

Nichtparametrische Zweistichproben-Tests für multivariate stochastische Integralordnungen

Von der Fakultät für Mathematik und Physik
der Gottfried Wilhelm Leibniz Universität Hannover
zur Erlangung des Grades
Doktor der Naturwissenschaften
Dr. rer. nat.
genehmigte Dissertation
von

Dipl.-Math. Dennis Behnke
geboren am 10.05.1984 in Burgwedel

2012

Referent: Prof. Dr. L. Baringhaus

Korreferent: Prof. Dr. N. Henze

Tag der Promotion: 2. Mai 2012

Zusammenfassung

Die stochastischen Integralordnungen bilden eine Klasse von partiellen Ordnungen auf der Menge der Verteilungen reeller Zufallsvektoren. Diese Ordnungen können jeweils mittels einer geeigneten Klasse messbarer Funktionen charakterisiert werden. Ist \mathcal{F} eine solche Klasse messbarer und reellwertiger Funktionen und sind X und Y Zufallsvektoren, so definieren wir $X \leq_{\mathcal{F}} Y \Leftrightarrow \forall f \in \mathcal{F} : Ef(X) \leq Ef(Y)$ und nennen $\leq_{\mathcal{F}}$ die von \mathcal{F} erzeugte stochastische Integralordnung. Ein klassisches Beispiel ist die von den Indikatorfunktionen $\mathbf{1}_{(t, \infty)}, t \in \mathbb{R}$, erzeugte gewöhnliche stochastische Ordnung.

Die vorliegende Arbeit befasst sich mit nichtparametrischen Zweistichproben-Tests zur Überprüfung der Hypothese $H : X \leq_{\mathcal{F}} Y$ gegen die Alternative $K : X \not\leq_{\mathcal{F}} Y$ im Fall der Laplace-Ordnung und der Konkordanz-Ordnung. Dabei werden die Grenzverteilungen der vorgestellten Testgrößen bei Gültigkeit der Hypothese und unter Folgen von benachbarten Alternativen bestimmt. Für Teilhypothesen werden Effizienzvergleiche durchgeführt.

Es zeigt sich, dass die Grenzverteilung der Testgrößen bei Gültigkeit der Hypothese sowohl für die Laplace-, als auch für die Konkordanz-Ordnung nicht verteilungsfrei ist. Ein Bootstrapverfahren wird vorgeschlagen und das Verhalten der Testgrößen unter dem Einfluss des Bootstrap untersucht. Die resultierenden Tests sind konsistent gegen alle Alternativen und halten asymptotisch das vorgegebene Testniveau ein. Die theoretischen Ergebnisse werden von einer Simulationsstudie begleitet.

Schlagwörter: Laplace-Ordnung, Konkordanz-Ordnung, Zweistichproben-Tests.

Abstract

Integral stochastic orders form an important class of partial orders on the set of distributions of real valued random vectors. These orders can be characterized by a suitable class of real valued measurable functions. To be more precise, given such a class \mathcal{F} of real valued measurable functions and two real valued random vectors X, Y , we define $X \leq_{\mathcal{F}} Y \Leftrightarrow \forall f \in \mathcal{F} : Ef(X) \leq Ef(Y)$ and call $\leq_{\mathcal{F}}$ the integral stochastic order generated by \mathcal{F} . A classical example is the usual stochastic order in the univariate case generated by the class of indicator functions $\mathbf{1}_{(t, \infty)}, t \in \mathbb{R}$.

This thesis deals with the multivariate nonparametric two-sample problem of testing the null hypothesis $H : X \leq_{\mathcal{F}} Y$ against the general alternative $K : X \not\leq_{\mathcal{F}} Y$ for both the Laplace transform order and the concordance order. Kolmogorov-Smirnov- and Cramér-von Mises-type test statistics are considered and the limiting distributions of the test statistics under the null hypothesis and under contiguous alternatives are obtained. For specific alternatives the relative approximate Bahadur efficiencies are computed.

Since the limiting distributions of the test statistics are not distribution-free under the null hypothesis a bootstrap procedure is proposed to estimate the critical values. The resulting tests are consistent against all alternatives and asymptotically of the preassigned level.

Numerical results are given to illustrate the theoretical results.

Keywords: Laplace transform order, concordance order, two-sample tests.

Inhaltsverzeichnis

1	Die Laplace-Ordnung	1
1.1	Einführung	1
1.2	Eigenschaften von $\leq_{\mathcal{L}}$	3
2	Test auf Laplace-Ordnung	11
2.1	Das Testproblem und die Testgrößen	11
2.2	Asymptotik	13
2.3	Kritische Werte über den Bootstrap	25
2.4	Invarianz des Testproblems	36
2.5	Simulationen	43
3	Benachbarte Alternativen	49
3.1	Asymptotisches Verhalten	50
3.2	Approximative Bahadursteigungen für eine Teilhypothese	60
4	Die Konkordanz-Ordnung	69
4.1	Grundbegriffe	69
4.2	Eigenschaften und Beziehungen zu anderen Ordnungen	71
4.3	Elliptisch symmetrische Verteilungen	76
5	Test auf Konkordanz-Ordnung	81
5.1	Supremums- und Integraltestgrößen	81
5.1.1	Ein parametrischer Test	82
5.2	Asymptotik	86
5.3	Konsistenz und Einhaltung des Testniveaus	94
5.4	Implementierung und numerische Ergebnisse	103

A	Stochastische Integralordnungen	115
B	Hoffmann-Jørgensen-Theorie	119
B.1	Schwache Konvergenz	119
B.2	Empirische Prozesse in $l^\infty(T)$	121
B.2.1	VC- und Donsker-Klassen	123
B.2.2	Konvergenzverhalten bei Dreiecksschemata	126
C	Benachbarte Alternativen	127
D	Verschiedenes	129
D.1	Gaußmaße	129
D.2	Separable Prozesse	130
D.3	Hilfssätze	130
	Literaturverzeichnis	135

Kapitel 1

Die Laplace-Ordnung

1.1 Einführung

Sei \mathbf{X} ein d -dimensionaler Zufallsvektor mit Werten in $\mathbb{R}_{\geq 0}^d = [0, \infty)^d$. Die *multivariate Laplacetransformierte* $\mathcal{L}_{\mathbf{X}}$ von $\mathbf{X} = (X_1, \dots, X_d)'$ ist definiert als

$$\mathcal{L}_{\mathbf{X}}(\mathbf{t}) = E\left(e^{-\mathbf{t}'\mathbf{X}}\right) = E\left(\prod_{i=1}^d e^{-t_i X_i}\right)$$

für $\mathbf{t} = (t_1, \dots, t_d)' \in \mathbb{R}_{\geq 0}^d$. Die Existenz der auftretenden Erwartungswerte für komponentenweise nicht-negative Zufallsvektoren ist dabei wegen der Beschränktheit von e^{-t} trivial. Mit Hilfe der (multivariaten) Laplacetransformierten lässt sich eine partielle Ordnung auf der Menge der Verteilungen d -dimensionaler Zufallsvektoren mit nicht-negativen Komponenten einführen (siehe auch [Sto83]). Wir geben die Definition dieser Ordnung sogleich an.

Zuvor vereinbaren wir zur Vereinfachung der Notation allerdings einige Schreibweisen: Fett gedruckte Buchstaben bezeichnen in der Folge stets Vektoren und für Vektoren $\mathbf{s} = (s_1, \dots, s_d)'$, $\mathbf{t} = (t_1, \dots, t_d)' \in \mathbb{R}^d$ setzen wir außerdem $\mathbf{t} \geq 0 \Leftrightarrow \mathbf{t} \in \mathbb{R}_{\geq 0}^d$ und $\mathbf{s} \leq \mathbf{t} \Leftrightarrow s_k \leq t_k \forall k \in \{1, \dots, d\}$. Hat der oben genannte Zufallsvektor \mathbf{X} die Verteilungsfunktion F , so schreiben wir für die (multivariate) Laplacetransformierte von \mathbf{X} auch \mathcal{L}_F .

1.1.1. Definition. Seien \mathbf{X} und \mathbf{Y} Zufallsvektoren mit Werten in $\mathbb{R}_{\geq 0}^d$. Dann heißt \mathbf{X} kleiner oder gleich \mathbf{Y} in der (multivariaten) Laplace-Ordnung, falls

$$\mathcal{L}_{\mathbf{X}}(\mathbf{t}) \geq \mathcal{L}_{\mathbf{Y}}(\mathbf{t}) \quad \text{für alle } \mathbf{t} \geq 0 \tag{1.1.1}$$

gilt. Wir schreiben in diesem Fall $\mathbf{X} \leq_{\mathcal{L}} \mathbf{Y}$ und verwenden synonym auch die Bezeichnung $F \leq_{\mathcal{L}} G$, falls \mathbf{X} die Verteilungsfunktion F und \mathbf{Y} die Verteilungsfunktion G hat.

Die Definition der Laplace-Ordnung ist in der Literatur nicht einheitlich. Einige Autoren fordern in (1.1.1) die umgekehrte Ungleichung, siehe z.B. [Kle83] und [Den01]. Neben der in 1.1.1 getroffenen Definition wird sich auch die äquivalente Definition der Laplace-Ordnung als *stochastische Integralordnung* als nützlich erweisen (vergleiche auch Kapitel A des Anhangs). Es gilt offenbar $\mathbf{X} \leq_{\mathcal{L}} \mathbf{Y}$ genau dann, wenn $Ef(\mathbf{X}) \leq Ef(\mathbf{Y})$ für alle $f \in \mathcal{F}$, mit $\mathcal{F} := \{f_{\mathbf{t}}, \mathbf{t} \in \mathbb{R}_{\geq 0}^d\}$ und

$$f_{\mathbf{t}} := \begin{cases} \mathbb{R}^d & \rightarrow \mathbb{R} \\ \mathbf{x} & \mapsto -e^{\mathbf{t}'\mathbf{x}} \end{cases}$$

ist.

Der Vergleich von Laplacetransformierten ist in vielen Bereichen der Stochastik von Interesse, wie etwa der Warteschlangentheorie, der Zuverlässigkeitstheorie oder der Nutzentheorie. Eine ausführliche Auflistung verschiedener Anwendungen der Laplace-Ordnung findet sich beispielsweise in [AKJ91] oder [Den01]. Wir beschränken uns auf ein motivierendes Beispiel im eindimensionalen Fall:

1.1.2. Beispiel. Wir betrachten das Wartesystem $G/M/1$ mit den unabhängigen und identisch verteilten Zwischenankunftszeiten $T_1 > 0, T_2 > 0, \dots$, und den davon und untereinander unabhängigen, je mit Parameter $\lambda > 0$ exponentialverteilten Bedienzeiten S_1, S_2, \dots . Es gelte $E(T_1) < \infty$. Sei weiter $Q := (Q(t), t \geq 0)$, wobei die Zufallsvariable $Q(t)$ für jedes $t \geq 0$ die Anzahl der Kunden im System zum Zeitpunkt t angebe. Ferner sei $A_n := \sum_{k=1}^n T_k$, $n \in \mathbb{N}$, der Zeitpunkt der Ankunft des n -ten Kunden im System. Bekanntlich ist dann $(Y_n, n \in \mathbb{N})$ mit

$$Y_n := \lim_{t \downarrow 0} Q(A_n - t)$$

eine homogene, irreduzible und aperiodische Markovkette, die genau dann eine stationäre Verteilung $\pi = (\pi_j, j \in \mathbb{N}_0)$ besitzt, falls $\frac{1}{\lambda E(T_1)} < 1$ ist. Genauer ist dann

$$\pi_j = (1 - \eta)\eta^j, \quad j \in \mathbb{N}_0, \tag{1.1.2}$$

also π eine geometrische Verteilung mit dem Parameter $\eta \in (0, 1)$, welcher sich als die im Intervall $(0, 1)$ eindeutig bestimmte Lösung der Gleichung

$$\eta = \mathcal{L}_{T_1}(\lambda(1 - \eta))$$

ergibt.

Seien nun $G/M/1$ und $\tilde{G}/M/1$ zwei Wartesysteme mit je derselben Bedienzeitverteilung, sowie verschiedenen Zwischenankunftsverteilungen mit den zugehörigen Verteilungsfunktionen F bzw. \tilde{F} . Die Laplace-Ordnung der Zwischenankunftszeiten lässt dann Rückschlüsse auf die stationären Verteilungen zu. Aufgrund der Monotonie der Laplace-transformierten folgt aus $F \leq_{\mathcal{L}} \tilde{F}$ nämlich sofort $\eta \geq \tilde{\eta}$. Existieren die entsprechenden stationären Verteilungen $\pi, \tilde{\pi}$, so implizieren die Ordnung der Parameter und (1.1.2) die stochastische Ordnung $\pi \geq_{st} \tilde{\pi}$.

1.2 Eigenschaften von $\leq_{\mathcal{L}}$

Bevor wir uns im Hauptteil dieser Arbeit mit statistischen Tests für das Testproblem der Laplace-Ordnung befassen, wollen wir zuvor der Vollständigkeit halber einige grundlegende Eigenschaften der Laplace-Ordnung angeben. Einige dieser Eigenschaften werden bei den späteren Betrachtungen benötigt, wohingegen andere von eigenständigem Interesse sind.

Wir orientieren uns bei den folgenden Ausführungen vorwiegend an den Ergebnissen in [Sto83] und [SS07]. Die meisten dieser Aussagen sind evident und werden hier nicht bewiesen, in den übrigen Fällen geben wir kurz eigene Argumente an. Die elementaren Gegenbeispiele, die in 1.2.6 gegeben werden, sind in dieser Form in der Literatur bisher scheinbar nicht bekannt.

Zunächst ist die Relation $\leq_{\mathcal{L}}$ reflexiv auf der Menge der Verteilungen nicht-negativer Zufallsvektoren und aufgrund der Endlichkeit der Erwartungswerte $E\left(e^{-t'X}\right)$ für alle $t \in \mathbb{R}_{\geq 0}^d$ und alle Zufallsvariablen mit Werten in $\mathbb{R}_{\geq 0}^d$ auch transitiv. Der Eindeigkeitsatz für Laplacetransformierte liefert ferner die Antisymmetrie von $\leq_{\mathcal{L}}$, so dass $\leq_{\mathcal{L}}$ in der Tat eine (partielle) Ordnung auf der Menge der Verteilungen von komponentenweise nicht-negativen Zufallsvektoren definiert. Viele Eigenschaften der multivariaten Laplace-Ordnung ergeben sich aus den entsprechenden Eigenschaften der univariaten Laplace-Ordnung. Der Grund dafür ist die folgende Charakterisierung von $\leq_{\mathcal{L}}$ über die Ordnung eindimensionaler Zufallsvariablen, die sich - wie auch die Aussage des darauf folgenden Lemmas - direkt anhand der Definition verifizieren lässt.

Im Folgenden seien die betrachteten Zufallsvektoren stets komponentenweise nicht-negativ.

1.2.1. Lemma. *Es gilt*

$$\mathbf{X} \leq_{\mathcal{L}} \mathbf{Y} \Leftrightarrow \mathbf{a}'\mathbf{X} \leq_{\mathcal{L}} \mathbf{a}'\mathbf{Y} \quad \text{für alle } \mathbf{a} \in \mathbb{R}_{\geq 0}^d.$$

□

1.2.2. Lemma.

a) Die Laplace-Ordnung ist abgeschlossen unter Marginalbildung: Für Zufallsvektoren $\mathbf{X} = (X_1, \dots, X_d)$, $\mathbf{Y} = (Y_1, \dots, Y_d)$ gilt

$$\mathbf{X} \leq_{\mathcal{L}} \mathbf{Y} \Rightarrow (X_{i_1}, \dots, X_{i_k}) \leq_{\mathcal{L}} (Y_{i_1}, \dots, Y_{i_k})$$

für alle paarweise verschiedenen $i_1, \dots, i_k \in \{1, \dots, d\}$, $k = 1, \dots, d$.

b) Die Laplace-Ordnung ist abgeschlossen unter Adjunktion: Für unabhängige Zufallsvektoren $\mathbf{X}_1, \dots, \mathbf{X}_n$ und unabhängige Zufallsvektoren $\mathbf{Y}_1, \dots, \mathbf{Y}_n$ mit $\dim(\mathbf{X}_i) = \dim(\mathbf{Y}_i) = k_i \in \mathbb{N}$ und $\mathbf{X}_i \leq_{\mathcal{L}} \mathbf{Y}_i$, $i = 1, \dots, n$ gilt

$$(\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n) \leq_{\mathcal{L}} (\mathbf{Y}_1, \mathbf{Y}_2, \dots, \mathbf{Y}_n).$$

c) Die Laplace-Ordnung ist abgeschlossen unter Faltungen: Für unabhängige, d -dimensionale Zufallsvektoren $\mathbf{X}_1, \dots, \mathbf{X}_n$ und unabhängige, d -dimensionale Zufallsvektoren $\mathbf{Y}_1, \dots, \mathbf{Y}_n$ mit $\mathbf{X}_i \leq_{\mathcal{L}} \mathbf{Y}_i$ für alle $i = 1, \dots, n$ gilt

$$\sum_{i=1}^n \mathbf{X}_i \leq_{\mathcal{L}} \sum_{i=1}^n \mathbf{Y}_i.$$

□

Die Laplace-Ordnung impliziert außerdem im folgenden Sinne die Ordnung zwischen den Komponenten der Erwartungswerte der betrachteten Zufallsvektoren:

1.2.3. Lemma. *Es gilt*

$$\mathbf{X} = (X_1, \dots, X_d) \leq_{\mathcal{L}} \mathbf{Y} = (Y_1, \dots, Y_d) \Rightarrow E(X_i) \leq E(Y_i), \quad i = 1, \dots, d,$$

sofern die Erwartungswerte endlich sind.

Beweis. Das vorige Lemma liefert zunächst $X_i \leq_{\mathcal{L}} Y_i$, $i = 1, \dots, d$, also auch

$$\forall t > 0: \quad \frac{1 - \mathcal{L}_{X_i}(t)}{t} \leq \frac{1 - \mathcal{L}_{Y_i}(t)}{t}.$$

Die Laplacetransformierten sind auf $(0, \infty)$ beliebig oft differenzierbar, so dass

$$\lim_{t \downarrow 0} \frac{1 - \mathcal{L}_{X_i}(t)}{t} = - \lim_{t \downarrow 0} \mathcal{L}'_{X_i}(t)$$

folgt. Mit $\mathcal{L}'_{X_i}(t) = -E(X_i e^{-tX_i})$ und dem Satz von der majorisierten Konvergenz ergibt sich schließlich die Behauptung. \square

Ähnliche Aussagen wie in Lemma 1.2.2 c) erhält man ebenfalls für Summen mit einer zufälligen Anzahl von Summanden.

1.2.4. Lemma. *Es seien $\mathbf{M} = (M_1, \dots, M_d), \mathbf{N} = (N_1, \dots, N_d)$ Zufallsvektoren mit Werten in \mathbb{N}_0^d .*

a) *Für jedes $j = 1, \dots, d$ sei $\{X_{ij}\}_{i \in \mathbb{N}}$ eine Folge von unabhängigen und identisch verteilten Zufallsvariablen. Für jedes $j = 1, \dots, d$ seien \mathbf{M} und $\{X_{ij}\}_{i \in \mathbb{N}}$ unabhängig und für jedes $j = 1, \dots, d$ seien \mathbf{N} und $\{X_{ij}\}_{i \in \mathbb{N}}$ unabhängig. Sei $\mathbf{M} \leq_{\mathcal{L}} \mathbf{N}$. Dann gilt*

$$\left(\sum_{i=1}^{M_1} X_{i1}, \dots, \sum_{i=1}^{M_d} X_{id} \right) \leq_{\mathcal{L}} \left(\sum_{i=1}^{N_1} X_{i1}, \dots, \sum_{i=1}^{N_d} X_{id} \right),$$

wobei die leere Summe gleich Null gesetzt wird.

b) *Seien $\{\mathbf{X}_i\}_{i \in \mathbb{N}}, \{\mathbf{Y}_i\}_{i \in \mathbb{N}}$ zwei Folgen d -dimensionaler Zufallsvektoren mit $\mathbf{X}_i \leq_{\mathcal{L}} \mathbf{Y}_i$ für alle $i \in \mathbb{N}$. Seien $\mathbf{N}, \{\mathbf{X}_i\}_{i \in \mathbb{N}}, \{\mathbf{Y}_i\}_{i \in \mathbb{N}}$ unabhängig. Es bezeichne X_{ij} bzw. Y_{ij} , $j = 1, \dots, d$, die j -te Komponente von \mathbf{X}_i bzw. \mathbf{Y}_i . Dann gilt*

$$\left(\sum_{i=1}^{N_1} X_{i1}, \dots, \sum_{i=1}^{N_d} X_{id} \right) \leq_{\mathcal{L}} \left(\sum_{i=1}^{N_1} Y_{i1}, \dots, \sum_{i=1}^{N_d} Y_{id} \right),$$

wobei die leere Summe gleich Null gesetzt wird.

Beweis.

a) Wir berechnen zunächst für $\mathbf{t} = (t_1, \dots, t_d)' \in \mathbb{R}_{\geq 0}^d$

$$\begin{aligned} \mathcal{L}_{\left(\sum_{i=1}^{M_1} X_{i1}, \dots, \sum_{i=1}^{M_d} X_{id}\right)}(\mathbf{t}) &= E \left(\exp \left(- \sum_{j=1}^d t_j \sum_{i=1}^{M_j} X_{ij} \right) \right) \\ &= \prod_{j=1}^d E \left(E \left(\prod_{i=1}^{M_j} \exp(-t_j X_{ij}) \middle| M_j \right) \right) \\ &= \prod_{j=1}^d E \left(\exp(-t_j X_{1j})^{M_j} \right) = E \left(\prod_{j=1}^d (\mathcal{L}_{X_{1j}}(t_j))^{M_j} \right) \\ &= E \left(\prod_{j=1}^d s_j^{M_j} \right), \end{aligned} \tag{1.2.1}$$

mit $s_j = \mathcal{L}_{X_{1j}}(t_j) \in (0, 1]$, $j = 1, \dots, d$. Nach Voraussetzung gilt $\mathbf{M} \leq_{\mathcal{L}} \mathbf{N}$, also $E\left(\prod_{j=1}^d e^{-t_j M_j}\right) \geq E\left(\prod_{j=1}^d e^{-t_j N_j}\right)$ für alle nicht-negativen Werte t_j , $j = 1, \dots, d$. Insbesondere folgt mit $t_j = -\ln s_j$, $j = 1, \dots, d$, dass

$$E\left(\prod_{j=1}^d s_j^{M_j}\right) \geq E\left(\prod_{j=1}^d s_j^{N_j}\right)$$

gilt. Zusammen mit (1.2.1) ergibt sich daher

$$\mathcal{L}_{\left(\sum_{i=1}^{M_1} X_{i1}, \dots, \sum_{i=1}^{M_d} X_{id}\right)}(\mathbf{t}) \geq \mathcal{L}_{\left(\sum_{i=1}^{N_1} X_{i1}, \dots, \sum_{i=1}^{N_d} X_{id}\right)}(\mathbf{t}), \quad \text{für alle } \mathbf{t} \geq 0.$$

b) Der Beweis kann ähnlich zu Teil a) geführt werden. □

Der Beweis von Lemma 1.2.4 a) zeigt, dass für nicht-negative ganzzahlige Zufallsvektoren \mathbf{X}, \mathbf{Y} die Ordnung bezüglich $\leq_{\mathcal{L}}$ die entsprechende Ordnung zwischen den erzeugenden Funktionen impliziert. Man überzeugt sich leicht, dass darüber hinaus dann ebenso die Rückrichtung, also insgesamt

$$\mathbf{X} \leq_{\mathcal{L}} \mathbf{Y} \Leftrightarrow E\left(\prod_{i=1}^d t_i^{X_i}\right) \geq E\left(\prod_{i=1}^d t_i^{Y_i}\right)$$

für alle $\mathbf{t} = (t_1, \dots, t_d)' \in (0, 1)^d$ zutrifft.

Wir wollen im Folgenden $\leq_{\mathcal{L}}$ in den Kontext anderer gängiger stochastischer Integralordnungen einordnen und weisen für die im nachstehenden Lemma verwendeten Notationen dazu erneut auf Kapitel A des Anhangs hin.

1.2.5. Lemma. *Es gilt*

a) $\mathbf{X} \leq_{st} \mathbf{Y} \Rightarrow \mathbf{X} \leq_{\mathcal{L}} \mathbf{Y}$,

b) $\mathbf{X} \leq_{icv} \mathbf{Y} \Rightarrow \mathbf{X} \leq_{\mathcal{L}} \mathbf{Y}$,

c) $\mathbf{X} \leq_{dir-cx} \mathbf{Y} \Rightarrow \mathbf{Y} \leq_{\mathcal{L}} \mathbf{X}$,

d) $\mathbf{X} \leq_{sm} \mathbf{Y} \Rightarrow \mathbf{Y} \leq_{\mathcal{L}} \mathbf{X}$.

Beweis. *Die Behauptungen in a) und b) folgen sofort, da die Funktionen $f_{\mathbf{t}}(\mathbf{x}) = -e^{-\mathbf{t}'\mathbf{x}}$ offensichtlich für jedes $\mathbf{t} \in \mathbb{R}_{\geq 0}^d$ monoton wachsend und konkav sind. Die Aussagen c) und d) folgen aus der Tatsache, dass die Funktionen $g_{\mathbf{t}}(\mathbf{x}) = e^{-\mathbf{t}'\mathbf{x}}$ für jedes $\mathbf{t} \in \mathbb{R}_{\geq 0}^d$ aufgrund von $\frac{\partial^2}{\partial x_i \partial x_j} g_{\mathbf{t}}(\mathbf{x}) \geq 0$, $i, j = 1, \dots, d$, komponentenweise konvex und supermodular sind. □*

Für bestimmte Verteilungsfamilien gilt in Teil b) des vorherigen Lemmas auch die umgekehrte Relation, also in diesen Fällen die Äquivalenz der beiden Ordnungen \leq_{icv} und $\leq_{\mathcal{L}}$. Sind beispielsweise $X \sim \text{Exp}(\lambda)$ und $Y \sim \text{Exp}(\mu)$ jeweils exponentialverteilt mit den Parametern $\mu, \lambda > 0$, so gilt $X \leq_{\mathcal{L}} Y \Leftrightarrow X \leq_{icv} Y$ (es gilt offenbar $X \leq_{st} Y \Leftrightarrow \lambda \geq \mu$).

Im Allgemeinen ist die Laplace-Ordnung allerdings schwächer als \leq_{icv} , wie die folgenden Gegenbeispiele illustrieren.

1.2.6. Beispiel ($X \leq_{\mathcal{L}} Y \not\Leftarrow X \leq_{icv} Y$). Es seien X und Y diskret verteilte Zufallsvariablen mit $P(X = 1) = P(X = 3) = 0.3$, $P(X = 4) = 0.4$ und $P(Y = 1) = 0.2$, $P(Y = 2) = 0.3$, $P(Y = 4) = 0.5$.

a) Mit $\Delta(t) := \mathcal{L}_X(t) - \mathcal{L}_Y(t)$ für $t \geq 0$ gilt

$$\Delta(t) = 0.1(e^{-t} - e^{-4t}) + 0.3(e^{-3t} - e^{-2t}),$$

und $\Delta(t) \geq 0$ für alle $t \geq 0$ genau dann, wenn $p(s) := 0.1s - 0.1s^4 + 0.3s^3 - 0.3s^2 \geq 0$ für alle $s \in (0, 1]$ gilt, was zu $q(s) := 0.1 - 0.1s^3 + 0.3s^2 - 0.3s \geq 0 \forall s \in (0, 1]$ äquivalent ist. Es ist $q(0) = 0.1 > 0$, $q(1) = 0$ und $q'(s) < 0$, $s \in (0, 1)$, und daher $X \leq_{\mathcal{L}} Y$. Die Funktion $f(x) := \min(0, x - 3)$ ist aber offensichtlich monoton wachsend und konkav mit $Ef(X) = -0.6 > -0.7 = Ef(Y)$, was auf $X \not\leq_{icv} Y$ führt.

(Wir merken außerdem an, dass ebenfalls $Y \not\leq_{icv} X$ gilt, denn aus $Y \leq_{icv} X$ würde aufgrund von Teil b) des vorherigen Lemmas bereits $Y \leq_{\mathcal{L}} X$, also $X \stackrel{v}{=} Y$ folgen).

b) Das Beispiel aus Teil a) lässt sich etwa auf die folgende Weise verallgemeinern. Sei $U \sim \mathfrak{R}([0, 1])$ gleichverteilt auf $[0, 1]$, U unabhängig von X und Y . Wir betrachten die Zufallsvariablen

$$V_{\varepsilon} := X + \varepsilon U, \quad W_{\varepsilon} := Y + \varepsilon U,$$

mit $\varepsilon \geq 0$. Mit Teil a) und der Aussage in Lemma 1.2.2 c) ergibt sich unmittelbar $V_{\varepsilon} \leq_{\mathcal{L}} W_{\varepsilon}$, $\varepsilon \geq 0$. Für die Funktion $f(x) := \min(0, x - 3)$ gilt dann

$$Ef(V_{\varepsilon}) = \sum_{k \in \{1, 3, 4\}} \left(\int_0^1 \min(0, k + \varepsilon y - 3) dy \right) P(X = k) = \begin{cases} -0.6 \varepsilon^{-1}, & \varepsilon \geq 2, \\ 0.15 \varepsilon - 0.6, & \varepsilon \leq 2, \end{cases}$$

und

$$\begin{aligned}
 Ef(W_\varepsilon) &= \sum_{k \in \{1,2,4\}} \left(\int_0^1 \min(0, k + \varepsilon y - 3) dy \right) P(Y = k) \\
 &= \begin{cases} -0.55 \varepsilon^{-1}, & \varepsilon \geq 2, \\ 0.1 \varepsilon - 0.15 \varepsilon^{-1} - 0.4, & \varepsilon \in [1, 2], \\ 0.25 \varepsilon - 0.7, & \varepsilon \leq 1. \end{cases}
 \end{aligned}$$

Insbesondere gilt $Ef(V_\varepsilon) > Ef(W_\varepsilon)$ für alle $\varepsilon < 1$ und somit insbesondere $V_\varepsilon \not\leq_{icv} W_\varepsilon$ für alle $\varepsilon < 1$.

Zum Abschluss dieses Abschnitts wollen wir noch einige Beispiele gängiger Verteilungsfamilien angeben, bei denen die Ordnung geeigneter Parameter Rückschlüsse auf die Laplace-Ordnung der entsprechenden Zufallsvariablen zulässt.

1.2.7. Beispiele.

a) Sei $X \sim \mathfrak{R}([a, b])$ gleichverteilt auf $[a, b]$, $0 \leq a < b$. Dann ist

$$\mathcal{L}_X(t) = \frac{e^{-ta} - e^{-tb}}{t(b-a)}, \quad t > 0,$$

und es gilt für $X \sim \mathfrak{R}([a, b])$ und $Y \sim \mathfrak{R}([c, d])$ die Implikation

$$a \leq c, b \leq d \Rightarrow X \leq_{\mathcal{L}} Y.$$

b) Sei $X \sim \Gamma(\alpha, \beta)$ gammaverteilt mit den Parametern $\alpha, \beta > 0$. Dann ist

$$\mathcal{L}_X(t) = \frac{\beta^\alpha}{(\beta + t)^\alpha}, \quad t \geq 0,$$

und es gilt für $X \sim \Gamma(\alpha, \beta)$ und $Y \sim \Gamma(\gamma, \delta)$ die Implikation

$$\alpha \leq \gamma, \beta \geq \delta \Rightarrow X \leq_{\mathcal{L}} Y.$$

Im Spezialfall $\alpha = \gamma = 1$ ergibt sich die entsprechende Aussage für die Exponentialverteilung und im Spezialfall $\alpha = \gamma = \frac{n}{2}, n \in \mathbb{N}, \beta = \delta = \frac{1}{2}$, folgt die entsprechende Aussage für χ_n^2 - die zentrale Chi-Quadrat-Verteilung mit n Freiheitsgraden.

c) Seien $X \sim \text{Wei}(\alpha, \beta), Y \sim \text{Wei}(\gamma, \delta)$ weibullverteilt mit den Parametern $\alpha, \beta, \gamma, \delta > 0$. Dann besitzt X die Dichte

$$f(x) = \alpha \beta x^{\alpha-1} e^{-\beta x^\alpha}, \quad x > 0,$$

und es gilt die Implikation

$$\alpha \leq \gamma, EX \leq EY \Rightarrow X \leq_{\mathcal{L}} Y.$$

Im Fall der Weibull-Verteilung ist die Laplacetransformierte nicht geschlossen darstellbar. Zum Beweis der getroffenen Aussage und ebenfalls der folgenden Aussage in Teil d) siehe [Tay83].

- d) Seien $X \sim \text{LogN}(\mu_1, \sigma_1^2), Y \sim \text{LogN}(\mu_2, \sigma_2^2)$ jeweils logarithmisch normalverteilt mit den Parametern $\mu_1, \mu_2 \in \mathbb{R}, \sigma_1^2, \sigma_2^2 > 0$. Dann besitzt X die Dichte

$$f(x) = \frac{1}{x\sigma_1\sqrt{2\pi}} \exp\left(-\frac{(\ln(x) - \mu_1)^2}{2\sigma_1^2}\right), \quad x > 0,$$

und es gilt die Implikation

$$\sigma_1 \geq \sigma_2, EX \leq EY \Rightarrow X \leq_{\mathcal{L}} Y.$$

- e) Sei \mathbf{X} vom *Gumbel-Morgenstern-Typ* mit der bivariaten Verteilungsfunktion

$$F_{\vartheta, \mu, \lambda}(x, y) = \text{Exp}_{\lambda}(x) \text{Exp}_{\mu}(y) (1 + \vartheta(1 - \text{Exp}_{\lambda}(x))(1 - \text{Exp}_{\mu}(y))),$$

$x, y \geq 0, \mu, \lambda > 0, \vartheta \in [-1, 1], \text{Exp}_{\lambda}(x) := 1 - \exp(-\lambda x), \text{Exp}_{\mu}(y) := 1 - \exp(-\mu y)$.

Dann ist

$$\mathcal{L}_{F_{\vartheta, \mu, \lambda}}(s, t) = \frac{\lambda}{\lambda + s} \frac{\mu}{\mu + t} + \vartheta \left(\frac{2\lambda}{2\lambda + s} - \frac{\lambda}{\lambda + s} \right) \left(\frac{2\mu}{2\mu + t} - \frac{\mu}{\mu + t} \right), \quad s, t \geq 0.$$

Wegen $\frac{2\lambda}{2\lambda + s} - \frac{\lambda}{\lambda + s} \geq 0$ für alle $s \geq 0$ und $\frac{2\mu}{2\mu + t} - \frac{\mu}{\mu + t} \geq 0$ für alle $t \geq 0$ ist $F_{\vartheta, \mu, \lambda} \leq_{\mathcal{L}} F_{\theta, \mu, \lambda}$ für $\vartheta \geq \theta$.

- f) Seien Y_0, Y_1, \dots, Y_d unabhängige Zufallsvariablen mit den Dichtefunktionen

$$f_{Y_k}(y) = \frac{1}{\Gamma(\vartheta_k)} y^{\vartheta_k - 1} e^{-y}, \quad y > 0, \vartheta_k > 0, k = 0, \dots, d.$$

Sei $X_k = Y_0 + Y_k, k = 1, \dots, d$. Dann ist $\mathbf{X} = (X_1, \dots, X_d)$ d -dimensional multivariat gammaverteilt mit den Parametern $\vartheta_0, \dots, \vartheta_d$ (siehe [KBJ00], S. 454 ff.) und es gilt

$$\mathcal{L}_{\mathbf{X}}(\mathbf{t}) = \left(1 + \sum_{k=1}^d t_k \right)^{-\vartheta_0} \prod_{k=1}^d (1 + t_k)^{-\vartheta_k}, \quad \mathbf{t} = (t_1, \dots, t_d) \geq 0.$$

Folglich ist $\mathcal{L}_{\mathbf{X}}(\mathbf{t})$ monoton fallend in $\vartheta_k, k = 0, \dots, d$.

Kapitel 2

Test auf Laplace-Ordnung

2.1 Das Testproblem und die Testgrößen

In diesem Abschnitt sollen Testverfahren zur Überprüfung der Hypothese der Laplace-Ordnung aufgestellt werden. Nichtparametrische Tests auf Laplace-Ordnung sind dabei nach Wissen des Autors in der Literatur bisher nicht untersucht worden.

Eine Ausnahme bilden die scheinbar unveröffentlichten Ansätze für nichtparametrische Tests für das Einstichproben-Testproblem auf Laplace-Ordnung im univariaten Fall in [Ben07]. Dort werden Testgrößen vorgestellt, die den hier untersuchten einseitigen Kolmogorov-Smirnov und Cramér-von Mises-Testgrößen im Eindimensionalen entsprechen (im Cramér-von Mises Fall wird allerdings auf Gewichtsfunktionen verzichtet). Das dort vorgestellte Bootstrapverfahren scheint für die vorliegende Situation und für den Fall der Supremumstestgröße im Allgemeinen allerdings unbrauchbar zu sein.

Es seien im Folgenden \mathbf{X}, \mathbf{Y} Zufallsvektoren mit Werten in $\mathbb{R}_{\geq 0}^d$. Wir betrachten das Zweistichproben-Testproblem

$$H : \mathbf{X} \leq_{\mathcal{L}} \mathbf{Y} \quad \text{vs.} \quad K : \mathbf{X} \not\leq_{\mathcal{L}} \mathbf{Y}, \quad (2.1.1)$$

was nach Definition 1.1.1 zu

$$H : \forall \mathbf{t} \geq 0 : \mathcal{L}_{\mathbf{X}}(\mathbf{t}) - \mathcal{L}_{\mathbf{Y}}(\mathbf{t}) \geq 0 \quad \text{vs.} \quad K : \exists \mathbf{t} \geq 0 : \mathcal{L}_{\mathbf{Y}}(\mathbf{t}) - \mathcal{L}_{\mathbf{X}}(\mathbf{t}) > 0$$

äquivalent ist. Es ist daher plausibel, die Testgrößen auf dem empirischen Äquivalent des Abstandes $\mathcal{L}_{\mathbf{Y}}(\mathbf{t}) - \mathcal{L}_{\mathbf{X}}(\mathbf{t})$ basieren zu lassen. Seien dazu $\mathbf{X}_1, \dots, \mathbf{X}_m, m \in \mathbb{N}$, unabhängige Zufallsvektoren mit je derselben Verteilung wie \mathbf{X} und $\mathbf{Y}_1, \dots, \mathbf{Y}_n, n \in \mathbb{N}$, unabhängige und von $\mathbf{X}_1, \dots, \mathbf{X}_m$ unabhängige Zufallsvektoren mit je derselben Verteilung wie \mathbf{Y} . Wir

definieren die *empirische Laplacetransformierte* von $\mathbf{X}_1, \dots, \mathbf{X}_m$ als

$$\mathcal{L}_{\mathbf{X}_1, \dots, \mathbf{X}_m}(\mathbf{t}) := \frac{1}{m} \sum_{i=1}^m \exp(-\mathbf{t}' \mathbf{X}_i), \quad \mathbf{t} = (t_1, \dots, t_d) \in \mathbb{R}_{\geq 0}^d,$$

und verwenden alternativ die Notation $\mathcal{L}_{\mathbf{X}, m}(\mathbf{t}) := \mathcal{L}_{\mathbf{X}_1, \dots, \mathbf{X}_m}(\mathbf{t})$. In analoger Weise bezeichnen wir mit $\mathcal{L}_{\mathbf{Y}, n}(\mathbf{t}) := \mathcal{L}_{\mathbf{Y}_1, \dots, \mathbf{Y}_n}(\mathbf{t})$ die empirische Laplacetransformierte der Zufallsvektoren $\mathbf{Y}_1, \dots, \mathbf{Y}_n$. Ist F die Verteilungsfunktion der Zufallsvariablen \mathbf{X}_i , so benutzen wir gelegentlich auch die Schreibweise $\mathcal{L}_{F, m} := \mathcal{L}_{\mathbf{X}, m}$.

Es sei ferner

$$\Delta_{m, n}(\mathbf{t}) := \sqrt{\frac{mn}{m+n}} (\mathcal{L}_{\mathbf{Y}, n}(\mathbf{t}) - \mathcal{L}_{\mathbf{X}, m}(\mathbf{t})), \quad m, n \in \mathbb{N}.$$

Um zu sinnvollen Testgrößen für das eingangs definierte Testproblem zu gelangen, bieten sich dann beispielsweise Supremums- und Integraltyp-Testgrößen an. Dabei sind im vorliegenden Fall lediglich einseitige Abweichungen von der Hypothese zu betrachten.

Diese Überlegungen führen in natürlicher Weise auf die folgenden einseitigen Varianten der bekannten Kolmogorov-Smirnov- und Cramér-von Mises-Testgrößen. Es sei

$$T_{m, n}^{KS} := \sup_{\mathbf{t} \geq 0} \Delta_{m, n}(\mathbf{t}) \quad (2.1.2)$$

und

$$\begin{aligned} T_{m, n}^{CvM} &:= \int_{\mathbb{R}_{\geq 0}^d} \Delta_{m, n}^+(\mathbf{t}) \omega(\mathbf{t}) d\mathbf{t} \\ &= \int_0^\infty \dots \int_0^\infty \Delta_{m, n}^+(t_1, \dots, t_d) \omega(t_1, \dots, t_d) dt_1 \dots dt_d, \end{aligned} \quad (2.1.3)$$

mit einer integrierbaren Gewichtsfunktion $\omega : \mathbb{R}_{\geq 0}^d \rightarrow \mathbb{R}_{> 0}$ und

$$\Delta_{m, n}^+(\mathbf{t}) := \max(0, \Delta_{m, n}(\mathbf{t}))$$

als Positivteil der Funktion $\mathbf{t} \mapsto \Delta_{m, n}(\mathbf{t})$, $\mathbf{t} \in \mathbb{R}_{\geq 0}^d$. Große Werte von $T_{m, n}^{KS}$ bzw. $T_{m, n}^{CvM}$ sprechen gegen die Hypothese.

Wir bemerken, dass die Wahl $\omega \equiv 1$ im Hinblick auf die gewünschte Endlichkeit des Integrals $T_{m, n}^{CvM}$ im Allgemeinen $\mathbf{X}_1, \dots, \mathbf{X}_m, \mathbf{Y}_1, \dots, \mathbf{Y}_n > 0$ f.s. nötig macht:

Für $d = 1$ und mit

$$c_k := \begin{cases} \frac{1}{n}, & k = 1, \dots, n \\ -\frac{1}{m}, & k = n+1, \dots, n+m \end{cases}, \quad Z_k := \begin{cases} Y_k, & k = 1, \dots, n \\ X_{k-n}, & k = n+1, \dots, n+m, \end{cases}$$

$k = 1, \dots, m + n$, gilt nämlich beispielsweise

$$\Delta_{m,n}(t) = \sqrt{\frac{mn}{m+n}} \sum_{k=1}^{m+n} c_k \exp(-tZ_k).$$

Die Funktion $t \mapsto \Delta_{m,n}(t)$ ist dann stetig auf $[0, \infty)$ und hat im Fall $\Delta_{m,n}(t) \not\equiv 0$ höchstens $m + n$ Nullstellen (siehe [PS70], Problem 76). Bezeichnen wir mit $a_1 < \dots < a_N$, $N \in \{1, \dots, m + n\}$, die paarweise verschiedenen und der Größe nach geordneten Nullstellen von $\Delta_{m,n}(t)$, so ist die Testgröße $T_{m,n}^{CvM}$ von der Form

$$T_{m,n}^{CvM} = \sum_{j \in I} \int_{a_j}^{a_{j+1}} \Delta_{m,n}(t) dt + \left(\int_{a_N}^{\infty} \Delta_{m,n}(t) dt \right)^+ \quad (2.1.4)$$

mit einer geeigneten gewählten Indexmenge $I \subset \{1, \dots, N\}$. Hinreichende Bedingungen für die fast sichere Endlichkeit von $T_{m,n}^{CvM}$ im Fall $\omega \equiv 1$ sind daher beispielsweise die Momentenbedingungen $E(X_1^{-1}) < \infty$, $E(Y_1^{-1}) < \infty$. Damit auf Bedingungen dieser Art an die Zufallsvektoren $\mathbf{X}_1, \mathbf{Y}_1$ verzichtet werden kann, beschränken wir uns im Folgenden auf Gewichtsfunktionen mit den genannten Eigenschaften.

2.2 Asymptotik

Es seien im Folgenden für $m, n \in \mathbb{N}$ stets $\mathbf{X}_1, \dots, \mathbf{X}_m, \mathbf{Y}_1, \dots, \mathbf{Y}_n$ unabhängige und d -dimensionale Zufallsvektoren mit je nicht-negativen Komponenten, wobei $\mathbf{X}_1, \dots, \mathbf{X}_m$ identisch verteilt mit der Verteilungsfunktion F und $\mathbf{Y}_1, \dots, \mathbf{Y}_n$ identisch verteilt mit der Verteilungsfunktion G sind. Wir setzen F und G als nicht-degeneriert voraus. Sei $\frac{m}{m+n} \rightarrow \tau \in [0, 1]$ für $m, n \rightarrow \infty$.

Wir wollen in diesem Abschnitt das asymptotische Verhalten der Testgrößen (2.1.2) und (2.1.3) untersuchen, die sich jeweils als Funktionen des Prozesses $(\Delta_{m,n}(\mathbf{t}), \mathbf{t} \geq 0)$ ergeben. Für $\Delta_{m,n}(\mathbf{t})$ erhalten wir dabei offenbar die Darstellung

$$\begin{aligned} \Delta_{m,n}(\mathbf{t}) &= \sqrt{\frac{m}{m+n}} \sqrt{n} (\mathcal{L}_{\mathbf{Y},n}(\mathbf{t}) - \mathcal{L}_{\mathbf{Y}}(\mathbf{t})) - \sqrt{\frac{n}{m+n}} \sqrt{m} (\mathcal{L}_{\mathbf{X},m}(\mathbf{t}) - \mathcal{L}_{\mathbf{X}}(\mathbf{t})) \\ &\quad + \sqrt{\frac{mn}{m+n}} (\mathcal{L}_{\mathbf{Y}}(\mathbf{t}) - \mathcal{L}_{\mathbf{X}}(\mathbf{t})), \quad \forall \mathbf{t} \geq 0. \end{aligned} \quad (2.2.1)$$

Setzen wir

$$\begin{aligned} \mathbb{L}_{\mathbf{Y},n}(\mathbf{t}) &:= \sqrt{n} (\mathcal{L}_{\mathbf{Y},n}(\mathbf{t}) - \mathcal{L}_{\mathbf{Y}}(\mathbf{t})) = \sqrt{n} \left(\frac{1}{n} \sum_{i=1}^n f_{\mathbf{t}}(\mathbf{Y}_i) - E f_{\mathbf{t}}(\mathbf{Y}_1) \right), \\ \mathbb{L}_{\mathbf{X},m}(\mathbf{t}) &:= \sqrt{m} (\mathcal{L}_{\mathbf{X},m}(\mathbf{t}) - \mathcal{L}_{\mathbf{X}}(\mathbf{t})) = \sqrt{m} \left(\frac{1}{m} \sum_{i=1}^m f_{\mathbf{t}}(\mathbf{X}_i) - E f_{\mathbf{t}}(\mathbf{X}_1) \right), \end{aligned}$$

mit $f_{\mathbf{t}} : \mathbb{R}_{\geq 0}^d \rightarrow \mathbb{R}$, $f_{\mathbf{t}}(\mathbf{x}) := \exp(-\mathbf{t}'\mathbf{x})$, so sind $(\mathbb{L}_{\mathbf{Y},n}(\mathbf{t}), \mathbf{t} \geq 0)$ und $(\mathbb{L}_{\mathbf{X},m}(\mathbf{t}), \mathbf{t} \geq 0)$ unabhängige und durch die Funktionenklasse $\mathcal{H} = \{f_{\mathbf{t}}; \mathbf{t} \geq 0\}$ indizierte empirische Prozesse (vergleiche Anhang B.2.1). Wegen $|f_{\mathbf{t}}| \leq 1$ für alle $\mathbf{t} \geq 0$ lassen sich diese empirischen Prozesse als Abbildungen mit Werten in $l^\infty(\mathcal{H})$ oder $l^\infty(\mathbb{R}_{\geq 0}^d)$ auffassen. Bekanntlich kann es bei der Interpretation des empirischen Prozesses als Abbildung mit Werten in speziellen Funktionenräumen zu Messbarkeitsproblemen kommen, wie etwa das klassische Beispiel des zur Gleichverteilung auf $[0, 1]$ gehörenden empirischen Prozesses, interpretiert als Abbildung in den Raum der *cádlág*-Funktionen auf $[0, 1]$, lehrt (siehe [Bil68], Kapitel 18). Im vorliegenden Fall können diese Probleme allerdings vermieden werden; ein Sachverhalt der im nächsten Lemma verdeutlicht werden soll.

