := \frac{n - \sum_r n_{rr}}{n(1 - \sum_s p_s^2)} = \frac{n - \sum_r n_{rr}}{n \sum_{t \neq u} p_t p_u}.$$

Anstelle von  $n$  steht bei  $\hat{\pi}^{(1)}$  die erwartete Anzahl von Übergängen zwischen verschiedenen Karten  $t$  und  $u$  im Nenner, die man erwartet, wenn bei jedem Übergang ein Intervallwechsel stattfinden würde, also wenn  $\pi = 1$  wäre, denn es gilt

$$\mathbb{E} \left[ \sum_{t \neq u} n_{tu} \mid \pi = 1 \right] = n \sum_{t \neq u} p_t p_u.$$

**2. Schritt:** In der Formel für den erwartungstreuen Schätzer  $\hat{\pi}^{(1)}$  werden die wahren Werte  $p_1, \dots, p_a$  verwendet. Da diese Wahrscheinlichkeiten jedoch nicht bekannt sind, sondern geschätzt werden müssen, werden die  $p_s$  durch  $\hat{p}_s$  ersetzt. Das führt zu dem Schätzer

$$\hat{\pi}^{(u)} := \frac{n - \sum_r n_{rr}}{n(1 - \sum_s \hat{p}_s^2)}.$$

**3. Schritt:** Mit einer positiven Wahrscheinlichkeit tritt der Fall ein, dass der Zähler größer als der Nenner wird. Das bedeutet, dass mehr Übergänge zwischen sich unterscheidenden Zuständen realisiert wurden, als man selbst für  $\pi = 1$  erwarten würde. Dann gilt  $\hat{\pi}^{(u)} > 1$ . Da  $\pi$  als Wahrscheinlichkeit interpretiert werden kann, wird der Schätzer bei 1 abgeschnitten. Damit ist die Herleitung des Schätzers  $\hat{\pi}$  für  $\pi$ , der für die Anpassung des SMI-Modells an die Daten angewandt wird, beendet und es sei

$$\hat{\pi} := \hat{\pi}^{(g)} := \frac{n - \sum_r n_{rr}}{n(1 - \sum_s \hat{p}_s^2)} \wedge 1.$$

**Bemerkung 4.4.** Interpretiert man  $\pi$  als Verhältnis zweier Größen, kann es sinnvoll sein, auch Werte  $> 1$  für einen Schätzer für  $\pi$  zuzulassen. In diesem Fall findet  $\hat{\pi}^{(u)}$  Anwendung. Im weiteren Verlauf der Arbeit wird stets angegeben, ob  $\hat{\pi} = \hat{\pi}^{(g)}$  oder  $\hat{\pi}^{(u)}$  als Schätzer für  $\pi$  verwendet wird.

#### 4.2.1 Vergleich mit ML-Schätzer auf der Basis von Hintergrundmodell-Variablen

Wären die Variablen  $(I_j)_{j \in \mathbb{N}}$  und  $(B_i)_{i \in \mathbb{N}_0}$  aus dem Hintergrundmodell gegeben, ließe sich der ML-Schätzer  $\hat{\pi}^{\text{ML,H}}$  für  $\pi$  auf der Basis der Ereignisse

$$A_i := \{\text{Intervallwechsel zwischen } B_{i-1} \text{ und } B_i\} \quad \forall i \in \{1, \dots, n\},$$

berechnen. Dieser ML-Schätzer wird nun bestimmt und mit  $\hat{\pi}$  verglichen.

Die Wahrscheinlichkeit für das Ereignis  $A_i$  beträgt

$$\mathbb{P}(A_i) = \pi \quad \forall i \in \{1, \dots, n\}$$

und  $A_1, \dots, A_n$  sind unabhängig voneinander. Somit ist die Anzahl

$$\sum_{i=1}^n \mathbf{1}_{A_i}$$

der Übergänge mit Intervallwechsel binomial( $n, \pi$ )-verteilt.

Sei

$$n_{rr} = \tilde{n}_{rr} + \bar{n}_{rr}$$

die Zerlegung aus Kapitel 4.1.1, mit der die Anzahl  $n_{rr}$  der Übergänge von einem Zustand  $r$  zum selben Zustand  $r$  zerlegt wird in die Anzahl der Übergänge von  $r$  zu  $r$  mit Intervallwechsel ( $\tilde{n}_{rr}$ ) und ohne Intervallwechsel ( $\bar{n}_{rr}$ ). Dann lässt sich die Anzahl der Übergänge zwischen beliebigen Zuständen mit Intervallwechsel beschreiben durch

$$\sum_{i=1}^n \mathbf{1}_{A_i} = \sum_{r \neq s} n_{rs} + \sum_r \tilde{n}_{rr} = n - \sum_r \bar{n}_{rr}.$$

Für den ML-Schätzer  $\hat{\pi}^{\text{ML,H}}$  auf Basis von  $(A_1, \dots, A_n)$  ergibt sich somit

$$\hat{\pi}^{\text{ML,H}} = \frac{\sum_{i=1}^n \mathbf{1}_{A_i}}{n} = \frac{n - \sum_r \bar{n}_{rr}}{n}.$$

### Fazit

Der Unterschied zwischen den Schätzern  $\hat{\pi}$  und  $\hat{\pi}^{\text{ML,H}}$  besteht darin, dass bei  $\hat{\pi}^{\text{ML,H}}$  die Übergänge von einem Zustand  $r$  zum selben Zustand  $r$  mit Intervallwechsel mitberücksichtigt werden, während sie bei  $\hat{\pi}$  vernachlässigt werden.

Dementsprechend werden beim ML-Schätzer für die Ermittlung der relativen Häufigkeit im Nenner alle  $n$  Übergänge herangezogen. Im Gegensatz dazu muss der Nenner für  $\hat{\pi}$  angepasst werden zur geschätzten erwarteten Anzahl von Übergängen zwischen zwei verschiedenen Zuständen bedingt darauf, dass  $\pi = 1$  ist.

### 4.2.2 Überprüfung von $\hat{\pi}^{(u)}$ durch Simulation

Die Abbildung 4.2 zeigt die durch  $\hat{\pi}^{(u)}$  geschätzten Werte für eine Auswahl von simulierten Markov-Ketten, die in Kapitel 4.1.2 beschrieben worden sind.

Der Mittelwert der simulierten  $\hat{\pi}^{(u)}$  überschätzt stets den wahren Wert von  $\pi$  um bis zu  $4/n$  mit  $n = 1000$ . Zudem erscheint es so, als gäbe es einen Trend für die Standardabweichung, kleiner zu sein, je „dichter“  $p_1, p_2, p_3, p_4$  beieinander liegen, z. B. je kleiner  $\sum_{r=1}^4 p_r^2$  ist. Genauere Betrachtungen für die Varianz von  $\hat{\pi}^{(u)}$  werden in Kapitel 5.3 angestellt.

Um die Werte von  $\hat{\pi} = \hat{\pi}^{(g)}$  anstatt von  $\hat{\pi}^{(u)}$  aus den Simulationen zu erhalten, müssen in Abbildung 4.2 lediglich alle Punkte mit  $\hat{\pi}^{(u)} > 1$  auf der x-Achse nach links auf die 1 verschoben werden.

### 4.3 Vergleich der Schätzer $\hat{p}_r$ und $\hat{\pi}^{(u)}$ mit den ML-Schätzern auf der Basis von $(X_0, \dots, X_n)$

Im Folgenden werden die ML-Schätzer  $\hat{p}_1^{\text{ML}}, \dots, \hat{p}_a^{\text{ML}}$  und  $\hat{\pi}^{\text{ML}}$  für  $p_1, \dots, p_a$  und  $\pi$  auf der Basis von  $(X_0, \dots, X_n)$  untersucht.

Dafür wird zunächst die Log-Likelihood-Funktion hergeleitet. Anschließend werden zwei Ansätze zur Berechnung der ML-Schätzer für  $\pi$  und  $p_r$  auf der Basis von  $(X_0, \dots, X_n)$  beschrieben, der Lagrange-Ansatz und der numerische Ansatz.

Durch Anwendung des Lagrange-Ansatzes wird sich herausstellen, wie bei diesen ML-Schätzern die Anzahl  $\bar{n}_{rr}$  der Übergänge von  $r$  zu  $r$  ohne Intervallwechsel geschätzt wird und wie sich daraus der ML-Schätzer für  $\pi$  bildet.

Für die konkrete Berechnung der ML-Schätzer wird jedoch der numerische Ansatz verfolgt. Die Log-Likelihood-Funktion hängt sowohl von  $(n_{rs})_{r,s \in S}$  als auch von  $x_0$  ab. Da  $x_0$  in den vorliegenden Daten nicht gegeben ist, wird geklärt, wie sich  $x_0$  bei der numerischen Approximation der ML-Schätzer vernachlässigen lässt.

## Überprüfung von $\hat{\pi}^{(u)}$ durch Simulation

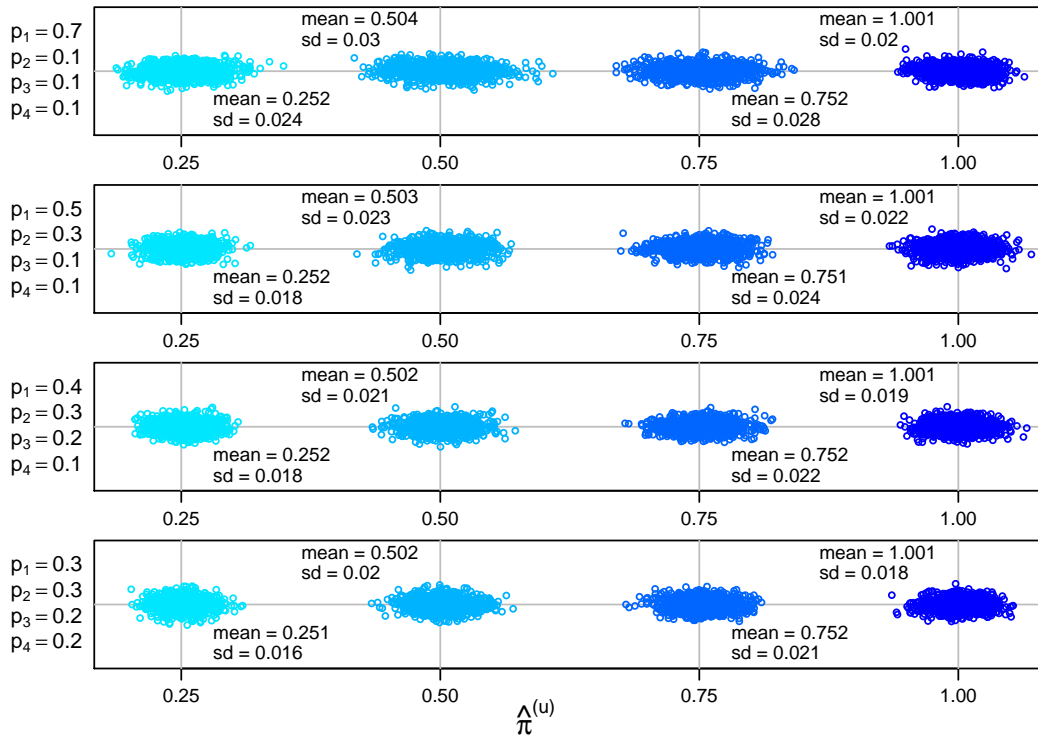


Abbildung 4.2: Schätzer  $\hat{\pi}^{(u)}$  für simulierte Daten.

Jede Punktwolke besteht aus 1000 Punkten, die für die Werte der Schätzer  $\hat{\pi}^{(u)}$  stehen, die aus den für eine ausgewählte Parameterkombination  $(p_1, p_2, p_3, p_4, \pi)$  simulierten Übergangshäufigkeiten  $(n_{rs})_{r,s \in \{1,2,3,4\}}$  berechnet wurden. Die vertikalen grauen Linien geben für die getroffene Punktwolke den Wert von  $\pi$  an, der für die jeweiligen Simulationen verwendet wurde. Die x-Koordinate eines Punktes entspricht dem Wert von  $\hat{\pi}^{(u)}$  für eine Simulation, die y-Koordinate ist eine normalverteilte Zufallsvariable (Punkte sind „gejittert“).

Mittelwerte (mean) und Standardabweichungen (sd) beziehen sich auf alle  $\hat{\pi}^{(u)}$  einer Punktwolke und sind auf drei Stellen gerundet.

Sei  $x = (x_0, \dots, x_n) \in S^{n+1}$ . Dann gilt für die Likelihood-Funktion

$$\mathbb{P}(X = x) = p_{x_0} \prod_{i=0}^{n-1} p_{x_i, x_{i+1}} = p_{x_0} \prod_{r,s \in S} p_{rs}^{n_{rs}} = p_{x_0} \prod_{\substack{r,s \in S \\ r \neq s}} p_{rs}^{n_{rs}} \prod_{t \in S} p_{tt}^{n_{tt}}.$$

Um die Likelihood-Funktion weiter zusammenzufassen, werden die Gleichungen

$$\sum_{\substack{r,s \in S \\ r \neq s}} n_{rs} = \sum_{r,s \in S} n_{rs} - \sum_{r \in S} n_{rr} = n - \sum_{r \in S} n_{rr}$$

und

$$\mathbf{1}_{\{x_0=s\}} + \sum_{\substack{r \in S \\ r \neq s}} n_{rs} = \mathbf{1}_{\{x_0=s\}} + \sum_{r \in S} n_{rs} - n_{ss} = n'_s - n_{ss} \quad \forall s \in S$$

verwendet, wobei  $n'_s$  in Kapitel 2.2 definiert wurde als

$$n'_s = \sum_{i=0}^n \mathbf{1}_{\{x_i=s\}} = \mathbf{1}_{\{x_0=s\}} + n_{.s} = \mathbf{1}_{\{x_0=s\}} + \sum_{r \in S} n_{rs} \quad \forall s \in S.$$

Mit der Formel

$$p_{rs} = \pi p_s + \delta_{r,s}(1 - \pi) \quad \forall r, s \in S$$

für die Übergangswahrscheinlichkeiten im SMI-Modell erhält man

$$\begin{aligned} \mathbb{P}(X = x) &= p_{x_0} \prod_{\substack{r,s \in S \\ r \neq s}} (\pi p_s)^{n_{rs}} \prod_{t \in S} (\pi p_t + 1 - \pi)^{n_{tt}} \\ &= \left( \sum_{s \in S} \mathbf{1}_{\{x_0=s\}} p_s \right) \pi^{n - \sum_r n_{rr}} \prod_{s \in S} p_s^{\sum_{r \neq s} n_{rs}} \prod_{t \in S} (\pi p_t + 1 - \pi)^{n_{tt}} \\ &= \pi^{n - \sum_r n_{rr}} \prod_{s \in S} p_s^{n'_s - n_{ss}} \prod_{t \in S} (\pi p_t + 1 - \pi)^{n_{tt}}. \end{aligned}$$

Für die Log-Likelihood-Funktion  $f$  folgt

$$f(\pi, p_1, \dots, p_a) = \left( n - \sum_{r \in S} n_{rr} \right) \log(\pi) + \sum_{s \in S} (n'_s - n_{ss}) \log(p_s) + \sum_{t \in S} n_{tt} \log(\pi p_t + 1 - \pi).$$

Diese Funktion ist für die Bestimmung der ML-Schätzer zu maximieren. Dabei ist zu beachten, dass  $n_{rs}$  für alle  $r, s \in S$  in den Daten vorliegt, jedoch  $x_0$  und somit auch  $n'_s$  nicht in allen Fällen zuverlässig bestimmt werden können (siehe Kapitel 2.2).

### 4.3.1 Lagrange-Ansatz

Eine mögliche Methode für die Maximierung der Log-Likelihood-Funktion stellt der Lagrange-Ansatz dar (siehe z. B. Forster, 2006) mit der Nebenbedingung  $\sum_s p_s = 1$ .

Für die Betrachtungen wird hier angenommen, dass neben  $n_{rs}$  für alle  $r, s \in S$  auch  $x_0 \in S$  gegeben ist und somit auch  $n'_s$  für alle  $s \in S$ .

Die Lagrange-Funktion mit Lagrange-Multiplikator  $\lambda$  lautet

$$\mathcal{L}(\pi, p_1, \dots, p_a, \lambda) := f(\pi, p_1, \dots, p_a) + \lambda \left( 1 - \sum_s p_s \right)$$

und die zu lösenden Gleichungen ergeben sich mit  $\vartheta := (\pi, p_1, \dots, p_a, \lambda)$  zu

$$\frac{\partial \mathcal{L}(\vartheta)}{\partial \pi} = \left( n - \sum_r n_{rr} \right) \frac{1}{\pi} + \sum_t n_{tt} \frac{p_t - 1}{\pi p_t + 1 - \pi} = 0 \quad (4.3)$$

$$\frac{\partial \mathcal{L}(\vartheta)}{\partial p_r} = (n'_r - n_{rr}) \frac{1}{p_r} + n_{rr} \frac{\pi}{\pi p_r + 1 - \pi} - \lambda = 0 \quad \forall r \in S \quad (4.4)$$

$$\frac{\partial \mathcal{L}(\vartheta)}{\partial \lambda} = 1 - \sum_s p_s = 0.$$

Angenommen,  $p_1, \dots, p_a$  sind gegeben, dann wäre für die Berechnung des ML-Schätzers  $\hat{\pi}^{\text{ML}}$  für  $\pi$  noch die Gleichung (4.3) zu lösen. Es müsste also die Nullstelle eines Polynoms  $a$ -ten Grades bestimmt werden, um eine geschlossene Formel für  $\hat{\pi}^{\text{ML}}$  zu finden. Dieser Weg hat sich insbesondere für  $a > 2$  als nicht praktikabel herausgestellt. Um dennoch eine Lösung für die ML-Schätzer  $\hat{p}_1^{\text{ML}}, \dots, \hat{p}_a^{\text{ML}}$  und  $\hat{\pi}^{\text{ML}}$  auf der Basis von  $(X_0, \dots, X_n)$  zu finden, lassen sich numerische Verfahren zur Approximation anwenden. In Kapitel 4.3.2 wird beschrieben, wie sich die numerische Approximation der ML-Schätzer mit R umsetzen lässt.

Bevor diese Beschreibung erfolgt, werden die Gleichungen des Lagrange-Ansatzes so umgeformt, dass sich Einblicke dazu ergeben, wie beim ML-Schätzer die Aufteilung

$$n_{rr} = \tilde{n}_{rr} + \bar{n}_{rr}$$

in Übergänge von  $r$  zu  $r$  mit Intervallwechsel ( $\tilde{n}_{rr}$ ) und ohne Intervallwechsel ( $\bar{n}_{rr}$ ) ermittelt wird. Zudem wird eine interpretierbare Formel für  $\hat{\pi}^{\text{ML}}$  hergeleitet, die jedoch versteckt selbst von  $\hat{\pi}^{\text{ML}}$  abhängt.

Aus Gleichung (4.3) folgt

$$\begin{aligned} \frac{n}{\pi} &= \sum_r \frac{n_{rr}}{\pi} \cdot \frac{\pi p_r + 1 - \pi}{p_{rr}} - \sum_t \frac{n_{tt}(p_t - 1)\pi}{\pi p_{tt}} \\ &= \sum_r \frac{n_{rr}(\pi p_r + 1 - \pi - \pi p_r + \pi)}{\pi p_{rr}} \\ &= \sum_r \frac{n_{rr}}{\pi p_{rr}}, \end{aligned}$$

also

$$n = \sum_r \frac{n_{rr}}{p_{rr}}, \quad (4.5)$$

und aus Gleichung (4.4) erhält man

$$\lambda = \frac{n'_r - n_{rr}}{p_r} + \frac{\pi n_{rr}}{p_{rr}} \quad \forall r \in S. \quad (4.6)$$

Es wird nun gezeigt, dass sich der Lagrange-Multiplikator  $\lambda$  interpretieren lässt als Schätzer für die Anzahl der Intervalle, die von  $B_0, \dots, B_n$  mindestens einmal getroffen werden.

Wie in Kapitel 4.1.1 erläutert wurde, lässt sich die Anzahl dieser Intervalle darstellen durch

$$(n+1) - \sum_r \bar{n}_{rr}.$$

Mit Gleichung (4.6) ergibt sich

$$\begin{aligned} \lambda &= \sum_r p_r \lambda \\ &= \sum_r n'_r - n_{rr} + \frac{\pi p_r n_{rr}}{p_{rr}} \\ &= \sum_r n'_r - \frac{\pi p_r + 1 - \pi}{p_{rr}} n_{rr} + \frac{\pi p_r}{p_{rr}} n_{rr} \\ &= (n+1) - \sum_r \frac{1 - \pi}{p_{rr}} n_{rr}. \end{aligned} \quad (4.7)$$

Dabei ist  $(1 - \pi)/p_{rr}$  die Wahrscheinlichkeit bedingt auf einen Übergang von  $r$  zu  $r$ , dass bei diesem Übergang kein Intervallwechsel stattfindet. Somit ist

$$\hat{\bar{n}}_{rr} := \frac{1 - \pi}{p_{rr}} n_{rr}$$

ein Schätzer für die Anzahl  $\bar{n}_{rr}$  der Übergänge von  $r$  zu  $r$  ohne Intervallwechsel und

$$\lambda = (n+1) - \sum_r \hat{\bar{n}}_{rr}$$

ein Schätzer für die Anzahl der Intervalle, die von mindestens einem  $B_i$  mit  $i \in \{0, \dots, n\}$  getroffen werden.

Ein weiteres Argument für diese Interpretation für  $\lambda$  liefert die Betrachtung des Erwartungswertes von  $\lambda$  unter der Bedingung, dass  $p_1, \dots, p_a$  und  $\pi$  bekannt sind. Nach



Gleichung (4.7) ist

$$\begin{aligned}
\mathbb{E}[\lambda | p_1, \dots, p_a, \pi] &= \mathbb{E} \left[ (n+1) - \sum_r \frac{1-\pi}{p_{rr}} n_{rr} \mid p_1, \dots, p_a, \pi \right] \\
&= n+1 - \sum_r \frac{1-\pi}{p_{rr}} n p_r p_{rr} \\
&= n+1 - (1-\pi)n \sum_r p_r \\
&= 1 + \pi n.
\end{aligned}$$

Der Erwartungswert von  $\lambda$  für gegebene  $p_1, \dots, p_a$  und  $\pi$  ist also gerade die erwartete Anzahl  $1 + \pi n$  der Intervalle, die von  $B_0, \dots, B_n$  getroffen werden, denn  $B_0$  ist der erste Abtastungszeitpunkt, der überhaupt ein Intervall trifft und somit sicher in einem noch nicht getroffenen Intervall liegt, und  $B_1, \dots, B_n$  liegen jeweils mit Wahrscheinlichkeit  $\pi$  in einem noch nicht getroffenen Intervall.

Analog zur Rechnung für Gleichung (4.7) lässt sich  $\lambda p_r$  bestimmen durch

$$\lambda p_r = n'_r - \frac{1-\pi}{p_{rr}} n_{rr} = n'_r - \hat{n}_{rr}$$

und interpretieren als Schätzer für die Anzahl der Intervalle, die von  $B_0, \dots, B_n$  getroffen werden und denen der Zustand  $r$  zugeordnet ist. Damit ergibt sich ein Schätzer  $\hat{n}_{rr}$  für die Anzahl  $\tilde{n}_{rr}$  der Übergänge von  $r$  zu  $r$  mit Intervallwechsel durch

$$\hat{n}_{rr} := n_{rr} - \hat{n}_{rr} = n_{rr} - (n'_r - \lambda p_r).$$

Dieser Schätzer wird nun bei der Interpretation von  $\hat{\pi}^{\text{ML}}$  wieder erscheinen.

Gleichung (4.6) lässt sich umformen zu

$$\frac{n_{rr}}{p_{rr}} = \frac{n_{rr} - (n'_r - \lambda p_r)}{\pi p_r} = \frac{\hat{n}_{rr}}{\pi p_r} \quad \forall r \in S$$

und durch Einsetzen der letzten Gleichung in Gleichung (4.5) sowie durch Multiplikation mit  $\pi/n$  ergibt sich

$$\hat{\pi}^{\text{ML}} = \sum_r \frac{n_{rr} - (n'_r - \lambda \hat{p}_r)}{n \hat{p}_r} = \sum_r \frac{\hat{n}_{rr}}{n \hat{p}_r}. \quad (4.8)$$

Dabei ist zu beachten, dass die rechte Seite der Gleichung (4.8) abhängig ist von  $\hat{\pi}^{\text{ML}}$ , sodass die Gleichung (4.8) nicht geeignet ist zur direkten Berechnung von  $\hat{\pi}^{\text{ML}}$ .

Stattdessen zeigt Gleichung (4.8), dass der ML-Schätzer  $\hat{\pi}^{\text{ML}}$  bestimmt wird, indem aus allen geschätzten Übergangshäufigkeiten von  $r$  zu  $r$  mit Intervallwechsel ein Schätzer für  $\pi$  gewichtet mit  $\hat{p}_r$  berechnet wird für alle  $r \in S$  und aus diesen Schätzern der gewichtete Mittelwert gebildet wird.

Während also für die Ermittlung des Schätzers  $\hat{\pi}$  aus Kapitel 4.2 die Übergangshäufigkeiten  $n_{rs}$  von einem Zustand  $r$  zu einem davon verschiedenen Zustand  $s$  herangezogen werden, dienen für die Bestimmung von  $\hat{\pi}^{\text{ML}}$  die geschätzten Übergangshäufigkeiten  $\hat{n}_{rr}$  von einem Zustand  $r$  zum selben Zustand  $r$  mit Intervallwechsel als Grundlage.

### 4.3.2 Numerischer Ansatz

Wie zu Beginn des Kapitels 4.3 gezeigt wurde, ist die Log-Likelihood-Funktion gegeben durch

$$f(\pi, p_1, \dots, p_a) = \left(n - \sum_r n_{rr}\right) \log(\pi) + \sum_s (n'_s - n_{ss}) \log(p_s) + \sum_t n_{tt} \log(\pi p_t + 1 - \pi).$$

Um die ML-Schätzer  $\hat{\pi}^{\text{ML}}, \hat{p}_1^{\text{ML}}, \dots, \hat{p}_a^{\text{ML}}$  zu bestimmen, müssen die Parameter gefunden werden, für die die Log-Likelihood-Funktion maximiert wird.

Neben  $\pi$  und  $p_1, \dots, p_a$  hängt diese Funktion nur von den gegebenen  $(n_{rs})_{r,s \in S}$  und vom unbekanntem  $x_0$  ab, wobei  $x_0$  für die Bestimmung von

$$n'_s = \mathbf{1}_{\{x_0=s\}} + \sum_r n_{rs}$$

benötigt wird. Da  $x_0$  in den gegebenen Daten nicht zur Verfügung steht, muss  $x_0$  vernachlässigt werden bei der numerischen Approximation der ML-Schätzer

$$\hat{\pi}^{\text{ML}}, \hat{p}_1^{\text{ML}}, \dots, \hat{p}_a^{\text{ML}}.$$

Es wird deshalb für diese Arbeit die Log-Likelihood-Funktion bedingt auf  $X_0 = x_0$  maximiert. Dadurch wird  $n'_s$  zu  $n_{.s} = \sum_r n_{rs}$  für alle  $s \in S$ .

Die numerische Approximation von  $\hat{\pi}^{\text{ML}}$  und  $\hat{p}_1^{\text{ML}}, \dots, \hat{p}_a^{\text{ML}}$  ist mit R durch die Funktion `optim` realisierbar. Die Funktion `optim` approximiert von einem Startpunkt aus durch ein iteratives Verfahren die Parameter, die eine gegebene Funktion maximieren.

Ziel ist es, die ML-Schätzer  $\hat{p}_r^{\text{ML}}$  und  $\hat{\pi}^{\text{ML}}$  auf der Basis von  $(X_0, \dots, X_n)$  mit den Schätzern  $\hat{p}_r$  und  $\hat{\pi}^{(u)}$  aus den Kapiteln 4.1 und 4.2 zu vergleichen.<sup>13</sup> Dafür wurden aus den in Kapitel 4.1.2 beschriebenen simulierten Datensätzen jeweils die Schätzer  $\hat{p}_r^{\text{ML}}$  und  $\hat{\pi}^{\text{ML}}$  ermittelt. Für die Ermittlung der ML-Schätzer mit der Funktion `optim` wurden zufällige Startwerte für die Iterationen gewählt, die mit der Methode „Nelder-Mead“ durchgeführt wurden.

In Abbildung 4.3 werden für ausgewählte Parameter  $(p_1, p_2, p_3, p_4)$  und für jedes  $\pi \in \{0.25, 0.50, 0.75, 1.00\}$  pro simulierter Markov-Kette die beiden Schätzer  $\hat{\pi}^{(u)}$  und  $\hat{\pi}^{\text{ML}}$  gegeneinander aufgetragen. Es wird deutlich, dass die beiden Schätzer eine Korrelation von annähernd 1 aufweisen. Je größer der wahre Wert von  $\pi$  ist, desto stärker weichen die beiden Schätzer tendenziell voneinander ab.

In Abbildung 4.4 werden für ausgewählte Parameter  $(p_1, p_2, p_3, p_4)$  und für jedes  $\pi \in \{0.25, 0.50, 0.75, 1.00\}$  pro simulierter Markov-Kette die beiden Schätzer  $\hat{p}_1$  und  $\hat{p}_1^{\text{ML}}$  gegeneinander aufgetragen. Dabei wurden die Parameter  $(p_1, p_2, p_3, p_4)$  so gewählt, dass  $p_1$  jeden Wert aus  $\{0.7, 0.5, 0.3, 0.1\}$  einmal annimmt. In der Abbildung zeigt sich, dass

<sup>13</sup>Der Vergleich erfolgt mit  $\hat{\pi}^{(u)}$  und nicht mit  $\hat{\pi} = \hat{\pi}^{(g)}$ , da  $\hat{\pi}^{(u)}$  mehr Informationen enthält.

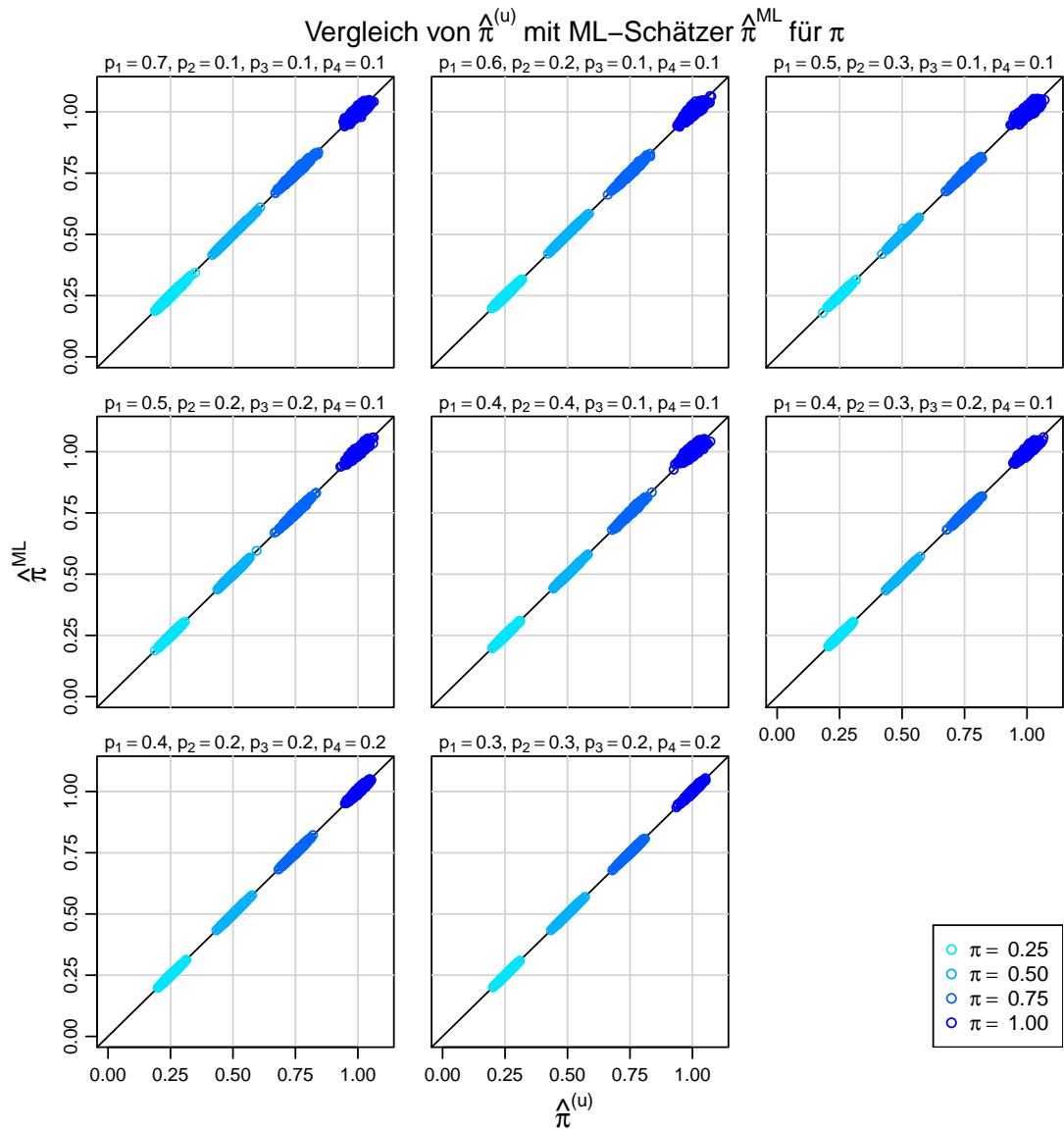


Abbildung 4.3: Vergleich des Schätzers  $\hat{\pi}^{(u)}$  mit dem ML-Schätzer  $\hat{\pi}^{\text{ML}}$  für verschiedene Parameter  $(p_1, p_2, p_3, p_4)$  jeweils für  $\pi = 0.25, 0.50, 0.75, 1.00$ . Für jede simulierte Markov-Kette sind die Werte der beiden Schätzer durch einen Punkt  $(\hat{\pi}^{(u)}, \hat{\pi}^{\text{ML}})$  dargestellt. Die eingezeichneten Diagonalen sind die Winkelhalbierenden des ersten Quadranten.

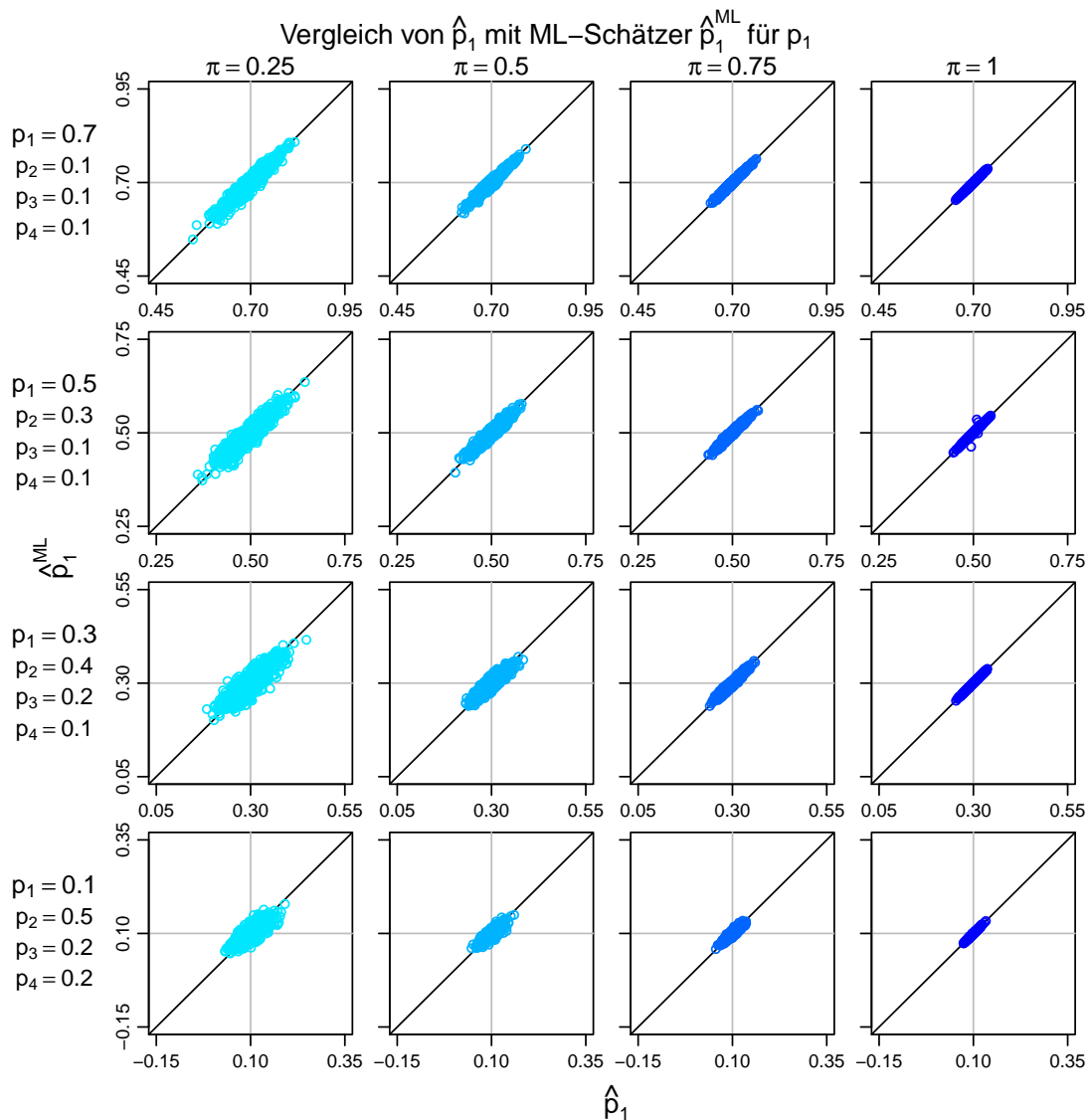


Abbildung 4.4: Vergleich des Schätzers  $\hat{p}_1$  mit dem ML-Schätzer  $\hat{p}_1^{\text{ML}}$  für verschiedene Parameter  $(p_1, p_2, p_3, p_4)$  mit  $p_1 \in \{0.7, 0.5, 0.3, 0.1\}$  jeweils für  $\pi = 0.25, 0.50, 0.75, 1.00$ . Für jede simulierte Markov-Kette sind die Werte der beiden Schätzer durch einen Punkt  $(\hat{p}_1, \hat{p}_1^{\text{ML}})$  dargestellt. Die eingezeichneten Diagonalen sind die Winkelhalbierenden des ersten Quadrants.

auch die Schätzer  $\hat{p}_1$  und  $\hat{p}_1^{\text{ML}}$  eine Korrelation von annähernd 1 aufweisen.<sup>14</sup> Je kleiner der wahre Wert von  $\pi$  ist, desto stärker weichen die Schätzer  $\hat{p}_1$  und  $\hat{p}_1^{\text{ML}}$  tendenziell voneinander ab.

Insgesamt lässt sich in den Abbildungen 4.3 und 4.4 erkennen, dass sich die Schätzer  $\hat{\pi}^{(u)}$  und  $\hat{\pi}^{\text{ML}}$  bzw.  $\hat{p}_1$  und  $\hat{p}_1^{\text{ML}}$  annähernd gleichen. Der Vorteil der Schätzer  $\hat{\pi}^{(u)}$  bzw.  $\hat{\pi} = \hat{\pi}^{(g)}$  und  $\hat{p}_r$  ist, dass sie sich mit weniger Rechenaufwand (und somit in kürzerer Zeit) berechnen lassen als die ML-Schätzer  $\hat{\pi}^{\text{ML}}$  und  $\hat{p}_r^{\text{ML}}$ . Die ML-Schätzer bieten den Vorteil, dass für sie die asymptotische Verteilung der Teststatistik bekannt ist bei einem in Kapitel 6 vorgestellten  $\chi^2$ -Anpassungstest für das SMI-Modell, während die Verteilung der Teststatistik bei Verwendung der Schätzer  $\hat{\pi}$  und  $\hat{p}_r$  durch (zeitintensive) Simulationen approximiert werden muss.<sup>15</sup>

In dieser Arbeit ist die Entscheidung auf die Verwendung der Schätzer  $\hat{\pi}$  und  $\hat{p}_r$  gefallen, die im Folgenden auch SMI-Modell-Schätzer genannt werden.

#### 4.4 Schätzer für $p_{rs}$

Es wird nun erläutert, wie aus den geschätzten Parametern  $\hat{p}_1, \dots, \hat{p}_a$  und  $\hat{\pi}$  letztendlich die Übergangswahrscheinlichkeiten  $p_{rs}$  für  $r, s \in S = \{1, \dots, a\}$  geschätzt werden.

Sei

$$\vartheta = (\pi, p_1, \dots, p_a)$$

ein Parameter-Vektor für das SMI-Modell. Die Übergangswahrscheinlichkeiten  $p_{rs}$  können dann als Funktionen aufgefasst werden, die von  $\vartheta$  abhängen durch

$$p_{rs} = p_{rs}(\vartheta) = \pi p_s + \delta_{r,s}(1 - \pi).$$

Schätzt man nun  $\vartheta$  durch

$$\hat{\vartheta} = (\hat{\pi}, \hat{p}_1, \dots, \hat{p}_a),$$

folgt für den Schätzer  $\hat{p}_{rs}$  für  $p_{rs}$  auf natürliche Weise

$$\hat{p}_{rs} := p_{rs}(\hat{\vartheta}) = \hat{\pi} \hat{p}_s + \delta_{r,s}(1 - \hat{\pi}).$$

#### 4.5 Die aus den Daten berechneten Schätzer

Nachfolgend werden die aus den vorliegenden Daten berechneten SMI-Modell-Schätzer

$$\hat{p}_A^{(J,Z)}, \hat{p}_B^{(J,Z)}, \hat{p}_C^{(J,Z)}, \hat{p}_D^{(J,Z)} \quad \text{und} \quad \hat{\pi}^{(J,Z)}$$

für alle Probanden  $J$  und alle Schlafstadien  $Z$  präsentiert.