Vorsehen mit der üblichen Metrik

$$\rho(s, t) = |\arctan(s) - \arctan(t)|, \quad s, t \in [-\infty, \infty],$$

$\arctan(-\infty) := -\pi/2$, $\arctan(\infty) := \pi/2$, fassen wir dabei im Folgenden die erweiterten reellen Zahlen $\overline{\mathbb{R}} = [-\infty, \infty]$ und auch die nicht-negativen erweiterten reellen Zahlen $\overline{\mathbb{R}}_{\geq 0} = [0, \infty]$ (vorsehen mit der entsprechenden Einschränkung von ρ auf $[0, \infty]$) jeweils als kompakten metrischen Raum auf. Dann sind auch die kartesischen Produkte $(\overline{\mathbb{R}}^d, \rho_d)$ und $(\overline{\mathbb{R}}_{\geq 0}^d, \rho_d)$ mit $\rho_d(\mathbf{s}, \mathbf{t}) := \sum_{k=1}^d \rho(s_k, t_k)$ für $\mathbf{s} = (s_1, \dots, s_d)$, $\mathbf{t} = (t_1, \dots, t_d) \in \overline{\mathbb{R}}^d$, kompakte metrische Räume. Für die folgenden Betrachtungen dieses Abschnitts definieren wir ferner für ein $\mathbf{t} = (t_1, \dots, t_d) \in \overline{\mathbb{R}}_{\geq 0}^d$ die Menge

$$I(\mathbf{t}) := \{j \in \{1, \dots, d\} : t_j = \infty\}. \quad (2.2.2)$$

2.2.1. Lemma. Für ein $\mathbf{t}^* = (t_1^*, \dots, t_d^*) \in \overline{\mathbb{R}}_{\geq 0}^d \setminus \mathbb{R}_{\geq 0}^d$ sei

$$\mathcal{L}_{\mathbf{X},m}(\mathbf{t}^*) := \lim_{\substack{t_j \rightarrow \infty, j \in I(\mathbf{t}^*) \\ t_j = t_j^*, j \notin I(\mathbf{t}^*)}} \mathcal{L}_{\mathbf{X},m}(\mathbf{t}), \quad \mathcal{L}_{\mathbf{X}}(\mathbf{t}^*) := \lim_{\substack{t_j \rightarrow \infty, j \in I(\mathbf{t}^*) \\ t_j = t_j^*, j \notin I(\mathbf{t}^*)}} \mathcal{L}_{\mathbf{X}}(\mathbf{t}) \quad (2.2.3)$$

und $(C(\overline{\mathbb{R}}_{\geq 0}^d), \|\cdot\|_\infty)$ sei der metrische Raum der stetigen Funktionen auf der kompakten Menge $\overline{\mathbb{R}}_{\geq 0}^d$, vorsehen mit der Supremumsmetrik. Es sei (Ω, \mathcal{A}, P) ein Wahrscheinlichkeitsraum und weiter $\{\mathbf{X}_k\}_{k \in \mathbb{N}}$ eine Folge unabhängiger und identisch verteilter Zufallsvektoren mit $\mathbf{X}_k : (\Omega, \mathcal{A}) \mapsto (\mathbb{R}_{\geq 0}^d, \mathfrak{B}(\mathbb{R}_{\geq 0}^d))$, $k \in \mathbb{N}$. Dann ist die Abbildung

$$\mathbb{L}_{\mathbf{X},m} : \Omega \rightarrow C(\overline{\mathbb{R}}_{\geq 0}^d) \\ \omega \mapsto \left(\sqrt{m} \left(\frac{1}{m} \sum_{i=1}^m f_{\mathbf{t}}(\mathbf{X}_i(\omega)) - E f_{\mathbf{t}}(\mathbf{X}_1(\omega)) \right) \right), \mathbf{t} \in \overline{\mathbb{R}}_{\geq 0}^d$$

für jedes $m \in \mathbb{N}$ $(\mathcal{A}, \mathfrak{B}(C(\overline{\mathbb{R}}_{\geq 0}^d)))$ -messbar.

Beweis. Die Funktion $\mathbf{t} \mapsto \mathcal{L}_{\mathbf{X},m}(\mathbf{t})$ ist stetig auf $\mathbb{R}_{\geq 0}^d$, monoton fallend in jeder Komponente und nach unten durch Null beschränkt, so dass die Limites in (2.2.3) existieren. Mit der getroffenen Definition der Fortsetzung auf $\overline{\mathbb{R}}_{\geq 0}^d$ sind die empirischen Laplacetransformierten $\mathcal{L}_{\mathbf{X},m}$ somit stetig auf $\overline{\mathbb{R}}_{\geq 0}^d$ für jedes $m \in \mathbb{N}$. Ebenso ist $\mathcal{L}_{\mathbf{X}}$ stetig auf $\overline{\mathbb{R}}_{\geq 0}^d$. Die Abbildung $\mathbb{L}_{\mathbf{X},m}$ nimmt also Werte in $(C(\overline{\mathbb{R}}_{\geq 0}^d), \|\cdot\|_{\infty})$ an. Es ist $(C(\overline{\mathbb{R}}_{\geq 0}^d), \|\cdot\|_{\infty})$ separabel, und $\mathbb{L}_{\mathbf{X},m}(\mathbf{t})$ ist für jedes feste \mathbf{t} eine reelle Zufallsvariable. Die Behauptung folgt dann mit [VW96], Problem 1.7.1. \square

2.2.2. Bemerkung. Bei den Untersuchungen der asymptotischen Eigenschaften der zuvor eingeführten stochastischen Prozesse und den damit verbundenen Testgrößen wollen wir im Folgenden den allgemeinen Ansatz der Hoffmann-Jørgensen-Theorie für schwache Konvergenz nicht messbarer Abbildungen verwenden (vergleiche Kapitel B des Anhangs). Setzen wir die empirischen Prozesse wie oben beschrieben stetig auf $\overline{\mathbb{R}}_{\geq 0}^d$ fort, so zeigt Lemma 2.2.1 allerdings, dass dieser Grad der Allgemeinheit hier im Grunde noch nicht nötig ist. Aufgrund der Borel-Messbarkeit der Abbildungen $\mathbb{L}_{\mathbf{X},m}$ und $\mathbb{L}_{\mathbf{Y},n}$ kann bei den Grenzbetrachtungen nämlich zunächst der gewöhnliche Begriff der schwachen Konvergenz verwendet und mit Standardmethoden der Verteilungskonvergenz stochastischer Prozesse argumentiert werden. Vorgreifend auf empirische Prozesse, die durch Funktionen mit analytisch schwächeren Eigenschaften indiziert werden - wie etwa die in Abschnitt 4 betrachteten - machen wir aber bereits hier von der allgemeineren Theorie Gebrauch.

Um Grenzwertaussagen für die empirischen Prozesse $\mathbb{L}_{\mathbf{X},m}$ und $\mathbb{L}_{\mathbf{Y},n}$ zu erhalten, wollen wir funktionale Grenzwertsätze benutzen. Im Hinblick auf das vorherige Lemma erscheint es dabei zunächst sinnvoll, die empirischen Prozesse $\mathbb{L}_{\mathbf{X},m}$ und $\mathbb{L}_{\mathbf{Y},n}$ mit der in (2.2.3) getroffenen Definition als Prozesse auf $\overline{\mathbb{R}}_{\geq 0}^d$ aufzufassen. Mit

$$\mathcal{F} := \left\{ f_{\mathbf{t}} : \mathbb{R}_{\geq 0}^d \rightarrow [0, 1] \mid \mathbf{t} \in \overline{\mathbb{R}}_{\geq 0}^d, f_{\mathbf{t}}(x_1, \dots, x_d) = \prod_{j=1}^d f_{t_j}(x_j) \right\}$$

und

$$f_{t_j}(x) := \begin{cases} \exp(-t_j x), & t_j \in \mathbb{R}, \\ \mathbf{1}_{\{0\}}(x), & t_j = \infty, \end{cases}$$

für $x \in \mathbb{R}_{\geq 0}$, sind dann $(\mathbb{L}_{\mathbf{X},m}(\mathbf{t}), \mathbf{t} \in \overline{\mathbb{R}}_{\geq 0}^d)$ bzw. $(\mathbb{L}_{\mathbf{Y},n}(\mathbf{t}), \mathbf{t} \in \overline{\mathbb{R}}_{\geq 0}^d)$ die durch \mathcal{F} indizierten empirischen Prozesse der Zufallsvektoren $\mathbf{X}_1, \dots, \mathbf{X}_m$ bzw. $\mathbf{Y}_1, \dots, \mathbf{Y}_n$. Das nächste Lemma sichert die asymptotische Straffheit dieser Prozesse.

2.2.3. Lemma. \mathcal{F} ist eine VC-Funktionenklasse und gleichmäßig Donsker.

Beweis. Wir betrachten die Funktionenklasse $\mathcal{G} = \{g_{\mathbf{t}}, \mathbf{t} \in \mathbb{R}_{\geq 0}^d\}$ mit

$$g_{\mathbf{t}} := \begin{cases} \mathbb{R}_{\geq 0}^d & \rightarrow \mathbb{R}_{\geq 0} \\ \mathbf{x} & \mapsto \mathbf{t}'\mathbf{x} \end{cases}.$$

Der Subgraph von $g_{\mathbf{t}}$ für ein festes $\mathbf{t} \in \mathbb{R}_{\geq 0}^d$ ist die Menge $\{(\mathbf{x}, s) \in \mathbb{R}_{\geq 0}^d \times \mathbb{R}_{\geq 0} : \mathbf{t}'\mathbf{x} \geq s\}$, die offenbar das Komplement eines Halbraums in $\mathbb{R}_{\geq 0}^{d+1}$ beschreibt. Die Menge aller Halbräume in \mathbb{R}^n , $n \in \mathbb{N}$, ist bekanntlich eine VC-Klasse (siehe [VW96], S. 152) und daher \mathcal{G} eine VC-Funktionenklasse. Die Funktion $\phi : \mathbb{R}_{\geq 0} \rightarrow [0, 1]$, $\phi(x) = \exp(-x)$, ist monoton und daher ist die Klasse $\phi(\mathcal{G})$, definiert durch

$$\phi(\mathcal{G}) := \{\phi \circ g_{\mathbf{t}}, \mathbf{t} \in \mathbb{R}_{\geq 0}^d\}, \quad (\phi \circ g_{\mathbf{t}})(\mathbf{x}) = \exp\left(-\sum_{j=1}^d t_j x_j\right),$$

gemäß [VW96], Lemma 2.6.18 (viii), ebenfalls eine VC-Funktionenklasse. Mit $I(\mathbf{t})$ wie in (2.2.2) und $\mathcal{H} := \mathcal{F} \setminus \phi(\mathcal{G})$ sowie

$$h_{\mathbf{t}}(x_1, \dots, x_d) := \prod_{j \in I(\mathbf{t})} \mathbf{1}_{\{0\}}(x_j) \exp\left(-\sum_{j \notin I(\mathbf{t})} t_j x_j\right), \quad (2.2.4)$$

gilt dann $\mathcal{H} = \{h_{\mathbf{t}}, \mathbf{t} \in \overline{\mathbb{R}_{\geq 0}^d} \setminus \mathbb{R}_{\geq 0}^d\}$. Offenbar gilt $I(\mathbf{t}) \neq \emptyset$ für jedes $\mathbf{t} \in \overline{\mathbb{R}_{\geq 0}^d} \setminus \mathbb{R}_{\geq 0}^d$. Bezeichnen wir mit \mathcal{I} die Menge sämtlicher nichtleerer Teilmengen der Menge $\{1, \dots, d\}$, erhalten wir außerdem die Darstellung

$$\mathcal{H} = \bigcup_{J \in \mathcal{I}} \mathcal{H}_J, \quad \mathcal{H}_J := \{h_{\mathbf{t}} \in \mathcal{H} : I(\mathbf{t}) = J\}.$$

Für jedes feste $J \subset \mathcal{I}$ gibt es dann aber aufgrund von (2.2.4) für jede Funktion $h_{\mathbf{t}} \in \mathcal{H}_J$ eine Funktion $\phi(g_{\mathbf{t}}) \in \phi(\mathcal{G})$ mit der Eigenschaft, dass $h_{\mathbf{t}}(\mathbf{x}) = \prod_{j \in J} \mathbf{1}_{\{0\}}(x_j) \phi(g_{\mathbf{t}})(\mathbf{x})$ für alle $\mathbf{x} \in \mathbb{R}_{\geq 0}^d$ zutrifft. Insbesondere hängt die Funktion $g(\mathbf{x}) = \prod_{j \in J} \mathbf{1}_{\{0\}}(x_j)$ in dieser Darstellung bei fest gewähltem $J \in \mathcal{I}$ nicht mehr von \mathbf{t} ab.

Mit [VW96], Lemma 2.6.18 (vi), folgt, dass jedes der \mathcal{H}_J , $J \in \mathcal{I}$, eine VC-Funktionenklasse ist. Somit ergibt sich \mathcal{F} als endliche Vereinigung von VC-Funktionenklassen, was schließlich die erste Behauptung des Lemmas zeigt.

Wegen $|f_{\mathbf{t}}| \leq 1$, Satz B.2.1.5 und Satz B.2.1.3 ist \mathcal{F} darüber hinaus Donsker für jedes zugrunde liegende Wahrscheinlichkeitsmaß Q . \square

Nach diesen vorbereitenden Überlegungen erhalten wir schließlich das folgende Resultat über das Grenzverhalten der zugrunde liegenden empirischen Prozesse.

2.2.4. Satz. *Es sei $(\mathbb{L}(\mathbf{t}), \mathbf{t} \in \overline{\mathbb{R}}_{\geq 0}^d)$ ein zentrierter Gaußprozess mit Pfaden in $C(\overline{\mathbb{R}}_{\geq 0}^d)$ und der Kovarianzfunktion*

$$\rho(\mathbf{s}, \mathbf{t}) := \text{Cov}(\mathbb{L}(\mathbf{s}), \mathbb{L}(\mathbf{t})) = \tau \rho_G(\mathbf{s}, \mathbf{t}) + (1 - \tau) \rho_F(\mathbf{s}, \mathbf{t}), \quad \mathbf{s}, \mathbf{t} \in \overline{\mathbb{R}}_{\geq 0}^d,$$

wobei für $\mathbf{s}, \mathbf{t} \in \overline{\mathbb{R}}_{\geq 0}^d$

$$\rho_F(\mathbf{s}, \mathbf{t}) := \mathcal{L}_F(\mathbf{s} + \mathbf{t}) - \mathcal{L}_F(\mathbf{s})\mathcal{L}_F(\mathbf{t}), \quad \rho_G(\mathbf{s}, \mathbf{t}) := \mathcal{L}_G(\mathbf{s} + \mathbf{t}) - \mathcal{L}_G(\mathbf{s})\mathcal{L}_G(\mathbf{t})$$

ist.

Dann gilt

$$\left(\sqrt{\frac{m}{m+n}} \mathbb{L}_{\mathbf{Y},n}(\mathbf{t}) - \sqrt{\frac{n}{m+n}} \mathbb{L}_{\mathbf{X},m}(\mathbf{t}), \mathbf{t} \in \overline{\mathbb{R}}_{\geq 0}^d \right) \xrightarrow{v} \left(\mathbb{L}(\mathbf{t}), \mathbf{t} \in \overline{\mathbb{R}}_{\geq 0}^d \right),$$

für $m, n \rightarrow \infty$.

Beweis. Nach Lemma 2.2.3 und Satz B.2.1.3 ist $(\mathbb{L}_{\mathbf{Y},n}(\mathbf{t}), \mathbf{t} \in \overline{\mathbb{R}}_{\geq 0}^d)$ eine asymptotisch straffe Folge von Abbildungen. Offenbar ist ferner für jedes $\mathbf{t} \in \overline{\mathbb{R}}_{\geq 0}^d$ $f_{\mathbf{t}} \in \mathcal{F}$ quadratintegrierbar bezüglich jedes zugrunde liegenden Wahrscheinlichkeitsmaßes und es sind $f_{\mathbf{t}}(\mathbf{Y}_1), \dots, f_{\mathbf{t}}(\mathbf{Y}_n)$ unabhängig und identisch verteilt. Wir erhalten deshalb mit dem mehrdimensionalen zentralen Grenzwertsatz, dass für beliebige $\mathbf{t}_1, \dots, \mathbf{t}_k \in \overline{\mathbb{R}}_{\geq 0}^d$

$$(\mathbb{L}_{\mathbf{Y},n}(\mathbf{t}_1), \dots, \mathbb{L}_{\mathbf{Y},n}(\mathbf{t}_k)) \xrightarrow{v} N_k(0, \Sigma) \quad (2.2.5)$$

mit $\Sigma = (\sigma_{ij})_{1 \leq i, j \leq k}$, $\sigma_{ij} = \text{Cov}(f_{\mathbf{t}_i}(\mathbf{Y}_1), f_{\mathbf{t}_j}(\mathbf{Y}_1))$ gilt. Gemäß Satz B.2.1 konvergiert $(\mathbb{L}_{\mathbf{Y},n}(\mathbf{t}), \mathbf{t} \in \overline{\mathbb{R}}_{\geq 0}^d)$ gegen einen straffen Grenzprozess $(\mathbb{L}_{\mathbf{Y}}(\mathbf{t}), \mathbf{t} \in \overline{\mathbb{R}}_{\geq 0}^d)$ mit Pfaden in $l^\infty(\overline{\mathbb{R}}_{\geq 0}^d)$, der wegen (2.2.5) ein zentrierter Gaußprozess mit der Kovarianzfunktion

$$\begin{aligned} \rho_G(\mathbf{s}, \mathbf{t}) &:= \text{Cov}(\mathbb{L}_{\mathbf{Y}}(\mathbf{s}), \mathbb{L}_{\mathbf{Y}}(\mathbf{t})) \\ &= \int_{\overline{\mathbb{R}}_{\geq 0}^d} f_{\mathbf{s}}(\mathbf{x}) f_{\mathbf{t}}(\mathbf{x}) dG(\mathbf{x}) - \int_{\overline{\mathbb{R}}_{\geq 0}^d} f_{\mathbf{s}}(\mathbf{x}) dG(\mathbf{x}) \int_{\overline{\mathbb{R}}_{\geq 0}^d} f_{\mathbf{t}}(\mathbf{x}) dG(\mathbf{x}) \\ &= \int_{\overline{\mathbb{R}}_{\geq 0}^d} e^{-(\mathbf{s}+\mathbf{t})'\mathbf{x}} dG(\mathbf{x}) - \int_{\overline{\mathbb{R}}_{\geq 0}^d} e^{-\mathbf{s}'\mathbf{x}} dG(\mathbf{x}) \int_{\overline{\mathbb{R}}_{\geq 0}^d} e^{-\mathbf{t}'\mathbf{x}} dG(\mathbf{x}) \\ &= \mathcal{L}_G(\mathbf{s} + \mathbf{t}) - \mathcal{L}_G(\mathbf{s})\mathcal{L}_G(\mathbf{t}), \quad \mathbf{s}, \mathbf{t} \in \overline{\mathbb{R}}_{\geq 0}^d, \end{aligned}$$

ist. Nach Lemma 2.2.1 kann der empirische Prozess $\mathbb{L}_{\mathbf{Y},n}$ als Zufallsvariable mit Werten in $C(\overline{\mathbb{R}}_{\geq 0}^d)$ aufgefasst werden. Da jede gleichmäßig konvergente Folge stetiger Funktionen auf

einem kompakten metrischen Raum einen stetigen Grenzwert besitzt, ist $(C(\overline{\mathbb{R}}_{\geq 0}^d), \|\cdot\|_{\infty})$ eine abgeschlossene Teilmenge von $(l^{\infty}(\overline{\mathbb{R}}_{\geq 0}^d), \|\cdot\|_{\infty})$.

Satz B.1.2 und $(\mathbb{L}_{\mathbf{Y},n}(\mathbf{t}), \mathbf{t} \in \overline{\mathbb{R}}_{\geq 0}^d) \xrightarrow{v} (\mathbb{L}_{\mathbf{Y}}(\mathbf{t}), \mathbf{t} \in \overline{\mathbb{R}}_{\geq 0}^d)$ implizieren daher

$$1 = \limsup_{n \rightarrow \infty} P\left(\mathbb{L}_{\mathbf{Y},n} \in C(\overline{\mathbb{R}}_{\geq 0}^d)\right) \leq P\left(\mathbb{L}_{\mathbf{Y}} \in C(\overline{\mathbb{R}}_{\geq 0}^d)\right),$$

so dass der Grenzprozess ebenfalls als Abbildung mit Werten in $C(\overline{\mathbb{R}}_{\geq 0}^d)$ aufgefasst werden kann.

Analog erhalten wir

$$(\mathbb{L}_{\mathbf{X},m}(\mathbf{t}), \mathbf{t} \in \overline{\mathbb{R}}_{\geq 0}^d) \xrightarrow{v} (\mathbb{L}_{\mathbf{X}}(\mathbf{t}), \mathbf{t} \in \overline{\mathbb{R}}_{\geq 0}^d),$$

mit einem zentrierten Gaußprozess $\mathbb{L}_{\mathbf{X}}$ mit Pfaden in $C(\overline{\mathbb{R}}_{\geq 0}^d)$ und der Kovarianzfunktion

$$\rho_F(\mathbf{s}, \mathbf{t}) = \text{Cov}(\mathbb{L}_{\mathbf{X}}(\mathbf{s}), \mathbb{L}_{\mathbf{X}}(\mathbf{t})) = \mathcal{L}_F(\mathbf{s} + \mathbf{t}) - \mathcal{L}_F(\mathbf{s})\mathcal{L}_F(\mathbf{t}), \quad \mathbf{s}, \mathbf{t} \in \overline{\mathbb{R}}_{\geq 0}^d.$$

Aufgrund der Unabhängigkeit von $\mathbb{L}_{\mathbf{X},m}(\mathbf{t})$ und $\mathbb{L}_{\mathbf{Y},n}(\mathbf{t})$ folgt schließlich wie behauptet

$$\left(\sqrt{\frac{m}{m+n}} \mathbb{L}_{\mathbf{Y},n}(\mathbf{t}) - \sqrt{\frac{n}{m+n}} \mathbb{L}_{\mathbf{X},m}(\mathbf{t}), \mathbf{t} \in \overline{\mathbb{R}}_{\geq 0}^d\right) \xrightarrow{v} (\mathbb{L}(\mathbf{t}), \mathbf{t} \in \overline{\mathbb{R}}_{\geq 0}^d),$$

für $m, n \rightarrow \infty$, mit

$$(\mathbb{L}(\mathbf{t}), \mathbf{t} \in \overline{\mathbb{R}}_{\geq 0}^d) := \left(\sqrt{\tau} \mathbb{L}_{\mathbf{Y}}(\mathbf{t}) - \sqrt{1-\tau} \mathbb{L}_{\mathbf{X}}(\mathbf{t}), \mathbf{t} \in \overline{\mathbb{R}}_{\geq 0}^d\right)$$

und

$$\begin{aligned} \rho(\mathbf{s}, \mathbf{t}) &:= \text{Cov}(\mathbb{L}(\mathbf{s}), \mathbb{L}(\mathbf{t})) = \tau E(\mathbb{L}_{\mathbf{Y}}(\mathbf{s})\mathbb{L}_{\mathbf{Y}}(\mathbf{t})) + (1-\tau)E(\mathbb{L}_{\mathbf{X}}(\mathbf{s})\mathbb{L}_{\mathbf{X}}(\mathbf{t})) \\ &= \tau \rho_G(\mathbf{s}, \mathbf{t}) + (1-\tau)\rho_F(\mathbf{s}, \mathbf{t}), \quad \mathbf{s}, \mathbf{t} \in \overline{\mathbb{R}}_{\geq 0}^d. \end{aligned} \quad (2.2.6)$$

□

Die Resultate des letzten Satzes und die Zerlegung des empirischen Prozesses $\Delta_{m,n}(\mathbf{t})$ in (2.2.1) zeigen, dass für ein festes $\mathbf{t} \in \overline{\mathbb{R}}_{\geq 0}^d$ das asymptotische Verhalten von $\Delta_{m,n}(\mathbf{t})$ von dem des Terms $\sqrt{\frac{mn}{m+n}}(\mathcal{L}_{\mathbf{Y}}(\mathbf{t}) - \mathcal{L}_{\mathbf{X}}(\mathbf{t}))$ beeinflusst wird. Bei Gültigkeit der Hypothese der Laplace-Ordnung von \mathbf{X} und \mathbf{Y} treten bei festem $\mathbf{t} \in \overline{\mathbb{R}}_{\geq 0}^d$ dabei offenbar die beiden Fälle

$$\sqrt{\frac{mn}{m+n}}(\mathcal{L}_{\mathbf{Y}}(\mathbf{t}) - \mathcal{L}_{\mathbf{X}}(\mathbf{t})) \longrightarrow \begin{cases} 0, & \mathcal{L}_{\mathbf{Y}}(\mathbf{t}) - \mathcal{L}_{\mathbf{X}}(\mathbf{t}) = 0, \\ -\infty, & \mathcal{L}_{\mathbf{Y}}(\mathbf{t}) - \mathcal{L}_{\mathbf{X}}(\mathbf{t}) < 0, \end{cases} \quad m, n \rightarrow \infty,$$

auf. Mit

$$A := \{\mathbf{t} \in \mathbb{R}_{\geq 0}^d; \mathcal{L}_{\mathbf{Y}}(\mathbf{t}) = \mathcal{L}_{\mathbf{X}}(\mathbf{t})\} \quad \text{und} \quad \bar{A} := \{\mathbf{t} \in \bar{\mathbb{R}}_{\geq 0}^d; \mathcal{L}_{\mathbf{Y}}(\mathbf{t}) = \mathcal{L}_{\mathbf{X}}(\mathbf{t})\}$$

erhalten wir die im folgenden Satz formulierten Grenzwertaussagen.

2.2.5. Satz. *Bei Gültigkeit der Hypothese (2.1.1) gilt für $m, n \rightarrow \infty$*

$$\begin{aligned} T_{m,n}^{KS} &\xrightarrow{v} \sup_{\mathbf{t} \in \bar{A}} \mathbb{L}(\mathbf{t}), \\ T_{m,n}^{CvM} &\xrightarrow{v} \int_A \mathbb{L}(\mathbf{t})^+ \omega(\mathbf{t}) \, d\mathbf{t}. \end{aligned}$$

Bei Gültigkeit der Alternative ist für $m, n \rightarrow \infty$

$$T_{m,n}^{KS} \xrightarrow{P} \infty, \quad T_{m,n}^{CvM} \xrightarrow{P} \infty.$$

Beweis. Setzt man

$$\mathbb{L}_{m,n}(\mathbf{t}) = \sqrt{\frac{m}{m+n}} \mathbb{L}_{\mathbf{Y},n}(\mathbf{t}) - \sqrt{\frac{n}{m+n}} \mathbb{L}_{\mathbf{X},m}(\mathbf{t}), \quad \mathbf{t} \in \bar{\mathbb{R}}_{\geq 0}^d,$$

so ist $(\mathbb{L}_{m,n}(\mathbf{t}), \mathbf{t} \in \bar{\mathbb{R}}_{\geq 0}^d)$ nach Satz 2.2.4 verteilungskonvergent mit dem straffen und Borel-messbaren Grenzwert $(\mathbb{L}(\mathbf{t}), \mathbf{t} \in \bar{\mathbb{R}}_{\geq 0}^d)$.

Es existieren daher nach Satz B.1.6 auf einem geeigneten Wahrscheinlichkeitsraum $(\tilde{\Omega}, \tilde{\mathcal{A}}, \tilde{P})$ Zufallsvariablen $\tilde{\mathbb{L}}_{m,n}, \tilde{\mathbb{L}} : \tilde{\Omega} \rightarrow C(\bar{\mathbb{R}}_{\geq 0}^d)$ mit $\tilde{\mathbb{L}}_{m,n} \stackrel{v}{=} \mathbb{L}_{m,n}, m, n \in \mathbb{N}$, sowie $\tilde{\mathbb{L}} \stackrel{v}{=} \mathbb{L}$ und $\tilde{\mathbb{L}}_{m,n} \rightarrow \tilde{\mathbb{L}}$ fast sicher für $m, n \rightarrow \infty$.

Wir betrachten den Kolmogorov-Smirnov-Fall: Offenbar gilt

$$T_{m,n}^{KS} = \sup_{\mathbf{t} \geq 0} \Delta_{m,n}(\mathbf{t}) = \sup_{\mathbf{t} \in \bar{\mathbb{R}}_{\geq 0}^d} \Delta_{m,n}(\mathbf{t})$$

und

$$\sup_{\mathbf{t} \in \bar{\mathbb{R}}_{\geq 0}^d} \Delta_{m,n}(\mathbf{t}) \stackrel{v}{=} \sup_{\mathbf{t} \in \bar{\mathbb{R}}_{\geq 0}^d} \left(\tilde{\mathbb{L}}_{m,n}(\mathbf{t}) + \sqrt{\frac{mn}{m+n}} (\mathcal{L}_{\mathbf{Y}}(\mathbf{t}) - \mathcal{L}_{\mathbf{X}}(\mathbf{t})) \right), \quad m, n \in \mathbb{N}.$$

Nach Konstruktion gilt ferner

$$\|\tilde{\mathbb{L}}_{m,n} - \tilde{\mathbb{L}}\|_{\infty} = \sup_{\mathbf{t} \in \bar{\mathbb{R}}_{\geq 0}^d} \left| \tilde{\mathbb{L}}_{m,n}(\mathbf{t}) - \tilde{\mathbb{L}}(\mathbf{t}) \right| \xrightarrow{f.s.} 0, \quad m, n \rightarrow \infty,$$

und die Funktion $\mathbf{t} \mapsto \tilde{\mathbb{L}}(\mathbf{t})$, $\mathbf{t} \in \overline{\mathbb{R}}_{\geq 0}^d$, ist stetig. Bei Gültigkeit der Hypothese der Laplace-Ordnung $\mathbf{X} \leq_{\mathcal{L}} \mathbf{Y}$ ist wie bereits bemerkt $h(\mathbf{t}) := \mathcal{L}_{\mathbf{Y}}(\mathbf{t}) - \mathcal{L}_{\mathbf{X}}(\mathbf{t}) \leq 0$ für alle $\mathbf{t} \in \overline{\mathbb{R}}_{\geq 0}^d$. Mit Satz D.3.2 folgt somit

$$\lim_{m,n \rightarrow \infty} \sup_{\mathbf{t} \in \overline{\mathbb{R}}_{\geq 0}^d} \left(\tilde{\mathbb{L}}_{m,n}(\mathbf{t}) + \sqrt{\frac{mn}{m+n}} (\mathcal{L}_{\mathbf{Y}}(\mathbf{t}) - \mathcal{L}_{\mathbf{X}}(\mathbf{t})) \right) = \sup_{\mathbf{t} \in \overline{A}} \tilde{\mathbb{L}}(\mathbf{t})$$

fast sicher. Daher ist

$$T_{m,n}^{KS} \xrightarrow{v} \sup_{\mathbf{t} \in \overline{A}} \mathbb{L}(\mathbf{t}), \quad m, n \rightarrow \infty,$$

wie behauptet.

Wir betrachten nun die Testgröße $T_{m,n}^{CvM}$. Mit den eingangs definierten Zufallsvariablen $\tilde{\mathbb{L}}_{m,n}, \tilde{\mathbb{L}}$ erhalten wir

$$T_{m,n}^{CvM} \stackrel{v}{=} \int_{\overline{\mathbb{R}}_{\geq 0}^d} \left(\tilde{\mathbb{L}}_{m,n}(\mathbf{t}) + \sqrt{\frac{mn}{m+n}} (\mathcal{L}_{\mathbf{Y}}(\mathbf{t}) - \mathcal{L}_{\mathbf{X}}(\mathbf{t})) \right)^+ \omega(\mathbf{t}) d\mathbf{t}.$$

Die Integranden sind für jedes $m, n \in \mathbb{N}$ fast sicher beschränkt und es gilt wegen

$$\lim_{m,n \rightarrow \infty} \left(\tilde{\mathbb{L}}_{m,n}(\mathbf{t}) + \sqrt{\frac{mn}{m+n}} h(\mathbf{t}) \right) \omega(\mathbf{t}) = \begin{cases} \tilde{\mathbb{L}}(\mathbf{t}) \omega(\mathbf{t}), & \mathbf{t} \in A, \\ -\infty, & \mathbf{t} \in \{\mathbf{t} \in \overline{\mathbb{R}}_{\geq 0}^d; h(\mathbf{t}) < 0\}, \end{cases}$$

fast sicher auch $\lim_{m,n \rightarrow \infty} \left(\tilde{\mathbb{L}}_{m,n}(\mathbf{t}) + \sqrt{\frac{mn}{m+n}} h(\mathbf{t}) \right)^+ \omega(\mathbf{t}) = \tilde{\mathbb{L}}^+(\mathbf{t}) \omega(\mathbf{t}) \mathbf{1}_A(\mathbf{t})$ fast sicher für jedes $\mathbf{t} \in \overline{\mathbb{R}}_{\geq 0}^d$. Bei Gültigkeit der Hypothese folgt daher mit dem Satz von der majorisierten Konvergenz schließlich

$$\begin{aligned} & \lim_{m,n \rightarrow \infty} \int_{\overline{\mathbb{R}}_{\geq 0}^d} \left(\tilde{\mathbb{L}}_{m,n}(\mathbf{t}) + \sqrt{\frac{mn}{m+n}} h(\mathbf{t}) \right)^+ \omega(\mathbf{t}) d\mathbf{t} \\ &= \int_{\overline{\mathbb{R}}_{\geq 0}^d} \lim_{m,n \rightarrow \infty} \left(\tilde{\mathbb{L}}_{m,n}(\mathbf{t}) + \sqrt{\frac{mn}{m+n}} h(\mathbf{t}) \right)^+ \omega(\mathbf{t}) d\mathbf{t} \\ &= \int_A \mathbb{L}(\mathbf{t})^+ \omega(\mathbf{t}) d\mathbf{t} \end{aligned}$$

fast sicher und daher die behauptete Verteilungskonvergenz für $T_{m,n}^{CvM}$ bei Gültigkeit der Hypothese.

Bei Gültigkeit der Alternative existiert zumindest ein $\mathbf{t} \in \overline{\mathbb{R}}_{\geq 0}^d$ mit $h(\mathbf{t}) > 0$, was die Behauptung $T_{m,n}^{KS} \xrightarrow{P} \infty$ für $m, n \rightarrow \infty$ im Kolmogorov-Smirnov-Fall zeigt. Aufgrund der Stetigkeit von h gilt dann außerdem $h(\mathbf{t}) > 0$ in einer Umgebung von \mathbf{t} , so dass schließlich für $m, n \rightarrow \infty$ die Behauptung $T_{m,n}^{CvM} \xrightarrow{P} \infty$ im Cramér-von Mises Fall folgt. \square

2.2.6. Bemerkung. Die Zufallsvariablen $\sup_{\mathbf{t} \in \bar{A}} \mathbb{L}(\mathbf{t})$ und $\sup_{\mathbf{t} \in A} \mathbb{L}(\mathbf{t})$ haben im Allgemeinen nicht dieselbe Verteilung, denn aufgrund von Satz 2.2.4 und der Definition der Fortsetzung der Laplacetransformierten auf $\bar{\mathbb{R}}_{\geq 0}^d$ in Lemma 2.2.1 wird deutlich, dass für ein $\mathbf{t} \in \bar{A} \setminus A$ auch $\text{Var}(\mathbb{L}(\mathbf{t})) \neq 0$ zutreffen kann. In der Tat, für jedes $\mathbf{t} \in \bar{\mathbb{R}}_{\geq 0}^d \setminus \mathbb{R}_{\geq 0}^d$ erhalten wir

$$\mathcal{L}_{\mathbf{X}}(\mathbf{t}) = E \left(\prod_{j \in I(\mathbf{t})} \mathbf{1}_{\{X_j=0\}} \exp \left(- \sum_{j \notin I(\mathbf{t})} t_j X_j \right) \right) \quad (2.2.7)$$

und somit im Allgemeinen nicht $\mathcal{L}_{\mathbf{X}}(\mathbf{t}) = 0$ für $\mathbf{t} \in \bar{\mathbb{R}}_{\geq 0}^d \setminus \mathbb{R}_{\geq 0}^d$.

Gilt $P(X_k = 0) = 0$ für alle Komponenten X_1, \dots, X_d von \mathbf{X} , so folgt aus (2.2.7) allerdings $\text{Var}(\mathbb{L}(\mathbf{t})) = 0$ für alle $\mathbf{t} \in \bar{A} \setminus A$. Insbesondere haben die Zufallsvariablen $\sup_{\mathbf{t} \in \bar{A}} \mathbb{L}(\mathbf{t})$ und $\sup_{\mathbf{t} \in A} \mathbb{L}(\mathbf{t})$ dann dieselbe Verteilung und es gilt bei Gültigkeit der Hypothese

$$T_{m,n}^{KS} \xrightarrow{v} \sup_{\mathbf{t} \in A} \mathbb{L}(\mathbf{t}), \quad m, n \rightarrow \infty.$$

Die Grenzverteilung der vorgestellten Testgrößen hängt also sowohl über die Kovarianzfunktion der Grenzprozesse als auch über die Mengen A bzw. \bar{A} von den zugrunde liegenden unbekanntten Verteilungen der Zufallsvektoren $\mathbf{X}_1, \dots, \mathbf{X}_m$ bzw. $\mathbf{Y}_1, \dots, \mathbf{Y}_n$ ab.

Die analytischen Eigenschaften der Laplacetransformierten liefern allerdings die im folgenden Satz formulierte Eigenschaft der Menge A in dem Fall, dass die Verteilungen der Zufallsvektoren \mathbf{X} und \mathbf{Y} nicht gleich sind. Diese Eigenschaft wird sich besonders im Zusammenhang mit dem Cramér-von Mises-Test als nützlich erweisen.

2.2.7. Satz. *Es sei $F \neq G$. Dann ist $A \subset \mathbb{R}_{\geq 0}^d$ eine Nullmenge bezüglich des d -dimensionalen Lebesgue-Borelschen Maßes.*

Beweis. Die Laplacetransformierten $\mathcal{L}_F(\mathbf{t})$ und $\mathcal{L}_G(\mathbf{t})$ sind jeweils analytische Funktionen auf $\mathbb{R}_{> 0}^d = (0, \infty)^d$ und somit ist die Differenzfunktion $\Delta(\mathbf{t}) = \mathcal{L}_F(\mathbf{t}) - \mathcal{L}_G(\mathbf{t})$ analytisch auf $\mathbb{R}_{> 0}^d$. Wegen $F \neq G$ ist $\Delta(\mathbf{t}) \not\equiv 0$. Es ist $\mathbb{R}_{> 0}^d \subset \mathbb{R}^d$ offen, so dass die Menge $A^\circ = \{\mathbf{t} \in \mathbb{R}_{> 0}^d : \Delta(\mathbf{t}) = 0\}$ nach Satz D.3.1 das d -dimensionale Lebesgue-Maß Null hat. Ferner ist $\mathbb{R}_{\geq 0}^d \setminus \mathbb{R}_{> 0}^d$ als endliche Vereinigung von Hyperflächen im \mathbb{R}^d ebenfalls vom d -dimensionalen Lebesgue-Maß Null, also auch A vom d -dimensionalen Lebesgue-Maß Null. \square

Der letzte Satz und die Aussagen in Satz 2.2.5 liefern daher das bemerkenswerte Resultat, dass die Grenzverteilung der Cramér-von Mises-Testgröße in sämtlichen Fällen, in

denen $F \leq_{\mathcal{L}} G$ und $F \neq G$ gilt, die Einpunktverteilung im Nullpunkt ist. Naturgemäß impliziert dieses Resultat ein spezielles Verhalten der Güte bei der Durchführung des Tests. Wir werden diesen Sachverhalt am Ende des nächsten Abschnittes präzisieren.

Wegen $\mathbb{L}(\mathbf{0}) \equiv 0$ fast sicher und $\mathbf{0} \in \bar{A}$ stellt sich ebenso die Frage, ob aus $F \leq_{\mathcal{L}} G$ und $F \neq G$ stets die Degeneriertheit der Grenzverteilung der Kolmogorov-Smirnov-Testgröße folgt. Zunächst ist klar, dass sich etwa für reelle Zufallsvariablen X, Y mit den Verteilungsfunktionen F, G und den Eigenschaften $X \leq_{\mathcal{L}} Y$ sowie $F \neq G$ sofort

$$\mathcal{L}_{(X,0)}(\mathbf{t}) - \mathcal{L}_{(Y,0)}(\mathbf{t}) = 0 \quad \text{für alle } \mathbf{t} \in \{(0, t) : t \in \mathbb{R}_{\geq 0}\}$$

ergibt. Das nachfolgende Beispiel zeigt, dass aber auch Zufallsvariablen mit $X \leq_{\mathcal{L}} Y$, $F \neq G$ und $P(X = 0) = P(Y = 0) = 0$ existieren, für die die Grenzverteilung der Kolmogorov-Smirnov-Testgröße nicht degeneriert ist.

2.2.8. Beispiel. Wir untersuchen die diskret verteilten Zufallsvariablen X, Y mit

$$X \sim \frac{2}{64}\delta_1 + \frac{1}{64}\delta_3 + \frac{24}{64}\delta_5 + \frac{2}{64}\delta_7 + \frac{35}{64}\delta_9,$$

$$Y \sim \frac{1}{64}\delta_1 + \frac{9}{64}\delta_3 + \frac{2}{64}\delta_5 + \frac{26}{64}\delta_7 + \frac{26}{64}\delta_9.$$

Dann gilt

$$\Delta(t) := \mathcal{L}_X(t) - \mathcal{L}_Y(t) = \exp(-t) \left(\frac{1}{8} - \frac{1}{2} \exp(-2t) + \frac{3}{8} \exp(-4t) \right)^2, \quad t \geq 0.$$

Es folgt $X \leq_{\mathcal{L}} Y$ und ferner gilt $\Delta(t) = 0$ für $t \geq 0$ genau dann, wenn $t \in \{0, \frac{1}{2} \ln(3)\}$ ist. Also ergibt sich $A = \{0, \frac{1}{2} \ln(3)\}$. Abbildung 2.1 zeigt den Plot von $\Delta(t)$ für $t \in [0, 4]$. Für Zufallsvariablen mit der angegebenen Verteilung konvergiert die Kolmogorov-Smirnov-Testgröße nach Satz 2.2.5 also nicht gegen eine degenerierte Verteilung, denn mit den dortigen Bezeichnungen gilt für jedes $t \in A$

$$\begin{aligned} \text{Var}(\mathbb{L}(t)) &= \tau (\mathcal{L}_Y(2t) - \mathcal{L}_Y(t)^2) + (1 - \tau) (\mathcal{L}_X(2t) - \mathcal{L}_X(t)^2) \\ &= \tau \mathcal{L}_Y(2t) + (1 - \tau) \mathcal{L}_X(2t) - \mathcal{L}_X(t)^2, \end{aligned}$$

so dass sich hier $\text{Var}(\mathbb{L}(\frac{1}{2} \ln(3))) \neq 0$ für alle $\tau \in [0, 1]$ ergibt. Folglich erhalten wir in diesem Fall wegen $P(X = 0) = P(Y = 0) = 0$, sowie mit $t_0 := \frac{1}{2} \ln(3)$

$$T_{m,n}^{KS} \rightarrow \sup_{t \in \{0, \frac{1}{2} \ln(3)\}} \mathbb{L}(t) \stackrel{f.s.}{=} \mathbb{L}^+(t_0), \quad m, n \rightarrow \infty.$$

Auf der Basis der Zufallsvariablen X, Y lassen sich außerdem sofort weitere Zufallsvariablen konstruieren, die nicht rein diskret verteilt sind und ebenfalls Beispiele für das oben genannte Verhalten liefern. Ist etwa Z eine Zufallsvariable mit stetiger Verteilung und Werten in $\mathbb{R}_{\geq 0}$, Z unabhängig von X und Y , so folgt aus Satz 1.2.2, Teil c), sofort $X + Z \leq_{\mathcal{L}} Y + Z$. Ferner berühren sich die Laplacetransformierten von $X + Z$ und $Y + Z$ in $t = \frac{1}{2} \ln(3)$.

Der oben beschriebene Sachverhalt zeigt erste Unterschiede im asymptotischen Verhalten der beiden vorgestellten Testgrößen bei Gültigkeit der Hypothese auf. Die Grenzverteilung der Cramér-von Mises-Testgröße ist in einem größeren Bereich der Hypothese degeneriert, denn die Grenzverteilung der Kolmogorov-Smirnov-Testgröße ist genau dann die Einpunktverteilung im Nullpunkt, wenn $\text{Var}(\mathbb{L}(\mathbf{t})) = 0$ für alle $\mathbf{t} \in \bar{A}$ gilt. Ist $F = G$ (also $\bar{A} = \overline{\mathbb{R}_{\geq 0}^d}$), so ist die Grenzverteilung der Kolmogorov-Smirnov-Testgröße aufgrund der vorausgesetzten Nichtdegeneriert der Verteilungsfunktionen F und G daher nicht degeneriert.

In den folgenden Abschnitten werden sich weitere Unterschiede im Verhalten der beiden Testgrößen herausstellen. Obwohl sich die Cramér-von Mises-Testgröße an dieser Stelle in dem zuvor dargelegten Sinne der Kolmogorov-Smirnov-Testgröße gegenüber in gewisser Weise als vorteilhaft erweist (die Degeneriertheit der Grenzverteilung der Cramér-von Mises-Testgröße in einem größeren Bereich der Hypothese lässt hoffen, dass der darauf basierende Test das Testniveau eher einhält), werden wir speziell im Hinblick auf eine praktische Umsetzung der Testverfahren später auch Vorteile der Supremumstestgröße erkennen.

Zuvor wollen wir im nächsten Abschnitt aber zunächst die für die Durchführung der Tests notwendigen kritischen Werte bereitstellen. Wie bereits erwähnt, hängen die Grenzverteilungen der Testgrößen bei Gültigkeit der Hypothese, also die Größen

$$T^{KS} := \sup_{\mathbf{t} \in \bar{A}} \mathbb{L}(\mathbf{t}) \quad \text{und} \quad T^{CvM} := \int_A \mathbb{L}(\mathbf{t})^+ \omega(\mathbf{t}) d\mathbf{t},$$

dabei über die Kovarianzfunktion des Prozesses $(\mathbb{L}(\mathbf{t}), \mathbf{t} \in \overline{\mathbb{R}_{\geq 0}^d})$ und über die Identitätsmengen A bzw. \bar{A} von den zugrunde liegenden unbekanntenen Verteilungen der Zufallsvektoren \mathbf{X} und \mathbf{Y} ab.

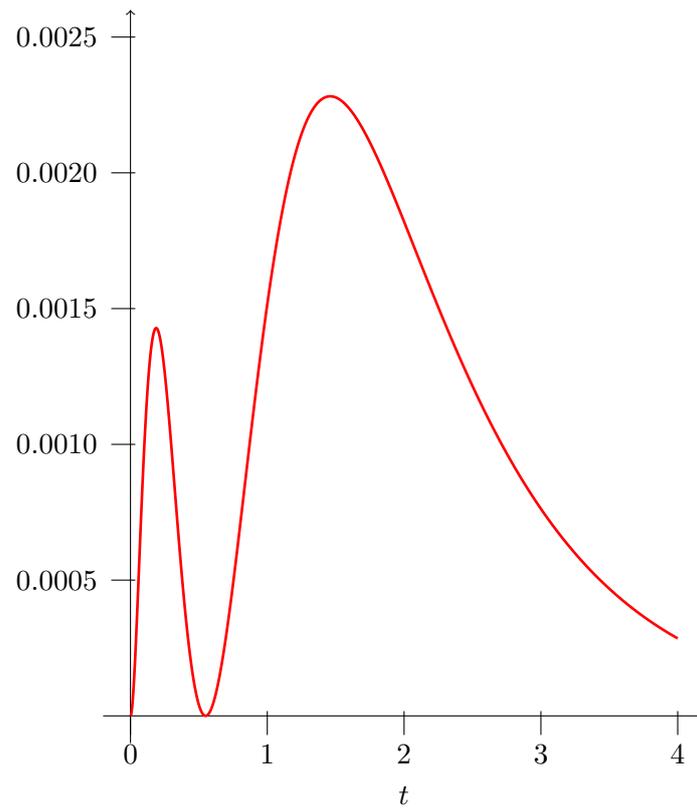


Abbildung 2.1: Plot der Differenz $\Delta(t)$ der Laplacetransformierten von X und Y . Es ist $X \leq_{\mathcal{L}} Y$ und $A = \{0, \frac{1}{2} \ln(3)\}$, die Laplacetransformierten haben daher einen von Null verschiedenen Berührungspunkt.

2.3 Kritische Werte über den Bootstrap

Die Aussagen des letzten Abschnitts zeigen, dass die Grenzverteilungen der Testgrößen bei Gültigkeit der Hypothese nicht verteilungsfrei sind. Die Quantile dieser Grenzverteilungen sind also nicht ohne Weiteres verfügbar, so dass es sinnvoll erscheint die kritischen Werte zur Durchführung des Tests über ein Bootstrap- und Simulationsverfahren zu ermitteln. In diesem Abschnitt soll dazu ein mögliches Bootstrapverfahren vorgestellt werden.

Wir bemerken zunächst, dass das Vorliegen von Ordnungsbeschränkungen im Bereich der Hypothese dazu führen kann, dass das Standard-Bootstrapverfahren im Zweistichprobenfall, also das Resampling anhand der Gleichverteilung auf der zusammengesetzten Stichprobe, kritische Werte liefert, die auch asymptotisch zu einer Nicht-Einhaltung des vorgegeben Testniveaus führen (siehe [BG09]). Seien

$$F_m(\mathbf{t}) = \frac{1}{m} \sum_{k=1}^m \mathbf{1}_{(-\infty, \mathbf{t}]}(\mathbf{X}_k), \quad G_n(\mathbf{t}) = \frac{1}{n} \sum_{k=1}^n \mathbf{1}_{(-\infty, \mathbf{t}]}(\mathbf{Y}_k)$$

die empirischen Verteilungsfunktionen zu $\mathbf{X}_1, \dots, \mathbf{X}_m$ bzw. $\mathbf{Y}_1, \dots, \mathbf{Y}_n$ und

$$H_{m,n}(\mathbf{t}) := \frac{n}{m+n} F_m(\mathbf{t}) + \frac{m}{m+n} G_n(\mathbf{t}), \quad \mathbf{t} \in \mathbb{R}^d. \quad (2.3.1)$$

Bei gegebenen $\mathbf{X}_1, \dots, \mathbf{X}_m, \mathbf{Y}_1, \dots, \mathbf{Y}_n$ seien $\hat{\mathbf{Z}}_{N1}, \dots, \hat{\mathbf{Z}}_{NN}$, $N = m+n$, unabhängige Zufallsvariablen mit der Verteilungsfunktion $H_{m,n}$. Dann ist $\hat{\mathbf{Z}}_{N1}, \dots, \hat{\mathbf{Z}}_{NN}$ für jedes $m, n \in \mathbb{N}$ ein Resample von $\mathbf{X}_1, \dots, \mathbf{X}_m, \mathbf{Y}_1, \dots, \mathbf{Y}_n$. Für feste $m, n \in \mathbb{N}$ bezeichnen wir mit $\hat{G}_{N,n}$ die empirische Verteilungsfunktion zu $\hat{\mathbf{Z}}_{N1}, \dots, \hat{\mathbf{Z}}_{Nn}$ und mit $\hat{F}_{N,m}$ die empirische Verteilungsfunktion zu $\hat{\mathbf{Z}}_{N(n+1)}, \dots, \hat{\mathbf{Z}}_{NN}$. Ferner seien

$$\mathcal{L}_{\hat{G}_{N,n}}(\mathbf{t}) = \frac{1}{n} \sum_{k=1}^n e^{-\mathbf{t}' \hat{\mathbf{Z}}_{Nk}}, \quad \mathcal{L}_{\hat{F}_{N,m}}(\mathbf{t}) = \frac{1}{m} \sum_{k=1}^m e^{-\mathbf{t}' \hat{\mathbf{Z}}_{N(n+k)}}$$

die zu $\hat{\mathbf{Z}}_{N1}, \dots, \hat{\mathbf{Z}}_{Nn}$ und $\hat{\mathbf{Z}}_{N(n+1)}, \dots, \hat{\mathbf{Z}}_{NN}$ gehörigen empirischen Laplacetransformierten. Wir definieren

$$\hat{\Delta}_{m,n}(\mathbf{t}) := \sqrt{\frac{mn}{m+n}} \left(\mathcal{L}_{\hat{G}_{N,n}}(\mathbf{t}) - \mathcal{L}_{\hat{F}_{N,m}}(\mathbf{t}) \right), \quad \mathbf{t} \geq 0,$$

sowie

$$\hat{T}_{m,n}^{KS} := \sup_{\mathbf{t} \geq 0} \hat{\Delta}_{m,n}(\mathbf{t}),$$

$$\hat{T}_{m,n}^{CuM} := \int_{\mathbb{R}_{\geq 0}^d} \hat{\Delta}_{m,n}^+(\mathbf{t}) \omega(\mathbf{t}) d\mathbf{t}.$$

Die nachstehende Grenzwertaussage liefert das asymptotische Verhalten der Testgrößen unter dem Einfluss des vorgeschlagenen Bootstrapverfahrens.

2.3.1. Satz. Sei $(\mathbb{L}_H(\mathbf{t}), \mathbf{t} \in \overline{\mathbb{R}}_{\geq 0}^d)$ ein zentrierter Gaußprozess mit Pfaden in $C(\overline{\mathbb{R}}_{\geq 0}^d)$ und der Kovarianzfunktion

$$\rho_H(\mathbf{s}, \mathbf{t}) := \mathcal{L}_H(\mathbf{s} + \mathbf{t}) - \mathcal{L}_H(\mathbf{s})\mathcal{L}_H(\mathbf{t}), \quad \mathbf{s}, \mathbf{t} \in \overline{\mathbb{R}}_{\geq 0}^d,$$

mit \mathcal{L}_H als der stetigen Fortsetzung der Laplacetransformierten der Verteilungsfunktion $H(\mathbf{t}) := (1 - \tau)F(\mathbf{t}) + \tau G(\mathbf{t})$, $\mathbf{t} \in \mathbb{R}_{\geq 0}^d$, auf den Bereich $\overline{\mathbb{R}}_{\geq 0}^d$. Dann gilt für fast alle $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{Y}_1, \mathbf{Y}_2, \dots$

$$\hat{T}_{m,n}^{KS} \xrightarrow{v} \sup_{\mathbf{t} \in \overline{\mathbb{R}}_{\geq 0}^d} \mathbb{L}_H(\mathbf{t}), \quad m, n \rightarrow \infty,$$

$$\hat{T}_{m,n}^{CvM} \xrightarrow{v} \int_{\overline{\mathbb{R}}_{\geq 0}^d} \mathbb{L}_H^+(\mathbf{t}) \omega(\mathbf{t}) d\mathbf{t}, \quad m, n \rightarrow \infty.$$

Beweis. Wir betrachten den Prozess $\hat{\Delta}_{m,n}$ erneut als Prozess auf $\overline{\mathbb{R}}_{\geq 0}^d$, indem wir die empirischen Laplacetransformierten gemäß (2.2.3) stetig fortsetzen. Für $\mathbf{t} \in \overline{\mathbb{R}}_{\geq 0}^d$ ist dann zunächst

$$\hat{\Delta}_{m,n}(\mathbf{t}) = \sqrt{\frac{m}{m+n}} \sqrt{n} \left(\mathcal{L}_{\hat{G}_{N,n}}(\mathbf{t}) - \mathcal{L}_{H_{m,n}}(\mathbf{t}) \right) - \sqrt{\frac{n}{m+n}} \sqrt{m} \left(\mathcal{L}_{\hat{F}_{N,m}}(\mathbf{t}) - \mathcal{L}_{H_{m,n}}(\mathbf{t}) \right).$$

Für festes $m, n \in \mathbb{N}$ und $f_{\mathbf{t}}(\mathbf{x}) := \exp(\mathbf{t}'\mathbf{x})$ ist

$$\sqrt{n} \left(\mathcal{L}_{\hat{G}_{N,n}}(\mathbf{t}) - \mathcal{L}_{H_{m,n}}(\mathbf{t}) \right) = \sqrt{n} \left(\frac{1}{n} \sum_{i=1}^n f_{\mathbf{t}}(\mathbf{Z}_{N_i}) - E f_{\mathbf{t}}(\mathbf{Z}_{N_1}) \right),$$

also $\left(\sqrt{n} \left(\mathcal{L}_{\hat{G}_{N,n}}(\mathbf{t}) - \mathcal{L}_{H_{m,n}}(\mathbf{t}) \right), \mathbf{t} \in \overline{\mathbb{R}}_{\geq 0}^d \right)$ der durch \mathcal{F} indizierte empirische Prozess der $\hat{\mathbf{Z}}_{N_1}, \dots, \hat{\mathbf{Z}}_{N_n}$. Nach Lemma 2.2.3 und Satz B.2.1.5 erfüllt \mathcal{F} die Entropiebedingung (B.2.1) aus Kapitel B.2.1. Es ist $E \equiv 1$ eine Envelope von \mathcal{F} . Offenbar ist E quadratintegrierbar bezüglich jedem $H_{m,n}$ mit der Eigenschaft, dass fast sicher

$$\forall \varepsilon \geq 0 : \limsup_{m,n \rightarrow \infty} \int_{\mathbb{R}^d} E^2(\mathbf{x}) \mathbf{1}_{\{E \geq \varepsilon \sqrt{n}\}} dH_{m,n}(\mathbf{x}) = 0,$$

für fast alle $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{Y}_1, \mathbf{Y}_2, \dots$ ist. Es gilt dann mit dem Satz von Glivenko-Cantelli in der Form von Satz B.2.1.6

$$\begin{aligned} \forall \mathbf{t} \in \overline{\mathbb{R}}_{\geq 0}^d : \lim_{m,n \rightarrow \infty} H_{m,n}(\mathbf{t}) &= \lim_{m,n \rightarrow \infty} \left(\frac{n}{m+n} F_m(\mathbf{t}) + \frac{m}{m+n} G_n(\mathbf{t}) \right) \\ &= (1 - \tau)F(\mathbf{t}) + \tau G(\mathbf{t}) \end{aligned}$$

für fast alle $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{Y}_1, \mathbf{Y}_2, \dots$, so dass im Hinblick auf Satz B.2.2.1 und Bemerkung B.2.2.2 zu zeigen bleibt, dass für fast alle $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{Y}_1, \mathbf{Y}_2, \dots$

$$\sup_{f, g \in \mathcal{F}} |L_2(H_{m,n})(f - g) - L_2(H)(f - g)| \rightarrow 0, \quad m, n \rightarrow \infty,$$

zutritt. Es ist

$$\begin{aligned} & \sup_{f, g \in \mathcal{F}} |L_2(H_{m,n})(f - g) - L_2(H)(f - g)| \\ &= \sup_{f, g \in \mathcal{F}} \left| \left(\int_{\mathbb{R}_{\geq 0}^d} (f(\mathbf{x}) - g(\mathbf{x}))^2 dH_{m,n}(\mathbf{x}) \right)^{\frac{1}{2}} - \left(\int_{\mathbb{R}_{\geq 0}^d} (f(\mathbf{x}) - g(\mathbf{x}))^2 dH(\mathbf{x}) \right)^{\frac{1}{2}} \right| \\ &= \sup_{\mathbf{s}, \mathbf{t} \in \overline{\mathbb{R}}_{\geq 0}^d} \left| \left(\int_{\mathbb{R}_{\geq 0}^d} (f_{\mathbf{s}}(\mathbf{x}) - f_{\mathbf{t}}(\mathbf{x}))^2 dH_{m,n}(\mathbf{x}) \right)^{\frac{1}{2}} - \left(\int_{\mathbb{R}_{\geq 0}^d} (f_{\mathbf{s}}(\mathbf{x}) - f_{\mathbf{t}}(\mathbf{x}))^2 dH(\mathbf{x}) \right)^{\frac{1}{2}} \right|. \end{aligned}$$

Wir setzen abkürzend $h_{\mathbf{s}, \mathbf{t}}(\mathbf{x}) = f_{\mathbf{s}}(\mathbf{x}) - f_{\mathbf{t}}(\mathbf{x})$ für alle $\mathbf{s}, \mathbf{t} \in \overline{\mathbb{R}}_{\geq 0}^d$. Dann gilt wegen $|h_{\mathbf{s}, \mathbf{t}}| \leq 1$ auch $\int_{\mathbb{R}_{\geq 0}^d} h_{\mathbf{s}, \mathbf{t}}^2(\mathbf{x}) dH_{m,n}(\mathbf{x}) \leq 1$ und $\int_{\mathbb{R}_{\geq 0}^d} h_{\mathbf{s}, \mathbf{t}}^2(\mathbf{x}) dH(\mathbf{x}) \leq 1$ für alle $\mathbf{s}, \mathbf{t} \in \overline{\mathbb{R}}_{\geq 0}^d$. Damit folgt

$$\begin{aligned} & \sup_{\mathbf{s}, \mathbf{t} \in \overline{\mathbb{R}}_{\geq 0}^d} \left| \left(\int_{\mathbb{R}_{\geq 0}^d} h_{\mathbf{s}, \mathbf{t}}^2(\mathbf{x}) dH_{m,n}(\mathbf{x}) \right)^{\frac{1}{2}} - \left(\int_{\mathbb{R}_{\geq 0}^d} h_{\mathbf{s}, \mathbf{t}}^2(\mathbf{x}) dH(\mathbf{x}) \right)^{\frac{1}{2}} \right| \\ & \leq \sup_{\mathbf{s}, \mathbf{t} \in \overline{\mathbb{R}}_{\geq 0}^d} \left(\left| \int_{\mathbb{R}_{\geq 0}^d} h_{\mathbf{s}, \mathbf{t}}^2(\mathbf{x}) dH_{m,n}(\mathbf{x}) - \int_{\mathbb{R}_{\geq 0}^d} h_{\mathbf{s}, \mathbf{t}}^2(\mathbf{x}) dH(\mathbf{x}) \right|^{\frac{1}{2}} \right) \\ & = \left(\sup_{\mathbf{s}, \mathbf{t} \in \overline{\mathbb{R}}_{\geq 0}^d} \left| \int_{\mathbb{R}_{\geq 0}^d} h_{\mathbf{s}, \mathbf{t}}^2(\mathbf{x}) dH_{m,n}(\mathbf{x}) - \int_{\mathbb{R}_{\geq 0}^d} h_{\mathbf{s}, \mathbf{t}}^2(\mathbf{x}) dH(\mathbf{x}) \right| \right)^{\frac{1}{2}}. \end{aligned}$$

Außerdem erhalten wir

$$\int_{\mathbb{R}_{\geq 0}^d} h_{\mathbf{s}, \mathbf{t}}^2(\mathbf{x}) dH_{m,n}(\mathbf{x}) = \mathcal{L}_{H_{m,n}}(2\mathbf{t}) - 2\mathcal{L}_{H_{m,n}}(\mathbf{t} + \mathbf{s}) + \mathcal{L}_{H_{m,n}}(2\mathbf{s})$$

und

$$\int_{\mathbb{R}_{\geq 0}^d} h_{\mathbf{s}, \mathbf{t}}^2(\mathbf{x}) dH(\mathbf{x}) = \mathcal{L}_H(2\mathbf{t}) - 2\mathcal{L}_H(\mathbf{t} + \mathbf{s}) + \mathcal{L}_H(2\mathbf{s}).$$

Es ergibt sich schließlich

$$\begin{aligned} & \sup_{\mathbf{s}, \mathbf{t} \in \overline{\mathbb{R}}_{\geq 0}^d} \left| \int_{\mathbb{R}_{\geq 0}^d} h_{\mathbf{s}, \mathbf{t}}^2(\mathbf{x}) dH_{m,n}(\mathbf{x}) - \int_{\mathbb{R}_{\geq 0}^d} h_{\mathbf{s}, \mathbf{t}}^2(\mathbf{x}) dH(\mathbf{x}) \right| \leq \sup_{\mathbf{t} \in \overline{\mathbb{R}}_{\geq 0}^d} |\mathcal{L}_{H_{m,n}}(2\mathbf{t}) - \mathcal{L}_H(2\mathbf{t})| \\ & \quad + 2 \sup_{\mathbf{s}, \mathbf{t} \in \overline{\mathbb{R}}_{\geq 0}^d} |\mathcal{L}_{H_{m,n}}(\mathbf{s} + \mathbf{t}) - \mathcal{L}_H(\mathbf{s} + \mathbf{t})| + \sup_{\mathbf{s} \in \overline{\mathbb{R}}_{\geq 0}^d} |\mathcal{L}_{H_{m,n}}(2\mathbf{s}) - \mathcal{L}_H(2\mathbf{s})|. \quad (2.3.2) \end{aligned}$$

Die Laplacetransformierte von $H_{m,n}$ konvergiert für $m, n \rightarrow \infty$ nach Satz B.2.1.6 fast sicher gleichmäßig gegen die Laplacetransformierte von H , so dass der Ausdruck in (2.3.2) für $m, n \rightarrow \infty$ fast sicher gegen Null konvergiert. Wir erhalten also mit Satz B.2.2.1, dass

$$\left(\sqrt{n} \left(\mathcal{L}_{\hat{G}_{N,n}}(\mathbf{t}) - \mathcal{L}_{H_{m,n}}(\mathbf{t})\right), \mathbf{t} \in \overline{\mathbb{R}}_{\geq 0}^d\right) \xrightarrow{v} (\mathbb{L}_H(\mathbf{t}), \mathbf{t} \in \overline{\mathbb{R}}_{\geq 0}^d), \quad m, n \rightarrow \infty,$$

für fast alle $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{Y}_1, \mathbf{Y}_2, \dots$ gilt und der Gaußprozess $(\mathbb{L}_H(\mathbf{t}), \mathbf{t} \in \overline{\mathbb{R}}_{\geq 0}^d)$ die Kovarianzfunktion

$$\begin{aligned} \rho_H(\mathbf{s}, \mathbf{t}) &= \int_{\mathbb{R}_{\geq 0}^d} f_{\mathbf{s}}(\mathbf{x}) f_{\mathbf{t}}(\mathbf{x}) dH(\mathbf{x}) - \int_{\mathbb{R}_{\geq 0}^d} f_{\mathbf{s}}(\mathbf{x}) dH(\mathbf{x}) \int_{\mathbb{R}_{\geq 0}^d} f_{\mathbf{t}}(\mathbf{x}) dH(\mathbf{x}) \\ &= \mathcal{L}_H(\mathbf{s} + \mathbf{t}) - \mathcal{L}_H(\mathbf{s}) \mathcal{L}_H(\mathbf{t}), \quad \mathbf{s}, \mathbf{t} \in \overline{\mathbb{R}}_{\geq 0}^d, \end{aligned}$$

hat. Analog erhalten wir für den zweiten empirischen Prozess

$$\left(\sqrt{m} \left(\mathcal{L}_{\hat{F}_{N,m}}(\mathbf{t}) - \mathcal{L}_{H_{m,n}}(\mathbf{t})\right), \mathbf{t} \in \overline{\mathbb{R}}_{\geq 0}^d\right) \xrightarrow{v} (\mathbb{L}_H(\mathbf{t}), \mathbf{t} \in \overline{\mathbb{R}}_{\geq 0}^d), \quad m, n \rightarrow \infty,$$

für fast alle $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{Y}_1, \mathbf{Y}_2, \dots$. Aufgrund der Unabhängigkeit der empirischen Prozesse bei gegebenen $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{Y}_1, \mathbf{Y}_2, \dots$ folgt

$$(\hat{\Delta}_{m,n}(\mathbf{t}), \mathbf{t} \in \overline{\mathbb{R}}_{\geq 0}^d) \xrightarrow{v} (\mathbb{L}_H(\mathbf{t}), \mathbf{t} \in \overline{\mathbb{R}}_{\geq 0}^d), \quad m, n \rightarrow \infty,$$

für fast alle $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{Y}_1, \mathbf{Y}_2, \dots$. Analog zu Satz 2.2.5 ergeben sich durch Heranziehen fast sicher konvergenter Versionen der betrachteten Prozesse dann die Behauptungen für $\hat{T}_{m,n}^{KS}$ und $\hat{T}_{m,n}^{CvM}$. Aufgrund der Stetigkeit der Pfade von \mathbb{L}_H und da $\mathbb{R}_{\geq 0}^d \subset \overline{\mathbb{R}}_{\geq 0}^d$ dicht ist, ergibt sich ferner $\sup_{\mathbf{t} \in \mathbb{R}_{\geq 0}^d} \mathbb{L}_H(\mathbf{t}) \stackrel{v}{=} \sup_{\mathbf{t} \in \overline{\mathbb{R}}_{\geq 0}^d} \mathbb{L}_H(\mathbf{t})$. \square

Für die im vorherigen Satz auftretenden Grenzverteilungen setzen wir im Folgenden

$$\hat{T}^{KS} := \sup_{\mathbf{t} \in \overline{\mathbb{R}}_{\geq 0}^d} \mathbb{L}_H(\mathbf{t}), \quad \hat{T}^{CvM} := \int_{\mathbb{R}_{\geq 0}^d} \mathbb{L}_H^+(\mathbf{t}) \omega(\mathbf{t}) d\mathbf{t}.$$

Die in der nächsten Bemerkung zusammengefassten Überlegungen rechtfertigen dann die getroffene Wahl des Bootstrapverfahrens in (2.3.1).

2.3.2. Bemerkung. Der Prozess $\mathbb{L}_H(\mathbf{t})$ aus Satz 2.3.1 hat die Kovarianzfunktion

$$\begin{aligned} \rho_H(\mathbf{s}, \mathbf{t}) &= (1 - \tau) \mathcal{L}_F(\mathbf{s} + \mathbf{t}) + \tau \mathcal{L}_G(\mathbf{s} + \mathbf{t}) \\ &\quad - \left((1 - \tau) \mathcal{L}_F(\mathbf{s}) + \tau \mathcal{L}_G(\mathbf{s}) \right) \left((1 - \tau) \mathcal{L}_F(\mathbf{t}) + \tau \mathcal{L}_G(\mathbf{t}) \right) \end{aligned}$$

für $\mathbf{s}, \mathbf{t} \in \overline{\mathbb{R}}_{\geq 0}^d$ und für die Kovarianzfunktion des Prozesses \mathbb{L} gilt nach Satz 2.2.4

$$\begin{aligned} \rho(\mathbf{s}, \mathbf{t}) &= (1 - \tau)\mathcal{L}_F(\mathbf{s} + \mathbf{t}) + \tau\mathcal{L}_G(\mathbf{s} + \mathbf{t}) \\ &\quad - \left((1 - \tau)\mathcal{L}_F(\mathbf{s})\mathcal{L}_F(\mathbf{t}) + \tau\mathcal{L}_G(\mathbf{s})\mathcal{L}_G(\mathbf{t}) \right), \quad \mathbf{s}, \mathbf{t} \in \overline{\mathbb{R}}_{\geq 0}^d. \end{aligned}$$

Aus diesen Darstellungen der Kovarianzfunktionen erhalten wir $\rho_H(\mathbf{s}, \mathbf{t})|_{\overline{A}} = \rho(\mathbf{s}, \mathbf{t})|_{\overline{A}}$, also mit Lemma B.2.2 ($\mathbb{L}(\mathbf{t}), \mathbf{t} \in \overline{A} \stackrel{v}{=} (\mathbb{L}_H(\mathbf{t}), \mathbf{t} \in \overline{A})$). Als Konsequenz ergibt sich, dass die in Satz 2.3.1 erhaltenen Grenzverteilungen der Testgrößen $\hat{T}_{m,n}^{KS}$ und $\hat{T}_{m,n}^{CvM}$ stochastisch größer sind als die Grenzverteilungen der Testgrößen $T_{m,n}^{KS}$ und $T_{m,n}^{CvM}$ bei Gültigkeit der Hypothese, denn es gilt

$$\sup_{\mathbf{t} \in \overline{A}} \mathbb{L}(\mathbf{t}) \stackrel{v}{=} \sup_{\mathbf{t} \in \overline{A}} \mathbb{L}_H(\mathbf{t}) \leq_{st} \sup_{\mathbf{t} \in \overline{\mathbb{R}}_{\geq 0}^d} \mathbb{L}_H(\mathbf{t}) \stackrel{v}{=} \sup_{\mathbf{t} \in \overline{\mathbb{R}}_{\geq 0}^d} \mathbb{L}_H(\mathbf{t})$$

und somit $T^{KS} \leq_{st} \hat{T}^{KS}$. Außerdem ist

$$T^{CvM} = \int_A \mathbb{L}^+(\mathbf{t}) \omega(\mathbf{t}) d\mathbf{t} \leq_{st} \int_{\overline{\mathbb{R}}_{\geq 0}^d} \mathbb{L}_H^+(\mathbf{t}) \omega(\mathbf{t}) d\mathbf{t} = \hat{T}^{CvM}.$$

Gilt $\mathbf{X} \stackrel{v}{=} \mathbf{Y}$, so haben \hat{T}^{KS} und T^{KS} bzw. \hat{T}^{CvM} und T^{CvM} insbesondere die gleiche Verteilung. Diese Überlegungen zeigen, dass die über das vorgeschlagene Bootstrapverfahren erhaltenen Quantile der Größen $\hat{T}_{m,n}^{KS}$ und $\hat{T}_{m,n}^{CvM}$ asymptotisch kritische Werte liefern, die die Einhaltung eines vorgegebenen Testniveaus bei der Durchführung der Tests gewährleisten.

Es zeigt sich ferner erneut eine Besonderheit des Cramér-von Mises-Tests. Aufgrund von Satz 2.2.7 erhalten wir bei Gültigkeit der Hypothese nur die beiden Fälle $T^{CvM} \equiv 0$ f.s. und $T^{CvM} \stackrel{v}{=} \hat{T}^{CvM}$ als mögliche Grenzverteilungen der Testgröße, so dass für den Cramér-von Mises-Test ebenso das eingangs erwähnte Standard-Bootstrapverfahren zulässige kritische Werte liefert. Wir formulieren diesen Sachverhalt in folgendem Korollar.

2.3.3. Korollar. Bei gegebenen $\mathbf{X}_1, \dots, \mathbf{X}_m, \mathbf{Y}_1, \dots, \mathbf{Y}_n$ sei

$$H_{m,n}^{Alt}(\mathbf{t}) := \frac{m}{m+n} F_m(\mathbf{t}) + \frac{n}{m+n} G_n(\mathbf{t}), \quad \mathbf{t} \in \mathbb{R}^d,$$

und für jedes $m, n \in \mathbb{N}$ seien $\hat{\mathbf{Z}}_{N1}^{Alt}, \dots, \hat{\mathbf{Z}}_{NN}^{Alt}$, $N = m+n$, unabhängige Zufallsvariablen mit der Verteilungsfunktion $H_{m,n}^{Alt}$. Es sei $\hat{T}_{m,n}^{CvM;Alt}$ die auf den Zufallsvariablen $\hat{\mathbf{Z}}_{N1}^{Alt}, \dots, \hat{\mathbf{Z}}_{NN}^{Alt}$ basierende Cramér-von Mises-Testgröße. Es sei $(\mathbb{L}_{H^{Alt}}(\mathbf{t}), \mathbf{t} \in \overline{\mathbb{R}}_{\geq 0}^d)$ ein zentrierter Gaußprozess mit Pfaden in $C(\overline{\mathbb{R}}_{\geq 0}^d)$ und der Kovarianzfunktion

$$\rho_{H^{Alt}}(\mathbf{s}, \mathbf{t}) = \mathcal{L}_{H^{Alt}}(\mathbf{s} + \mathbf{t}) - \mathcal{L}_{H^{Alt}}(\mathbf{s})\mathcal{L}_{H^{Alt}}(\mathbf{t}), \quad \mathbf{s}, \mathbf{t} \in \overline{\mathbb{R}}_{\geq 0}^d,$$

und $\mathcal{L}_{H^{Alt}}$ die Laplacetransformierte der Verteilungsfunktion

$$H^{Alt}(\mathbf{t}) := \tau F(\mathbf{t}) + (1 - \tau)G(\mathbf{t}), \quad \mathbf{t} \in \mathbb{R}_{\geq 0}^d.$$

Dann gilt

$$\hat{T}_{m,n}^{CvM;Alt} \xrightarrow{v} \int_{\mathbb{R}_{\geq 0}^d} \mathbb{L}_{H^{Alt}}^+(\mathbf{t}) \omega(\mathbf{t}) dt, \quad m, n \rightarrow \infty,$$

für fast alle $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{Y}_1, \mathbf{Y}_2, \dots$ □

Der Beweis der Grenzwertaussage verläuft vollkommen analog zu Satz 2.3.1 und wird an dieser Stelle nicht geführt. Die Prozesse $\mathbb{L}_{H^{Alt}}$ und \mathbb{L} haben im Fall der Gleichheit der Verteilungen von \mathbf{X} und \mathbf{Y} dieselbe Kovarianzfunktion und daher dieselbe Verteilung. Wegen $T_{m,n}^{CvM} \xrightarrow{P} 0$ für $m, n \rightarrow \infty$, falls $F \leq_{\mathcal{L}} G$ und $F \neq G$ gilt, ergibt sich so bei Gültigkeit der Hypothese

$$T^{CvM} \leq_{st} \int_{\mathbb{R}_{\geq 0}^d} \mathbb{L}_{H^{Alt}}^+(\mathbf{t}) \omega(\mathbf{t}) dt. \quad (2.3.3)$$

Aufgrund der vorangegangenen Betrachtungen für den Cramér-von Mises-Test liegt die Frage nahe, ob das im vorherigen Korollar verwendete Bootstrapverfahren, basierend auf dem Resampling anhand der Gleichverteilung auf der zusammengelegten Stichprobe, auch für die Supremumstestgröße im Allgemeinen zu brauchbaren kritischen Werten führt. Anders als im Fall von T^{CvM} ist die Supremumstestgröße $T_{m,n}^{KS}$ gemäß Beispiel 2.2.8 allerdings bei Gültigkeit der Hypothese und dem Vorliegen verschiedener Verteilungen der Zufallsvektoren \mathbf{X} und \mathbf{Y} asymptotisch nicht zwangsläufig die Einpunktverteilung im Nullpunkt. Ein stochastischer Vergleich im Stile von (2.3.3), der die Grenzverteilung von $T_{m,n}^{KS}$ unter dem Einfluss der durch $H_{m,n}^{Alt}$ beschriebenen Resamplingvorschrift mit der Grenzverteilung T^{KS} in Beziehung setzt, ist daher nicht ohne Weiteres möglich. In der Tat ist nicht zu sehen, warum die so gewonnenen kritischen Werte stets eine Einhaltung des Testniveaus gewährleisten sollen. In [Ben07] wird im Einstichprobenfall und für $d = 1$ ein Moving-Block-Bootstrapverfahren für beide Testgrößen vorgeschlagen.

Die Resultate in Satz 2.3.1 und die anschließende Bemerkung rechtfertigen zumindest, dass die Quantile, die mit Hilfe des durch (2.3.1) beschriebenen Bootstrapverfahrens gewonnen werden, brauchbare kritische Werte für beide Testgrößen liefern. Wir beschränken uns bei den weiteren Untersuchungen daher auf diese kritischen Werte und das damit verbundene Bootstrapverfahren und untersuchen im Folgenden Eigenschaften der so erhaltenen Tests. Es seien dazu

$$\hat{q}_{m,n}^{KS}(\alpha) := \inf\{q \in \mathbb{R} : P(\hat{T}_{m,n}^{KS} > q) \leq \alpha\}, \quad \alpha \in (0, 1),$$

$$\hat{q}^{KS}(\alpha) := \inf\{q \in \mathbb{R} : P(\hat{T}^{KS} > q) \leq \alpha\}, \quad \alpha \in (0, 1),$$

die $(1 - \alpha)$ -Quantile von $\hat{T}_{m,n}^{KS}$ und \hat{T}^{KS} , sowie

$$\hat{q}_{m,n}^{CvM}(\alpha) := \inf\{q \in \mathbb{R} : P(\hat{T}_{m,n}^{CvM} > q) \leq \alpha\}, \quad \alpha \in (0, 1),$$

$$\hat{q}^{CvM}(\alpha) := \inf\{q \in \mathbb{R} : P(\hat{T}^{CvM} > q) \leq \alpha\}, \quad \alpha \in (0, 1),$$

die $(1 - \alpha)$ -Quantile von $\hat{T}_{m,n}^{CvM}$ und \hat{T}^{CvM} . Es seien

$$\varphi_{m,n}^{KS}(\mathbf{X}_1, \dots, \mathbf{X}_m, \mathbf{Y}_1, \dots, \mathbf{Y}_n) := \begin{cases} 1, & T_{m,n}^{KS} > \hat{q}_{m,n}^{KS}(\alpha), \\ 0, & T_{m,n}^{KS} \leq \hat{q}_{m,n}^{KS}(\alpha), \end{cases}$$

sowie

$$\varphi_{m,n}^{CvM}(\mathbf{X}_1, \dots, \mathbf{X}_m, \mathbf{Y}_1, \dots, \mathbf{Y}_n) := \begin{cases} 1, & T_{m,n}^{CvM} > \hat{q}_{m,n}^{CvM}(\alpha), \\ 0, & T_{m,n}^{CvM} \leq \hat{q}_{m,n}^{CvM}(\alpha). \end{cases}$$

Wir verwerfen die Hypothese (2.1.1) also für ein gegebenes Testniveau $\alpha \in (0, 1)$, falls die Werte der Testgrößen die über das Bootstrapverfahren gewonnenen $(1 - \alpha)$ -Quantile der Größen $\hat{T}_{m,n}^{KS}$ bzw. $\hat{T}_{m,n}^{CvM}$ überschreiten. Der folgende Satz liefert die ersten Resultate für die vorgeschlagenen Tests.

2.3.4. Satz. *Die Tests $\varphi_{m,n}^{KS}$ und $\varphi_{m,n}^{CvM}$ sind für das Testproblem (2.1.1) konsistent gegen alle Alternativen und halten asymptotisch jedes vorgegebene Testniveau $\alpha \in (0, \frac{1}{2})$ ein.*

Beweis. Wir betrachten zunächst den Fall des Tests $\varphi_{m,n}^{KS}$. Aufgrund von Satz 2.3.1 gilt

$$P(\mathbb{L}_H \in C(\bar{\mathbb{R}}_{\geq 0}^d)) = 1. \quad (2.3.4)$$

Da die Pfade von \mathbb{L}_H fast sicher stetig sind, ist \mathbb{L}_H ein separabler Prozess (vergleiche D.2.2) und

$$\hat{T}^{KS} = \sup_{\mathbf{t} \in \mathbb{R}_{\geq 0}^d} \mathbb{L}_H(\mathbf{t}) = \sup_{\mathbf{t} \in \bar{\mathbb{R}}_{\geq 0}^d} \mathbb{L}_H(\mathbf{t})$$

kann als Supremum abzählbar vieler normalverteilter Zufallsvariablen angesehen werden, die nicht sämtlich die Varianz Null besitzen. Nach Satz D.1.1 ist die Funktion $x \mapsto P\left(\sup_{\mathbf{t} \in \bar{\mathbb{R}}_{\geq 0}^d} \mathbb{L}_H(\mathbf{t}) < x\right)$ für

$$a_0 := \inf \left\{ x \in \mathbb{R} : P\left(\sup_{\mathbf{t} \in \bar{\mathbb{R}}_{\geq 0}^d} \mathbb{L}_H(\mathbf{t}) < x\right) > 0 \right\}$$

stetig auf (a_0, ∞) und die Dichte zu $\sup_{\mathbf{t} \in \overline{\mathbb{R}}_{\geq 0}^d} \mathbb{L}_H(\mathbf{t})$ ist auf (a_0, ∞) fast überall definiert und positiv.

Wir wollen a_0 bestimmen. Mit den Sätzen D.1.2 und D.1.3 folgt zunächst, dass die Zufallsvariable $\|\mathbb{L}_H\|_\infty = \sup_{\mathbf{t} \in \overline{\mathbb{R}}_{\geq 0}^d} |\mathbb{L}_H(\mathbf{t})|$ eine streng monoton wachsende Verteilungsfunktion auf $[0, \infty)$ hat, also ist

$$P(\|\mathbb{L}_H\|_\infty < x) > 0, \quad x > 0.$$

Es ist $\hat{T}^{KS} \leq_{st} \|\mathbb{L}_H\|_\infty$, so dass $a_0 \leq 0$ folgt. Für die Testgröße $\hat{T}_{m,n}^{KS} = \sup_{\mathbf{t} \geq 0} \hat{\Delta}_{m,n}(\mathbf{t})$ gilt wegen $\hat{\Delta}_{m,n}(\mathbf{0}) = 0$ fast sicher offenbar $\hat{T}_{m,n}^{KS} \geq 0$ fast sicher und daher mit dem Portmanteau-Theorem B.1.2 $P(\hat{T}^{KS} \geq 0) \geq \limsup_{m,n \rightarrow \infty} P(\hat{T}_{m,n}^{KS} \geq 0) = 1$. Es folgt

$$P\left(\sup_{\mathbf{t} \in \overline{\mathbb{R}}_{\geq 0}^d} \mathbb{L}_H(\mathbf{t}) < x\right) = 0, \quad x < 0,$$

und somit schließlich $a_0 = 0$. Mit den Vorbetrachtungen erhalten wir daraus, dass die Verteilungsfunktion von \hat{T}^{KS} streng monoton wachsend auf $[0, \infty)$ ist.

Es gilt $\hat{T}_{m,n}^{KS} \xrightarrow{v} \hat{T}^{KS}$, $m, n \rightarrow \infty$, für fast alle $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{Y}_1, \mathbf{Y}_2, \dots$. Zusammen mit der gezeigten strengen Monotonie der Verteilungsfunktion der Grenzverteilung auf $[0, \infty)$ und

$$\forall a \in (0, 1) \forall m, n \in \mathbb{N}: \quad \hat{q}_{m,n}^{KS} \geq 0,$$

für fast alle $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{Y}_1, \mathbf{Y}_2, \dots$, folgt daher

$$\lim_{m,n \rightarrow \infty} \hat{q}_{m,n}^{KS}(\alpha) = \hat{q}^{KS}(\alpha), \quad (2.3.5)$$

fast sicher (siehe [WMF95], Satz 6.66). Nach Satz 2.2.4 gilt bei Gültigkeit der Alternative $T_{m,n}^{KS} \xrightarrow{P} \infty$ für $m, n \rightarrow \infty$ und mit (2.3.5) folgt daher

$$\lim_{m,n \rightarrow \infty} P(\varphi_{m,n}^{KS}(\mathbf{X}_1, \dots, \mathbf{X}_m, \mathbf{Y}_1, \dots, \mathbf{Y}_n) = 1 \mid K) = 1,$$

also die behauptete Konsistenz von $\varphi_{m,n}^{KS}$ gegen alle Alternativen.

Wir untersuchen die Einhaltung des Testniveaus: Der Grenzprozess \mathbb{L}_H hat die Varianzfunktion

$$\text{Var}(\mathbb{L}_H(\mathbf{t})) = \rho_H(\mathbf{t}, \mathbf{t}) = \mathcal{L}_H(2\mathbf{t}) - \mathcal{L}_H(\mathbf{t})^2, \quad \mathbf{t} \in \overline{\mathbb{R}}_{\geq 0}^d.$$

Nach Voraussetzung sind die Verteilungsfunktionen F und G nicht-degeneriert. Wir finden folglich ein $\mathbf{t}_0 \in \overline{\mathbb{R}}_{\geq 0}^d$ mit $\sigma_0^2 := \text{Var}(\mathbb{L}_H(\mathbf{t}_0)) \neq 0$. Die Zufallsvariable $\mathbb{L}_H(\mathbf{t}_0)$ ist daher von der Form $\mathbb{L}_H(\mathbf{t}_0) = U$ mit $U \sim N(0, \sigma_0^2)$, $\sigma_0^2 > 0$. Wegen $\hat{T}^{KS} \geq \mathbb{L}_H(\mathbf{t}_0)$ erhalten wir so

$$\hat{q}^{KS}(\alpha) \geq \inf\{q \in \mathbb{R} : P(U > q) \leq \alpha\}, \quad \alpha \in (0, 1),$$

wobei die rechte Seite für $\alpha \in (0, \frac{1}{2})$ größer als Null ist. Bei Gültigkeit der Hypothese konvergiert $T_{m,n}^{KS}$ in Verteilung gegen den Grenzwert $T^{KS} = \sup_{\mathbf{t} \in \bar{A}} \mathbb{L}(\mathbf{t})$. Mit dem gleichen Argument wie für \hat{T}^{KS} erhalten wir aus Satz D.1.1 für den Fall, dass die Grenzverteilung T^{KS} nicht-degeneriert im Nullpunkt ist, dass T^{KS} eine auf (a_0, ∞) , $a_0 := \inf \left\{ x : P \left(\sup_{\mathbf{t} \in \bar{A}} \mathbb{L}(\mathbf{t}) < x \right) > 0 \right\}$, stetige Verteilungsfunktion hat. Da die Verteilungen der Prozesse \mathbb{L} und \mathbb{L}_H nach Bemerkung 2.3.2 auf \bar{A} gleich sind folgt

$$\begin{aligned} T^{KS} &= \sup_{\mathbf{t} \in \bar{A}} \mathbb{L}(\mathbf{t}) \stackrel{v}{=} \sup_{\mathbf{t} \in \bar{A}} \mathbb{L}_H(\mathbf{t}) \\ &\leq \sup_{\mathbf{t} \in \bar{A}} |\mathbb{L}_H(\mathbf{t})| \leq \sup_{\mathbf{t} \in \mathbb{R}_{\geq 0}^d} |\mathbb{L}_H(\mathbf{t})| \\ &= \|\mathbb{L}_H\|_{\infty}, \end{aligned}$$

und daher wie oben $a_0 = 0$. Die so im Fall der Nichtdegeneriertheit von T^{KS} erhaltene Stetigkeit der Verteilungsfunktion von T^{KS} auf $(0, \infty)$ liefert mit (2.3.5) dann

$$\lim_{m,n \rightarrow \infty} P(T_{m,n}^{KS} > \hat{q}_{m,n}^{KS}(\alpha) | H) = P(T^{KS} > \hat{q}^{KS}(\alpha)) \leq P(\hat{T}^{KS} > \hat{q}^{KS}(\alpha)) = \alpha,$$

für alle $\alpha \in (0, \frac{1}{2})$ und somit in diesen Fällen die Einhaltung des Testniveaus. Ist $T^{KS} \equiv 0$, so gilt

$$\lim_{m,n \rightarrow \infty} P(T_{m,n}^{KS} > \hat{q}_{m,n}^{KS}(\alpha) | H) = P(T^{KS} > \hat{q}^{KS}(\alpha)) = 0,$$

für alle $\alpha \in (0, \frac{1}{2})$, so dass $\varphi_{m,n}^{KS}$ auch in diesen Fällen das Testniveau einhält.

Wir betrachten nun den Cramér-von Mises-Test $\varphi_{m,n}^{CvM}$. Die Pfade des Grenzprozesses \mathbb{L}_H sind fast sicher beschränkt und daher ist die Funktion $\mathbf{t} \mapsto \mathbb{L}_H(\mathbf{t}) \omega(\mathbf{t})$, $\mathbf{t} \in \mathbb{R}_{\geq 0}^d$, fast sicher integrierbar bezüglich des Lebesgue-Maßes auf \mathbb{R}^d . Ist Q ein Wahrscheinlichkeitsmaß mit der Lebesgue-Dichte ω , so können wir den stochastischen Prozess \mathbb{L}_H , durch Einschränkung der Pfade von \mathbb{L}_H auf den Bereich $\mathbb{R}_{\geq 0}^d$, daher als Abbildung mit Werten in $(L_Q^1(\mathbb{R}_{\geq 0}^d), \|\cdot\|_{L_Q^1})$ auffassen. Dabei bezeichnet $(L_Q^1(\mathbb{R}_{\geq 0}^d), \|\cdot\|_{L_Q^1})$ den Raum der bezüglich Q integrierbaren, reellwertigen Funktionen auf $\mathbb{R}_{\geq 0}^d$ versehen mit der L_Q^1 -Norm $\|x\|_{L_Q^1} := \int_{\mathbb{R}_{\geq 0}^d} |x(\mathbf{t})| dQ(\mathbf{t})$. Wir wollen analog zum Fall des Tests $\varphi_{m,n}^{KS}$ verfahren und stellen

$$\hat{T}^{CvM} = \int_{\mathbb{R}_{\geq 0}^d} \mathbb{L}_H^+(\mathbf{t}) \omega(\mathbf{t}) d\mathbf{t}$$

dazu zunächst als Supremum abzählbar vieler normalverteilter Zufallsvariablen dar. Es ist

$(L_Q^1(\mathbb{R}_{\geq 0}^d), \|\cdot\|_{L_Q^1})$ ein separabler Banachraum mit dem Dualraum $(L_Q^\infty(\mathbb{R}_{\geq 0}^d), \|\cdot\|_{L_Q^\infty})$,

$$L_Q^\infty(\mathbb{R}_{\geq 0}^d) := \{f : \mathbb{R}_{\geq 0}^d \rightarrow \mathbb{R} : f \text{ Borel-messbar und } \operatorname{ess\,sup}_{\mathbb{R}_{\geq 0}^d} |f| < \infty\},$$

$$\operatorname{ess\,sup}_{\mathbb{R}_{\geq 0}^d} |f| := \inf\{x \in \mathbb{R}_{\geq 0} : |f| \leq x \text{ } Q\text{-f.s.}\},$$

$$\|f\|_{L_Q^\infty} := \operatorname{ess\,sup}_{\mathbb{R}_{\geq 0}^d} |f|,$$

wobei jedes $g \in L_Q^\infty(\mathbb{R}_{\geq 0}^d)$ via $g(f) := \int_{\mathbb{R}_{\geq 0}^d} f(\mathbf{t})g(\mathbf{t}) dQ(\mathbf{t})$, $f \in L_Q^1(\mathbb{R}_{\geq 0}^d)$, ein lineares und stetiges Funktional auf $L_Q^1(\mathbb{R}_{\geq 0}^d)$ definiert. Als Folgerung des Satzes von Hahn-Banach (siehe z.B. [Rud91], Corollary 3.3) erhalten wir für die Norm eines Elements $f \in L_Q^1(\mathbb{R}_{\geq 0}^d)$ die Darstellung

$$\|f\|_{L_Q^1} = \int_{\mathbb{R}_{\geq 0}^d} |f(\mathbf{t})| dQ(\mathbf{t}) = \sup_{g \in L_Q^\infty(\mathbb{R}_{\geq 0}^d), \|g\|_{L_Q^\infty} = 1} \left\{ \int_{\mathbb{R}_{\geq 0}^d} f(\mathbf{t})g(\mathbf{t}) dQ(\mathbf{t}) \right\}. \quad (2.3.6)$$

Die einfache Identität $2 \max(0, x) = |x| + x$, $x \in \mathbb{R}$, und (2.3.6) liefern dann

$$\begin{aligned} 2\hat{T}^{CvM} &= 2 \int_{\mathbb{R}_{\geq 0}^d} \mathbb{L}_H^+(\mathbf{t}) dQ(\mathbf{t}) = \int_{\mathbb{R}_{\geq 0}^d} (|\mathbb{L}_H(\mathbf{t})| + \mathbb{L}_H(\mathbf{t})) dQ(\mathbf{t}) \\ &= \sup_{g \in L_Q^\infty(\mathbb{R}_{\geq 0}^d), \|g\|_{L_Q^\infty} = 1} \left\{ \int_{\mathbb{R}_{\geq 0}^d} \mathbb{L}_H(\mathbf{t})g(\mathbf{t}) dQ(\mathbf{t}) \right\} + \int_{\mathbb{R}_{\geq 0}^d} \mathbb{L}_H(\mathbf{t}) dQ(\mathbf{t}) \\ &= \sup_{g \in L_Q^\infty(\mathbb{R}_{\geq 0}^d), \|g\|_{L_Q^\infty} = 1} \left\{ \int_{\mathbb{R}_{\geq 0}^d} (1 + g(\mathbf{t}))\mathbb{L}_H(\mathbf{t}) dQ(\mathbf{t}) \right\}. \end{aligned}$$

Nach Satz D.3.3 kann dieses Supremum als ein Supremum über abzählbar viele Elemente aus $L_Q^\infty(\mathbb{R}_{\geq 0}^d)$ angesehen werden. Die Zufallsvariablen $\int_{\mathbb{R}_{\geq 0}^d} (1 + g(\mathbf{t}))\mathbb{L}_H(\mathbf{t}) dQ(\mathbf{t})$ sind ferner normalverteilt (siehe z.B. [Vak79], Theorem 1 für einen Beweis) und haben offenbar den Erwartungswert Null. Erneut mit Satz D.1.1 ergibt sich, dass $P\left(\int_{\mathbb{R}_{\geq 0}^d} \mathbb{L}_H^+(\mathbf{t}) \omega(\mathbf{t}) dt < x\right)$ stetig auf (a_0, ∞) ,

$$a_0 := \inf \left\{ x \in \mathbb{R} : P\left(\int_{\mathbb{R}_{\geq 0}^d} \mathbb{L}_H^+(\mathbf{t}) \omega(\mathbf{t}) dt < x\right) > 0 \right\},$$

ist, wobei mit Satz D.1.2 und $\hat{T}_{m,n}^{CvM} \geq 0$ fast sicher, wieder $a_0 = 0$ folgt. Nach Satz 2.2.4 gilt bei Gültigkeit der Alternative $T_{m,n}^{CvM} \xrightarrow{P} \infty$ für $m, n \rightarrow \infty$ und mit der Konvergenz der Quantile erhalten wir daher analog zu oben

$$\lim_{m, n \rightarrow \infty} P(\varphi_{m,n}^{CvM}(\mathbf{X}_1, \dots, \mathbf{X}_m, \mathbf{Y}_1, \dots, \mathbf{Y}_n) = 1 \mid K) = 1,$$

also die behauptete Konsistenz von $\varphi_{m,n}^{CvM}$ gegen alle Alternativen. Ebenso erhalten wir dann in Analogie zu oben, dass T^{CvM} im Falle der Nichtdegeneriertheit eine auf $(0, \infty)$ stetige Verteilungsfunktion hat. Die Stetigkeit der Verteilungsfunktion der Grenzverteilung und die Konvergenz der Quantile in (2.3.5) liefert dann erneut

$$\lim_{m,n \rightarrow \infty} P(T_{m,n}^{CvM} > \hat{q}_{m,n}^{CvM}(\alpha) | H) = P(T^{CvM} > \hat{q}^{CvM}(\alpha)) \leq P(\hat{T}^{CvM} > \hat{q}^{CvM}(\alpha)) = \alpha,$$

für alle $\alpha \in (0, \frac{1}{2})$, also die behauptete Einhaltung des Testniveaus von $\varphi_{m,n}^{CvM}$. Ist $T^{CvM} \equiv 0$ degeneriert, so gilt schließlich

$$\lim_{m,n \rightarrow \infty} P(T_{m,n}^{CvM} > \hat{q}_{m,n}^{CvM}(\alpha) | H) = P(T^{CvM} > \hat{q}^{CvM}(\alpha)) = 0,$$

für alle $\alpha \in (0, \frac{1}{2})$. □

Bei Gültigkeit der Hypothese lässt sich das Verhalten der Tests mit den im letzten Beweis verwendeten Argumenten noch genauer beschreiben. Es ist klar, dass die Mengen A und \bar{A} die Gestalt der Grenzverteilung der Testgrößen maßgeblich bestimmt. Mit den Ergebnissen in Satz 2.2.7 und den anschließenden Ausführungen liefert Teil c) des nächsten Korollars insbesondere das bereits zuvor erwähnte Resultat, dass der Cramér-von Mises-Test im Fall $F \leq_{\mathcal{L}} G$, $F \neq G$, asymptotisch den Fehler erster Art Null hat.