<sup>14</sup>Beim Vergleich der beiden Abbildungen 4.3 und 4.4 ist zu beachten, dass die gezeigten Ausschnitte aus der x- und y-Achse in Abbildung 4.4 halb so groß ist wie in Abbildung 4.3.

<sup>15</sup>Auf diesen Umstand wird in Kapitel 6.1.1 genauer eingegangen.

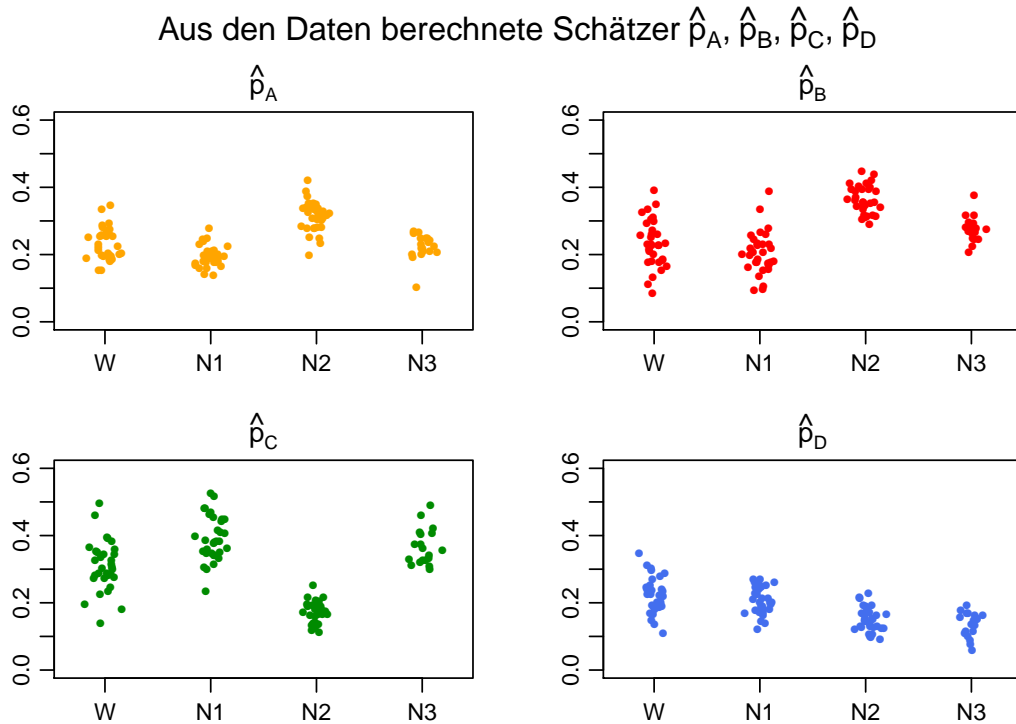


Abbildung 4.5: Aus den Daten berechnete Schätzer  $\hat{p}_r^{(J,Z)}$  getrennt nach Zustand  $r \in S$  und nach Schlafstadium  $Z$ . Die y-Koordinate eines Punktes entspricht dem Wert von  $\hat{p}_r^{(J,Z)}$  für einen Probanden  $J$  und ein Schlafstadium  $Z$ , die x-Koordinate ist eine normalverteilte Zufallsvariable (Punkte sind „gejittert“).

In Abbildung 4.5 sind die SMI-Modell-Schätzer  $\hat{p}_A^{(J,Z)}, \hat{p}_B^{(J,Z)}, \hat{p}_C^{(J,Z)}, \hat{p}_D^{(J,Z)}$  jeweils nach Schlafstadium  $Z$  gruppiert dargestellt. Es ist zu erkennen, dass sich pro festen Zustand  $r$  und festes Schlafstadium  $Z$  die Werte  $\hat{p}_r^{(J,Z)}$  der verschiedenen Probanden stets um weniger als 0.4 untereinander unterscheiden. Der kleinste aus den Daten berechnete Wert für einen Schätzer  $\hat{p}_r$  ist  $> 0.08$  und der größte berechnete Wert ist  $< 0.53$ .

Für einen festen Zustand  $r \in \{A, B, C\}$  haben jeweils die Schätzer  $\hat{p}_r^{(J,Z)}$  für die Schlafstadien W, N1, N3 ungefähr die gleiche Größenordnung. Die Schätzer zum Schlafstadium N2 weichen ab von dieser Größenordnung, für die Zustände A und B nach oben und für den Zustand C nach unten.

Die Schätzer  $\hat{p}_D^{(J,Z)}$  werden im Durchschnitt kleiner mit der Tiefe des Schlafes.

In Abbildung 4.6 sind die Schätzer  $\hat{\pi}_{(J,Z)}^{(u)}$  gruppiert nach Schlafstadium zu sehen. Die graue gestrichelte Linie in dieser Abbildung gibt die Höhe 1 an. Im Schlafstadium W liegen die Schätzer von sieben Probanden über der gestrichelten Linie, für das Schlafstadium N1 liegt der Schätzer eines Probanden darüber und für die Schlafstadien N2

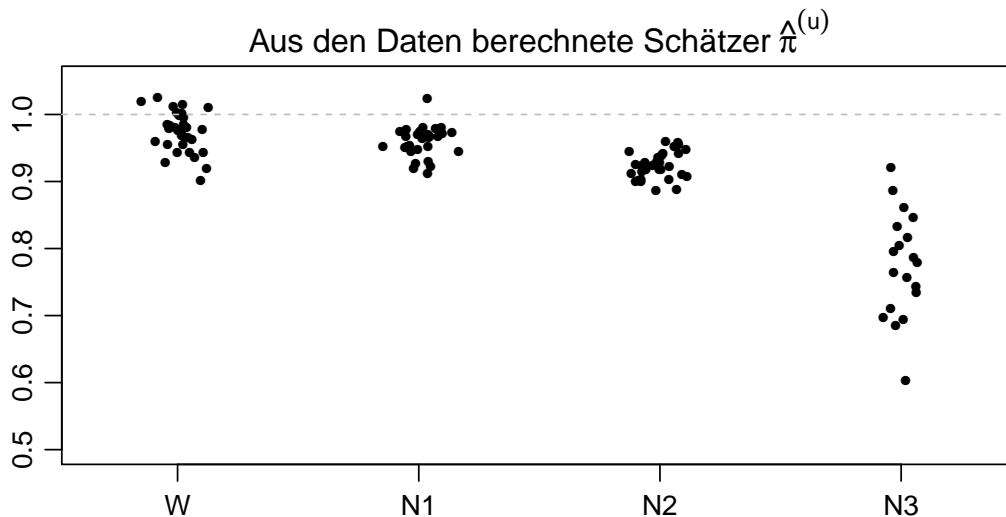


Abbildung 4.6: Aus den Daten berechnete Schätzer  $\hat{\pi}_{(J,Z)}^{(u)}$  getrennt nach Schlafstadium  $Z$ . Die y-Koordinate eines Punktes entspricht dem Wert von  $\hat{\pi}_{(J,Z)}^{(u)}$  für einen Probanden  $J$  und ein Schlafstadium  $Z$ , die x-Koordinate ist eine normalverteilte Zufallsvariable (Punkte sind „gejittert“).

und N3 sind alle Schätzer kleiner als 1. Um aus  $\hat{\pi}_{(J,Z)}^{(u)}$  den SMI-Modell-Schätzer  $\hat{\pi}^{(J,Z)}$  zu gewinnen, wird das Minimum

$$\hat{\pi}^{(J,Z)} = \hat{\pi}_{(J,Z)}^{(u)} \wedge 1$$

gebildet.

Im Durchschnitt nehmen die Werte der Schätzer  $\hat{\pi}_{(J,Z)}^{(u)}$  mit der Tiefe des Schlafes ab. Während in den Schlafstadien W, N1 und N2 die Werte im Bereich von ungefähr 0.90 bis 1.02 liegen, beobachtet man für das Schlafstadium N3 Werte zwischen 0.60 und 0.93, wobei der Durchschnitt hier bei ca. 0.77 liegt.

Je kleiner der Wert für  $\pi$  bzw. für  $\hat{\pi}$  ist, desto wahrscheinlicher ist es im Hintergrundmodell, dass bei einem Übergang ein Intervallwechsel stattfindet. Nach Proposition 3.5 ist

$$\pi = \frac{\lambda_1}{\lambda_1 + \lambda_2},$$

wobei  $\lambda_1$  der Parameter des homogenen Poisson'schen Punktprozesses ist, durch den die zufälligen Intervalle  $(I_j)_{j \in \mathbb{N}}$  gegeben werden, und  $\lambda_2$  der Parameter des homogenen Poisson'schen Punktprozesses ist, durch den die zufälligen Abtastungszeitpunkte  $(B_i)_{i \in \mathbb{N}_0}$  gegeben werden (siehe Definition von Modell 3 in Kapitel 3.2). Ein kleines  $\pi$  bedeutet, dass  $\lambda_1$  klein ist relativ zu  $\lambda_2$ , also dass in Erwartung in jedem Intervall  $I_j$  mehr Abtastungszeitpunkte liegen als für ein großes  $\pi$ .

Zum Abschluss dieses Kapitels werden in Abbildung 4.7 die geschätzten Übergangswahrscheinlichkeiten  $\hat{p}_{rs}$  des SMI-Modells (Kreuze) als Beispiel für die Probanden 2 und 5 im Schlafstadium N2 gezeigt<sup>16</sup> und mit den geschätzten Übergangswahrscheinlichkeiten  $\hat{p}_{rs}^{(v)}$  des vollen Modells (Kreise) verglichen.

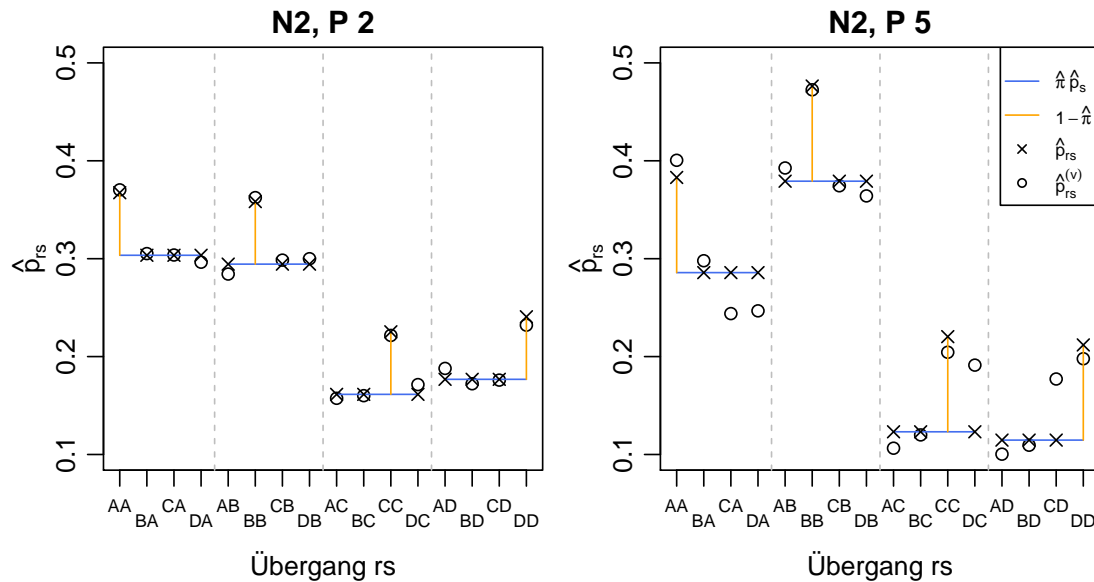


Abbildung 4.7: Geschätzte Übergangswahrscheinlichkeiten  $\hat{p}_{rs}$  im SMI-Modell (Kreuze) für die Probanden 2 (links) und 5 (rechts) im Schlafstadium N2 im Vergleich zu den geschätzten Übergangswahrscheinlichkeiten  $\hat{p}_{rs}^{(v)}$  des vollen Modells (Kreise). Die blauen Linien geben die Höhe von  $\hat{\pi} \hat{p}_s$  an, die Längen der orangenen Linien entsprechen jeweils  $1 - \hat{\pi}$ .

Man sieht in Abbildung 4.7, dass sich die Schätzer des SMI-Modells und des vollen Modells für Proband 2 im Schlafstadium N2 annähernd gleichen. Dagegen weichen die Schätzer des SMI-Modells und des vollen Modells für Proband 5 im Schlafstadium N2 zum Beispiel bei den Übergängen  $(C, A)$ ,  $(D, A)$ ,  $(C, D)$  und  $(D, C)$  relativ deutlich voneinander ab.

Bevor die Anpassungsgüte des SMI-Modells in Kapitel 6 untersucht wird, beschäftigt sich das folgende Kapitel 5 zunächst mit den Varianzen der Schätzer  $\hat{p}_r$  und  $\hat{\pi}^{(u)}$ .

<sup>16</sup>Die Probanden 2 und 5 im Schlafstadium N2 wurden auch schon für das Beispiel 3.1 herangezogen.



## 5 Varianzen der Schätzer

Dieses Kapitel beschäftigt sich mit den Varianzen der beiden Schätzer  $\hat{p}_r$  und  $\hat{\pi}^{(u)}$ , die im vorangegangenen Kapitel für das SMI-Modell definiert wurden durch

$$\hat{p}_r = \frac{n_r}{n}$$

und

$$\hat{\pi}^{(u)} = \frac{n - \sum_r n_{rr}}{n(1 - \sum_s \hat{p}_s^2)}.$$

Für beide Varianzen werden Formeln hergeleitet und anhand dieser Formeln das Verhalten der beiden Varianzen in Abhängigkeit der Parameter  $\pi$  und  $p_1, \dots, p_a$  untersucht. Die Varianz des Schätzers  $\hat{\pi} = \hat{\pi}^{(g)}$  ist höchstens so groß wie die Varianz von  $\hat{\pi}^{(u)}$ , da  $\hat{\pi}$  das Minimum von  $\hat{\pi}^{(u)}$  und 1 ist.

In Kapitel 5.1 wird die Varianz von  $\hat{p}_r$  behandelt. Die Varianz von  $\hat{\pi}^{(u)}$  wird in Kapitel 5.3 mit der Delta-Methode<sup>17</sup> approximiert. Für diese Methode werden Formeln für die Kovarianzen  $\text{Cov}(n_{rs}, n_{tu})$  benötigt, deren Herleitung in Kapitel 5.2 erfolgt.

### 5.1 Varianz von $\hat{p}_r$

Ziel in diesem Kapitel ist es, die folgende Proposition zu beweisen.

**Proposition 5.1.** *Sei  $r \in S$ . Dann gilt für*

$$\hat{p}_r = \sum_{s \in S} n_{rs}/n$$

*im Kontext des SMI-Modells*

$$\text{Var}(\hat{p}_r) = \frac{2 - \pi}{\pi} \cdot \frac{p_r(1 - p_r)}{n} + \mathcal{O}(n^{-2}).$$

Zunächst wird in Kapitel 5.1.1 die Intuition zur Formel für  $\text{Var}(\hat{p}_r)$  im Fall  $n \rightarrow \infty$  erläutert, danach folgt in Kapitel 5.1.2 der Beweis zu Proposition 5.1.

Anschließend wird in Kapitel 5.1.3 das Verhalten der obigen Formel für  $\text{Var}(\hat{p}_r)$  in Abhängigkeit der Parameter  $p_1, \dots, p_a$  und  $\pi$  untersucht.

#### 5.1.1 Intuition

Geht man davon aus, dass  $\pi = 1$  gilt, ist die Auftrittshäufigkeit

$$n_r = \sum_{s \in S} n_{rs} = \sum_{i=0}^{n-1} \mathbf{1}_{\{X_i=r\}}$$

<sup>17</sup>Die Delta-Methode wird in Anhang B beschrieben.

des Zustandes  $r \in S$  in  $(X_0, \dots, X_{n-1})$  binomial( $n, p_r$ )-verteilt. Denn in Bildern des Hintergrundmodells findet bei jedem Übergang von  $X_i$  zu  $X_{i+1}$  mindestens ein Intervallwechsel statt zwischen den entsprechenden  $B_i$  und  $B_{i+1}$ , sodass alle  $X_i$  mit  $i \in \mathbb{N}_0$  unabhängig voneinander sind und jeweils mit Wahrscheinlichkeit  $p_r$  den Wert  $r$  annehmen. Damit folgt direkt die für den Binomial-Fall bekannte Varianz

$$\text{Var}(\hat{p}_r) = \frac{p_r(1-p_r)}{n}.$$

Für  $0 < \pi < 1$  sind  $X_0, \dots, X_{n-1}$  abhängig voneinander, weil zwei  $X_{i_1}$  und  $X_{i_2}$  denselben Wert annehmen müssen, wenn die entsprechenden Abtastungszeitpunkte  $B_{i_1}$  und  $B_{i_2}$  im selben Intervall liegen. Bedingt auf eine feste Aufteilung der Abtastungszeitpunkte  $(B_i)_{i \in \mathbb{N}_0}$  auf die Intervalle  $(I_j)_{j \in \mathbb{N}}$  lässt sich aus jedem von  $B_0, \dots, B_{n-1}$  getroffenen Intervall  $I_{j_k}$  genau ein  $B_{i_k} \in I_{j_k}$  auswählen, sodass die Zufallsvariablen  $X_{i_1}, \dots, X_{i_l}$  der verkleinerten Stichprobe unabhängig sind, wobei  $l \in \mathbb{N}$  die Anzahl der getroffenen Intervalle sei.

Wie in Kapitel 4.3 dargestellt, ist die erwartete Anzahl der von  $B_0, \dots, B_{n-1}$  getroffenen Intervalle gerade  $1 + \pi(n-1) = 1 - \pi + \pi n$ . Asymptotisch lässt sich der Term  $1 - \pi$  vernachlässigen, sodass man in Erwartung auf  $\pi n$  getroffene Intervalle kommt. Bedingt darauf, dass  $\pi n = l \in \mathbb{N}$  die Anzahl der getroffenen Intervalle und somit der unabhängigen  $X_{i_1}, \dots, X_{i_l}$  ist, ergibt sich die Varianz

$$\text{Var}\left(\frac{\sum_{k=1}^l \mathbf{1}_{\{X_{i_k}=r\}}}{l}\right) = \frac{p_r(1-p_r)}{l} = \frac{p_r(1-p_r)}{\pi n}.$$

Im Vergleich zum Fall  $\pi = 1$  erhält man also für den Fall  $0 < \pi < 1$  durch den kleineren Stichprobenumfang von  $\pi n$  statt  $n$  unabhängiger Zufallsvariablen eine um den Faktor  $1/\pi$  größere Varianz.

Da beim Schätzer  $\hat{p}_r$  für  $p_r$  nicht nur die ausgewählten  $X_{i_1}, \dots, X_{i_l}$ , sondern alle Zufallsvariablen  $X_0, \dots, X_{n-1}$  verwendet werden, vergrößert sich die Varianz für  $0 < \pi < 1$  im Vergleich zum Binomial-Fall  $\pi = 1$  durch einen weiteren Effekt.

Für  $0 < \pi < 1$  treten Sequenzen  $(B_i, \dots, B_{i+m})$  der Länge  $m+1 > 1$  von Abtastungszeitpunkten, die alle im selben Intervall  $I_j$  liegen, mit positiver Wahrscheinlichkeit auf und dementsprechend auch Sequenzen  $(X_i, \dots, X_{i+m})$  von abhängigen Beobachtungen gleicher Ausprägung. Dadurch sind im Vergleich zum Binomial-Fall  $\pi = 1$  höhere Konzentrationen einzelner Zustände möglich.

Besonders deutlich wird dies für  $\pi \rightarrow 0$ . In diesem Fall gilt

$$\mathbb{P}(X_0 = \dots = X_{n-1}) \geq (1-\pi)^{n-1} \rightarrow 1.$$

Die höheren Konzentrationen lassen  $\hat{p}_r$  stärker um seinen Erwartungswert  $p_r$  streuen.

Tatsächlich vergrößert sich die Varianz asymptotisch noch einmal um den Faktor  $2 - \pi$ , d.h. insgesamt um den Faktor  $(2 - \pi)/\pi$ , wie in Proposition 5.1 gezeigt wird.

### 5.1.2 Mathematische Herleitung

Nachdem die intuitiven Einflussfaktoren beschrieben worden sind, die für  $0 < \pi < 1$  im Vergleich zu  $\pi = 1$  zu einer Vergrößerung der Varianz von  $\hat{p}_r$  führen, wird nun der Beweis für Proposition 5.1 geführt.

*Beweis (Proposition 5.1).* Sei  $r \in S$ . Mit  $Z_i := Z_{i,r} := \mathbf{1}_{\{X_i=r\}}$  erhält man

$$\hat{p}_r = \frac{n_r}{n} = \frac{1}{n} \sum_{i=0}^{n-1} Z_i.$$

Wegen

$$\text{Var}(\hat{p}_r) = \frac{1}{n^2} \text{Var} \left( \sum_{i=0}^{n-1} Z_i \right) = \frac{1}{n^2} \left( \sum_{i=0}^{n-1} \text{Var}(Z_i) + 2 \sum_{\substack{i,j=0 \\ i < j}}^{n-1} \text{Cov}(Z_i, Z_j) \right) \quad (5.1)$$

genügt es zunächst,  $\text{Cov}(Z_i, Z_j)$  und  $\text{Var}(Z_i) = \text{Cov}(Z_i, Z_i)$  zu berechnen.

Seien  $i, j \in \{0, \dots, n-1\}$  nicht notwendigerweise verschieden, o.B.d.A sei  $i \leq j$ . Da  $Z_i, Z_j$  Indikatorvariablen sind, gilt

$$\begin{aligned} \text{Cov}(Z_i, Z_j) &= \mathbb{E}[Z_i Z_j] - \mathbb{E}[Z_i] \mathbb{E}[Z_j] \\ &= \mathbb{P}(X_i = r, X_j = r) - \mathbb{P}(X_i = r) \mathbb{P}(X_j = r) \\ &= \mathbb{P}(X_i = r, X_j = r) - p_r^2. \end{aligned}$$

Unter der Bedingung, dass die zu  $X_i$  und  $X_j$  gehörenden Abtastungszeitpunkte  $B_i$  und  $B_j$  in verschiedenen Intervallen liegen, sind  $X_i$  und  $X_j$  unabhängig und insbesondere ist

$$\mathbb{P}(X_i = r, X_j = r) = p_r^2.$$

In diesem Fall ist also  $\text{Cov}(Z_i, Z_j) = 0$ .

Bedingt man dagegen darauf, dass die entsprechenden  $B_i$  und  $B_j$  sich im selben Intervall befinden, so nehmen  $X_i$  und  $X_j$  den gleichen Wert an, sodass

$$\mathbb{P}(X_i = r, X_j = r) = p_r$$

folgt. Dieser Fall tritt genau dann ein, wenn bei allen  $j - i$  Übergängen

$$(X_i, X_{i+1}), \dots, (X_{j-1}, X_j)$$

zwischen  $X_i$  und  $X_j$  kein Intervallwechsel stattfindet. Dies geschieht pro Übergang unabhängig mit Wahrscheinlichkeit  $1 - \pi$ . Insgesamt erhält man also

$$\mathbb{P}(B_i \text{ und } B_j \text{ liegen im selben Intervall}) = (1 - \pi)^{j-i}.$$

Diese Überlegungen führen mit dem Satz von der totalen Wahrscheinlichkeit zu

$$\begin{aligned}\text{Cov}(Z_i, Z_j) &= \mathbb{P}(X_i = r, X_j = r) - p_r^2 \\ &= (1 - \pi)^{j-i}(p_r - p_r^2) + (1 - (1 - \pi)^{j-i}) \cdot 0 \\ &= (1 - \pi)^{j-i}p_r(1 - p_r).\end{aligned}$$

Insbesondere impliziert dies im Fall  $i = j$  die bekannte Varianz

$$\text{Var}(Z_i) = p_r(1 - p_r)$$

für die Indikatorvariable  $Z_i$ .

Entscheidend für den Wert der Kovarianz  $\text{Cov}(Z_i, Z_j)$  ist die Differenz  $j - i$ . Es gibt genau  $n - d$  Paare  $(X_i, X_j)$  mit  $j - i = d$  für  $i, j \in \{0, \dots, n - 1\}$ ,  $i < j$ . Der kleinste Wert, den  $d$  annehmen kann, ist 1, der größte  $n - 1$ . Für die Summe der Kovarianzen aus Gleichung (5.1) ergibt sich daher

$$\begin{aligned}\sum_{\substack{i, j=0 \\ i < j}}^{n-1} \text{Cov}(Z_i, Z_j) &= \sum_{d=1}^{n-1} (n - d)(1 - \pi)^d p_r(1 - p_r) \\ &= \sum_{k=1}^{n-1} \sum_{d=1}^k (1 - \pi)^d p_r(1 - p_r) \\ &= \sum_{k=1}^{n-1} \left( \frac{1 - (1 - \pi)^{k+1}}{1 - (1 - \pi)} - 1 \right) p_r(1 - p_r) \\ &= \sum_{k=1}^{n-1} \left( \frac{1 - \pi - (1 - \pi)^{k+1}}{\pi} \right) p_r(1 - p_r) \\ &= \frac{1 - \pi}{\pi} \left( (n - 1) - \sum_{k=1}^{n-1} (1 - \pi)^k \right) p_r(1 - p_r) \\ &= \frac{1 - \pi}{\pi} \left( n - \frac{1 - (1 - \pi)^n}{\pi} \right) p_r(1 - p_r),\end{aligned}$$

wobei in der zweiten Zeile die Summanden umsortiert werden und in der dritten und sechsten Zeile jeweils die Formel für Partialsummen von Geometrischen Reihen verwendet wird.

Mit diesem Ergebnis sowie mit  $\text{Var}(Z_i) = p_r(1 - p_r)$  folgt aus Gleichung (5.1) die gesuchte Formel

$$\begin{aligned}\text{Var}(\hat{p}_r) &= \frac{1}{n^2} \left( np_r(1 - p_r) + 2 \frac{1 - \pi}{\pi} \left( n - \frac{1 - (1 - \pi)^n}{\pi} \right) p_r(1 - p_r) \right) \\ &= \frac{(\pi + 2 - 2\pi)p_r(1 - p_r)}{\pi n} - 2(1 - \pi)(1 - (1 - \pi)^n) \frac{p_r(1 - p_r)}{\pi^2 n^2} \\ &= \frac{(2 - \pi)p_r(1 - p_r)}{\pi n} + \mathcal{O}(n^{-2}).\end{aligned} \quad \square$$

**Bemerkung 5.2.** Der durch  $\mathcal{O}(n^{-2})$  zusammengefasste, asymptotisch vernachlässigbare Term

$$-2(1-\pi)(1-(1-\pi)^n)\frac{p_r(1-p_r)}{\pi^2 n^2}$$

von  $\text{Var}(\hat{p}_r)$  ist negativ, d.h. die Varianz wird durch  $\text{Var}(\hat{p}_r) \approx (2-\pi)p_r(1-p_r)/(\pi n)$  überschätzt.

Da  $p_r(1-p_r)$  sein Maximum  $1/4$  für  $p_r = 1/2$  annimmt und  $1-\pi < 1$  gilt, lässt sich der absolute Fehler abschätzen durch

$$\left| -2(1-\pi)(1-(1-\pi)^n)\frac{p_r(1-p_r)}{\pi^2 n^2} \right| < \frac{1}{2\pi^2 n^2}.$$

Je größer  $n$  ist bzw. je näher  $\pi$  bei der 1 liegt, desto kleiner ist die angegebene obere Schranke  $(2\pi^2 n^2)^{-1}$  des absoluten Fehlers.

### 5.1.3 Einfluss der Parameter des SMI-Modells auf $\text{Var}(\hat{p}_r)$

In diesem Abschnitt wird der Einfluss der Parameter  $p_1, \dots, p_a$  und  $\pi$  des SMI-Modells auf die Varianz von  $\hat{p}_r$  gemäß der Formel

$$\text{Var}(\hat{p}_r) = \frac{2-\pi}{\pi} \cdot \frac{p_r(1-p_r)}{n} + \mathcal{O}(n^{-2})$$

von Proposition 5.1 untersucht, wobei der asymptotisch vernachlässigbare Term  $\mathcal{O}(n^{-2})$  nicht berücksichtigt wird.

#### **Einfluss von $p_1, \dots, p_a$**

Seien  $\pi \in (0, 1]$  und  $n \in \mathbb{N}$  fest gegeben. Dann ist  $p_r(1-p_r)$  der verbleibende variable Term in der Formel von  $\text{Var}(\hat{p}_r)$ . Die Varianz von  $\hat{p}_r$  hängt in diesem Fall nur von  $p_r$  ab und insbesondere bei gegebenem  $p_r$  nicht von den anderen Wahrscheinlichkeiten  $p_1, \dots, p_{r-1}, p_{r+1}, \dots, p_a$ , solange

$$\sum_{\substack{t=1 \\ t \neq r}}^a p_t = 1 - p_r$$

erfüllt wird.

Durch  $p_r(1-p_r)$  wird eine nach unten geöffnete Parabel beschrieben mit einem Maximum von  $1/4$  in  $p_r = 1/2$ . Die Parabel ist in Abbildung 5.1 auf der linken Seite dargestellt.

In den entarteten Fällen  $p_r = 0$  und  $p_r = 1$  gilt

$$p_r(1-p_r) = 0$$

und somit nach der Formel von Proposition 5.1

$$\text{Var}(\hat{p}_r) = 0.$$

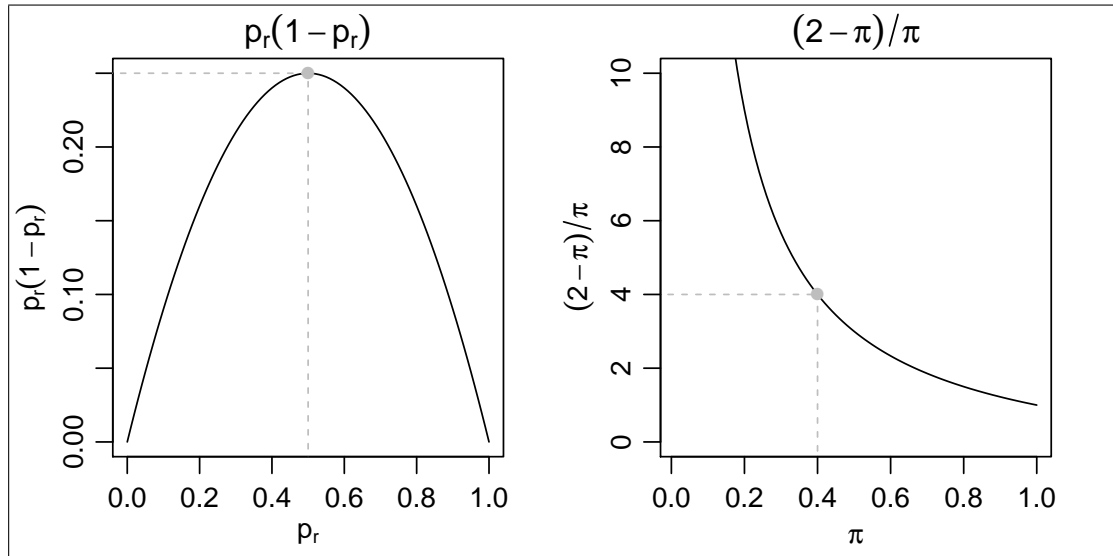


Abbildung 5.1: Graphen von  $p_r(1 - p_r)$  und  $(2 - \pi)/\pi$ .

Dieses Ergebnis für die Varianz deckt sich damit, dass für  $p_r = 0$  alle  $X_i \neq r$  und für  $p_r = 1$  alle  $X_i = r$  sind, also der Schätzer  $\hat{p}_r$  gegeben ist durch  $\hat{p}_r = 0$  bzw.  $\hat{p}_r = 1$ .

#### Einfluss von $\pi$

Seien nun  $n \in \mathbb{N}$  und  $p_1, \dots, p_a \in (0, 1)$  mit  $\sum_t p_t = 1$  fest gegeben. Der variable Term von  $\text{Var}(\hat{p}_r)$  lautet dann  $(2 - \pi)/\pi$  und ist in Abbildung 5.1 auf der rechten Seite für  $0 < \pi \leq 1$  illustriert.

Für  $\pi = 1$  sind alle  $X_i$  unabhängig voneinander, sodass man sich für  $n_r$  im Binomial- bzw. Multinomial-Fall befindet, der in Kapitel 3.2 beschrieben wurde. Dementsprechend gilt

$$\frac{2 - \pi}{\pi} = 1$$

mit der Konsequenz, dass die bekannte Formel

$$\text{Var}(\hat{p}_r) = \frac{p_r(1 - p_r)}{n}$$

folgt. Für  $\pi = 1$  nimmt  $(2 - \pi)/\pi$  zudem sein Minimum auf  $(0, 1]$  an.

Der Term  $(2 - \pi)/\pi$  ist streng monoton fallend für  $\pi > 0$ . D.h. je kleiner  $\pi$  ist, desto größer ist auch die Varianz von  $\hat{p}_r$ .<sup>18</sup> Insbesondere gilt

$$\lim_{\pi \rightarrow 0} \frac{2 - \pi}{\pi} = \infty.$$

<sup>18</sup>Die Intuition zu diesem Umstand wurde bereits im Kapitel 5.1.1 beschrieben.

## Fazit

Zusammenfassend bestätigt sich der Eindruck aus Kapitel 4.1.2, dass die Varianz von  $\hat{p}_r$  umso größer ist, je näher  $p_r$  bei 0.5 liegt und je kleiner  $\pi$  ist. Dabei ist der Einfluss von  $p_r$  auf die Varianz von  $\hat{p}_r$  beschränkt, während die Varianz durch den Einfluss von  $\pi$  beliebig groß werden kann.

## 5.2 Kovarianzen von Auftritts- und Übergangshäufigkeiten in Markov-Ketten

Für die Approximation von  $\text{Var}(\hat{\pi})$  mithilfe der Delta-Methode, die in Anhang B erläutert wird, werden die Kovarianzen  $\text{Cov}(n_{rs}, n_{tu})$  zweier Übergangshäufigkeiten  $n_{rs}$  und  $n_{tu}$  mit  $r, s, t, u \in S$  benötigt, deren Berechnung im Folgenden das Ziel ist.

Dafür werden zunächst in Kapitel 5.2.1 die Kovarianzen  $\text{Cov}(n_r, n_s)$  zweier Auftritts-häufigkeiten  $n_r$  und  $n_s$  der Zustände  $r$  bzw.  $s$  in den ersten  $n$  Zuständen  $X_0, \dots, X_{n-1}$  bestimmt für eine beliebige homogene Markov-Kette  $(X_i)_{i \in \mathbb{N}_0}$  1. Ordnung mit endlichem Zustandsraum und einer Übergangsmatrix, die den Eigenwert 1 mit Multiplizität 1 und sonst keine anderen Eigenwerte auf dem Einheitskreis besitzt.

Der Beweis für die Formel für die Kovarianzen  $\text{Cov}(n_r, n_s)$  zweier Auftrittshäufigkeiten wurde von Good (1961) erbracht.

Aus der Formel von Good lässt sich dann in Kapitel 5.2.2 die Kovarianz  $\text{Cov}(n_{rs}, n_{tu})$  zweier Übergangshäufigkeiten herleiten. Die Idee ist dabei, dass man die Paare

$$Y_i := (X_{i-1}, X_i) \quad \forall i \in \{1, \dots, n\}$$

als Markov-Kette mit Zustandsraum  $S \times S$  betrachtet und die Formel für  $\text{Cov}(n_{r'}, n_{s'})$  anwendet mit

$$r' := (r, s) \quad \text{und} \quad s' := (t, u).$$

### 5.2.1 Formel für $\text{Cov}(n_r, n_s)$

Es wird nun die bereits erwähnte Formel für die Kovarianz zweier Auftrittshäufigkeiten von Zuständen in einer Markov-Kette angegeben und bewiesen. Dabei wird der Darstellung von Good (1961) gefolgt.

**Proposition 5.3** (Good). *Sei  $(X_i)_{i \in \mathbb{N}_0}$  eine Markov-Kette mit endlichem Zustandsraum  $S$ , deren Übergangsmatrix  $P$  den Eigenwert 1 mit algebraischer Multiplizität 1 und sonst nur Eigenwerte mit Betrag  $< 1$  besitze. Es bezeichne  $e := (1, \dots, 1)^\top \in \mathbb{R}^{|S|}$  den rechten und  $p = (p_1, \dots, p_{|S|})^\top \in \mathbb{R}^{|S|}$  den linken Eigenvektor zum Eigenwert 1. Die Startverteilung sei zudem durch  $p^\top$  gegeben. Mit*

$$R := P - ep^\top,$$

der Einheitsmatrix  $I$  und der Auftrittshäufigkeit

$$n_r = \sum_{i=0}^{n-1} \mathbf{1}_{\{X_i=r\}}$$

eines Zustandes  $r \in S$  unter den ersten  $n$  Zuständen  $X_0, \dots, X_{n-1}$  der Markov-Kette gilt für alle  $r, s \in S$  dann

$$\text{Cov}(n_r, n_s) = n \left( \delta_{r,s} p_r - p_r p_s + p_r (R(I - R)^{-1})_{r,s} + p_s (R(I - R)^{-1})_{s,r} \right) + \mathcal{O}(1).$$

*Beweis.* Seien  $r, s \in S$ . Der Ansatz für die Bestimmung der Kovarianz von  $n_r$  und  $n_s$  ist in diesem Beweis

$$\text{Cov}(n_r, n_s) = \mathbb{E}[n_r n_s] - \mathbb{E}[n_r] \mathbb{E}[n_s]. \quad (5.2)$$

Zunächst wird für den zweiten Summanden eine Formel hergeleitet und danach für den ersten Summanden.

Da  $p$  der linke Eigenvektor von  $P$  zum Eigenwert 1 mit algebraischer Multiplizität 1 ist, ist die durch  $p^\top$  gegebene Startverteilung die stationäre Verteilung zur Übergangsmatrix  $P$ . Daher gilt

$$\mathbb{P}(X_i = r) = p_r \quad \forall i \in \mathbb{N}_0$$

und weiterhin

$$\mathbb{E}[n_r] = \mathbb{E} \left[ \sum_{i=0}^{n-1} Z_{i,r} \right] = \sum_{i=0}^{n-1} \mathbb{E}[Z_{i,r}] = np_r. \quad (5.3)$$

mit  $Z_{i,t} := \mathbf{1}_{\{X_i=t\}}$  für alle  $i \in \mathbb{N}_0$  und  $t \in S$ .

Für den zweiten Summanden von (5.2) ergibt sich also

$$-\mathbb{E}[n_r] \mathbb{E}[n_s] = -n^2 p_r p_s. \quad (5.4)$$

Im Folgenden wird der erste Summand  $\mathbb{E}[n_r n_s]$  bestimmt.

Wegen

$$n_r n_s = \left( \sum_{i=0}^{n-1} Z_{i,r} \right) \left( \sum_{j=0}^{n-1} Z_{j,s} \right) = \sum_{i,j=0}^{n-1} Z_{i,r} Z_{j,s}$$

genügt es,  $\mathbb{E}[Z_{i,r} Z_{j,s}]$  für alle  $(i, j) \in \{0, \dots, n-1\}^2$  zu betrachten.

Sei zunächst  $i = j$ , dann ist

$$Z_{i,r} Z_{i,s} = \begin{cases} Z_{i,r} & , \text{ falls } r = s, \\ 0 & , \text{ falls } r \neq s \end{cases}$$

und somit

$$\mathbb{E}[Z_{i,r} Z_{i,s}] = \mathbb{E}[\delta_{r,s} Z_{i,r}] = \delta_{r,s} p_r.$$



Da es genau  $n$  Möglichkeiten für  $(i, j) \in \{0, \dots, n-1\}^2$  mit  $i = j$  gibt, folgt

$$\mathbb{E}[n_r n_s] = \sum_{i,j=0}^{n-1} \mathbb{E}[Z_{i,r} Z_{j,s}] = n\delta_{r,s} p_r + \sum_{\substack{i,j=0 \\ i \neq j}}^{n-1} \mathbb{E}[Z_{i,r} Z_{j,s}]. \quad (5.5)$$

Sei nun  $i < j$  und  $d := j - i$ . Dann ergibt sich

$$\mathbb{E}[Z_{i,r} Z_{j,s}] = \mathbb{E}[Z_{i,r} Z_{i+d,s}] = p_r (P^d)_{r,s}.$$

Für ein  $d \in \{1, \dots, n-1\}$  gibt es genau  $n-d$  mögliche Paare  $(i, j) \in \{0, \dots, n-1\}^2$  mit  $d = j - i$ . Damit folgt

$$\sum_{\substack{i,j=0 \\ i < j}}^{n-1} \mathbb{E}[Z_{i,r} Z_{j,s}] = \sum_{d=1}^{n-1} (n-d) p_r (P^d)_{r,s} = p_r \left( \sum_{d=1}^{n-1} (n-d) P^d \right)_{r,s}. \quad (5.6)$$

Für den Fall  $i > j$  gilt analog

$$\sum_{\substack{i,j=0 \\ i > j}}^{n-1} \mathbb{E}[Z_{i,r} Z_{j,s}] = p_s \left( \sum_{d=1}^{n-1} (n-d) P^d \right)_{s,r}. \quad (5.7)$$

Setzt man die Gleichungen (5.6) und (5.7) in Gleichung (5.5) ein, erhält man

$$\begin{aligned} \mathbb{E}[n_r n_s] &= n\delta_{r,s} p_r + \sum_{\substack{i,j=0 \\ i < j}}^{n-1} \mathbb{E}[Z_{i,r} Z_{j,s}] + \sum_{\substack{i,j=0 \\ i > j}}^{n-1} \mathbb{E}[Z_{i,r} Z_{j,s}] \\ &= n\delta_{r,s} p_r + p_r \left( \sum_{d=1}^{n-1} (n-d) P^d \right)_{r,s} + p_s \left( \sum_{d=1}^{n-1} (n-d) P^d \right)_{s,r}. \end{aligned} \quad (5.8)$$

Die Matrix  $\sum_{d=1}^{n-1} (n-d) P^d$  lässt sich durch Umsortieren und mithilfe der Formel für Partialsummen von Geometrischen Reihen vereinfachen. Da sich diese Formel nur auf Matrizen anwenden lässt, deren Eigenwerte alle  $< 1$  sind, wird  $P$  in zwei Summanden  $R$  und  $ep^\top$  zerlegt, sodass der Summand  $R$  nur Eigenwerte  $< 1$  besitzt.

Der Summand  $ep^\top$  hat zwar den Eigenwert 1, wird sich aber ohne die Formel für Partialsummen Geometrischer Reihen auf einfache Weise auswerten lassen.

Es sei also

$$R := P - ep^\top.$$

Der einzige Eigenwert von  $P$  vom Betrag 1 ist 1 und gehört zum rechten Eigenvektor  $e$  bzw. zum linken Eigenvektor  $p$ . Alle anderen Eigenwerte von  $P$  sind  $< 1$ . Da die Zeilen der Matrix  $ep^\top$  alle gleich sind, hat sie Rang 1. Deshalb hat  $ep^\top$  nur einen Eigenwert

$> 0$ . Dieser Eigenwert ist die 1 mit dem zugehörigen rechten Eigenvektor  $e$  bzw. dem zugehörigen linken Eigenvektor  $p$ , denn wegen

$$p^\top e = p_1 + \cdots + p_{|S|} = 1$$

gelten

$$ep^\top e = e \quad \text{und} \quad p^\top ep^\top = p^\top.$$

Daraus lässt sich folgern, dass die Differenz  $R = P - ep^\top$  der beiden Matrizen nur Eigenwerte vom Betrag  $< 1$  besitzt. Insbesondere ergibt sich

$$Re = Pe - ep^\top e = e - e = 0$$

und

$$p^\top R = p^\top P - p^\top ep^\top = p^\top - p^\top = 0.$$

Es wird nun gezeigt, dass  $P^d$  in Gleichung (5.8) durch  $ep^\top + R^d$  ersetzt werden kann.

Multipliziert man  $P^d = (ep^\top + R)^d$  aus, treten die Terme  $(ep^\top)^d$  und  $R^d$  jeweils genau einmal auf und sonst nur gemischte Terme, die mindestens einen der Faktoren  $ep^\top R$  und  $Rep^\top$  enthalten und jeweils 0 ergeben, da  $p^\top R = 0$  und  $Re = 0$  gelten.

Wegen  $p^\top e = 1$  folgt zudem

$$(ep^\top)^d = e(p^\top e)^{d-1}p^\top = ep^\top \quad \forall d \geq 1.$$

Somit ist gezeigt, dass

$$P^d = ep^\top + R^d \quad \forall d \geq 1 \tag{5.9}$$

ist.

Die Formel für Partialsummen von Geometrischen Reihen für die Matrix  $R$  lautet

$$\sum_{d=0}^{n-1} R^d = (I - R^n)(I - R)^{-1}.$$

Dabei ist die Invertierbarkeit von  $I - R$  dadurch sichergestellt, dass  $R$  nur Eigenwerte  $< 1$  besitzt. Denn wäre  $I - R$  nicht invertierbar, gäbe es einen Vektor  $v \in \mathbb{R}^{|S|}$  mit  $(I - R)v = 0$ , also mit  $v = Rv$ . Dies ist ein Widerspruch dazu, dass  $R$  keinen Eigenwert 1 hat.

Mit den obigen Überlegungen lässt sich der Ausdruck  $\sum_{d=1}^{n-1} (n-d)P^d$  aus (5.8) auswerten.

Es gilt

$$\begin{aligned} \sum_{d=1}^{n-1} (n-d)P^d &= \sum_{d=1}^{n-1} (n-d)(ep^\top + R^d) \\ &= \frac{n(n-1)}{2}ep^\top + \sum_{d=1}^{n-1} (n-d)R^d. \end{aligned}$$

Der Term  $\sum_{d=1}^{n-1} (n-d)R^d$  kann nun mithilfe mehrfacher Anwendung der Formel für Partialsummen von Geometrischen Reihen und durch Umsortieren der Summanden vereinfacht werden:

$$\begin{aligned}
\sum_{d=1}^{n-1} (n-d)R^d &= nR \sum_{d=0}^{n-2} R^d - R \sum_{d=0}^{n-2} (d+1)R^d \\
&= nR(I - R^{n-1})(I - R)^{-1} - R \sum_{k=0}^{n-2} \sum_{d=k}^{n-2} R^d \\
&= n(R - R^n)(I - R)^{-1} - R \sum_{k=0}^{n-2} (R^k - R^{n-1})(I - R)^{-1} \\
&= n(R - R^n)(I - R)^{-1} + (n-1)R^n(I - R)^{-1} - R \sum_{k=0}^{n-2} R^k (I - R)^{-1} \\
&= nR(I - R)^{-1} - R^n(I - R)^{-1} - (R - R^n)(I - R)^{-2} \\
&= nR(I - R)^{-1} + \mathcal{O}(1).
\end{aligned}$$

Es folgt also

$$\sum_{d=1}^{n-1} (n-d)P^d = \frac{n(n-1)}{2} ep^\top + nR(I - R)^{-1} + \mathcal{O}(1). \quad (5.10)$$

Für Gleichung (5.8) wird der  $(r, s)$ -te bzw.  $(s, r)$ -te Eintrag der Matrix  $\sum_{d=1}^{n-1} (n-d)P^d$  benötigt. Mit  $(ep^\top)_{r,s} = p_s$  und Gleichung (5.10) erhält man

$$p_r \left( \sum_{d=1}^{n-1} (n-d)P^d \right)_{r,s} = \frac{n(n-1)}{2} p_r p_s + n p_r (R(I - R)^{-1})_{r,s} \quad (5.11)$$

und mit  $(ep^\top)_{s,r} = p_r$  analog

$$p_s \left( \sum_{d=1}^{n-1} (n-d)P^d \right)_{s,r} = \frac{n(n-1)}{2} p_r p_s + n p_s (R(I - R)^{-1})_{s,r}. \quad (5.12)$$

Die Behauptung der Proposition folgt nun, indem man in die Gleichung (5.2) zunächst die Gleichungen (5.4) und (5.8) einsetzt und anschließend die Gleichungen (5.11) und

(5.12):

$$\begin{aligned}
\text{Cov}(n_r, n_s) &= \mathbb{E}[n_r n_s] - \mathbb{E}[n_r] \mathbb{E}[n_s] \\
&= n\delta_{r,s}p_r + p_r \left( \sum_{d=1}^{n-1} (n-d)P^d \right)_{r,s} + p_s \left( \sum_{d=1}^{n-1} (n-d)P^d \right)_{s,r} - n^2 p_r p_s \\
&= n \left( \delta_{r,s}p_r + 2 \frac{n-1}{2} p_r p_s + p_r (R(I-R)^{-1})_{r,s} \right. \\
&\quad \left. + p_s (R(I-R)^{-1})_{s,r} - n p_r p_s \right) + \mathcal{O}(1) \\
&= n \left( \delta_{r,s}p_r - p_r p_s + p_r (R(I-R)^{-1})_{r,s} + p_s (R(I-R)^{-1})_{s,r} \right) + \mathcal{O}(1).
\end{aligned}$$

□

**Bemerkung 5.4.** Der Beweis zu Proposition 5.3 zeigt, dass der Fehlerterm, der durch  $\mathcal{O}(1)$  dargestellt wird, sich beläuft auf

$$\begin{aligned}
&- \left( p_r (R^n(I-R)^{-1} + (R+R^n)(I-R)^{-2})_{r,s} \right. \\
&\quad \left. + p_s (R^n(I-R)^{-1} + (R+R^n)(I-R)^{-2})_{s,r} \right).
\end{aligned}$$

Als Beispiel wird die Proposition 5.3 auf das SMI-Modell angewendet. Dabei wird sich ein alternativer Beweis für Proposition 5.1 zeigen.

**Beispiel 5.5.** Seien  $r, s \in S = \{1, \dots, a\}$ . Es wird nun für das SMI-Modell gezeigt, dass

$$\text{Cov}(n_r, n_s) = n \left( \frac{2-\pi}{\pi} p_r (\delta_{r,s} - p_s) \right) + \mathcal{O}(1) \tag{5.13}$$

gilt, woraus direkt die Formel

$$\text{Var}(\hat{p}_r) = \frac{2-\pi}{\pi n} p_r (1-p_r) + \mathcal{O}(n^{-2})$$

von Proposition 5.1 folgt.

Im SMI-Modell haben die Übergangswahrscheinlichkeiten die Form

$$p_{rs} = \pi p_s + \delta_{r,s}(1-\pi),$$

sodass man die Übergangsmatrix

$$P = \pi e p^\top + (1-\pi)I$$

erhält, wobei  $p^\top = (p_1, \dots, p_a)$  durch die Startverteilung gegeben sei.

Die Eigenwerte von  $P$  sind 1 mit algebraischer Multiplizität 1 und  $1-\pi$  mit algebraischer Multiplizität  $a-1$ . Dabei gilt  $1-\pi < 1$ , weil  $0 < \pi \leq 1$  ist. Der linke Eigenvektor zum

Eigenwert 1 ist  $p$  mit algebraischer Multiplizität 1, da  $p^\top$  die stationäre Verteilung der Markov-Kette ist nach Proposition 3.7.

Damit werden die Voraussetzungen für Proposition 5.3 erfüllt.

Es sei

$$R := P - ep^\top = \pi ep^\top + (1 - \pi)I - ep^\top = (1 - \pi)(I - ep^\top).$$

Nun lassen sich die Matrixeinträge  $(R(I - R)^{-1})_{r,s}$  und  $(R(I - R)^{-1})_{s,r}$  aus der Formel von Proposition 5.3 berechnen.

Wegen der Beziehung  $P^d = ep^\top + R^d$  für alle  $d \geq 1$  aus Gleichung (5.9) gilt

$$P(I - P) = P - P^2 = R - R^2 = R(I - R),$$

sodass sich mit  $Pe = e$  direkt

$$\pi R = \pi(P - ep^\top) = \pi P(I - ep^\top) = P(I - (1 - \pi)I - \pi ep^\top) = P(I - P) = R(I - R)$$

ergibt. Da nach dem Beweis zu Proposition 5.3 die Matrix  $I - R$  invertierbar ist, impliziert die letzte Gleichung

$$R(I - R)^{-1} = \frac{1}{\pi}R = \frac{1 - \pi}{\pi}(I - ep^\top)$$

und somit

$$(R(I - R)^{-1})_{r,s} = \frac{1 - \pi}{\pi}(\delta_{r,s} - p_s) \quad \text{sowie} \quad (R(I - R)^{-1})_{s,r} = \frac{1 - \pi}{\pi}(\delta_{r,s} - p_r).$$

Weiterhin gilt

$$p_r (R(I - R)^{-1})_{r,s} = \frac{1 - \pi}{\pi}(\delta_{r,s} p_r - p_r p_s) = \frac{1 - \pi}{\pi}(\delta_{r,s} p_s - p_r p_s) = p_s (R(I - R)^{-1})_{s,r}.$$

Aus Proposition 5.3 erhält man damit

$$\begin{aligned} \text{Cov}(n_r, n_s) &= n \left( \delta_{r,s} p_r - p_r p_s + p_r (R(I - R)^{-1})_{r,s} + p_s (R(I - R)^{-1})_{s,r} \right) + \mathcal{O}(1) \\ &= n \left( p_r (\delta_{r,s} - p_s) + 2 \frac{1 - \pi}{\pi} p_r (\delta_{r,s} - p_s) \right) + \mathcal{O}(1) \\ &= n \left( \frac{2 - \pi}{\pi} p_r (\delta_{r,s} - p_s) \right) + \mathcal{O}(1), \end{aligned}$$

also insbesondere im Fall  $r = s$  die in Proposition 5.1 berechnete Varianz

$$\text{Var}(\hat{p}_r) = \frac{1}{n^2} \text{Var}(n_r) = \frac{2 - \pi}{\pi} \cdot \frac{p_r(1 - p_r)}{n} + \mathcal{O}(n^{-2}).$$

### 5.2.2 Formel für $\text{Cov}(n_{rs}, n_{tu})$

In diesem Abschnitt wird für das SMI-Modell die Kovarianz der Übergangshäufigkeiten  $n_{rs}$  und  $n_{tu}$  bestimmt für Zustände  $r, s, t, u \in S = \{1, \dots, a\}$ . Zu diesem Zweck wird die Markov-Kette  $(Y_i)_{i \in \mathbb{N}}$  mit Zustandsraum  $S \times S$  betrachtet, die definiert sei durch

$$Y_i := (X_{i-1}, X_i) \quad \forall i \in \mathbb{N}.$$

Die Übergangshäufigkeit  $n_{rs}$  von einem Zustand  $r$  zu einem Zustand  $s$  in  $(X_0, \dots, X_n)$  ist dann gerade die Auftrittshäufigkeit des Zustandes  $(r, s)$  in  $(Y_1, \dots, Y_n)$ . Die Kovarianzen der Auftrittshäufigkeiten können daher mit Proposition 5.3 angewandt auf  $(Y_1, \dots, Y_n)$  bestimmt werden, sofern die Markov-Kette  $(Y_i)_{i \in \mathbb{N}}$  die Voraussetzungen dieser Proposition erfüllt.

Konkret ist es das Ziel, folgende Proposition zu beweisen.

**Proposition 5.6.** *Sei  $(X_0, \dots, X_n)$  die Sequenz der ersten  $n+1$  Zustände einer Markov-Kette  $(X_i)_{i \in \mathbb{N}_0}$ , die gemäß dem SMI-Modell gegeben sei mit Zustandsraum  $S = \{1, \dots, a\}$  für ein  $a \in \mathbb{N}$  und der stationären Verteilung*

$$p^\top = (p_1, \dots, p_a)$$

*sowie der Übergangsmatrix  $P = (p_{rs})_{r,s \in S}$ . Für die Übergangshäufigkeiten  $n_{rs}$  und  $n_{tu}$  von  $r$  zu  $s$  bzw. von  $t$  zu  $u$  in den ersten  $n$  Übergängen gilt*

$$\begin{aligned} & \text{Cov}(n_{rs}, n_{tu}) \\ &= n \left( \delta_{rs,tu} p_r p_{rs} - \frac{2 + \pi}{\pi} p_r p_{rs} p_t p_{tu} + \frac{1}{\pi} (\delta_{s,t} p_r p_{rs} p_{su} + \delta_{u,r} p_t p_{tu} p_{us}) \right) + \mathcal{O}(1). \end{aligned} \quad (5.14)$$

Der Beweis zur Proposition 5.6 wird sich auf drei Lemmata stützen. Diese Lemmata behandeln die Aussagen, dass

- (a) die oben definierte Folge  $(Y_i)_{i \in \mathbb{N}}$  eine Markov-Kette bildet mit zu bestimmender Übergangsmatrix  $Q$  und Startverteilung  $q^\top$ ,
- (b) die Markov-Kette  $(Y_i)_{i \in \mathbb{N}}$  die Voraussetzungen von Proposition 5.3 erfüllt mit stationärer Verteilung  $q^\top$  und
- (c) die Formel

$$R(I - R)^{-1} = \frac{1}{\pi} R$$

mit  $R := Q - eq^\top$  gilt.

Damit wird sich aus Proposition 5.3 die zu zeigende Formel (5.14) für  $\text{Cov}(n_{rs}, n_{tu})$  herleiten lassen.

**Lemma 5.7.** Sei  $(X_i)_{i \in \mathbb{N}_0}$  wie in Proposition 5.6 gegeben. Durch

$$Y_i := (X_{i-1}, X_i) \quad \forall i \in \mathbb{N}$$

wird dann eine Markov-Kette  $(Y_i)_{i \in \mathbb{N}}$  1. Ordnung definiert mit Zustandsraum  $S \times S$  und Startverteilung  $q^\top = (q_{(rs)})_{(r,s) \in S \times S}$ , gegeben durch

$$q_{(rs)} = p_r p_{rs} \quad \forall (r, s) \in S \times S,$$

sowie mit der Übergangsmatrix  $Q = (Q_{rs,tu})_{(r,s),(t,u) \in S \times S}$ , gegeben durch

$$Q_{rs,tu} = \delta_{s,t} p_{tu} \quad \forall (r, s), (t, u) \in S \times S.$$

Da die Zustände der  $Y_i$  die Übergänge  $(r, s)$  von einem  $r \in S$  zu einem  $s \in S$  in der Markov-Kette  $(X_i)_{i \in \mathbb{N}_0}$  sind, wurde in Lemma 5.7 eine neue Notation für Matrix- und Vektor-Einträge verwendet, die nun kurz erläutert wird. Danach folgt der Beweis zum Lemma.

Die Übergangswahrscheinlichkeit von  $(r, s)$  zu  $(t, u)$  stehe in der Übergangsmatrix  $Q$  in der  $((r-1)a + s)$ -ten Zeile und der  $((t-1)a + u)$ -ten Spalte, wobei  $a = |S|$  ist. Die Reihenfolge der Zustände in den Zeilen und Spalten sei also

$$(1, 1), (1, 2), \dots, (1, a), (2, 1), (2, 2), \dots, (2, a), \dots, (a, 1), (a, 2), \dots, (a, a).$$

Für den Eintrag in der zu  $(r, s)$  gehörenden Zeile und der zu  $(t, u)$  gehörenden Spalte der Matrix  $M$  wird die Notation

$$M_{(r,s),(t,u)} \text{ oder } M_{rs,tu}$$

genutzt. Analog wird der zu  $(r, s)$  gehörende Eintrag eines Vektors  $v$  mit  $v_{(r,s)}$  oder  $v_{rs}$  bezeichnet und in die  $((r-1)a + s)$ -te Zeile geschrieben.

Außerdem sei das Kronecker-Delta  $\delta_{rs,tu}$  definiert durch

$$\delta_{rs,tu} := \delta_{(r,s),(t,u)} := \begin{cases} 1 & , \text{ falls } (r, s) = (t, u), \\ 0 & , \text{ falls } (r, s) \neq (t, u). \end{cases}$$

*Beweis (Lemma 5.7).* Für den Nachweis der Markov-Eigenschaft 1. Ordnung seien  $i \in \mathbb{N}$  und  $(m_1, m'_1), \dots, (m_i, m'_i) \in S \times S$ , o.B.d.A. mit  $m'_k = m_{k+1}$  für alle  $k \in \{1, \dots, i-2\}$ , damit das Ereignis

$$\{Y_1 = (m_1, m'_1), \dots, Y_{i-1} = (m_{i-1}, m'_{i-1})\}$$

möglich ist. Ausdrücklich sei  $m'_{i-1} \neq m_i$  zugelassen. Dann gilt

$$\begin{aligned} & \mathbb{P}(Y_i = (m_i, m'_i) \mid Y_{i-1} = (m_{i-1}, m'_{i-1}), \dots, Y_1 = (m_1, m'_1)) \\ &= \mathbb{P}((X_{i-1}, X_i) = (m_i, m'_i) \mid X_{i-1} = m'_{i-1}, X_{i-2} = m'_{i-2}, \dots, X_1 = m'_1, X_0 = m_1) \\ &= \mathbb{P}((X_{i-1} = m_i, X_i = m'_i \mid X_{i-1} = m'_{i-1}, X_{i-2} = m'_{i-2}) \\ &= \mathbb{P}(Y_i = (m_i, m'_i) \mid Y_{i-1} = (m_{i-1}, m'_{i-1})), \end{aligned}$$

wobei in der dritten Zeile die Markov-Eigenschaft von  $(X_i)_{i \in \mathbb{N}_0}$  eingeht.

Die Startverteilung berechnet sich unmittelbar zu

$$q_{(rs)} = \mathbb{P}(Y_1 = (r, s)) = \mathbb{P}(X_0 = r, X_1 = s) = p_r p_{rs} \quad \forall (r, s) \in S \times S.$$

Ein Übergang in  $(Y_i)_{i \in \mathbb{N}}$  von  $(r, s)$  zu  $(t, u)$  ist nur möglich, wenn  $s = t$  ist, und geschieht dann mit der Wahrscheinlichkeit  $p_{tu}$ . Für die Übergangswahrscheinlichkeiten gilt also

$$Q_{rs,tu} = \delta_{s,t} p_{tu} \quad \forall (r, s), (t, u) \in S \times S.$$

□

Das nächste Beispiel soll die Struktur der Übergangsmatrix  $Q$  verdeutlichen.

**Beispiel 5.8.** Für  $a = 2$  sieht die Übergangsmatrix  $Q$  folgendermaßen aus:

$$\begin{array}{cccc} & (1,1) & (1,2) & (2,1) & (2,2) \\ \begin{array}{l} (1,1) \\ (1,2) \\ (2,1) \\ (2,2) \end{array} & \begin{pmatrix} p_{11} & p_{12} & 0 & 0 \\ 0 & 0 & p_{21} & p_{22} \\ p_{11} & p_{12} & 0 & 0 \\ 0 & 0 & p_{21} & p_{22} \end{pmatrix} \end{array}.$$

**Lemma 5.9.** Die Übergangsmatrix  $Q$  der Markov-Kette  $(Y_i)_{i \in \mathbb{N}}$  besitzt den Eigenwert 1 mit algebraischer Multiplizität 1 und sonst nur Eigenwerte mit Betrag  $< 1$ . D.h.  $(Y_i)_{i \in \mathbb{N}}$  erfüllt die Voraussetzungen von Proposition 5.3. Insbesondere ist  $q^\top$  die stationäre Verteilung.

*Beweis.* Es wird die folgende Behauptung gezeigt: Die Eigenwerte von  $Q$  sind 1 mit Multiplizität 1,  $1 - \pi$  mit Multiplizität  $a - 1$  sowie 0 mit Multiplizität  $a^2 - a$  für  $0 < \pi < 1$ . Ist  $\pi = 1$ , so hat der Eigenwert  $1 - \pi = 0$  die Multiplizität  $a^2 - 1$  und der Eigenwert 1 die Multiplizität 1, wobei alle angegebenen Multiplizitäten die algebraischen Multiplizitäten sind.

Sei zunächst  $0 < \pi < 1$ . Der rechte Eigenvektor zum Eigenwert 1 ist  $e$  und der transponierte linke Eigenvektor ist die Startverteilung  $q^\top$  aus Lemma 5.7, die gegeben ist durch

$$q_{(rs)} = p_r p_{rs} \quad \forall (r, s) \in S \times S.$$

Denn  $(p_1, \dots, p_a)$  ist nach Proposition 3.7 die stationäre Verteilung von  $(X_i)_{i \in \mathbb{N}_0}$  und für alle  $i \in \mathbb{N}$  gilt somit

$$\begin{aligned} \mathbb{P}(Y_i = (r, s)) &= \mathbb{P}(X_{i-1} = r, X_i = s) \\ &= \mathbb{P}(X_{i-1} = r) \mathbb{P}(X_i = s | X_{i-1} = r) \\ &= p_r p_{rs}. \end{aligned}$$



Also ist  $q^\top$  die stationäre Verteilung von  $(Y_i)_{i \in \mathbb{N}}$  und  $q$  ein linker Eigenvektor von  $Q$  zum Eigenwert 1.

Es wird nun gezeigt, dass alle anderen Eigenwerte von  $Q$  kleiner als 1 sind. Dafür wird nachgerechnet, dass sich  $a - 1$  linear unabhängige rechte Eigenvektoren  $v_1, \dots, v_{a-1}$  zum Eigenwert  $1 - \pi$  beschreiben lassen durch

$$(v_i)_{(rs)} = \delta_{i,s}p_a - \delta_{a,s}p_i \quad \forall i \in \{1, \dots, a - 1\}, (r, s) \in S \times S.$$

Beispielsweise haben für  $a = 3$  die Eigenvektoren  $v_1$  und  $v_2$  die Form

$$\begin{aligned} v_1^\top &= (p_3, 0, -p_1, p_3, 0, -p_1, p_3, 0, -p_1), \\ v_2^\top &= (0, p_3, -p_2, 0, p_3, -p_2, 0, p_3, -p_2). \end{aligned}$$

Für alle  $i, j \in \{1, \dots, a - 1\}$  ist

$$(v_i)_{1j} = \begin{cases} p_a & , \text{ falls } i = j, \\ 0 & , \text{ falls } i \neq j. \end{cases}$$

Daraus folgt wegen  $p_a > 0$ , dass  $v_1, \dots, v_{a-1}$  linear unabhängig sind

Dass es sich bei  $v_1, \dots, v_{a-1}$  um Eigenvektoren zum Eigenwert  $1 - \pi$  handelt, zeigt die folgende Rechnung:

$$\begin{aligned} (Qv_i)_{(rs)} &= \sum_{t=1}^a \sum_{u=1}^a Q_{rs,tu} (v_i)_{(tu)} \\ &= \sum_{t=1}^a \sum_{u=1}^a \delta_{s,t} (\pi p_u + \delta_{t,u} (1 - \pi)) (\delta_{i,u} p_a - \delta_{a,u} p_i) \\ &= \sum_{u=1}^a (\pi p_u + \delta_{s,u} (1 - \pi)) (\delta_{i,u} p_a - \delta_{a,u} p_i) \\ &= \pi p_i p_a - \pi p_a p_i + (1 - \pi) (\delta_{i,s} p_a - \delta_{a,s} p_i) \\ &= (1 - \pi) ((v_i)_{(rs)}). \end{aligned}$$

Die verbleibenden  $a^2 - a$  Eigenwerte müssen 0 sein, da die Übergangswahrscheinlichkeiten  $Q_{rs,tu} = \delta_{s,t} p_{tu}$  für alle  $r = 1, \dots, a$  gleich sind und  $Q$  somit nur genau  $a$  paarweise verschiedene Zeilen besitzt, d.h.

$$\dim(\text{Kern}(Q)) = a^2 - a.$$

Da die ermittelten geometrischen Multiplizitäten der Eigenwerte 1,  $1 - \pi$  und 0 in der Summe  $a^2$  ergeben, entsprechen die geometrischen den algebraischen Multiplizitäten.

Setzt man  $\pi = 1$ , gilt  $1 - \pi = 0$ , sodass der Eigenwert 0 in diesem Fall die algebraische Multiplizität  $(a - 1) + (a^2 - a) = a^2 - 1$  hat. Der Eigenwert 1 bleibt mit algebraischer Multiplizität 1 erhalten.

Damit ist gezeigt, dass  $(Y_i)_{i \in \mathbb{N}}$  die Voraussetzungen von Proposition 5.3 erfüllt.  $\square$

**Lemma 5.10.** Für die Übergangsmatrix  $Q$  und die stationäre Verteilung  $q^\top$  der Markov-Kette  $(Y_i)_{i \in \mathbb{N}}$  sei

$$R := Q - eq^\top.$$

Dann gilt

$$(R(I - R)^{-1})_{rs,tu} = \frac{1}{\pi} R_{rs,tu} = \frac{1}{\pi} (\delta_{s,t} - p_t) p_{tu} \quad \forall (r, s), (t, u) \in S \times S.$$

*Beweis.* Es gilt

$$R_{rs,tu} = Q_{rs,tu} - q_{tu} = \delta_{s,t} p_{tu} - p_t p_{tu} = (\delta_{s,t} - p_t) p_{tu}.$$

Nun wird

$$(I - R)^{-1} = I + \frac{1}{\pi} R$$

bewiesen, denn daraus folgt direkt

$$R(I - R)^{-1} = (I - (I - R))(I - R)^{-1} = (I - R)^{-1} - I = I + \frac{1}{\pi} R - I = \frac{1}{\pi} R$$

und somit die gesuchte Formel

$$(R(I - R)^{-1})_{rs,tu} = \frac{1}{\pi} (\delta_{s,t} - p_t) p_{tu}.$$

Zu diesem Zweck beobachtet man zunächst

$$(Q^2)_{rs,tu} = p_{st} p_{tu}.$$

Diese Formel lässt sich dadurch erklären, dass die Markov-Kette  $(Y_i)_{i \in \mathbb{N}}$  nur dann in zwei Schritten von  $(r, s)$  nach  $(t, u)$  gelangt, wenn der erste Schritt zu  $(s, t)$  und der zweite zu  $(t, u)$  führt.

Für den Beweis von

$$(I - R)^{-1} = I + \frac{1}{\pi} R$$

genügt es wegen

$$(I - R) \left( I + \frac{1}{\pi} R \right) = I + \frac{1}{\pi} ((1 - \pi)R - R^2),$$

zu zeigen, dass

$$(1 - \pi)R - R^2 = 0$$

ist. Unter Berücksichtigung von  $R^2 = Q^2 - eq^\top$  und der oben hergeleiteten Formel für  $(Q^2)_{rs,tu}$  erhält man

$$\begin{aligned} ((1 - \pi)R - R^2)_{rs,tu} &= \left( (1 - \pi)(Q - eq^\top) - Q^2 + eq^\top \right)_{rs,tu} \\ &= \left( (1 - \pi)Q + \pi eq^\top - Q^2 \right)_{rs,tu} \\ &= (1 - \pi)\delta_{s,t} p_{tu} + \pi p_t p_{tu} - p_{st} p_{tu} \\ &= (\pi p_t + \delta_{s,t}(1 - \pi)) p_{tu} - p_{st} p_{tu} \\ &= p_{st} p_{tu} - p_{st} p_{tu} \\ &= 0. \end{aligned}$$

Es folgt also insgesamt

$$(I - R)^{-1} = I + \frac{1}{\pi}R.$$

□

Damit sind alle Lemmata, die für den Beweis von Proposition 5.6 benötigt werden, bewiesen worden.

*Beweis (Proposition 5.6).* Die Proposition 5.3 gibt eine Formel an für die Kovarianzen der Auftrittshäufigkeiten zweier Zustände unter den ersten  $n$  Zuständen einer Markov-Kette, die gewisse Voraussetzungen erfüllt. Die Markov-Kette  $(Y_i)_{i \in \mathbb{N}}$  erfüllt nach Lemma 5.9 gerade diese Voraussetzungen.

Für  $(Y_1, \dots, Y_n)$  ist die Auftrittshäufigkeit eines Zustandes  $(r, s)$  gerade die Übergangshäufigkeit  $n_{rs}$  von  $r$  zu  $s$  in  $(X_0, \dots, X_n)$ , wobei  $(X_i)_{i \in \mathbb{N}_0}$  eine Markov-Kette gemäß dem SMI-Modell ist. Mit Proposition 5.3 gilt also

$$\begin{aligned} & \text{Cov}(n_{rs}, n_{tu}) \\ &= n \left( \delta_{rs,tu} q_{(rs)} - q_{(rs)} q_{(tu)} + q_{(rs)} (R(I - R)^{-1})_{rs,tu} + q_{(tu)} (R(I - R)^{-1})_{tu,rs} \right) + \mathcal{O}(1). \end{aligned}$$

Für  $(R(I - R)^{-1})_{rs,tu}$  lässt sich nach Lemma 5.10

$$(R(I - R)^{-1})_{rs,tu} = \frac{1}{\pi}(\delta_{s,t} - p_t)p_{tu} = \frac{1}{\pi}(\delta_{r,s}p_{tu} - q_{(tu)})$$

einsetzen, sodass sich folgende Rechnung ergibt:

$$\begin{aligned} & \text{Cov}(n_{rs}, n_{tu}) \\ &= n \left( \delta_{rs,tu} q_{(rs)} - q_{(rs)} q_{(tu)} + \frac{1}{\pi} q_{(rs)} (\delta_{s,t} p_{tu} - q_{(tu)}) + \frac{1}{\pi} q_{(tu)} (\delta_{u,r} p_{rs} - q_{(rs)}) \right) + \mathcal{O}(1) \\ &= n \left( \delta_{rs,tu} q_{(rs)} - q_{(rs)} q_{(tu)} - \frac{2}{\pi} q_{(rs)} q_{(tu)} + \frac{1}{\pi} (\delta_{s,t} q_{(rs)} p_{tu} + \delta_{u,r} q_{(tu)} p_{rs}) \right) + \mathcal{O}(1) \\ &= n \left( \delta_{rs,tu} p_r p_{rs} - \frac{2 + \pi}{\pi} p_r p_{rs} p_t p_{tu} + \frac{1}{\pi} (\delta_{s,t} p_r p_{rs} p_{su} + \delta_{u,r} p_t p_{tu} p_{us}) \right) + \mathcal{O}(1). \end{aligned}$$

□

**Bemerkung 5.11.** Durch Aufsummieren über alle  $s \in S$  und  $u \in S$  liefert Gleichung (5.14) von Proposition 5.6 asymptotisch die gleiche Formel für  $\text{Cov}(n_r, n_t)$ , die schon durch Gleichung (5.13) in Beispiel 5.5 angegeben wurde.

Für den Nachweis wird verwendet, dass für alle  $r, t \in S$  die Formeln

$$\sum_{s=1}^a \sum_{u=1}^a \delta_{rs,tu} p_r p_{rs} = \sum_{s=1}^a \delta_{r,t} p_r p_{rs}, \quad \sum_{s=1}^a \delta_{s,t} p_r p_{rs} = p_r p_{rt}, \quad \sum_{s=1}^a p_{rs} = 1$$

und

$$p_{rt} = \pi p_t + \delta_{r,t}(1 - \pi)$$

gelten.

Damit folgt

$$\begin{aligned} \text{Cov}(n_r, n_t) &= \text{Cov}\left(\sum_{s=1}^a n_{rs}, \sum_{u=1}^a n_{tu}\right) = \sum_{s=1}^a \sum_{u=1}^a \text{Cov}(n_{rs}, n_{tu}) \\ &= \sum_{s=1}^a \sum_{u=1}^a n \left( \delta_{rs,tu} p_r p_{rs} - \frac{2+\pi}{\pi} p_r p_{rs} p_t p_{tu} + \frac{1}{\pi} (\delta_{s,t} p_r p_{rs} p_{su} + \delta_{u,r} p_t p_{tu} p_{us}) \right) + \mathcal{O}(1) \\ &= \sum_{s=1}^a n \left( \delta_{r,t} p_r p_{rs} - \frac{2+\pi}{\pi} p_r p_{rs} p_t + \frac{1}{\pi} \delta_{s,t} p_r p_{rs} \right) + \sum_{u=1}^a n \left( \frac{1}{\pi} \delta_{u,r} p_t p_{tu} \right) + \mathcal{O}(1) \\ &= n \left( \delta_{r,t} p_r - \frac{2+\pi}{\pi} p_r p_t + \frac{1}{\pi} (p_r p_{rt} + p_t p_{tr}) \right) + \mathcal{O}(1) \\ &= n \left( \delta_{r,t} p_r - \frac{2+\pi}{\pi} p_r p_t + \frac{1}{\pi} (p_r (\pi p_t + \delta_{r,t}(1-\pi)) + p_t (\pi p_r + \delta_{r,t}(1-\pi))) \right) + \mathcal{O}(1) \\ &= n \left( \left(1 + 2 \frac{1-\pi}{\pi}\right) \delta_{r,t} p_r - \left(\frac{2+\pi-2\pi}{\pi}\right) p_r p_t \right) + \mathcal{O}(1) \\ &= n \left( \frac{2-\pi}{\pi} p_r (\delta_{r,t} - p_t) \right) + \mathcal{O}(1). \end{aligned}$$

### 5.3 Varianz von $\hat{\pi}^{(u)}$

In diesem Unterkapitel wird die Varianz vom Schätzer  $\hat{\pi}^{(u)}$  untersucht. Der Schätzer  $\hat{\pi}^{(u)}$  wurde in Kapitel 4.2 angegeben durch

$$\hat{\pi}^{(u)} = \frac{n - \sum_r n_{rr}}{n(1 - \sum_k \hat{p}_k^2)} = \frac{\sum_{r \neq s} n_{rs} / n}{1 - \sum_k \hat{p}_k^2}$$

und unterscheidet sich vom Schätzer  $\hat{\pi} = \hat{\pi}^{(g)}$  dadurch, dass  $\hat{\pi}^{(u)}$  nicht bei 1 abgeschnitten wird. Es muss also bei der Bestimmung der Varianz von  $\hat{\pi}^{(u)}$  kein Gewicht auf der 1 berücksichtigt werden. Die Varianz von  $\hat{\pi}$  ist durch die Varianz von  $\hat{\pi}^{(u)}$  nach oben beschränkt.

Da sich im Zähler und im Nenner von  $\hat{\pi}^{(u)}$  Zufallsvariablen befinden, die abhängig voneinander sind, wird die Delta-Methode zur Approximation der Varianz von  $\hat{\pi}^{(u)}$  verwendet. Die Delta-Methode wird in Anhang B beschrieben.

Die nachstehende Proposition gibt eine Formel für die Varianz von  $\hat{\pi}^{(u)}$  an.

**Proposition 5.12.** *Im Kontext des SMI-Modells gilt*

$$\text{Var}\left(\hat{\pi}^{(u)}\right) \approx \frac{-\pi^2 + \pi + \sum_r p_r^2 \left(2\pi^2 - \pi + (-4\pi^2 + 2\pi)p_r + (3\pi^2 - 2\pi) \sum_t p_t^2\right)}{n \left(1 - \sum_k p_k^2\right)^2} + \mathcal{O}(n^{-2}).$$

Diese Proposition wird in Kapitel 5.3.1 hergeleitet. Anschließend erfolgt in Kapitel 5.3.2 die Untersuchung des Einflusses von den Parametern  $p_1, \dots, p_a$  und  $\pi$  auf die Varianz von  $\hat{\pi}^{(u)}$ .

### 5.3.1 Herleitung

Für den Beweis zur Proposition 5.12 wird zunächst durch das nachstehende Lemma 5.13 eine Approximation für  $\text{Var}(\pi^{(u)})$  in Abhängigkeit von Kovarianzen von Auftretts- und Übergangshäufigkeiten angegeben, die mithilfe der Delta-Methode ermittelt wird. Danach werden die in Kapitel 5.2.1 berechneten Formeln für die Kovarianzen in diese Approximation eingesetzt.

**Lemma 5.13.** *Im Kontext des SMI-Modells gilt*

$$\text{Var}\left(\hat{\pi}^{(u)}\right) \approx \frac{\sum_{r=1}^a \sum_{t=1}^a \frac{\text{Cov}(n_{rr}, n_{tt}) - 4\pi p_r \text{Cov}(n_r, n_{tt}) + 4\pi^2 p_r p_t \text{Cov}(n_r, n_t)}{n^2 \left(1 - \sum_k p_k^2\right)^2}}{n^2 \left(1 - \sum_k p_k^2\right)^2}. \quad (5.15)$$

*Beweis.* Für den Beweis wird die Delta-Methode in vier Schritten angewendet:

1. Ermittlung eines Parameter-Vektors  $T = (T_1, \dots, T_l)$  minimaler Dimension,
2. Berechnung der partiellen Ableitungen

$$\frac{\partial \hat{\pi}^{(u)}(T)}{\partial T_i},$$

3. Einsetzen von  $\vartheta = \mathbb{E}[T]$  in die partiellen Ableitungen,
4. Anwendung der Formel

$$\text{Var}(\hat{\pi}^{(u)}) = \sum_{i=1}^l \sum_{j=1}^l \frac{\partial \hat{\pi}^{(u)}(T)}{\partial T_i} \Big|_{T=\vartheta} \cdot \frac{\partial \hat{\pi}^{(u)}(T)}{\partial T_j} \Big|_{T=\vartheta} \cdot \text{Cov}(T_i, T_j).$$

**1. Schritt:** Die Anzahl  $n \in \mathbb{N}$  der betrachteten Übergänge in der Markov-Kette  $(X_i)_{i \in \mathbb{N}_0}$  des SMI-Modells sei fest gegeben.

Wegen

$$n = \sum_{r,s=1}^a n_{rs}$$

kann o.B.d.A.  $n_{aa}$  dargestellt werden durch

$$n_{aa} = n - \sum_{(r,s) \neq (a,a)} n_{rs}.$$

Als Konsequenz kann die Formel für  $\hat{p}_k$  folgendermaßen angegeben werden:

$$\hat{p}_k = \begin{cases} \frac{1}{n} \sum_{l=1}^a n_{kl} & , \text{ falls } k \neq a, \\ \frac{1}{n} \left( n - \sum_{t=1}^{a-1} \sum_{u=1}^a n_{tu} \right) & , \text{ falls } k = a. \end{cases}$$

Es hängt also

$$\hat{\pi}^{(u)} = \frac{\sum_{i \neq j} n_{ij}/n}{1 - \sum_k \hat{p}_k^2}$$

ab von  $a^2 - 1$  Parametern, deren Vektor mit  $T$  bezeichnet werde:

$$T := (n_{11}, n_{12}, \dots, n_{1a}, \dots, n_{a-1,1}, n_{a-1,2}, \dots, n_{a-1,a}, n_{a1}, n_{a2}, \dots, n_{a,a-1}).$$

Die Formel für  $\hat{\pi}^{(u)}$  in Abhängigkeit von  $T$  ist gegeben durch

$$\hat{\pi}^{(u)}(T) = \frac{\sum_{i \neq j} n_{ij}/n}{1 - \frac{1}{n^2} \left( \sum_{k=1}^{a-1} \left( \sum_{l=1}^a n_{kl} \right)^2 + \left( n - \sum_{t=1}^{a-1} \sum_{u=1}^a n_{tu} \right)^2 \right)}.$$

Für die Anwendung der Delta-Methode auf  $\hat{\pi}^{(u)}$  zur Approximation der Varianz wird vorausgesetzt, dass  $\hat{p}_k \neq 1$  und  $p_k \neq 1$  sind für alle  $k \in S$ . Dadurch ist sichergestellt, dass  $\hat{\pi}^{(u)}$  zweifach stetig differenzierbar ist, weil in diesem Fall der Nenner von  $\hat{\pi}^{(u)}(T)$  sowie die Nenner der 1. partiellen Ableitungen  $\neq 0$  sind.

Da in den vorliegenden Daten für jeden Probanden und jedes Schlafstadium alle vier Zustände  $A, B, C, D$  mindestens einmal auftreten, stellt die getroffene Voraussetzung keine Einschränkung dar.

**2. Schritt:** Nun werden Formeln für die partiellen Ableitungen

$$\frac{\partial \hat{\pi}^{(u)}(T)}{\partial n_{rs}} \text{ mit } (r, s) \neq (a, a)$$

bestimmt, die für die Anwendung der Delta-Methode benötigt werden.