2.3.5. Korollar. Sei $\alpha \in (0, \frac{1}{2})$. Bei Gültigkeit der Hypothese der Laplace-Ordnung $\mathbf{X} \leq_{\mathcal{L}} \mathbf{Y}$ erhalten wir:

a) Ist $F = G$, so gilt

$$\lim_{m,n \rightarrow \infty} P(\varphi_{m,n}^{KS}(\mathbf{X}_1, \dots, \mathbf{X}_m, \mathbf{Y}_1, \dots, \mathbf{Y}_n) = 1) = \alpha$$

und

$$\lim_{m,n \rightarrow \infty} P(\varphi_{m,n}^{CvM}(\mathbf{X}_1, \dots, \mathbf{X}_m, \mathbf{Y}_1, \dots, \mathbf{Y}_n) = 1) = \alpha.$$

b) Ist $\bar{A} = \{\mathbf{0}\}$, so gilt

$$\lim_{m,n \rightarrow \infty} P(\varphi_{m,n}^{KS}(\mathbf{X}_1, \dots, \mathbf{X}_m, \mathbf{Y}_1, \dots, \mathbf{Y}_n) = 1) = 0$$

c) Ist $F \neq G$, also A nach Satz 2.2.7 eine Lebesgue-Nullmenge im $\mathbb{R}_{\geq 0}^d$, so gilt

$$\lim_{m,n \rightarrow \infty} P(\varphi_{m,n}^{CvM}(\mathbf{X}_1, \dots, \mathbf{X}_m, \mathbf{Y}_1, \dots, \mathbf{Y}_n) = 1) = 0$$

Beweis. Die Aussagen in b) und c) folgen sofort aus der Tatsache, dass in den angesprochenen Fällen aufgrund von Satz 2.2.4 die Grenzverteilungen der Testgrößen Einpunktverteilungen im Nullpunkt sind. Die Aussage in Teil a) folgt, da nach Bemerkung 2.3.2 die Grenzverteilungen T^{KS} und \hat{T}^{KS} bzw. T^{CvM} und \hat{T}^{CvM} im Fall $F = G$ gleich sind. Die im letzten Satz bewiesene Stetigkeit der Verteilungsfunktion von \hat{T}^{KS} auf $(0, \infty)$ liefert dann

$$\lim_{m,n \rightarrow \infty} P(\varphi_{m,n}^{KS}(\mathbf{X}_1, \dots, \mathbf{X}_m, \mathbf{Y}_1, \dots, \mathbf{Y}_n) = 1) = P(\hat{T}^{KS} > \hat{q}^{KS}(\alpha)) = \alpha.$$

Analog lässt sich für T^{CvM} argumentieren. \square

2.4 Invarianz des Testproblems

Invarianzeigenschaften des Testproblems sind insbesondere im Hinblick auf die Praktikabilität der hier bereitgestellten statistischen Verfahren von Interesse. Dieser Sachverhalt führt in natürlicher Weise auf die Frage nach geeigneten Transformationen $\phi : \mathbb{R}_{\geq 0}^d \rightarrow \mathbb{R}_{\geq 0}^d$ der Zufallsvektoren \mathbf{X}, \mathbf{Y} , bezüglich denen das ursprünglich betrachtete Testproblem $H : \mathbf{X} \leq_{\mathcal{L}} \mathbf{Y}$ invariant ist, für die also

$$\mathbf{X} \leq_{\mathcal{L}} \mathbf{Y} \Leftrightarrow \phi(\mathbf{X}) \leq_{\mathcal{L}} \phi(\mathbf{Y}) \quad (2.4.1)$$

für alle $\mathbb{R}_{\geq 0}^d$ -wertige Zufallsvektoren \mathbf{X}, \mathbf{Y} zutrifft. Das Resultat dieses Abschnitts liefert eine Klasse von Transformationen mit dieser Eigenschaft. Der folgende Satz gibt dabei sogar eine Charakterisierung aller Homöomorphismen $\phi : \mathbb{R}_{\geq 0}^d \rightarrow \mathbb{R}_{\geq 0}^d$ an, die die Bedingung (2.4.1) für alle $\mathbb{R}_{\geq 0}^d$ -wertigen Zufallsvektoren \mathbf{X}, \mathbf{Y} erfüllen.

Bevor wir die Aussage formulieren, erinnern wir zuvor an den Begriff der *Permutationsmatrix*. Für $n \in \mathbb{N}$ sei dazu \mathfrak{S}_n die symmetrische Gruppe vom Grad n , bestehend aus sämtlichen Permutationen der Menge $\{1, \dots, n\}$. Es seien ferner $\mathbf{e}_k, k = 1, \dots, n$, die n -dimensionalen Einheitsvektoren. Dann setzen wir für jedes $\pi \in \mathfrak{S}_n$

$$\mathbb{P}_\pi := [\mathbf{e}_{\pi(1)}, \dots, \mathbf{e}_{\pi(n)}] \in \mathbb{R}^{n \times n},$$

und nennen \mathbb{P}_π die zu π gehörige Permutationsmatrix.

2.4.1. Satz. *Es sei $\phi : \mathbb{R}_{\geq 0}^d \rightarrow \mathbb{R}_{\geq 0}^d$ ein Homöomorphismus mit der Eigenschaft, dass $\mathbf{X} \leq_{\mathcal{L}} \mathbf{Y} \Leftrightarrow \phi(\mathbf{X}) \leq_{\mathcal{L}} \phi(\mathbf{Y})$ für alle $\mathbb{R}_{\geq 0}^d$ -wertigen Zufallsvektoren \mathbf{X}, \mathbf{Y} zutrifft. Dann ist ϕ von der Form*

$$\phi(\mathbf{x}) = \mathbf{A}\mathbf{x}, \quad \mathbf{x} \in \mathbb{R}_{\geq 0}^d, \quad (2.4.2)$$

mit einer geeigneten $d \times d$ -Matrix A . Ferner existiert für jede Matrix A in (2.4.2) eine Diagonalmatrix $D \in \mathbb{R}^{d \times d}$ mit positiven Diagonaleinträgen und eine Permutation $\pi \in \mathfrak{S}_d$, so dass

$$A = DP_\pi$$

gilt. Insbesondere hat A in jeder Zeile und in jeder Spalte genau einen von Null verschiedenen (und damit positiven) Eintrag.

Beweis. Es sei $\phi : \mathbb{R}_{\geq 0}^d \rightarrow \mathbb{R}_{\geq 0}^d$ ein Homöomorphismus mit der Eigenschaft, dass

$$\mathbf{X} \leq_{\mathcal{L}} \mathbf{Y} \Leftrightarrow \phi(\mathbf{X}) \leq_{\mathcal{L}} \phi(\mathbf{Y}) \quad (2.4.3)$$

für alle $\mathbb{R}_{\geq 0}^d$ -wertigen Zufallsvektoren \mathbf{X}, \mathbf{Y} gilt.

Zunächst folgt mit der Jensenschen Ungleichung $\mathbf{X} \leq_{\mathcal{L}} E(\mathbf{X})$ für alle komponentenweise beschränkten $\mathbb{R}_{\geq 0}^d$ -wertigen Zufallsvektoren \mathbf{X} und somit $\phi(\mathbf{X}) \leq_{\mathcal{L}} \phi(E(\mathbf{X}))$. Nach Lemma 1.2.1 ist $\phi(\mathbf{X}) \leq_{\mathcal{L}} \phi(E(\mathbf{X}))$ äquivalent zu $\mathbf{a}'\phi(\mathbf{X}) \leq_{\mathcal{L}} \mathbf{a}'\phi(E(\mathbf{X}))$ für alle $\mathbf{a} \in \mathbb{R}_{\geq 0}^d$. Wir erhalten

$$\forall t \geq 0 \forall \mathbf{a} \in \mathbb{R}_{\geq 0}^d : \quad E\left(e^{-t\mathbf{a}'\phi(\mathbf{X})}\right) \geq E\left(e^{-t\mathbf{a}'\phi(E(\mathbf{X}))}\right),$$

was wiederum

$$\forall \mathbf{a} \in \mathbb{R}_{\geq 0}^d : \quad \lim_{t \rightarrow 0} \frac{1}{t} \left(E\left(e^{-t\mathbf{a}'(\phi(\mathbf{X}) - \phi(E(\mathbf{X})))}\right) - 1 \right) \geq 0,$$

impliziert. Der Grenzübergang $t \rightarrow 0$ liefert mit dem Satz von der majorisierten Konvergenz daher

$$\forall \mathbf{a} \in \mathbb{R}_{\geq 0}^d : \quad \mathbf{a}'(\phi(E(\mathbf{X})) - E(\phi(\mathbf{X}))) \geq 0,$$

und folglich $\phi(E(\mathbf{X})) \geq E(\phi(\mathbf{X}))$ für alle beschränkten $\mathbb{R}_{\geq 0}^d$ -wertigen Zufallsvektoren \mathbf{X} . Wir wenden dieses Resultat auf Zufallsvektoren \mathbf{X} mit $P(\mathbf{X} = \mathbf{x}_1) = p$, $P(\mathbf{X} = \mathbf{x}_2) = 1-p$ für $p \in [0, 1]$ und $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}_{\geq 0}^d$ an. Dann folgt

$$\phi(p\mathbf{x}_1 + (1-p)\mathbf{x}_2) \geq p\phi(\mathbf{x}_1) + (1-p)\phi(\mathbf{x}_2), \quad p \in [0, 1]. \quad (2.4.4)$$

Es sei $\phi = (\phi_1, \dots, \phi_d)$, mit den Komponentenfunktionen $\phi_k : \mathbb{R}_{\geq 0}^d \rightarrow \mathbb{R}_{\geq 0}$, $k = 1, \dots, d$. Aufgrund von (2.4.4) sind die Funktionen ϕ_k , $k = 1, \dots, d$, konkav. Wir zeigen, dass sämtliche ϕ_k , $k = 1, \dots, d$, außerdem monoton wachsend sind in dem Sinne, dass aus $\mathbf{x} \leq \mathbf{y} \in \mathbb{R}_{\geq 0}^d$ bereits $\phi_k(\mathbf{x}) \leq \phi_k(\mathbf{y})$ folgt. Angenommen, eine der Funktionen ϕ_1, \dots, ϕ_k - etwa ϕ_1 - ist nicht monoton wachsend. Dann existieren $\mathbf{x}_0, \mathbf{y}_0 \in \mathbb{R}_{\geq 0}^d$ mit $\mathbf{x}_0 \leq \mathbf{y}_0$ und $\phi_1(\mathbf{x}_0) > \phi_1(\mathbf{y}_0)$. Aufgrund der Konkavität von ϕ_1 gilt zunächst

$$\forall \mathbf{x}, \mathbf{y} \in \mathbb{R}_{\geq 0}^d \forall p \in [0, 1] : \quad \phi_1((1-p)\mathbf{x} + p\mathbf{y}) \geq (1-p)\phi_1(\mathbf{x}) + p\phi_1(\mathbf{y}).$$

Ist für ein $p > 1$ und $\mathbf{x}, \mathbf{y} \in \mathbb{R}_{\geq 0}^d$ auch $\mathbf{z} := (1-p)\mathbf{x} + p\mathbf{y} \in \mathbb{R}_{\geq 0}^d$, so ergibt sich daraus

$$\phi_1(\mathbf{y}) = \phi_1\left(\frac{1}{p}\mathbf{z} + \frac{p-1}{p}\mathbf{x}\right) \geq \frac{1}{p}\phi_1(\mathbf{z}) + \frac{p-1}{p}\phi_1(\mathbf{x}).$$

Somit erhalten wir, dass für jedes feste $p > 1$ die Ungleichung

$$\phi_1(\mathbf{z}) = \phi_1((1-p)\mathbf{x} + p\mathbf{y}) \leq (1-p)\phi_1(\mathbf{x}) + p\phi_1(\mathbf{y})$$

für alle $\mathbf{x}, \mathbf{y} \in \mathbb{R}_{\geq 0}^d$ mit Eigenschaft $(1-p)\mathbf{x} + p\mathbf{y} \in \mathbb{R}_{\geq 0}^d$ zutrifft. Für jedes $p > 1$ ist aber wegen $\mathbf{x}_0 \leq \mathbf{y}_0$ auch $(1-p)\mathbf{x}_0 + p\mathbf{y}_0 \in \mathbb{R}_{\geq 0}^d$. Wir erhalten daher für jedes $p > 1$

$$\begin{aligned} \phi_1((1-p)\mathbf{x}_0 + p\mathbf{y}_0) &\leq (1-p)\phi_1(\mathbf{x}_0) + p\phi_1(\mathbf{y}_0) \\ &= \phi_1(\mathbf{x}_0) + p(\phi_1(\mathbf{y}_0) - \phi_1(\mathbf{x}_0)) \rightarrow -\infty, \quad p \rightarrow \infty, \end{aligned}$$

im Widerspruch zur Bijektivität von ϕ auf $\mathbb{R}_{\geq 0}^d$. Somit sind alle Komponentenfunktionen $\phi_k : \mathbb{R}_{\geq 0}^d \rightarrow \mathbb{R}_{\geq 0}$, $k = 1, \dots, d$, monoton wachsend. Folglich ist auch ϕ monoton wachsend in dem Sinne, dass ϕ - bei jeweils festgehaltenen übrigen Komponenten - monoton wachsend in jeder Komponente ist.

Gilt (2.4.3) für eine Funktion ϕ mit den genannten Eigenschaften, so gilt diese Aussage offenbar auch für die Umkehrfunktion ϕ^{-1} . Für einen beschränkten Zufallsvektor \mathbf{X} ist ferner auch der Zufallsvektor $\phi^{-1}(\mathbf{X})$ beschränkt, so dass analoge Argumente wie eingangs für ϕ , angewendet auf ϕ^{-1} , schließlich $\phi^{-1}(E(\mathbf{X})) \geq E(\phi^{-1}(\mathbf{X}))$ für alle beschränkten $\mathbb{R}_{\geq 0}^d$ -wertigen Zufallsvektoren \mathbf{X} und somit insbesondere

$$\phi^{-1}(E(\phi(\mathbf{X}))) \geq E(\phi^{-1}(\phi(\mathbf{X}))) = E(\mathbf{X})$$

liefern. Wenden wir auf diese Ungleichung die monoton wachsende Funktion ϕ an, so ergibt sich daraus $\phi(E(\mathbf{X})) \leq E(\phi(\mathbf{X}))$ für alle beschränkten $\mathbb{R}_{\geq 0}^d$ -wertigen Zufallsvektoren \mathbf{X} , was für diese Zufallsvektoren schließlich die Identität $\phi(E(\mathbf{X})) = E(\phi(\mathbf{X}))$ impliziert.

Wir untersuchen nun die Funktionalgleichung

$$\phi(E(\mathbf{X})) = E(\phi(\mathbf{X})), \tag{2.4.5}$$

für alle beschränkten $\mathbb{R}_{\geq 0}^d$ -wertigen Zufallsvektoren \mathbf{X} und für Homöomorphismen $\phi : \mathbb{R}_{\geq 0}^d \rightarrow \mathbb{R}_{\geq 0}^d$: Zunächst ist aufgrund der Bijektivität und da ϕ monoton wachsend ist offenbar $\phi(\mathbf{0}) = \mathbf{0}$. Bezeichnet man ferner erneut mit \mathbf{e}_k , $k = 1, \dots, d$, den k -ten Einheitsvektor, so liefert (2.4.5) angewendet auf diskrete Zufallsvektoren \mathbf{X} mit $P(\mathbf{X} = \mathbf{e}_k) = p_k$, $p_k \geq 0$, $k = 1, \dots, d$, $\sum_{k=1}^d p_k = 1$, dass

$$\phi\left(\sum_{k=1}^d p_k \mathbf{e}_k\right) = \sum_{k=1}^d p_k \phi(\mathbf{e}_k) = A\mathbf{p},$$

mit $\mathbf{p} = (p_1, \dots, p_d)'$ und der Matrix $A = (\phi(\mathbf{e}_1), \dots, \phi(\mathbf{e}_d)) \in \mathbb{R}^{d \times d}$ gilt. Ferner folgt aus (2.4.5), dass für $\alpha \in [0, 1]$ und $\mathbf{x} \in \mathbb{R}_{\geq 0}^d$

$$\phi(\alpha \mathbf{x}) = \phi(\alpha \mathbf{x} + (1 - \alpha)\mathbf{0}) = \alpha\phi(\mathbf{x}) + (1 - \alpha)\phi(\mathbf{0}) = \alpha\phi(\mathbf{x})$$

zutritt. Ist $\alpha > 1$, so ergibt sich für $\mathbf{x} \in \mathbb{R}_{\geq 0}^d$ offenbar $\phi(\mathbf{x}) = \phi\left(\frac{1}{\alpha}\alpha\mathbf{x}\right) = \frac{1}{\alpha}\phi(\alpha\mathbf{x})$, also schließlich

$$\forall \alpha \geq 0 \forall \mathbf{x} \in \mathbb{R}_{\geq 0}^d : \quad \phi(\alpha \mathbf{x}) = \alpha\phi(\mathbf{x}).$$

Sei nun $\mathbf{x} = (x_1, \dots, x_d)' \in \mathbb{R}_{\geq 0}^d$, $\mathbf{x} \neq \mathbf{0}$, beliebig. Dann hat der Vektor

$$\mathbf{p} = (p_1, \dots, p_d)' := \frac{1}{\sum_{k=1}^d x_k} \mathbf{x}$$

die Eigenschaft $p_k \geq 0$, $k = 1, \dots, d$, $\sum_{k=1}^d p_k = 1$. Es folgt

$$\phi(\mathbf{x}) = \phi\left(\left(\sum_{k=1}^d x_k\right) \mathbf{p}\right) = \left(\sum_{k=1}^d x_k\right) \phi(\mathbf{p}) = \left(\sum_{k=1}^d x_k\right) A\mathbf{p} = A\mathbf{x},$$

für alle $\mathbf{x} \in \mathbb{R}_{\geq 0}^d$. Somit hat ϕ die Gestalt $\phi(\mathbf{x}) = A\mathbf{x}$ und A hat aufgrund der Bijektivität von ϕ nicht-negative Einträge. Um die im Satz behaupteten Eigenschaften von A zu zeigen, betrachten wir zur Matrix $A = (a_{ij})_{1 \leq i, j \leq d}$ mit nicht-negativen Einträgen die inverse Matrix $A^{-1} = (b_{lk})_{1 \leq l, k \leq d}$. Offenbar gilt aufgrund der Bijektivität dann auch $b_{lk} \geq 0$ für alle $1 \leq l, k \leq d$. Wir nehmen an, dass A in einer Zeile oder einer Spalte mehr als einen positiven Eintrag hat. Ohne Einschränkung seien $a_{11}, a_{12} > 0$. Aus

$$\forall k \in \{2, \dots, d\} : \quad \sum_{j=1}^d a_{1j} b_{jk} = 0,$$

folgt dann aber $b_{1k} = 0$ und $b_{2k} = 0$ jeweils für alle $k = 2, \dots, d$, so dass die ersten beiden Zeilenvektoren von A^{-1} linear abhängig sind - im Widerspruch zur Bijektivität von ϕ . Also hat A in jeder Zeile und in jeder Spalte genau einen von Null verschiedenen positiven Eintrag. Jede derartige Matrix ergibt sich durch Vertauschung der Spalten einer Diagonalmatrix mit positiven Einträgen, was schlussendlich die Behauptung zeigt. \square

Die Matrizen A mit der Eigenschaft, dass in jeder Zeile und in jeder Spalte genau ein von Null verschiedener, positiver Eintrag steht, lassen sich geometrisch interpretieren. Im Spezialfall $d = 2$ etwa kommen nur die Matrizen

$$A = \begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix} \quad \text{und} \quad A = \begin{pmatrix} 0 & \alpha \\ \beta & 0 \end{pmatrix},$$

mit $\alpha, \beta > 0$ in Frage, die zu Streckungen bzw. Spiegelungen in \mathbb{R}^2 korrespondieren.

Als Folgerung ergibt sich außerdem, dass mit den genannten Abbildungen in der Tat eine Gruppe von Transformationen vorliegt, bezüglich der das Testproblem auf Laplace-Ordnung invariant ist. Wir bezeichnen dazu für $a_1, \dots, a_d \in \mathbb{R}$ mit $\text{diag}(a_1, \dots, a_d)$ die $d \times d$ -Diagonalmatrix mit den Diagonaleinträgen a_1, \dots, a_d .

2.4.2. Korollar. Die Klasse

$$\mathcal{G} := \left\{ \phi : \mathbb{R}_{\geq 0}^d \rightarrow \mathbb{R}_{\geq 0}^d \mid \phi(\mathbf{x}) = \mathbb{D}\mathbb{P}_\pi \mathbf{x}, \pi \in \mathfrak{S}_d, \mathbb{D} = \text{diag}(a_1, \dots, a_d), a_k > 0, k = 1, \dots, d \right\}$$

der im letzten Satz betrachteten Transformationen ist bezüglich der Komposition von Abbildungen eine nicht abelsche Gruppe.

Beweis. Die Assoziativität und die Existenz eines neutralen Elements sind offensichtlich. Um die Existenz eines inversen Elements in \mathcal{G} zu zeigen, sei $\phi \in \mathcal{G}$ beliebig. Wir definieren dann zu ϕ mit der Darstellung $\phi(\mathbf{x}) = \mathbb{D}\mathbb{P}_\pi \mathbf{x}$, $\mathbf{x} \in \mathbb{R}_{\geq 0}^d$, mit $\mathbb{D} = \text{diag}(a_1, \dots, a_d)$, $a_k > 0, k = 1, \dots, d$, und $\pi \in \mathfrak{S}_d$ die Abbildung $\psi(\mathbf{x}) := \tilde{\mathbb{D}}\mathbb{P}_{\pi^{-1}} \mathbf{x}$, $\mathbf{x} \in \mathbb{R}_{\geq 0}^d$, mit $\tilde{\mathbb{D}} = \text{diag}(a_{\pi(1)}^{-1}, \dots, a_{\pi(d)}^{-1})$. Offenbar ist $\psi \in \mathcal{G}$.

Für einen beliebigen Vektor $\mathbf{x} = (x_1, \dots, x_d)' \in \mathbb{R}_{\geq 0}^d$ und eine beliebige Permutation $\pi \in \mathfrak{S}_d$ gilt

$$\mathbb{P}_\pi \mathbf{x} = [\mathbf{e}_{\pi(1)}, \dots, \mathbf{e}_{\pi(d)}] \begin{pmatrix} x_1 \\ \vdots \\ x_d \end{pmatrix} = \begin{pmatrix} x_{\pi^{-1}(1)} \\ \vdots \\ x_{\pi^{-1}(d)} \end{pmatrix}.$$

Daher erhalten wir für beliebiges $\mathbf{x} = (x_1, \dots, x_d)' \in \mathbb{R}_{\geq 0}^d$

$$\begin{aligned} \psi(\phi(\mathbf{x})) &= (\tilde{\mathbb{D}}\mathbb{P}_{\pi^{-1}}) (\mathbb{D}\mathbb{P}_\pi \mathbf{x}) = \tilde{\mathbb{D}}\mathbb{P}_{\pi^{-1}}\mathbb{D} \begin{pmatrix} x_{\pi^{-1}(1)} \\ \vdots \\ x_{\pi^{-1}(d)} \end{pmatrix} \\ &= \tilde{\mathbb{D}}\mathbb{P}_{\pi^{-1}} \begin{pmatrix} a_1 x_{\pi^{-1}(1)} \\ \vdots \\ a_d x_{\pi^{-1}(d)} \end{pmatrix} = \tilde{\mathbb{D}} \begin{pmatrix} a_{\pi(1)} x_1 \\ \vdots \\ a_{\pi(d)} x_d \end{pmatrix} \\ &= \mathbf{x}. \end{aligned}$$

Es folgt $\psi(\phi(\mathbf{x})) = \mathbf{x}$ für alle $\mathbf{x} \in \mathbb{R}_{\geq 0}^d$ und somit $\psi = \phi^{-1}$. Mit ähnlichen Überlegungen überzeugt man sich schließlich davon, dass für beliebige $\phi, \psi \in \mathcal{G}$ auch $\phi \circ \psi \in \mathcal{G}$ gilt.

Also ist \mathcal{G} wie behauptet eine Gruppe. Im Allgemeinen ist allerdings $\mathbb{D}\mathbb{P}_\pi \neq \mathbb{P}_\pi\mathbb{D}$, also ist \mathcal{G} nicht abelsch. (Für $d > 2$ ergibt sich dieser Sachverhalt alternativ aus der

Tatsache, dass die symmetrische Gruppe für $d > 2$ bekanntlich nicht abelsch ist und somit im Allgemeinen $\mathbb{P}_\pi \mathbb{P}_\sigma = \mathbb{P}_{\pi \circ \sigma} \neq \mathbb{P}_{\sigma \circ \pi} = \mathbb{P}_\sigma \mathbb{P}_\pi$, $\pi, \sigma \in \mathfrak{S}_d$ gilt.) \square

2.4.3. Bemerkung. Das Testproblem auf Laplace-Ordnung der Zufallsvariablen \mathbf{X}, \mathbf{Y} ist zwar gemäß Satz 2.4.1 invariant unter den dort angegebenen Transformationen, dieser Sachverhalt trifft aber bei fest gewählter Gewichtsfunktion ω offenbar nicht auf den Cramér-von Mises-Test zu. Wir wollen einige mögliche Modifikationen der bisher untersuchten Cramér-von Mises-Testgröße erwähnen, die die Invarianz des Tests unter den genannten Transformationen gewährleisten können.

- a) Unter zusätzlichen Integrabilitätsvoraussetzungen an die Zufallsvariablen \mathbf{X}, \mathbf{Y} kann als Integrator beispielsweise das Lebesgue-Maß, also die Gewichtsfunktion $\omega \equiv 1$ gewählt werden (vergleiche Abschnitt 2.1). Nach dem Transformationssatz erhalten wir dann die Invarianz des Cramér-von Mises-Tests bezüglich der in Satz 2.4.1 genannten Transformationen ϕ in dem Sinne, dass

$$\varphi_{m,n}^{CvM}(\mathbf{X}_1, \dots, \mathbf{X}_m, \mathbf{Y}_1, \dots, \mathbf{Y}_n) = \varphi_{m,n}^{CvM}(\phi(\mathbf{X}_1), \dots, \phi(\mathbf{X}_m), \phi(\mathbf{Y}_1), \dots, \phi(\mathbf{Y}_n))$$

fast sicher gilt. Die Argumente und Beweisführungen die Asymptotik der Testgröße betreffend, müssen in diesem Fall an einigen Stellen modifiziert werden (vergleiche z.B. [BG09] für eine ähnliche Situation im univariaten Fall).

- b) Sei $d = 1$ und für $k = 1, \dots, m + n$ sei

$$Z_k := \begin{cases} Y_k, & k = 1, \dots, n, \\ X_{k-n}, & k = n + 1, \dots, n + m. \end{cases}$$

Wir setzen ferner $\bar{Z} = \frac{1}{m+n} \sum_{k=1}^{m+n} Z_k$ bzw. $\overline{\phi(Z)} = \frac{1}{m+n} \sum_{k=1}^{m+n} \phi(Z_k)$, sowie bei gegebenen $X_1, \dots, X_m, Y_1, \dots, Y_n$

$$\omega_{m,n}(t) := \bar{Z} \exp(-t\bar{Z}), \quad t \geq 0.$$

Im eindimensionalen Fall sind die Transformationen aus Satz 2.4.1 von der Form $\phi(x) = cx$, mit einem geeigneten $c > 0$. Wählen wir als Gewichtsfunktion der Cramér-von Mises-Testgröße die anhand der empirischen Daten gebildete Funktion $\omega_{m,n}$, so erhalten

wir durch eine einfache Substitution offenbar

$$\begin{aligned} T_{m,n}^* &:= \sqrt{\frac{mn}{m+n}} \int_0^\infty \left(\frac{1}{n} \sum_{k=1}^n e^{-tY_k} - \frac{1}{m} \sum_{k=1}^m e^{-tX_k} \right)^+ \bar{Z} e^{-t\bar{Z}} dt \\ &= \sqrt{\frac{mn}{m+n}} \int_0^\infty \left(\frac{1}{n} \sum_{k=1}^n e^{-t\phi(Y_k)} - \frac{1}{m} \sum_{k=1}^m e^{-t\phi(X_k)} \right)^+ \frac{\phi(\bar{Z})}{\phi(Z)} e^{-t\phi(\bar{Z})} dt, \quad m, n \in \mathbb{N}, \end{aligned}$$

für jedes beliebige $c > 0$, also die Invarianz der Testgröße

$$T_{m,n}^* := \int_0^\infty \Delta_{m,n}^+(t) \omega_{m,n}(t) dt$$

unter den genannten Transformationen. Bei asymptotischen Betrachtungen ist hierbei die Zufälligkeit der Gewichtsfunktion $\omega_{m,n}$ zu beachten.

Es gilt mit $\sigma := (\tau E(X) + (1 - \tau)E(Y))$

$$\forall t \geq 0: \quad \lim_{m,n \rightarrow \infty} \omega_{m,n}(t) = \omega(t) := \sigma \exp(-\sigma t)$$

fast sicher.

Abschließend sei bemerkt, dass sich in der $d \geq 2$ -dimensionalen Situation die Angabe eines geeigneten empirischen Maßes, welches im oben genannten Sinne eine Kompensation der Transformationen der Form $\phi(\mathbf{x}) = A\mathbf{x}$ bewirkt und so zu einer unter diesen Transformationen invarianten Testgröße führt, ohne Weiteres aus der eindimensionalen Situation verallgemeinern lässt.

- c) Statt wie in Teil a) bezüglich eines festen Maßes zu integrieren, oder wie in Teil b) eine empirische Modifikation der Exponentialverteilung als Integrator zu verwenden, bietet sich auch die - in Zweistichproben-Problemen häufig übliche - Integration bezüglich einer diskreten Verteilung auf der zusammengesetzten Stichprobe an.

Für gegebene $\mathbf{X}_1, \dots, \mathbf{X}_m, \mathbf{Y}_1, \dots, \mathbf{Y}_n$ sei dazu

$$H_{m,n} = \frac{n}{m+n} \frac{1}{m} \sum_{k=1}^m \delta_{\mathbf{X}_k} + \frac{m}{m+n} \frac{1}{n} \sum_{k=1}^n \delta_{\mathbf{Y}_k}, \quad m, n \in \mathbb{N},$$

und

$$\begin{aligned} T_{m,n}^{**} &= \int_{\mathbb{R}_{\geq 0}^d} \Delta_{m,n}^+(\mathbf{t}) dH_{m,n}(\mathbf{t}) \\ &= \frac{n}{m(m+n)} \sum_{k=1}^m \Delta_{m,n}^+(\mathbf{X}_k) + \frac{m}{n(m+n)} \sum_{k=1}^n \Delta_{m,n}^+(\mathbf{Y}_k). \end{aligned} \quad (2.4.6)$$

Offensichtlich liefert (2.4.6) in diesem Fall eine einfach zu berechnende Darstellung der Testgröße. Im Hinblick auf die Konsistenz dieses Verfahrens ist allerdings zu bemerken, dass das betrachtete Abstandsmaß

$$\int_{\mathbb{R}_{\geq 0}^d} (\mathcal{L}_{\mathbf{Y},n}(\mathbf{t}) - \mathcal{L}_{\mathbf{X},m}(\mathbf{t}))^+ dH_{m,n}(\mathbf{t})$$

auch bei Gültigkeit der Alternative asymptotisch den Wert Null annehmen kann. Beispielsweise gilt im Fall $d = 1$ und für Zufallsvariablen X und Y mit den Verteilungsfunktionen F bzw. G und den Zweipunktverteilungen $P^Y = \frac{1}{2}\delta_0 + \frac{1}{2}\delta_1$ bzw. $P^X = \frac{8}{9}\delta_{\frac{1}{3}} + \frac{1}{9}\delta_{\frac{2}{3}}$, dass mit

$$t_0 := -3 \ln \left(\frac{1}{18} (\sqrt{373} - 7) \right) \approx 1.13,$$

zunächst $\mathcal{L}_Y(t) - \mathcal{L}_X(t) \leq 0$ für $t \in (0, t_0]$ und ferner $\mathcal{L}_Y(t) - \mathcal{L}_X(t) > 0$ für $t > t_0$ ist. Somit ergibt sich $X \not\leq_{\mathcal{L}} Y$. Wegen $t_0 > 1$ ergibt sich allerdings $(\mathcal{L}_Y(t) - \mathcal{L}_X(t))^+ = 0$ für alle t im Träger des zu $H = (1 - \tau)F + \tau G$ gehörenden Maßes. Folglich gilt in diesem Fall

$$\int_0^\infty (\mathcal{L}_Y(t) - \mathcal{L}_X(t))^+ dH(t) = 0.$$

2.5 Simulationen

Bei der Durchführung der Tests mit konkreten Daten ist aufgrund des vorgeschlagenen Bootstrapverfahrens und der theoretischen Resultate in Satz 2.3.4 asymptotisch im Allgemeinen mit einer Unterschreitung des vorgegebenen Testniveaus, also einem konservativen Verhalten der Tests zu rechnen. Wir wollen in diesem Abschnitt einige Simulationsergebnisse aufführen.

Zur Berechnung der Testgrößen $T_{m,n}^{KS}$ und $T_{m,n}^{CvM}$ anhand von Beobachtungen der Zufallsvektoren $\mathbf{X}_1, \dots, \mathbf{X}_m, \mathbf{Y}_1, \dots, \mathbf{Y}_n$ muss zum Einen das Supremum und zum Anderen das Integral der Differenz der empirischen Laplacetransformierten

$$\Delta_{m,n}(\mathbf{t}) := \sqrt{\frac{mn}{m+n}} (\mathcal{L}_{\mathbf{Y},n}(\mathbf{t}) - \mathcal{L}_{\mathbf{X},m}(\mathbf{t})), \quad m, n \in \mathbb{N},$$

auf $\mathbb{R}_{\geq 0}^d$ bestimmt werden. Die Funktion $\mathbf{t} \mapsto \Delta_{m,n}(\mathbf{t})$ ist stetig und differenzierbar, so dass bekannte und implementierte Verfahren zur numerischen Berechnung mehrdimensionaler Integrale und Algorithmen zum Auffinden von Maximalstellen stetiger Funktionen verwendet werden können um die Testgrößen zu berechnen.

Die Simulationsergebnisse dieses Abschnitts wurden im Fall des Cramér-von Mises-Tests für die Cramér-von Mises-Testgröße mit Gewichtsfunktion $\omega(\mathbf{t}) = \exp(-\sum_{k=1}^d t_k)$ beobachtet.

Die Bestimmung der empirischen Güten beider Tests wurde anhand der folgenden Vorgehensweise durchgeführt.

i) Es werden m Zufallszahlen $\mathbf{x}_1, \dots, \mathbf{x}_m$ zu einer vorgegeben Verteilungsfunktion F von \mathbf{X} und n Zufallszahlen $\mathbf{y}_1, \dots, \mathbf{y}_n$ zu einer vorgegebenen Verteilungsfunktion G von \mathbf{Y} erzeugt und daraus die Werte der Testgrößen $T_{m,n}^{KS}$ (bzw. $T_{m,n}^{CvM}$) berechnet.

ii) Anschließend werden $m + n$ Zufallszahlen $\mathbf{z}_1, \dots, \mathbf{z}_{m+n}$ zur Verteilung

$$\frac{n}{m+n} \frac{1}{m} \sum_{k=1}^m \delta_{\mathbf{x}_k} + \frac{m}{m+n} \frac{1}{n} \sum_{k=1}^n \delta_{\mathbf{y}_k}$$

erzeugt und der Wert der Testgröße $\hat{T}_{m,n}^{KS}$ (bzw. $\hat{T}_{m,n}^{CvM}$) basierend auf $\mathbf{z}_1, \dots, \mathbf{z}_{m+n}$ berechnet.

iii) Schritt ii) wird $R \in \mathbb{N}$ Mal wiederholt, so dass sich R Werte $(\hat{T}_{m,n}^{KS})^{(1)}, \dots, (\hat{T}_{m,n}^{KS})^{(R)}$ (bzw. $(\hat{T}_{m,n}^{CvM})^{(1)}, \dots, (\hat{T}_{m,n}^{CvM})^{(R)}$) ergeben.

iv) Für ein vorgegebenes Testniveau $\alpha \in (0, \frac{1}{2})$ wird das empirische $(1 - \alpha)$ -Quantil $\hat{q}_{m,n}^{KS}(\alpha)$ von $(\hat{T}_{m,n}^{KS})^{(1)}, \dots, (\hat{T}_{m,n}^{KS})^{(R)}$ (bzw. das empirische $(1 - \alpha)$ -Quantil $\hat{q}_{m,n}^{CvM}(\alpha)$ von $(\hat{T}_{m,n}^{CvM})^{(1)}, \dots, (\hat{T}_{m,n}^{CvM})^{(R)}$) bestimmt.

v) Die Hypothese wird verworfen, falls $T_{m,n}^{KS} > \hat{q}_{m,n}^{KS}(\alpha)$ (bzw. $T_{m,n}^{CvM} > \hat{q}_{m,n}^{CvM}(\alpha)$) ist.

vi) Die Schritte i)-v) werden anschließend $S \in \mathbb{N}$ Mal wiederholt und es wird die relative Häufigkeit der Ablehnungen der Hypothese notiert.

Die Tabellen 2.2 und 2.3 enthalten Simulationsergebnisse der empirischen Güten der Tests $\varphi_{m,n}^{KS}$ und $\varphi_{m,n}^{CvM}$ bei Gültigkeit der Hypothese (Tabelle 2.2) und bei Gültigkeit der Alternative (Tabelle 2.3) für jeweils ausgewählte eindimensionale Verteilungen. Es wurde bei den Simulationen stets $S = 1000$ und $\alpha = 0.05$ gewählt. Erste Simulationen zeigen dabei, dass sich im Hinblick auf eine Einhaltung des Testniveaus ein relativ geringer Umfang R von Bootstrap-Wiederholungen als ausreichend erweist und ein höherer Umfang nicht zu signifikanten Verbesserungen führt. Die Ergebnisse in den Tabellen 2.2 und 2.3 ergaben sich stets für $R = 50$.

Beide Tests zeigen bei Gültigkeit der Hypothese vergleichbare Resultate. Bei Gleichheit der Verteilungen F und G erhalten wir für die empirischen Güten Werte in der Größenordnung des vorgeschriebenen Niveaus. Gilt $F \leq_{\mathcal{L}} G$ und $F \neq G$, so zeigt sich eine deutlichere Unterschreitung des Testniveaus, was in vielen Fällen die theoretischen Ergebnisse aus Korollar 2.3.5 widerspiegelt. Auch bei Gültigkeit der Alternative verhalten sich beide Tests in den simulierten Fällen ähnlich. Die benötigte Zeit für die Berechnung der Testgröße ist im Fall des Kolmogorov-Smirnov-Tests allerdings deutlich geringer als die numerisch aufwändigere Berechnung der Integraltestgröße $T_{m,n}^{CvM}$.

Um ein unterschiedliches Verhalten der beiden Tests zu beobachten, kommen wir nochmals auf die in Beispiel 2.2.8 untersuchten Zufallsvariablen X, Y mit

$$P^X = \frac{2}{64}\delta_1 + \frac{1}{64}\delta_3 + \frac{24}{64}\delta_5 + \frac{2}{64}\delta_7 + \frac{35}{64}\delta_9,$$

$$P^Y = \frac{1}{64}\delta_1 + \frac{9}{64}\delta_3 + \frac{2}{64}\delta_5 + \frac{26}{64}\delta_7 + \frac{26}{64}\delta_9$$

zurück. Wir hatten dort gesehen, dass $X \leq_{\mathcal{L}} Y$ und $A = \{t \geq 0 : \mathcal{L}_X(t) = \mathcal{L}_Y(t)\} = \{0, \frac{1}{2} \ln(3)\}$ gilt, und die Grenzverteilungen der beiden Testgrößen $T_{m,n}^{KS}$ und $T_{m,n}^{CvM}$ somit nach Satz 2.2.4 verschieden sind. Tabelle 2.1 zeigt die empirischen Güten der beiden Tests für diese Verteilungen für $R = 50$ und $S = 1000$.

F	G	$\varphi_{m,n}^{KS}$			$\varphi_{m,n}^{CvM}$		
		a)	b)	c)	a)	b)	c)
P^X	P^Y	0.040	0.040	0.036	0.023	0.027	0.022

Tabelle 2.1: Empirische Güten anhand der Verteilungen von X und Y aus Beispiel 2.2.8 für $\alpha = 0.05$ und a) $m = n = 1000$, b) $m = 800, n = 1200$, c) $m = 1200, n = 800$.

F	G	$\varphi_{m,n}^{KS}$			$\varphi_{m,n}^{CvM}$		
		a)	b)	c)	a)	b)	c)
$\mathfrak{R}([0, 1])$	$\mathfrak{R}([0, 1])$	0.057	0.052	0.064	0.076	0.062	0.062
$\mathfrak{R}([1, 2])$	$\mathfrak{R}([1, 2])$	0.062	0.059	0.073	0.071	0.063	0.062
$\mathfrak{R}([10, 20])$	$\mathfrak{R}([10, 20])$	0.054	0.064	0.056	0.045	0.060	0.059
$\mathfrak{R}([\pi, 2\pi])$	$\mathfrak{R}([\pi, 2\pi])$	0.058	0.067	0.061	0.057	0.055	0.058
$\mathfrak{R}([0, 1])$	$\mathfrak{R}([\frac{1}{2}, 1])$	0.000	0.000	0.000	0.000	0.000	0.000
$\mathfrak{R}([0, 1])$	$\mathfrak{R}([3, 4])$	0.000	0.000	0.000	0.000	0.000	0.000
$\mathfrak{R}([3, 4])$	$\mathfrak{R}([5, 6])$	0.000	0.000	0.000	0.000	0.000	0.000
$\mathfrak{R}([\pi, 2\pi])$	$\mathfrak{R}([4, 10])$	0.000	0.000	0.000	0.000	0.000	0.000
Exp(1)	Exp(1)	0.049	0.039	0.080	0.060	0.063	0.054
Exp(2)	Exp(2)	0.076	0.056	0.055	0.069	0.085	0.063
Exp(10)	Exp(10)	0.063	0.051	0.071	0.049	0.076	0.043
Exp(1.1)	Exp(1)	0.034	0.032	0.030	0.032	0.027	0.038
Exp(1.5)	Exp(1)	0.008	0.008	0.011	0.007	0.003	0.001
Exp(2)	Exp(1)	0.001	0.000	0.000	0.000	0.001	0.000
Exp(10)	Exp(1)	0.000	0.000	0.000	0.000	0.000	0.000
$\Gamma(2, 2)$	$\Gamma(2, 2)$	0.068	0.051	0.075	0.065	0.068	0.066
$\Gamma(2, 1)$	$\Gamma(2, 1)$	0.063	0.057	0.058	0.073	0.066	0.074
χ_1^2	Exp(1)	0.010	0.005	0.007	0.017	0.010	0.010
χ_2^2	$\Gamma(3, 2)$	0.000	0.000	0.000	0.000	0.000	0.000
$\Gamma(2, 2)$	$\Gamma(2, 3)$	0.000	0.002	0.001	0.000	0.000	0.000
$\Gamma(2, 2)$	$\Gamma(2.1, 2.1)$	0.057	0.046	0.058	0.032	0.024	0.035
Weib(2, 2)	Weib(2, 2)	0.069	0.053	0.059	0.057	0.070	0.082
Weib(2, 3)	Weib(2, 3)	0.067	0.051	0.070	0.057	0.054	0.081
Weib(10, 10)	Weib(10, 10)	0.059	0.061	0.071	0.053	0.058	0.084
Weib(2, 2)	Weib(2.1, 2)	0.053	0.045	0.064	0.059	0.045	0.061
Weib(2, 2)	Weib(2.5, 2)	0.037	0.024	0.040	0.022	0.032	0.033
Weib(2, 2)	Weib(10, 2)	0.002	0.001	0.001	0.001	0.000	0.000
LogN(0, 1)	LogN(0, 1)	0.059	0.052	0.052	0.063	0.069	0.058
LogN(1, 1)	LogN(1, 1)	0.063	0.054	0.046	0.060	0.068	0.083
LogN(1, 1.1)	LogN(2, 1)	0.000	0.000	0.000	0.000	0.000	0.000
LogN(1, 1.5)	LogN(2, 1)	0.000	0.000	0.000	0.000	0.000	0.000
LogN(1, 2)	LogN(2, 1)	0.003	0.000	0.001	0.000	0.000	0.000
LogN(1, 2)	LogN(3, 4)	0.015	0.015	0.011	0.007	0.007	0.003

Tabelle 2.2: Empirische Güten von $\varphi_{m,n}^{KS}$ und $\varphi_{m,n}^{CvM}$ bei Gültigkeit der Hypothese für $\alpha = 0.05$ und a) $m = 20, n = 20$, b) $m = 20, n = 30$, c) $m = 30, n = 20$.

F	G	$\varphi_{m,n}^{KS}$			$\varphi_{m,n}^{CvM}$		
		a)	b)	c)	a)	b)	c)
$\mathfrak{R}([0, 1])$	$\mathfrak{R}([0, \frac{1}{2}])$	0.812	0.856	0.849	0.903	0.964	0.925
$\mathfrak{R}([0, 1])$	$\mathfrak{R}([0, 0.8])$	0.211	0.195	0.205	0.278	0.320	0.307
$\mathfrak{R}([0, 1])$	$\mathfrak{R}([0, 0.99])$	0.054	0.060	0.075	0.048	0.065	0.074
$\mathfrak{R}([2, 3])$	$\mathfrak{R}([1, 4])$	0.210	0.210	0.283	0.229	0.240	0.317
$\mathfrak{R}([2, 3])$	$\mathfrak{R}([1.5, 3.5])$	0.120	0.118	0.134	0.135	0.120	0.150
$\mathfrak{R}([2, 3])$	$\mathfrak{R}([1.9, 3.1])$	0.076	0.064	0.076	0.066	0.055	0.076
Exp(1)	Exp(1.5)	0.276	0.320	0.320	0.316	0.331	0.357
Exp(1)	Exp(1.4)	0.238	0.236	0.295	0.239	0.268	0.255
Exp(1)	Exp(1.3)	0.268	0.172	0.192	0.192	0.221	0.217
Exp(1)	Exp(1.2)	0.128	0.145	0.139	0.138	0.146	0.159
Exp(1)	Exp(1.1)	0.087	0.094	0.099	0.098	0.115	0.108
$\Gamma(2.1, 1)$	$\Gamma(2, 1)$	0.108	0.092	0.102	0.027	0.090	0.111
$\Gamma(2.2, 1)$	$\Gamma(2, 1)$	0.130	0.137	0.151	0.145	0.153	0.196
$\Gamma(2.5, 1)$	$\Gamma(2, 1)$	0.320	0.350	0.389	0.337	0.354	0.396
Exp(1)	χ_1^2	0.330	0.359	0.441	0.201	0.263	0.217
Weib(2.1, 1)	Weib(2, 1)	0.066	0.065	0.082	0.068	0.071	0.085
Weib(2.2, 1)	Weib(2, 1)	0.089	0.073	0.088	0.051	0.066	0.072
Weib(2.3, 1)	Weib(2, 1)	0.085	0.080	0.102	0.069	0.086	0.092
Weib(2.4, 1)	Weib(2, 1)	0.083	0.095	0.097	0.071	0.076	0.095
Weib(2.5, 1)	Weib(2, 1)	0.101	0.087	0.112	0.090	0.094	0.123
Weib(1, 2)	Weib(1, 1.9)	0.069	0.076	0.076	0.073	0.077	0.093
Weib(1, 2)	Weib(1, 1.5)	0.187	0.188	0.226	0.197	0.248	0.227
Weib(1, 2)	Weib(1, 1)	0.560	0.632	0.647	0.578	0.647	0.629
LogN(0, 1)	LogN(0, 1.1)	0.057	0.068	0.070	0.064	0.061	0.074
LogN(0, 1)	LogN(0, 1.5)	0.094	0.108	0.114	0.053	0.071	0.064
LogN(0, 1)	LogN(0, 2)	0.185	0.222	0.265	0.268	0.244	0.169

Tabelle 2.3: Empirische Güten von $\varphi_{m,n}^{KS}$ und $\varphi_{m,n}^{CvM}$ bei Gültigkeit der Alternative für $\alpha = 0.05$ und a) $m = 20, n = 20$, b) $m = 20, n = 30$, c) $m = 30, n = 20$.

Kapitel 3

Benachbarte Alternativen

Neben dem in Kapitel 2.2 untersuchten Verhalten der Testgröße bei Vorliegen einer festen Alternative $F \not\prec_{\mathcal{L}} G$, stellt sich auch die Frage nach dem Verhalten der Testgrößen bei Vorliegen einer Folge von Alternativverteilungen, die - in noch zu präzisierendem Sinne - gegen eine Hypothesenverteilung konvergieren. Dies führt auf den bekannten Begriff der benachbarten Alternative (vgl. Kapitel C).

Es seien im Folgenden für jedes $m \in \mathbb{N}$ $\mathbf{X}_{m1}, \dots, \mathbf{X}_{mm}$ unabhängige, identisch verteilte und $\mathbb{R}_{\geq 0}^d$ -wertige Zufallsvektoren mit der Verteilungsfunktion F und für jedes $n \in \mathbb{N}$ seien $\mathbf{Y}_{n1}, \dots, \mathbf{Y}_{nn}$, unabhängige, identisch verteilte und $\mathbb{R}_{\geq 0}^d$ -wertige Zufallsvektoren mit der Verteilungsfunktion G . Seien $\mathbf{X}_{m1}, \dots, \mathbf{X}_{mm}, \mathbf{Y}_{n1}, \dots, \mathbf{Y}_{nn}$ unabhängig. Wir fassen dabei $\mathbf{X}_{mi}, \mathbf{Y}_{nj}$, $i = 1, \dots, m, j = 1, \dots, n$, für feste $m, n \in \mathbb{N}$ als die auf

$$\bigotimes_{i=1}^{m+n} \left(\mathbb{R}_{\geq 0}^d, \mathfrak{B}(\mathbb{R}_{\geq 0}^d) \right) = \left((\mathbb{R}_{\geq 0}^d)^{m+n}, (\mathfrak{B}(\mathbb{R}_{\geq 0}^d))^{m+n} \right)$$

definierten Koordiantenprojektionen

$$\begin{aligned} \mathbf{X}_{mi} : \quad & \left((\mathbb{R}_{\geq 0}^d)^{m+n}, (\mathfrak{B}(\mathbb{R}_{\geq 0}^d))^{m+n} \right) \rightarrow \left(\mathbb{R}_{\geq 0}^d, \mathfrak{B}(\mathbb{R}_{\geq 0}^d) \right) \\ & (\mathbf{x}_1, \dots, \mathbf{x}_m, \mathbf{y}_1, \dots, \mathbf{y}_n) \mapsto \mathbf{x}_i \end{aligned}$$

und

$$\begin{aligned} \mathbf{Y}_{nj} : \quad & \left((\mathbb{R}_{\geq 0}^d)^{m+n}, (\mathfrak{B}(\mathbb{R}_{\geq 0}^d))^{m+n} \right) \rightarrow \left(\mathbb{R}_{\geq 0}^d, \mathfrak{B}(\mathbb{R}_{\geq 0}^d) \right) \\ & (\mathbf{x}_1, \dots, \mathbf{x}_m, \mathbf{y}_1, \dots, \mathbf{y}_n) \mapsto \mathbf{y}_j \end{aligned}$$

auf.