Es werden drei Fälle unterschieden:

- (a)  $s \neq r = a$ ,
- (b)  $s = r \neq a$ ,
- (c)  $s \neq r \neq a$ .

Diese Unterscheidung rührt daher, dass für  $s \neq r = a$  der Nenner von  $\hat{\pi}^{(u)}$  nicht von  $n_{rs}$  abhängt und für  $s = r \neq a$  der Zähler nicht von  $n_{rr}$  abhängt.

**Fall (a)**  $s \neq r = a$ : Der Nenner wird für die partielle Ableitung nach  $n_{rs} = n_{as}$  als Konstante behandelt. Im Zähler kommt  $n_{as}$  genau einmal als Summand vor. Also ist

$$\frac{\partial \hat{\pi}^{(u)}(T)}{\partial n_{rs}} = \frac{\partial \hat{\pi}^{(u)}(T)}{\partial n_{as}} = \frac{1/n}{1 - \sum_k \hat{p}_k^2} \quad \text{für } s \neq r = a.$$

**Fall (b)**  $s = r \neq a$ : In diesem Fall kann der Zähler für die partielle Ableitung nach  $n_{rr}$  als Konstante angesehen werden. Im Nenner tritt  $n_{rr}$  für festes  $r$  in der vorderen quadrierten Summe für  $k = l = r$  auf und in der hinteren quadrierten Summe für  $t = u = r$ .

Für die partielle Ableitung des Nenners von  $\hat{\pi}^{(u)}(T)$  nach  $n_{rr}$  gilt somit

$$\begin{aligned} & \frac{\partial}{\partial n_{rr}} \left( 1 - \frac{1}{n^2} \sum_{k=1}^{a-1} \left( \sum_{l=1}^a n_{kl} \right)^2 - \frac{1}{n^2} \left( n - \sum_{t=1}^{a-1} \sum_{u=1}^a n_{tu} \right)^2 \right) \\ &= -\frac{2}{n^2} \left( \sum_{l=1}^a n_{rl} \right) - \frac{2}{n^2} \left( n - \sum_{t=1}^{a-1} \sum_{u=1}^a n_{tu} \right) (-1) \\ &= -\frac{2}{n} \left( \hat{p}_r - 1 + \sum_{t=1}^{a-1} \hat{p}_t \right) \\ &= -\frac{2}{n} (\hat{p}_r - \hat{p}_a). \end{aligned}$$

Daraus resultiert

$$\frac{\partial \hat{\pi}^{(u)}(T)}{\partial n_{rs}} = \frac{\partial \hat{\pi}^{(u)}(T)}{\partial n_{rr}} = \frac{2}{n} \cdot \frac{\sum_{i \neq j} n_{ij}/n}{\left( 1 - \sum_k \hat{p}_k^2 \right)^2} (\hat{p}_r - \hat{p}_a) \quad \text{für } s = r \neq a.$$

**Fall (c)  $s \neq r \neq a$ :** Für den letzten Fall weist  $\hat{\pi}^{(u)}(T)$  den Term  $n_{rs}$  sowohl im Zähler als auch im Nenner auf. Durch die Produktformel lassen sich Zähler und Kehrwert des Nenners jeweils einzeln ableiten, mit dem jeweils anderen Term multiplizieren und die Produkte addieren.

Daher folgt mit den Argumenten der Fälle (a) und (b) direkt

$$\frac{\partial \hat{\pi}^{(u)}(T)}{\partial n_{rs}} = \frac{1/n}{1 - \sum_k \hat{p}_k^2} + \frac{2}{n} \cdot \frac{\sum_{i \neq j} n_{ij}/n}{\left(1 - \sum_k \hat{p}_k^2\right)^2} (\hat{p}_r - \hat{p}_a) \quad \text{für } s \neq r \neq a.$$

**3. Schritt:** In diesem Schritt wird in die partiellen Ableitungen der Vektor

$$\vartheta := \mathbb{E}[T]$$

des Erwartungswertes von  $T$  eingesetzt, wobei

$$\mathbb{E}[n_{rs}] = np_r p_{rs} = np_r (\pi p_s + \delta_{r,s}(1 - \pi)) \quad \forall r, s \in S$$

nach Korollar 3.8 gilt.

In allen drei oben beschriebenen Fällen tritt  $\hat{p}_r$  für ein  $r \in S$  auf. Beim Einsetzen von  $\vartheta$  in die partiellen Ableitungen erhält man statt  $\hat{p}_r$  den Ausdruck

$$\hat{p}_k(\vartheta) = \frac{1}{n} \sum_{l=1}^a np_k p_{kl} = p_k \sum_{l=1}^a p_{kl} = p_k \quad \text{für } k \neq a$$

und ebenso

$$\hat{p}_a(\vartheta) = \frac{1}{n} \left( n - \sum_{t=1}^{a-1} \sum_{u=1}^a np_t p_{tu} \right) = 1 - \sum_{t=1}^{a-1} p_t = p_a.$$

Für  $\sum_{i \neq j} n_{ij}/n$  erhält man durch Einsetzen der Erwartungswerte

$$\left( \sum_{i \neq j} n_{ij}/n \right) (\vartheta) = \sum_{i \neq j} np_i \pi p_j / n = \pi \left( 1 - \sum_k p_k^2 \right).$$

Mit den letzten Ergebnissen lässt sich  $\frac{\partial \hat{\pi}^{(u)}(T)}{\partial n_{rs}} \Big|_{T=\vartheta}$  für alle  $(r, s) \neq (a, a)$  herleiten:



$$\frac{\partial \hat{\pi}^{(u)}(T)}{\partial n_{rs}} \Big|_{T=\vartheta} = \begin{cases} \frac{1}{n \left(1 - \sum_k p_k^2\right)} & \text{für } s \neq r = a, \\ \frac{2\pi(p_r - p_a)}{n \left(1 - \sum_k p_k^2\right)} & \text{für } s = r \neq a, \\ \frac{1 + 2\pi(p_r - p_a)}{n \left(1 - \sum_k p_k^2\right)} & \text{für } s \neq r \neq a. \end{cases}$$

Die drei Formeln für die verschiedenen Fälle lassen sich zusammenfassen zu einer einzigen Formel durch

$$\frac{\partial \hat{\pi}^{(u)}(T)}{\partial n_{rs}} \Big|_{T=\vartheta} = \frac{(1 - \delta_{r,s}) + 2\pi(p_r - p_a)}{n \left(1 - \sum_k p_k^2\right)} \quad \forall (r, s) \neq (a, a). \quad (5.16)$$

Damit ist die Formel für die partiellen Ableitungen von  $\hat{\pi}^{(u)}(T)$  an der Stelle  $T = \vartheta$  bestimmt.

**4. Schritt:** Nach der Delta-Methode (siehe Anhang B) lässt sich  $\text{Var}(\hat{\pi}^{(u)})$  auf folgende Weise approximieren:

$$\text{Var}(\hat{\pi}^{(u)}) \approx \sum_{\substack{r,s=1 \\ (r,s) \neq (a,a)}}^a \sum_{\substack{t,u=1 \\ (t,u) \neq (a,a)}}^a \frac{\partial \hat{\pi}^{(u)}(T)}{\partial n_{rs}} \Big|_{T=\vartheta} \cdot \frac{\partial \hat{\pi}^{(u)}(T)}{\partial n_{tu}} \Big|_{T=\vartheta} \cdot \text{Cov}(n_{rs}, n_{tu}). \quad (5.17)$$

Durch das Einsetzen der Formel für die partiellen Ableitungen aus Gleichung (5.16) in die obige Approximation (5.17) erhält man

$$\text{Var}(\hat{\pi}^{(u)}) \approx \sum_{\substack{r,s=1 \\ (r,s) \neq (a,a)}}^a \sum_{\substack{t,u=1 \\ (t,u) \neq (a,a)}}^a \frac{(1 - \delta_{r,s})(1 - \delta_{t,u})}{n^2 \left(1 - \sum_k p_k^2\right)^2} \text{Cov}(n_{rs}, n_{tu}) \quad (5.18)$$

$$+ \sum_{\substack{r,s=1 \\ (r,s) \neq (a,a)}}^a \sum_{\substack{t,u=1 \\ (t,u) \neq (a,a)}}^a \frac{2\pi(p_t - p_a)(1 - \delta_{r,s})}{n^2 \left(1 - \sum_k p_k^2\right)^2} \text{Cov}(n_{rs}, n_{tu}) \quad (5.19)$$

$$+ \sum_{\substack{r,s=1 \\ (r,s) \neq (a,a)}}^a \sum_{\substack{t,u=1 \\ (t,u) \neq (a,a)}}^a \frac{2\pi(p_r - p_a)(1 - \delta_{t,u})}{n^2 \left(1 - \sum_k p_k^2\right)^2} \text{Cov}(n_{rs}, n_{tu}) \quad (5.20)$$

$$+ \sum_{\substack{r,s=1 \\ (r,s) \neq (a,a)}}^a \sum_{\substack{t,u=1 \\ (t,u) \neq (a,a)}}^a \frac{2\pi(p_r - p_a)2\pi(p_t - p_a)}{n^2 \left(1 - \sum_k p_k^2\right)^2} \text{Cov}(n_{rs}, n_{tu}). \quad (5.21)$$

Die Zeile (5.18), die Summe der Zeilen (5.19) und (5.20) sowie die Zeile (5.21) werden im Folgenden nacheinander weiter umgeformt. Zur besseren Übersicht wird der Nenner  $n^2 (1 - \sum_k p_k^2)^2$  dabei nicht mitgeschrieben.

Der Ausdruck  $(1 - \delta_{t,u})$  aus der ersten Zeile ist genau dann  $\neq 0$ , wenn  $t = u$  ist, und sonst nimmt er den Wert 1 an. Somit folgt für den Term aus der ersten Zeile (ohne Nenner)

$$\begin{aligned}
& \sum_{\substack{r,s=1 \\ (r,s) \neq (a,a)}}^a \sum_{\substack{t,u=1 \\ (t,u) \neq (a,a)}}^a (1 - \delta_{r,s}) (1 - \delta_{t,u}) \text{Cov}(n_{rs}, n_{tu}) \\
&= \sum_{\substack{r,s=1 \\ (r,s) \neq (a,a)}}^a (1 - \delta_{r,s}) \text{Cov} \left( n_{rs}, \sum_{\substack{t,u=1 \\ t \neq u}}^a n_{tu} \right) = \text{Cov} \left( \sum_{\substack{r,s=1 \\ r \neq s}}^a n_{rs}, \sum_{\substack{t,u=1 \\ t \neq u}}^a n_{tu} \right) \\
&= \text{Cov} \left( n - \sum_{r=1}^a n_{rr}, n - \sum_{t=1}^a n_{tt} \right) = \text{Cov} \left( \sum_{r=1}^a n_{rr}, \sum_{t=1}^a n_{tt} \right) \\
&= \sum_{r=1}^a \sum_{t=1}^a \text{Cov}(n_{rr}, n_{tt}).
\end{aligned}$$

Für die Summe der Ausdrücke aus den Zeilen (5.19) und (5.20) gilt (ohne Nenner)

$$\begin{aligned}
& 4\pi \sum_{\substack{r,s=1 \\ (r,s) \neq (a,a)}}^a \sum_{\substack{t,u=1 \\ (t,u) \neq (a,a)}}^a (p_r - p_a) (1 - \delta_{t,u}) \text{Cov}(n_{rs}, n_{tu}) \\
&= 4\pi \sum_{r=1}^{a-1} (p_r - p_a) \text{Cov} \left( \sum_{s=1}^a n_{rs}, \sum_{\substack{t,u=1 \\ (t,u) \neq (a,a)}}^a n_{tu} \right) \\
&= 4\pi \sum_{r=1}^{a-1} (p_r - p_a) \text{Cov} \left( n_r, n - \sum_{t=1}^a n_{tt} \right) \\
&= 4\pi \sum_{r=1}^{a-1} p_r \text{Cov} \left( n_r, n - \sum_{t=1}^a n_{tt} \right) - 4\pi p_a \text{Cov} \left( \sum_{r=1}^{a-1} n_r, n - \sum_{t=1}^a n_{tt} \right) \\
&= 4\pi \sum_{r=1}^{a-1} p_r \text{Cov} \left( n_r, n - \sum_{t=1}^a n_{tt} \right) - 4\pi p_a \text{Cov} \left( n - n_a, n - \sum_{t=1}^a n_{tt} \right) \\
&= 4\pi \sum_{r=1}^a p_r \text{Cov} \left( n_r, n - \sum_{t=1}^a n_{tt} \right) \\
&= -4\pi \sum_{r=1}^a \sum_{t=1}^a p_r \text{Cov}(n_r, n_{tt}).
\end{aligned}$$

Für den Ausdruck aus der Zeile (5.21) folgt entsprechend (ohne Nenner)

$$\begin{aligned}
& 4\pi^2 \sum_{\substack{r,s=1 \\ (r,s) \neq (a,a)}}^a \sum_{\substack{t,u=1 \\ (t,u) \neq (a,a)}}^a (p_r - p_a)(p_t - p_a) \text{Cov}(n_{rs}, n_{tu}) \\
&= 4\pi^2 \sum_{r=1}^a \sum_{\substack{t,u=1 \\ (t,u) \neq (a,a)}}^a p_r (p_t - p_a) \text{Cov}(n_r, n_{tu}) \\
&= 4\pi^2 \sum_{r=1}^a \sum_{t=1}^a p_r p_t \text{Cov}(n_r, n_t).
\end{aligned}$$

Damit folgt direkt die Aussage von Lemma 5.13:

$$\text{Var}(\hat{\pi}^{(u)}) \approx \sum_{r=1}^a \sum_{t=1}^a \frac{\text{Cov}(n_{rr}, n_{tt}) - 4\pi p_r \text{Cov}(n_r, n_{tt}) + 4\pi^2 p_r p_t \text{Cov}(n_r, n_t)}{n^2 \left(1 - \sum_k p_k^2\right)^2}.$$

□

**Bemerkung 5.14.** In der Approximation (5.15) von  $\text{Var}(\hat{\pi}^{(u)})$  aus Lemma 5.13 ist der vordere Summand

$$\sum_{r=1}^a \sum_{t=1}^a \frac{\text{Cov}(n_{rr}, n_{tt})}{n^2 \left(1 - \sum_k p_k^2\right)^2}$$

gerade die Varianz des Schätzers

$$\hat{\pi}^{(1)} = \frac{n - \sum_r n_{rr}}{n(1 - \sum_k p_k^2)}$$

aus Kapitel 4.2, dessen Unterschied zum Schätzer  $\hat{\pi}^{(u)}$  darin besteht, dass bei  $\hat{\pi}^{(1)}$  im Nenner für alle  $k \in S$  der wahre Wert  $p_k$  und nicht der Schätzer  $\hat{p}_k$  steht. Da der Nenner von  $\hat{\pi}^{(1)}$  konstant ist, lässt sich die Varianz von  $\hat{\pi}^{(1)}$  berechnen durch

$$\text{Var}(\hat{\pi}^{(1)}) = \frac{\text{Var}(n - \sum_r n_{rr})}{n^2 (1 - \sum_k p_k^2)^2} = \frac{\text{Cov}(\sum_r n_{rr}, \sum_t n_{tt})}{n^2 (1 - \sum_k p_k^2)^2} = \sum_r \sum_t \frac{\text{Cov}(n_{rr}, n_{tt})}{n^2 (1 - \sum_k p_k^2)^2}.$$

Somit geben in der Approximation (5.15) von  $\text{Var}(\hat{\pi}^{(u)})$  aus Lemma 5.13 die beiden hinteren Summanden

$$\sum_{r=1}^a \sum_{t=1}^a \frac{-4\pi p_r \text{Cov}(n_r, n_{tt}) + 4\pi^2 p_r p_t \text{Cov}(n_r, n_t)}{n^2 \left(1 - \sum_k p_k^2\right)^2}$$

die Änderung der Variabilität an, die dadurch entsteht, dass beim Schätzer  $\hat{\pi}^{(u)}$  im Nenner durch  $\hat{p}_k$  geschätzte Werte für  $p_k$  stehen.

Im Folgenden wird der Beweis zu Proposition 5.12 geführt, indem in die Approximation (5.15) von  $\text{Var}(\hat{\pi}^{(u)})$  aus Lemma 5.13 die Formeln für die Kovarianzen aus den Kapiteln 5.2.1 und 5.2.2 eingesetzt werden.

*Beweis (Proposition 5.12).* Nach Lemma 5.13 gilt

$$\text{Var}(\hat{\pi}^{(u)}) \approx \sum_{r=1}^a \sum_{t=1}^a \frac{\overbrace{\text{Cov}(n_{rr}, n_{tt})}^{(I)} - 4\pi p_r \overbrace{\text{Cov}(n_r, n_{tt})}^{(II)} + 4\pi^2 p_r p_t \overbrace{\text{Cov}(n_r, n_t)}^{(III)}}{n^2 \left(1 - \sum_k p_k^2\right)^2}. \quad (5.22)$$

Es werden nun die Formeln für die Kovarianzen aus den Kapiteln 5.2.1 (Gleichung (5.13) in Beispiel 5.5) und 5.2.2 (Proposition 5.6) in die obige Approximation eingesetzt. Diese Formeln lauten im Kontext des SMI-Modells

$$\text{Cov}(n_r, n_t) = n \left( \frac{2 - \pi}{\pi} p_r (\delta_{r,t} - p_t) \right) + \mathcal{O}(1) \quad (5.23)$$

und

$$\begin{aligned} & \text{Cov}(n_{rs}, n_{tu}) \\ &= n \left( \delta_{rs,tu} p_r p_{rs} - \frac{2 + \pi}{\pi} p_r p_{rs} p_t p_{tu} + \frac{1}{\pi} (\delta_{s,t} p_r p_{rs} p_{su} + \delta_{u,r} p_t p_{tu} p_{us}) \right) + \mathcal{O}(1). \end{aligned} \quad (5.24)$$

Für alle  $r, t \in S$  folgt aus der letzten Gleichung direkt

$$\begin{aligned} & \text{Cov}(n_r, n_{tt}) \\ &= \sum_{s=1}^a \text{Cov}(n_{rs}, n_{tt}) \\ &= \sum_{s=1}^a n \left( \delta_{rs,tt} p_r p_{rs} - \frac{2 + \pi}{\pi} p_r p_{rs} p_t p_{tt} + \frac{1}{\pi} (\delta_{s,t} p_r p_{rs} p_{st} + \delta_{t,r} p_t p_{tt} p_{ts}) \right) + \mathcal{O}(1) \\ &= n \left( \delta_{r,t} p_r p_{rr} - \frac{2 + \pi}{\pi} p_r p_t p_{tt} + \frac{1}{\pi} (p_r p_{rt} p_{tt} + \delta_{t,r} p_r p_{rr}) \right) + \mathcal{O}(1) \\ &= n \left( \delta_{r,t} \frac{1 + \pi}{\pi} p_r p_{rr} - \frac{2 + \pi}{\pi} p_r p_t p_{tt} + \frac{1}{\pi} p_r (\pi p_t + \delta_{r,t} (1 - \pi)) p_{tt} \right) + \mathcal{O}(1) \\ &= n \left( \delta_{r,t} \frac{2}{\pi} p_r p_{rr} - \frac{2}{\pi} p_r p_t p_{tt} \right) + \mathcal{O}(1) \\ &= \frac{2n}{\pi} p_r p_{tt} (\delta_{r,t} - p_t) + \mathcal{O}(1). \end{aligned} \quad (5.25)$$

Für den Term unter (II) in der Approximation (5.22) ergibt sich aus Gleichung (5.25) somit

$$\begin{aligned}
& -4\pi \sum_r \sum_t p_r \text{Cov}(n_r, n_{tt}) \\
&= -8n \sum_r \sum_t p_r^2 p_{tt} (\delta_{r,t} - p_t) + \mathcal{O}(1) \\
&= -8n \sum_r p_r^2 \left( p_{rr} - \sum_t p_t p_{tt} \right) + \mathcal{O}(1) \\
&= -8n \sum_r p_r^2 \left( \pi p_r + 1 - \pi - \sum_t p_t (\pi p_t + 1 - \pi) \right) + \mathcal{O}(1) \\
&= -8n \sum_r p_r^2 \left( \pi p_r - \left( \pi \sum_t p_t^2 \right) + (1 - \pi) - (1 - \pi) \sum_t p_t \right) + \mathcal{O}(1) \\
&= -8\pi n \sum_r p_r^2 \left( p_r - \sum_t p_t^2 \right) + \mathcal{O}(1). \tag{5.26}
\end{aligned}$$

Der Term unter (III) in der Approximation (5.22) lässt sich mit Gleichung (5.23) schreiben als

$$\begin{aligned}
4\pi^2 \sum_r \sum_t p_r p_t \text{Cov}(n_r, n_t) &= 4(2 - \pi)\pi n \sum_r \sum_t p_r^2 p_t (\delta_{r,t} - p_t) + \mathcal{O}(1) \\
&= 4(2 - \pi)\pi n \sum_r p_r^2 \left( p_r - \sum_t p_t^2 \right) + \mathcal{O}(1). \tag{5.27}
\end{aligned}$$

In der Summe ergeben die Terme unter (II) und (III) in der Approximation (5.22) gemäß den Gleichungen (5.26) und (5.27) also

$$\begin{aligned}
& -4\pi \sum_r \sum_t p_r \text{Cov}(n_r, n_{tt}) + 4\pi^2 \sum_r \sum_t p_r p_t \text{Cov}(n_r, n_t) \\
&= -8\pi n \sum_r p_r^2 \left( p_r - \sum_t p_t^2 \right) + 4(2 - \pi)\pi n \sum_r p_r^2 \left( p_r - \sum_t p_t^2 \right) + \mathcal{O}(1) \\
&= -4\pi^2 n \sum_r p_r^2 \left( p_r - \sum_t p_t^2 \right) + \mathcal{O}(1). \tag{5.28}
\end{aligned}$$

Es fehlt noch die Rechnung für den Term unter (I) in der Approximation (5.22). Hier lässt sich Gleichung (5.24) einsetzen und man erhält

$$\begin{aligned}
& \sum_r \sum_t \text{Cov}(n_{rr}, n_{tt}) \\
&= n \sum_r \sum_t \left( \delta_{rr,tt} p_r p_{rr} - \frac{2+\pi}{\pi} p_r p_{rr} p_t p_{tt} + \frac{1}{\pi} (\delta_{r,t} p_r p_{rr}^2 + \delta_{r,t} p_r p_{rr}^2) \right) + \mathcal{O}(1) \\
&= n \sum_r \sum_t p_r p_{rr} \left( \delta_{r,t} \left( 1 + \frac{2}{\pi} p_{rr} \right) - \frac{2+\pi}{\pi} p_t p_{tt} \right) + \mathcal{O}(1) \\
&= n \sum_r p_r p_{rr} \left( \left( 1 + 2 \frac{\pi p_r + 1 - \pi}{\pi} \right) - \frac{2+\pi}{\pi} \sum_t p_t (\pi p_t + 1 - \pi) \right) + \mathcal{O}(1) \\
&= n \sum_r p_r p_{rr} \left( 2p_r - 1 + \frac{2}{\pi} - \frac{(2+\pi)(1-\pi)}{\pi} - (2+\pi) \sum_t p_t^2 \right) + \mathcal{O}(1) \\
&= n \sum_r p_r (\pi p_r + 1 - \pi) \left( 2p_r + \pi - (2+\pi) \sum_t p_t^2 \right) + \mathcal{O}(1). \tag{5.29}
\end{aligned}$$

Die letzte Zeile lässt sich durch Ausmultiplizieren vereinfachen zu

$$\begin{aligned}
& n \sum_r p_r (\pi p_r + 1 - \pi) \left( 2p_r + \pi - (2+\pi) \sum_t p_t^2 \right) + \mathcal{O}(1) \\
&= n \left( \sum_r p_r^2 \left( 2p_r \pi + \pi^2 - (2+\pi) \pi \sum_t p_t^2 \right) \right. \\
&\quad \left. + \sum_r p_r (1-\pi) \left( 2p_r + \pi - (2+\pi) \sum_t p_t^2 \right) \right) + \mathcal{O}(1) \\
&= n \left( \sum_r p_r^2 \left( 2p_r \pi + \pi^2 - (2+\pi) \pi \sum_t p_t^2 + 2(1-\pi) \right) \right. \\
&\quad \left. + (1-\pi) \left( \pi - (2+\pi) \sum_r p_r^2 \right) \right) + \mathcal{O}(1) \\
&= n \left( (1-\pi) \pi + \sum_r p_r^2 \left( 2\pi^2 - \pi + 2p_r \pi - (2+\pi) \pi \sum_t p_t^2 \right) \right) + \mathcal{O}(1).
\end{aligned}$$

Aus Gleichung (5.29) wird dann

$$\begin{aligned}
& \sum_r \sum_t \text{Cov}(n_{rr}, n_{tt}) \\
&= n \left( (1-\pi) \pi + \sum_r p_r^2 \left( 2\pi^2 - \pi + 2p_r \pi - (2+\pi) \pi \sum_t p_t^2 \right) \right) + \mathcal{O}(1). \tag{5.30}
\end{aligned}$$

Setzt man die Gleichungen (5.30) und (5.28) in die Approximation (5.15) ein, folgt

$$\begin{aligned}
\text{Var}(\hat{\pi}^{(u)}) &\approx \sum_{r=1}^a \sum_{t=1}^a \frac{\text{Cov}(n_{rr}, n_{tt}) - 4\pi p_r \text{Cov}(n_r, n_{tt}) + 4\pi^2 p_r p_t \text{Cov}(n_r, n_t)}{n^2 \left(1 - \sum_k p_k^2\right)^2} \\
&= \frac{n \left( (1-\pi)\pi + \sum_r p_r^2 \left( 2\pi^2 - \pi + 2p_r\pi - (2+\pi)\pi \sum_t p_t^2 \right) \right)}{n^2 \left(1 - \sum_k p_k^2\right)^2} \\
&\quad - \frac{4\pi^2 n \sum_r p_r^2 \left( p_r - \sum_t p_t^2 \right)}{n^2 \left(1 - \sum_k p_k^2\right)^2} + \mathcal{O}(n^{-2}) \\
&= \frac{(1-\pi)\pi + \sum_r p_r^2 \left( 2\pi^2 - \pi + (2\pi - 4\pi^2)p_r + (3\pi^2 - 2\pi) \sum_t p_t^2 \right)}{n \left(1 - \sum_k p_k^2\right)^2} + \mathcal{O}(n^{-2}).
\end{aligned}$$

Damit ist Proposition 5.12 bewiesen. □

### 5.3.2 Einfluss der Parameter des SMI-Modells auf die Varianz von $\hat{\pi}^{(u)}$

In diesem Kapitel wird der Einfluss der Parameter  $p_1, \dots, p_a$  und  $\pi$  des SMI-Modells auf die Varianz von  $\hat{\pi}^{(u)}$  gemäß der Formel

$$\text{Var}(\hat{\pi}^{(u)}) \approx \frac{-\pi^2 + \pi + \sum_r p_r^2 \left( 2\pi^2 - \pi + (-4\pi^2 + 2\pi)p_r + (3\pi^2 - 2\pi) \sum_t p_t^2 \right)}{n \left(1 - \sum_k p_k^2\right)^2} + \mathcal{O}(n^{-2})$$

von Proposition 5.12 untersucht, wobei die Varianz von  $\hat{\pi}^{(u)}$  mit der rechten Seite gleichgesetzt wird und der asymptotisch vernachlässigbare Term  $\mathcal{O}(n^{-2})$  nicht berücksichtigt wird.

Es wird also

$$f(\pi, p_1, \dots, p_a) := \frac{-\pi^2 + \pi + \sum_r p_r^2 \left( 2\pi^2 - \pi + (-4\pi^2 + 2\pi)p_r + (3\pi^2 - 2\pi) \sum_t p_t^2 \right)}{n \left(1 - \sum_k p_k^2\right)^2}$$

als Näherung für  $\text{Var}(\hat{\pi}^{(u)})$  betrachtet.

Für  $a > 1$  lässt sich o.B.d.A.  $p_a$  ersetzen durch  $1 - \sum_{s=1}^{a-1} p_s$ , sodass  $f$  nicht von  $p_a$  abhängt. Daher wird von nun an  $f(\pi, p_1, \dots, p_{a-1})$  anstelle von  $f(\pi, p_1, \dots, p_a)$  geschrieben. Sollte in Ausnahmefällen doch der Parameter  $p_a$  verwendet werden, ist dieser je nach Kontext als  $p_a$  oder  $1 - \sum_{s=1}^{a-1} p_s$  zu lesen.

Werden zudem ein oder mehrere Parameter fest gewählt, werden im weiteren Verlauf des Kapitels auch diese in der Regel aus dem Parameter-Vektor  $(\pi, p_1, \dots, p_{a-1})$  gestrichen, um zu verdeutlichen, von welchen variablen Größen  $f$  in diesem Fall abhängt.

Begonnen wird bei der Untersuchung mit dem Fall  $a = 2$ , bei dem nur die Parameter  $\pi$  und o.B.d.A.  $p_1$  frei wählbar sind, da  $p_2 = 1 - p_1$  ist. Anschließend wird der Fall  $a = 3$  behandelt, gefolgt vom Fall  $a = 4$ , der für die in dieser Arbeit untersuchten Daten zutrifft ( $S = \{A, B, C, D\}$ , also  $|S| = 4$ ).

In allen Fällen wird die Varianz von  $\hat{\pi}^{(u)}$  mithilfe von  $f$  insbesondere auf die folgenden Punkte hin untersucht:

- (a) Form des Graphen von  $f$  für beliebige  $p_1, \dots, p_a$  und  $\pi$ ,
- (b) Werte für  $p_1, \dots, p_a$ , für die  $f(p_1, \dots, p_a)$  minimal bzw. maximal wird für festes  $\pi$ ,
- (c) Werte für  $\pi$ , für die  $f(\pi)$  minimal bzw. maximal wird für feste  $p_1, \dots, p_a$ .

Zum Schluss dieses Kapitels werden die für  $a \in \{2, 3, 4\}$  und teilweise für beliebige  $a > 1$  gewonnenen Erkenntnisse zusammengefasst.

In allen folgenden Abbildungen von Graphen von  $f$  wird  $n = 1$  gesetzt. Um die in den Abbildungen abzulesenden Funktionswerte von  $f$  auf beliebige  $n \in \mathbb{N}$  zu übertragen, müssen diese Funktionswerte mit  $1/n$  multipliziert werden.

Da der Nenner

$$n \left( 1 - \sum_k p_k^2 \right)^2$$

von  $f$  den Wert 0 annimmt, wenn  $p_k = 1$  ist für ein  $k \in \{1, \dots, a\}$ , wird in diesem Kapitel stets

$$p_1, \dots, p_a < 1$$

angenommen, auch für den Ausdruck  $p_k \nearrow 1$ .

### Der Fall $a = 2$

Zunächst wird  $f$  in Abhängigkeit von  $p_1$  für festes  $\pi$  betrachtet. Abbildung 5.2 zeigt Graphen von  $f(p_1)$  für verschiedene feste Werte von  $\pi$ .

Die Graphen sind symmetrisch mit Spiegelachse bei  $p_1 = 0.5$ , da  $f$  für  $p_1$  und  $1 - p_1$  stets dieselben Werte annimmt bei festem  $\pi$ .

Für  $\pi = 0$  ist  $f(p_1) = 0$ , da in jedem Summanden des Zählers von  $f$  der Faktor  $\pi$  steckt und  $p_1, p_2 < 1$  sind. Diese Aussage besitzt Gültigkeit für alle  $a > 1$ .



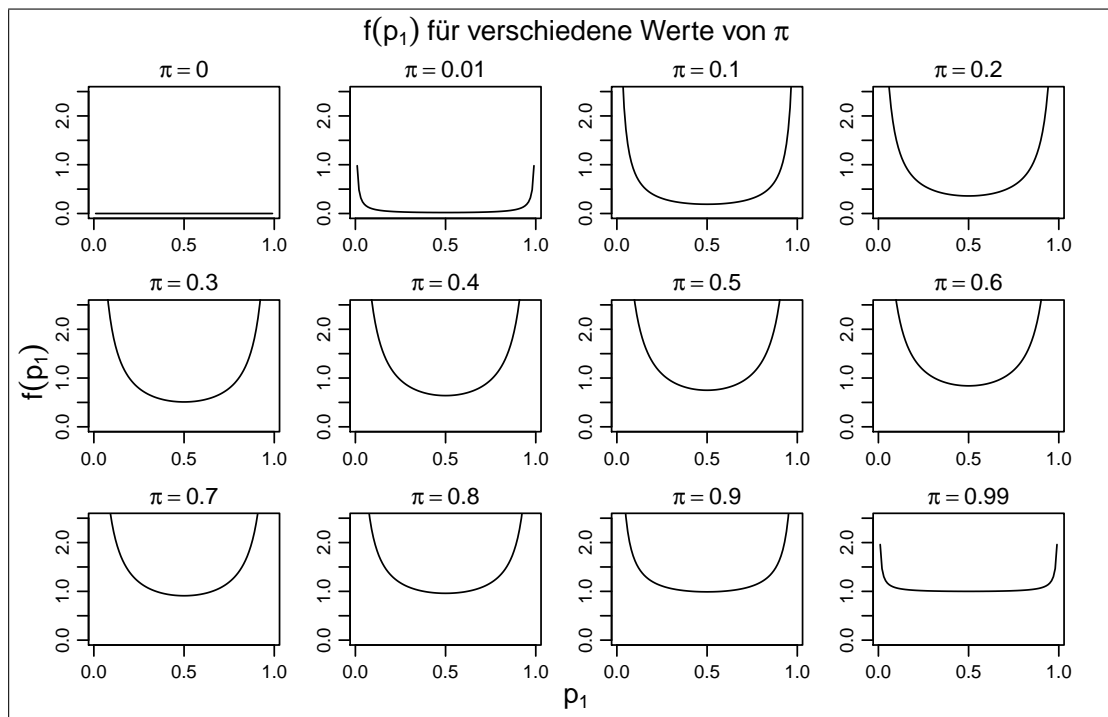


Abbildung 5.2: Graphen von  $f(p_1)$  auf  $[0.01, 0.99]$  für verschiedene feste Werte von  $\pi$ .

Die Graphen für festes  $\pi > 0$  haben die Form eines breiten Tals mit steilen Seitenwänden. An den Rändern des Definitionsbereichs von  $f$  gilt dann

$$\lim_{p_1 \searrow 0} f(p_1) = \lim_{p_1 \nearrow 1} f(p_1) = \infty.$$

Je näher sich  $\pi$  am Bereich von ca. 0.4 bis 0.6 befindet, desto schmaler wird das Tal.

Aus Symmetriegründen nimmt  $f(p_1)$  für festes  $\pi > 0$  ein Extremum an, wenn  $p_1 = 0.5$  ist. Genauer sind die partiellen Ableitungen von  $f(p_1, p_2)$  nach  $p_1$  und  $p_2$  im Lagrange-Ansatz (siehe z. B. Forster, 2006) gleich und ihre Formeln unterscheiden sich nur durch Vertauschen der Bezeichnungen  $p_1$  und  $p_2$ . Somit ist eine Lösung des Gleichungssystems für  $p_1$  und  $p_2$  beim Lagrange-Ansatz gerade gegeben, wenn  $p_1 = p_2$  gilt.

**Bemerkung 5.15.** Sei  $a > 1$  beliebig. Dann folgt wiederum aus Symmetriegründen (bzw. über den Lagrange-Ansatz), dass  $f(p_1, \dots, p_{a-1})$  ein Extremum annimmt an der Stelle

$$p_1 = \dots = p_{a-1} (= p_a) = \frac{1}{a}.$$

Die Graphen aus Abbildung 5.2 legen nahe, dass es sich für  $a = 2$  bei dem Extremum in  $p_1 = 0.5$  um ein globales Minimum handelt.

Der Funktionswert von  $f(\pi, p_1)$  an der Stelle  $p_1 = 0.5$  lässt sich berechnen durch

$$f(\pi, 0.5) = \frac{\pi(2 - \pi)}{n}$$

und ist umso größer, je größer  $\pi$  ist. Dass  $f(\pi, 0.5)$  monoton wachsend ist in  $\pi$ , kann man auch in Abbildung 5.2 beobachten.

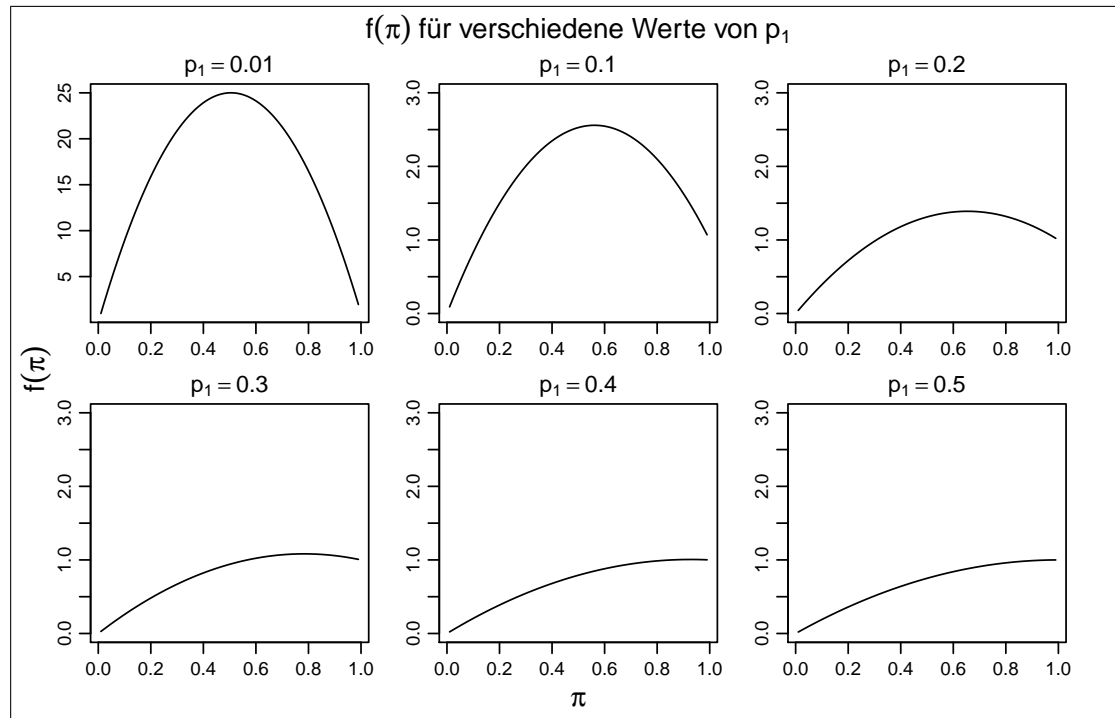


Abbildung 5.3: Graphen von  $f(\pi)$  auf  $[0.01, 0.99]$  für verschiedene feste Werte von  $p_1$ . Zu beachten ist, dass für  $p_1 = 0.01$  die y-Achse den Bereich von 0 bis 25 abdeckt, während die y-Achsen für die anderen Werte von  $p_1$  jeweils den Bereich von 0 bis 3 abdecken. Für  $p_1 > 0.5$  ergeben sich die gleichen Graphen wie für  $1 - p_1$ .

Hält man  $p_1$  fest und lässt  $\pi$  variabel, so beobachtet man parabel-förmige Graphen für  $f(\pi)$ , wie in Abbildung 5.3 zu sehen ist für verschiedene feste Werte von  $p_1$ .

Wegen  $f(\pi, p_1) = f(0, p_1) = 0$  für  $\pi = 0$  startet der Graph von  $f(\pi)$  in 0 für alle festen  $0 < p_1 < 1$ . Je näher der feste Wert von  $p_1$  bei 0.5 liegt, desto flacher wird die Parabel und desto weiter rechts liegt der Scheitelpunkt. Für  $p_1 \searrow 0$  bzw.  $p_1 \nearrow 1$  befindet sich der Scheitelpunkt bei  $\pi \searrow 0.5$ , für  $p_1 = 0.5$  befindet er sich bei  $\pi = 1$ .

**Bemerkung 5.16.** Das  $\pi$ , für das  $f(\pi)$  maximal wird bei festem  $p_1$ , lässt sich formal bestimmen durch partielles Ableiten von  $f(\pi, p_1)$  nach  $\pi$  und Nullsetzen der Ableitung. Für allgemeines  $a \in \mathbb{N}$  gilt

$$\frac{\partial f(\pi, p_1, \dots, p_{a-1})}{\partial \pi} = \frac{-2\pi + 1 + \sum_r p_r^2 (4\pi - 1 + (-8\pi + 2)p_r + (6\pi - 2) \sum_t p_t^2)}{n(1 - \sum_k p_k^2)^2}.$$

Unter der Voraussetzung  $p_1, \dots, p_a < 1$  wird der Ausdruck 0 für

$$\begin{aligned} \pi_{\max}(p_1, \dots, p_{a-1}) = \pi_{\max} &= \frac{-1 + \sum_r p_r^2 (1 - 2p_r + 2 \sum_t p_t^2)}{-2 + \sum_r p_r^2 (4 - 8p_r + 6 \sum_t p_t^2)} \\ &= \frac{1}{2} \left( 1 + \frac{\sum_r p_r^2 (-1 + 2p_r - \sum_t p_t^2)}{-1 + \sum_r p_r^2 (2 - 4p_r + 3 \sum_t p_t^2)} \right), \end{aligned}$$

wobei  $r$  und  $t$  für alle Summen  $1, \dots, a$  durchlaufen.

Für den Fall  $a = 2$  ist  $\pi_{\max}(p_1)$  in Abhängigkeit von  $p_1$  als grüne Kurve eingetragen in Abbildung 5.4, in der zudem die Funktion  $f(\pi, p_1)$  als Konturplot dargestellt wird. In dieser Abbildung sieht man die Gültigkeit von

$$0.5 < \pi_{\max}(p_1) \leq 1 \quad \forall p_1 \in (0, 1).$$

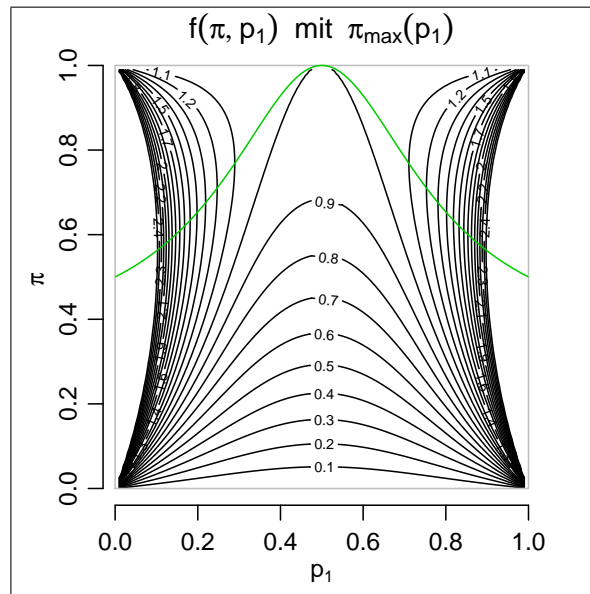


Abbildung 5.4: Konturplot von  $f(\pi, p_1)$  mit eingezeichnetem Graphen von  $\pi_{\max}(p_1)$  (grün). Konturlinien sind für den Bereich von 0.1 bis 2.5 in 0.1-Schritten eingetragen. Die grauen Linien an den Seiten begrenzen den Definitionsbereich von  $f$ .

**Der Fall  $a = 3$**

Wieder werde zunächst  $\pi$  fest gewählt und dabei sei  $\pi > 0$  (für  $\pi = 0$  ist  $f \equiv 0$  für  $p_1, p_2, p_3 \in (0, 1)$ ). Es wird nun  $f$  in Abhängigkeit von  $p_1$  und  $p_2$  untersucht, wobei  $p_3 = 1 - p_1 - p_2$  gilt.

In Abbildung 5.5 ist auf der linken Seite als Beispiel der Konturplot von  $f(p_1, p_2)$  für  $\pi = 0.5$  gegeben.

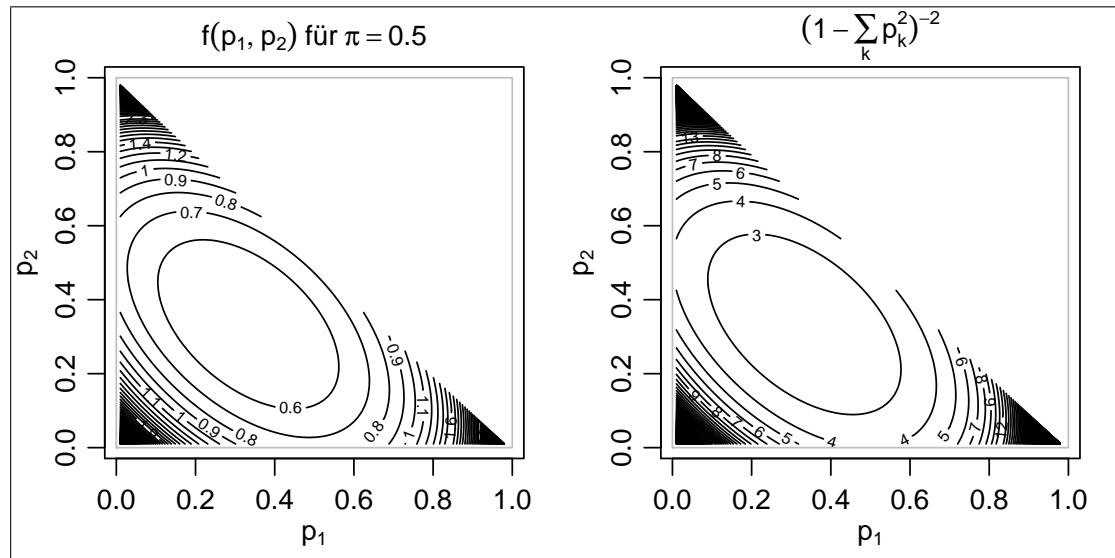


Abbildung 5.5: Konturplot von  $f(p_1, p_2)$  für  $\pi = 0.5$  (links) und von  $(1 - \sum_k p_k^2)^{-2}$  (rechts).

Die grundlegende Form des Graphen von  $f(p_1, p_2)$  lässt sich, wie auch schon im Fall  $a = 2$ , für  $\pi > 0$  beschreiben als Tal, wobei das Tal nun dreidimensional ist. Die Tal-Form des Graphen von  $f(p_1, p_2)$  wird im Wesentlichen vom Faktor

$$\frac{1}{n(1 - \sum_k p_k^2)^2}$$

bestimmt, der von  $\pi$  unabhängig ist und dessen Graph in Abbildung 5.5 auf der rechten Seite zu sehen ist. Lediglich die Richtung der Ausdehnung und die Tiefe des Tals werden durch den Zähler von  $f$  beeinflusst.

So dehnt sich das Tal für kleine  $\pi$  stärker in Richtung der Punkte  $(0.5, 0.5)$ ,  $(0, 0.5)$  und  $(0.5, 0)$  aus, während sich das Tal für große  $\pi$  stärker in Richtung der Punkte  $(1, 0)$ ,  $(0, 1)$  und  $(0, 0)$  ausdehnt. Dies lässt sich erkennen beim Vergleich von Konturplots für  $f$  und Konturplots für den Zähler von  $f$ , die für verschiedene Werte von  $\pi$  in Abbildung 5.6 gezeigt werden.

Der Term

$$\frac{1}{n(1 - \sum_k p_k^2)^2}$$

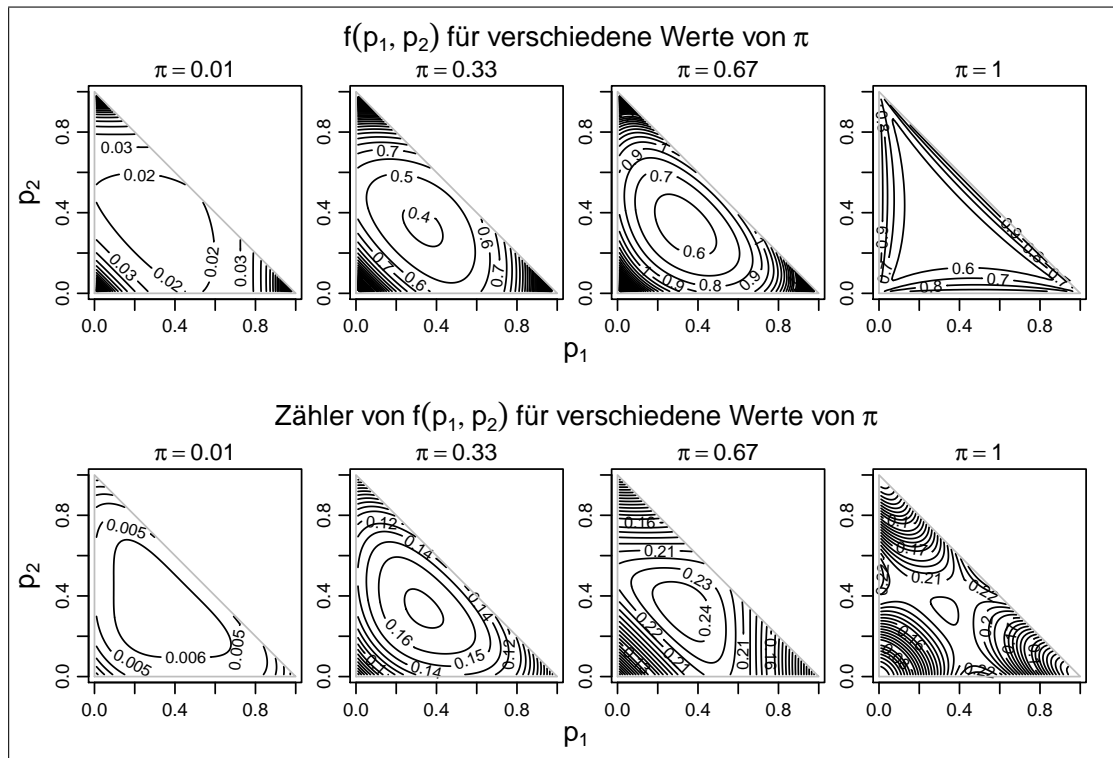


Abbildung 5.6: Konturplot von  $f(p_1, p_2)$  (oben) und vom Zähler von  $f(p_1, p_2)$  (unten) für verschiedene Werte von  $\pi$ .

nimmt sein Minimum an, wenn  $\sum_k p_k^2$  sein Minimum annimmt. Die partielle Ableitung von  $\sum_k p_k^2$  nach  $p_r$  lautet  $2p_r$  für alle  $r$ , sodass nach dem Lagrange-Ansatz ein Extremum angenommen wird, wenn  $p_r = p_s$  gilt für alle  $r, s \in \{1, 2, 3\}$ , d.h. wenn

$$p_1 = p_2 = p_3 = \frac{1}{3}$$

ist. Da auf der Diagonalen der Hesse-Matrix in jeder Zeile eine 2 eingetragen ist und alle Einträge außerhalb der Diagonalen 0 sind, ist die Hesse-Matrix positiv definit und somit ist das Extremum ein Minimum. Zusammenfassend wird

$$\frac{1}{n \left(1 - \sum_k p_k^2\right)^2}$$

minimal für  $p_1 = p_2 = p_3 = \frac{1}{3}$ .

**Bemerkung 5.17.** Die Bestimmung der Minimalstelle von

$$\frac{1}{n \left(1 - \sum_k p_k^2\right)^2}$$

lässt sich verallgemeinern für beliebige  $a > 1$ . Analog zur obigen Argumentation folgt, dass das Minimum angenommen wird für

$$p_1 = \dots = p_a = \frac{1}{a}.$$

Wie in Bemerkung 5.15 beschrieben, lässt sich aus Symmetriegründen feststellen, dass durch

$$p_1 = p_2 = p_3 = \frac{1}{3}$$

eine Extremalstelle von  $f(p_1, p_2)$  beschrieben wird. Die berechneten Werte legen, wie im Fall  $a = 2$ , auch im Fall  $a = 3$  nahe, dass in dieser Stelle ein globales Minimum angenommen wird.

Der Wert von  $\pi$ , für den  $f(\pi)$  maximal wird bei gegebenem  $(p_1, p_2)$ , lässt sich mit der Formel für  $\pi_{\max}$  aus Bemerkung 5.16 bestimmen. Der Konturplot von  $\pi_{\max}(p_1, p_2)$  ist in Abbildung 5.7 auf der linken Seite gegeben. Auf der rechten Seite zeigt Abbildung 5.7 Graphen von  $\pi_{\max}(p_1) = \pi_{\max}(p_1, p_2)$  für verschiedene feste Werte von  $p_2$ .

Die für den Konturplot berechneten Werte von  $\pi_{\max}(p_1, p_2)$ , also mit

$$p_1, p_2 \in \{0.01, 0.02, \dots, 0.99\},$$

liegen im Intervall  $[0.506, 0.956]$ . Die kleinsten Werte werden angenommen für  $p_r \nearrow 1$  für ein  $r \in \{1, 2, 3\}$ , die größten Werte werden angenommen für  $p_r = p_s \nearrow 0.5$  für  $r, s \in \{1, 2, 3\}$  mit  $r \neq s$ .

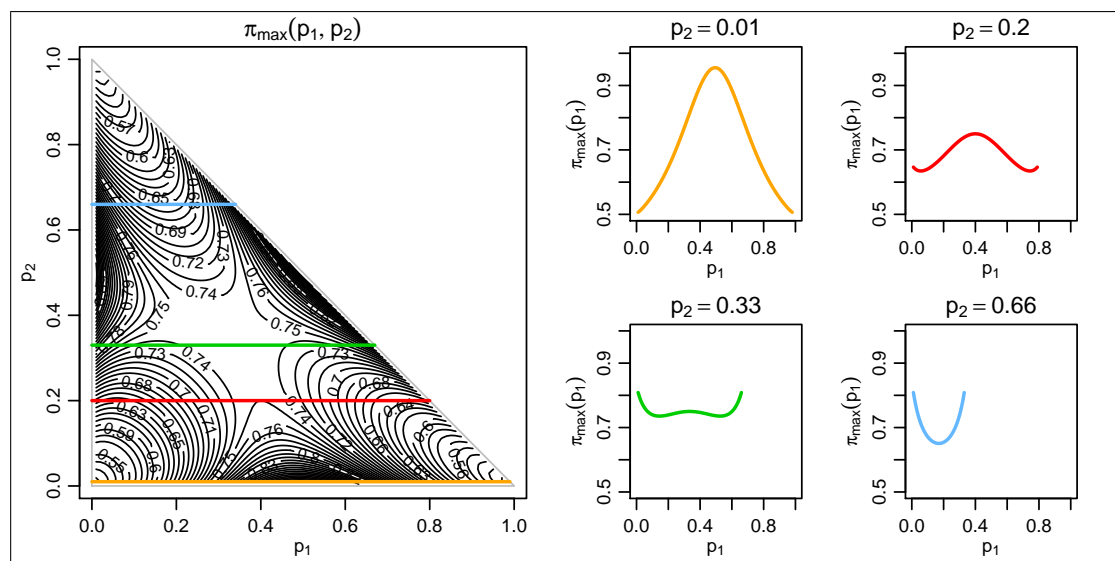


Abbildung 5.7: Konturplot von  $\pi_{\max}(p_1, p_2)$  (links) mit Querschnitten für verschiedene Werte von  $p_2$  (rechts).

**Der Fall  $a = 4$**

Betrachtet man die Konturplots von  $f(p_2, p_3)$  für verschiedene feste Werte von  $p_1$  und  $\pi$ , wie in Abbildung 5.8, so stellt man fest, dass jeweils die Talförmigkeit der Graphen erhalten bleibt.

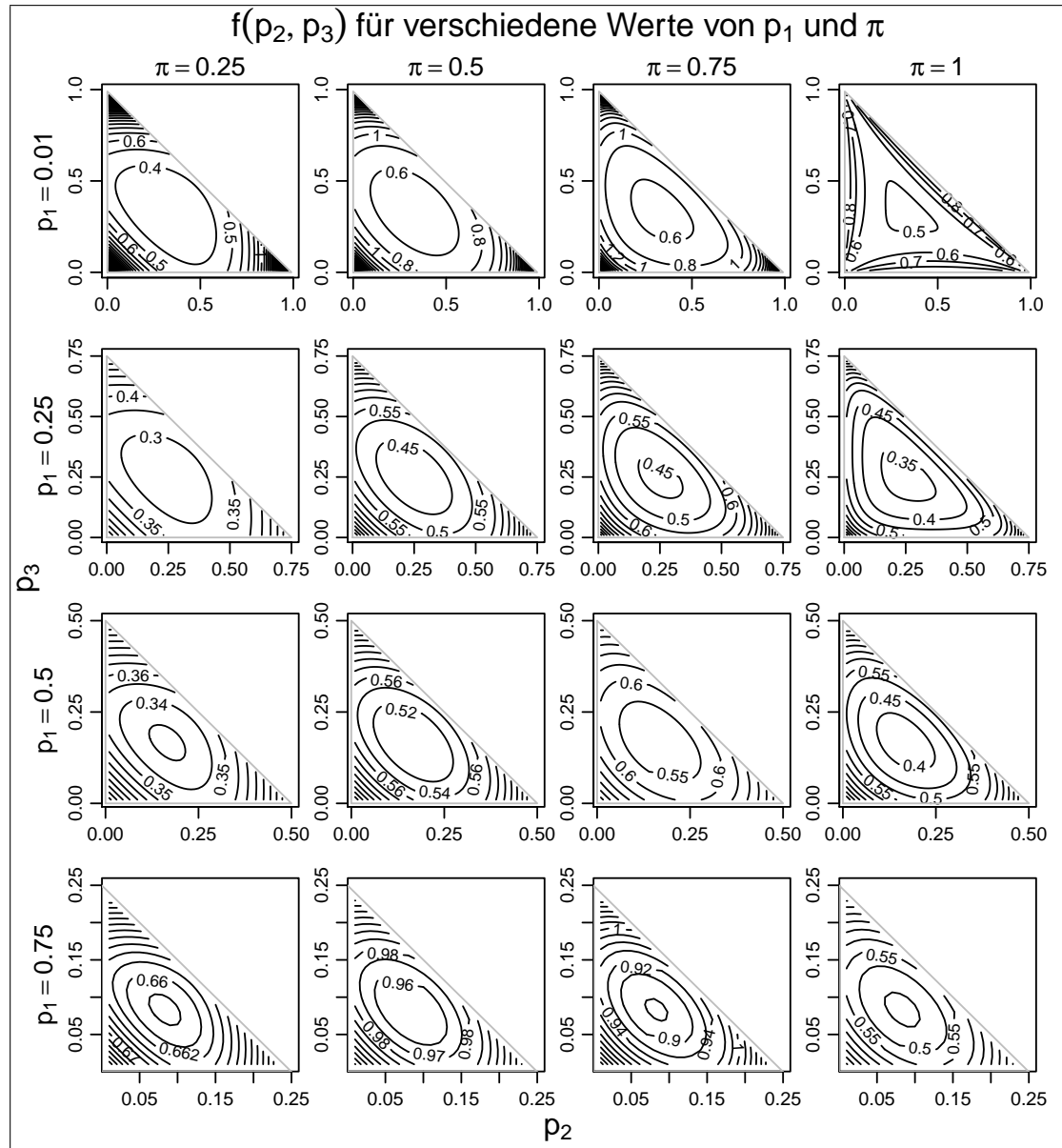


Abbildung 5.8: Konturplot von  $f(p_2, p_3)$  für verschiedene Werte von  $p_1$  (pro Zeile) und von  $\pi$  (pro Spalte).

Der maximale Wert, den  $p_2$ ,  $p_3$  und  $p_4$  für gegebenes  $p_1$  annehmen können, ist  $1 - p_1$ . Eine Extremalstelle von  $f(p_2, p_3)$  befindet sich dann aus Symmetriegründen im Punkt

mit

$$p_2 = p_3 = p_4 = \frac{1 - p_1}{3}.$$

Die für die Konturplots berechneten Funktionswerte deuten darauf hin, dass es sich bei der Extremalstelle jeweils um eine globale Minimalstelle handelt für ein festes  $p_1$ .

Für jedes  $\pi \in \{0.25, 0.50, 0.75, 1.00\}$  ist der geringste Funktionswert von  $f$  im jeweiligen Konturplot für den Fall  $p_1 = 0.25$  zu finden, und zwar an der Stelle

$$p_1 = p_2 = p_3 = p_4 = \frac{1}{4}.$$

### Zusammenfassung der Beobachtungen

Für  $\pi = 0$  ist  $f \equiv 0$  für alle  $p_1, \dots, p_a \in [0, 1)$  mit  $\sum_r p_r = 1$ .

Ist  $\pi > 0$  gegeben, hat der Graph von  $f(p_1, \dots, p_{a-1})$  die Form eines  $a$ -dimensionalen Tals. Die Tal-Form wird durch den Faktor

$$\frac{1}{n \left(1 - \sum_k p_k^2\right)^2}$$

herbeigeführt (unabhängig von  $\pi$ ). Die Richtung der Ausdehnung des Tals wird durch den Zähler von  $f(p_1, \dots, p_{a-1})$  bestimmt in Abhängigkeit von  $\pi$ .

Darüber hinaus nimmt der Faktor

$$\frac{1}{n \left(1 - \sum_k p_k^2\right)^2}$$

sein globales Minimum an für

$$p_r = \frac{1}{a} \quad \forall r \in \{1, \dots, a\}.$$

Die Funktion  $f(p_1, \dots, p_{a-1})$  nimmt ein Extremum an für

$$p_r = \frac{1}{a} \quad \forall r \in \{1, \dots, a\}.$$

Die für  $a = 2, 3, 4$  berechneten Werte lassen darauf schließen, dass es sich um ein globales Minimum handelt.

Sind neben  $\pi$  auch  $p_1, \dots, p_t$  fest gegeben, wird von  $f(p_{t+1}, \dots, p_{a-1})$  aus Symmetriegründen ein Extremum angenommen für

$$p_{t+1} = \dots = p_a = \frac{1 - \sum_{r=1}^t p_r}{a - t}.$$

Die für  $a = 2, 3, 4$  berechneten Werte lassen darauf schließen, dass es sich in den meisten Fällen um ein globales Minimum handelt auf dem eingeschränkten Definitionsbereich von  $f$ .

Für feste  $p_1, \dots, p_a$  lässt sich der Wert  $\pi_{\max}$  von  $\pi$ , für den  $f(\pi)$  maximal wird, berechnen mithilfe der Formel aus Bemerkung 5.16. Die für  $a = 2, 3$  berechneten Werte für  $\pi_{\max}$  lassen darauf schließen, dass  $\pi_{\max}$  im Intervall  $[0.5, 1]$  liegt.



## 6 Überprüfung des SMI-Modells

Nachdem die Modell-Schätzer  $\hat{p}_1, \dots, \hat{p}_a$  und  $\hat{\pi} = \hat{\pi}^{(g)}$  für die Anpassung des SMI-Modells an die gegebenen Daten in Kapitel 4 entwickelt worden sind, stellt sich nun die Frage, ob sich die Daten durch das SMI-Modell beschreiben lassen oder sich die Daten signifikant vom SMI-Modell unterscheiden.

Um für einen Probanden  $J$  zu einem festen Schlafstadium  $Z$  die jeweilige Anpassungsgüte des SMI-Modells zu untersuchen, wird in Kapitel 6.1 ein Anpassungstest angewendet, der sich an Pearsons  $\chi^2$ -Anpassungstest orientiert. Die zugehörigen p-Werte werden dabei durch Simulationen von Markov-Ketten gemäß dem SMI-Modell und die daraus resultierenden empirischen Verteilungsfunktionen der  $\chi^2$ -Statistiken bestimmt.

Es wird sich zeigen, dass auf einem Signifikanz-Niveau von 5% für weit mehr als die Hälfte aller Paare  $(J, Z)$  die Nullhypothese verworfen wird, dass die gegebenen Daten  $(n_{rs}^{(J,Z)})_{r,s}$  einer Markov-Kette der Länge  $n^{(J,Z)} + 1$  entstammen, die die Bedingungen des SMI-Modells erfüllt.

Aus diesem Grund wird in Kapitel 6.2 untersucht, inwiefern das SMI-Modell von den gegebenen Daten abweicht. Dazu wird die Eigenschaft ausgenutzt, dass im SMI-Modell mit Wahrscheinlichkeit  $\pi$  bei einem Übergang von  $X_k$  zu  $X_{k+1}$  in der Markov-Kette  $(X_i)_{i \in \mathbb{N}_0}$  mindestens ein Intervallwechsel im Hintergrundmodell stattfindet unabhängig vom Ausgangszustand  $X_k$ . Für die geschätzten Parameter des SMI-Modells gilt

$$\hat{\pi} = \frac{\hat{p}_{rs}}{\hat{p}_s} = \frac{1 - \hat{p}_{rr}}{1 - \hat{p}_r} \quad \forall r, s \in S, r \neq s.$$

Geht man davon aus, dass das SMI-Modell die Verteilung der gegebenen beobachteten Daten  $(n_{rs}^{(J,Z)})_{r,s}$  beschreibt, dann könnte man in den obigen Gleichungen den Schätzer  $\hat{p}_{rs}$  durch den Schätzer  $\hat{p}_{rs}^{(v)}$  des vollen Modells ersetzen, ohne dass die daraus resultierenden Werte von

$$\frac{\hat{p}_{rs}^{(v)}}{\hat{p}_s} \quad \text{bzw.} \quad \frac{1 - \hat{p}_{rr}^{(v)}}{1 - \hat{p}_r},$$

die dann mit positiver Wahrscheinlichkeit verschieden sind für jedes Paar  $(r, s)$  bzw.  $(r, r)$ , zu stark von  $\hat{\pi}$  abweichen.

Auch hier werden gemäß dem SMI-Modell simulierte Markov-Ketten betrachtet, um für den Übergang von  $r$  zu  $s$  bzw. von  $r$  zu  $r$  die typische Abweichung der obigen Brüche von  $\hat{\pi}$  zu ermitteln. Durch den Vergleich der typischen Abweichung mit der tatsächlichen, aus den gegebenen Daten ermittelten Abweichung der entsprechenden Brüche von  $\hat{\pi}$  lassen sich Rückschlüsse darauf ziehen, welche Übergangswahrscheinlichkeiten  $p_{rs}$  durch das SMI-Modell nicht ausreichend erfasst werden und in welche Richtung das SMI-Modell korrigiert werden müsste.

Für die Übergangswahrscheinlichkeiten des SMI-Modells, bei denen pro Schlafstadium die größte Anzahl von Probanden signifikante Abweichungen von den geschätzten Übergangswahrscheinlichkeiten des vollen Modell zeigt, wird in Kapitel 6.3 der Versuch unternommen, die entsprechenden Übergangswahrscheinlichkeiten durch eine um drei Parameter erweiterte Version des SMI-Modells (Modell 4) zu korrigieren.

## 6.1 Ein $\chi^2$ -Anpassungstest für das SMI-Modell mithilfe von Simulation

In diesem Unterkapitel wird ein Anpassungstest beschrieben und durchgeführt, der sich an Pearsons  $\chi^2$ -Anpassungstest orientiert und mit dem für einen Probanden  $J$  und ein Schlafstadium  $Z$  die folgende Nullhypothese getestet wird:

**Nullhypothese  $H_0^{(J,Z)}$ :** Die beobachteten Übergangshäufigkeiten  $(n_{rs}^{(J,Z)})_{r,s}$  sind so verteilt wie die Übergangshäufigkeiten einer Markov-Kette der Länge  $n^{(J,Z)} + 1$  mit Zustandsraum  $S = \{A, B, C, D\}$ , die die Bedingungen des SMI-Modells erfüllt.

In Kapitel 6.1.1 wird zunächst ein kurzer Überblick über die Entwicklung von  $\chi^2$ -Anpassungstests für Markov-Ketten gegeben ausgehend von Pearsons  $\chi^2$ -Anpassungstest. Die bei den  $\chi^2$ -Anpassungstests betrachteten Test-Statistiken  $\mathcal{X}^2$  sind jeweils asymptotisch  $\chi^2$ -verteilt. Werden bei der Berechnung von  $\mathcal{X}^2$  Parameter geschätzt, beeinflusst dies die Anzahl der Freiheitsgrade der  $\chi^2$ -Verteilungen. Die Bestimmung der Anzahl der Freiheitsgrade der jeweiligen  $\chi^2$ -Verteilungen wird in diesem Kapitel ebenfalls thematisiert.

Da, wie sich zeigen wird, bei der Verwendung der Schätzer  $\hat{p}_1, \dots, \hat{p}_a$  und  $\hat{\pi}$  die Anzahl der Freiheitsgrade nicht bekannt ist, werden die Verteilungen der Test-Statistiken  $\mathcal{X}_{(J,Z)}^2$  für jeden Probanden  $J$  und jedes Schlafstadium  $Z$  beim Test der Nullhypothese  $H_0^{(J,Z)}$  mithilfe von durch Simulationen erzeugten empirischen Verteilungsfunktionen approximiert. Die Beschreibung des genauen Vorgehens beim Testen von  $H_0^{(J,Z)}$  erfolgt in Kapitel 6.1.2.

Die Ergebnisse der einzelnen Tests zu den Nullhypothesen  $H_0^{(J,Z)}$  werden in Kapitel 6.1.3 vorgestellt, ebenso wie die Ergebnisse der simultanen Tests der Nullhypothesen aller Probanden in einem Schlafstadium  $Z$ .

Um zu überprüfen, dass durch das Testverfahren aus Kapitel 6.1.2 die Nullhypothesen  $H_0^{(J,Z)}$  jeweils nur mit einer Wahrscheinlichkeit von 5% abgelehnt werden unter der Voraussetzung, dass  $H_0^{(J,Z)}$  richtig ist, werden in Kapitel 6.1.4 drei Datensätze

$$(n_{rs}^{\{k\}})_{rs}, \quad k \in \{1, 2, 3\},$$

gemäß dem SMI-Modell mit zufälligen Parametern

$$p_1^{\{k\}}, p_2^{\{k\}}, p_3^{\{k\}}, p_4^{\{k\}} \quad \text{und} \quad \pi^{\{k\}}$$

erzeugt und anschließend für diese die p-Werte zu den analogen Nullhypothesen  $H_0^{\{k\}}$  bestimmt. Daraufhin wird für diese p-Werte die Nullhypothese getestet, dass die p-Werte uniform auf  $[0, 1]$  verteilt sind.

In Kapitel 6.1.5 werden mögliche Gründe dafür erläutert, dass eine der Nullhypothesen  $H_0^{(J,Z)}$  abgelehnt wird.

### 6.1.1 Mathematischer Hintergrund

Im Jahre 1900 wurde von Pearson der  $\chi^2$ -Anpassungstest vorgestellt (Pearson, 1900). Dieser Anpassungstest prüft die Nullhypothese, dass die zufällige Aufteilung  $(n_1, \dots, n_k)$  von  $n$  unabhängigen Beobachtungen  $X_1, \dots, X_n$  einer Zufallsvariablen auf  $k$  verschiedenen Kategorien zu einer Verteilung  $F_0$  konsistent ist.

Für die Anzahl  $n_i$  der Beobachtungen, die in Kategorie  $i$  fallen, mit  $\sum_{i=1}^k n_i = n$ , wird die  $\chi^2$ -Statistik beschrieben durch

$$\chi^2 = \sum_{i=1}^k \frac{(n_i - \mathbb{E}_0[n_i])^2}{\mathbb{E}_0[n_i]},$$

wobei mit  $\mathbb{E}_0[n_i]$  die unter der Nullhypothese erwartete Anzahl von Beobachtungen der Kategorie  $i$  bezeichnet werde. Die Teststatistik  $\chi^2$  ist unter der Nullhypothese asymptotisch  $\chi^2(k-1)$ -verteilt. Die Nullhypothese wird auf einem Signifikanzniveau von  $\alpha$  verworfen, wenn  $\chi^2$  größer ist als das  $(1-\alpha)$ -Quantil der  $\chi^2(k-1)$ -Verteilung (siehe z. B. Wasserman, 2004).

Möchte man dieses Testverfahren auf eine endliche Markov-Kette  $(X_i)_{i \in \{0, \dots, n\}}$  mit endlichem Zustandsraum  $S$  übertragen, kann man alle geordneten Paare von Zuständen  $(r, s)$  als Kategorien auffassen, auf die sich die Paare  $(X_{i-1}, X_i)$  aufteilen,  $i \in \{1, \dots, n\}$ . Ist  $|S| = a$ , gibt es also  $a^2$  Kategorien.

Während jedoch beim von Pearson behandelten Fall die Beobachtungen unabhängig voneinander sind, hängen die nun betrachteten  $((X_{i-1}, X_i))_{i \in \{1, \dots, n\}}$  voneinander ab. Außerdem ist es etwa für reduzible Markov-Ketten  $(X_i)_{i \in \mathbb{N}_0}$  möglich, dass ein Zustand  $r$  angenommen wird, von dem aus sich ein Zustand  $s$  nicht mit positiver Wahrscheinlichkeit in endlichen vielen Schritten erreichen lässt.

Dieser beiden Probleme hat sich Bartlett im Jahr 1951 angenommen und Voraussetzungen an die Markov-Kette angegeben, unter denen sich der  $\chi^2$ -Anpassungstest auf Markov-Ketten übertragen lässt (Bartlett, 1951):

**Satz 6.1.** *Sei  $X$  eine Markov-Kette der Länge  $n+1$  mit endlichem Zustandsraum  $S = \{1, \dots, a\}$  und Übergangsmatrix  $P = (p_{rs})_{rs}$ . Für alle  $r, s \in S$  sei  $n_r$  die Auftrittshäufigkeit des Zustandes  $r$  in  $X$  und  $n_{rs}$  die Übergangshäufigkeit vom Zustand  $r$  zum Zustand  $s$  in  $X$ . Ist die Markov-Kette regulär, d.h. es existiert ein  $m \in \mathbb{N}$ , sodass*

sämtliche Einträge von  $P^m$  größer 0 sind, dann ist

$$\mathcal{X}^2 = \sum_{r,s=1}^a \frac{(n_{rs} - n_r p_{rs})^2}{n_r p_{rs}}$$

asymptotisch  $\chi^2$ -verteilt mit  $c - a$  Freiheitsgraden, wobei  $c$  die Anzahl der Einträge in  $P$  ist, die  $> 0$  sind. Sind alle  $p_{rs} > 0$ , impliziert das also  $a^2 - a$  Freiheitsgrade.

*Beweisskizze.* Im Wesentlichen muss gezeigt werden, dass die Übergangshäufigkeiten  $n_{rs}$  asymptotisch normalverteilt sind. Genauer gilt (nach Billingsley, 1961):

$$((n_{rs} - n_r p_{rs})/\sqrt{n_r})_{r,s \in S} \xrightarrow{n \rightarrow \infty} \mathcal{N}(0, \Sigma)$$

mit Kovarianzmatrix  $\Sigma$  gegeben durch

$$\Sigma_{rs,tu} = \delta_{r,t}(\delta_{s,u} p_{rs} - p_{rs} p_{ru}).$$

Sei  $r \in S$  fest gewählt und  $c_r := |\{s \in S : p_{rs} > 0\}|$ . Dann folgt, dass

$$\mathcal{X}_r^2 := \sum_{s=1}^a \frac{(n_{rs} - n_r p_{rs})^2}{n_r p_{rs}}$$

asymptotisch  $\chi^2$ -verteilt ist mit  $c_r - 1$  Freiheitsgraden.

Da die Statistiken  $\mathcal{X}_1^2, \dots, \mathcal{X}_a^2$  asymptotisch unabhängig sind, ist

$$\mathcal{X}^2 = \sum_{r=1}^a \mathcal{X}_r^2 = \sum_{r,s=1}^a \frac{(n_{rs} - n_r p_{rs})^2}{n_r p_{rs}}$$

asymptotisch  $\chi^2$ -verteilt mit  $\sum_{r=1}^a (c_r - 1) = c - a$  Freiheitsgraden, wobei  $c = \sum_{r=1}^a c_r$  ist.  $\square$

**Bemerkung 6.2.** Wegen der asymptotischen Betrachtung in Satz 6.1 kann die Startverteilung der Markov-Kette  $X$  beliebig gewählt werden.

Mit Satz 6.1 lässt sich die Nullhypothese untersuchen, dass die Übergangsmatrix der Markov-Kette  $(X_i)_{i \in \{0, \dots, n\}}$ , die den Übergangshäufigkeiten  $n_{rs}$  zugrunde liegt, die Form  $P = (p_{rs})_{r,s \in S}$  hat. Die entsprechende Teststatistik ist dann

$$\mathcal{X}^2 = \sum_{r,s \in S} \frac{(n_{rs} - n_r p_{rs})^2}{n_r p_{rs}}.$$

### $\chi^2$ -Anpassungstest mit geschätzten Parametern

Da die Werte von  $p_{rs}$  nicht bekannt sind, müssen sie für das SMI-Modell geschätzt werden. Es ist dann zu klären, ob und wie sich das Schätzen von Parametern auf die asymptotische Verteilung von  $\mathcal{X}^2$  auswirkt.

Der folgende Satz, der eine Adaption eines Satzes von Cramér (1946) an den vorliegenden Markov-Ketten-Fall mit Zustandsraum  $S = \{1, \dots, a\}$  ist und ursprünglich auf Fisher (1924) zurückgeht, nennt Voraussetzungen, unter denen die asymptotische Verteilung von  $\mathcal{X}^2$  bekannt ist.

**Satz 6.3.** *Seien die Übergangswahrscheinlichkeiten  $p_{rs} = p_{rs}(\vartheta) > 0$  Funktionen in Abhängigkeit eines  $d$ -dimensionalen Parametervektors  $\vartheta = (\vartheta_1, \dots, \vartheta_d)$ , der durch  $\hat{\vartheta}$  geschätzt werde. Dann ist*

$$\mathcal{X}^2 = \sum_{r,s \in S} \frac{(n_{rs} - n_r p_{rs}(\hat{\vartheta}))^2}{n_r p_{rs}(\hat{\vartheta})}$$

*asymptotisch  $\chi^2$ -verteilt mit  $a^2 - a - d$  Freiheitsgraden, falls die Schätzer  $\hat{\vartheta}_1, \dots, \hat{\vartheta}_d$  asymptotisch normalverteilt und effizient sind.*

Dabei heißt ein Schätzer  $\hat{\vartheta}$  effizient, wenn er erwartungstreu ist und  $\text{Var}(\hat{\vartheta}) = I(\vartheta)^{-1}$  gilt mit der Fisher-Information  $I(\vartheta)$  von  $\vartheta$  (siehe z. B. Held, 2008).

**Bemerkung 6.4.** Unter bestimmten Regularitätsvoraussetzungen, die von Billingsley (1961) für den betrachteten Markov-Ketten-Fall gegeben werden, lässt sich die Formel  $a^2 - a - d$  für die Freiheitsgrade beispielsweise für den ML-Schätzer von  $\vartheta$  anwenden.

Die zu berechnende Teststatistik  $\mathcal{X}^2$  hat für das SMI-Modell die Form

$$\mathcal{X}^2 = \sum_{r,s \in S} \frac{(n_{rs} - n_r \hat{p}_{rs})^2}{n_r \hat{p}_{rs}} \quad (6.1)$$

mit  $\hat{p}_{rs} = \hat{\pi} \hat{p}_s + \delta_{r,s}(1 - \hat{\pi})$ .

Die Schätzer  $\hat{p}_{rs}$  hängen per Definition in Kapitel 4.4 von  $a$  Parametern ab, nämlich von  $\hat{\pi}$  und o.B.d.A. von  $\hat{p}_1, \dots, \hat{p}_{a-1}$  mit  $\hat{p}_a = 1 - \sum_{r \neq a} \hat{p}_r$ . Der Satz 6.3 lässt sich hier jedoch nicht anwenden, da nicht bekannt ist, ob die Modell-Schätzer  $\hat{p}_1, \dots, \hat{p}_a$  und  $\hat{\pi}$  effizient sind.

Um trotzdem Aussagen über die Verteilung von  $\mathcal{X}^2$  treffen zu können, wird im folgenden Abschnitt die empirische Verteilungsfunktion von  $\mathcal{X}^2$  durch Simulation bestimmt.

### 6.1.2 Beschreibung der Durchführung des Anpassungstests

In diesem Abschnitt wird die  $\mathcal{X}^2$ -Statistik aus Gleichung (6.1) für jeden Probanden  $J$  und jedes Schlafstadium  $Z$  aus den gegebenen Daten berechnet, um die Anpassungsgüte des SMI-Modells zu untersuchen. Da die (asymptotische) Verteilung von  $\mathcal{X}^2$  nicht bekannt ist, wird die empirische Verteilungsfunktion von  $\mathcal{X}^2$  mittels Simulation ermittelt.

Dazu werden die Übergangshäufigkeiten  $n_{rs}$  für 10000 Markov-Ketten der Länge  $n + 1$  gemäß der Verteilung simuliert, die durch das SMI-Modell mit den Parametern

$$\hat{p}_1, \hat{p}_2, \hat{p}_3, \hat{p}_4 \quad \text{und} \quad \hat{\pi}$$

gegeben ist ( $\hat{p}_4 = 1 - \hat{p}_1 - \hat{p}_2 - \hat{p}_3$ ). Dabei werden die benötigten Schätzer aus den Daten berechnet nach den Formeln aus Kapitel 4.

Um die Variablen zu kennzeichnen, die zu den verschiedenen simulierten Markov-Ketten gehören, wird der Index  $[i]$  für die  $i$ -te simulierte Markov-Kette verwendet.

Es wird nun der genaue Vorgang für den Anpassungstest durch Simulation beschrieben für einen Probanden  $J$  im Schlafstadium  $Z$ .

Gegeben sind die Daten  $n_{rs} = n_{rs}^{(J,Z)}$  für alle  $r, s \in S = \{1, 2, 3, 4\}$ .

1. Berechne  $n = \sum_{r,s \in S} n_{rs}$  und  $n_t = \sum_{u \in S} n_{tu}$  für alle  $t \in S$ .
2. Berechne die Schätzer  $\hat{p}_1, \hat{p}_2, \hat{p}_3, \hat{p}_4$  gemäß Gleichung (4.1) und  $\hat{\pi} = \hat{\pi}^{(g)}$  gemäß Gleichung (4.2) aus den  $n_{rs}$ .
3. Berechne die geschätzten Übergangswahrscheinlichkeiten  $\hat{p}_{rs} = \hat{\pi} \hat{p}_s + \delta_{r,s}(1 - \hat{\pi})$ .
4. Berechne  $\mathcal{X}_{(J,Z)}^2 = \mathcal{X}^2$  gemäß Gleichung (6.1).
5. Simuliere 10000 Markov-Ketten  $X^{[i],(J,Z)} = X^{[i]} = (X_0^{[i]}, \dots, X_n^{[i]})$  der Länge  $n+1$  mit Startverteilung  $(\hat{p}_1, \hat{p}_2, \hat{p}_3, \hat{p}_4)$  und Übergangsmatrix  $\hat{P} = (\hat{p}_{rs})_{r,s \in S}$  und führe die Schritte 6 bis 9 für jedes  $X^{[i]}$  aus,  $i = 1, \dots, 10000$ .
6. Bestimme  $n_{rs}^{[i]} = n_{rs}^{[i],(J,Z)}$  und  $n_r^{[i]}$  für alle  $r, s \in S$ .
7. Berechne die Schätzer  $\hat{p}_1^{[i]}, \hat{p}_2^{[i]}, \hat{p}_3^{[i]}, \hat{p}_4^{[i]}$  und  $\hat{\pi}^{[i]}$  aus den  $n_{rs}^{[i]}$ .
8. Berechne die Übergangswahrscheinlichkeiten  $\hat{p}_{rs}^{[i]} = \hat{\pi}^{[i]} \hat{p}_s^{[i]} + \delta_{r,s}(1 - \hat{\pi}^{[i]})$ .

$$9. \text{ Berechne } \mathcal{X}_{[i],(J,Z)}^2 = \mathcal{X}_{[i]}^2 := \sum_{r,s \in S} \frac{(n_{rs}^{[i]} - n_r^{[i]} \hat{p}_{rs}^{[i]})^2}{n_r^{[i]} \hat{p}_{rs}^{[i]}}.$$

10. Bestimme aus der empirischen Verteilung von  $(\mathcal{X}_{[i]}^2)_{i \in \{1, \dots, 10000\}}$  den empirischen p-Wert durch

$$\mathbf{p}^{(J,Z)} := \frac{|\{i \in \{1, \dots, 10000\} : \mathcal{X}_{[i]}^2 > \mathcal{X}^2\}|}{10000}.$$

**Nullhypothese**  $H_0^{(J,Z)}$ : Die beobachteten Übergangshäufigkeiten  $(n_{rs}^{(J,Z)})_{r,s}$  sind so verteilt wie die Übergangshäufigkeiten einer Markov-Kette der Länge  $n^{(J,Z)} + 1$  mit Zustandsraum  $S = \{A, B, C, D\}$ , die die Bedingungen des SMI-Modells erfüllt.