Es sei $F \leq_{\mathcal{L}} G$. Die gemeinsame Verteilung von $\mathbf{X}_{m1}, \dots, \mathbf{X}_{mm}, \mathbf{Y}_{n1}, \dots, \mathbf{Y}_{nn}$ auf $((\mathbb{R}_{\geq 0}^d)^{m+n}, (\mathfrak{B}(\mathbb{R}_{\geq 0}^d))^{m+n})$ ist dann gegeben durch

$$P^{(m+n)} = \bigotimes_{i=1}^m F \otimes \bigotimes_{i=1}^n G.$$

Wir wollen Folgen von Alternativen der Form

$$Q^{(m+n)} = \bigotimes_{i=1}^m \hat{F}_{m,n} \otimes \bigotimes_{i=1}^n \hat{G}_{m,n}, \quad m, n \in \mathbb{N}, \quad (3.0.1)$$

mit

$$d\hat{F}_{m,n} = \left(1 + \frac{h_{m,n}}{\sqrt{\frac{mn}{m+n}}} \right) dF, \quad m, n \in \mathbb{N}, \quad (3.0.2)$$

und

$$d\hat{G}_{m,n} = \left(1 + \frac{g_{m,n}}{\sqrt{\frac{mn}{m+n}}} \right) dG, \quad m, n \in \mathbb{N}, \quad (3.0.3)$$

betrachten. Dabei seien $h_{m,n}, g_{m,n} : (\mathbb{R}_{\geq 0}^d, \mathfrak{B}(\mathbb{R}_{\geq 0}^d)) \rightarrow \mathbb{R}$ Folgen integrierbarer Funktionen mit den fast sicheren Grenzwerten h und g unter $P^{(m+n)}$. Ohne Einschränkung sei $h_{m,n}, g_{m,n} \neq 0$ fast sicher für alle $m, n \in \mathbb{N}$. Offenbar gilt

$$\int_{\mathbb{R}_{\geq 0}^d} h_{m,n}(\mathbf{x}) dF(\mathbf{x}) = \int_{\mathbb{R}_{\geq 0}^d} g_{m,n}(\mathbf{x}) dG(\mathbf{x}) = 0 \quad (3.0.4)$$

für alle $m, n \in \mathbb{N}$. Außerdem halten wir an dieser Stelle fest, dass $\hat{F}_{m,n} \not\leq_{\mathcal{L}} \hat{G}_{m,n}$ genau dann der Fall ist, wenn ein $\mathbf{t}_{m,n} \in \mathbb{R}_{\geq 0}^d$ existiert mit

$$\int_{\mathbb{R}_{\geq 0}^d} g_{m,n}(\mathbf{y}) e^{-\mathbf{t}_{m,n}' \mathbf{y}} dG(\mathbf{y}) - \int_{\mathbb{R}_{\geq 0}^d} h_{m,n}(\mathbf{x}) e^{-\mathbf{t}_{m,n}' \mathbf{x}} dF(\mathbf{x}) > \sqrt{\frac{mn}{m+n}} (\mathcal{L}_F(\mathbf{t}_{m,n}) - \mathcal{L}_G(\mathbf{t}_{m,n})).$$

3.1 Asymptotisches Verhalten

Das zu untersuchende Konvergenzverhalten der Testgrößen in der einleitend beschriebenen Situation hängt von der Art der in (3.0.1) definierten Alternativen, also insbesondere von den Funktionen $h_{m,n}$ und $g_{m,n}$ ab. Bei geeigneter Wahl dieser Funktionen ist zu hoffen, dass die Folgen der Verteilungen $\{P^{(m+n)}\}_{m,n \in \mathbb{N}}$ und $\{Q^{(m+n)}\}_{m,n \in \mathbb{N}}$ benachbart sind. Eine in diesem Zusammenhang üblicherweise getroffene Voraussetzung ist die gleichmäßige Beschränktheit der Funktionenfolgen $\{h_{m,n}\}_{m,n \in \mathbb{N}}$ und $\{g_{m,n}\}_{m,n \in \mathbb{N}}$, eine Bedingung die in vielen typischen Anwendungsbeispielen allerdings nicht gegeben ist.

Wir unterstellen im Folgenden die schwächere gleichmäßige Beschränktheit der vierten absoluten Momente der Folgen $\{h_{m,n}\}_{m,n \in \mathbb{N}}$ und $\{g_{m,n}\}_{m,n \in \mathbb{N}}$, also

$$\sup_{m,n \in \mathbb{N}} E(h_{m,n}^4(\mathbf{X})) < \infty, \quad \sup_{m,n \in \mathbb{N}} E(g_{m,n}^4(\mathbf{Y})) < \infty \quad (3.1.1)$$

für $\mathbf{X} \sim F$ bzw. $\mathbf{Y} \sim G$. Für viele gängige Verteilungsfamilien sind diese Bedingungen erfüllt (siehe [Tah09], 4.2, für den Einstichprobenfall).

3.1.1. Satz. *Es gelte (3.1.1) und $\frac{m}{m+n} \rightarrow \tau \in (0, 1)$. Dann sind die Folgen der Wahrscheinlichkeitsmaße $\{P^{(m+n)}\}_{m,n \in \mathbb{N}}$ und $\{Q^{(m+n)}\}_{m,n \in \mathbb{N}}$ definiert durch*

$$P^{(m+n)} = \bigotimes_{i=1}^m F \otimes \bigotimes_{i=1}^n G, \quad Q^{(m+n)} = \bigotimes_{i=1}^m \hat{F}_{m,n} \otimes \bigotimes_{i=1}^n \hat{G}_{m,n},$$

und (3.0.2) sowie (3.0.3) benachbart.

Beweis. Wir betrachten den Log-Likelihood-Quotienten $\Lambda_{m,n} = \log \frac{dQ^{(m+n)}}{dP^{(m+n)}}$. Aufgrund der Unabhängigkeit der $\mathbf{X}_{mi}, \mathbf{Y}_{nj}$ für feste $m, n \in \mathbb{N}$ ist

$$\begin{aligned} \Lambda_{m,n}(\mathbf{X}_{m1}, \dots, \mathbf{X}_{mm}, \mathbf{Y}_{n1}, \dots, \mathbf{Y}_{nn}) &= \log \frac{dQ^{(m+n)}(\mathbf{X}_{m1}, \dots, \mathbf{X}_{mm}, \mathbf{Y}_{n1}, \dots, \mathbf{Y}_{nn})}{dP^{(m+n)}(\mathbf{X}_{m1}, \dots, \mathbf{X}_{mm}, \mathbf{Y}_{n1}, \dots, \mathbf{Y}_{nn})} \\ &= \log \left(\prod_{i=1}^m \frac{d\hat{F}_{m,n}(\mathbf{X}_{mi})}{dF(\mathbf{X}_{mi})} \cdot \prod_{j=1}^n \frac{d\hat{G}_{m,n}(\mathbf{Y}_{nj})}{dG(\mathbf{Y}_{nj})} \right) \\ &= \sum_{i=1}^m \log \left(1 + \frac{h_{m,n}(\mathbf{X}_{mi})}{\sqrt{\frac{mn}{m+n}}} \right) + \sum_{j=1}^n \log \left(1 + \frac{g_{m,n}(\mathbf{Y}_{nj})}{\sqrt{\frac{mn}{m+n}}} \right). \end{aligned}$$

Wegen (3.1.1) können wir ohne Einschränkung annehmen, dass $1 + \frac{h_{m,n}(\mathbf{X}_{mi})}{\sqrt{\frac{mn}{m+n}}}$ und $1 + \frac{g_{m,n}(\mathbf{Y}_{nj})}{\sqrt{\frac{mn}{m+n}}}$ für alle $m, n \in \mathbb{N}$ jeweils positiv sind. Eine Taylorentwicklung liefert dann für die Summanden in der letzten Gleichung die Darstellung

$$\log \left(1 + \frac{h_{m,n}(\mathbf{X}_{mi})}{\sqrt{\frac{mn}{m+n}}} \right) = \frac{h_{m,n}(\mathbf{X}_{mi})}{\sqrt{\frac{mn}{m+n}}} - \frac{1}{2} \frac{h_{m,n}^2(\mathbf{X}_{mi})}{\frac{mn}{m+n}} + R_2(h_{m,n}(\mathbf{X}_{mi})), \quad m, n \in \mathbb{N},$$

mit dem Restglied $R_2(h_{m,n}(\mathbf{X}_{mi}))$. Unter den getroffenen Annahmen an die Funktionen $h_{m,n}$ gilt mit entsprechenden Überlegungen wie in [Tah09]

$$\sum_{i=1}^m R_2(h_{m,n}(\mathbf{X}_{mi})) = o_{P^{(m+n)}}(1), \quad \text{für } m, n \rightarrow \infty.$$

Analog erhalten wir durch eine Taylorentwicklung der Summanden $\log\left(1 + \frac{g_{m,n}(\mathbf{Y}_{nj})}{\sqrt{\frac{mn}{m+n}}}\right)$

$$\log\left(1 + \frac{g_{m,n}(\mathbf{Y}_{nj})}{\sqrt{\frac{mn}{m+n}}}\right) = \frac{g_{m,n}(\mathbf{Y}_{nj})}{\sqrt{\frac{mn}{m+n}}} - \frac{1}{2} \frac{g_{m,n}^2(\mathbf{Y}_{nj})}{\frac{mn}{m+n}} + R_2(g_{m,n}(\mathbf{Y}_{nj})),$$

mit $R_2(g_{m,n}(\mathbf{Y}_{nj}))$ als Restglied der Taylorentwicklung mit der Eigenschaft

$$\sum_{j=1}^n R_2(g_{m,n}(\mathbf{Y}_{nj})) = o_{P(m+n)}(1), \quad \text{für } m, n \rightarrow \infty,$$

und daraus somit die Darstellung

$$\begin{aligned} \Lambda_{m,n}(\mathbf{X}_{m1} \dots, \mathbf{Y}_{nn}) &= \sum_{i=1}^m \left(\frac{h_{m,n}(\mathbf{X}_{mi})}{\sqrt{\frac{mn}{m+n}}} - \frac{1}{2} \frac{h_{m,n}^2(\mathbf{X}_{mi})}{\frac{mn}{m+n}} \right) \\ &\quad + \sum_{j=1}^n \left(\frac{g_{m,n}(\mathbf{Y}_{nj})}{\sqrt{\frac{mn}{m+n}}} - \frac{1}{2} \frac{g_{m,n}^2(\mathbf{Y}_{nj})}{\frac{mn}{m+n}} \right) + o_{P(m+n)}(1). \end{aligned}$$

Wir betrachten den ersten Summanden. Zunächst implizieren die fast sichere Konvergenz $h_{m,n} \rightarrow h, m, n \rightarrow \infty$, und (3.1.1) die Konvergenz von $h_{m,n}$ gegen h in L^2 , d.h., es gilt

$$\lim_{m,n \rightarrow \infty} \text{Var}(h_{m,n}(\mathbf{X}_{mi})) = \text{Var}(h(\mathbf{X}_1)) = \int_{\mathbb{R}_{\geq 0}^d} h^2(\mathbf{x}) dF(\mathbf{x}).$$

Die Bedingung (3.1.1) impliziert für die Zufallsvariablen $h_{m,n}(\mathbf{X}_{mi}), i = 1, \dots, m$, ferner die Ljapunovbedingung, denn

$$\lim_{m,n \rightarrow \infty} \left(\sum_{i=1}^m \text{Var}(h_{m,n}(\mathbf{X}_{mi})) \right)^{-2} \sum_{i=1}^m E(h_{m,n}^4(\mathbf{X}_{mi})) = 0.$$

Mit $\frac{m}{m+n} \rightarrow \tau$ und $\sigma_2^2 := \int_{\mathbb{R}_{\geq 0}^d} h^2(\mathbf{x}) dF(\mathbf{x})$ liefert der zentrale Grenzwertsatz von Lindeberg-Feller für $m, n \rightarrow \infty$

$$\sqrt{\frac{m+n}{n}} \frac{1}{\sqrt{m}} \sum_{i=1}^m h_{m,n}(\mathbf{X}_{mi}) \xrightarrow{v} \sqrt{\frac{1}{1-\tau}} N(0, \sigma_2^2) = N\left(0, \frac{1}{1-\tau} \sigma_2^2\right),$$

unter $P^{(m+n)}$. Außerdem ergibt sich

$$-\frac{1}{2} \frac{m+n}{n} \sum_{i=1}^m \frac{h_{m,n}^2(\mathbf{X}_{mi})}{m} \xrightarrow{P^{(m+n)}} -\frac{1}{2} \frac{1}{1-\tau} \sigma_2^2, \quad m, n \rightarrow \infty.$$

Mit dem Lemma von Slutsky ergibt sich daraus zusammenfassend, dass

$$\sum_{i=1}^m \left(\frac{h_{m,n}(\mathbf{X}_{mi})}{\sqrt{\frac{mn}{m+n}}} - \frac{1}{2} \frac{h_{m,n}^2(\mathbf{X}_{mi})}{\frac{mn}{m+n}} \right) \xrightarrow{v} N \left(-\frac{1}{2} \frac{1}{1-\tau} \sigma_2^2, \frac{1}{1-\tau} \sigma_2^2 \right), \quad m, n \rightarrow \infty,$$

unter $P^{(m+n)}$ gilt. Analog schließt man

$$\sum_{j=1}^n \left(\frac{g_{m,n}(\mathbf{Y}_{nj})}{\sqrt{\frac{mn}{m+n}}} - \frac{1}{2} \frac{g_{m,n}^2(\mathbf{Y}_{nj})}{\frac{mn}{m+n}} \right) \xrightarrow{v} N \left(-\frac{1}{2} \frac{1}{\tau} \sigma_1^2, \frac{1}{\tau} \sigma_1^2 \right), \quad m, n \rightarrow \infty,$$

unter $P^{(m+n)}$ mit $\sigma_1^2 := \int_{\mathbb{R}_{\geq 0}^d} g(\mathbf{x})^2 dG(\mathbf{x})$. Aufgrund der Unabhängigkeit liefert dies

$$\Lambda_{m,n}(\mathbf{X}_{m1} \dots, \mathbf{Y}_{nn}) \xrightarrow{v} N \left(-\frac{1}{2} \sigma^2, \sigma^2 \right), \quad m, n \rightarrow \infty, \quad (3.1.2)$$

unter $P^{(m+n)}$, mit

$$\sigma^2 = \frac{1}{\tau} \sigma_1^2 + \frac{1}{1-\tau} \sigma_2^2. \quad (3.1.3)$$

Wir wenden schließlich Satz C.2 an und erhalten aus (3.1.2) so

$$\Lambda_{m,n}(\mathbf{X}_{m1} \dots, \mathbf{Y}_{nn}) \xrightarrow{v} N \left(\frac{1}{2} \sigma^2, \sigma^2 \right), \quad m, n \rightarrow \infty,$$

unter $Q^{(m+n)}$ sowie die behauptete Benachbarkeit von $\{P^{(m+n)}\}_{m,n \in \mathbb{N}}$ und $\{Q^{(m+n)}\}_{m,n \in \mathbb{N}}$. \square

Wir wollen nun die Grenzverteilung der Testgrößen

$$T_{m,n}^{KS} = \sup_{\mathbf{t} \geq 0} \Delta_{m,n}(\mathbf{t})$$

und

$$T_{m,n}^{CvM} = \int_{\mathbb{R}_{\geq 0}^d} \Delta_{m,n}^+(\mathbf{t}) \omega(\mathbf{t}) d\mathbf{t},$$

unter der Folge der benachbarten Alternativen $\{Q^{(m+n)}\}_{m,n \in \mathbb{N}}$ untersuchen. Es bezeichne erneut $\Lambda_{m,n} = \Lambda_{m,n}(\mathbf{X}_{m1} \dots, \mathbf{Y}_{nn})$ den Log-Likelihoodquotienten aus dem vorherigen Satz. Ferner unterstellen wir im Folgenden den Spezialfall $F = G$.

3.1.2. Satz. Für $\mathbf{t}_1, \dots, \mathbf{t}_k \in \mathbb{R}_{\geq 0}^d$, $k \in \mathbb{N}$, sei $\mathbf{c}' = (c(\mathbf{t}_1), \dots, c(\mathbf{t}_k)) \in \mathbb{R}^k$, mit

$$c(\mathbf{t}_i) := \int_{\mathbb{R}_{\geq 0}^d} e^{-\mathbf{t}_i' \mathbf{x}} (g(\mathbf{x}) - h(\mathbf{x})) dF(\mathbf{x}), \quad i = 1, \dots, k.$$

Weiter sei $\sigma^2 = \frac{1}{\tau}\sigma_1^2 + \frac{1}{1-\tau}\sigma_2^2$ definiert wie in (3.1.3) und $\Sigma = (\sigma_{ij})_{1 \leq i, j \leq k} \in \mathbb{R}^{k \times k}$ habe die Einträge $\sigma_{ij} = \rho(\mathbf{t}_i, \mathbf{t}_j)$, wobei

$$\rho(\mathbf{s}, \mathbf{t}) = \rho_F(\mathbf{s}, \mathbf{t}), \quad \mathbf{s}, \mathbf{t} \in \mathbb{R}_{\geq 0}^d,$$

die Kovarianzfunktion des Grenzprozesses \mathbb{L} aus Satz 2.2.4 im Fall $F = G$ bezeichnet. Dann gilt unter der Folge der Verteilungen $\{P^{(m+n)}\}_{m, n \in \mathbb{N}}$

$$(\Delta_{m,n}(\mathbf{t}_1), \dots, \Delta_{m,n}(\mathbf{t}_k), \Lambda_{m,n}) \xrightarrow{v} N_{k+1} \left(\begin{pmatrix} 0 \\ \vdots \\ 0 \\ -\frac{1}{2}\sigma^2 \end{pmatrix}, \begin{pmatrix} \Sigma & \mathbf{c} \\ \mathbf{c}' & \sigma^2 \end{pmatrix} \right), \quad m, n \rightarrow \infty.$$

Beweis. Nach Satz 3.1.1 hat $\Lambda_{m,n}$ dieselbe Grenzverteilung wie

$$\tilde{\Lambda}_{m,n} := \sum_{i=1}^m \left(\frac{h_{m,n}(\mathbf{X}_{mi})}{\sqrt{\frac{mn}{m+n}}} - \frac{1}{2} \frac{h_{m,n}^2(\mathbf{X}_{mi})}{\frac{mn}{m+n}} \right) + \sum_{j=1}^n \left(\frac{g_{m,n}(\mathbf{Y}_{nj})}{\sqrt{\frac{mn}{m+n}}} - \frac{1}{2} \frac{g_{m,n}^2(\mathbf{Y}_{nj})}{\frac{mn}{m+n}} \right).$$

Die Behauptung ist nach dem Lemma von Cramér-Wold (vgl. z.B. [WMF95], 5.69) äquivalent dazu, dass für jedes $\mathbf{x}' = (x_1, \dots, x_{k+1}) \in \mathbb{R}^{k+1}$ und $\mathbf{t}_1, \dots, \mathbf{t}_k \in \mathbb{R}_{\geq 0}^d$

$$\sum_{i=1}^k x_i \Delta_{m,n}(\mathbf{t}_i) + x_{k+1} \tilde{\Lambda}_{m,n} \xrightarrow{v} N \left(\mathbf{x}' \begin{pmatrix} 0 \\ \vdots \\ 0 \\ -\frac{1}{2}\sigma^2 \end{pmatrix}, \mathbf{x}' \begin{pmatrix} \Sigma & \mathbf{c} \\ \mathbf{c}' & \sigma^2 \end{pmatrix} \mathbf{x} \right) \quad (3.1.4)$$

unter $P^{(m+n)}$ gilt. Es ist

$$\begin{aligned} & \sum_{i=1}^k x_i \Delta_{m,n}(\mathbf{t}_i) + x_{k+1} \tilde{\Lambda}_{m,n} \\ &= \sum_{i=1}^k x_i \sqrt{\frac{mn}{m+n}} \left(\frac{1}{n} \sum_{j=1}^n e^{-\mathbf{t}'_i \mathbf{Y}_{nj}} - \mathcal{L}_{\mathbf{Y}_{n1}}(\mathbf{t}_i) \right) + x_{k+1} \sum_{j=1}^n \left(\frac{g_{m,n}(\mathbf{Y}_{nj})}{\sqrt{\frac{mn}{m+n}}} - \frac{1}{2} \frac{g_{m,n}^2(\mathbf{Y}_{nj})}{\frac{mn}{m+n}} \right) \\ &+ \sum_{i=1}^k x_i \sqrt{\frac{mn}{m+n}} \left(\mathcal{L}_{\mathbf{X}_{m1}}(\mathbf{t}_i) - \frac{1}{m} \sum_{\nu=1}^m e^{-\mathbf{t}'_i \mathbf{X}_{m\nu}} \right) + x_{k+1} \sum_{\nu=1}^m \left(\frac{h_{m,n}(\mathbf{X}_{m\nu})}{\sqrt{\frac{mn}{m+n}}} - \frac{1}{2} \frac{h_{m,n}^2(\mathbf{X}_{m\nu})}{\frac{mn}{m+n}} \right) \\ &+ \sum_{i=1}^k x_i \sqrt{\frac{mn}{m+n}} (\mathcal{L}_{\mathbf{Y}_{n1}}(\mathbf{t}_i) - \mathcal{L}_{\mathbf{X}_{m1}}(\mathbf{t}_i)). \end{aligned}$$

für alle $\mathbf{t}_1, \dots, \mathbf{t}_k \in \mathbb{R}_{\geq 0}^d$. Da nach Voraussetzung $F = G$ ist, verschwindet der letzte Term im obigen Ausdruck für alle $\mathbf{t}_i \in \mathbb{R}_{\geq 0}^d$. Wir erhalten daher die Darstellung

$$\sum_{i=1}^k x_i \Delta_{m,n}(\mathbf{t}_i) + x_{k+1} \tilde{\Lambda}_{m,n} = \frac{1}{\sqrt{n}} \sum_{j=1}^n C_{m,n}^{(1)}(\mathbf{Y}_{nj}) + \frac{1}{\sqrt{m}} \sum_{\nu=1}^m C_{m,n}^{(2)}(\mathbf{X}_{m\nu})$$

mit

$$\begin{aligned} C_{m,n}^{(1)}(\mathbf{Y}_{nj}) &= \sum_{i=1}^k x_i \sqrt{\frac{m}{m+n}} (\exp(-\mathbf{t}'_i \mathbf{Y}_{nj}) - \mathcal{L}_{\mathbf{Y}_{n1}}(\mathbf{t}_i)) \\ &\quad + x_{k+1} \left(\sqrt{\frac{m+n}{m}} g_{m,n}(\mathbf{Y}_{nj}) - \frac{1}{2} \frac{m+n}{m\sqrt{n}} g_{m,n}^2(\mathbf{Y}_{nj}) \right) \end{aligned}$$

und

$$\begin{aligned} C_{m,n}^{(2)}(\mathbf{X}_{m\nu}) &= \sum_{i=1}^k x_i \sqrt{\frac{n}{m+n}} (\mathcal{L}_{\mathbf{X}_{m1}}(\mathbf{t}_i) - \exp(-\mathbf{t}'_i \mathbf{X}_{m\nu})) \\ &\quad + x_{k+1} \left(\sqrt{\frac{m+n}{n}} h_{m,n}(\mathbf{X}_{m\nu}) - \frac{1}{2} \frac{m+n}{\sqrt{mn}} h_{m,n}^2(\mathbf{X}_{m\nu}) \right). \end{aligned}$$

Die Ausdrücke $\frac{1}{\sqrt{n}} \sum_{j=1}^n C_{m,n}^{(1)}(\mathbf{Y}_{nj})$ und $\frac{1}{\sqrt{m}} \sum_{\nu=1}^m C_{m,n}^{(2)}(\mathbf{X}_{m\nu})$ sind stochastisch unabhängig.

Wir betrachten zunächst $C_{m,n}^{(1)}(\mathbf{Y}_{nj})$. Mit (3.0.4) ist

$$E \left(C_{m,n}^{(1)}(\mathbf{Y}_{nj}) \right) = -x_{k+1} \frac{1}{2} \frac{m+n}{m\sqrt{n}} E g_{m,n}^2(\mathbf{Y}_{n1}), \quad \forall j = 1, \dots, n,$$

so dass mit der L^2 -Konvergenz von $g_{m,n}$ für jedes $j = 1, \dots, n$

$$\lim_{m,n \rightarrow \infty} \sqrt{n} E \left(C_{m,n}^{(1)}(\mathbf{Y}_{nj}) \right) = -\frac{x_{k+1}}{2} \frac{1}{\tau} E g^2(\mathbf{Y}_{n1})$$

folgt. Außerdem ist für jedes $j = 1, \dots, n$

$$\begin{aligned} \text{Var}(C_{m,n}^{(1)}(\mathbf{Y}_{nj})) &= \sum_{r,s=1}^k x_r x_s \frac{m}{m+n} \text{Cov} \left(e^{-\mathbf{t}'_r \mathbf{Y}_{nj}} - \mathcal{L}_{\mathbf{Y}_{n1}}(\mathbf{t}_r), e^{-\mathbf{t}'_s \mathbf{Y}_{nj}} - \mathcal{L}_{\mathbf{Y}_{n1}}(\mathbf{t}_s) \right) \\ &\quad + x_{k+1}^2 \frac{m+n}{m} \text{Var}(g_{m,n}(\mathbf{Y}_{nj})) + x_{k+1}^2 \frac{1}{4} \frac{(m+n)^2}{m^2 n} \text{Var}(g_{m,n}^2(\mathbf{Y}_{nj})) \\ &\quad - x_{k+1}^2 \sqrt{\frac{m+n}{m}} \frac{m+n}{m\sqrt{n}} \text{Cov}(g_{m,n}(\mathbf{Y}_{nj}), g_{m,n}^2(\mathbf{Y}_{nj})) \\ &\quad + 2 \sum_{r=1}^k x_r x_{k+1} \sqrt{\frac{m}{m+n}} \sqrt{\frac{m+n}{m}} \text{Cov} \left(e^{-\mathbf{t}'_r \mathbf{Y}_{nj}} - \mathcal{L}_{\mathbf{Y}_{n1}}(\mathbf{t}_r), g_{m,n}(\mathbf{Y}_{nj}) \right) \\ &\quad - \sum_{r=1}^k x_r x_{k+1} \sqrt{\frac{m}{m+n}} \frac{m+n}{m\sqrt{n}} \text{Cov} \left(e^{-\mathbf{t}'_r \mathbf{Y}_{nj}} - \mathcal{L}_{\mathbf{Y}_{n1}}(\mathbf{t}_r), g_{m,n}^2(\mathbf{Y}_{nj}) \right). \end{aligned} \tag{3.1.5}$$

Es gilt

$$\begin{aligned} \text{Cov} \left(e^{-\mathbf{t}'_r \mathbf{Y}_{nj}} - \mathcal{L}_{\mathbf{Y}_{n1}}(\mathbf{t}_r), e^{-\mathbf{t}'_s \mathbf{Y}_{nj}} - \mathcal{L}_{\mathbf{Y}_{n1}}(\mathbf{t}_s) \right) &= \mathcal{L}_F(\mathbf{t}_r + \mathbf{t}_s) - \mathcal{L}_F(\mathbf{t}_r) \mathcal{L}_F(\mathbf{t}_s) \\ &= \rho_F(\mathbf{t}_r, \mathbf{t}_s) \end{aligned}$$

für alle $n \in \mathbb{N}, j = 1, \dots, n$, und außerdem ist

$$\text{Cov} \left(e^{-\mathbf{t}' \mathbf{Y}_{nj}} - \mathcal{L}_{\mathbf{Y}_{n1}}(\mathbf{t}), g_{m,n}(\mathbf{Y}_{nj}) \right) = E \left(\exp(-\mathbf{t}' \mathbf{Y}_{nj}) g_{m,n}(\mathbf{Y}_{nj}) \right).$$

Mit dem Satz von der majorisierten Konvergenz folgt

$$\sigma_{(1)}^2(\mathbf{t}) := \lim_{m,n \rightarrow \infty} E \left(\exp(-\mathbf{t}' \mathbf{Y}_{nj}) g_{m,n}(\mathbf{Y}_{nj}) \right) = \int_{\mathbb{R}_{\geq 0}^d} \exp(-\mathbf{t}' \mathbf{y}) g(\mathbf{y}) dF(\mathbf{y}).$$

Die gleichmäßige Beschränktheit der vierten Momente von $g_{m,n}$ und $\frac{m}{m+n} \rightarrow \tau$, ergeben mit (3.1.5) und $\sigma_1^2 = E g^2(\mathbf{Y}_{n1})$ dann

$$\begin{aligned} \alpha_1 := \lim_{m,n \rightarrow \infty} \text{Var}(C_{m,n}^{(1)}(\mathbf{Y}_{n1})) &= \sum_{r,s=1}^k x_r x_s \tau \rho_F(\mathbf{t}_r, \mathbf{t}_s) + x_{k+1}^2 \frac{1}{\tau} \sigma_1^2 \\ &\quad + 2 \sum_{r=1}^k x_r x_{k+1} \sigma_{(1)}^2(\mathbf{t}_r) \end{aligned}$$

unter $P^{(m+n)}$. Mit dem zentralen Grenzwertsatz von Lindeberg-Feller folgt

$$\frac{1}{\sqrt{n}} \sum_{j=1}^n C_{m,n}^{(1)}(\mathbf{Y}_{nj}) \xrightarrow{v} N \left(-\frac{x_{k+1}}{2} \frac{1}{\tau} \sigma_1^2, \alpha_1 \right), \quad m, n \rightarrow \infty,$$

unter $P^{(m+n)}$. Analoge Betrachtungen für $C_{m,n}^{(2)}(\mathbf{X}_{m\nu})$ führen mit $\sigma_2^2 = E h^2(\mathbf{X}_{m1})$,

$$\sigma_{(2)}^2(\mathbf{t}) := \lim_{m,n \rightarrow \infty} E \left(\exp(-\mathbf{t}' \mathbf{X}_{m\nu}) h_{m,n}(\mathbf{X}_{m\nu}) \right) = \int_{\mathbb{R}_{\geq 0}^d} \exp(-\mathbf{t}' \mathbf{x}) h(\mathbf{x}) dF(\mathbf{x})$$

und

$$\begin{aligned} \alpha_2 := \lim_{m,n \rightarrow \infty} \text{Var}(C_{m,n}^{(2)}(\mathbf{X}_{m\nu})) &= \sum_{r,s=1}^k x_r x_s (1 - \tau) \rho_F(\mathbf{t}_r, \mathbf{t}_s) + x_{k+1}^2 \frac{1}{1 - \tau} \sigma_2^2 \\ &\quad - 2 \sum_{r=1}^k x_r x_{k+1} \sigma_{(2)}^2(\mathbf{t}_r), \end{aligned}$$

auf

$$\frac{1}{\sqrt{m}} \sum_{\nu=1}^m C_{m,n}^{(2)}(\mathbf{X}_{m\nu}) \xrightarrow{v} N \left(-\frac{x_{k+1}}{2} \frac{1}{1 - \tau} \sigma_2^2, \alpha_2 \right), \quad m, n \rightarrow \infty,$$

unter $P^{(m+n)}$. Mit $c(\mathbf{t})$ wie im Satz definiert, erhalten wir die Darstellung $c(\mathbf{t}) = \sigma_{(1)}^2(\mathbf{t}) - \sigma_{(2)}^2(\mathbf{t})$, $\mathbf{t} \in \mathbb{R}_{\geq 0}^d$. Dann ist

$$\alpha_1 + \alpha_2 = \sum_{r,s=1}^k x_r x_s \rho_F(\mathbf{t}_r, \mathbf{t}_s) + x_{k+1}^2 \left(\frac{1}{\tau} \sigma_1^2 + \frac{1}{1-\tau} \sigma_2^2 \right) + 2x_{k+1} \sum_{r=1}^k x_r c(\mathbf{t}_r)$$

und es folgt daraus

$$\sum_{i=1}^k x_i \Delta_{m,n}(\mathbf{t}_i) + x_{k+1} \tilde{\Lambda}_{m,n} \xrightarrow{v} N \left(-\frac{x_{k+1}}{2} \left(\frac{1}{\tau} \sigma_1^2 + \frac{1}{1-\tau} \sigma_2^2 \right), \alpha_1 + \alpha_2 \right), \quad m, n \rightarrow \infty,$$

unter $P^{(m+n)}$. Mit der im Satz definierten Matrix Σ rechnet man schließlich nach, dass für alle $\mathbf{x}' = (x_1, \dots, x_{k+1}) \in \mathbb{R}^{k+1}$

$$\mathbf{x}' \begin{pmatrix} 0 \\ \vdots \\ 0 \\ -\frac{1}{2} \sigma^2 \end{pmatrix} = -\frac{x_{k+1}}{2} \left(\frac{1}{\tau} \sigma_1^2 + \frac{1}{1-\tau} \sigma_2^2 \right),$$

$$\mathbf{x}' \begin{pmatrix} \Sigma & \mathbf{c} \\ \mathbf{c}' & \sigma^2 \end{pmatrix} \mathbf{x} = \alpha_1 + \alpha_2,$$

gilt, was die Behauptung in (3.1.4) zeigt. \square

Nach diesen Vorüberlegungen können wir das Hauptresultat dieses Abschnitts formulieren.

3.1.3. Satz. *Unter der Folge $\{Q^{(m+n)}\}_{m,n \in \mathbb{N}}$ von benachbarten Alternativen gilt*

$$T_{m,n}^{KS} \xrightarrow{v} \sup_{\mathbf{t} \in \mathbb{R}_{\geq 0}^d} (\mathbb{L}(\mathbf{t}) + c(\mathbf{t})), \quad m, n \rightarrow \infty,$$

und

$$T_{m,n}^{CvM} \xrightarrow{v} \int_{\mathbb{R}_{\geq 0}^d} (\mathbb{L}(\mathbf{t}) + c(\mathbf{t}))^+ \omega(\mathbf{t}) d\mathbf{t}, \quad m, n \rightarrow \infty,$$

mit

$$c(\mathbf{t}) := \int_{\mathbb{R}_{\geq 0}^d} e^{-\mathbf{t}'\mathbf{x}} (g(\mathbf{x}) - h(\mathbf{x})) dF(\mathbf{x}), \quad \mathbf{t} \in \mathbb{R}_{\geq 0}^d.$$

Beweis.

Wie bereits gezeigt sind $\{P^{(m+n)}\}_{m,n \in \mathbb{N}}$ und $\{Q^{(m+n)}\}_{m,n \in \mathbb{N}}$ benachbarte Folgen von Wahrscheinlichkeitsmaßen und es ist

$$N_{k+1} \left(\begin{pmatrix} 0 \\ \vdots \\ 0 \\ -\frac{1}{2}\sigma^2 \end{pmatrix}, \begin{pmatrix} \Sigma & \mathbf{c} \\ \mathbf{c}' & \sigma^2 \end{pmatrix} \right),$$

die Grenzverteilung des Vektors $(\Delta_{m,n}(\mathbf{t}_1), \dots, \Delta_{m,n}(\mathbf{t}_k), \Lambda_{m,n})$, $\mathbf{t}_1, \dots, \mathbf{t}_k \in \mathbb{R}_{\geq 0}^d$, unter $\{P^{(m+n)}\}_{m,n \in \mathbb{N}}$. Nach Satz C.3 konvergiert daher $(\Delta_{m,n}(\mathbf{t}_1), \dots, \Delta_{m,n}(\mathbf{t}_k))$ für $\mathbf{t}_1, \dots, \mathbf{t}_k \in \mathbb{R}_{\geq 0}^d$ unter $Q^{(m+n)}$ gegen die Normalverteilung

$$N_k \left(\begin{pmatrix} c(\mathbf{t}_1) \\ \vdots \\ c(\mathbf{t}_k) \end{pmatrix}, \Sigma \right), \quad \mathbf{t}_1, \dots, \mathbf{t}_k \in \mathbb{R}_{\geq 0}^d. \quad (3.1.6)$$

Die Konvergenz der Randverteilungen und ein Straffheitsargument wie in Satz 2.2.4 liefert dann mit den dortigen Bezeichnungen und der Funktion $c(\mathbf{t})$, $\mathbf{t} \in \mathbb{R}_{\geq 0}^d$, aus dem vorigen Satz

$$\left(\Delta_{m,n}(\mathbf{t}), \mathbf{t} \in \mathbb{R}_{\geq 0}^d \right) \xrightarrow{v} \left(\mathbb{L}(\mathbf{t}) + c(\mathbf{t}), \mathbf{t} \in \mathbb{R}_{\geq 0}^d \right), \quad m, n \rightarrow \infty,$$

unter $Q^{(m+n)}$. Mit dem Satz von der stetigen Abbildung folgt die Behauptung für die Testgrößen. \square

Im Kapitel 2.3 hatten wir bereits gesehen, dass die $(1-\alpha)$ -Quantile $\hat{q}_{m,n}^{KS}(\alpha)$ von $\hat{T}_{m,n}^{KS}$ für $F = G$ gegen den Grenzwert $\hat{q}^{KS}(\alpha)$ konvergieren, wobei $\hat{q}^{KS}(\alpha)$ das entsprechende Quantil der Grenzverteilung \hat{T}^{KS} ist. Ebenso konvergiert die Folge der Quantile $\hat{q}_{m,n}^{CvM}(\alpha)$ gegen den Grenzwert $\hat{q}^{CvM}(\alpha)$. Die Benachbarkeit der Folgen der Verteilungen $\{P^{(m+n)}\}_{m,n \in \mathbb{N}}$ und $\{Q^{(m+n)}\}_{m,n \in \mathbb{N}}$ liefert dann sofort die entsprechende Konvergenz der Quantile unter $\{Q^{(m+n)}\}_{m,n \in \mathbb{N}}$.

3.1.4. Korollar. Es seien $\hat{q}_{m,n}^{KS}(\alpha)$ bzw. $\hat{q}_{m,n}^{CvM}(\alpha)$ die $(1-\alpha)$ -Quantile der Testgrößen $\hat{T}_{m,n}^{KS}$ bzw. $\hat{T}_{m,n}^{CvM}$. Sei $F = G$. Dann gilt

$$\hat{q}_{m,n}^{KS}(\alpha) \xrightarrow{Q^{(m+n)}} \hat{q}^{KS}(\alpha) \quad \text{und} \quad \hat{q}_{m,n}^{CvM}(\alpha) \xrightarrow{Q^{(m+n)}} \hat{q}^{CvM}(\alpha),$$

für $m, n \rightarrow \infty$ und $\alpha \in (0, 1)$.

Beweis. Wie bereits gesehen gilt unter $P^{(m+n)}$

$$\hat{q}_{m,n}^{KS}(\alpha) - \hat{q}^{KS}(\alpha) \xrightarrow{P^{(m+n)}} 0 \quad \text{und} \quad \hat{q}_{m,n}^{CvM}(\alpha) - \hat{q}^{CvM}(\alpha) \xrightarrow{P^{(m+n)}} 0.$$

Da $P^{(m+n)}$ und $Q^{(m+n)}$ benachbart sind, gilt dies nach Definition auch unter $Q^{(m+n)}$. \square

3.1.5. Bemerkung.

a) Für die Shiftfunktion

$$c(\mathbf{t}) := \int_{\mathbb{R}_{\geq 0}^d} e^{-\mathbf{t}'\mathbf{x}} (g(\mathbf{x}) - h(\mathbf{x})) dF(\mathbf{x}), \quad \mathbf{t} \in \mathbb{R}_{\geq 0}^d,$$

aus dem vorherigen Satz gilt offenbar $c(\mathbf{t}) \geq 0$, falls $g(\mathbf{x}) \geq h(\mathbf{x})$ für alle $\mathbf{x} \in \mathbb{R}^d$ ist. Somit gilt dann für $\alpha \in (0, \frac{1}{2})$ unter $Q^{(m+n)}$

$$\begin{aligned} \lim_{m,n \rightarrow \infty} Q^{(m+n)}(T_{m,n}^{KS} > \hat{q}_{m,n}^{KS}(\alpha)) &= P\left(\sup_{\mathbf{t} \in \mathbb{R}^d} (\mathbb{L}(\mathbf{t}) + c(\mathbf{t})) > \hat{q}^{KS}(\alpha)\right) \\ &\geq P\left(\sup_{\mathbf{t} \in \mathbb{R}^d} \mathbb{L}(\mathbf{t}) > \hat{q}^{KS}(\alpha)\right) \\ &= \alpha. \end{aligned}$$

Der Test ist für diese benachbarten Alternativen also asymptotisch unverfälscht.

b) Wir betrachten exemplarisch ein Kontaminationsmodell. Es seien dazu F und G gegeben und die Folge von benachbarten Alternativen sei definiert durch

$$\hat{F}_{m,n} = \left(1 - \frac{1}{\sqrt{\frac{mn}{m+n}}}\right) F + \frac{1}{\sqrt{\frac{mn}{m+n}}} H_1, \quad m, n \in \mathbb{N},$$

und

$$\hat{G}_{m,n} = \left(1 - \frac{1}{\sqrt{\frac{mn}{m+n}}}\right) G + \frac{1}{\sqrt{\frac{mn}{m+n}}} H_2, \quad m, n \in \mathbb{N},$$

für beliebige Verteilungen H_1, H_2 mit jeweils existierenden Radon-Nikodym-Ableitungen $\frac{dH_1}{dF}, \frac{dH_2}{dG}$. Dann gilt

$$d\hat{F}_{m,n} = \left(1 + \frac{1}{\sqrt{\frac{mn}{m+n}}} \left(\frac{dH_1}{dF} - 1\right)\right) dF,$$

$$d\hat{G}_{m,n} = \left(1 + \frac{1}{\sqrt{\frac{mn}{m+n}}} \left(\frac{dH_2}{dG} - 1 \right) \right) dG.$$

Setzen wir voraus, dass die vierten Momente von $h_{m,n}(\mathbf{x}) := \frac{dH_1}{dF}(\mathbf{x}) - 1$, $\mathbf{x} \in \mathbb{R}_{\geq 0}^d$, und $g_{m,n}(\mathbf{x}) := \frac{dH_2}{dG}(\mathbf{x}) - 1$, $\mathbf{x} \in \mathbb{R}_{\geq 0}^d$, beschränkt sind und ferner $F = G$ gilt, so ergibt sich für die Shiftfunktion $c(\mathbf{t})$ offenbar

$$\begin{aligned} c(\mathbf{t}) &= \int \exp(-\mathbf{t}'\mathbf{x}) \left(\frac{dH_2}{dG}(\mathbf{x}) - \frac{dH_1}{dF}(\mathbf{x}) \right) dF(\mathbf{x}) \\ &= \mathcal{L}_{H_2}(\mathbf{t}) - \mathcal{L}_{H_1}(\mathbf{t}). \end{aligned}$$

Wie wir in Teil a) eingesehen haben, ist $c(\mathbf{t}) \geq 0$, $\mathbf{t} \geq 0$, eine hinreichende Bedingung dafür, dass der Kolmogorov-Smirnov-Test unter den hier definierten benachbarten Alternativen unverfälscht ist. In diesem Fall ist $c(\mathbf{t}) \geq 0$, $\mathbf{t} \geq 0$, äquivalent zu $\mathcal{L}_{H_2}(\mathbf{t}) \geq \mathcal{L}_{H_1}(\mathbf{t})$ für alle $\mathbf{t} \geq 0$, was auf $H_2 \leq_{\mathcal{L}} H_1$ und daher auf $\hat{F}_{m,n} \geq_{\mathcal{L}} \hat{G}_{m,n}$ für alle $m, n \in \mathbb{N}$ führt.

3.2 Approximative Bahadursteigungen für eine Teilhypothese

In diesem Abschnitt wollen wir exemplarisch eine Effizienzuntersuchung für den auf der Supremumstestgröße

$$T_{m,n}^{KS} = \sup_{\mathbf{t} \geq 0} \Delta_{m,n}(\mathbf{t})$$

basierenden Test auf Laplace-Ordnung durchführen. Wir legen zu diesem Zweck kurz den auf Bahadur (1960) zurückgehenden Begriff der *relativen approximativen Bahadur-Effizienz* im Zweistichprobenfall dar.

Es seien $\mathbf{X}_1, \mathbf{X}_2, \dots$ unabhängig und identisch verteilt mit der Verteilungsfunktion F und $\mathbf{Y}_1, \mathbf{Y}_2, \dots$ unabhängig und identisch verteilt mit der Verteilungsfunktion G . Es seien $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{Y}_1, \mathbf{Y}_2, \dots$ unabhängig. Ferner sei $(m_k, n_k)_{k \in \mathbb{N}}$, $m_k, n_k \in \mathbb{N}$, derart, dass $\lim_{k \rightarrow \infty} \frac{m_k}{m_k + n_k} = \tau \in (0, 1)$ gilt. Wir betrachten Teilmengen $\mathfrak{P}_1, \mathfrak{P}_2 \subset \{F : \mathbb{R}^d \rightarrow [0, 1] \mid F \text{ Verteilungsfunktion}\}$ und setzen $\Theta := \mathfrak{P}_1 \times \mathfrak{P}_2$. Für jedes $k \in \mathbb{N}$ sei T_{m_k, n_k} eine auf den Zufallsvariablen $\mathbf{X}_1, \dots, \mathbf{X}_{m_k}, \mathbf{Y}_1, \dots, \mathbf{Y}_{n_k}$ basierende Teststatistik für eine Hypothese $H : (F, G) \in \Theta_0$, mit geeignetem $\Theta_0 \subset \Theta$. Große Werte von T_{m_k, n_k} sprechen dabei wie üblich gegen die Hypothese. Die folgende Definition geht für den Einstichprobenfall auf Bahadur selbst zurück (siehe [Bah60b]).

3.2.1. Definition. Die Folge der Teststatistiken $(T_{m_k, n_k})_{k \in \mathbb{N}}$ heißt *Standardfolge* für die Hypothese $H : (F, G) \in \Theta_0$, falls die folgenden Bedingungen gelten:

i) Für alle $\vartheta \in \Theta_0$ gilt

$$\lim_{k \rightarrow \infty} P(T_{m_k, n_k} \leq x) = \Psi(x), \quad x \in \mathbb{R},$$

mit einer nicht von ϑ abhängenden Verteilungsfunktion Ψ auf \mathbb{R} .

ii) Es gibt ein $a \in (0, \infty)$ mit

$$\lim_{x \rightarrow \infty} \frac{-2 \log(1 - \Psi(x))}{x^2} = a.$$

iii) Es gibt eine Funktion $b : \Theta \setminus \Theta_0 \rightarrow (0, \infty)$ mit der Eigenschaft, dass für alle $\vartheta \in \Theta_1$

$$\lim_{k \rightarrow \infty} P\left(\left|\sqrt{\frac{m_k + n_k}{m_k n_k}} T_{m_k, n_k} - b(\vartheta)\right| > x\right) = 0, \quad x > 0,$$

gilt.

In der ursprünglichen Definition der Standardfolge wird zusätzlich die Stetigkeit der in Bedingung i) genannten Verteilungsfunktion gefordert. Auf die Stetigkeit der Verteilungsfunktion Ψ kann aber für die folgenden Betrachtungen verzichtet werden (vergleiche [KK95], Bemerkung zur Definition 2).

Wir definieren dann die *approximative Bahadursteigung* der Folge $(T_{m_k, n_k})_{k \in \mathbb{N}}$ für die Hypothese H als

$$c(\vartheta) := \begin{cases} 0, & \vartheta \in \Theta_0, \\ ab^2(\vartheta), & \vartheta \in \Theta \setminus \Theta_0. \end{cases}$$

Dann lassen sich die Tests $(T_{m_k, n_k})_{k \in \mathbb{N}}$ für die Hypothese H mittels des Ausdrucks $c(\vartheta)$, $\vartheta \in \Theta \setminus \Theta_0$, bewerten. (Wir verzichten an dieser Stelle auf eine detailliertere Darstellung sowie auf anschauliche Interpretationen der Bahadursteigungen und verweisen für weitere Darlegungen der Begriffe auf die Standardarbeiten [Bah60b] und [Bah60a]. Für eine Diskussion darüber, inwieweit approximative Bahadur-Effizienzen gute Approximationen der exakten Bahadur-Effizienzen liefern, siehe [Bah67]).

Aufgrund der nicht vorhandenen Verteilungsfreiheit der Grenzverteilung von $T_{m, n}^{KS}$ bei Gültigkeit der Hypothese beschränken wir uns an dieser Stelle bei der Wahl von Θ_0 auf eine fest gewählte Verteilungsfamilie und betrachten dazu ausgewählte Skalenalternativen. Exemplarisch gehen wir davon von der Familie der Exponentialverteilungen auf $\mathbb{R}_{\geq 0}$ aus.

Der Fall anderer Verteilungsfamilien und der zugehörigen Skalenalternativen kann vielfach ähnlich behandelt werden.

Sei $\Theta := \{(F, G) : F = \text{Exp}(1), G = \text{Exp}(1 + \vartheta), \vartheta \in [0, \infty)\}$ und $\Theta_0 = \{(F, G) : F = G = \text{Exp}(1)\}$. Dann ist $\Theta_1 := \Theta \setminus \Theta_0 = \{(F, G) : F = \text{Exp}(1), G = \text{Exp}(1 + \vartheta), \vartheta \in (0, \infty)\}$. Dabei sei $\text{Exp}(\lambda), \lambda > 0$, erneut die Verteilungsfunktion der Exponentialverteilung mit Parameter λ . Wir identifizieren Θ_1 mit der Parametermenge $(0, \infty)$ und untersuchen das Testproblem

$$H : (F, G) \in \Theta_0 \quad \text{vs.} \quad K : (F, G) \in \Theta_1.$$

Es gilt für $t \geq 0$

$$\mathcal{L}_{\text{Exp}(1)}(t) = \frac{1}{1+t}, \quad \mathcal{L}_{\text{Exp}(1+\vartheta)}(t) = \frac{1+\vartheta}{t+(1+\vartheta)},$$

so dass für $(F, G) \in \Theta_0$ auch $F \leq_{\mathcal{L}} G$ und für $(F, G) \in \Theta_1$ entsprechend $F \not\leq_{\mathcal{L}} G$ gilt.