Große Werte von  $\mathcal{X}_{(J,Z)}^2$  sprechen dafür, dass die beobachteten Übergangshäufigkeiten zumindest teilweise stark von den durch das SMI-Modell erwarteten Übergangshäufigkeiten abweichen. Die Nullhypothese  $H_0^{(J,Z)}$  wird auf einem Signifikanzniveau von  $\alpha$  abgelehnt, wenn weniger als  $\alpha \cdot 100\%$  der simulierten  $\mathcal{X}_{[i],(J,Z)}^2$  größer sind als  $\mathcal{X}_{(J,Z)}^2$ , d.h. wenn der p-Wert  $\mathbf{p}^{(J,Z)} < \alpha$  ist.

### 6.1.3 Ergebnisse des Anpassungstests bei Anwendung auf die gegebenen Daten

In diesem Abschnitt werden die Ergebnisse des Anpassungstests bei einem Signifikanzniveau von 5% vorgestellt. Der Anpassungstest wurde entsprechend dem in Kapitel 6.1.2 beschriebenen Vorgehen auf die gegebenen Daten angewendet für jeden Probanden  $J$  und jedes Schlafstadium  $Z$ .

#### Einzelne Tests

Um eine Vorstellung von der Größe und Verteilung der errechneten  $\chi^2_{(J,Z)}$  und  $\chi^2_{(J,Z),[i]}$  zu bekommen, werden zunächst die Probanden 2 und 5 betrachtet, die schon in den Kapiteln 3.1 und 4.5 für Beispiele herangezogen wurden.

In Abbildung 6.1 werden die Ergebnisse für Proband 2 dargestellt und in Abbildung 6.2 die für Proband 5. Die Abbildungen zeigen für jedes Schlafstadium die Histogramme der 10000 simulierten  $\chi^2_{[i]}$  mit den zugehörigen p-Werten des Anpassungstests. Dabei ist zu beachten, dass Proband 2 das Schlafstadium N3 nicht erreicht hat. Die Balken der Histogramme sind so normiert, dass die Flächen der Balken in der Summe 1 ergeben. Damit lassen sich die Histogramme als empirische Dichtefunktion interpretieren. In jedem Histogramm sind sowohl der durch die Daten beobachtete Wert  $\chi^2_{(2,Z)}$  bzw.  $\chi^2_{(5,Z)}$  (grüne Linie) als auch das 95%-Quantil der simulierten  $\chi^2_{[i),(2,Z)}$  bzw.  $\chi^2_{[i),(5,Z)}$  (rote Linie) auf der x-Achse markiert.

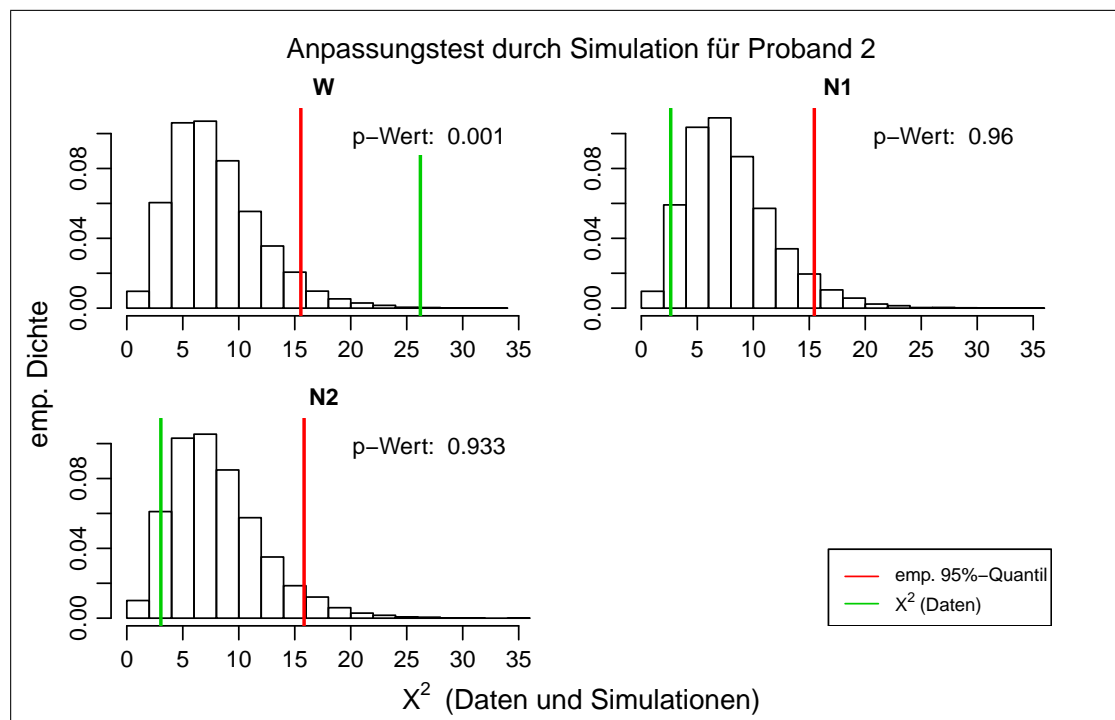


Abbildung 6.1: Anpassungstests durch Simulation für Proband 2, alle erreichten Schlafstadien. Die p-Werte sind auf 3 Stellen gerundet.

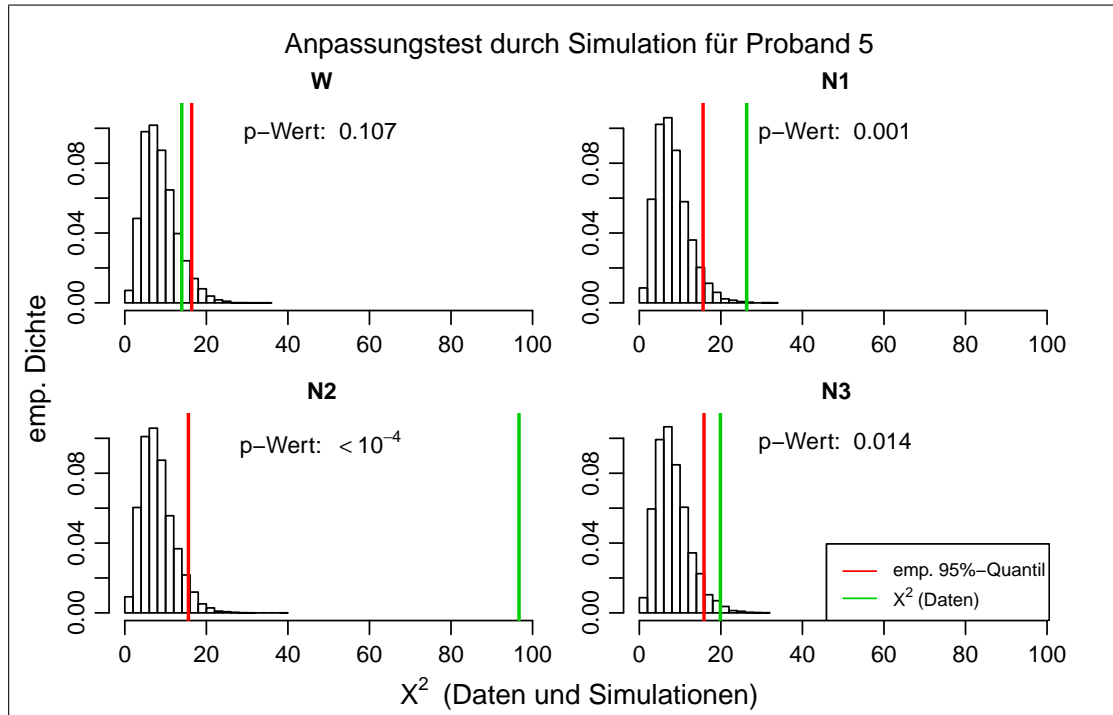


Abbildung 6.2: Anpassungstests durch Simulation für Proband 5, alle erreichten Schlafstadien. Die p-Werte sind auf 3 Stellen gerundet.

Es fällt bei der Betrachtung mit bloßem Auge auf, dass sich die empirischen Dichtefunktionen eines Probanden zu den verschiedenen Schlafstadien sehr ähnlich sind. Dementsprechend liegen die eingezeichneten 95%-Quantile dicht beieinander zwischen 15.45 und 15.84. Bei Berücksichtigung der verschiedenen Maßstäbe auf der x-Achse lässt sich zudem feststellen, dass die empirischen Dichtefunktionen auch für die beiden Probanden 2 und 5 annähernd gleich sind. Dies entspricht den in Kapitel 6.1.1 eingeführten theoretischen Betrachtungen der asymptotischen Verteilung von  $\chi^2$ , dass diese Verteilung nur von der Anzahl der Freiheitsgrade abhängt, die durch die Anzahl  $a$  der Zustände und die Wahl der Schätzer gegeben und für alle Probanden und Schlafstadien gleich ist.

Für Proband 2 ergeben zu den verschiedenen Schlafstadien die p-Werte

$$p^{(2,W)} \approx 0.001, \quad p^{(2,N1)} \approx 0.960, \quad p^{(2,N2)} \approx 0.933.$$

Passend dazu, dass Proband 2 im Schlafstadium N2 in dieser Arbeit bisher als Beispiel aufgeführt worden ist für einen Fall, bei dem die geschätzten Übergangswahrscheinlichkeiten  $\hat{p}_{rs}$  des SMI-Modells sehr nah an den entsprechenden Schätzern  $\hat{p}_{rs}^{(v)} = n_{rs}/n_r$  des vollen Modells liegen (siehe Kapitel 4.5), wird die Nullhypothese  $H_0^{(2,N2)}$  bei einem relativ hohen p-Wert von  $p^{(2,N2)} \approx 0.933$  nicht verworfen. Während auch für N1 die Nullhypothese  $H_0^{(2,N1)}$  nicht verworfen wird, zeigen die Daten für das Schlafstadium W signifikante Abweichungen von den Modellannahmen des SMI-Modells.



Bei Proband 5 zeigt sich ein umgekehrtes Bild. Für das Schlafstadium W wird die Nullhypothese nicht abgelehnt, für die Schlafstadien N1, N2 und N3 wird sie verworfen. Auffällig ist dabei, dass  $\chi^2_{(5,N2)} \approx 96$  einen extrem großen Wert annimmt. In dieser Arbeit wurde Proband 5 im Schlafstadium N2 als Beispiel genannt für einen Fall, bei dem durch das bloße Betrachten der Unterschiede zwischen den Schätzern  $\hat{p}_{rs}$  und  $\hat{p}_{rs}^{(v)}$  in Abbildung 4.7 vermutet werden konnte, dass sich die Struktur der Übergangswahrscheinlichkeiten des SMI-Modells nicht mit den beobachteten Übergangswahrscheinlichkeiten  $\hat{p}_{rs}^{(v)}$  des vollen Modells vereinbaren lässt.

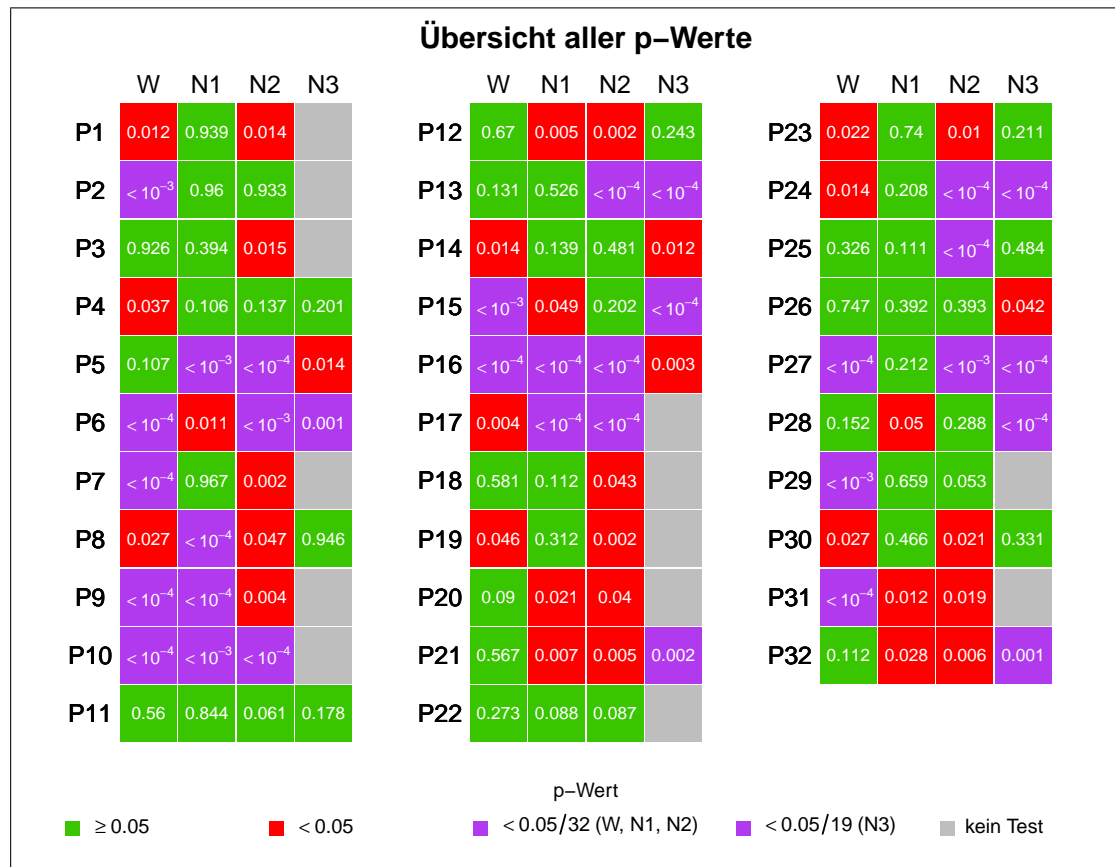


Abbildung 6.3: Übersicht über die p-Werte aller mithilfe von Simulation ausgeführten Anpassungstests für das SMI-Modell. (Die von der Legende abweichende Färbung für Proband 28 in N1 erklärt sich durch den Rundungsfehler.)

Abbildung 6.3 zeigt eine Übersicht der p-Werte  $p^{(J,Z)}$  aller Probanden  $J$  und Schlafstadien  $Z$ . Die grünen Quadrate geben an, dass die Nullhypothese für die zugehörigen Proband-Schlafstadium-Kombinationen  $(J, Z)$  durch den Anpassungstest auf einem Signifikanzniveau von 5% nicht abgelehnt wird. Bei einem roten oder violetten Quadrat wird die entsprechende Nullhypothese verworfen.

So lässt sich auf einen Blick erkennen, dass für einen Probanden die Nullhypothesen

zu N1, N2 und N3 nicht immer geschlossen verworfen bzw. geschlossen nicht verworfen werden (im Unterschied zu den Probanden 2 und 5) und dass das Testergebnis zu W nicht immer das Gegenteil der Ergebnisse zu N1, N2 und N3 liefern muss.

Weiterhin sieht man in Abbildung 6.3, dass es Probanden gibt, für die zu keinem Schlafstadium die Nullhypothese verworfen wird, nämlich für die Probanden 11 und 22. Dagegen wird für die Probanden 6, 9, 10, 16, 17 und 31 jede der Nullhypothesen verworfen.

Insgesamt dominieren p-Werte  $< 5\%$  die Ergebnisse der Anpassungstests. Für das Schlafstadium W werden bei 19 von 32 Probanden die Nullhypothesen verworfen, für N1 bei 14 von 32, für N2 bei 23 von 32 und für N3 bei 12 von 19 Probanden.

### Simultanes Testen

Für ein festes Schlafstadium  $Z$  wurde bisher für jeden Probanden  $J$  die zugehörige Nullhypothese  $H_0^{(J,Z)}$  einzeln auf einem Signifikanzniveau von  $5\%$  überprüft. Es wird nun die Nullhypothese  $H_0^{(Z)}$  formuliert, dass für alle Probanden  $J$  die Nullhypothese  $H_0^{(J,Z)}$  richtig ist:

**Nullhypothese  $H_0^{(Z)}$ :** Die beobachteten Übergangshäufigkeiten  $(n_{rs}^{(J,Z)})_{r,s}$  sind für alle Probanden  $J \in \mathcal{J}_Z$  so verteilt wie die Übergangshäufigkeiten einer Markov-Kette der Länge  $n^{(J,Z)} + 1$  mit Zustandsraum  $S = \{A, B, C, D\}$ , die die Bedingungen des SMI-Modells erfüllt.

Um  $H_0^{(Z)}$  nur mit einer Wahrscheinlichkeit von höchstens  $5\%$  irrtümlich abzulehnen, wird die obere Grenze (0.05) für die p-Werte, die zur Ablehnung einer der Nullhypothesen  $H_0^{(J,Z)}$  führen, entsprechend nach unten angepasst. Durch die Bonferroni-Korrektur wird diese Grenze konkret von 0.05 auf  $0.05/32$  für W, N1 und N2 bzw. auf  $0.05/19$  für N3 korrigiert, da es für N3 19 Probanden und somit 19 Einzel-Nullhypothesen gibt und für die anderen Schlafstadien entsprechend 32 (für Bonferroni-Korrektur siehe z. B. Wasserman, 2004).

In Abbildung 6.3 geben die violetten Quadrate an, dass die zum entsprechenden Schlafstadium  $Z$  gehörende Bonferroni-korrigierte Grenze vom p-Wert unterschritten wird und daher  $H_0^{(Z)}$  abgelehnt wird auf einem Signifikanzniveau von  $5\%$ . Wie in der Abbildung außerdem zu erkennen ist, gibt es für jedes Schlafstadium mindestens sechs Probanden, deren p-Werte kleiner sind als die jeweilige Bonferroni-korrigierte Grenze  $0.05/32$  bzw.  $0.05/19$ , sodass die Hypothesen  $H_0^{(W)}$ ,  $H_0^{(N1)}$ ,  $H_0^{(N2)}$  und  $H_0^{(N3)}$  alle verworfen werden.

Die Nullhypothese  $H_0^{(Z)}$  lässt sich für jedes Schlafstadium  $Z$  auch untersuchen über die Teststatistik

$$\mathcal{X}_{(Z)}^2 := \sum_{J \in \mathcal{J}_Z} \mathcal{X}_{(J,Z)}^2.$$

Da die Teststatistiken  $\mathcal{X}_{(J,Z)}^2$  unabhängig und asymptotisch  $\chi^2$ -verteilt sind, folgt direkt, dass  $\mathcal{X}_{(Z)}^2$  ebenfalls asymptotisch  $\chi^2$ -verteilt ist.

Soll die Nullhypothese wieder mittels Simulation überprüft werden, kann man auf die bereits ermittelten und unabhängigen  $\chi^2_{[i],(J,Z)}$  zurückgreifen durch

$$\chi^2_{[i],(Z)} := \sum_{J \in \mathcal{J}_Z} \chi^2_{[i],(J,Z)} \quad \forall i \in \{1, \dots, 10000\}.$$

Die daraus resultierenden empirischen Dichten werden in Abbildung 6.4 dargestellt. Die empirische Dichtefunktion in N3 nimmt eine andere Form an und liegt weiter links als die empirischen Dichtefunktionen in den anderen drei Schlafstadien, da für N3 nur die Daten von 19 und nicht von 32 Probanden zur Verfügung stehen.

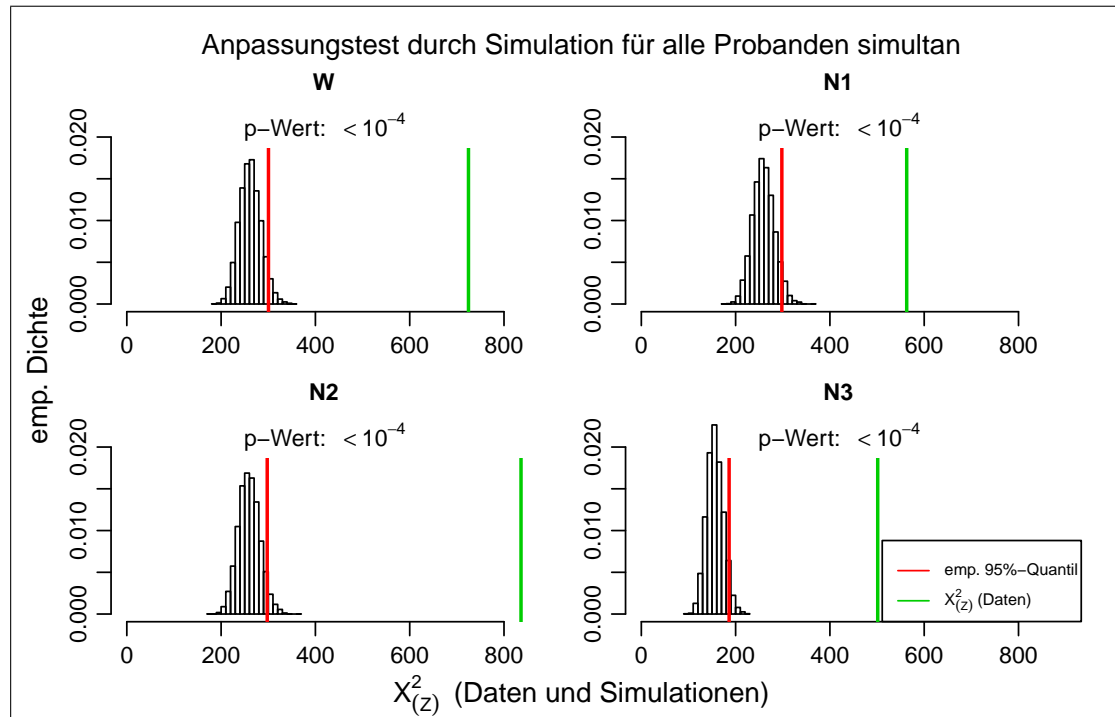


Abbildung 6.4: Anpassungstests durch Simulation für alle Probanden simultan, alle Schlafstadien.

Durch die Simulationen wurde für jede der vier Nullhypothesen  $H_0^{(Z)}$  ein p-Wert  $< 10^{-4}$  bestimmt, sodass die Nullhypothesen  $H_0^{(Z)}$  auch durch diesen Test verworfen werden.

Insgesamt zeigen die Daten in jedem Schlafstadium signifikante Abweichungen von den Modellannahmen des SMI-Modells. Lediglich für jeweils weniger als die Hälfte aller Probanden in den Schlafstadien W, N2 und N3 werden die Nullhypothesen  $H_0^{(J,Z)}$  auf einem Signifikanzniveau von 5% nicht verworfen sowie bei weniger als 60% aller Probanden in Schlafstadium N1.

### 6.1.4 Überprüfung des Fehlers 1. Art für die Durchführung des Anpassungstests

In Kapitel 6.1.2 wurde die Durchführung der Hypothesentests für die Nullhypothesen  $H_0^{(J,Z)}$  angegeben. Bei dieser Durchführung werden für einen festen Probanden  $J$  und ein festes Schlafstadium  $Z$  Markov-Ketten gemäß dem SMI-Modell simuliert, um eine empirische Verteilung der  $\chi^2_{(J,Z)}$ -Statistik und somit einen entsprechenden empirischen p-Wert zu ermitteln. Als Parameter für das SMI-Modell werden dabei die SMI-Modell-Schätzer  $\hat{p}_1^{(J,Z)}, \hat{p}_2^{(J,Z)}, \hat{p}_3^{(J,Z)}, \hat{p}_4^{(J,Z)}$  und  $\hat{\pi}^{(J,Z)}$  verwendet. Die Nullhypothese  $H_0^{(J,Z)}$  wird nach der in Kapitel 6.1.2 beschriebenen Durchführung verworfen, wenn der aus den simulierten Markov-Ketten ermittelte empirische p-Wert  $p^{(J,Z)} < 0.05$  ist. Es ist jedoch nicht klar, dass dieser p-Wert unter der Annahme, dass die Nullhypothese  $H_0^{(J,Z)}$  wahr ist, nur mit einer Wahrscheinlichkeit von 5% kleiner als 0.05 ist (Fehler 1. Art ist nicht bekannt).

Daher werden nun dreimal jeweils 1000 Markov-Ketten

$$X^{\{k,1\}}, \dots, X^{\{k,1000\}} \quad , k \in \{1, 2, 3\},$$

der Länge 5001 gemäß dem SMI-Modell mit zufälligen Parametern

$$p_1^{\{k\}}, p_2^{\{k\}}, p_3^{\{k\}}, p_4^{\{k\}} \quad \text{und} \quad \pi^{\{k\}}$$

simuliert und für jedes  $k \in \{1, 2, 3\}$  wird getestet, ob die jeweils 1000 p-Werte

$$p^{\{k,1\}}, \dots, p^{\{k,1000\}},$$

die zu der nachstehenden Nullhypothese  $H_0^{\{k,l\}}$  durch die in Kapitel 6.1.2 beschriebene Durchführung für alle  $l \in \{1, \dots, 1000\}$  ermittelt werden, uniform auf  $[0, 1]$  verteilt sind.

**Nullhypothese  $H_0^{\{k,l\}}$ :** Die Übergangshäufigkeiten  $n_{rs}^{\{k,l\}}$  der Markov-Kette  $X^{\{k,l\}}$  sind so verteilt wie die Übergangshäufigkeiten einer Markov-Kette der Länge 5001 mit Zustandsraum  $S = \{A, B, C, D\}$ , die die Bedingungen des SMI-Modells erfüllt.

Die Parameter  $p_r^{\{k,l\}}$  und  $\pi^{\{k,l\}}$  werden für die Durchführung der Anpassungstests geschätzt durch die Schätzer  $\hat{p}_r^{\{k,l\}}$  und  $\hat{\pi}^{\{k,l\}}$ , die nach Gleichung (4.1) bzw. (4.2) berechnet werden.

Für den Test der Nullhypothese  $H_0^{\{k,l\}}$  wird für festes  $l \in \{1, \dots, 1000\}$  analog zu der in Kapitel 6.1.2 beschriebenen Durchführung die empirische Verteilung der Teststatistik

$$\chi^2_{\{k,l\}} = \sum_{r,s=1}^4 \frac{\left( n_{rs}^{\{k,l\}} - n_r^{\{k,l\}} \hat{p}_{rs}^{\{k,l\}} \right)^2}{n_r^{\{k,l\}} \hat{p}_{rs}^{\{k,l\}}}$$

ermittelt mithilfe weiterer 1000 Markov-Ketten

$$X^{[1],\{k,l\}}, \dots, X^{[1000],\{k,l\}},$$

die gemäß dem SMI-Modell mit Parametern  $\hat{p}_1^{\{k,l\}}, \hat{p}_2^{\{k,l\}}, \hat{p}_3^{\{k,l\}}, \hat{p}_4^{\{k,l\}}$  und  $\hat{\pi}^{\{k,l\}}$  simuliert werden.<sup>19</sup>

In Abbildung 6.5 ist für jedes  $k \in \{1, 2, 3\}$  das Histogramm der p-Werte  $p^{\{k,l\}}$  dargestellt, das jeweils vermuten lässt, dass die p-Werte uniform-verteilt sind.

Prüfen lässt sich die Nullhypothese  $H_0^{\{k\}}$ , dass die p-Werte auf  $[0, 1]$  uniform-verteilt sind, mit Fishers Kombinationstest (siehe z. B. Wolf, 1986). Bei diesem Test wird die Teststatistik

$$-2 \sum_{l=1}^{1000} \log(p^{\{k,l\}})$$

betrachtet, die unter der Nullhypothese  $H_0^{\{k\}}$  gerade  $\chi^2$ -verteilt ist mit  $2 \cdot 1000$  Freiheitsgraden.

Die zu  $H_0^{\{k\}}$  gehörenden p-Werte  $p^{\{k\}}$  lauten für die simulierten Datensätze

$$p^{\{1\}} \approx 0.48, \quad p^{\{2\}} \approx 0.13, \quad p^{\{3\}} \approx 0.91.$$

Die Nullhypothesen  $H_0^{\{k\}}$  werden also auf einem Signifikanzniveau von 5% nicht verworfen.

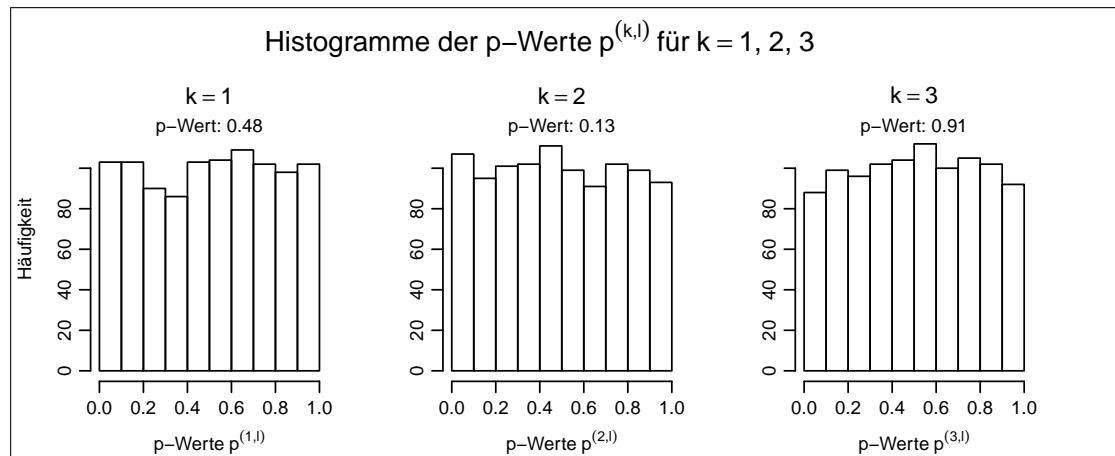


Abbildung 6.5: Histogramme der p-Werte  $p^{\{k,l\}}$  getrennt nach  $k = 1, 2, 3$ . Die angegebenen p-Werte beziehen sich auf Fishers Kombinationstest.

Es wird daher davon ausgegangen, dass der  $\alpha$ -Fehler 5% beträgt, falls eine der Nullhypothesen  $H_0^{(J,Z)}$  verworfen wird, wenn der zugehörige p-Wert  $p^{(J,Z)} < 0.05$  ist.

<sup>19</sup>Aus Laufzeitgründen werden entgegen der in Kapitel 6.1.2 beschriebenen Durchführung nur 1000 und nicht 10000 Markov-Ketten simuliert für die Ermittlung der empirischen Verteilung von  $\chi_{\{k,l\}}^2$ .

### 6.1.5 Gründe für die Ablehnung des SMI-Modells

Die in Kapitel 6.1.3 angewendeten Anpassungstests haben zwar als Ergebnis, dass das SMI-Modell zu verwerfen ist, sie liefern jedoch keinen konkreten Hinweis darauf, weshalb das SMI-Modell signifikant von den Daten abweicht.

Ein naheliegender Erklärungsansatz ist, dass sich die spezielle Struktur der Übergangswahrscheinlichkeiten des SMI-Modells nicht in erforderlichem Maße mit den gegebenen Daten verträgt. Denn ein großer Wert von  $\chi^2_{(J,Z)}$ , der zu einer Ablehnung der Nullhypothese führt, deutet darauf hin, dass mindestens für ein Paar  $(r, s) \in S \times S$  die beobachtete Übergangshäufigkeit  $n_{rs}$  stark von dem durch das SMI-Modell erwarteten Wert  $n_r \hat{p}_{rs}$  abweicht. Je größer diese Abweichung ist, desto unwahrscheinlicher ist es, dass sie durch Zufall beobachtet worden ist, unter der Annahme, dass das SMI-Modell richtig ist.

Es können in diesem Zusammenhang also zum einen Fälle auftreten, bei denen eine (und damit mindestens eine zweite) Übergangswahrscheinlichkeit abgeändert werden müsste und ansonsten die durch das SMI-Modell vorgegebene Struktur der Übergangswahrscheinlichkeiten richtig ist. Zum anderen ist es möglich, dass die gesamte Struktur verworfen werden muss.

Darüber hinaus gibt es beispielsweise die Möglichkeiten, dass die beobachtete Sequenz  $(X_i)_{i \in \{0, \dots, n\}}$  einzelne oder alle Bedingungen an eine homogene Markov-Kette 1. Ordnung nicht erfüllt.

Obwohl die Daten durch die Anpassungstests für die Mehrheit der Probanden und beim simultanen Testen signifikante Abweichungen von den Modellannahmen des SMI-Modells zeigen, kann das SMI-Modell trotzdem nützlich sein und einen Erkenntnisgewinn bringen, wenn man der oben beschriebenen Frage nachgeht, inwiefern das SMI-Modell die Struktur der Übergangswahrscheinlichkeiten des vollen Modells nicht richtig beschreibt. Dabei ist auch von Interesse, ob die Abweichung der gegebenen Daten vom SMI-Modell für jeden Probanden bzw. für jedes Schlafstadium eine andere Form annimmt oder konsistent ist.

Dieser Fragestellung wird in Kapitel 6.2 nachgegangen.

## 6.2 Identifizierung der Abweichungen des SMI-Modells von den Daten

Im Folgenden soll die Frage geklärt werden, bei welchen Übergängen  $(r, s)$  das SMI-Modell am stärksten von den gegebenen Daten abweicht.

Dazu wird die Formel für die im SMI-Modell geschätzten Übergangswahrscheinlichkeiten

$$\hat{p}_{rs} = \hat{\pi} \hat{p}_s + \delta_{r,s}(1 - \hat{\pi})$$

nach  $\hat{\pi}$  aufgelöst, sodass man

$$\hat{\pi} = \frac{\hat{p}_{rs}}{\hat{p}_s} = \frac{1 - \hat{p}_{rr}}{1 - \hat{p}_r} \quad \forall r, s \in S, r \neq s, \quad (6.2)$$

erhält.

Wenn die Nullhypothese  $H_0^{(J,Z)}$  wahr ist, dass die gegebenen Übergangshäufigkeiten  $n_{rs}^{(J,Z)}$  eines Probanden  $J$  in einem Schlafstadium  $Z$  einer Markov-Kette der Länge  $n^{(J,Z)} + 1$  entstammen, die die Bedingungen des SMI-Modells erfüllt, dann erwartet man, dass die Schätzer  $\hat{p}_{rs}$  des SMI-Modells und die Schätzer  $\hat{p}_{rs}^{(v)}$  des vollen Modells für alle  $r, s \in S$  in den meisten Fällen relativ dicht beieinander liegen, denn es gilt mit der Delta-Methode<sup>20</sup>

$$\mathbb{E}[\hat{p}_{rs}] = \mathbb{E}[\hat{\pi}\hat{p}_s + \delta_{r,s}(1 - \hat{\pi})] \approx \pi p_s + \delta_{r,s}(1 - \pi) = p_{rs} = \frac{np_r p_{rs}}{np_r} \approx \mathbb{E}[\hat{p}_{rs}^{(v)}].$$

Ersetzt man für  $r, s \in S$  in Gleichung (6.2) die Schätzer  $\hat{p}_{rs}$  bzw.  $\hat{p}_{rr}$  des SMI-Modells durch die Schätzer  $\hat{p}_{rs}^{(v)}$  bzw.  $\hat{p}_{rr}^{(v)}$  des vollen Modells, dann erwartet man, dass sich die entsprechenden resultierenden Werte von

$$\hat{\pi}_{rs} := \frac{\hat{p}_{rs}^{(v)}}{\hat{p}_s} \quad \text{und} \quad \hat{\pi}_{rr} := \frac{1 - \hat{p}_{rr}^{(v)}}{1 - \hat{p}_r}$$

nicht stark von  $\hat{\pi}$  unterscheiden.

Im Folgenden werden für alle Probanden  $J$  und Schlafstadien  $Z$  die Differenzen

$$\hat{\pi}_{rs}^{(J,Z)} - \hat{\pi}^{(J,Z)} \quad \forall r, s \in S$$

dahin gehend untersucht, ob ihre Werte unter der Nullhypothese  $H_0^{(J,Z)}$  signifikant von der 0 abweichen bei einem Signifikanzniveau von 5%. Dazu werden für die Differenzen 95%-Konfidenzintervalle mithilfe von simulierten Markov-Ketten approximiert, sodass darauf geschlossen werden kann, für welche Übergänge  $(r, s)$  die Werte von  $\hat{\pi}^{(J,Z)}$  und  $\hat{\pi}_{rs}^{(J,Z)}$  signifikant voneinander abweichen für einen Probanden  $J$  in einem Schlafstadium  $Z$ .

In Abbildung 6.6 sind die Differenzen  $\hat{\pi}_{rs}^{(J,Z)} - \hat{\pi}^{(J,Z)}$  in studentisierter Form (standardisiert mit geschätztem Mittelwert und geschätzter Standardabweichung) für alle Probanden  $J$  und Schlafstadien  $Z$  eingetragen. Unter anderem zeigt sich in dieser Abbildung, dass im Schlafstadium N2 bei den Übergängen  $(r, s) \in \{(C, D), (D, C)\}$  die Abweichungen zwischen  $\hat{\pi}_{rs}^{(J,N2)}$  und  $\hat{\pi}^{(J,N2)}$  für mehr als die Hälfte aller Probanden so groß ist, dass sich die Differenzen  $\hat{\pi}_{rs}^{(J,N2)} - \hat{\pi}^{(J,N2)}$  (studentisiert) außerhalb ihres approximativen 95%-Konfidenzintervalls befinden. Dies könnte, wie in Kapitel 6.2.1 näher ausgeführt wird, ein Hinweis darauf sein, dass die Übergangswahrscheinlichkeiten für die Übergänge  $(C, D)$  und  $(D, C)$  im Schlafstadium N2 systematisch nicht richtig erfasst werden durch das SMI-Modell.

Genauere theoretische Vorüberlegungen zu den Differenzen  $\hat{\pi}_{rs} - \hat{\pi}$  werden in Kapitel 6.2.1 erläutert. Die Beschreibung der Durchführung der Ermittlung der approximativen 95%-Konfidenzintervalle mithilfe von Simulationen erfolgt Schritt für Schritt in Kapitel 6.2.2. Zum Abschluss werden die Ergebnisse in Kapitel 6.2.3 diskutiert.

<sup>20</sup>Die Delta-Methode wird in Anhang B beschrieben.

## 6.2.1 Vorüberlegungen

Das SMI-Modell ist dadurch charakterisiert, dass die Wahrscheinlichkeit dafür, dass bei einem Übergang von  $X_i$  zu  $X_{i+1}$  in der Markov-Kette  $(X_i)_{i \in \mathbb{N}_0}$  im Hintergrundmodell mindestens ein Intervallwechsel auftritt, unabhängig von der Zeit und dem Ausgangszustand  $X_i$  konstant  $\pi$  beträgt. Aus der Definition der Übergangswahrscheinlichkeiten

$$p_{rs} = \pi p_s + \delta_{r,s}(1 - \pi)$$

folgen direkt

$$\pi = \frac{p_{rs}}{p_s} \quad \forall r, s \in S, r \neq s,$$

und

$$\pi = \frac{1 - p_{rr}}{1 - p_r} \quad \forall r \in S.$$

Für die Schätzer  $\hat{p}_1, \hat{p}_2, \hat{p}_3, \hat{p}_4$  und  $\hat{\pi}$  gilt analog

$$\hat{\pi} = \frac{\hat{p}_{rs}}{\hat{p}_s} = \frac{1 - \hat{p}_{rr}}{1 - \hat{p}_r} \quad \forall r, s \in S, r \neq s,$$

wobei  $\hat{\pi} = \hat{\pi}^{(g)}$  ist, also bei der 1 abgeschnitten wird.

Ersetzt man  $\hat{p}_{rs}$  im Zähler für  $r \neq s$  durch den Schätzer  $p_{rs}^{(v)} = n_{rs}/n_r$  des vollen Modells, werde der dadurch entstehende Term bezeichnet mit

$$\hat{\pi}_{rs} := \frac{\hat{p}_{rs}^{(v)}}{\hat{p}_s} = \frac{n_{rs}}{n_r} \cdot \frac{n}{n_s}. \quad (6.3)$$

Analog sei

$$\hat{\pi}_{rr} := \frac{1 - \hat{p}_{rr}^{(v)}}{1 - \hat{p}_r} = \frac{n_r - n_{rr}}{n_r} \cdot \frac{n}{n - n_r}, \quad (6.4)$$

wobei hier  $\hat{p}_{rr}$  durch  $\hat{p}_{rr}^{(v)} = n_{rr}/n_r$  ersetzt wird.

Die Werte von  $\hat{\pi}_{rs}$  bzw.  $\hat{\pi}_{rr}$  werden hier als Verhältnisse zweier Zahlen und nicht als Wahrscheinlichkeiten aufgefasst, sodass  $\hat{\pi}_{rs}$  und  $\hat{\pi}_{rr}$  nicht bei 1 abgeschnitten werden.