Wir wollen nun die approximative Bahadursteigung von $T_{m,n}^{KS}$ berechnen und überprüfen zu diesem Zweck, ob es sich bei $(T_{m,n}^{KS})_{m,n \in \mathbb{N}}$ um eine Standardfolge handelt. Zunächst erfüllt $T_{m,n}^{KS}$ aufgrund von Satz 2.2.5 offensichtlich Bedingung i). Für jedes $\vartheta \in (0, \infty)$ erhalten wir darüber hinaus bei Gültigkeit der Alternative

$$\begin{aligned} \sqrt{\frac{m+n}{mn}} T_{m,n}^{KS} &\xrightarrow{P_\vartheta} \sup_{t \geq 0} (\mathcal{L}_{\text{Exp}(1+\vartheta)}(t) - \mathcal{L}_{\text{Exp}(1)}(t)) \\ &= \sup_{t \geq 0} \left(\frac{1+\vartheta}{(1+\vartheta)+t} - \frac{1}{1+t} \right), \quad m, n \rightarrow \infty. \end{aligned}$$

Für $f(t) := \frac{1+\vartheta}{1+\vartheta+t} - \frac{1}{1+t}$, $t \geq 0$, folgt

$$\frac{d}{dt} f(t) = \frac{(1+\vartheta-t^2)\vartheta}{(1+\vartheta+t)^2(1+t)^2},$$

so dass $f(t)$ für $t = \sqrt{1+\vartheta}$ auf $(0, \infty)$ maximal wird. Insgesamt erhalten wir so für alle $\vartheta \in \Theta_1$

$$\sqrt{\frac{mn}{m+n}} T_{m,n}^{KS} \xrightarrow{P_\vartheta} b_1(\vartheta), \quad m, n \rightarrow \infty,$$

mit der Funktion $b_1(\vartheta) := f(\sqrt{1+\vartheta})$. Für den Nachweis von Bedingung (iii) verwenden wir das folgende klassische Resultat über die Verteilung der Suprema von Gaußprozessen.

3.2.2. Satz ([Adl90], S. 43). *Sei $\emptyset \neq T \subset \mathbb{R}^d$ und $(X_t, t \in T)$ ein zentrierter und separabler Gaußprozess mit fast sicher beschränkten Pfaden und $\sigma^2 := \sup_{t \in T} \text{Var}(X_t) > 0$. Dann gilt*

$$\lim_{x \rightarrow \infty} -2 \frac{\log P(\sup_{t \in T} X_t > x)}{x^2} = \frac{1}{\sigma^2}.$$

□

Nach Satz 2.2.5 hat $T_{m,n}^{KS}$ bei Gültigkeit der Hypothese die Grenzverteilung $\sup_{t \geq 0} \mathbb{L}(t)$, wobei der zentrierte Gaußprozess $(\mathbb{L}(t), t \geq 0)$ fast sicher beschränkte Pfade und die Varianzfunktion

$$\rho_{\text{Exp}(1)}(t, t) = \mathcal{L}_{\text{Exp}(1)}(2t) - (\mathcal{L}_{\text{Exp}(1)}(t))^2 = \frac{1}{1+2t} - \frac{1}{(1+t)^2}$$

hat. Es folgt

$$\tau := \operatorname{argmax}_{t \geq 0} \rho_{\text{Exp}(1)}(t, t) = \frac{1}{2} + \frac{1}{2}\sqrt{5}$$

und daher für σ aus dem letzten Satz

$$\sigma^2 = \rho_{\text{Exp}(1)}(\tau) = \frac{1}{1+2\tau} - \frac{1}{(1+\tau)^2} \approx 0.09017.$$

Somit ist $(T_{m,n}^{KS})_{m,n \in \mathbb{N}}$ eine Standardfolge. Mit $a_1 := \sigma^{-2}$ ist dann $c_1(\vartheta) := a_1 [b_1(\vartheta)]^2$ die approximative Bahadursteigung von $T_{m,n}^{KS}$ für die Hypothese H .

Wir wollen die approximative Bahadursteigung von $T_{m,n}^{KS}$ mit der approximativen Bahadursteigung von anderen Tests für das beschriebene Testproblem vergleichen. Seien $U := X_1 + \dots + X_m$, $V := Y_1 + \dots + Y_n$ und

$$T_{m,n}^{(1)} := \sqrt{\frac{mn}{m+n}} \left(\frac{\frac{1}{m}U}{\frac{1}{n}V} - 1 \right), \quad m, n \in \mathbb{N}.$$

Dann ist der Test

$$\phi(X_1, \dots, X_m, Y_1, \dots, Y_n) = \begin{cases} 1, & T_{m,n}^{(1)} > c_{m,n}, \\ 0, & T_{m,n}^{(1)} \leq c_{m,n}, \end{cases}$$

mit geeignet gewählten Konstanten $c_{m,n} > 0$, $m, n \in \mathbb{N}$, der gleichmäßig beste unverfälschte Test für das allgemeinere Testproblem

$$H^* : (F, G) \in \Theta_0 \quad \text{vs.} \quad K^* : (F, G) \in \Theta_1,$$

wobei $\Theta_0 = \{(F, G) : F = \text{Exp}(\lambda), G = \text{Exp}(\vartheta), \lambda \leq \vartheta\}$ und $\Theta_1 = \{(F, G) : F = \text{Exp}(\lambda), G = \text{Exp}(\vartheta), \lambda > \vartheta\}$ ist (siehe [Leh86], Section 3, 14.6. Dort wird ein äquivalenter Test für den Fall vorgeschlagen, dass zusätzlich Lageparameter zugelassen sind. Die Konstanten $c_{m,n}$ ergeben sich dann über die Quantile von $F_{2m,2n}$ - der F -Verteilung mit $2m$ Freiheitsgraden im Zähler und $2n$ Freiheitsgraden im Nenner. Im Fall $F = G$ ist $F_{2m,2n}$ die Verteilung des Quotienten $(\frac{1}{m}U) / (\frac{1}{n}V)$).

Wir verwenden diesen Test als Test für das eingangs beschriebene Testproblem. Es folgt dann bei Gültigkeit der Hypothese H

$$\begin{aligned} \sqrt{\frac{mn}{m+n}} \left(\frac{\frac{1}{m}U}{\frac{1}{n}V} - 1 \right) &= \frac{\sqrt{\frac{n}{m+n}} \sqrt{m} \left(\frac{1}{m}U - 1 \right) - \sqrt{\frac{m}{m+n}} \sqrt{n} \left(\frac{1}{n}V - 1 \right)}{\frac{1}{n}V} \\ &\xrightarrow{v} \frac{N(0,1)}{1} = N(0,1), \quad m, n \rightarrow \infty, \end{aligned}$$

also $T_{m,n}^{(1)} \xrightarrow{v} N(0,1)$ für $m, n \rightarrow \infty$. Außerdem folgt mit dem starken Gesetz der großen Zahlen bei Gültigkeit der Alternative K für jedes $\vartheta \in (0, \infty)$

$$\sqrt{\frac{m+n}{mn}} T_{m,n}^{(1)} \xrightarrow{f.s} \frac{1}{1+\vartheta} - 1 = \vartheta, \quad m, n \rightarrow \infty.$$

Setzen wir also $b_2(\vartheta) := \vartheta, \theta \in (0, \infty)$ und

$$a_2 := \lim_{x \rightarrow \infty} \frac{-2 \log \left(P \left(T_{m,n}^{(1)} > x \right) \right)}{x^2} = 1,$$

so ergibt sich $c_2(\vartheta) := a_2 (b_2(\vartheta))^2 = \vartheta^2$ als approximative Bahadursteigung für $T_{m,n}^{(1)}$. Wir definieren die relative approximative Bahadur-Effizienz der Testgröße $T_{m,n}^{KS}$ relativ zur Testgröße $T_{m,n}^{(1)}$ als

$$\epsilon_{\text{Rel}}(T_{m,n}^{KS}, T_{m,n}^{(1)})(\vartheta) := \frac{c_1(\vartheta)}{c_2(\vartheta)} = \frac{a_1 [b_1(\vartheta)]^2}{\vartheta^2} = \frac{a_1 \left(\frac{\vartheta \sqrt{1+\vartheta}}{(1+\vartheta + \sqrt{1+\vartheta})(1+\sqrt{1+\vartheta})} \right)^2}{\vartheta^2}.$$

Die folgende Tabelle zeigt die relative Effizienz $\epsilon_{\text{Rel}}(T_{m,n}^{KS}, T_{m,n}^{(1)})$ für ausgewählte Werte von ϑ . Dabei wird

$$\epsilon_{\text{Rel}}(T_{m,n}^{KS}, T_{m,n}^{(1)})(0) := \lim_{\vartheta \rightarrow 0} \epsilon_{\text{Rel}}(T_{m,n}^{KS}, T_{m,n}^{(1)})(\vartheta)$$

gesetzt.

ϑ	0	1	2	5	10	∞
$\epsilon_{\text{Rel}}(T_{m,n}^{KS}, T_{m,n}^{(1)})(\vartheta)$	$\frac{11}{32} + \frac{5}{32}\sqrt{5}$	0.326	0.199	0.078	0.031	0

Abbildung 3.1 zeigt darüber hinaus die Entwicklung von $\epsilon_{\text{Rel}}(T_{m,n}^{KS}, T_{m,n}^{(1)})(\vartheta)$ für $\vartheta \in [0, 10]$. Erwartungsgemäß schneidet der gleichmäßig beste unverfälschte Test bezüglich der approximativen Bahadursteigungen besser ab als $T_{m,n}^{KS}$.

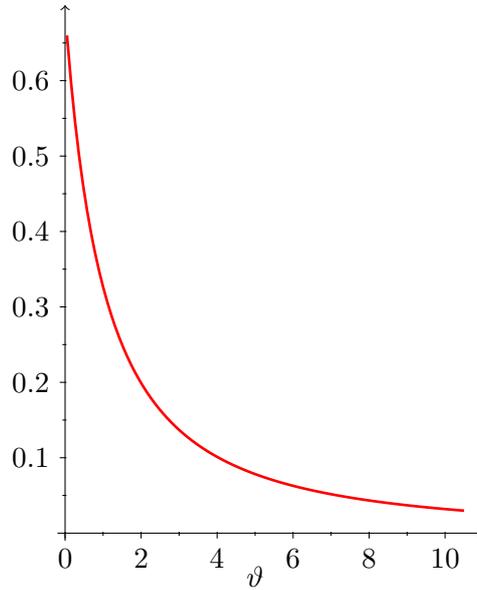


Abbildung 3.1: Relative approximative Bahadur-Effizienz $\epsilon_{\text{Rel}}(T_{m,n}^{KS}, T_{m,n}^{(1)})(\vartheta)$.

Wir wollen eine weitere Vergleichs-Testgröße heranziehen, und betrachten zu diesem Zweck die klassische einseitige Kolmogorov-Smirnov-Testgröße

$$T_{m,n}^{(2)} = \sqrt{\frac{mn}{m+n}} \sup_{t \geq 0} (G_n(t) - F_m(t)).$$

Bekanntlich gilt bei Gültigkeit der Hypothese

$$T_{m,n}^{(2)} \xrightarrow{v} \sup_{t \geq 0} \mathbb{L}(t), \quad m, n \rightarrow \infty,$$

in $l^\infty(\mathbb{R}_{\geq 0})$, mit dem zentrierten Gaußprozess \mathbb{L} mit der Kovarianzfunktion

$$\rho(s, t) = F(\min(s, t)) - F(s)F(t), \quad s, t \geq 0.$$

Alternativ ergibt sich diese Grenzwertaussage als Spezialfall der im Abschnitt 5.1 präsentierten Ergebnisse beim Test auf die sogenannte Konkordanz-Ordnung. Bei Gültigkeit der Alternative erhalten wir für jedes $\vartheta \in (0, \infty)$ dann

$$\sqrt{\frac{m+n}{mn}} T_{m,n}^{(2)} \xrightarrow{P_\vartheta} \sup_{t \geq 0} (\exp(-t) - \exp(-(1+\vartheta)t)), \quad m, n \rightarrow \infty.$$

Mit $g(t) = \exp(-t) - \exp(-(1+\vartheta)t)$ folgt

$$\frac{d}{dt} g(t) = 0 \quad \Leftrightarrow \quad t = \frac{\ln(1+\vartheta)}{\vartheta}, \quad \vartheta \in (0, \infty),$$

und daher

$$\begin{aligned} b_3(\vartheta) &:= \sup_{t \geq 0} g(t) = g\left(\frac{\ln(1+\vartheta)}{\vartheta}\right) \\ &= (1+\vartheta)^{-\frac{1}{\vartheta}} (1 - (1+\vartheta)^{-1}), \quad \vartheta \in (0, \infty). \end{aligned}$$

Aus

$$\text{Var}(\mathbb{L}(t)) = F(\min(t, t)) - F(t)F(t) = \exp(-t) - \exp(-2t)$$

für alle $t \geq 0$ folgt $\sup_{t \geq 0} \text{Var}(\mathbb{L}(t)) = \text{Var}(\mathbb{L}(\ln(2))) = \frac{1}{4}$ und mit Satz 3.2.2 schließlich

$$a_3 := \lim_{x \rightarrow \infty} \frac{-2 \log \left(P \left(T_{m,n}^{(2)} > x \right) \right)}{x^2} = 4.$$

Wir erhalten, dass $(T_{m,n}^{(2)})_{m,n \in \mathbb{N}}$ eine Standardfolge ist und somit

$$c_3(\vartheta) = a_3 (b_3(\vartheta))^2 = 4 \left((1+\vartheta)^{-\frac{1}{\vartheta}} (1 - (1+\vartheta)^{-1}) \right)^2$$

als die approximative Bahadursteigung für $(T_{m,n}^{(2)})_{m,n \in \mathbb{N}}$. Die relative approximative Bahadur-Effizienz der Testgröße $T_{m,n}^{(2)}$ relativ zu $T_{m,n}^{KS}$ ist dann der Quotient

$$\epsilon_{\text{Rel}}(T_{m,n}^{(2)}, T_{m,n}^{KS})(\vartheta) = \frac{c_3(\vartheta)}{c_1(\vartheta)} = \frac{4 \left((1+\vartheta)^{-\frac{1}{\vartheta}} (1 - (1+\vartheta)^{-1}) \right)^2}{\left(\frac{\vartheta \sqrt{1+\vartheta}}{(1+\vartheta + \sqrt{1+\vartheta})(1 + \sqrt{1+\vartheta})} \right)^2 a_1}.$$

Für einige Werte von ϑ ist die relative Effizienz in der folgenden Tabelle vertafelt.

ϑ	0	1	2	5	10	∞
$\epsilon_{\text{Rel}}(T_{m,n}^{(2)}, T_{m,n}^{KS})(\vartheta)$	0.781	0.7675	0.744	0.692	0.640	0.360

Der auf den Laplacetransformierten basierende Kolmogorov-Smirnov-Test $T_{m,n}^{KS}$ weist also für die untersuchte Teilhypothese eine bessere approximative Bahadur-Effizienz auf als der klassische einseitige Kolmogorov-Smirnov-Test.

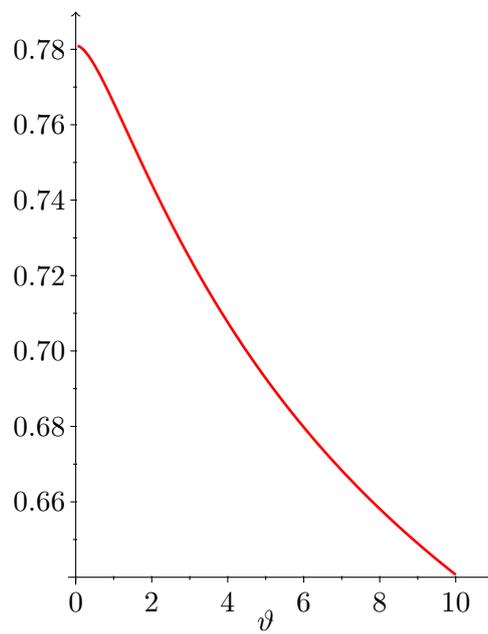


Abbildung 3.2: Relative approximative Bahadur-Effizienz $\epsilon_{\text{Rel}}(T_{m,n}^{(2)}, T_{m,n}^{KS})(\vartheta)$.

Kapitel 4

Die Konkordanz-Ordnung

4.1 Grundbegriffe

Bei der mathematischen Modellierung von Risiken in höheren Dimensionen spielt häufig der Grad der stochastischen Abhängigkeit der betrachteten Zufallsvariablen eine zentrale Rolle. Die in diesem Abschnitt zu definierende Konkordanz-Ordnung liefert eine Ordnung von Zufallsvektoren, die einen Vergleich der Abhängigkeiten der Komponenten dieser Zufallsvektoren ermöglicht.

Es sei $\mathbf{X} = (X_1, \dots, X_d)$ ein d -dimensionaler Zufallsvektor mit reellen Komponenten und der Verteilungsfunktion $F_{\mathbf{X}}$, sowie der *Überlebens-Funktion*

$$\bar{F}_{\mathbf{X}}(\mathbf{t}) = P(X_1 > t_1, \dots, X_d > t_d), \quad \mathbf{t} = (t_1, \dots, t_d) \in \mathbb{R}^d.$$

Betrachtet man die gewöhnliche stochastische Ordnung \leq_{st} von Zufallsvariablen im Ein-dimensionalen, so ist $X \leq_{st} Y$ wegen $F_X(t) = 1 - \bar{F}_X(t)$, $t \in \mathbb{R}$, äquivalent zu $\bar{F}_X(t) \leq \bar{F}_Y(t)$, $t \in \mathbb{R}$. Für $d \geq 2$ besteht dieser Zusammenhang offenbar nicht mehr. Vielmehr gilt mit $A_j := \{X_j > t_j\}$ für $j = 1, \dots, d$, und $F_{\mathbf{X}}^{(i_1, \dots, i_k)}$ als Verteilungsfunktion von $(X_{i_1}, \dots, X_{i_k})$ mit paarweise verschiedenen $i_1, \dots, i_k \in \{1, \dots, d\}$, $k \leq d$, dass

$$\begin{aligned} \bar{F}_{\mathbf{X}}(t_1, \dots, t_d) &= P(A_1 \cap \dots \cap A_d) = 1 - P(A_1^c \cup \dots \cup A_d^c) \\ &= 1 - \sum_{k=1}^d (-1)^{k-1} \sum_{1 \leq i_1 < \dots < i_k \leq d} P(A_{i_1}^c \cap \dots \cap A_{i_k}^c) \\ &= 1 - \sum_{k=1}^d (-1)^{k-1} \sum_{1 \leq i_1 < \dots < i_k \leq d} F_{\mathbf{X}}^{(i_1, \dots, i_k)}(t_1, \dots, t_d) \end{aligned} \quad (4.1.1)$$

ist. Diese Überlegungen führen auf die folgende Definition.

4.1.1. Definition. Es seien $\mathbf{X} = (X_1, \dots, X_d)$ und $\mathbf{Y} = (Y_1, \dots, Y_d)$ Zufallsvektoren mit Werten in \mathbb{R}^d , $d \geq 2$, und den Verteilungsfunktionen F und G . Dann heißt \mathbf{X} kleiner oder gleich \mathbf{Y} in der *Konkordanz-Ordnung*, in Zeichen $\mathbf{X} \leq_c \mathbf{Y}$, falls jeweils für alle $\mathbf{t} \in \mathbb{R}^d$

$$F(\mathbf{t}) \leq G(\mathbf{t}) \quad \text{und} \quad \bar{F}(\mathbf{t}) \leq \bar{G}(\mathbf{t})$$

zutritt.

Anders als die zuvor betrachtete Laplace-Ordnung, die sich in natürlicher Weise als Verallgemeinerung einer univariaten stochastischen Ordnung ergibt, macht die getroffene Definition der Konkordanz-Ordnung erst in einer multivariaten Situation Sinn.

Es gilt $\mathbf{X} \leq_c \mathbf{Y}$ genau dann, wenn $Ef(\mathbf{X}) \leq Ef(\mathbf{Y})$ für alle $f \in \{\mathbf{1}_{(-\infty, \mathbf{t}]}, \mathbf{t} \in \mathbb{R}^d\} \cup \{\mathbf{1}_{(\mathbf{t}, \infty)}, \mathbf{t} \in \mathbb{R}^d\}$ ist, wobei die Indikatoren wieder in der Weise $\mathbf{1}_{(-\infty, \mathbf{t}]} := \mathbf{1}_{(-\infty, t_1] \times \dots \times (-\infty, t_d]}$ zu lesen sind. Wir schreiben synonym $F \leq_c G$ statt $\mathbf{X} \leq_c \mathbf{Y}$, falls \mathbf{X} bzw. \mathbf{Y} die Verteilungsfunktionen F bzw. G haben. Beziehungen zu verschiedenen anderen Ordnungen werden in Abschnitt 4.2 aufgezeigt. Ein elementares Beispiel zweier bezüglich \leq_c geordneter Verteilungen geben wir sogleich an. Weitere Beispiele und eine Klassifikation von *elliptisch symmetrischen Verteilungen* mittels der Konkordanz-Ordnung werden in Abschnitt 4.3 betrachtet.

4.1.2. Beispiel ([Joe90]).

Es seien $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^d$ und $\mathbf{x}_3 := \mathbf{x}_1 \wedge \mathbf{x}_2$ bzw. $\mathbf{x}_4 := \mathbf{x}_1 \vee \mathbf{x}_2$ das (komponentenweise gebildete) Minimum bzw. Maximum von $\mathbf{x}_1, \mathbf{x}_2$. Es sei F die Verteilungsfunktion des diskreten Wahrscheinlichkeitsmaßes

$$P = \sum_{i=1}^4 p_i \delta_{\mathbf{x}_i},$$

mit $p_i > 0$, $\sum_{i=1}^4 p_i = 1$. Für ein $0 < \varepsilon < \min(p_1, p_2, 1 - p_3, 1 - p_4)$ sei G die Verteilungsfunktion von

$$Q = (p_1 - \varepsilon)\delta_{\mathbf{x}_1} + (p_2 - \varepsilon)\delta_{\mathbf{x}_2} + (p_3 + \varepsilon)\delta_{\mathbf{x}_3} + (p_4 + \varepsilon)\delta_{\mathbf{x}_4}.$$

Es folgt $F(\mathbf{t}) \leq G(\mathbf{t})$ und $\bar{F}(\mathbf{t}) \leq \bar{G}(\mathbf{t})$ jeweils für alle $\mathbf{t} \in \mathbb{R}^d$, so dass $F \leq_c G$ gilt.

Abbildung 4.1 verdeutlicht diesen Sachverhalt am Beispiel der Verteilungsfunktionen im Fall $d = 2$ und für Punkte $\mathbf{x}_1 = (x_{11}, x_{12})'$ und $\mathbf{x}_2 = (x_{21}, x_{22})'$ mit $x_{11} < x_{21}$ und $x_{12} > x_{22}$. In diesem Fall nimmt eine Zufallsvariable mit der Verteilungsfunktion G die Werte \mathbf{x}_3 und \mathbf{x}_4 jeweils mit größerer Wahrscheinlichkeit an, als eine Zufallsvariable mit

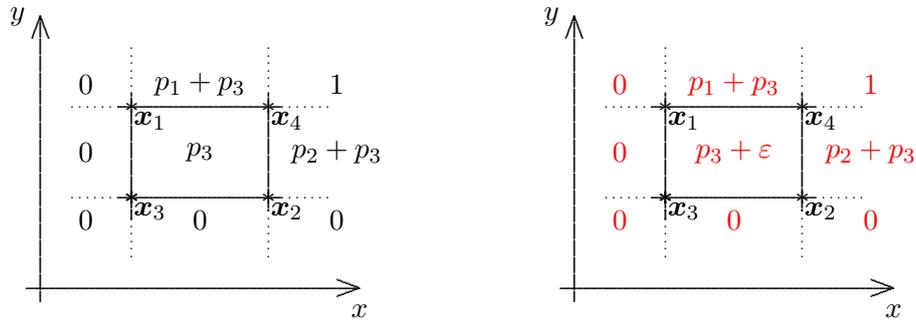


Abbildung 4.1: Werte von $F(x, y)$ (links) und $G(x, y)$ (rechts) für $d = 2$.

der Verteilungsfunktion F . Die Abbildung zeigt die Werte, die die Verteilungsfunktionen F und G in den entsprechend markierten Bereichen annehmen. Beispielsweise ist die Verteilungsfunktion F konstant auf dem Rechteck

$$R := \{(x, y) \in \mathbb{R}^2 : x_{11} \leq x \leq x_{21}, x_{22} \leq y \leq x_{12}\} \setminus \{x_4\}$$

mit dem Funktionswert p_3 . Das Verhalten der Überlebensfunktionen \bar{F} und \bar{G} macht man sich auf analoge Weise klar.

4.2 Eigenschaften und Beziehungen zu anderen Ordnungen

Wir wollen hier einige wichtige Eigenschaften der Konkordanz-Ordnung aufführen, die die einleitend bereits angedeutete Charakterisierung von \leq_c als *Multivariate Abhängigkeitsordnung* (eingeführt in [KS87]) begründen. Es zeigt sich, dass der bivariate Fall dabei von gesondertem Interesse ist.

Eine erste offensichtliche Beobachtung ist, dass im Falle von $F(\mathbf{t}) \leq G(\mathbf{t}), \mathbf{t} \in \mathbb{R}^d$, für die eindimensionalen Randverteilungsfunktionen auch $F^{(i)}(t) \leq G^{(i)}(t), t \in \mathbb{R}$, für jedes $i = 1, \dots, d$ gilt. Analog folgt aus $\bar{F}(\mathbf{t}) \leq \bar{G}(\mathbf{t}), \mathbf{t} \in \mathbb{R}^d$, dass $F^{(i)}(t) \geq G^{(i)}(t), t \in \mathbb{R}$, für jedes $i = 1, \dots, d$ ist. Zusammengefasst ergeben diese Überlegungen das nachstehende Resultat.

4.2.1. Lemma. *Es sei $\mathbf{X} \leq_c \mathbf{Y}$. Dann gilt $X_i \stackrel{v}{=} Y_i, i = 1, \dots, d$.*

□

Die folgenden elementaren Eigenschaften ergeben sich durch ähnliche Überlegungen oder direkt anhand von Definition 4.1.1.

4.2.2. Lemma.

- a) Die Konkordanz-Ordnung \leq_c ist reflexiv, transitiv und antisymmetrisch.
 b) Die Konkordanz-Ordnung \leq_c ist invariant unter wachsenden Transformationen, es gilt also

$$(X_1, \dots, X_d) \leq_c (Y_1, \dots, Y_d) \Rightarrow (f_1(X_1), \dots, f_d(X_d)) \leq_c (f_1(Y_1), \dots, f_d(Y_d))$$

für alle reellen und monoton wachsenden Funktionen f_1, \dots, f_d .

- c) Die Konkordanz-Ordnung \leq_c ist abgeschlossen unter Marginalbildung, es gilt also

$$\mathbf{X} \leq_c \mathbf{Y} \Rightarrow (X_{i_1}, \dots, X_{i_k}) \leq_c (Y_{i_1}, \dots, Y_{i_k})$$

für alle paarweise verschieden $i_1, \dots, i_k \in \{1, \dots, d\}$, $k = 2, \dots, d$.

- d) Seien $\mathbf{X} \leq_c \mathbf{Y}$, $\mathbf{U} \leq_c \mathbf{V}$, \mathbf{X}, \mathbf{U} unabhängig, sowie \mathbf{Y}, \mathbf{V} unabhängig. Dann ist $\mathbf{X} + \mathbf{U} \leq_c \mathbf{Y} + \mathbf{V}$.

Beweis. Die Aussagen a) - c) sind offensichtlich. Eigenschaft d) folgt aus

$$\begin{aligned} P(\mathbf{X} + \mathbf{U} \leq \mathbf{t}) &= \int_{\mathbb{R}^d} P(\mathbf{X} \leq \mathbf{t} - \mathbf{u}) dP^{\mathbf{U}}(\mathbf{u}) \leq \int_{\mathbb{R}^d} P(\mathbf{Y} \leq \mathbf{t} - \mathbf{u}) dP^{\mathbf{U}}(\mathbf{u}) \\ &= P(\mathbf{Y} + \mathbf{U} \leq \mathbf{t}) = \int_{\mathbb{R}^d} P(\mathbf{U} \leq \mathbf{t} - \mathbf{y}) dP^{\mathbf{Y}}(\mathbf{y}) \\ &\leq \int_{\mathbb{R}^d} P(\mathbf{V} \leq \mathbf{t} - \mathbf{y}) dP^{\mathbf{Y}}(\mathbf{y}) = P(\mathbf{Y} + \mathbf{V} \leq \mathbf{t}), \quad \mathbf{t} \in \mathbb{R}^d. \end{aligned}$$

Der Fall der Überlebensfunktion kann analog behandelt werden. □

Die folgende Charakterisierung der Konkordanz-Ordnung geht auf Bergmann ([Ber91]) zurück.

4.2.3. Satz ([Ber91], Prop. 3.1. und 3.2.). Es gilt $\mathbf{X} \leq_c \mathbf{Y}$ genau dann, wenn

$$E \left(\prod_{k=1}^d f_k(X_k) \right) \leq E \left(\prod_{k=1}^d f_k(Y_k) \right)$$

für sämtliche monoton wachsenden reellen Funktionen f_1, \dots, f_d , $f_k \geq 0$, $k = 1, \dots, d$, und sämtliche monoton fallenden reellen Funktionen f_1, \dots, f_d , $f_k \geq 0$, $k = 1, \dots, d$, gilt.

Ist $d = 2$, so folgt aus $\mathbf{X} \leq_c \mathbf{Y}$, dass $E(f_1(X_1)f_2(X_2)) \leq E(f_1(Y_1)f_2(Y_2))$ für alle monoton wachsenden reellen Funktionen f_1, f_2 ist.

Aus dem letzten Satz ergeben sich erste Konsequenzen im Hinblick auf den Vergleich der Kovarianzen der Komponenten von Zufallsvektoren. Im folgenden Korollar formulieren wir die zu Beginn dieses Abschnitts angedeutete Monotonieeigenschaft der Kovarianzen.

4.2.4. Korollar. Es sei $\mathfrak{M} := \{f : \mathbb{R} \rightarrow \mathbb{R}; f \text{ monoton wachsend}\}$ die Menge der monoton wachsenden reellen Funktionen. Dann gilt

- a) $\mathbf{X} \leq_c \mathbf{Y} \implies \text{Cov}(f(X_i), g(X_j)) \leq \text{Cov}(f(Y_i), g(Y_j))$, für alle $i \neq j$ und alle $f, g \in \mathfrak{M}$ für die die Kovarianzen existieren.
- b) Ist $d = 2$ und haben \mathbf{X}, \mathbf{Y} die gleichen Randverteilungen, so ist die Implikation in a) sogar eine Äquivalenz, es gilt dann also

$$\mathbf{X} \leq_c \mathbf{Y} \iff \text{Cov}(f(X_1), g(X_2)) \leq \text{Cov}(f(Y_1), g(Y_2)) \quad \forall f, g \in \mathfrak{M}$$

für die die Kovarianzen existieren.

Beweis. a) Aus $\mathbf{X} \leq_c \mathbf{Y}$ folgt wegen Lemma 4.2.2 b) und c) zunächst $(X_i, X_j) \leq_c (Y_i, Y_j)$ für alle $i \neq j$. Mit Satz 4.2.3 und der Tatsache, dass \mathbf{X} und \mathbf{Y} dann die gleichen Randverteilung haben, folgt die Behauptung.

b) Im Falle der Gleichheit der Randverteilungen ist wegen (4.1.1) $\mathbf{X} \leq_c \mathbf{Y}$ äquivalent zu $F(t_1, t_2) \leq G(t_1, t_2)$ für alle $t_1, t_2 \in \mathbb{R}$. Die Beziehung $F(t_1, t_2) \leq G(t_1, t_2)$ folgt dann aus der Eigenschaft

$$\forall f, g \in \mathfrak{M} : \quad \text{Cov}(f(X_1), g(X_2)) \leq \text{Cov}(f(Y_1), g(Y_2)),$$

durch Wahl von $f := \mathbf{1}_{(-\infty, t_1]}$ und $g := \mathbf{1}_{(-\infty, t_2]}$. □

Insbesondere folgt aus $\mathbf{X} \leq_c \mathbf{Y}$ also $\text{Cov}(X_i, X_j) \leq \text{Cov}(Y_i, Y_j) \quad \forall i \neq j$ falls die Kovarianzen existieren. Die Konkordanz-Ordnung \leq_c liefert also den eingangs behaupteten Vergleich der Abhängigkeiten der Komponenten von Zufallsvektoren mit gleichen Randverteilungen.

Im bivariaten Fall ist die Konkordanz-Ordnung unter einer Zusatzvoraussetzung darüber hinaus sogar die einzige stochastische Integralordnung zwischen Zufallsvektoren mit gleichen Randverteilungen die sowohl invariant unter monoton wachsenden Transformationen ist, als auch $\text{Cov}(X_i, X_j) \leq \text{Cov}(Y_i, Y_j) \quad \forall i \neq j$ impliziert (vergleiche [MS02], Theorem 3.8.3).

4.2.5. Bemerkung. Monotonien in der Form von Korollar 4.2.4 bleiben unter der Konkordanz-Ordnung auch für andere Korrelationsmaße erhalten. Die Gleichheit der Randverteilungen liefert zunächst, dass - Existenz der entsprechenden Größen vorausgesetzt - aus $X \leq_c Y$ bereits $\rho_{X_i, X_j} \leq \rho_{Y_i, Y_j}$ für alle $i \neq j$ folgt, wobei

$$\rho_{X_i, X_j} = \frac{\text{Cov}(X_i, X_j)}{\sqrt{\text{Var}(X_i)\text{Var}(X_j)}}$$

den Korrelationskoeffizienten von X_i und X_j und analog ρ_{Y_i, Y_j} den Korrelationskoeffizienten von Y_i und Y_j bezeichnet.

Wir werden diese Eigenschaften im Hinblick auf Korollar 4.2.4 b) später dazu verwenden können, um in einigen parametrischen Spezialfällen alternative Testgrößen genauer zu untersuchen. Weitere Korrelationsmaße für die sich ähnliche Resultate ergeben sind etwa Kendall's τ und Spearman's ρ (siehe [YO69]). Verallgemeinerungen dieses Sachverhalts auf multivariate Varianten von Kendall's τ und Spearman's ρ , und entsprechende Monotonieeigenschaften im Bezug auf die Konkordanz-Ordnung finden sich in [Joe90].

Auch im Fall der Konkordanz-Ordnung wollen wir Beziehungen zu einigen anderen gängigen multivariaten stochastischen Integralordnungen anführen. Satz 4.2.3 liefert dabei zunächst die Beziehung zur eingangs betrachteten Laplace-Ordnung:

4.2.6. Korollar. Seien \mathbf{X}, \mathbf{Y} komponentenweise nicht-negativ. Dann gilt $\mathbf{X} \leq_c \mathbf{Y} \implies \mathbf{Y} \leq_{\mathcal{L}} \mathbf{X}$.

Beweis. Für feste $t_i \geq 0, i = 1, \dots, d$, sind die Funktionen $f_{t_i} : x \mapsto \exp(-t_i x)$ nicht-negativ und monoton fallend. Satz 4.2.3 liefert die Behauptung. \square

Wie wir einleitend bereits festgehalten haben, unterscheidet sich die Definition der Konkordanz-Ordnung für $d \geq 2$ von der üblichen gewöhnlichen stochastischen Ordnung. Das folgende Beispiel zeigt, dass im Allgemeinen weder die Konkordanz-Ordnung die gewöhnliche stochastische Ordnung, noch umgekehrt die gewöhnliche stochastische Ordnung die Konkordanz-Ordnung impliziert.

4.2.7. Beispiel.

a) Seien $\mathbf{x}_1 = (1, 0), \mathbf{x}_2 = (0, 1), \mathbf{y}_1 = (0, 0), \mathbf{y}_2 = (1, 1)$ und \mathbf{X}, \mathbf{Y} Zufallsvektoren mit $P(\mathbf{X} = \mathbf{x}_1) = P(\mathbf{X} = \mathbf{x}_2) = P(\mathbf{Y} = \mathbf{y}_1) = P(\mathbf{Y} = \mathbf{y}_2) = \frac{1}{2}$. Dann gilt $\mathbf{X} \leq_c \mathbf{Y}$. Wir betrachten die Funktion $f : \mathbb{R}^2 \rightarrow \mathbb{R}, f(x, y) = \mathbf{1}_M(x, y)$, mit

$$M := \{(x, y) \in \mathbb{R}^2; x, y \geq 0\} \setminus \{(0, 0)\}.$$

Es ist f monoton wachsend in jeder Komponente (bei jeweils festgehaltener anderer Komponente) und außerdem beschränkt. Ferner ist $Ef(\mathbf{X}) = 1$, $Ef(\mathbf{Y}) = \frac{1}{2}$ und somit $\mathbf{X} \not\leq_{st} \mathbf{Y}$.

- b) Seien $\mathbf{x}_1 = (0, 0)$, $\mathbf{x}_2 = (1, 1)$, $\mathbf{y}_1 = (2, 2)$, $\mathbf{y}_2 = (3, 3)$ und \mathbf{X} , \mathbf{Y} Zufallsvektoren mit $P(\mathbf{X} = \mathbf{x}_1) = P(\mathbf{X} = \mathbf{x}_2) = P(\mathbf{Y} = \mathbf{y}_1) = P(\mathbf{Y} = \mathbf{y}_2) = \frac{1}{2}$. Dann gilt für alle monotonen und beschränkten Funktionen $Ef(\mathbf{X}) = \frac{1}{2}(f((0, 0)) + f(1, 1)) \leq \frac{1}{2}(f((2, 2)) + f((3, 3))) = Ef(\mathbf{Y})$ und daher $\mathbf{X} \leq_{st} \mathbf{Y}$. Es ist aber z.B. $P(\mathbf{X} \leq \mathbf{x}_1) = \frac{1}{2} > P(\mathbf{Y} \leq \mathbf{x}_1) = 0$ also insbesondere $\mathbf{X} \not\leq_c \mathbf{Y}$. Ebenso erhält man mit diesen Überlegungen $\mathbf{Y} \not\leq_c \mathbf{X}$.

Neben der erwähnten Beziehung zur Laplace-Ordnung ist besonders der enge Zusammenhang der Konkordanz-Ordnung zur supermodularen Ordnung \leq_{sm} interessant, die vor allem bei der Bewertung und dem Vergleich von Risiken und in der Nutzentheorie eine Rolle spielt (siehe [Den05]).

4.2.8. Satz. *Im bivariaten Fall ($d = 2$) stimmen die supermodulare Ordnung und die Konkordanz-Ordnung überein. Für $d > 2$ gilt*

$$\mathbf{X} \leq_{sm} \mathbf{Y} \implies \mathbf{X} \leq_c \mathbf{Y}$$

Beweis. Für den ersten Teil der Behauptung verweisen wir auf [MS02] Theorem 3.8.2. Der zweite Teil folgt aus der Charakterisierung der Konkordanz-Ordnung als Integralordnung und der Tatsache, dass die Funktionen $\mathbf{1}_{(-\infty, \mathbf{t}]}$ und $\mathbf{1}_{(\mathbf{t}, \infty)}$ offensichtlich für jedes feste $\mathbf{t} \in \mathbb{R}^d$ supermodular sind. \square

4.2.9. Bemerkung.

- a) Für $d \geq 3$ stimmen \leq_{sm} und \leq_c nicht mehr überein, d.h., im Allgemeinen folgt aus $\mathbf{X} \leq_c \mathbf{Y}$ nicht $\mathbf{X} \leq_{sm} \mathbf{Y}$. Siehe [Sca00] für ein Beispiel.
- b) Ein zentraler Grund für die Bedeutung der supermodularen Ordnung \leq_{sm} im Risikomanagement ist die Implikation

$$\mathbf{X} \leq_{sm} \mathbf{Y} \implies \sum_{i=1}^d X_i \leq_{icx} \sum_{i=1}^d Y_i. \quad (4.2.1)$$

Im Kontext der Nutzentheorie lässt sich auf diese Weise das Entscheidungsverhalten *risikoaverser* Entscheidungsträger beschreiben (vergleiche etwa [DG96] für Details).

Ein statistisches Verfahren zur direkten Überprüfung der Hypothese $\sum_{i=1}^d X_i \leq_{icx} \sum_{i=1}^d Y_i$ für nicht-negative Zufallsvariablen findet sich in [BG09].

- c) Im bivariaten Fall gilt für die Konkordanz-Ordnung wegen Satz 4.2.8 auch die Implikation in (4.2.1). Für $d \geq 3$ folgt aus $\mathbf{X} \leq_c \mathbf{Y}$ im Allgemeinen aber nicht $\sum_{i=1}^d X_i \leq_{icx} \sum_{i=1}^d Y_i$. Siehe dazu [Mül97].

4.3 Elliptisch symmetrische Verteilungen

Für Verteilungen, die gewissen Symmetrien unterliegen, lässt sich die Konkordanz-Ordnung häufig durch die Ordnung bestimmter Parameter beschreiben. Die folgende Aussage ist auch als *Slepiansche Ungleichung* bekannt.

4.3.1. Lemma ([Ton80], S. 8). *Es seien $\mathbf{X} \sim N_d(\mu_1, \Sigma_1)$ und $\mathbf{Y} \sim N_d(\mu_2, \Sigma_2)$ d -dimensional normalverteilt mit Erwartungswertvektoren $\mu_1, \mu_2 \in \mathbb{R}^d$ und $\Sigma_1 = (\sigma_{ij}^1)_{1 \leq i, j \leq d}$, $\Sigma_2 = (\sigma_{ij}^2)_{1 \leq i, j \leq d}$ als Kovarianzmatrizen. Sei $\mu_1 = \mu_2$ und $\sigma_{ii}^1 = \sigma_{ii}^2$, $i = 1, \dots, d$. Ist $\sigma_{ij}^1 \leq \sigma_{ij}^2$ für alle $1 \leq i, j \leq d$, so gilt $\mathbf{X} \leq_c \mathbf{Y}$.*

Lemma 4.3.1 ergibt sich als Spezialfall eines ähnlichen Resultates für *elliptisch symmetrischen Verteilungen* (zuerst eingeführt in [Kel70]), welches wir in Satz 4.3.3 formulieren werden. Zur Bedeutung von elliptisch symmetrischen Verteilungen im Risikomanagement siehe z.B. [EMS02].

4.3.2. Definition. Sei \mathbf{X} ein d -dimensionaler Zufallsvektor mit Werten in \mathbb{R}^d , $\mu \in \mathbb{R}^d$, $\Sigma \in \mathbb{R}^{d \times d}$ symmetrisch und positiv definit und $\phi : [0, \infty) \rightarrow \mathbb{R}$ eine Funktion mit der Eigenschaft, dass die Funktion $\psi : (t_1, \dots, t_d) \mapsto \phi\left(\sum_{i=1}^d t_i^2\right)$ die Fouriertransformierte eines d -dimensionalen Zufallsvektors ist. Es heißt \mathbf{X} dann *d -dimensional elliptisch symmetrisch verteilt* mit den Parametern μ , Σ und dem *charakteristischen Generator* ϕ , in Zeichen $\mathbf{X} \sim E_d(\mu, \Sigma, \phi)$, falls \mathbf{X} die Fouriertransformierte

$$\varphi_{\mathbf{X}}(\mathbf{t}) = e^{it'\mu} \phi(\mathbf{t}'\Sigma\mathbf{t}), \quad \mathbf{t} \in \mathbb{R}^d, \quad (4.3.1)$$

hat.

Elliptisch symmetrische Verteilungen ergeben sich als lineare Transformationen sphärisch symmetrischer Verteilung: Sei dazu \mathbf{Y} sphärisch symmetrisch, d.h., es gelte $\mathbf{Y} \stackrel{v}{=} P\mathbf{Y}$ für alle orthogonalen Matrizen P . Für $\mathbf{t}_1, \mathbf{t}_2 \in \mathbb{R}^d$, $\mathbf{t}_1 \neq \mathbf{t}_2$, mit $\|\mathbf{t}_1\| = \|\mathbf{t}_2\|$, existiert dann eine orthogonale Matrix Q mit $Q\mathbf{t}_1 = \mathbf{t}_2$. Es folgt für die Fouriertransformierte $\psi_{\mathbf{Y}}$ von \mathbf{Y} , dass

$$\psi_{\mathbf{Y}}(\mathbf{t}_2) = E\left(e^{it_2'\mathbf{Y}}\right) = E\left(e^{i(Q\mathbf{t}_1)'\mathbf{Y}}\right) = E\left(e^{it_1'Q'\mathbf{Y}}\right) = E\left(e^{it_1'\mathbf{Y}}\right) = \psi_{\mathbf{Y}}(\mathbf{t}_1).$$

Somit hängt die Funktion $\psi_{\mathbf{Y}}(\mathbf{t})$ nur von der Norm von \mathbf{t} ab. Wegen $\mathbf{Y} \stackrel{v}{=} -\mathbf{Y}$ ist $\psi_{\mathbf{Y}}$ außerdem reellwertig, so dass $\psi_{\mathbf{Y}}$ von der Form $\psi_{\mathbf{Y}}(\mathbf{t}) = \phi(\|\mathbf{t}\|^2)$ mit einer wie in Definition 4.3.2 betrachteten Funktion ϕ ist.

Ist dann $\mu \in \mathbb{R}^d$ und $\Sigma = AA'$ die Cholesky-Zerlegung der positiv definiten Matrix Σ , so hat der transformierte Zufallsvektor

$$\mathbf{X} = \mu + A\mathbf{Y} \quad (4.3.2)$$

die Fouriertransformierte

$$\varphi_{\mathbf{X}}(\mathbf{t}) = e^{it'\mu} E\left(e^{it'AY}\right) = e^{it'\mu} \phi(\mathbf{t}'AA'\mathbf{t}) = e^{it'\mu} \phi(\mathbf{t}'\Sigma\mathbf{t}).$$

Also ist $\mathbf{X} \sim E_d(\mu, \Sigma, \phi)$. Umgekehrt kann jede elliptisch symmetrische Verteilung als Verteilung eines transformierten Zufallsvektors mit sphärisch symmetrischer Verteilung gemäß (4.3.2) erhalten werden, siehe dazu [FKN87], Kapitel 2.5.

Falls $\mathbf{X} \sim E_d(\mu, \Sigma, \phi)$ eine Dichte $f_{\mathbf{X}}$ besitzt, so ist diese von der Form

$$f_{\mathbf{X}}(\mathbf{t}) = c_d |\Sigma|^{-\frac{1}{2}} g((\mathbf{t} - \mu)' \Sigma^{-1} (\mathbf{t} - \mu)), \quad \mathbf{t} \in \mathbb{R}^d, \quad (4.3.3)$$

mit einer Funktion $g : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$ mit der Eigenschaft $\int_0^\infty y^{\frac{d}{2}-1} g(y) dy < \infty$ und einer Normierungskonstanten

$$c_d = \frac{\Gamma(d/2)}{\pi^{d/2} \int_0^\infty y^{d/2-1} g(y) dy}$$

(vergleiche [FA90], (iii)). Die Funktion g heißt auch *der Dichtegenerator* von \mathbf{X} .

Den Zusammenhang zur Konkordanz-Ordnung liefert der folgende Satz und das sich anschließende Korollar. Wir bezeichnen dabei erneut eine reellwertige Funktion $f : \mathbb{R}^d \rightarrow \mathbb{R}$, $(x_1, \dots, x_d) \mapsto f(x_1, \dots, x_d)$, als monoton wachsend in x_j , falls die Funktion $x_j \mapsto f(x_1, \dots, x_d)$ für festgehaltene $x_k, k \neq j$, monoton wachsend ist.

4.3.3. Satz ([Joe90], Theorem A.2). *Sei $\mathbf{Y} = (Y_1, \dots, Y_d)$ sphärisch symmetrisch und $\Sigma = (\sigma_{ij})_{1 \leq i, j \leq d}$ eine symmetrische und positiv definite Matrix mit $\sigma_{ii} = 1$ für $i = 1, \dots, d$ und der Cholesky-Zerlegung $\Sigma = AA'$. Sei $\mathbf{X} = (X_1, \dots, X_d)$ von der Form $\mathbf{X} = A\mathbf{Y}$. Dann ist die Funktion $P(X_1 \leq t_1, \dots, X_d \leq t_d)$ für jedes feste $\mathbf{t} \in \mathbb{R}^d$ monoton wachsend in jeder Komponente $\sigma_{ij}, i \neq j$.*

4.3.4. Korollar. Seien $\mathbf{X} \sim E_d(\mu, \Sigma_1, \phi)$ und $\mathbf{Y} \sim E_d(\mu, \Sigma_2, \phi)$ mit $\Sigma_1 = (\sigma_{ij}^1)_{1 \leq i, j \leq d}$, $\Sigma_2 = (\sigma_{ij}^2)_{1 \leq i, j \leq d}$, sowie $\sigma_{ii}^1 = \sigma_{ii}^2 := \sigma_{ii}$, $i = 1, \dots, d$, und $\sigma_{ij}^1 \leq \sigma_{ij}^2$ für alle $i \neq j$. Dann ist $\mathbf{X} \leq_c \mathbf{Y}$.