Unter der Annahme, dass die gegebenen Daten  $n_{rs} = n_{rs}^{(J,Z)}$  eines Probanden  $J$  im Schlafstadium  $Z$  dem SMI-Modell mit Parametern  $\hat{p}_1, \hat{p}_2, \hat{p}_3, \hat{p}_4$  und  $\hat{\pi}$  entstammen, ergibt sich für den Erwartungswert von  $\hat{\pi}_{rs}$ ,  $r \neq s$ , mithilfe der Delta-Methode

$$\mathbb{E}[\hat{\pi}_{rs}] \approx \frac{n\hat{p}_r\hat{p}_{rs}}{n\hat{p}_r} \cdot \frac{n}{n\hat{p}_s} = \frac{\hat{p}_{rs}}{\hat{p}_s} = \hat{\pi}$$

und für den Erwartungswert von  $\hat{\pi}_{rr}$  entsprechend

$$\mathbb{E}[\hat{\pi}_{rr}] \approx \frac{n\hat{p}_r - n\hat{p}_r\hat{p}_{rr}}{n\hat{p}_r} \cdot \frac{n}{n - n\hat{p}_r} = \frac{1 - \hat{p}_{rr}}{1 - \hat{p}_r} = \hat{\pi}.$$



Der Erwartungswert der Differenzen  $\hat{\pi}_{rs} - \hat{\pi}$  und  $\hat{\pi}_{rr} - \hat{\pi}$  beträgt also ungefähr 0.

Ziel ist es, 95%-Konfidenzintervalle für die Differenzen  $\hat{\pi}_{rs} - \hat{\pi}$  bzw.  $\hat{\pi}_{rr} - \hat{\pi}$  für einen festen Probanden  $J$  in einem festen Schlafstadium  $Z$  mithilfe von simulierten Markov-Ketten zu approximieren, unter der oben getroffenen Annahme, dass das SMI-Modell die Verteilung der Übergangshäufigkeiten beschreibt. Die simulierten Markov-Ketten werden dabei verwendet, um den Erwartungswert von  $\hat{\pi}_{rs} - \hat{\pi}$ , der nur ungefähr 0 beträgt, und die Standardabweichung dieser Differenz zu schätzen, sodass sich das 95%-Konfidenzintervall für die studentisierte Form von  $\hat{\pi}_{rs} - \hat{\pi}$  durch das Intervall  $[-2, 2]$  approximieren lässt (durch Simulationen hat sich gezeigt, dass die Differenzen  $\hat{\pi}_{rs} - \hat{\pi}$  approximativ normalverteilt sind). Liegt eine der Differenzen (in studentisierter Form) nicht in ihrem approximierten 95%-Konfidenzintervall, weist das darauf hin, dass sich  $\hat{\pi}_{rs}$  und  $\hat{\pi}$  bzw.  $\hat{\pi}_{rr}$  und  $\hat{\pi}$  signifikant voneinander unterscheiden.

Es wird nun erörtert, wie sich signifikante Unterschiede zwischen  $\hat{\pi}_{rs}$  und  $\hat{\pi}$  bzw. zwischen  $\hat{\pi}_{rr}$  und  $\hat{\pi}$  deuten lassen. Dafür wird zur besseren Lesbarkeit der wahrscheinliche Fall angenommen, dass sich die 0 im 95%-Konfidenzintervall von  $\hat{\pi}_{rs} - \hat{\pi}$  bzw. von  $\hat{\pi}_{rr} - \hat{\pi}$  befindet.

Zunächst wird der Fall  $r \neq s$  betrachtet, für den

$$\hat{\pi}_{rs} - \hat{\pi} = \frac{\hat{p}_{rs}^{(v)} - \hat{p}_{rs}}{\hat{p}_s} \quad (6.5)$$

gilt. Wenn die 0 im 95%-Konfidenzintervall liegt und die Differenz  $\hat{\pi}_{rs} - \hat{\pi}$  größer (bzw. kleiner) als die obere (bzw. untere) Grenze des 95%-Konfidenzintervalls ist, dann folgt nach Gleichung (6.5)

$$\hat{p}_{rs}^{(v)} > \hat{p}_{rs} \quad (\text{bzw. } \hat{p}_{rs}^{(v)} < \hat{p}_{rs}).$$

Die Schätzer  $\hat{p}_{rs}^{(v)}$  und  $\hat{p}_{rs}$  weichen aber nicht notwendigerweise signifikant voneinander ab, da auch der Nenner  $\hat{p}_s$  in Gleichung (6.5) den Wert der Differenz  $\hat{\pi}_{rs} - \hat{\pi}$  beeinflusst.

Das Konfidenzintervall für  $\hat{\pi}_{rs} - \hat{\pi}$  wird ermittelt unter der Annahme, dass die in den Daten gegebenen Übergangshäufigkeiten  $n_{rs}$  verteilt sind wie die Übergangshäufigkeiten einer Markov-Kette gemäß dem SMI-Modell, wobei für die Simulationen zur Ermittlung des 95%-Konfidenzintervalls die Parameter  $\hat{p}_1, \dots, \hat{p}_a$  und  $\hat{\pi}$  verwendet werden. Unter dieser Annahme gilt für den Erwartungswert des Schätzer  $\hat{p}_s^{[i]}$  für  $p_s$  aus den simulierten Daten

$$\mathbb{E} \left[ \hat{p}_s^{[i]} \right] = \mathbb{E} \left[ n_s^{[i]} / n \right] = \hat{p}_s.$$

Ist die Nullhypothese nicht richtig und wird als Alternativhypothese das volle Modell mit Parametern  $(\hat{p}_{tu}^{(v)})_{t,u \in S}$  als richtig angenommen, gilt auch in diesem Fall

$$\mathbb{E} \left[ \hat{p}_s^{[i]} \right] = \mathbb{E} \left[ n_s^{[i]} / n \right] = \hat{p}_s,$$

denn  $n_s^{[i]}/n$  ist auch im vollen Modell der Schätzer für  $\hat{p}_s$  und erwartungstreu. Daher wird der Einfluss des Nenners  $\hat{p}_s$  der Gleichung (6.5) darauf, ob die Differenz  $\hat{\pi}_{rs} - \hat{\pi}$  außerhalb ihres 95%-Konfidenzintervalls liegt, als gering erachtet.

Für den Fall  $r = s$  ergibt sich

$$\hat{\pi}_{rr} - \hat{\pi} = \frac{(1 - \hat{p}_{rr}^{(v)}) - (1 - \hat{p}_{rr})}{1 - \hat{p}_r} = \frac{\hat{p}_{rr} - \hat{p}_{rr}^{(v)}}{1 - \hat{p}_r} = \frac{\sum_{t \neq r} \hat{p}_{rt}^{(v)} - \sum_{t \neq r} \hat{p}_{rt}}{\sum_{t \neq r} \hat{p}_t}. \quad (6.6)$$

Wenn die 0 im 95%-Konfidenzintervall liegt und die Differenz  $\hat{\pi}_{rr} - \hat{\pi}$  größer (bzw. kleiner) als die obere (bzw. untere) Grenze des 95%-Konfidenzintervalls ist, dann folgt nach Gleichung (6.6)

$$\hat{p}_{rr}^{(v)} < \hat{p}_{rr} \quad (\text{bzw. } \hat{p}_{rr}^{(v)} > \hat{p}_{rr}),$$

gerade andersherum als im Fall  $r \neq s$ .

Die Schätzer  $\hat{p}_{rr}^{(v)}$  und  $\hat{p}_{rr}$  weichen wie im Fall  $r \neq s$  nicht notwendigerweise signifikant voneinander ab, da auch der Nenner  $1 - \hat{p}_r$  in Gleichung (6.6) den Wert der Differenz  $\hat{\pi}_{rr} - \hat{\pi}$  beeinflusst.

Analog zum Fall  $r \neq s$  wird der Einfluss des Nenners  $1 - \hat{p}_r$  darauf, ob die Differenz  $\hat{\pi}_{rr} - \hat{\pi}$  außerhalb oder innerhalb ihres 95%-Konfidenzintervalls liegt, als gering erachtet.

**Bemerkung 6.5.** Anhand der letzten Umformung in Gleichung (6.6) lässt sich erkennen, wie die Differenz  $\hat{\pi}_{rr} - \hat{\pi}$  für ein  $r \in S$  abhängt von

$$\hat{p}_{rs}^{(v)} - \hat{p}_{rs} \quad \forall s \neq r.$$

Beobachtet man für ein  $s \neq r$  oder für mehrere  $s \neq r$ , dass die Differenz  $\hat{\pi}_{rs} - \hat{\pi}$  außerhalb ihres 95%-Konfidenzintervalls liegt, dann wird in den meisten Fällen auch  $\hat{p}_{rs}^{(v)} - \hat{p}_{rs}$  betragsweise groß sein für die entsprechenden  $s \neq r$ , sodass tendenziell davon ausgegangen werden kann, dass auch  $\hat{\pi}_{rr} - \hat{\pi}$  im Absolutbetrag große Werte annimmt, die ggf. außerhalb des 95%-Konfidenzintervalls liegen.

## 6.2.2 Durchführung der Ermittlung des 95%-Konfidenzintervalls durch Simulationen

Nach den theoretischen Vorüberlegungen wird nun das genaue Vorgehen für die Simulationen für einen Probanden  $J$  im Schlafstadium  $Z$  beschrieben, mit deren Hilfe die 95%-Konfidenzintervalle für  $\hat{\pi}_{rs} - \hat{\pi}$  approximativ bestimmt werden.

Gegeben sind die Daten  $n_{rs} = n_{rs}^{(J,Z)}$  für alle  $r, s \in S = \{1, 2, 3, 4\}$ .

1. Berechne  $n = \sum_{r,s \in S} n_{rs}$  und  $n_t = \sum_{u \in S} n_{tu}$  für alle  $t \in S$ .
2. Berechne  $\hat{\pi}_{tu}$  für alle  $t, u \in S$  aus den  $n_{rs}$  gemäß den Gleichungen (6.3) und (6.4).
3. Berechne die Schätzer  $\hat{p}_1, \hat{p}_2, \hat{p}_3, \hat{p}_4$  gemäß Gleichung 4.1 und  $\hat{\pi}$  gemäß Gleichung 4.2 aus den Übergangshäufigkeiten  $n_{rs}$ .

4. Berechne die geschätzten Übergangswahrscheinlichkeiten  $\hat{p}_{rs} = \hat{\pi}\hat{p}_s + \delta_{r,s}(1 - \hat{\pi})$ .
5. Berechne die Differenz  $D_{tu} := \hat{\pi}_{tu} - \hat{\pi}$  für alle  $t, u \in S$ .
6. Simuliere 10000 Markov-Ketten  $X^{[i]} = (X_0^{[i]}, \dots, X_n^{[i]})$  der Länge  $n + 1$  mit Startverteilung  $(\hat{p}_1, \hat{p}_2, \hat{p}_3, \hat{p}_4)$  und Übergangsmatrix  $\hat{P} = (\hat{p}_{rs})_{r,s \in S}$  und führe die Schritte 7 bis 10 für jedes  $X^{[i]}$  aus,  $i = 1, \dots, 10000$ .
7. Bestimme  $n_{rs}^{[i]}$  für alle  $r, s \in S$ .
8. Berechne  $\hat{\pi}_{tu}^{[i]}$  für alle  $t, u \in S$  aus den  $n_{rs}^{[i]}$ .
9. Berechne den Schätzer  $\hat{\pi}^{[i]}$  aus den  $n_{rs}^{[i]}$ .
10. Berechne die Differenz  $D_{tu}^{[i]} = \hat{\pi}_{tu}^{[i]} - \hat{\pi}^{[i]}$  für alle  $t, u \in S$ .
11. Berechne für alle  $t, u \in S$  die empirische Standardabweichung  $\hat{s}_{tu}$  der Differenzen  $D_{tu}^{[i]}$  über alle  $i \in \{1, \dots, 10000\}$  durch

$$\hat{s}_{tu} := \left( \frac{1}{10000 - 1} \sum_{i=1}^{10000} \left( D_{tu}^{[i]} - \frac{1}{10000} \sum_{i=1}^{10000} D_{tu}^{[i]} \right)^2 \right)^{1/2}.$$

12. Bestimme  $\Delta_{tu}^{(J,Z)} = \Delta_{tu} := (\hat{\pi}_{tu} - \hat{\pi}) / \hat{s}_{tu}$  für alle  $t, u \in S$ .

Ist für  $t, u \in S$  der Absolutbetrag  $|\Delta_{tu}^{(J,Z)}| > 2$ , deutet dies auf Abweichungen von der Nullhypothese hin mit den in Kapitel 6.2.1 diskutierten Folgerungen.

### 6.2.3 Ergebnisse

In Abbildung 6.6 sind die berechneten  $\Delta_{tu}^{(J,Z)}$  für alle Probanden  $J$  nach Schlafstadium  $Z$  getrennt eingetragen. Zu jedem Übergang  $(r, s)$  wird in der Abbildung angegeben, wie viele Probanden außerhalb des approximativen 95%-Konfidenzintervalls liegen.

Besonders auffällig in der Abbildung 6.6 sind die starken Ausschläge nach oben von  $\Delta_{CD}^{(J,N2)}$  und  $\Delta_{DC}^{(J,N2)}$  für das Schlafstadium N2 bei jeweils 20 Probanden. Auch für die anderen Schlafstadien  $Z \in \mathcal{Z} \setminus \{N2\}$  stechen jeweils zwei bis drei Übergänge  $(r, s)$  hervor, die im jeweiligen Schlafstadium die größte Anzahl von Probanden  $J$  aufweisen, für die  $\Delta_{rs}^{(J,Z)} > 2$  ist, sodass  $\Delta_{rs}^{(J,Z)}$  oberhalb seiner oberen 95%-Konfidenzintervallgrenze liegt.

Zusammenfassend lassen sich in den verschiedenen Schlafstadien folgende Übergänge  $(r, s)$  hervorheben, für die  $\Delta_{rs}^{(J,Z)} > 2$  bei möglichst vielen Probanden gilt (die jeweiligen Zustände  $r$  und  $s$  werden für jedes Schlafstadium  $Z$  im weiteren Verlauf der Arbeit in alphabetischer Ordnung mit  $z_1^{(Z)}$  bzw.  $z_2^{(Z)}$  bezeichnet):

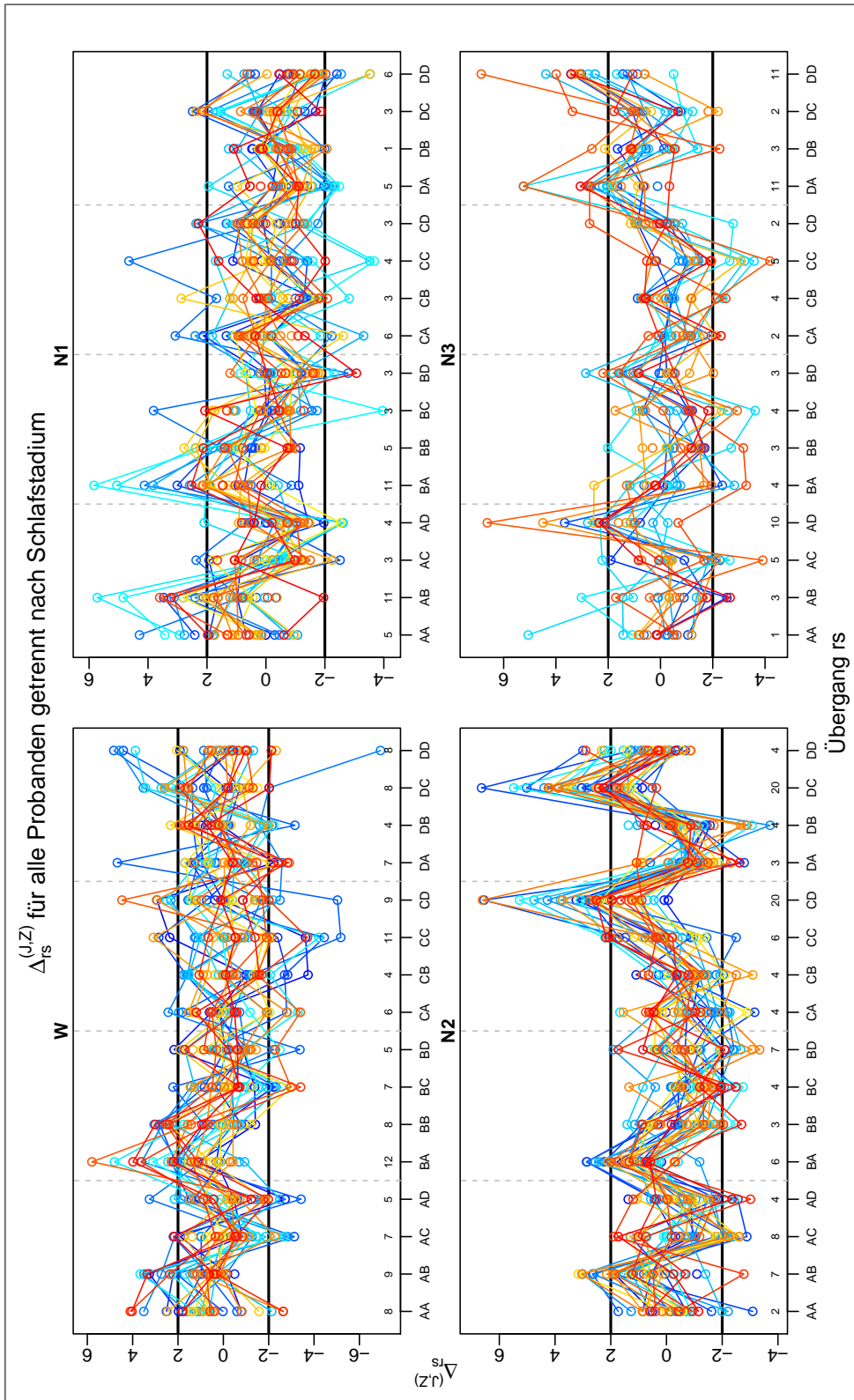


Abbildung 6.6: Übersicht über alle  $\Delta_{rs}^{(J,Z)}$  für jeden Übergang  $(r, s)$  aller Probanden  $J$  getrennt nach Schlafstadium  $Z$ . Für einen festen Probanden  $J$  in einem festen Schlafstadium  $Z$  sind die  $\Delta_{rs}^{(J,Z)}$  als Punkte markiert und durch Linien verbunden. Die horizontalen schwarzen Linien markieren die approximativen 95%-Konfidenzintervalle  $[-2, 2]$  der  $\Delta_{rs}^{(J,Z)}$ . Für jeden Probanden sind die  $\Delta_{rs}$  als Punkte markiert und durch Linien verbunden. Die horizontalen roten Linien markieren das approximative 95%-Konfidenzintervall.

W:  $(A, B)$  und  $(B, A)$  (bei 9 und 12 Probanden)<sup>21</sup>, d.h.

$$z_1^{(W)} := A, \quad z_2^{(W)} := B.$$

N1:  $(A, B)$  und  $(B, A)$  (bei jeweils 11 Probanden), d.h.

$$z_1^{(N1)} := A, \quad z_2^{(N1)} := B.$$

N2:  $(C, D)$  und  $(D, C)$  (bei jeweils 20 Probanden), d.h.

$$z_1^{(N2)} := C, \quad z_2^{(N2)} := D.$$

N3:  $(A, D)$ ,  $(D, A)$  und  $(D, D)$  (bei 10, 11 und 11 Probanden), d.h.

$$z_1^{(N3)} := A, \quad z_2^{(N3)} := D.$$

Für jedes Schlafstadium  $Z$  treten die meisten  $\Delta_{rs}^{(J,Z)} > 2$  für die Übergänge

$$(r, s) \in \left\{ (z_1^{(Z)}, z_2^{(Z)}), (z_2^{(Z)}, z_1^{(Z)}) \right\}$$

auf, die einander entgegengesetzt sind, wobei der Übergang  $(D, D)$  in N3 einen Spezialfall darstellt.

Wegen der in Bemerkung 6.5 beschriebenen Abhängigkeit von  $\hat{\pi}_{DD} - \hat{\pi}$  und  $\hat{\pi}_{DA} - \hat{\pi}$  hängt auch  $\Delta_{DD}$  von  $\Delta_{DA}$  ab. Daher wird der Übergang  $(D, D)$  im Schlafstadium N3 im Folgenden nicht mituntersucht.

Liegt  $\Delta_{rs}$  außerhalb seines 95%-Konfidenzintervalls für einen Übergang  $(r, s)$ , unterscheiden sich die geschätzten Übergangswahrscheinlichkeiten  $\hat{p}_{rs}$  und  $\hat{p}_{rs}^{(v)}$  (ggf. signifikant) voneinander, wie in Kapitel 6.2.1 erläutert wurde. Wenn sich  $\Delta_{rs}$  dabei oberhalb der oberen Grenze des 95%-Konfidenzintervalls befindet, also  $\Delta_{rs} > 2$  ist, dann gilt

$$\hat{p}_{rs} < \hat{p}_{rs}^{(v)}.$$

Es besteht also die Möglichkeit, dass  $\hat{p}_{rs}$  im SMI-Modell zu klein geschätzt wird, sodass die gegebenen Daten nicht ausreichend genau beschrieben werden. Für ein Schlafstadium  $Z$  könnte somit zu einem der oben angegebenen Übergänge

$$(r, s) \in \left\{ (z_1^{(Z)}, z_2^{(Z)}), (z_2^{(Z)}, z_1^{(Z)}) \right\}$$

(oder zu beiden Übergängen) die Übergangswahrscheinlichkeit  $\hat{p}_{rs}^{(v)}$  des vollen Modells durch den Schätzer  $\hat{p}_{rs}$  des SMI-Modells für gewisse Probanden unterschätzt worden

<sup>21</sup>Im Schlafstadium W treten für die Übergänge  $(C, C)$  und  $(C, D)$  bei 11 bzw. bei 9 Probanden Werte von  $\Delta_{r,s}^{(J,Z)}$  auf mit  $|\Delta_{r,s}^{(J,Z)}| > 2$ , von denen aber nur 3 bzw. 5 Werte größer als 2 sind.

sein. Dies könnte dann ein Hinweis darauf sein, dass es in jedem Schlafstadium gewisse dominierende Abfolgen von Zuständen<sup>22</sup> gibt, die mit höheren Wahrscheinlichkeiten auftreten als andere. Diese Abfolgen könnten beispielsweise der Form

$$z_1^{(Z)} \rightarrow z_2^{(Z)} \rightarrow z_1^{(Z)} \quad \text{oder} \quad z_2^{(Z)} \rightarrow z_1^{(Z)} \rightarrow z_2^{(Z)}$$

sein.

Eine weitere mögliche Erklärung dafür, dass das SMI-Modell die Übergangswahrscheinlichkeiten für bestimmte entgegengesetzte Übergänge  $(z_1^{(Z)}, z_2^{(Z)})$  und  $(z_2^{(Z)}, z_1^{(Z)})$  unterschätzt, könnte bei der Methode der Datenerhebung gefunden werden. Wie in Kapitel 1 beschrieben wurde, werden für die Ermittlung der Sequenzen  $x^{(J,Z)}$  die Messwerte eines EEGs, die zu bestimmten Auswertungszeitpunkten beobachtet werden, durch eine Clusteranalyse in vier Cluster eingeteilt. Liegen die Messwerte zu hintereinander folgenden Auswertungszeitpunkten bildlich gesprochen auf der Grenze zwischen zwei Clustern, die den Zuständen  $z_1^{(Z)}$  und  $z_2^{(Z)}$  entsprechen, wäre es möglich, dass die Messwerte abwechselnd den Clustern bzw. Zuständen  $z_1^{(Z)}$  und  $z_2^{(Z)}$  zugeordnet werden, obwohl sich die Messwerte kaum unterscheiden und eigentlich alle denselben EEG-Microstate repräsentieren, also im selben Cluster liegen müssten.

Im nächsten Unterkapitel wird das SMI-Modell dahingehend abgeändert, dass die Übergangswahrscheinlichkeiten für die oben getrennt nach Schlafstadium  $Z$  aufgelisteten Übergänge

$$(z_1^{(Z)}, z_2^{(Z)}) \quad \text{und} \quad (z_2^{(Z)}, z_1^{(Z)})$$

nach oben korrigiert werden, jedoch nicht für den Übergang

$$(z_2^{(N3)}, z_2^{(N3)}) = (D, D)$$

in N3.

### 6.3 Das Erweiterte SMI-Modell

In Kapitel 6.2.3 wurde ausgeführt, dass die Möglichkeit besteht, dass die Übergangswahrscheinlichkeit für den Übergang  $(r, s)$  durch den Schätzer  $\hat{p}_{rs}$  des SMI-Modells zu klein geschätzt wird, wenn  $\Delta_{rs}$  größer ist als die obere Grenze seines 95%-Konfidenzintervalls, also wenn  $\Delta_{rs} > 2$  ist.

Für jedes Schlafstadium  $Z$  wurden zudem zwei entgegengesetzte Übergänge

$$(z_1^{(Z)}, z_2^{(Z)}) \quad \text{und} \quad (z_2^{(Z)}, z_1^{(Z)})$$

zwischen zwei Zuständen

$$z_1 = z_1^{(Z)} \quad \text{und} \quad z_2 = z_2^{(Z)}$$

<sup>22</sup>Im EEG-Microstate-Kontext entsprechen die Abfolgen von Zuständen der sogenannten Syntax, also der Abfolge der EEG-Microstates, die z. B. von Lehmann et al. (2005) und Schlegel et al. (2012) mithilfe von Markov-Ketten untersucht wurde.

hervorgehoben, für die  $\Delta_{z_1 z_2}^{(J,Z)} > 2$  bzw.  $\Delta_{z_2 z_1}^{(J,Z)} > 2$  für möglichst viele Probanden  $J$  gilt. Für die Übergänge  $(z_1, z_2)$  und  $(z_2, z_1)$  wird also nach den Ausführungen in Kapitel 6.2.3 in Erwägung gezogen, dass die zugehörigen Übergangswahrscheinlichkeiten im SMI-Modell für viele Probanden, eventuell systematisch, zu klein definiert werden und somit auch die entsprechenden Schätzer des SMI-Modells zu kleine Werte ergeben.

Die beiden Zustände  $z_1$  und  $z_2$  werden in Tabelle 1 in Abhängigkeit vom jeweiligen Schlafstadium  $Z$  noch einmal angegeben.

Schlafstadium	W	N1	N2	N3
$z_1$	A	A	C	A
$z_2$	B	B	D	D

Tabelle 1: Zustände  $z_1, z_2$ , für die  $\Delta_{z_1 z_2}^{(J,Z)} > 2$  bzw.  $\Delta_{z_2 z_1}^{(J,Z)} > 2$  für möglichst viele Probanden  $J$  gilt.

Das SMI-Modell wird im Folgenden dahin gehend abgeändert, dass die Übergangswahrscheinlichkeiten  $p_{z_1 z_2}$  und  $p_{z_2 z_1}$  größer definiert werden, damit sie bzw. ihre Schätzer näher an den Übergangswahrscheinlichkeiten  $\hat{p}_{z_1 z_2}^{(v)}$  und  $\hat{p}_{z_2 z_1}^{(v)}$  des vollen Modells liegen. Werden die Übergangswahrscheinlichkeiten  $p_{z_1 z_2}$  und  $p_{z_2 z_1}$  vergrößert, müssen andere Übergangswahrscheinlichkeiten entsprechend verkleinert werden.

### 6.3.1 Idee und Definition

Nachstehend wird zunächst die Idee für die neue Definition der Übergangswahrscheinlichkeiten vorgestellt. Anschließend erfolgt die formale Definition des abgeänderten SMI-Modells, das als „Erweitertes SMI-Modell“ bezeichnet wird.

Ausgehend vom SMI-Modell mit Zustandsraum  $S$  in Verbindung mit dem Hintergrundmodell lässt sich ein Übergang von  $r \in S$  nach  $s \in S$  als zweistufiges Zufallsexperiment auffassen, das in Abbildung 6.7 illustriert wird.

In der ersten Stufe findet im Sinne des Hintergrundmodells mit Wahrscheinlichkeit  $1 - \pi$  kein Intervallwechsel statt, sodass ein Übergang von  $r$  zu  $r$  erfolgt. Mit der Gegenwahrscheinlichkeit  $\pi$  tritt ein Intervallwechsel auf und in der nächsten Stufe wird mit Wahrscheinlichkeit  $p_s$  der Zustand  $s$  ausgewählt, sodass ein Übergang von  $r$  zu  $s$  erfolgt, wobei  $r = s$  möglich ist.

Die Idee für die neue Definition der Übergangswahrscheinlichkeiten ist, dass man in den Fällen, in denen man von  $z_1$  oder von  $z_2$  aus startet, zwischen der ersten und zweiten Stufe eine weitere Stufe einfügt, durch die sich die Wahrscheinlichkeit erhöht,  $z_2$  bzw.  $z_1$  zu erreichen. Bei Start in einem Zustand in  $r \in S \setminus \{z_1, z_2\}$  bleibt es bei den oben beschriebenen zwei Schritten des SMI-Modells.

Das dreistufige Zufallsexperiment bei Start in  $z_1$  ist in Abbildung 6.8 dargestellt. Vertauscht man in Abbildung 6.8 die Variablen  $z_1$  und  $z_2$ , erhält man eine Darstellung des dreistufigen Zufallsexperiments bei Start in  $z_2$ .

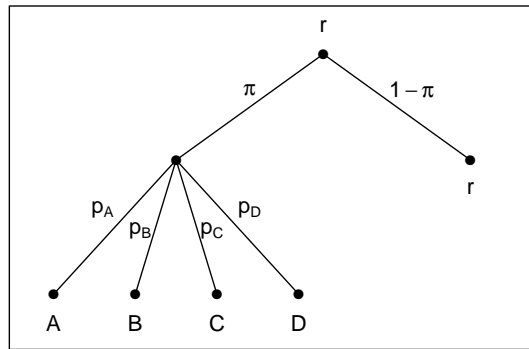


Abbildung 6.7: Zerlegung eines Übergangs im SMI-Modell in ein zweistufiges Zufallsexperiment.

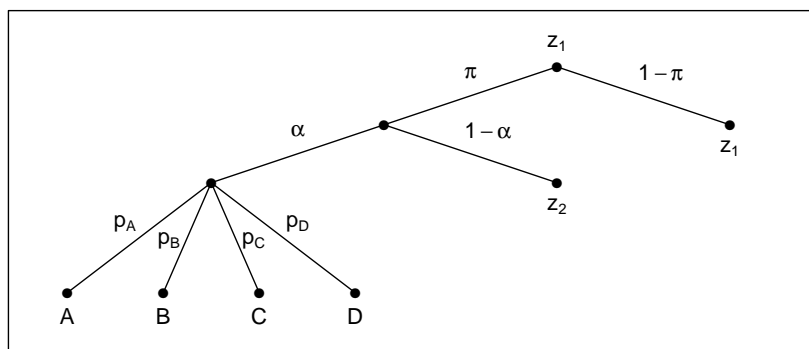


Abbildung 6.8: Zerlegung eines Übergangs im Erweiterten SMI-Modell in ein dreistufiges Zufallsexperiment für den Ausgangszustand  $z_1$ .

Startend bei  $z_1$  (bzw.  $z_2$ ) wird im dreistufigen Zufallsexperiment in der ersten Stufe mit Wahrscheinlichkeit  $1-\pi$  der Zustand  $z_1$  (bzw.  $z_2$ ) ausgewählt und mit Wahrscheinlichkeit  $\pi$  geht es mit der neu eingefügten zweiten Stufe weiter. In der zweiten Stufe wird mit Wahrscheinlichkeit  $1-\alpha$  für ein festes

$$\alpha \in (0, 1]$$

gerade der andere Zustand  $z_2$  (bzw.  $z_1$ ) gewählt und mit Wahrscheinlichkeit  $\alpha$  gelangt man zur dritten Stufe, bei der mit Wahrscheinlichkeit  $p_s$  der Zustand  $s \in S$  ausgewählt wird.

Durch die Auswahl von zwei speziellen Zuständen  $z_1$  und  $z_2$  sowie durch die Einführung einer neuen Wahrscheinlichkeit  $\alpha$  wird das SMI-Modell durch die neue Definition der Übergangswahrscheinlichkeiten um drei Parameter erweitert. Das resultierende Modell wird daher „Erweitertes SMI-Modell (für die Zustände  $z_1$  und  $z_2$ )“ genannt.

Es folgt die formale Definition des Erweiterten SMI-Modells.



**Modell 4** (Erweitertes SMI-Modell (für die Zustände  $z_1$  und  $z_2$ )). Seien  $a \geq 2$  und  $S := \{1, \dots, a\}$ , sodass  $z_1 = z_1^{(Z)}, z_2 = z_2^{(Z)} \in S, z_1 \neq z_2$ , gilt. Die Sequenz  $(x_0^{(J,Z)}, \dots, x_n^{(J,Z)})$  von Zuständen aus  $S$  werde aufgefasst als die ersten  $n + 1 = n^{(J,Z)} + 1$  Zustände  $(X_0, \dots, X_n)$  einer Markov-Kette  $(X_i)_{i \in \mathbb{N}_0}$  mit Zustandsraum  $S$  und der Startverteilung

$$p^\top = (p_1, \dots, p_a)$$

sowie der Übergangsmatrix  $P^{(\text{erw})} = (p_{rs}^{(\text{erw})})_{r,s \in S}$ , gegeben durch

$$p_{rs}^{(\text{erw})} := \begin{cases} \pi p_s & , r \neq s, \\ \pi p_s + 1 - \pi & , r = s, \end{cases} \quad \text{für } r \in S \setminus \{z_1, z_2\},$$

$$p_{z_1 s}^{(\text{erw})} := \begin{cases} \pi \alpha p_s & , s \in S \setminus \{z_1, z_2\}, \\ \pi \alpha p_s + 1 - \pi & , s = z_1, \\ \pi(\alpha p_s + 1 - \alpha) & , s = z_2, \end{cases}$$

und analog

$$p_{z_2 s}^{(\text{erw})} := \begin{cases} \pi \alpha p_s & , s \in S \setminus \{z_1, z_2\}, \\ \pi \alpha p_s + 1 - \pi & , s = z_2, \\ \pi(\alpha p_s + 1 - \alpha) & , s = z_1, \end{cases}$$

wobei

$$\alpha, \pi \in (0, 1]$$

und

$$p_1, \dots, p_a \in (0, 1) \text{ mit } \sum_{r=1}^a p_r = 1.$$

**Bemerkung 6.6.** (i) Für  $\alpha = 1$  entspricht das erweiterte SMI-Modell dem SMI-Modell.

(ii) Die Übergangswahrscheinlichkeiten  $p_{rs}^{(\text{erw})}$  des erweiterten SMI-Modells für die Zustände  $z_1, z_2$  lassen sich für alle  $r, s \in S$  ausdrücken durch

$$\begin{aligned} p_{rs}^{(\text{erw})} &= (1 - \delta_{r,z_1} - \delta_{r,z_2}) \left( \pi p_s + \delta_{r,s} (1 - \pi) \right) \\ &\quad + (\delta_{r,z_1} + \delta_{r,z_2}) \left( (\alpha \pi p_s + \delta_{r,s} (1 - \pi) + (1 - \delta_{r,s}) (\delta_{s,z_1} + \delta_{s,z_2}) \pi (1 - \alpha)) \right) \\ &= \pi p_s + \delta_{r,s} (1 - \pi) \\ &\quad + (\delta_{r,z_1} + \delta_{r,z_2}) \left( (\alpha - 1) \pi p_s + (1 - \delta_{r,s}) (\delta_{s,z_1} + \delta_{s,z_2}) \pi (1 - \alpha) \right) \\ &= p_{rs} + (\delta_{r,z_1} + \delta_{r,z_2}) \left( (\alpha - 1) \pi p_s + (1 - \delta_{r,s}) (\delta_{s,z_1} + \delta_{s,z_2}) \pi (1 - \alpha) \right). \end{aligned}$$

Für die Begründung beobachtet man im Fall  $r \in S \setminus \{z_1, z_2\}$ , dass

$$p_{rs}^{(\text{erw})} = p_{rs} = \pi p_s + \delta_{r,s} (1 - \pi)$$

gilt. Ist dagegen  $r \in \{z_1, z_2\}$ , so setzt sich  $p_{rs}^{(\text{erw})}$  stets aus  $\pi\alpha p_s$ , im Fall  $r = s$  zusätzlich aus  $1 - \pi$  und im Fall  $r \neq s \in \{z_1, z_2\}$  zusätzlich aus  $\pi(1 - \alpha)$  zusammen.

Im Gegensatz zum SMI-Modell ist  $(p_1, \dots, p_a)$  beim Erweiterten SMI-Modell mit  $\alpha < 1$  nicht die stationäre Verteilung, wie die folgende Proposition zeigt.

**Proposition 6.7.** *Seien  $a \geq 2$ ,  $z_1, z_2 \in S = \{1, \dots, a\}$  mit  $z_1 \neq z_2$  und  $T := \{z_1, z_2\}$ . Sei außerdem  $(X_i)_{i \in \mathbb{N}_0}$  eine Markov-Kette gemäß dem Erweiterten SMI-Modell für die Zustände  $z_1$  und  $z_2$ .*

*Die stationäre Verteilung*

$$\bar{p}^\top := (\bar{p}_1, \dots, \bar{p}_a)$$

*von  $(X_i)_{i \in \mathbb{N}_0}$  ist dann gegeben durch*

$$\bar{p}_r := \begin{cases} \frac{\alpha p_r}{\alpha \sum_{t \in T} p_t + \sum_{u \in S \setminus T} p_u} & , \text{ falls } r \in S \setminus T, \\ \frac{p_{z_1} + (1 - \alpha)p_{z_2}}{(2 - \alpha) \left( \alpha \sum_{t \in T} p_t + \sum_{u \in S \setminus T} p_u \right)} & , \text{ falls } r = z_1, \\ \frac{p_{z_2} + (1 - \alpha)p_{z_1}}{(2 - \alpha) \left( \alpha \sum_{t \in T} p_t + \sum_{u \in S \setminus T} p_u \right)} & , \text{ falls } r = z_2. \end{cases}$$

*Beweis.* Da  $p_{rs}^{(\text{erw})} > 0$  gilt für alle  $r, s \in S$ , ist die Markov-Kette  $(X_i)_{i \in \mathbb{N}_0}$  irreduzibel und somit besitzt sie eine eindeutige stationäre Verteilung. Es lässt sich nachrechnen, dass  $\bar{p} = (\bar{p}_1, \dots, \bar{p}_a)^\top$  ein linker Eigenvektor von  $P^{(\text{erw})}$  zum Eigenwert 1 ist, also die stationäre Verteilung beschreibt.  $\square$

### 6.3.2 Anpassung des Erweiterten SMI-Modells

Als Schätzer für die Parameter  $\alpha$ ,  $\pi$  und  $p_1, \dots, p_a$  des Erweiterten SMI-Modells werden in dieser Arbeit die ML-Schätzer

$$\hat{\alpha}^{(\text{erw}), \text{ML}}, \hat{\pi}^{(\text{erw}), \text{ML}}, \hat{p}_1^{(\text{erw}), \text{ML}}, \dots, \hat{p}_a^{(\text{erw}), \text{ML}}$$

auf der Basis von  $(X_0, \dots, X_n)$  verwendet. Der Ansatz zur Ermittlung der ML-Schätzer gestaltet sich analog zu dem in Kapitel 4.3.2 beschriebenen Ansatz für die ML-Schätzer auf der Basis von  $(X_0, \dots, X_n)$  im SMI-Modell.

Die ML-Schätzer wurden für die gegebenen Daten numerisch in R mit der Funktion `optim` und der Methode „Nelder-Mead“ approximiert. Als Startwerte für die Iterationen, die von der Funktion `optim` durchgeführt wurden, wurden

$$\alpha^{(\text{Start})} = \pi^{(\text{Start})} = 0.80, \quad p_A^{(\text{Start})} = \dots = p_D^{(\text{Start})} = 0.25$$

gewählt.

Die Schätzer  $\hat{p}_{rs}^{(\text{erw}),\text{ML}}$  für die Übergangswahrscheinlichkeiten  $p_{rs}^{(\text{erw})}$  werden dadurch bestimmt, dass  $p_A, p_B, p_C, p_D, \pi$  und  $\alpha$  in der Formel für  $p_{rs}^{(\text{erw})}$  durch ihre ML-Schätzer ersetzt werden (analog zu Kapitel 4.4).

Als Beispiel für die Anpassung des Erweiterten SMI-Modells durch die ML-Schätzer werden die geschätzten Übergangswahrscheinlichkeiten  $\hat{p}_{rs}^{(\text{erw}),\text{ML}}$  für die Probanden 2 und 5 im Schlafstadium N2 in Abbildung 6.9 dargestellt zusammen mit den Schätzern  $\hat{p}_{rs}$  des SMI-Modells und den Schätzern  $\hat{p}_{rs}^{(v)}$  des vollen Modells.

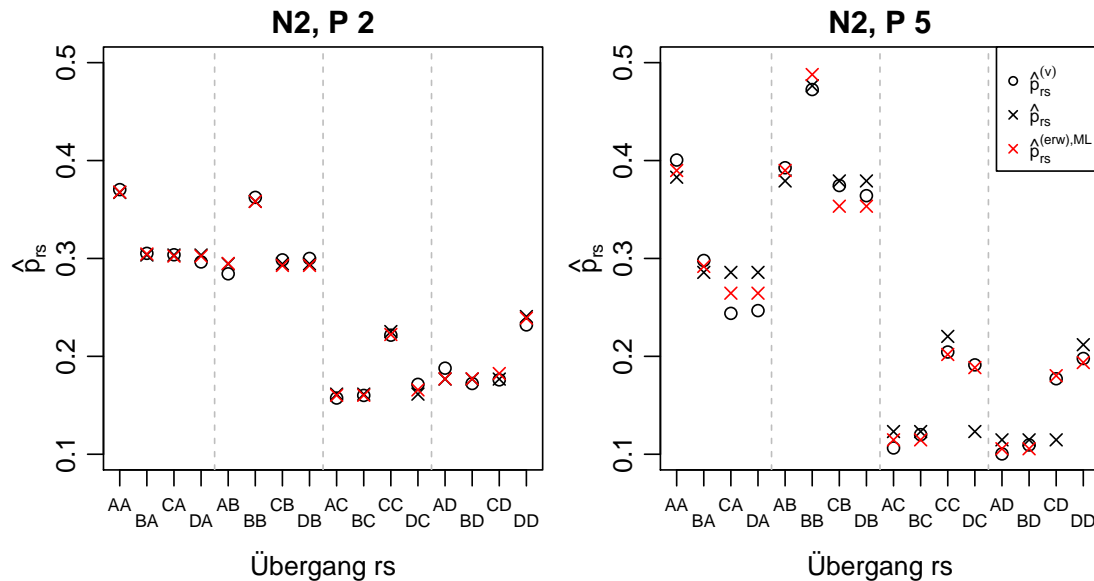


Abbildung 6.9: Geschätzte Übergangswahrscheinlichkeiten  $\hat{p}_{rs}^{(\text{erw}),\text{ML}}$  im Erweiterten SMI-Modell (rote Kreuze) für die Probanden 2 (links) und 5 (rechts) im Schlafstadium N2 im Vergleich zu den geschätzten Übergangswahrscheinlichkeiten  $\hat{p}_{rs}$  des SMI-Modells (schwarze Kreuze) und  $\hat{p}_{rs}^{(v)}$  des vollen Modells (Kreise).