Beweis. Wir definieren die Diagonalmatrix

$$D := \begin{pmatrix} \sigma_{11}^{-\frac{1}{2}} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_{dd}^{-\frac{1}{2}} \end{pmatrix}$$

und die Zufallsvektoren $\hat{\mathbf{X}} = D(\mathbf{X} - \mu)$, $\hat{\mathbf{Y}} = D(\mathbf{Y} - \mu)$. Es gilt

$$\varphi_{\hat{\mathbf{X}}}(\mathbf{t}) = E\left(e^{i\mathbf{t}'D(\mathbf{X}-\mu)}\right) = e^{-i(D'\mathbf{t})'\mu}\varphi_{\mathbf{X}}(D'\mathbf{t}) = \phi(\mathbf{t}'\hat{\Sigma}_1\mathbf{t}), \quad \mathbf{t} \in \mathbb{R}^d,$$

mit $\hat{\Sigma}_1 = (\hat{\sigma}_{ij}^1)_{1 \leq i, j \leq d} = D\Sigma_1 D$. Es ist $\hat{\Sigma}_1$ positiv definit, so dass $\hat{\mathbf{X}} \sim E_d(0, \hat{\Sigma}_1, \phi)$ folgt. Analog erhält man $\hat{\mathbf{Y}} \sim E_d(0, \hat{\Sigma}_2, \phi)$ mit $\hat{\Sigma}_2 = (\hat{\sigma}_{ij}^2)_{1 \leq i, j \leq d} = D\Sigma_2 D$. Es ist $\hat{\sigma}_{ii}^1 = \hat{\sigma}_{ii}^2 = 1$, $i = 1, \dots, d$, und ferner für $i \neq j$ $\hat{\sigma}_{ij}^1 \leq \hat{\sigma}_{ij}^2$ genau dann, wenn $\sigma_{ij}^1 \leq \sigma_{ij}^2$ ist. Satz 4.3.3 liefert dann zunächst $P(\hat{\mathbf{X}} \leq \mathbf{t}) \leq P(\hat{\mathbf{Y}} \leq \mathbf{t})$ für alle $\mathbf{t} \in \mathbb{R}^d$. Wegen

$$\varphi_{-\hat{\mathbf{X}}}(\mathbf{t}) = \varphi_{\hat{\mathbf{X}}}(-\mathbf{t}) = \phi(\mathbf{t}'\hat{\Sigma}_1\mathbf{t})$$

folgt $-\hat{\mathbf{X}} \sim E_d(0, \hat{\Sigma}_1, \phi)$ und ebenso $-\hat{\mathbf{Y}} \sim E_d(0, \hat{\Sigma}_2, \phi)$. Daher ist die Funktion $P(-\hat{\mathbf{X}} \leq \mathbf{t}) = P(\hat{\mathbf{X}} \geq \mathbf{t})$, wieder aufgrund von Satz 4.3.3, für jedes feste $\mathbf{t} \in \mathbb{R}^d$ monoton wachsend in $\hat{\sigma}_{ij}^1$. Wegen

$$\lim_{\varepsilon \downarrow 0} P(\hat{\mathbf{X}} \geq \mathbf{t} - \varepsilon \mathbf{1}) = P(\hat{\mathbf{X}} > \mathbf{t}),$$

ist außerdem ebenfalls $P(\hat{\mathbf{X}} > \mathbf{t})$ monoton wachsend in $\hat{\sigma}_{ij}^1$ und somit gilt $P(\hat{\mathbf{X}} > \mathbf{t}) \leq P(\hat{\mathbf{Y}} > \mathbf{t})$ für alle $\mathbf{t} \in \mathbb{R}^d$. Es folgt $\hat{\mathbf{X}} \leq_c \hat{\mathbf{Y}}$. Schließlich liefert Lemma 4.2.2 b) die Behauptung für \mathbf{X} und \mathbf{Y} . \square

4.3.5. Beispiele.

- a) *Gleichverteilungen:* Wir definieren für $c > 0$ den Dichtegenerator $g(t) := c_d \mathbf{1}_{[0, c^2]}(t)$ und die zugehörige Familie von Dichtefunktionen

$$f(\mathbf{t}) = c_d |\Sigma|^{-\frac{1}{2}} \mathbf{1}_{[0, c^2]}((\mathbf{t} - \mu)' \Sigma^{-1} (\mathbf{t} - \mu)), \quad \mathbf{t} \in \mathbb{R}^d,$$

der Gleichverteilungen auf der Menge $\{\mathbf{t} \in \mathbb{R}^d; (\mathbf{t} - \mu)' \Sigma^{-1} (\mathbf{t} - \mu) \leq c^2\}$. Für $d = 2$ und $\mu = 0$ erhält man mit Korollar 4.3.4 insbesondere das anschauliche Resultat, dass eine Gleichverteilung auf dem Kreis mit Radius c im Sinne der Konkordanz-Ordnung kleiner ist, als die Gleichverteilung auf den Ellipsen $\{(t_1, t_2) \in \mathbb{R}^2; t_1^2 + t_2^2 - 2bt_1t_2 \leq (1 - b^2)c^2\}$, $b \in (0, 1)$, deren Hauptachsen auf der Ursprungsgeraden mit dem Richtungsvektor $(1, 1)$ liegen.

- b) *Kotz-Typ*: Wir betrachten den Dichtegenerator $g(t) = c_d t^{n-1} \exp(-rt^s)$, $r, s > 0$, $2n + d > 2$ und die zugehörige Familie von Dichtefunktionen

$$f(\mathbf{t}) = c_d [(\mathbf{t} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{t} - \boldsymbol{\mu})]^{n-1} |\boldsymbol{\Sigma}|^{-\frac{1}{2}} \exp(-r[(\mathbf{t} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{t} - \boldsymbol{\mu})]^s), \quad \mathbf{t} \in \mathbb{R}^d,$$

mit

$$c_d := \frac{s \Gamma(d/2) r^{(2n+d-2)/2s}}{\pi^{d/2} \Gamma((2n+d-2)/2s)}.$$

Eine Zufallsvariable \mathbf{X} mit der Dichte f heißt vom *Kotz-Typ mit Parametern n, r, s* . Speziell für $n = s = 1$ und $r = \frac{1}{2}$ ergibt sich die multivariate Normalverteilung mit Erwartungswertvektor $\boldsymbol{\mu}$ und Kovarianzmatrix $\boldsymbol{\Sigma}$. Satz 4.3.3 ist somit eine Verallgemeinerung der Slepianischen Ungleichung in Satz 4.3.1.

- c) *Pearson-Typ VII*: Der Dichtegenerator

$$g(t) = \frac{\Gamma(n)}{(\pi m)^{d/2} \Gamma(n - d/2)} \left(1 + \frac{t}{m}\right)^{-n}, \quad n > \frac{d}{2}, m > 0,$$

führt zu symmetrischen Verteilungen von Pearson-Typ VII mit den Freiheitsgraden m, n . Speziell für $m \in \mathbb{N}$ und $n = \frac{1}{2}(d + m)$ ergibt sich die multivariate t -Verteilung mit m Freiheitsgraden und der Dichte

$$f(\mathbf{t}) = \frac{\Gamma((d+m)/2)}{(\pi m)^{d/2} \Gamma(m/2)} |\boldsymbol{\Sigma}|^{-1/2} \left(1 + \frac{1}{m} (\mathbf{t} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{t} - \boldsymbol{\mu})\right)^{-(d+m)/2}, \quad \mathbf{t} \in \mathbb{R}^d.$$

Für $m = 1$ und $n = \frac{1}{2}(d + 1)$ führt das auf die multivariate Cauchyverteilung.

- d) *Logistische Verteilung*: Der Dichtegenerator $g(t) = c_d \frac{e^{-t}}{(1+e^{-t})^2}$ führt auf multivariate symmetrische logistische Verteilungen.

Weitere Beispiele zu elliptisch symmetrischen Verteilungen finden sich in [FC84].

Kapitel 5

Test auf Konkordanz-Ordnung

5.1 Supremums- und Integraltestgrößen

Wir wollen in diesem Abschnitt nichtparametrische Testverfahren für die Hypothese der Konkordanz-Ordnung bereitstellen und erneut deren asymptotische Eigenschaften untersuchen. Die im letzten Kapitel begründete Interpretation von \leq_c als multivariate Abhängigkeitsordnung auf der einen Seite und deren Charakterisierung über die Ordnung von Parametern bestimmter Verteilungsfamilien auf der anderen Seite, legen verschiedene Ansätze für parametrische Verfahren nahe. In [CDS04] findet sich beispielsweise im Einstichprobenfall ein parametrisches Verfahren um die Hypothese der Konkordanz-Ordnung in dem Fall zu überprüfen, dass die parametrische Klasse der zugrunde liegenden Kopulas der Zufallsvektoren \mathbf{X} , \mathbf{Y} bekannt ist. Bisher ist der Literatur nach Wissen des Autors kein nichtparametrischer Test auf Konkordanz-Ordnung bekannt.

Es seien im Folgenden \mathbf{X} und \mathbf{Y} d -dimensionale und komponentenweise reellwertige Zufallsvektoren mit den Verteilungsfunktionen F und G und den Überlebensfunktionen \bar{F} und \bar{G} . Wir untersuchen das Zweistichproben-Testproblem

$$H : \mathbf{X} \leq_c \mathbf{Y} \quad \text{vs.} \quad K : \mathbf{X} \not\leq_c \mathbf{Y}. \quad (5.1.1)$$

Äquivalent dazu liegt die Hypothese vor, dass die Verteilungsfunktionen und die Überlebensfunktionen jeweils punktweise geordnet sind, also

$$\begin{aligned} H : \forall \mathbf{t} \in \mathbb{R}^d : \quad & F(\mathbf{t}) \leq G(\mathbf{t}) \quad \wedge \quad \bar{F}(\mathbf{t}) \leq \bar{G}(\mathbf{t}) \\ K : \exists \mathbf{t} \in \mathbb{R}^d : \quad & F(\mathbf{t}) > G(\mathbf{t}) \quad \vee \quad \bar{F}(\mathbf{t}) > \bar{G}(\mathbf{t}), \end{aligned}$$

gilt. Seien $\mathbf{X}_1, \dots, \mathbf{X}_m$, $m \in \mathbb{N}$, unabhängige und identisch verteilte Zufallsvektoren mit

der Verteilungsfunktion F und $\mathbf{Y}_1, \dots, \mathbf{Y}_n$, $n \in \mathbb{N}$, unabhängige und identisch verteilte Zufallsvektoren mit der Verteilungsfunktion G . Es seien $\mathbf{X}_1, \dots, \mathbf{X}_m, \mathbf{Y}_1, \dots, \mathbf{Y}_n$ unabhängig, sowie F, G stetig. Es seien F_m, G_n die empirischen Verteilungsfunktionen zu F, G und

$$\bar{F}_m(\mathbf{t}) := \frac{1}{m} \sum_{i=1}^m \mathbf{1}_{(t, \infty)}(\mathbf{X}_i), \quad \bar{G}_n(\mathbf{t}) := \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{(t, \infty)}(\mathbf{Y}_i), \quad \mathbf{t} \in \mathbb{R}^d,$$

die *empirischen Überlebensfunktionen* zu F und G .

Wir setzen ferner für $\mathbf{t} \in \mathbb{R}^d$ und $m, n \in \mathbb{N}$

$$\Delta_{m,n}(\mathbf{t}) := \sqrt{\frac{mn}{m+n}} (F_m(\mathbf{t}) - G_n(\mathbf{t})), \quad \bar{\Delta}_{m,n}(\mathbf{t}) := \sqrt{\frac{mn}{m+n}} (\bar{F}_m(\mathbf{t}) - \bar{G}_n(\mathbf{t})).$$

Ähnlich zum Test auf Laplace-Ordnung scheinen auch hier einseitige Varianten von Supremums- und Integraltestgrößen, die auf diesen empirischen Abständen beruhen, plausible Kandidaten für Testgrößen zu sein. Wir definieren daher die Testgrößen

$$C_{m,n}^{Kui} := \sup_{\mathbf{t} \in \mathbb{R}^d} \Delta_{m,n}(\mathbf{t}) + \sup_{\mathbf{t} \in \mathbb{R}^d} \bar{\Delta}_{m,n}(\mathbf{t}),$$

und

$$C_{m,n}^{KS} := \max \left(\sup_{\mathbf{t} \in \mathbb{R}^d} \Delta_{m,n}(\mathbf{t}), \sup_{\mathbf{t} \in \mathbb{R}^d} \bar{\Delta}_{m,n}(\mathbf{t}) \right)$$

sowie

$$C_{m,n}^{CvM} := \int_{\mathbb{R}^d} \left(\Delta_{m,n}^+(\mathbf{t}) + \bar{\Delta}_{m,n}^+(\mathbf{t}) \right) \omega(\mathbf{t}) d\mathbf{t},$$

erneut mit einer integrierbaren Gewichtsfunktion $\omega : \mathbb{R}^d \rightarrow \mathbb{R}_{>0}$. Die definierten Testgrößen können als einseitige Varianten der klassischen Kuiper-, Kolmogorov-Smirnov- und Cramér-von Mises-Testgrößen angesehen werden. Große Werte von $C_{m,n}^{Kui}$ respektive $C_{m,n}^{KS}$ oder $C_{m,n}^{CvM}$ sprechen gegen die Hypothese.

5.1.1 Ein parametrischer Test

In parametrischen Situationen liefern die Ergebnisse des letzten Kapitels Ansätze für alternative Testgrößen. Wir wollen diesen Sachverhalt anhand eines Beispiels illustrieren. Seien $\mathbf{X} = (X_1, X_2), \mathbf{Y} = (Y_1, Y_2)$ zweidimensionale Zufallsvektoren mit den Verteilungen

$$\mathbf{X} \sim N_2(\mu, \Sigma_1), \quad \mathbf{Y} \sim N_2(\mu, \Sigma_2), \quad (5.1.2)$$

für ein $\mu \in \mathbb{R}^2$ und den positiv definiten Kovarianzmatrizen

$$\Sigma_1 = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{pmatrix}, \quad \Sigma_2 = \begin{pmatrix} \sigma_1^2 & \tau_{12} \\ \tau_{21} & \sigma_2^2 \end{pmatrix},$$

$\sigma_1^2, \sigma_2^2 > 0$, $\sigma_{12} = \text{Cov}(X_1, X_2)$, $\tau_{12} = \text{Cov}(Y_1, Y_2)$. Dann haben \mathbf{X} und \mathbf{Y} elliptisch symmetrische Verteilungen mit denselben Randverteilungen. Nach Korollar 4.2.4 a) und Satz 4.3.3 ist daher die Ordnung der Kovarianzen eine äquivalente Bedingung für die Konkordanz-Ordnung von \mathbf{X} und \mathbf{Y} , das heißt es gilt $\sigma_{12} \leq \tau_{12} \Leftrightarrow \mathbf{X} \leq_c \mathbf{Y}$.

Es seien $\rho_{\mathbf{X}} = \rho_{X_1, X_2}$ und $\rho_{\mathbf{Y}} = \rho_{Y_1, Y_2}$ die Korrelationskoeffizienten von \mathbf{X} bzw. von \mathbf{Y} . Da \mathbf{X} und \mathbf{Y} die selben Randverteilungen haben, gilt $\sigma_{12} \leq \tau_{12}$ genau dann, wenn $\rho_{\mathbf{X}} \leq \rho_{\mathbf{Y}}$ ist. Ferner ist $\rho_{\mathbf{X}} \leq \rho_{\mathbf{Y}}$ genau dann, wenn

$$\frac{1}{2} \log \left(\frac{1 + \rho_{\mathbf{X}}}{1 - \rho_{\mathbf{X}}} \right) \leq \frac{1}{2} \log \left(\frac{1 + \rho_{\mathbf{Y}}}{1 - \rho_{\mathbf{Y}}} \right)$$

ist. Wir definieren

$$\rho_{\mathbf{X},m} := \frac{\sum_{k=1}^m (X_{k1} - \bar{X}_{\cdot 1})(X_{k2} - \bar{X}_{\cdot 2})}{\sqrt{\sum_{k=1}^m (X_{k1} - \bar{X}_{\cdot 1})^2 \sum_{k=1}^m (X_{k2} - \bar{X}_{\cdot 2})^2}}, \quad \bar{X}_{\cdot j} := \frac{1}{m} \sum_{k=1}^m X_{kj}, \quad j = 1, 2,$$

mit unabhängigen Zufallsvariablen $\mathbf{X}_1 = (X_{11}, X_{12}), \dots, \mathbf{X}_m = (X_{m1}, X_{m2})$, $m \in \mathbb{N}$, je mit derselben Verteilung wie \mathbf{X} . Es ist $\rho_{\mathbf{X},m}$ der *empirische Korrelationskoeffizient* der Zufallsvariablen $\mathbf{X}_1, \dots, \mathbf{X}_m$. Analog bezeichnen wir mit $\rho_{\mathbf{Y},n}$ den empirischen Korrelationskoeffizienten der Zufallsvariablen $\mathbf{Y}_1, \dots, \mathbf{Y}_n$.

Zur Überprüfung der Hypothese $H : \rho_{\mathbf{X}} \leq \rho_{\mathbf{Y}}$ gegen die Alternative $K : \rho_{\mathbf{X}} > \rho_{\mathbf{Y}}$ scheint dann die Testgröße

$$T_{m,n} = \sqrt{\frac{mn}{m+n}} \left(\frac{1}{2} \log \left(\frac{1 + \rho_{\mathbf{X},m}}{1 - \rho_{\mathbf{X},m}} \right) - \frac{1}{2} \log \left(\frac{1 + \rho_{\mathbf{Y},n}}{1 - \rho_{\mathbf{Y},n}} \right) \right)$$

plausibel. Wir schreiben

$$\begin{aligned} T_{m,n} &= \sqrt{\frac{n}{m+n}} \sqrt{m} \left(\frac{1}{2} \log \left(\frac{1 + \rho_{\mathbf{X},m}}{1 - \rho_{\mathbf{X},m}} \right) - \frac{1}{2} \log \left(\frac{1 + \rho_{\mathbf{X}}}{1 - \rho_{\mathbf{X}}} \right) \right) \\ &\quad - \sqrt{\frac{m}{m+n}} \sqrt{n} \left(\frac{1}{2} \log \left(\frac{1 + \rho_{\mathbf{Y},n}}{1 - \rho_{\mathbf{Y},n}} \right) - \frac{1}{2} \log \left(\frac{1 + \rho_{\mathbf{Y}}}{1 - \rho_{\mathbf{Y}}} \right) \right) \\ &\quad + \sqrt{\frac{mn}{m+n}} \left(\frac{1}{2} \log \left(\frac{1 + \rho_{\mathbf{X}}}{1 - \rho_{\mathbf{X}}} \right) - \frac{1}{2} \log \left(\frac{1 + \rho_{\mathbf{Y}}}{1 - \rho_{\mathbf{Y}}} \right) \right). \end{aligned} \quad (5.1.3)$$

Bekanntlich gilt

$$\sqrt{m} \left(\frac{1}{2} \log \left(\frac{1 + \rho_{\mathbf{X},m}}{1 - \rho_{\mathbf{X},m}} \right) - \frac{1}{2} \log \left(\frac{1 + \rho_{\mathbf{X}}}{1 - \rho_{\mathbf{X}}} \right) \right) \xrightarrow{v} N(0, 1), \quad m \rightarrow \infty,$$

(siehe [WMF95], Beispiel 5.111). Das Konvergenzverhalten der Testgröße wird daher vorwiegend durch den letzten Term in (5.1.3) bestimmt. Wir erhalten direkt die nachstehenden asymptotischen Aussagen.

5.1.1. Lemma.

- a) Sei $\rho_{\mathbf{X}} = \rho_{\mathbf{Y}}$. Dann gilt $T_{m,n} \xrightarrow{v} N(0, 1)$ für $m, n \rightarrow \infty$.
- b) Sei $\rho_{\mathbf{X}} < \rho_{\mathbf{Y}}$. Dann gilt $T_{m,n} \xrightarrow{P} -\infty$ für $m, n \rightarrow \infty$.
- c) Sei $\rho_{\mathbf{X}} > \rho_{\mathbf{Y}}$. Dann gilt $T_{m,n} \xrightarrow{P} \infty$ für $m, n \rightarrow \infty$.

Bezeichnet man mit u_α , $\alpha \in (0, 1)$, das $(1 - \alpha)$ -Quantil der Standard-Normalverteilung, so liefert das letzte Lemma, dass der Test

$$\varphi_{m,n}^{Par}(\mathbf{X}_1, \dots, \mathbf{X}_m, \mathbf{Y}_1, \dots, \mathbf{Y}_n) := \begin{cases} 1, & T_{m,n} > u_\alpha, \\ 0, & T_{m,n} \leq u_\alpha, \end{cases}$$

asymptotisch jedes vorgegebene Testniveau $\alpha \in (0, 1)$ einhält und im zugrunde liegenden Modell konsistent gegen alle Alternativen ist. Die Tabellen 5.1 und 5.2 zeigen die bei einigen Simulationen erhaltenen empirischen Güten dieses Tests. Es bezeichnet F die Verteilungsfunktion von \mathbf{X} und G die Verteilungsfunktion von \mathbf{Y} .

Die Diagonalelemente in Tabelle 5.1 und 5.2 repräsentieren den Fall, dass $F = G$ ist; unterhalb der Diagonalen wird der Fall $F \not\leq_c G$ behandelt. Oberhalb der Diagonalen ist $F \leq_c G$ und $F \neq G$.

In Tabelle 5.1 wurde für die Verteilungen in (5.1.2) stets $\mu = (0, 0)$ und $\sigma_1^2 = \sigma_2^2 = 1$ gewählt. In Tabelle 5.2 ist für die Verteilungen in (5.1.2) stets $\mu = (15, -10)$ und $\sigma_1^2 = 0.5, \sigma_2^2 = 3$.

Es zeigt sich in beiden Fällen eine leichte Überschreitung des Testniveaus für die Situation $F = G$, während das Testniveau im Fall $F \leq_c G$, $F \neq G$, deutlich unterschritten wird.

$\alpha = 0.05$		G				
		$\tau_{12} = 0$	$\tau_{12} = \frac{1}{5}$	$\tau_{12} = \frac{1}{2}$	$\tau_{12} = \frac{7}{10}$	$\tau_{12} = \frac{9}{10}$
F	$\sigma_{12} = 0$	0.056	0.005	0.000	0.000	0.000
	$\sigma_{12} = \frac{1}{5}$	0.298	0.056	0.000	0.000	0.000
	$\sigma_{12} = \frac{1}{2}$	0.855	0.552	0.062	0.000	0.000
	$\sigma_{12} = \frac{7}{10}$	0.995	0.947	0.467	0.059	0.000
	$\sigma_{12} = \frac{9}{10}$	1.000	1.000	0.996	0.926	0.056

Tabelle 5.1: Empirische Güte für $\varphi_{m,n}^{Par}$ für die Verteilungen in (5.1.2) mit $\mu = (0, 0)$ und $\sigma_1^2 = \sigma_2^2 = 1$ bei 1000 Durchläufen. Es ist $m = n = 50$.

$\alpha = 0.05$		G				
		$\tau_{12} = 0$	$\tau_{12} = \frac{1}{5}$	$\tau_{12} = \frac{1}{2}$	$\tau_{12} = \frac{7}{10}$	$\tau_{12} = \frac{9}{10}$
F	$\sigma_{12} = 0$	0.061	0.010	0.000	0.000	0.000
	$\sigma_{12} = \frac{1}{5}$	0.162	0.054	0.005	0.000	0.000
	$\sigma_{12} = \frac{1}{2}$	0.593	0.345	0.055	0.005	0.000
	$\sigma_{12} = \frac{7}{10}$	0.896	0.666	0.241	0.060	0.004
	$\sigma_{12} = \frac{9}{10}$	0.993	0.948	0.677	0.350	0.056

Tabelle 5.2: Empirische Güte für $\varphi_{m,n}^{Par}$ für die Verteilungen in (5.1.2) mit $\mu = (15, -10)$ und $\sigma_1^2 = 0.5, \sigma_2^2 = 3$ bei 1000 Durchläufen. Es ist $m = 50, n = 30$.

5.2 Asymptotik

Wir kehren zur eingangs beschriebenen nichtparametrischen Ausgangssituation und den vorgestellten einseitigen Testgrößen zurück. Die Testgrößen $C_{m,n}^{Kui}$, $C_{m,n}^{KS}$ und $C_{m,n}^{CvM}$ aus Abschnitt 5.1 ergeben sich dabei offenbar jeweils als Funktionale von $(\Delta_{m,n}(\mathbf{t}), \mathbf{t} \in \mathbb{R}^d)$ und $(\bar{\Delta}_{m,n}(\mathbf{t}), \mathbf{t} \in \mathbb{R}^d)$.

Gemäß der im letzten Kapitel angemerkten Charakterisierung von \leq_c als stochastische Integralordnung gilt $\mathbf{X} \leq_c \mathbf{Y}$ genau dann, wenn $Ef(\mathbf{X}) \leq Ef(\mathbf{Y})$ für alle $f \in \mathcal{F} := \mathcal{F}_1 \cup \mathcal{F}_2$ mit

$$\mathcal{F}_1 := \{\mathbf{1}_{(-\infty, \mathbf{t}]}, \mathbf{t} \in \mathbb{R}^d\}, \quad \mathcal{F}_2 := \{\mathbf{1}_{(\mathbf{t}, \infty)}, \mathbf{t} \in \mathbb{R}^d\} \quad (5.2.1)$$

ist. Seien $P_m = \frac{1}{m} \sum_{k=1}^m \delta_{\mathbf{X}_k}$ und $Q_n = \frac{1}{n} \sum_{k=1}^n \delta_{\mathbf{Y}_k}$ die empirischen Maße der Zufallsvektoren $\mathbf{X}_1, \dots, \mathbf{X}_m$ bzw. $\mathbf{Y}_1, \dots, \mathbf{Y}_n$. Wir betrachten den durch \mathcal{F} indizierten empirischen Prozess

$$\mathbb{G}_{m,n} := (\mathbb{G}_{m,n}f, f \in \mathcal{F}) = \left(\sqrt{\frac{mn}{m+n}} (P_m - Q_n)f, f \in \mathcal{F} \right).$$

Offenbar lässt sich $\mathbb{G}_{m,n}|_{\mathcal{F}_1}$ mit $\Delta_{m,n}$ und $\mathbb{G}_{m,n}|_{\mathcal{F}_2}$ mit $\bar{\Delta}_{m,n}$ identifizieren. Der empirische Prozess wird durch die oberen und unteren Indikatorfunktionen indiziert, also durch Funktionenklassen, die bekanntlich VC-Funktionenklassen sind. Wir formulieren dieses Resultat der Vollständigkeit halber in folgendem Lemma.

5.2.1. Lemma. *Die Funktionenklasse $\mathcal{F} = \mathcal{F}_1 \cup \mathcal{F}_2$ mit $\mathcal{F}_1, \mathcal{F}_2$ wie in (5.2.1) ist eine VC-Funktionenklasse und außerdem gleichmäßig Donsker.*

Beweis. Die Mengensysteme $\{(-\infty, \mathbf{t}], \mathbf{t} \in \mathbb{R}^d\}$ und $\{(\mathbf{t}, \infty), \mathbf{t} \in \mathbb{R}^d\}$ sind VC-Klassen (vgl. [VW96], 2.6.1 und 2.6.17 i)) und die Funktionenklassen $\mathcal{F}_1, \mathcal{F}_2$ daher VC-Funktionenklassen. Aufgrund der Beschränktheit sämtlicher Funktionen $f \in \mathcal{F}_1 \cup \mathcal{F}_2$ sind \mathcal{F}_1 und \mathcal{F}_2 insbesondere gleichmäßig Donsker. Dann ist \mathcal{F} als Vereinigung von \mathcal{F}_1 und \mathcal{F}_2 ebenfalls eine VC-Funktionenklasse und wegen [VW96], 2.10.7, ist \mathcal{F} ferner gleichmäßig Donsker. \square

5.2.2. Bemerkung. Bei den weiteren Überlegungen werden teilweise auch die durch die Funktionenklasse $\bar{\mathcal{F}} = \bar{\mathcal{F}}_1 \cup \bar{\mathcal{F}}_2$,

$$\bar{\mathcal{F}}_1 := \{\mathbf{1}_{(-\infty, \mathbf{t}]}, \mathbf{t} \in \bar{\mathbb{R}}^d\}, \quad \bar{\mathcal{F}}_2 := \{\mathbf{1}_{(\mathbf{t}, \infty)}, \mathbf{t} \in \bar{\mathbb{R}}^d\},$$

indizierten empirischen Prozesse betrachtet. Mit ähnlichen Überlegungen wie in Lemma 5.2.1 folgt, dass $\bar{\mathcal{F}}$ ebenfalls eine VC-Funktionenklasse und gleichmäßig Donsker ist.

Im Folgenden sei wieder $\frac{m}{m+n} \rightarrow \tau \in [0, 1]$. Bei Gültigkeit der Hypothese werden sich die Bereiche $A = \{\mathbf{t} \in \mathbb{R}^d : F(\mathbf{t}) = G(\mathbf{t})\}$ und $B = \{\mathbf{t} \in \mathbb{R}^d : \bar{F}(\mathbf{t}) = \bar{G}(\mathbf{t})\}$ erneut als relevant für das Grenzverhalten der Testgrößen herausstellen. Bevor wir die Resultate zum asymptotischen Verhalten der Testgrößen formulieren, führen wir zunächst noch einige Notationen ein:

Es seien im Folgenden P bzw. Q die Verteilungen der Zufallsvariablen \mathbf{X}_1 bzw. \mathbf{Y}_1 . Es seien $C_F := (C_F f, f \in \mathcal{F})$ und $C_G := (C_G f, f \in \mathcal{F})$ stochastisch unabhängige und jeweils zentrierte und straffe Gaußprozesse mit Pfaden in $l^\infty(\mathcal{F})$ und den zugehörigen Kovarianzfunktionen

$$\text{Cov}(C_F f, C_F g) := Pfg - PfPg, \quad \text{Cov}(C_G f, C_G g) := Qfg - QfQg,$$

für $f, g \in \mathcal{F}$. Wir setzen

$$\mathcal{C} := (\mathcal{C}f, f \in \mathcal{F}) := (\sqrt{1-\tau} C_F f - \sqrt{\tau} C_G f, f \in \mathcal{F})$$

und bezeichnen die Restriktion des Prozesses \mathcal{C} auf die Teilmengen \mathcal{F}_1 bzw. \mathcal{F}_2 mit C bzw. \bar{C} . Genauer sei

$$C = (C(\mathbf{t}), \mathbf{t} \in \mathbb{R}^d), \quad \text{mit } C(\mathbf{t}) := \mathcal{C}\mathbf{1}_{(-\infty, \mathbf{t}]},$$

sowie

$$\bar{C} = (\bar{C}(\mathbf{t}), \mathbf{t} \in \mathbb{R}^d), \quad \text{mit } \bar{C}(\mathbf{t}) := \mathcal{C}\mathbf{1}_{(\mathbf{t}, \infty)}.$$

Dann ist $C(\mathbf{t}) = (C(\mathbf{t}), \mathbf{t} \in \mathbb{R}^d)$ ein zentrierter Gaußprozess in $l^\infty(\mathbb{R}^d)$ mit der Kovarianzfunktion

$$\begin{aligned} \text{Cov}(C(\mathbf{s}), C(\mathbf{t})) &= (1-\tau) [F(\min(\mathbf{s}, \mathbf{t})) - F(\mathbf{s})F(\mathbf{t})] \\ &\quad + \tau [G(\min(\mathbf{s}, \mathbf{t})) - G(\mathbf{s})G(\mathbf{t})], \quad \mathbf{s}, \mathbf{t} \in \mathbb{R}^d, \end{aligned}$$

und $(\bar{C}(\mathbf{t}), \mathbf{t} \in \mathbb{R}^d)$ ein zentrierter Gaußprozess in $l^\infty(\mathbb{R}^d)$ mit der Kovarianzfunktion

$$\begin{aligned} \text{Cov}(\bar{C}(\mathbf{s}), \bar{C}(\mathbf{t})) &= (1-\tau) [\bar{F}(\max(\mathbf{s}, \mathbf{t})) - \bar{F}(\mathbf{s})\bar{F}(\mathbf{t})] \\ &\quad + \tau [\bar{G}(\max(\mathbf{s}, \mathbf{t})) - \bar{G}(\mathbf{s})\bar{G}(\mathbf{t})], \quad \mathbf{s}, \mathbf{t} \in \mathbb{R}^d. \end{aligned}$$

5.2.3. Satz. *Bei Gültigkeit der Hypothese gilt für $m, n \rightarrow \infty$*

$$C_{m,n}^{CvM} \xrightarrow{v} C^{CvM} := \int_A C^+(\mathbf{t}) \omega(\mathbf{t}) d\mathbf{t} + \int_B \bar{C}^+(\mathbf{t}) \omega(\mathbf{t}) d\mathbf{t}.$$

Bei Gültigkeit der Alternative ist

$$C_{m,n}^{CvM} \xrightarrow{P} \infty, \quad m, n \rightarrow \infty.$$

Beweis. Für ein $f \in \mathcal{F}$ ist

$$\begin{aligned} \mathbb{G}_{m,n}f &= \sqrt{\frac{mn}{m+n}}(P_m - P)f - \sqrt{\frac{mn}{m+n}}(Q_n - Q)f \\ &\quad + \sqrt{\frac{mn}{m+n}}(P - Q)f. \end{aligned} \quad (5.2.2)$$

Aufgrund des vorhergehenden Lemmas sind die durch \mathcal{F} indizierten empirischen Prozesse $\left(\sqrt{\frac{mn}{m+n}}(P_m - P)f, f \in \mathcal{F}\right)$ und $\left(\sqrt{\frac{mn}{m+n}}(Q_n - Q)f, f \in \mathcal{F}\right)$ jeweils asymptotisch straffe Folgen von Abbildungen mit Werten in $(l^\infty(\mathcal{F}), \|\cdot\|_\infty)$. Die d -dimensionalen Randverteilungen dieser empirischen Prozesse konvergieren aufgrund der Quadratintegrierbarkeit der Indikatorfunktionen außerdem nach dem mehrdimensionalen zentralen Grenzwertsatz gegen d -dimensionale Normalverteilungen.

Etwas genauer formuliert, konvergieren die d -dimensionalen Randverteilungen von $\left(\sqrt{\frac{mn}{m+n}}(P_m - P)f, f \in \mathcal{F}\right)$ in Verteilung gegen zentrierte Normalverteilungen $N_d(0, \Sigma)$, mit $\Sigma = (\sigma_{ij})_{1 \leq i, j \leq d}$ und $\sigma_{ij} = \text{Cov}(f_i(\mathbf{X}_1), f_j(\mathbf{X}_1))$ für geeignete $f_i, f_j \in \mathcal{F}$, und es konvergieren die d -dimensionalen Randverteilungen von $\left(\sqrt{\frac{mn}{m+n}}(Q_n - Q)f, f \in \mathcal{F}\right)$ in Verteilung gegen zentrierte Normalverteilung $N_d(0, \Sigma)$, mit $\Sigma = (\sigma_{ij})_{1 \leq i, j \leq d}$ und $\sigma_{ij} = \text{Cov}(f_i(\mathbf{Y}_1), f_j(\mathbf{Y}_1))$ für geeignete $f_i, f_j \in \mathcal{F}$.

Die asymptotische Straffheit der empirischen Prozesse ergibt dann

$$\left(\sqrt{\frac{mn}{m+n}}(P_m - P)f, f \in \mathcal{F}\right) \xrightarrow{v} (\sqrt{1-\tau}C_F f, f \in \mathcal{F}), \quad m, n \rightarrow \infty,$$

in $l^\infty(\mathcal{F})$.

Für Funktionen $f_s = \mathbf{1}_{(-\infty, s]}$, $f_t = \mathbf{1}_{(-\infty, t]}$ $\in \mathcal{F}_1$ ist insbesondere

$$\begin{aligned} \text{Cov}(C_F f_s, C_F f_t) &= \int_{\mathbb{R}^d} f_s(\mathbf{x})f_t(\mathbf{x})dF(\mathbf{x}) - \int_{\mathbb{R}^d} f_s(\mathbf{x})dF(\mathbf{x}) \int_{\mathbb{R}^d} f_t(\mathbf{x})dF(\mathbf{x}) \\ &= \int_{\mathbb{R}^d} \mathbf{1}_{(-\infty, t]}(\mathbf{x})\mathbf{1}_{(-\infty, s]}(\mathbf{x})dF(\mathbf{x}) \\ &\quad - \int_{\mathbb{R}^d} \mathbf{1}_{(-\infty, s]}(\mathbf{x})dF(\mathbf{x}) \int_{\mathbb{R}^d} \mathbf{1}_{(-\infty, t]}(\mathbf{x})dF(\mathbf{x}) \\ &= F(\min(\mathbf{s}, \mathbf{t})) - F(\mathbf{s})F(\mathbf{t}), \end{aligned} \quad (5.2.3)$$

wobei das Minimum komponentenweise zu bilden ist. Analog ist für Funktionen $g_s = \mathbf{1}_{(s, \infty)}$, $g_t = \mathbf{1}_{(t, \infty)} \in \mathcal{F}_2$

$$\text{Cov}(C_F g_s, C_F g_t) = \bar{F}(\max(\mathbf{s}, \mathbf{t})) - \bar{F}(\mathbf{s})\bar{F}(\mathbf{t}). \quad (5.2.4)$$

Mit der gleichen Argumentation ergibt sich für den zweiten empirischen Prozess

$$\left(\sqrt{\frac{mn}{m+n}} (Q_n - Q) f, f \in \mathcal{F} \right) \xrightarrow{v} (\sqrt{\tau} C_G f, f \in \mathcal{F}), \quad m, n \rightarrow \infty,$$

in $l^\infty(\mathcal{F})$ mit

$$\text{Cov}(C_G f_s, C_G f_t) = G(\min(\mathbf{s}, \mathbf{t})) - G(\mathbf{s})G(\mathbf{t}), \quad \text{falls } f_s, f_t \in \mathcal{F}_1, \quad (5.2.5)$$

und

$$\text{Cov}(C_G g_s, C_G g_t) = \bar{G}(\max(\mathbf{s}, \mathbf{t})) - \bar{G}(\mathbf{s})\bar{G}(\mathbf{t}), \quad \text{falls } g_s, g_t \in \mathcal{F}_2. \quad (5.2.6)$$

Aufgrund der Unabhängigkeit folgt schließlich

$$\begin{aligned} \left(\sqrt{\frac{mn}{m+n}} \left((P_m - P) f - (Q_n - Q) f \right), f \in \mathcal{F} \right) &\xrightarrow{v} (\sqrt{1-\tau} C_F f - \sqrt{\tau} C_G f, f \in \mathcal{F}) \\ &= (\mathcal{C}f, f \in \mathcal{F}), \quad m, n \rightarrow \infty, \end{aligned}$$

in $l^\infty(\mathcal{F})$. Es ist

$$\begin{aligned} \text{Cov}(\mathcal{C}f, \mathcal{C}g) &= \text{Cov}(\sqrt{1-\tau} C_F f - \sqrt{\tau} C_G f, \sqrt{1-\tau} C_F g - \sqrt{\tau} C_G g) \\ &= (1-\tau) \text{Cov}(C_F f, C_F g) + \tau \text{Cov}(C_G f, C_G g), \quad f, g \in \mathcal{F}. \end{aligned}$$

Eingeschränkt auf $\mathcal{F}_1 = \{f_t, \mathbf{t} \in \mathbb{R}^d : f_t = \mathbf{1}_{(-\infty, \mathbf{t}]}\}$ hat der Prozess $(\mathcal{C}f, f \in \mathcal{F})$ also dieselbe Kovarianzfunktion wie $(C(\mathbf{t}), \mathbf{t} \in \mathbb{R}^d)$ und somit nach Lemma B.2.2 dieselbe Verteilung. Außerdem hat $(\mathcal{C}f, f \in \mathcal{F})$ eingeschränkt auf $\mathcal{F}_2 = \{f_t, \mathbf{t} \in \mathbb{R}^d : f_t = \mathbf{1}_{(\mathbf{t}, \infty)}\}$ dieselbe Kovarianzfunktion wie $(\bar{C}(\mathbf{t}), \mathbf{t} \in \mathbb{R}^d)$ und somit dieselbe Verteilung.

Wir wollen nun das Verhalten der Testgröße $C_{m,n}^{CvM}$ selbst untersuchen und definieren zu diesem Zweck zunächst für $\mathbf{t} \in \mathbb{R}^d$ und $m, n \in \mathbb{N}$ die Hilfsprozesse $(C_{m,n}(\mathbf{t}), \mathbf{t} \in \mathbb{R}^d)$ und $(\bar{C}_{m,n}(\mathbf{t}), \mathbf{t} \in \mathbb{R}^d)$ durch

$$\begin{aligned} C_{m,n}(\mathbf{t}) &:= \sqrt{\frac{mn}{m+n}} (F_m(\mathbf{t}) - F(\mathbf{t})) - \sqrt{\frac{mn}{m+n}} (G_n(\mathbf{t}) - G(\mathbf{t})) \\ \bar{C}_{m,n}(\mathbf{t}) &:= \sqrt{\frac{mn}{m+n}} (\bar{F}_m(\mathbf{t}) - \bar{F}(\mathbf{t})) - \sqrt{\frac{mn}{m+n}} (\bar{G}_n(\mathbf{t}) - \bar{G}(\mathbf{t})). \end{aligned}$$

Die Prozesse $C_{m,n}$ und $\bar{C}_{m,n}$ können dann als Einschränkungen von

$$\left(\sqrt{\frac{mn}{m+n}} \left((P_m - P) f - (Q_n - Q) f \right), f \in \mathcal{F} \right)$$

auf die Bereiche \mathcal{F}_1 bzw. \mathcal{F}_2 angesehen werden und konvergieren nach den Vorbetrachtungen in Verteilung gegen die Gaußprozesse C und \bar{C} .

Ferner betrachten wir das Funktional

$$T^{CvM} : l^\infty(\mathcal{F}) \longrightarrow \mathbb{R}_{\geq 0}$$

$$x \mapsto \int_{\mathbb{R}^d}^* x^+(\mathbf{1}_{(-\infty, \mathbf{t}]}) d\mu(\mathbf{t}) + \int_{\mathbb{R}^d}^* x^+(\mathbf{1}_{(\mathbf{t}, \infty)}) d\mu(\mathbf{t}),$$

wobei μ ein Wahrscheinlichkeitsmaß auf \mathbb{R}^d ist, welches die eingangs definierte Gewichtsfunktion ω als Lebesgue-Dichte hat, und $\int^* d\mu$ das äußere Integral bezüglich μ bezeichnet. Dann ist $T^{CvM}(\mathbb{G}_{m,n}) = C_{m,n}^{CvM}$. Bei Gültigkeit der Hypothese H gilt außerdem $F(\mathbf{t}) \leq G(\mathbf{t})$ und $\bar{F}(\mathbf{t}) \leq \bar{G}(\mathbf{t})$ jeweils für alle $\mathbf{t} \in \mathbb{R}^d$, so dass dann

$$\begin{aligned} T^{CvM}(\mathbb{G}_{m,n}) &= \int_{\mathbb{R}^d}^* \mathbb{G}_{m,n}^+(\mathbf{1}_{(-\infty, \mathbf{t}]}) d\mu(\mathbf{t}) + \int_{\mathbb{R}^d}^* \mathbb{G}_{m,n}^+(\mathbf{1}_{(\mathbf{t}, \infty)}) d\mu(\mathbf{t}) \\ &= \int_{\mathbb{R}^d} \Delta_{m,n}^+(\mathbf{t}) d\mu(\mathbf{t}) + \int_{\mathbb{R}^d} \bar{\Delta}_{m,n}^+(\mathbf{t}) d\mu(\mathbf{t}) \\ &= \int_{\{F=G\}} C_{m,n}^+(\mathbf{t}) d\mu(\mathbf{t}) + \int_{\{F<G\}} \left(C_{m,n}(\mathbf{t}) + \sqrt{\frac{mn}{m+n}}(F(\mathbf{t}) - G(\mathbf{t})) \right)^+ d\mu(\mathbf{t}) \\ &\quad + \int_{\{\bar{F}=\bar{G}\}} \bar{C}_{m,n}^+(\mathbf{t}) d\mu(\mathbf{t}) + \int_{\{\bar{F}<\bar{G}\}} \left(\bar{C}_{m,n}(\mathbf{t}) + \sqrt{\frac{mn}{m+n}}(\bar{F}(\mathbf{t}) - \bar{G}(\mathbf{t})) \right)^+ d\mu(\mathbf{t}) \end{aligned}$$

gilt. Aufgrund von Satz B.1.6 finden wir auf einem geeigneten Wahrscheinlichkeitsraum Abbildungen $\tilde{C}_{m,n}, \tilde{\bar{C}}$ mit $\tilde{C}_{m,n} \sim C_{m,n}$, $\tilde{C} \sim C$ und $\|\tilde{C}_{m,n} - \tilde{C}\|_\infty^* \rightarrow 0$ fast sicher für $m, n \rightarrow \infty$ und ebenso Abbildungen $\tilde{\bar{C}}_{m,n}, \tilde{\bar{C}}$ mit $\tilde{\bar{C}}_{m,n} \sim \bar{C}_{m,n}$, $\tilde{\bar{C}} \sim \bar{C}$ und $\|\tilde{\bar{C}}_{m,n} - \tilde{\bar{C}}\|_\infty^* \rightarrow 0$ fast sicher für $m, n \rightarrow \infty$.

Ferner ist

$$\begin{aligned} 0 &\leq \int_{\{F<G\}}^* \left(\tilde{C}_{m,n}(\mathbf{t}) + \sqrt{\frac{mn}{m+n}}(F(\mathbf{t}) - G(\mathbf{t})) \right)^+ d\mu(\mathbf{t}) \\ &\leq \int_{\{F<G\}}^* \left(\sup_{\mathbf{s} \in \mathbb{R}^d} \tilde{C}_{m,n}(\mathbf{s}) + \sqrt{\frac{mn}{m+n}}(F(\mathbf{t}) - G(\mathbf{t})) \right)^+ d\mu(\mathbf{t}) \end{aligned}$$

und $\sup_{\mathbf{s} \in \mathbb{R}^d} \tilde{C}_{m,n}(\mathbf{s}) \xrightarrow{f.s.*} \sup_{\mathbf{s} \in \mathbb{R}^d} \tilde{C}(\mathbf{s})$ für $m, n \rightarrow \infty$ (zur Definition von ' $\xrightarrow{f.s.*}$ ', siehe Definition B.1.5). Mit $\sqrt{\frac{mn}{m+n}} \rightarrow \infty$ folgt dann mit dem Satz von der majorisierten Konvergenz für das äußere Integral (siehe auch [VW96], S. 13)

$$\int_{\{F<G\}}^* \left(\sup_{\mathbf{s} \in \mathbb{R}^d} \tilde{C}_{m,n}(\mathbf{s}) + \sqrt{\frac{mn}{m+n}}(F(\mathbf{t}) - G(\mathbf{t})) \right)^+ d\mu(\mathbf{t}) \xrightarrow{f.s.*} 0,$$

für $m, n \rightarrow \infty$. Analog erhalten wir

$$\int_{\{\bar{F} < \bar{G}\}}^* \left(\sup_{\mathbf{s} \in \mathbb{R}^d} \tilde{C}_{m,n}(\mathbf{s}) + \sqrt{\frac{mn}{m+n}} (\bar{F}(\mathbf{t}) - \bar{G}(\mathbf{t})) \right)^+ d\mu(\mathbf{t}) \xrightarrow{f.s.*} 0,$$

für $m, n \rightarrow \infty$.