Für das Schlafstadium N2 ist  $z_1 = z_1^{(\text{N2})} = C$  und  $z_2 = z_2^{(\text{N2})} = D$ .

Bei Proband 2 zeigen sich in Abbildung 6.9 für das Schlafstadium N2 kaum Unterschiede zwischen den Schätzern der Übergangswahrscheinlichkeiten des SMI-Modells und des Erweiterten SMI-Modells (die roten Kreuze überdecken die schwarzen Kreuze). Das spricht zum einen dafür, dass  $\alpha^{(\text{erw}),\text{ML}} \approx 1$  ist, und zum anderen dafür, dass die ML-Schätzer annähernd die gleichen Werte liefern wie die SMI-Modell-Schätzer  $\hat{p}_A, \hat{p}_B, \hat{p}_C, \hat{p}_D$  und  $\hat{\pi}$ .

Bei Proband 5 lässt sich in Abbildung 6.9 für das Schlafstadium N2 erkennen, dass sich die geschätzten Übergangswahrscheinlichkeiten  $\hat{p}_{rs}^{(\text{erw}),\text{ML}}$  und  $\hat{p}_{rs}^{(v)}$  für die Übergänge

$$(r, s) \in \left\{ \left( z_1^{(\text{N2})}, z_2^{(\text{N2})} \right), \left( z_2^{(\text{N2})}, z_1^{(\text{N2})} \right) \right\} = \{(C, D), (D, C)\}$$

annähernd gleichen, während der Schätzer  $\hat{p}_{rs}$  des SMI-Modells sich relativ stark von  $\hat{p}_{rs}^{(v)}$  unterscheidet für  $(r, s) \in \{(C, D), (D, C)\}$ . Das Ziel, die Übergangswahrscheinlichkeiten für genau diese Übergänge den Schätzern des vollen Modells anzugleichen, wurde hier also durch das Erweiterte SMI-Modell erreicht.

Im Vergleich mit den Schätzern des SMI-Modells liegen die Schätzer des Erweiterten SMI-Modells (für die Zustände  $C$  und  $D$ ) für alle Übergänge näher oder ähnlich nah an den Schätzern des vollen Modells außer für die Übergänge  $(B, B)$ ,  $(C, B)$  und  $(B, C)$ .

Insgesamt zeigt sich in den Beispielen für Proband 2 und 5, dass die Schätzer des Erweiterten SMI-Modells (für die Zustände  $C$  und  $D$ ) für die meisten Übergänge mindestens so dicht wie die Schätzer des SMI-Modells an den Schätzern des vollen Modells  $\hat{p}_{rs}^{(v)}$  liegen.

### 6.3.3 Ein $\chi^2$ -Anpassungstest für das Erweiterte SMI-Modell

Der  $\chi^2$ -Anpassungstest aus Kapitel 6.1 wird nun für jeden Probanden  $J$  und jedes Schlafstadium  $Z$  auf das Erweiterte SMI-Modell angewendet.

Nach Satz 6.3 und Bemerkung 6.4 ist die Teststatistik

$$\mathcal{X}^2 = \sum_{r,s \in S} \frac{\left( n_{rs} - n_r \hat{p}_{rs}^{(\text{erw}), \text{ML}} \right)^2}{n_r \hat{p}_{rs}^{(\text{erw}), \text{ML}}} \quad (6.7)$$

asymptotisch  $\chi^2$ -verteilt mit  $a^2 - a - d$  Freiheitsgraden, wobei  $d$  die Anzahl der zu schätzenden Parameter sei und  $a = |S|$ .

Im Erweiterten SMI-Modell für die Zustände  $z_1$  und  $z_2$  müssen die Parameter  $\pi$ ,  $\alpha$  sowie o.B.d.A.  $p_1, \dots, p_{a-1}$  geschätzt werden. Die Parameter  $z_1$  und  $z_2$  sind fest vorgegeben und müssen somit nicht geschätzt werden. Für das Erweiterte SMI-Modell gilt also

$$d = a + 1.$$

Für die gegebenen Daten ist  $a = 4$ , sodass  $\mathcal{X}^2$  asymptotisch  $\chi^2$ -verteilt ist mit

$$4^2 - 4 - (4 + 1) = 7$$

Freiheitsgraden.

Analog zu Kapitel 6.1 wird für jeden Probanden  $J$  und jedes Schlafstadium  $Z$  mithilfe der  $\mathcal{X}^2$ -Statistik die folgende Nullhypothese getestet:

**Nullhypothese  $H_0^{(\text{erw}), (J, Z)}$ :** Die beobachteten Übergangshäufigkeiten  $(n_{rs}^{(J, Z)})_{r, s}$  sind so verteilt wie die Übergangshäufigkeiten einer Markov-Kette der Länge  $n^{(J, Z)} + 1$  mit Zustandsraum  $S = \{A, B, C, D\}$ , die die Bedingungen des Erweiterten SMI-Modells für die Zustände  $z_1^{(Z)}$  und  $z_2^{(Z)}$  erfüllt.

In Abbildung 6.10 sind die p-Werte dargestellt, die berechnet wurden durch

$$p_{(\text{erw})}^{(J,Z)} := \mathbb{P} \left( \mathcal{X}_{(J,Z)}^2 < \mathfrak{x}^2 \right),$$

wobei  $\mathcal{X}_{(J,Z)}^2$  der aus den Daten für Proband  $J$  und Schlafstadium  $Z$  nach Gleichung (6.7) berechnete feste Wert sei und  $\mathfrak{x}^2$  eine  $\chi^2$ -verteilte Zufallsvariable mit 7 Freiheitsgraden. Wie schon in Abbildung 6.1 zeigen auch in Abbildung 6.10 die grünen Quadrate an, dass die zugehörige Nullhypothese  $H_0^{(\text{erw}), (J,Z)}$  auf einem Signifikanzniveau von 5% nicht verworfen wird, im Gegensatz zu den roten und violetten Quadraten.

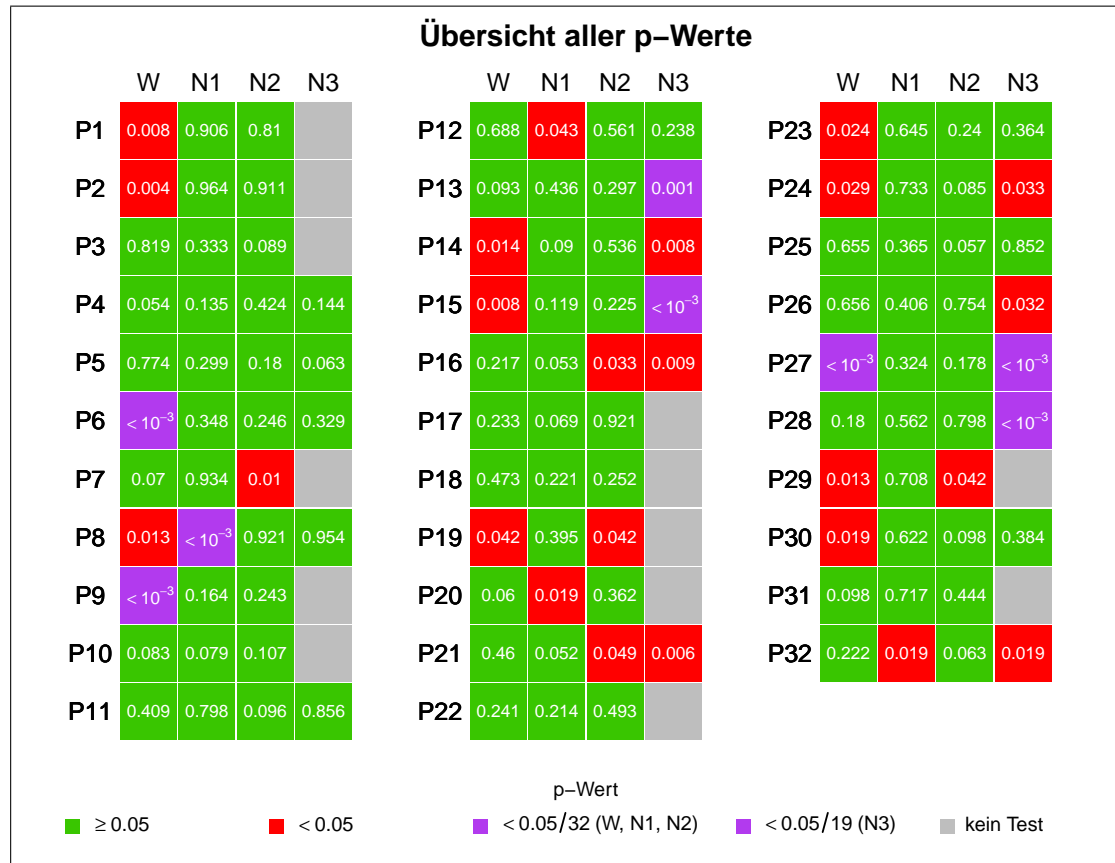


Abbildung 6.10: Übersicht über die p-Werte aller Anpassungstests für das Erweiterte SMI-Modell.

Im Vergleich zum SMI-Modell wird die Nullhypothese  $H_0^{(\text{erw}), (J,Z)}$  für das Erweiterte SMI-Modell für merklich weniger Proband-Schlafstadium-Kombinationen  $(J, Z)$  auf einem Signifikanzniveau von 5% verworfen, nämlich

- für W bei 13 statt 19 Probanden,

- für N1 bei 4 statt 14 Probanden,
- für N2 bei 5 statt 23 Probanden und
- für N3 bei 10 statt 12 Probanden.

Vor allem für die Schlafstadien N1 und N2 lassen sich die Daten deutlich besser an das Erweiterte SMI-Modell als an das SMI-Modell anpassen.

Ermittelt man die p-Werte durch Simulationen analog zur Durchführung des Anpassungstestes, die in Kapitel 6.1.2 beschrieben wurde, wird für die Schlafstadien W und N1 jeweils eine Nullhypothese mehr abgelehnt, für N2 werden 3 Nullhypothesen weniger abgelehnt und für N3 wird eine Nullhypothese weniger abgelehnt.

Eine weitere Möglichkeit besteht darin, die Nullhypothese  $H_0^{(\text{erw}), (Z)}$  für ein festes Schlafstadium  $Z$  mittels Bonferroni-Korrektur zu testen, dass alle Nullhypothesen  $H_0^{(\text{erw}), (J, Z)}$  simultan wahr sind. Während die analoge Nullhypothese  $H_0^{(Z)}$  für das SMI-Modell für alle Schlafstadien  $Z$  bei Anwendung der Bonferroni-Korrektur verworfen wird (siehe 6.1.3), wird die Nullhypothese  $H_0^{(\text{erw}), (Z)}$  für das Erweiterte SMI-Modell nur für die Schlafstadien W, N1 und N3 nach Bonferroni-Korrektur auf einem Signifikanz-Niveau von 5% verworfen. Durch den Anpassungstest zeigen die Daten für das Schlafstadium N2 keine signifikanten Abweichungen von den Modellannahmen des Erweiterten SMI-Modells.

Insgesamt wird deutlich, dass sich die gegebenen Daten durch das Erweiterte SMI-Modell besser beschreiben lassen als durch das SMI-Modell. Dieses Ergebnis würde man auch erwarten, da das SMI-Modell für  $\alpha = 1$  ein Spezialfall des Erweiterten SMI-Modells ist und die Struktur der Übergangswahrscheinlichkeiten für das Erweiterte SMI-Modell aufgrund der drei zusätzlichen Parameter  $z_1, z_2$  und  $\alpha$  flexibler gewählt werden kann.

Das SMI-Modell weist jedoch den Vorteil auf, dass es sich durch das Hintergrundmodell beschreiben lässt, das den Prozessen bei der Ermittlung von EEG-Microstate-Sequenzen nachempfunden ist (siehe Kapitel 1.2). Soll das Hintergrundmodell erhalten werden, ist daher das SMI-Modell dem Erweiterten SMI-Modell vorzuziehen.

## 7 Schlussbetrachtung

### 7.1 Zusammenfassung

In dieser Arbeit wurde ein Modell vorgestellt für Sequenzen von Karten (im Kontext von EEG-Microstates), die sich formal beschreiben lassen als endliche Folgen von Zuständen aus einem endlichen Zustandsraum.

Unter der Annahme, dass sich diese Sequenzen als homogene Markov-Kette 1. Ordnung modellieren lassen, wurde anhand der für diese Arbeit vorliegenden Daten eine spezielle Struktur für die Übergangsmatrix der Markov-Kette motiviert.

Das daraus resultierende SMI-Modell für die Sequenzen von Karten, lässt sich dadurch konstruieren, dass man zufällige markierte Intervalle zu zufälligen Zeitpunkten abtastet, wobei die Intervalle und Abtastungszeitpunkte durch unabhängige Poisson'sche Punktprozesse gegeben sind. Im Vergleich zum vollen Modell mit 12 Parametern ist die Anzahl der Parameter des SMI-Modells auf 4 Parameter reduziert.

Für die Parameter des SMI-Modells wurden Schätzer zur Anpassung des SMI-Modells an die Daten vorgestellt und diskutiert. Anschließend wurden Formeln für die asymptotischen Varianzen dieser Schätzer hergeleitet. Dabei wurde auch eine Formel für Kovarianzen von Auftrittshäufigkeiten von Zuständen in allgemeineren Markov-Ketten nach einem Ergebnis von Good (1961) bewiesen.

Die Anpassungsgüte des SMI-Modells an die gegebenen Daten wurde mit einem Anpassungstest überprüft, der sich am  $\chi^2$ -Anpassungstest von Pearson orientiert. Es zeigte sich, dass für einen Großteil der Probanden in den jeweiligen Schlafstadien das SMI-Modell von den Daten signifikant abweicht.

Ausgehend von diesem Ergebnis wurde untersucht, für welche Übergänge von einem Zustand  $r$  zu einem Zustand  $s$  in den als Markov-Kette modellierten Sequenzen von Karten die Abweichungen zwischen den gegebenen Daten und dem SMI-Modell am deutlichsten sind. Dabei wurden für jedes Schlafstadium zwei einander entgegengesetzte Übergänge  $(z_1, z_2)$  und  $(z_2, z_1)$  identifiziert, die von allen Übergängen, die größte Anzahl von Probanden aufweisen, bei deren Daten sich für die jeweiligen Übergänge signifikante Abweichungen vom SMI-Modell zeigen.

Abschließend wurde das Erweiterte SMI-Modell vorgestellt und untersucht, bei dem diese Abweichungen für die zwei entgegengesetzten Übergänge  $(z_1, z_2)$  und  $(z_2, z_1)$  durch die Einführung von drei weiteren Parametern möglichst vermieden werden.

### 7.2 Ausblick

Aufbauend auf die Ergebnisse und Beobachtungen dieser Arbeit schließen sich weiterführende Untersuchungsmöglichkeiten in Bezug auf das SMI-Modell an.

In Kapitel 6.2.3 wurde die Möglichkeit erörtert, dass die Übergänge  $(z_1, z_2)$  und  $(z_2, z_1)$  für zwei bestimmte Zustände  $z_1, z_2$  häufiger in den Sequenzen  $x^{(J,Z)}$  auftreten, als man

für das SMI-Modell erwarten würde. Als eine mögliche Erklärung dafür wurde genannt, dass bei der Erhebung der Sequenzen  $x^{(J,Z)}$  die Topografien zu den Auswertungszeitpunkten teilweise keinem der ermittelten Cluster eindeutig zuzuordnen ist, sondern bildlich gesprochen auf der Grenze zwischen zwei Clustern liegen. Dadurch könnte es passieren, dass etwa der Zustand  $z_1$  mit einer gewissen Wahrscheinlichkeit dem  $i$ -ten Auswertungszeitpunkt zugeordnet wird, obwohl  $z_2$  der nicht beobachtbare wahre Zustand wäre, und umgekehrt. Zu diesem Umstand ließe sich ein Hidden Markov Model (HMM) untersuchen, bei dem die nicht beobachtbaren wahren Zustände als Markov-Kette (z. B. durch das SMI-Modell) modelliert werden und abhängig vom wahren Zustand mit gewissen Wahrscheinlichkeiten andere Zustände beobachtet werden.

Es lässt sich aufgrund von Abbildung 4.6 auf Seite 55 vermuten, dass der Parameter  $\pi^{(J,N3)}$  für jeden Probanden  $J$  und das Schlafstadium N3 signifikant kleiner ist als die Parameter  $\hat{\pi}^{(J,Z)}$  in den anderen drei Schlafstadien  $Z$ . Dieser Vermutung könnte man beispielsweise mit einem Permutations- oder t-Test nachgehen.

Schließlich wäre es interessant, im Hintergrundmodell zum SMI-Modell die zufälligen Intervalle und Abtastungszeitpunkte nicht durch Poisson'sche Punktprozesse zu definieren, sondern beispielsweise durch Gamma-Prozesse, und die Eigenschaften der dadurch resultierenden Sequenzen  $(X_0, \dots, X_n)$  zu analysieren.



## A Ermittlung von EEG-Microstate-Sequenzen

Für jeden Probanden  $J$  wurde in verschiedenen Schlafstadien<sup>23</sup>  $Z$  das elektrische Feld an der Kopfoberfläche aufgezeichnet durch die Methode der Elektroenzephalografie mit 30 Elektroden über einen Zeitraum von mehreren Minuten.

Im Folgenden wird erläutert, wie aus den aufgezeichneten Elektroenzephalogrammen (EEGs) für jeden Probanden  $J$  zu den verschiedenen Schlafstadien  $Z$  die Sequenzen

$$x^{(J,Z)} = \left( x_0^{(J,Z)}, \dots, x_{n^{(J,Z)}}^{(J,Z)} \right)$$

ermittelt wurden, aus denen wiederum die für diese Arbeit vorliegenden Übergangshäufigkeiten  $n_{rs}^{(J,Z)}$  bestimmt wurden. Dabei wird auch auf die Theorie der EEG-Microstates eingegangen.

Für jeden Probanden liegen in einem zeitlichen Abstand von jeweils 4 Millisekunden Messwerte eines mit 30 Elektroden aufgezeichneten EEGs vor. Die Messwerte eines Probanden  $J$  in einem Schlafstadium  $Z$  zum Zeitpunkt

$$t \in \mathcal{T}_{(J,Z)} := \left\{ t_1^{(J,Z)}, \dots, t_{l^{(J,Z)}}^{(J,Z)} \right\}$$

seien bezeichnet durch

$$v_{(J,Z)}^{(t)} = \left( v_{(J,Z),1}^{(t)}, \dots, v_{(J,Z),30}^{(t)} \right).$$

Dabei hängen die genauen Messzeitpunkte  $t_k^{(J,Z)}$  ebenso von  $(J, Z)$  ab wie die Anzahl  $l^{(J,Z)}$  der Messzeitpunkte. Außerdem seien für jeden Probanden  $J$  die Mengen

$$\mathcal{T}_{(J,W)}, \mathcal{T}_{(J,N1)}, \mathcal{T}_{(J,N2)}, \mathcal{T}_{(J,N3)}$$

der Messzeitpunkte paarweise disjunkt.

Es ist möglich, dass Messwerte zu gewissen Zeitpunkten aus verschiedenen Gründen unbrauchbar sind, etwa weil sich der Proband zu heftig bewegt hat, und deshalb verworfen werden. Außerdem werden die Messwerte und -zeitpunkte von einem Probanden  $J$  zu einem Schlafstadium  $Z$  nicht verwendet, wenn sich der Proband  $J$  nicht mindestens 105 Sekunden in diesem Schlafstadium befunden hat während der Aufzeichnung des EEGs.

Im Folgenden wird davon ausgegangen, dass für jeden Probanden  $J$  jeweils die Messwerte zu allen Zeitpunkten

$$t \in \bigcup_Z \mathcal{T}_{(J,Z)}$$

verwendet werden.

Von Lehmann et al. (1987) wurde beobachtet, dass die Topographie (hochdimensionale Ausprägung) des durch eine Elektroenzephalografie gemessenen elektrischen Feldes an

<sup>23</sup>Informationen zu den Probanden und Schlafstadien in Kapitel 2.1.

der Kopfoberfläche über gewisse Zeiträume annähernd stabil ist, abgesehen von Polarität und der Stärke der Ausprägung. Das bedeutet, dass die Vektoren  $v_{(J,Z)}^{(t)}$  für gewisse direkt aufeinanderfolgende Zeitpunkte  $t_1, \dots, t_n \in \mathcal{T}_{(J,Z)}$  annähernd auf einer Geraden liegen. Nach einem dieser Zeiträume mit annähernd stabiler Topographie erfolgt ein diskontinuierlicher Wechsel zum nächsten Zeitraum mit annähernd stabiler Topographie. Diese Beobachtung führt zur Definition von EEG-Microstates:

**Definition A.1.** „EEG microstates are time epochs in the sub-second range in which the scalp potential topographies appear as discrete segments of electrical stability, separated by sharp and short transitions, switching abruptly into the next period of stability.“ (Brodbeck et al., 2012)

Weiterhin schlägt Lehmann et al. (1987) vor, die EEG-Microstates zu bestimmen, indem man nur die Messwerte in den Zeitpunkten betrachtet, in denen das elektrische Feld lokal maximal stark ausgeprägt ist. Denn um diese Zeitpunkte herum sei die Topographie des elektrischen Feldes besonders stabil und durch die starke Ausprägung des elektrischen Feldes werde der Einfluss von Messfehlern („Noise“) geringer.

Gemessen wird die Stärke der Ausprägung des elektrischen Feldes durch die „Global Field Power“ (GFP), die gerade der empirischen Standardabweichung über alle 30 Messwerte zu einem festen Zeitpunkt  $t \in \mathcal{T}_{(J,Z)}$  entspricht. D.h. für einen Probanden  $J$  und ein Schlafstadium  $Z$  wird die GFP berechnet durch

$$\text{GFP}_{(J,Z)}(t) := \left( \frac{1}{30} \sum_{i=1}^{30} \left( v_{(J,Z),i}^{(t)} - \bar{v}_{(J,Z)}^{(t)} \right)^2 \right)^{1/2} \quad \forall t \in \mathcal{T}_{(J,Z)},$$

wobei  $\bar{v}_{(J,Z)}^{(t)}$  gerade das arithmetische Mittel

$$\bar{v}_{(J,Z)}^{(t)} = \frac{1}{30} \sum_{i=1}^{30} v_{(J,Z),i}^{(t)}$$

der Komponenten von  $v_{(J,Z)}^{(t)}$  sei.

Die lokalen Maxima von GFP werden GFP-Peaks genannt. Die Zeitpunkte, in denen ein GFP-Peak für Proband  $J$  und Schlafstadium  $Z$  auftritt, seien bezeichnet mit

$$b_0^{(J,Z)}, \dots, b_{n_{(J,Z)}}^{(J,Z)},$$

wobei auch die Anzahl  $n_{(J,Z)} + 1$  der GFP-Peaks für jedes Paar  $(J, Z)$  unterschiedlich sein kann.

Im nächsten Schritt wird eine Menge prototypischer Topographien gesucht, die jeweils eine Gruppe ähnlicher Messwerte zu den Zeitpunkten der GFP-Peaks repräsentieren. Dazu werden die Messwerte

$$v_{(J,Z)}^{(b_0^{(J,Z)})}, \dots, v_{(J,Z)}^{(b_{n_{(J,Z)}}^{(J,Z)})}$$

zu den Zeitpunkten der GFP-Peaks mithilfe einer Clusteranalyse in Cluster eingeteilt. Das für die Erhebung der vorliegenden Daten verwendete Verfahren der Clusteranalyse wird „Topographical Atomize and Agglomerate Hierarchical Clustering“ (T-AAHC) genannt (Brodbeck et al. (2012), für technische Details Handy (2009)). Für jeden Probanden  $J$  und jedes Schlafstadium  $Z$  werden durch die Clusteranalyse  $k_{(J,Z)}$  Cluster ermittelt. Zu jedem dieser Cluster

$$c \in \mathcal{C}_{(J,Z)} := \left\{ c_1^{(J,Z)}, \dots, c_{k_{(J,Z)}}^{(J,Z)} \right\}$$

wird ein 30-dimensionaler Vektor  $m^{(c)}$  als Repräsentant über die Durchschnittsbildung aller Messwerte in diesem Cluster berechnet.

In einem anschließenden Schritt werden für ein festes Schlafstadium  $Z$  die Vektoren  $m^{(c)}$  aller Cluster

$$c \in \bigcup_{J \in \mathcal{J}_Z} \mathcal{C}_{(J,Z)}$$

sämtlicher Probanden  $J$  wieder durch das T-AAHC in neue Cluster eingeteilt. Dadurch wurden in den für diese Arbeit vorliegenden Daten für jedes Schlafstadium  $Z$  vier Cluster

$$c_A^{(Z)}, c_B^{(Z)}, c_C^{(Z)}, c_D^{(Z)}$$

gefunden. Die Menge der Cluster im Schlafstadium  $Z$  sei  $\mathcal{C}^{(Z)}$ . Die 30-dimensionalen Vektoren, die jeweils einen der Cluster  $c \in \mathcal{C}^{(Z)}$  repräsentieren und Karten (maps) genannt werden, seien mit

$$A_Z, B_Z, C_Z, D_Z$$

und die Menge dieser Karten mit  $S_Z$  bezeichnet.

Für jeden Probanden  $J$  und jedes Schlafstadium  $Z$  wird nun allen Zeitpunkten

$$b_0^{(J,Z)}, \dots, b_{n_{(J,Z)}}^{(J,Z)}$$

der GFP-Peaks diejenige Karte  $m \in S_Z$  zugeordnet, die die größte Ähnlichkeit zu dem jeweiligen Messwert  $v_{b_k^{(J,Z)}}^{(J,Z)}$  aufweist. Die Ähnlichkeit wird gemessen über den Absolutbetrag des empirischen Korrelationskoeffizienten von  $m$  und  $v_{b_k^{(J,Z)}}^{(J,Z)}$ . Für jedes  $k \in \{0, \dots, n_{(J,Z)}\}$  werde  $x_k^{(J,Z)}$  definiert als die auf diese Weise ermittelte Karte  $m$ . Formal bedeutet das

$$x_k^{(J,Z)} := \arg \max_{m \in S_Z} \left| \frac{\sum_{i=1}^{30} \left( v_{(J,Z),i}^{(b_k^{(J,Z)})} - \bar{v}_{(J,Z)}^{(b_k^{(J,Z)})} \right) (m_i - \bar{m})}{\left( \sum_{i=1}^{30} \left( v_{(J,Z),i}^{(b_k^{(J,Z)})} - \bar{v}_{(J,Z)}^{(b_k^{(J,Z)})} \right)^2 \right)^{1/2} \left( \sum_{i=1}^{30} (m_i - \bar{m})^2 \right)^{1/2}} \right|.$$

Somit erhält man für jeden Probanden  $J$  und jedes Schlafstadium  $Z$  eine Sequenz

$$x^{(J,Z)} = \left( x_0^{(J,Z)}, \dots, x_{n^{(J,Z)}}^{(J,Z)} \right)$$

von Karten aus  $S_Z$ .

Die Zuweisung der Kartenbezeichnungen

$$A_Z, B_Z, C_Z, D_Z$$

erfolgte bei der Erhebung der vorliegenden Daten für jedes Schlafstadium  $Z$  so, dass die Karten möglichst ähnlich sind zu den jeweiligen entsprechenden Referenzkarten

$$A, B, C, D$$

aus anderen Studien.

In den gegebenen Daten wird bei den Sequenzen  $x^{(J,Z)}$  daher beispielsweise nicht zwischen den Karten

$$A_W, A_{N1}, A_{N2}, A_{N3}$$

aus den verschiedenen Schlafstadien unterschieden, sondern diese Karten alle mit  $A$  bezeichnet (analog für  $B, C, D$ ).

Somit ist der Zustandsraum für alle  $x^{(J,Z)}$  gegeben durch

$$S := \{A, B, C, D\}.$$

## B Die Delta-Methode

Sei  $(Z_1, \dots, Z_m)$  ein reell-wertiger Vektor von Zufallsvariablen und  $f$  eine Funktion auf  $\mathbb{R}^m$ . Erfüllt  $f$  in Korollar B.2 näher bestimmte Differenzierbarkeitsbedingungen, lässt sich die Delta-Methode anwenden. Die Delta-Methode liefert eine Approximation für  $\mathbb{E}[f(X)]$  mithilfe der Taylor-Approximation erster Ordnung an der Entwicklungsstelle  $\mathbb{E}[X]$  und darauf aufbauend eine Approximation für  $\text{Var}(f(X))$ . Für die Taylor'sche Formel wird hier die benötigte Version 1. Ordnung angegeben.

**Satz B.1** (Taylor'sche Formel 1. Ordnung). *Sei  $U \subset \mathbb{R}^m$  offen,  $a \in U$  und  $f : U \rightarrow \mathbb{R}$  eine zweifach stetig differenzierbare Funktion. Dann gilt für alle  $x \in U$  mit  $a + \lambda x \in U$  für alle  $\lambda \in [0, 1]$ :*

$$f(x) = f(a) + \langle \text{grad}f(a), x - a \rangle + R_2(x)$$

mit

$$R_2(x) = \frac{1}{2} \langle \text{Hess}f(a + \lambda_0(x - a))(x - a), x - a \rangle \quad \text{für ein } \lambda_0 \in (0, 1).$$

Dabei bezeichne  $\text{grad}f$  den Gradienten und  $\text{Hess}f$  die Hesse-Matrix von  $f$ .

*Beweis.* Siehe z. B. Forster (2006). □

**Korollar B.2** (Delta-Methode). Sei  $Z = (Z_1, \dots, Z_m)$  ein Vektor von Zufallsvariablen mit Werten in  $U \subset \mathbb{R}^m$  offen und sei  $\vartheta = (\vartheta_1, \dots, \vartheta_m)$  mit  $\vartheta_i := \mathbb{E}[Z_i]$ . Sei weiterhin  $f$  eine auf  $U$  zweifach stetig differenzierbare Funktion und bezeichne

$$f'_i(\vartheta) := \left. \frac{\partial f(z)}{\partial z_i} \right|_{z_1=\vartheta_1, \dots, z_m=\vartheta_m}$$

die erste partielle Ableitung von  $f$  nach  $z_i$  an der Stelle  $\vartheta$ . Dann gelten

$$\mathbb{E}[f(Z)] \approx f(\vartheta)$$

und

$$\text{Var}(f(Z)) \approx \sum_{i=1}^m (f'_i(\vartheta))^2 \text{Var}(Z_i) + 2 \sum_{\substack{i,j=1 \\ i < j}}^m f'_i(\vartheta) f'_j(\vartheta) \text{Cov}(Z_i, Z_j).$$

*Beweis.* Der Beweis folgt den Ausführungen von Casella und Berger (2002). Nach der Taylor'schen Formel 1. Ordnung erhält man

$$f(z) = f(\vartheta) + \langle \text{grad} f(\vartheta), z - \vartheta \rangle + R_2(z) \quad \forall z \in U$$

und somit

$$\mathbb{E}[f(Z)] = f(\vartheta) + \sum_{i=1}^m f'_i(\vartheta) \mathbb{E}[Z_i - \vartheta_i] + \mathbb{E}[R_2(Z)].$$

Da  $\mathbb{E}[Z_i] = \vartheta_i$  ist, folgt  $\mathbb{E}[Z_i - \vartheta_i] = 0$  und somit

$$\mathbb{E}[f(Z)] = f(\vartheta) + \mathbb{E}[R_2(Z)] \approx f(\vartheta).$$

Mit dieser Approximation für  $\mathbb{E}[f(Z)]$  und erneuter Anwendung der Taylor'schen Formel ergibt sich

$$\begin{aligned} \text{Var}(f(Z)) &= \mathbb{E}[(f(Z) - \mathbb{E}[f(Z)])^2] \\ &\approx \mathbb{E}[(f(Z) - f(\vartheta))^2] \\ &= \mathbb{E} \left[ \left( \sum_{i=1}^m f'_i(\vartheta)(Z_i - \vartheta_i) + R_2(Z) \right)^2 \right] \\ &\approx \mathbb{E} \left[ \left( \sum_{i=1}^m f'_i(\vartheta)(Z_i - \vartheta_i) \right)^2 \right] \\ &= \sum_{i=1}^m (f'_i(\vartheta))^2 \text{Var}(Z_i) + 2 \sum_{\substack{i,j=1 \\ i < j}}^m f'_i(\vartheta) f'_j(\vartheta) \text{Cov}(Z_i, Z_j) \end{aligned}$$

□

## Literaturverzeichnis

- Bartlett, M. S. (1951). The frequency goodness of fit test for probability chains. *Mathematical Proceedings of the Cambridge Philosophical Society*, 47(01):86.
- Billingsley, P. (1961). Statistical Methods in Markov Chains. *The Annals of Mathematical Statistics*, 32(1):12–40.
- Brodbeck, V., Kuhn, A., Wegner, F. v., Morzelewski, A., Tagliazucchi, E., Borisov, S., Michel, C. M., und Laufs, H. (2012). EEG microstates of wakefulness and NREM sleep. *NeuroImage*, 62(3):2129–2139.
- Casella, G. und Berger, R. L. (2002). *Statistical inference*. Thomson Learning, Pacific Grove, 2. Auflage.
- Cramér, H. (1946). *Mathematical methods of statistics*. Princeton University Press, Princeton.
- Fisher, R. A. (1924). The Conditions Under Which  $\chi^2$  Measures the Discrepancy Between Observation and Hypothesis. *Journal of the Royal Statistical Society*, 87(3):442–450.
- Forster, O. (2006). *Analysis 2*. Vieweg, Wiesbaden, 7. Auflage.
- Good, I. J. (1961). The Frequency Count of a Markov Chain and the Transition to Continuous Time. *The Annals of Mathematical Statistics*, 32(1):41–48.
- Handy, T. C. (2009). *Brain signal analysis: Advances in neuroelectric and neuromagnetic methods*. MIT Press, Cambridge.
- Held, L. (2008). *Methoden der statistischen Inferenz: Likelihood und Bayes*. Spektrum Akademischer Verlag, Heidelberg.
- Kersting, G. und Wakolbinger, A. (2008). *Elementare Stochastik*. Birkhäuser, Basel.
- Lehmann, D., Faber, P. L., Galderisi, S., Herrmann, W. M., Kinoshita, T., Koukkou, M., Mucci, A., Pascual-Marqui, R. D., Saito, N., Wackermann, J., Winterer, G., und Koenig, T. (2005). EEG microstate duration and syntax in acute, medication-naïve, first-episode schizophrenia: a multi-center study. *Psychiatry Research: Neuroimaging*, 138(2):141–156.
- Lehmann, D. und Michel, C. M. (2011). EEG-defined functional microstates as basic building blocks of mental processes. *Clinical Neurophysiology*, 122(6):1073–1074.
- Lehmann, D., Ozaki, H., und Pal, I. (1987). EEG alpha map series: brain micro-states by space-oriented adaptive segmentation. *Electroencephalography and Clinical Neurophysiology*, 67(3):271–288.

- Lehmann, D., Strik, W., Henggeler, B., Koenig, T., und Koukkou, M. (1998). Brain electric microstates and momentary conscious mind states as building blocks of spontaneous thinking: I. Visual imagery and abstract thoughts. *International Journal of Psychophysiology*, 29(1):1–11.
- Meintrup, D. und Schäffler, S. (2005). *Stochastik: Theorie und Anwendungen*. Springer, Berlin.
- Nunez, P. und Srinivasan, R. (2007). Electroencephalogram. *Scholarpedia*, 2(2):1348.
- Pearson, K. (1900). X. On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *Philosophical Magazine Series 5*, 50(302):157–175.
- Schlegel, F., Lehmann, D., Faber, P. L., Milz, P., und Gianotti, L. R. R. (2012). EEG Microstates During Resting Represent Personality Differences. *Brain Topography*, 25(1):20–26.
- Wackermann, J., Lehmann, D., Michel, C., und Strik, W. (1993). Adaptive segmentation of spontaneous EEG map series into spatially defined microstates. *International Journal of Psychophysiology*, 14(3):269–283.
- Wasserman, L. (2004). *All of statistics: A concise course in statistical inference*. Springer, New York.
- Weeß, H.-G. und Landwehr, R. (2009). Phänomenologie, Funktion und Physiologie des Schlafes. *PiD - Psychotherapie im Dialog*, 10(02):101–106.
- Wolf, F. M. (1986). *Meta-analysis: Quantitative methods for research synthesis*. Sage Publications, Beverly Hills.

## Abbildungsverzeichnis

1.1	Exemplarische Darstellung der Ermittlung einer Sequenz $x$ anhand eines EEGs und der Ausprägungsstärke des elektrischen Feldes sowie der Ermittlung einer entsprechenden Markov-Kette $X$ in der parallelen Konstruktion zum SMI-Modell . . . . .	10
3.1	Exemplarische Darstellung des Hintergrundmodells . . . . .	19
3.2	Übergangswahrscheinlichkeiten aller Probanden getrennt nach Schlafstadium . . . . .	20
3.3	Maximale Abstände $h^{(J,Z)}$ pro Proband getrennt nach Schlafstadium . . . . .	22
3.4	Maximale Abweichung der Differenzen $d_A, d_B, d_C, d_D$ voneinander pro Proband getrennt nach Schlafstadium . . . . .	23
3.5	Geschätzte Übergangswahrscheinlichkeiten im Schlafstadium N2 für Proband 2 und Proband 5 . . . . .	24
3.6	Exemplarische Darstellung des Hintergrundmodells mit Variablenbezeichnungen . . . . .	28
4.1	Schätzer $\hat{p}_1$ für simulierte Daten . . . . .	40
4.2	Schätzer $\hat{\pi}^{(u)}$ für simulierte Daten . . . . .	45
4.3	Vergleich des Schätzers $\hat{\pi}^{(u)}$ mit dem ML-Schätzer $\hat{\pi}^{\text{ML}}$ . . . . .	51
4.4	Vergleich des Schätzers $\hat{p}_1$ mit dem ML-Schätzer $\hat{p}_1^{\text{ML}}$ . . . . .	52
4.5	Aus den Daten berechnete Schätzer $\hat{p}_r^{(J,Z)}$ getrennt nach Zustand $r \in S$ und nach Schlafstadium $Z$ . . . . .	54
4.6	Aus den Daten berechnete Schätzer $\hat{\pi}_{(J,Z)}^{(u)}$ getrennt nach Schlafstadium $Z$ . . . . .	55
4.7	Geschätzte Übergangswahrscheinlichkeiten $\hat{p}_{rs}$ im SMI-Modell für die Probanden 2 und 5 im Schlafstadium N2 im Vergleich zu den geschätzten Übergangswahrscheinlichkeiten $\hat{p}_{rs}^{(v)}$ des vollen Modells . . . . .	56
5.1	Graphen von $p_r(1 - p_r)$ und $(2 - \pi)/\pi$ . . . . .	62
5.2	Graphen von $f(p_1)$ auf $[0.01, 0.99]$ für verschiedene feste Werte von $\pi$ . . . . .	89
5.3	Graphen von $f(\pi)$ auf $[0.01, 0.99]$ für verschiedene feste Werte von $p_1$ . . . . .	90
5.4	Konturplot von $f(\pi, p_1)$ mit eingezeichnetem Graphen von $\pi_{\max}(p_1)$ . . . . .	91
5.5	Konturplot von $f(p_1, p_2)$ für $\pi = 0.5$ und von $(1 - \sum_k p_k^2)^{-2}$ . . . . .	92
5.6	Konturplot von $f(p_1, p_2)$ und vom Zähler von $f(p_1, p_2)$ für verschiedene Werte von $\pi$ . . . . .	93
5.7	Konturplot von $\pi_{\max}(p_1, p_2)$ mit Querschnitten für verschiedene Werte von $p_2$ . . . . .	94
5.8	Konturplot von $f(p_2, p_3)$ für verschiedene Werte von $p_1$ und von $\pi$ . . . . .	95
6.1	Anpassungstests durch Simulation für Proband 2, alle erreichten Schlafstadien . . . . .	103
6.2	Anpassungstests durch Simulation für Proband 5, alle erreichten Schlafstadien . . . . .	104
6.3	Übersicht über die p-Werte aller mithilfe von Simulation ausgeführten Anpassungstests für das SMI-Modell . . . . .	105



6.4	Anpassungstests durch Simulation für alle Probanden simultan, alle Schlafstadien . . . . .	107
6.5	Histogramme der p-Werte $\mathbf{p}^{\{k,l\}}$ getrennt nach $k = 1, 2, 3$ . . . . .	109
6.6	Übersicht über alle $\Delta_{rs}^{(J,Z)}$ für jeden Übergang $(r, s)$ aller Probanden $J$ getrennt nach Schlafstadium $Z$ . . . . .	116
6.7	Zerlegung eines Übergangs im SMI-Modell in ein zweistufiges Zufallsexperiment . . . . .	120
6.8	Zerlegung eines Übergangs im erweiterten Modell in ein dreistufiges Zufallsexperiment für den Ausgangszustand $z_1$ . . . . .	120
6.9	Geschätzte Übergangswahrscheinlichkeiten $\hat{p}_{rs}^{(\text{erw}),\text{ML}}$ im Erweiterten SMI-Modell für die Probanden 2 und 5 im Schlafstadium N2 im Vergleich zu den geschätzten Übergangswahrscheinlichkeiten $\hat{p}_{rs}$ des SMI-Modells und $\hat{p}_{rs}^{(v)}$ des vollen Modells . . . . .	123
6.10	Übersicht über die p-Werte aller Anpassungstests für das Erweiterte SMI-Modell . . . . .	125