Die Funktionen $\mathbf{t} \mapsto \tilde{C}_{m,n}^+(\mathbf{t})$ und $\mathbf{t} \mapsto \tilde{\bar{C}}_{m,n}^+(\mathbf{t})$ sind nach Konstruktion f.s. gleichmäßig beschränkt in $m, n \in \mathbb{N}$, so dass wir erneut mit dem Satz von der majorisierten Konvergenz

$$\left(\int_A^* \tilde{C}_{m,n}^+(\mathbf{t}) d\mu(\mathbf{t}) + \int_B^* \tilde{\bar{C}}_{m,n}^+(\mathbf{t}) d\mu(\mathbf{t}) \right) \xrightarrow{f.s.*} \int_A^* \tilde{C}^+(\mathbf{t}) d\mu(\mathbf{t}) + \int_B^* \tilde{\bar{C}}^+(\mathbf{t}) d\mu(\mathbf{t})$$

für $m, n \rightarrow \infty$ erhalten. Zusammengefasst gilt

$$\begin{aligned} & \int_{\{F=G\}}^* \tilde{C}_{m,n}^+(\mathbf{t}) d\mu(\mathbf{t}) + \int_{\{F < G\}}^* \left(\tilde{C}_{m,n}(\mathbf{t}) + \sqrt{\frac{mn}{m+n}} (F(\mathbf{t}) - G(\mathbf{t})) \right)^+ d\mu(\mathbf{t}) \\ & + \int_{\{\bar{F}=\bar{G}\}}^* \tilde{\bar{C}}_{m,n}^+(\mathbf{t}) d\mu(\mathbf{t}) + \int_{\{\bar{F} < \bar{G}\}}^* \left(\tilde{\bar{C}}_{m,n}(\mathbf{t}) + \sqrt{\frac{mn}{m+n}} (\bar{F}(\mathbf{t}) - \bar{G}(\mathbf{t})) \right)^+ d\mu(\mathbf{t}) \\ & \xrightarrow{f.s.*} \int_A^* \tilde{C}^+(\mathbf{t}) d\mu(\mathbf{t}) + \int_B^* \tilde{\bar{C}}^+(\mathbf{t}) d\mu(\mathbf{t}), \end{aligned}$$

für $m, n \rightarrow \infty$. Es folgt

$$T^{CvM}(\mathbb{G}_{m,n}) \xrightarrow{v} \int_A^* C^+(\mathbf{t}) d\mu(\mathbf{t}) + \int_B^* \bar{C}^+(\mathbf{t}) d\mu(\mathbf{t}), \quad m, n \rightarrow \infty.$$

Bei Gültigkeit der Alternative existiert mindestens ein $\mathbf{t}_0 \in \mathbb{R}^d$ mit $F(\mathbf{t}_0) > G(\mathbf{t}_0)$ oder $\bar{F}(\mathbf{t}_0) > \bar{G}(\mathbf{t}_0)$. Aufgrund der Stetigkeit von F und G und der daraus resultierenden Stetigkeit der jeweiligen Überlebensfunktionen gilt diese Beziehung dann in einer Umgebung von \mathbf{t}_0 , so dass für mindestens eines der Integrale

$$\int_{\mathbb{R}^d} \Delta_{m,n}^+(\mathbf{t}) \omega(\mathbf{t}) d\mathbf{t}, \quad \int_{\mathbb{R}^d} \bar{\Delta}_{m,n}^+(\mathbf{t}) \omega(\mathbf{t}) d\mathbf{t},$$

die stochastische Konvergenz gegen unendlich folgt. □

Als Folgerungen ergeben sich die entsprechenden Grenzwertaussagen für die anderen Testgrößen. Wir setzen dazu

$$\bar{A} := \{\mathbf{t} \in \bar{\mathbb{R}}^d : F(\mathbf{t}) = G(\mathbf{t})\}, \quad \bar{B} := \{\mathbf{t} \in \bar{\mathbb{R}}^d : \bar{F}(\mathbf{t}) = \bar{G}(\mathbf{t})\}.$$

Offenbar sind \bar{A} und \bar{B} stets nicht leer und es gilt außerdem

$$\sup_{\mathbf{t} \in \mathbb{R}^d} \Delta_{m,n}(\mathbf{t}) = \sup_{\mathbf{t} \in \bar{\mathbb{R}}^d} \Delta_{m,n}(\mathbf{t}) \quad \text{und} \quad \sup_{\mathbf{t} \in \mathbb{R}^d} \bar{\Delta}_{m,n}(\mathbf{t}) = \sup_{\mathbf{t} \in \bar{\mathbb{R}}^d} \bar{\Delta}_{m,n}(\mathbf{t}).$$

5.2.4. Satz. *Bei Gültigkeit der Hypothese gilt*

$$C_{m,n}^{Kui} \xrightarrow{v} C^{Kui} := \sup_{\mathbf{t} \in A} C(\mathbf{t}) + \sup_{\mathbf{t} \in B} \bar{C}(\mathbf{t}), \quad m, n \rightarrow \infty.$$

Bei Gültigkeit der Alternative ist

$$C_{m,n}^{Kui} \xrightarrow{P} \infty, \quad m, n \rightarrow \infty.$$

Beweis. Es seien $C_{m,n}(\mathbf{t}), \bar{C}_{m,n}(\mathbf{t})$ wie im Beweis des vorherigen Satzes und

$$\begin{aligned} T^{Kui} : l^\infty(\mathcal{F}) &\longrightarrow \mathbb{R} \\ x &\mapsto \sup_{\mathcal{F}_1} x + \sup_{\mathcal{F}_2} x. \end{aligned}$$

Dann ist $C_{m,n}^{Kui} = T^{Kui}(\mathbb{G}_{m,n})$ und es gilt

$$\begin{aligned} T^{Kui}(\mathbb{G}_{m,n}) &= \sup_{f \in \mathcal{F}_1} \mathbb{G}_{m,n}f + \sup_{f \in \mathcal{F}_2} \mathbb{G}_{m,n}f \\ &= \sup_{\mathbf{t} \in \mathbb{R}^d} \left(C_{m,n}(\mathbf{t}) + \sqrt{\frac{mn}{m+n}} (F(\mathbf{t}) - G(\mathbf{t})) \right) \\ &\quad + \sup_{\mathbf{t} \in \mathbb{R}^d} \left(\bar{C}_{m,n}(\mathbf{t}) + \sqrt{\frac{mn}{m+n}} (\bar{F}(\mathbf{t}) - \bar{G}(\mathbf{t})) \right) \end{aligned}$$

Die auftretenden (empirischen) Verteilungs- und Überlebensfunktionen lassen sich in natürlicher Weise auf $\bar{\mathbb{R}}^d$ fortsetzen. Mit dieser natürlichen Fortsetzung auf $\bar{\mathbb{R}}^d$ ist aufgrund von Bemerkung 5.2.2 offensichtlich, dass der entsprechend fortgesetzte Prozess $\bar{\mathbb{G}}_{m,n} := (\mathbb{G}_{m,n}f, f \in \bar{\mathcal{F}})$ außerdem in $l^\infty(\bar{\mathbb{R}}^d)$ in Verteilung gegen einen straffen Grenzwert konvergiert. Die auf $\bar{\mathbb{R}}^d$ fortgesetzten Prozesse $(C_{m,n}(\mathbf{t}), \mathbf{t} \in \bar{\mathbb{R}}^d)$ und $(\bar{C}_{m,n}(\mathbf{t}), \mathbf{t} \in \bar{\mathbb{R}}^d)$ konvergieren dann als entsprechende Einschränkungen von $\bar{\mathbb{G}}_{m,n}$ auf die Teilmengen $\bar{\mathcal{F}}_1$ und $\bar{\mathcal{F}}_2$ ebenfalls in $l^\infty(\bar{\mathbb{R}}^d)$ gegen die straffen und Borel-messbaren Grenzwerte $(C(\mathbf{t}), \mathbf{t} \in \bar{\mathbb{R}}^d)$ und $(\bar{C}(\mathbf{t}), \mathbf{t} \in \bar{\mathbb{R}}^d)$. Der Grenzprozess C hat wegen Satz B.2.5 fast sicher gleichmäßig stetige Pfade bezüglich der Semimetrik

$$\rho_2(\mathbf{s}, \mathbf{t}) = (E(C(\mathbf{s}) - C(\mathbf{t}))^2)^{\frac{1}{2}}, \quad \mathbf{s}, \mathbf{t} \in \bar{\mathbb{R}}^d,$$

auf $\bar{\mathbb{R}}^d$.

Die Kovarianzfunktion von C aus Satz 5.2.3 ist aufgrund der Stetigkeit von F und G stetig auf $\bar{\mathbb{R}}^d \times \bar{\mathbb{R}}^d$. Die Stetigkeit der Pfade bezüglich der Semimetrik ρ_2 und die Stetigkeit der Kovarianzfunktion liefern zusammen, dass die Pfade $\mathbf{t} \mapsto C(\mathbf{t})$ fast sicher stetige Funktionen auf $\bar{\mathbb{R}}^d$ sind (vergleiche dazu [Adl90], S. 2-3). Analog lässt sich für den Grenzprozess \bar{C} argumentieren.

Erneut nach Satz B.1.6 finden wir dann auf einem geeigneten Wahrscheinlichkeitsraum $l^\infty(\overline{\mathbb{R}}^d)$ -wertige Abbildungen $\tilde{C}_{m,n}, \tilde{C}$ mit der Eigenschaft $\tilde{C}_{m,n} \sim C_{m,n}, \tilde{C} \sim C$ und $\|\tilde{C}_{m,n} - \tilde{C}\|_\infty^* \rightarrow 0$ fast sicher, sowie Abbildungen $\tilde{\bar{C}}_{m,n}, \tilde{\bar{C}}$ mit $\tilde{\bar{C}}_{m,n} \sim \bar{C}_{m,n}, \tilde{\bar{C}} \sim \bar{C}$ und $\|\tilde{\bar{C}}_{m,n} - \tilde{\bar{C}}\|_\infty^* \rightarrow 0$ fast sicher. Da die fast sicher konvergenten Versionen $\tilde{C}_{m,n}$ und $\tilde{\bar{C}}$ nach Satz B.1.6 so gewählt werden können, dass sie keine anderen Werte als $C_{m,n}$ und C annehmen, übertragen sich die zuvor genannten Stetigkeitseigenschaften auf die Pfade von $\tilde{C}_{m,n}, \tilde{C}$.

Bei Gültigkeit der Hypothese gilt $F(\mathbf{t}) - G(\mathbf{t}) \leq 0, \mathbf{t} \in \overline{\mathbb{R}}^d$, und es ist etwa $\infty \cdot \mathbf{1} \in \{\mathbf{t} \in \overline{\mathbb{R}}^d : F(\mathbf{t}) = G(\mathbf{t})\}$. Wir wenden Satz D.3.2 an und erhalten

$$\sup_{\mathbf{t} \in \overline{\mathbb{R}}^d} \left(\tilde{C}_{m,n}(\mathbf{t}) + \sqrt{\frac{mn}{m+n}} (F(\mathbf{t}) - G(\mathbf{t})) \right) \xrightarrow{f.s.*} \sup_{\{\mathbf{t} \in \overline{\mathbb{R}}^d : F(\mathbf{t}) = G(\mathbf{t})\}} \tilde{C}(\mathbf{t}),$$

für $m, n \rightarrow \infty$. Analog folgern wir

$$\sup_{\mathbf{t} \in \overline{\mathbb{R}}^d} \left(\tilde{\bar{C}}_{m,n}(\mathbf{t}) + \sqrt{\frac{mn}{m+n}} (\bar{F}(\mathbf{t}) - \bar{G}(\mathbf{t})) \right) \xrightarrow{f.s.*} \sup_{\{\mathbf{t} \in \overline{\mathbb{R}}^d : \bar{F}(\mathbf{t}) = \bar{G}(\mathbf{t})\}} \tilde{\bar{C}}(\mathbf{t}),$$

für $m, n \rightarrow \infty$ und somit

$$C_{m,n}^{Kui} \xrightarrow{v} \sup_{\{\mathbf{t} \in \overline{\mathbb{R}}^d : F(\mathbf{t}) = G(\mathbf{t})\}} C(\mathbf{t}) + \sup_{\{\mathbf{t} \in \overline{\mathbb{R}}^d : \bar{F}(\mathbf{t}) = \bar{G}(\mathbf{t})\}} \bar{C}(\mathbf{t}), \quad m, n \rightarrow \infty.$$

Bei Gültigkeit der Alternative finden wir wieder ein \mathbf{t}_0 mit $F(\mathbf{t}_0) > G(\mathbf{t}_0)$ oder $\bar{F}(\mathbf{t}_0) > \bar{G}(\mathbf{t}_0)$, so dass mindestens einer der Summanden $\sup_{\mathbf{t} \in \overline{\mathbb{R}}^d} \Delta_{m,n}(\mathbf{t})$ oder $\sup_{\mathbf{t} \in \overline{\mathbb{R}}^d} \bar{\Delta}_{m,n}(\mathbf{t})$ für $m, n \rightarrow \infty$ gegen unendlich konvergiert. \square

Schließlich erhalten wir für den Kolmogorov-Smirnov Fall die folgenden Grenzwertaussagen:

5.2.5. Satz. *Bei Gültigkeit der Hypothese gilt*

$$C_{m,n}^{KS} \xrightarrow{v} C^{KS} := \max \left(\sup_{\mathbf{t} \in \bar{A}} C(\mathbf{t}), \sup_{\mathbf{t} \in \bar{B}} \bar{C}(\mathbf{t}) \right), \quad m, n \rightarrow \infty.$$

Bei Gültigkeit der Alternative ist

$$C_{m,n}^{KS} \xrightarrow{P} \infty, \quad m, n \rightarrow \infty.$$

Beweis. Es bezeichne $x \vee y$ das Maximum von x und y . Aus dem Beweis zum vorherigen Satz, ergibt sich dann bei Gültigkeit der Hypothese sofort

$$\sup_{\mathbf{t} \in \mathbb{R}^d} \left(\tilde{C}_{m,n}(\mathbf{t}) + \sqrt{\frac{mn}{m+n}} (F(\mathbf{t}) - G(\mathbf{t})) \right) \vee \sup_{\mathbf{t} \in \mathbb{R}^d} \left(\tilde{\bar{C}}_{m,n}(\mathbf{t}) + \sqrt{\frac{mn}{m+n}} (\bar{F}(\mathbf{t}) - \bar{G}(\mathbf{t})) \right) \\ \xrightarrow{f.s.*} \max \left(\sup_{\mathbf{t} \in \bar{A}} \tilde{C}(\mathbf{t}), \sup_{\mathbf{t} \in \bar{B}} \tilde{\bar{C}}(\mathbf{t}) \right),$$

für $m, n \rightarrow \infty$. Daraus folgt ebenfalls bei Gültigkeit der Hypothese

$$\max \left(\sup_{\mathbf{t} \in \mathbb{R}^d} \Delta_{m,n}(\mathbf{t}), \sup_{\mathbf{t} \in \mathbb{R}^d} \bar{\Delta}_{m,n}(\mathbf{t}) \right) \xrightarrow{v} \max \left(\sup_{\mathbf{t} \in \bar{A}} C(\mathbf{t}), \sup_{\mathbf{t} \in \bar{B}} \bar{C}(\mathbf{t}) \right), \quad m, n \rightarrow \infty,$$

und daher die Behauptung. □

5.2.6. Bemerkung. Die Ergebnisse dieses Abschnitts zur Asymptotik der Testgrößen ergeben sich alternativ auch durch alleinige Betrachtung der Asymptotik des empirischen Prozesses $\mathbb{G}_{m,n}|_{\mathcal{F}_1}$ und unter Zuhilfenahme der Darstellung der Überlebensfunktion mittels der Verteilungsfunktion in (4.1.1). Mit ähnlichen Argumenten lässt sich die empirische Überlebensfunktion mittels der empirischen Verteilungsfunktion ausdrücken. Demnach ergibt sich der Prozess $(\bar{\Delta}_{m,n}(\mathbf{t}), \mathbf{t} \in \mathbb{R}^d)$ als Funktion des Prozesses $(\Delta_{m,n}(\mathbf{t}), \mathbf{t} \in \mathbb{R}^d)$.

5.3 Konsistenz und Einhaltung des Testniveaus

Analog zum Test auf Laplace-Ordnung wollen wir auch für die Tests auf Konkordanz-Ordnung ein Bootstrapverfahren angeben, um die zur Durchführung des Tests nötigen kritischen Werte zu simulieren. Sei dazu

$$H_{m,n}(\mathbf{t}) := \frac{n}{m+n} F_m(\mathbf{t}) + \frac{m}{m+n} G_n(\mathbf{t}), \quad \mathbf{t} \in \mathbb{R}^d,$$

und bei gegebenen $\mathbf{X}_1, \dots, \mathbf{X}_m, \mathbf{Y}_1, \dots, \mathbf{Y}_m$ seien $\hat{\mathbf{Z}}_{N1}, \dots, \hat{\mathbf{Z}}_{NN}$, $N = m+n$, unabhängige Zufallsvariablen mit der Verteilungsfunktion $H_{m,n}$. Es sei $H(\mathbf{t}) := (1 - \tau)F(\mathbf{t}) + \tau G(\mathbf{t})$, $\mathbf{t} \in \mathbb{R}^d$. Weiter setzen wir

$$\hat{F}_m(\mathbf{t}) = \frac{1}{m} \sum_{\nu=1}^m \mathbf{1}_{(-\infty, \mathbf{t}]}(\hat{\mathbf{Z}}_{N\nu}), \quad \hat{G}_n(\mathbf{t}) = \frac{1}{n} \sum_{\nu=m+1}^N \mathbf{1}_{(-\infty, \mathbf{t}]}(\hat{\mathbf{Z}}_{N\nu})$$

und

$$\hat{\bar{F}}_m(\mathbf{t}) = \frac{1}{m} \sum_{\nu=1}^m \mathbf{1}_{(\mathbf{t}, \infty)}(\hat{\mathbf{Z}}_{N\nu}), \quad \hat{\bar{G}}_n(\mathbf{t}) = \frac{1}{n} \sum_{\nu=m+1}^N \mathbf{1}_{(\mathbf{t}, \infty)}(\hat{\mathbf{Z}}_{N\nu}).$$

Für $\mathbf{t} \in \mathbb{R}^d$ sei außerdem

$$\hat{\Delta}_{m,n}(\mathbf{t}) := \sqrt{\frac{mn}{m+n}} \left(\hat{F}_m(\mathbf{t}) - \hat{G}_n(\mathbf{t}) \right), \quad \hat{\bar{\Delta}}_{m,n}(\mathbf{t}) := \sqrt{\frac{mn}{m+n}} \left(\hat{\bar{F}}_m(\mathbf{t}) - \hat{\bar{G}}_n(\mathbf{t}) \right)$$

und für $m, n \in \mathbb{N}$ definieren wir

$$\begin{aligned} \hat{C}_{m,n}^{Kui} &:= \sup_{\mathbf{t} \in \mathbb{R}^d} \hat{\Delta}_{m,n}(\mathbf{t}) + \sup_{\mathbf{t} \in \mathbb{R}^d} \hat{\bar{\Delta}}_{m,n}(\mathbf{t}), \\ \hat{C}_{m,n}^{KS} &:= \max \left(\sup_{\mathbf{t} \in \mathbb{R}^d} \hat{\Delta}_{m,n}(\mathbf{t}), \sup_{\mathbf{t} \in \mathbb{R}^d} \hat{\bar{\Delta}}_{m,n}(\mathbf{t}) \right), \\ \hat{C}_{m,n}^{CvM} &:= \int_{\mathbb{R}^d} \left(\hat{\Delta}_{m,n}^+(\mathbf{t}) + \hat{\bar{\Delta}}_{m,n}^+(\mathbf{t}) \right) \omega(\mathbf{t}) dt. \end{aligned}$$

Das Verhalten der empirischen Prozesse $(\hat{\Delta}_{m,n}(\mathbf{t}), \mathbf{t} \in \mathbb{R}^d)$ und $(\hat{\bar{\Delta}}_{m,n}(\mathbf{t}), \mathbf{t} \in \mathbb{R}^d)$ für wachsende $m, n \in \mathbb{N}$ liefert die folgenden Grenzwertaussagen für die Testgrößen unter dem Einfluss des Bootstrapverfahrens. Ein Vergleich dieser Aussagen mit den Resultaten des letzten Abschnittes zeigt, dass die über das vorgeschlagene Bootstrapverfahren erhaltenen kritischen Werte verwendet werden können, um eine Einhaltung des Testniveaus zu gewährleisten (vergleiche auch Bemerkung 2.3.2 für den Fall der Laplace-Ordnung).

5.3.1. Satz. *Es sei $(\hat{C}f, f \in \mathcal{F})$ ein zentrierter Gaußprozess mit Pfaden in $l^\infty(\mathcal{F})$ und der Kovarianzfunktion $\text{Cov}(\hat{C}f, \hat{C}g) = Hfg - HfHg$, $f, g \in \mathcal{F}$. Sei*

$$C_H = (C_H(\mathbf{t}), \mathbf{t} \in \mathbb{R}^d), \quad C_H(\mathbf{t}) := \hat{C}\mathbf{1}_{(-\infty, \mathbf{t}]}, \quad \mathbf{t} \in \mathbb{R}^d,$$

und

$$\bar{C}_H = (\bar{C}_H(\mathbf{t}), \mathbf{t} \in \mathbb{R}^d), \quad \bar{C}_H(\mathbf{t}) := \hat{C}\mathbf{1}_{(\mathbf{t}, \infty)}, \quad \mathbf{t} \in \mathbb{R}^d.$$

Dann ist $(C_H(\mathbf{t}), \mathbf{t} \in \mathbb{R}^d)$ ein zentrierter Gaußprozess mit fast sicher beschränkten Pfaden und der Kovarianzfunktion $\rho_{C_H}(\mathbf{s}, \mathbf{t}) := H(\min(\mathbf{s}, \mathbf{t})) - H(\mathbf{s})H(\mathbf{t})$ und $(\bar{C}_H(\mathbf{t}), \mathbf{t} \in \mathbb{R}^d)$ ist ein zentrierter Gaußprozess mit fast sicher beschränkten Pfaden und der Kovarianzfunktion $\rho_{\bar{C}_H}(\mathbf{s}, \mathbf{t}) = \bar{H}(\max(\mathbf{s}, \mathbf{t})) - \bar{H}(\mathbf{s})\bar{H}(\mathbf{t})$ und es gilt für fast alle $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{Y}_1, \mathbf{Y}_2, \dots$

a)

$$\hat{C}_{m,n}^{Kui} \xrightarrow{v} \hat{C}^{Kui} := \sup_{\mathbf{t} \in \mathbb{R}^d} C_H(\mathbf{t}) + \sup_{\mathbf{t} \in \mathbb{R}^d} \bar{C}_H(\mathbf{t}), \quad m, n \rightarrow \infty.$$

b)

$$\hat{C}_{m,n}^{KS} \xrightarrow{v} \hat{C}^{KS} := \max \left(\sup_{\mathbf{t} \in \mathbb{R}^d} C_H(\mathbf{t}), \sup_{\mathbf{t} \in \mathbb{R}^d} \bar{C}_H(\mathbf{t}) \right), \quad m, n \rightarrow \infty.$$

c)

$$\hat{C}_{m,n}^{CvM} \xrightarrow{v} \hat{C}^{CvM} := \int_{\mathbb{R}^d} (C_H^+(\mathbf{t}) + \overline{C}_H^+(\mathbf{t})) \omega(\mathbf{t}) dt, \quad m, n \rightarrow \infty.$$

Beweis.

Wir definieren bei gegebenen $\mathbf{X}_1, \dots, \mathbf{X}_m, \mathbf{Y}_1, \dots, \mathbf{Y}_n$ die empirischen Maße $\hat{P}_m = \frac{1}{m} \sum_{\nu=1}^m \delta_{\hat{\mathbf{Z}}_{N\nu}}$ und $\hat{Q}_n = \frac{1}{n} \sum_{\nu=m+1}^N \delta_{\hat{\mathbf{Z}}_{N\nu}}$ und den Prozess

$$\begin{aligned} \hat{\mathbb{G}}_{m,n} &= (\hat{\mathbb{G}}_{m,n} f, f \in \mathcal{F}) := \left(\sqrt{\frac{mn}{m+n}} (\hat{P}_m - \hat{Q}_n) f, f \in \mathcal{F} \right) \\ &= \left(\sqrt{\frac{mn}{m+n}} (\hat{P}_m - H_{m,n}) f - \sqrt{\frac{mn}{m+n}} (\hat{Q}_n - H_{m,n}) f, f \in \mathcal{F} \right). \end{aligned}$$

Wegen $|f| \leq 1$ für alle $f \in \mathcal{F}$ ist $E \equiv 1$ eine Envelope von \mathcal{F} und nach Lemma 5.2.1 erfüllt \mathcal{F} daher die Entropiebedingung (B.2.1). In Analogie zu Satz 2.3.1 und im Hinblick auf Satz B.2.2.1 und Bemerkung B.2.2.2 untersuchen wir den Abstand

$$\begin{aligned} &\sup_{f,g \in \mathcal{F}} |L_2(H_{m,n})(f-g) - L_2(H)(f-g)| \\ &= \sup_{f,g \in \mathcal{F}} \left| \left(\int_{\mathbb{R}^d} (f(\mathbf{x}) - g(\mathbf{x}))^2 dH_{m,n}(\mathbf{x}) \right)^{\frac{1}{2}} - \left(\int_{\mathbb{R}^d} (f(\mathbf{x}) - g(\mathbf{x}))^2 dH(\mathbf{x}) \right)^{\frac{1}{2}} \right|, \end{aligned}$$

und zeigen, dass für fast alle $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{Y}_1, \mathbf{Y}_2, \dots$

$$\lim_{m,n \rightarrow \infty} \sup_{f,g \in \mathcal{F}} |L_2(H_{m,n})(f-g) - L_2(H)(f-g)| = 0$$

gilt. Die Funktionen $f, g \in \mathcal{F}$ sind Indikatorfunktionen. Mit $|\mathbf{1}_A - \mathbf{1}_B| = \mathbf{1}_{A \Delta B}$ für beliebige Ereignisse A, B , folgt daher $(\mathbf{1}_A - \mathbf{1}_B)^2 = \mathbf{1}_{A \Delta B}^2 = \mathbf{1}_{A \Delta B}$. Ist $\hat{\mathbf{Z}}$ eine Zufallsvariable mit der Verteilungsfunktion H , ergibt sich daher

$$\begin{aligned} \sup_{f,g \in \mathcal{F}} |L_2(H_{m,n})(f-g) - L_2(H)(f-g)| &= \sup_{A \in \mathfrak{D}} \left| P(\hat{\mathbf{Z}}_{N1} \in A)^{\frac{1}{2}} - P(\hat{\mathbf{Z}} \in A)^{\frac{1}{2}} \right| \\ &\leq \left(\sup_{A \in \mathfrak{D}} \left| P(\hat{\mathbf{Z}}_{N1} \in A) - P(\hat{\mathbf{Z}} \in A) \right| \right)^{\frac{1}{2}}, \end{aligned}$$

mit

$$\mathfrak{D} = \left\{ A \Delta B : A, B \in \{(-\infty, \mathbf{t}], \mathbf{t} \in \mathbb{R}^d\} \cup \{(\mathbf{t}, \infty), \mathbf{t} \in \mathbb{R}^d\} \right\}.$$

Für ein beliebiges $A \in \mathcal{B}(\mathbb{R}^d)$ gilt bei gegebenen $\mathbf{X}_1, \dots, \mathbf{X}_m, \mathbf{Y}_1, \dots, \mathbf{Y}_n$

$$P(\hat{\mathbf{Z}}_{N1} \in A) = \frac{n}{m+n} \frac{1}{m} \sum_{\nu=1}^m \mathbf{1}_A(\mathbf{X}_\nu) + \frac{m}{m+n} \frac{1}{n} \sum_{\nu=1}^n \mathbf{1}_A(\mathbf{Y}_\nu)$$

und

$$P(\hat{\mathbf{Z}} \in A) = (1 - \tau)P(\mathbf{X}_1 \in A) + \tau P(\mathbf{Y}_1 \in A).$$

Das Mengensystem \mathfrak{D} ist eine VC-Klasse (vergleiche etwa [VW96], Lemma 2.6.17 i)-iii)) und die Menge der Indikatorfunktionen $\{\mathbf{1}_A, A \in \mathfrak{D}\}$ somit eine VC-Funktionenklasse. Der Satz von Gilvenko Cantelli in der Form von Satz B.2.1.6 liefert dann, dass für fast alle $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{Y}_1, \mathbf{Y}_2, \dots$

$$\begin{aligned} \sup_{A \in \mathfrak{D}} \left| P(\hat{\mathbf{Z}}_{N_1} \in A) - P(\hat{\mathbf{Z}} \in A) \right| &\leq \sup_{A \in \mathfrak{D}} \left| \frac{n}{m+n} \frac{1}{m} \sum_{\nu=1}^m \mathbf{1}_A(\mathbf{X}_\nu) - (1 - \tau)P(\mathbf{X}_1 \in A) \right| \\ &\quad + \sup_{A \in \mathfrak{D}} \left| \frac{m}{m+n} \frac{1}{n} \sum_{\nu=1}^n \mathbf{1}_A(\mathbf{Y}_\nu) - \tau P(\mathbf{Y}_1 \in A) \right| \\ &\longrightarrow 0, \quad m, n \rightarrow \infty, \end{aligned}$$

gilt. Mit Satz B.2.2.1 und den Vorbetrachtungen ergibt sich daraus schließlich für $m, n \rightarrow \infty$ und fast alle $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{Y}_1, \mathbf{Y}_2, \dots$

$$\left(\sqrt{\frac{mn}{m+n}} \left(\hat{P}_m - H_{m,n} \right) f, f \in \mathcal{F} \right) \xrightarrow{v} \left((1 - \tau) \hat{C}f, f \in \mathcal{F} \right), \quad m, n \rightarrow \infty,$$

mit dem zentrierten Gaußprozess $(\hat{C}f, f \in \mathcal{F})$ mit Pfaden in $l^\infty(\mathcal{F})$ und der Kovarianzfunktion $\text{Cov}(\hat{C}f, \hat{C}g) = Hfg - HfHg, f, g \in \mathcal{F}$. Analoge Argumente führen auf

$$\left(\sqrt{\frac{mn}{m+n}} \left(\hat{Q}_n - H_{m,n} \right) f, f \in \mathcal{F} \right) \xrightarrow{v} \left(\tau \hat{C}f, f \in \mathcal{F} \right), \quad m, n \rightarrow \infty,$$

für fast alle $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{Y}_1, \mathbf{Y}_2, \dots$. Aufgrund der bedingten Unabhängigkeit der empirischen Prozesse bei gegebenen $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{Y}_1, \mathbf{Y}_2, \dots$ gilt daher $\hat{\mathbb{G}}_{m,n} \xrightarrow{v} \hat{C}$ in $l^\infty(\mathcal{F})$ für $m, n \rightarrow \infty$ und fast alle $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{Y}_1, \mathbf{Y}_2, \dots$.

Eingeschränkt auf \mathcal{F}_1 hat \hat{C} dieselbe Kovarianzfunktion wie C_H und eingeschränkt auf \mathcal{F}_2 hat \hat{C} dieselbe Kovarianzfunktion wie \overline{C}_H . Um die behaupteten Aussagen für $\hat{C}_{m,n}^{Kui}, \hat{C}_{m,n}^{KS}$ und $\hat{C}_{m,n}^{CvM}$ zu erhalten, gehen wir schließlich wie in den Beweisen zu den Sätzen 5.2.3 - 5.2.5 vor, indem wir die entsprechenden Funktionale auf fast sicher konvergente Versionen der Prozesse $\hat{\mathbb{G}}_{m,n}$ und \hat{C} anwenden. \square

Um Resultate bezüglich der Konsistenz der vorgestellten Verfahren zu erhalten, benötigen wir wieder analytische Aussagen über die Verteilungsfunktionen der Grenzprozesse unter dem Einfluss des Bootstrapverfahrens. Begründet durch die Ergebnisse des vorherigen

Satzes definieren wir die Tests

$$\varphi_{m,n}^{KS}(\mathbf{X}_1, \dots, \mathbf{X}_m, \mathbf{Y}_1, \dots, \mathbf{Y}_n) = \begin{cases} 1, & C_{m,n}^{KS} > \hat{c}_{m,n}^{KS}(\alpha), \\ 0, & C_{m,n}^{KS} \leq \hat{c}_{m,n}^{KS}(\alpha), \end{cases}$$

$$\varphi_{m,n}^{Kui}(\mathbf{X}_1, \dots, \mathbf{X}_m, \mathbf{Y}_1, \dots, \mathbf{Y}_n) = \begin{cases} 1, & C_{m,n}^{Kui} > \hat{c}_{m,n}^{Kui}(\alpha), \\ 0, & C_{m,n}^{Kui} \leq \hat{c}_{m,n}^{Kui}(\alpha), \end{cases}$$

und

$$\varphi_{m,n}^{CvM}(\mathbf{X}_1, \dots, \mathbf{X}_m, \mathbf{Y}_1, \dots, \mathbf{Y}_n) = \begin{cases} 1, & C_{m,n}^{CvM} > \hat{c}_{m,n}^{CvM}(\alpha), \\ 0, & C_{m,n}^{CvM} \leq \hat{c}_{m,n}^{CvM}(\alpha). \end{cases}$$

mit den jeweiligen $(1 - \alpha)$ -Quantilen

$$\hat{c}_{m,n}^{KS}(\alpha) := \inf\{c \in \mathbb{R} : P(\hat{C}_{m,n}^{KS} > c) \leq \alpha\}, \quad \alpha \in (0, 1),$$

$$\hat{c}^{KS}(\alpha) := \inf\{c \in \mathbb{R} : P(\hat{C}^{KS} > c) \leq \alpha\}, \quad \alpha \in (0, 1),$$

sowie

$$\hat{c}_{m,n}^{Kui}(\alpha) := \inf\{c \in \mathbb{R} : P(\hat{C}_{m,n}^{Kui} > c) \leq \alpha\}, \quad \alpha \in (0, 1),$$

$$\hat{c}^{Kui}(\alpha) := \inf\{c \in \mathbb{R} : P(\hat{C}^{Kui} > c) \leq \alpha\}, \quad \alpha \in (0, 1),$$

und

$$\hat{c}_{m,n}^{CvM}(\alpha) := \inf\{c \in \mathbb{R} : P(\hat{C}_{m,n}^{CvM} > c) \leq \alpha\}, \quad \alpha \in (0, 1),$$

$$\hat{c}^{CvM}(\alpha) := \inf\{c \in \mathbb{R} : P(\hat{C}^{CvM} > c) \leq \alpha\}, \quad \alpha \in (0, 1).$$

Die folgenden Sätze liefern die gewünschten Eigenschaften der Tests.

5.3.2. Satz. Für $\alpha \in (0, \frac{1}{2})$ ist $\varphi_{m,n}^{KS}$ ein konsistenter Test zum Niveau α für das Testproblem (5.1.1).

Beweis.

Wir betrachten wieder den Gaußprozess $(\hat{C}f, f \in \mathcal{F})$ aus Satz 5.3.1 mit der Kovarianzfunktion $\text{Cov}(\hat{C}f, \hat{C}g) = Hfg - HfHg$, $f, g \in \mathcal{F}$. Da sich \hat{C} als Grenzwert eines durch eine VC-Klasse indizierten empirischen Prozesses ergibt, ist \hat{C} eine straffe und Borel-messbare Zufallsgröße mit Werten in $l^\infty(\mathcal{F})$. Die Parametermenge \mathcal{F} ist nach Satz B.2.5 totalbeschränkt bezüglich der Semimetrik ρ_2 und die Pfade von \hat{C} sind fast sicher gleichmäßig

stetig bezüglich ρ_2 . Wir können \hat{C} somit als Zufallsgröße mit Werten in $(C^u(\mathcal{F}), \|\cdot\|_\infty)$ auffassen. Nach Satz D.2.2 ist \hat{C} ein separabler Gaußprozess. Ferner gilt

$$\hat{C}^{KS} = \max \left(\sup_{f \in \mathcal{F}_1} \hat{C}f, \sup_{f \in \mathcal{F}_2} \hat{C}f \right) = \sup_{f \in \mathcal{F}} \hat{C}f.$$

Aufgrund der Separabilität von \hat{C} können wir dieses Supremum als Supremum von abzählbar vielen Zufallsvariablen ansehen, wobei die Varianzen $\text{Var}(\hat{C}f)$, $f \in \mathcal{F}$, aufgrund der vorausgesetzten Nichtdegeneriertheit von F und G nicht sämtlich Null sind. Die Anwendung von Satz D.1.1 gibt dann die Stetigkeit der Verteilungsfunktion von \hat{C}^{KS} im Bereich (a_0, ∞) , mit

$$a_0 := \inf \left\{ x \in \mathbb{R} : P \left(\sup_{f \in \mathcal{F}} \hat{C}f < x \right) > 0 \right\}.$$

Außerdem ist die Wahrscheinlichkeitsdichte von \hat{C}^{KS} nach Satz D.1.1 in diesem Bereich positiv.

Wir verfahren wie im Beweis zu Satz 2.3.4: Es ist $(C^u(\mathcal{F}), \|\cdot\|_\infty)$ ein separabler Banachraum (siehe Satz D.3.4). Mit Satz D.1.2 erhalten wir dann, dass die Verteilungsfunktion der Norm $\|\hat{C}\|_\infty = \sup_{f \in \mathcal{F}} |\hat{C}f|$ streng monoton wachsend auf $[0, \infty)$ ist. Damit ergibt sich

$$P \left(\|\hat{C}\|_\infty < x \right) > 0,$$

für alle $x > 0$. Mit $\hat{C}^{KS} \leq_{st} \|\hat{C}\|_\infty$ folgt daraus zunächst $a_0 \leq 0$ und zusammen mit $P(\hat{C}^{KS} \geq 0) = 1$ also $a_0 = 0$ und die Stetigkeit der Verteilungsfunktion von \hat{C}^{KS} im Bereich $(0, \infty)$. Wegen $\hat{C}_{m,n}^{KS} \xrightarrow{v} \sup_{f \in \mathcal{F}} \hat{C}f$ für $m, n \rightarrow \infty$ und fast alle $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{Y}_1, \mathbf{Y}_2, \dots$, sowie

$$\forall a \in (0, 1) \forall m, n \in \mathbb{N} : \hat{c}_{m,n}^{KS} \geq 0, \quad (5.3.1)$$

für fast alle $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{Y}_1, \mathbf{Y}_2, \dots$, ergibt sich aus der strengen Monotonie der Verteilungsfunktion von \hat{C}^{KS} auf $[0, \infty)$, dass für fast alle $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{Y}_1, \mathbf{Y}_2, \dots$

$$\lim_{m, n \rightarrow \infty} \hat{c}_{m,n}^{KS}(\alpha) = \hat{c}^{KS}(\alpha), \quad (5.3.2)$$

ist. Das Verhalten der Testgröße $C_{m,n}^{KS}$ bei Gültigkeit der Alternative liefert somit die Konsistenz von $\varphi_{m,n}^{KS}$ gegen alle Alternativen.

Sei $\alpha \in (0, \frac{1}{2})$. Dann ist $\hat{c}^{KS}(\alpha) > 0$, da die Dichte von \hat{C}^{KS} nach Satz D.1.1 nach unten durch die Dichte einer normalverteilten Zufallsvariablen mit nicht-negativem Erwartungswert beschränkt ist. Wir betrachten dann den Grenzwert

$$C^{KS} = \max \left(\sup_{t \in A} C(t), \sup_{t \in B} \bar{C}(t) \right) \quad (5.3.3)$$

der Kolmogorov-Smirnov-Testgröße aus Satz 5.2.5 bei Gültigkeit der Hypothese. Ist $C^{KS} \equiv 0$ degeneriert, so hat C^{KS} die Verteilungsfunktion $P(C^{KS} \leq x) = \mathbf{1}_{[0, \infty)}(x)$, $x \in \mathbb{R}$, so dass aufgrund der fast sicheren Konvergenz der Quantile $\hat{c}_{m,n}^{KS}(\alpha)$ gegen einen positiven Grenzwert

$$\lim_{m,n \rightarrow \infty} P(C_{m,n}^{KS} > \hat{c}_{m,n}^{KS}(\alpha) | H) = P(C^{KS} > \hat{c}^{KS}(\alpha)) = 0,$$

für alle $\alpha \in (0, \frac{1}{2})$ folgt. Ist die Verteilungsfunktion von C^{KS} nicht-degeneriert, so ist C^{KS} wegen (5.3.3) von der Form

$$C^{KS} = \sup_{f \in \mathcal{F}^*} \mathcal{C}f,$$

mit $\mathcal{F}^* = \{f \in \mathcal{F} : Pf = Qf\}$. Ein Separabilitätsargument für $(\mathcal{C}f, f \in \mathcal{F})$ und Satz D.1.1 liefern dann analog zu oben die Stetigkeit der Verteilungsfunktion von C^{KS} auf $(0, \infty)$. Wegen $\hat{c}^{KS}(\alpha) > 0$ für $\alpha \in (0, \frac{1}{2})$ und (5.3.2) ergibt sich somit erneut

$$\lim_{m,n \rightarrow \infty} P(C_{m,n}^{KS} > \hat{c}_{m,n}^{KS}(\alpha) | H) = P(C^{KS} > \hat{c}^{KS}(\alpha)) \leq P(\hat{C}^{KS} > \hat{c}^{KS}(\alpha)) = \alpha,$$

für alle $\alpha \in (0, \frac{1}{2})$. □

5.3.3. Satz. Für $\alpha \in (0, \frac{1}{2})$ ist $\varphi_{m,n}^{Kui}$ ein konsistenter Test zum Niveau α für das Testproblem (5.1.1).

Beweis. Wir betrachten erneut den Gaußprozess $(\hat{\mathcal{C}}f, f \in \mathcal{F})$ und dessen Einschränkungen auf die Bereiche \mathcal{F}_1 und \mathcal{F}_2 . Dann ist der Prozess $\hat{\mathcal{C}}|_{\mathcal{F}_1} + \hat{\mathcal{C}}|_{\mathcal{F}_2} := (\hat{\mathcal{C}}f + \hat{\mathcal{C}}g, (f, g) \in \mathcal{F}_1 \times \mathcal{F}_2)$ ein straffer Gaußprozess, der wegen Satz B.2.5 als Zufallsvariable mit dem Wertebereich $C^u(\mathcal{F}_1 \times \mathcal{F}_2)$ interpretiert werden kann. Es ist

$$\hat{C}^{Kui} = \sup_{f \in \mathcal{F}_1} \hat{\mathcal{C}}f + \sup_{g \in \mathcal{F}_2} \hat{\mathcal{C}}g = \sup_{(f,g) \in \mathcal{F}_1 \times \mathcal{F}_2} (\hat{\mathcal{C}}f + \hat{\mathcal{C}}g),$$

und dieses Supremum kann aufgrund der Separabilität von $(\hat{\mathcal{C}}f, f \in \mathcal{F})$ als Supremum abzählbar vieler Elemente angesehen werden. Es gilt offenbar

$$\hat{C}^{Kui} \leq \sup_{(f,g) \in \mathcal{F}_1 \times \mathcal{F}_2} |\hat{\mathcal{C}}f + \hat{\mathcal{C}}g| = \|\hat{\mathcal{C}}|_{\mathcal{F}_1} + \hat{\mathcal{C}}|_{\mathcal{F}_2}\|_\infty,$$

mit $\|\cdot\|_\infty$ als Supremumsnorm auf $C^u(\mathcal{F}_1 \times \mathcal{F}_2)$.

Mit Überlegungen wie im vorherigen Satz, erhalten wir dann mit den Sätzen D.1.1 und D.1.2 die Stetigkeit der Verteilungsfunktion von \hat{C}^{Kui} auf $(0, \infty)$ und die strenge Monotonie auf $[0, \infty)$, woraus die Konsistenz des Tests gegen alle Alternativen folgt. Die Darstellung

$$C^{Kui} = \sup_{(f,g) \in \mathcal{F}_1^* \times \mathcal{F}_2^*} (\hat{\mathcal{C}}f + \hat{\mathcal{C}}g),$$

mit den Teilmengen $\mathcal{F}_1^* = \{f \in \overline{\mathcal{F}}_1 : Pf = Qf\}$, $\mathcal{F}_2^* = \{f \in \overline{\mathcal{F}}_2 : Pf = Qf\}$, liefert ferner erneut die Stetigkeit der Verteilungsfunktion von C^{Kui} auf $(0, \infty)$ im Fall der Nichtdegeneriertheit der Grenzverteilung analog zum Beweis des letzten Satzes und somit die Einhaltung des Testniveaus für $\alpha \in (0, \frac{1}{2})$. \square

Schließlich erhalten wir für den Cramér-von Mises-Test:

5.3.4. Satz. Für $\alpha \in (0, \frac{1}{2})$ ist $\varphi_{m,n}^{CvM}$ ein konsistenter Test zum Niveau α für das Testproblem (5.1.1).

Beweis. Es ist

$$\hat{C}^{CvM} = \int_{\mathbb{R}^d} (C_H^+(\mathbf{t}) + \overline{C}_H^+(\mathbf{t})) \omega(\mathbf{t}) d\mathbf{t}.$$

Die Pfade $\mathbf{t} \mapsto C_H^+(\mathbf{t}) \omega(\mathbf{t})$, $\mathbf{t} \in \mathbb{R}^d$, und $\mathbf{t} \mapsto \overline{C}_H^+(\mathbf{t}) \omega(\mathbf{t})$, $\mathbf{t} \in \mathbb{R}^d$, sind integrierbar bezüglich des Lebesgue-Borelschen Maßes auf \mathbb{R}^d . Bezeichnet Q ein Wahrscheinlichkeitsmaß auf \mathbb{R}^d mit der Lebesgue-Dichte ω , so können wir C_H und \overline{C}_H jeweils als Zufallsvariable mit Werten in $(L_Q^1(\mathbb{R}^d), \|\cdot\|_{L_Q^1})$ auffassen. Wir argumentieren dann ähnlich wie im Beweis zu Satz 2.3.4 indem wir mit Hilfe des Satzes von Hahn-Banach und der Identität $2x^+ = 2 \max(0, x) = |x| + x$, $x \in \mathbb{R}$, die Darstellung

$$\begin{aligned} 2\hat{C}^{CvM} &= \int_{\mathbb{R}^d} (|C_H(\mathbf{t})| + C_H(\mathbf{t})) dQ(\mathbf{t}) + \int_{\mathbb{R}^d} (|\overline{C}_H(\mathbf{t})| + \overline{C}_H(\mathbf{t})) dQ(\mathbf{t}) \\ &= \sup_{f \in L_Q^\infty(\mathbb{R}^d), \|f\|_{L_Q^\infty} = 1} \left\{ \int_{\mathbb{R}^d} (1 + f(\mathbf{t})) C_H(\mathbf{t}) dQ(\mathbf{t}) \right\} \\ &\quad + \sup_{g \in L_Q^\infty(\mathbb{R}^d), \|g\|_{L_Q^\infty} = 1} \left\{ \int_{\mathbb{R}^d} (1 + g(\mathbf{t})) \overline{C}_H(\mathbf{t}) dQ(\mathbf{t}) \right\} \\ &= \sup_{\substack{(f,g) \in L_Q^\infty(\mathbb{R}^d) \times L_Q^\infty(\mathbb{R}^d) \\ \|f\|_{L_Q^\infty} = \|g\|_{L_Q^\infty} = 1}} \left\{ \int_{\mathbb{R}^d} ((1 + f(\mathbf{t})) C_H(\mathbf{t}) + (1 + g(\mathbf{t})) \overline{C}_H(\mathbf{t})) dQ(\mathbf{t}) \right\} \end{aligned}$$

erhalten, so dass sich \hat{C}^{CvM} als Supremum über abzählbar viele normalverteilte Zufallsvariablen ergibt.

Mit dem Maß $\mu := \delta_0 + \delta_1$ auf $\mathfrak{P}(\{0, 1\})$ definieren wir ferner den Maßraum

$$(T, \mathfrak{G}, \tilde{Q}) = (\mathbb{R}^d \times \{0, 1\}, \mathcal{B}(\mathbb{R}^d) \otimes \mathfrak{P}(\{0, 1\}), Q \otimes \mu),$$

sowie den Prozess $Z = (Z(\mathbf{s}, t), (\mathbf{s}, t) \in T)$ mit $Z(\mathbf{s}, 0) := C_H(\mathbf{s})$, $Z(\mathbf{s}, 1) := \overline{C}_H(\mathbf{s})$. Dann ist Z ein Gaußprozess mit Werten im separablen Banachraum $L_Q^1(T)$ und es gilt mit dem

Satz von Fubini

$$\|Z\|_{L^1_{\hat{Q}}(T)} = \int_{\mathbb{R}^d} (|C_H(\mathbf{t})| + |\overline{C}_H(\mathbf{t})|) dQ(\mathbf{t}) \geq_{st} \hat{C}^{CvM}.$$

Also ist \hat{C}^{CvM} stochastisch nach oben durch die Norm eines Banachraum-wertigen gaußschen Zufallselements beschränkt. Schließlich liefert dann die Kombination der Sätze D.3.3, D.1.1 und D.1.2 für die Verteilungsfunktion von \hat{C}^{CvM} erneut die Stetigkeit auf $(0, \infty)$ und die strenge Monotonie auf $[0, \infty)$.

Analog wie zuvor schließen wir aus

$$\forall a \in (0, 1) \forall m, n \in \mathbb{N} : \quad \hat{c}_{m,n}^{CvM} \geq 0,$$

für fast alle $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{Y}_1, \mathbf{Y}_2, \dots$ und

$$\lim_{m,n \rightarrow \infty} \hat{c}_{m,n}^{CvM}(\alpha) = \hat{c}^{CvM}(\alpha),$$

fast sicher und für alle $\alpha \in (0, 1)$, dass bei Gültigkeit der Alternative $K : \mathbf{X} \not\prec_c \mathbf{Y}$

$$\lim_{m,n \rightarrow \infty} P(\varphi_{m,n}^{CvM}(\mathbf{X}_1, \dots, \mathbf{X}_m, \mathbf{Y}_1, \dots, \mathbf{Y}_n) = 1 | K) = 1$$

ist. Schließlich ist die Grenzverteilung der Testgröße $C_{m,n}^{CvM}$ entweder degeneriert im Nullpunkt, oder von der Form

$$C^{CvM} = \int_A C^+(\mathbf{t}) \omega(\mathbf{t}) d\mathbf{t} + \int_B \overline{C}^+(\mathbf{t}) \omega(\mathbf{t}) d\mathbf{t},$$

mit $\emptyset \neq A \cup B \subset \mathbb{R}^d$. Wir verwenden dann die gleichen Argumente wie zuvor um die Stetigkeit der Verteilungsfunktion der Grenzverteilung auf $(0, \infty)$ einzusehen und erhalten wiederum

$$\lim_{m,n \rightarrow \infty} P(C_{m,n}^{CvM} > \hat{c}_{m,n}^{CvM}(\alpha) | H) = P(C^{CvM} > \hat{c}^{CvM}(\alpha)) \leq P(\hat{C}^{CvM} > \hat{c}^{CvM}(\alpha)) = \alpha,$$

für alle $\alpha \in (0, \frac{1}{2})$. □

Alle drei Tests halten asymptotisch jedes vorgegebene Testniveau $\alpha \in (0, \frac{1}{2})$ ein. Im Fall der Degeneriertheit der Grenzverteilungen der Testgrößen C^{KS} , C^{Kui} oder C^{CvM} ergibt sich mit den vorherigen Ergebnissen ferner, dass die Tests in diesen Fällen darüber hinaus sogar asymptotisch einen Fehler erster Art von Null haben, denn es ist dann etwa

$$\lim_{m,n \rightarrow \infty} P(C_{m,n}^{CvM} > \hat{c}_{m,n}^{CvM}(\alpha) | H) = P(C^{CvM} > \hat{c}^{CvM}(\alpha)) = 0,$$

für $\alpha \in (0, \frac{1}{2})$.

5.4 Implementierung und numerische Ergebnisse

Die Berechnung der Testgrößen erfordert vor allem die Berechnung der Differenz der empirischen Verteilungsfunktionen

$$\Delta_{m,n}(\mathbf{t}) = \sqrt{\frac{mn}{m+n}} (F_m(\mathbf{t}) - G_n(\mathbf{t})), \quad \mathbf{t} \in \mathbb{R}^d,$$

und die Differenz der empirischen Überlebensfunktionen

$$\bar{\Delta}_{m,n}(\mathbf{t}) := \sqrt{\frac{mn}{m+n}} (\bar{F}_m(\mathbf{t}) - \bar{G}_n(\mathbf{t})), \quad \mathbf{t} \in \mathbb{R}^d.$$

Offenbar sind die Funktionen $\Delta_{m,n}(\mathbf{t})$ und $\bar{\Delta}_{m,n}(\mathbf{t})$ "stückweise" konstant und nehmen jeweils nur endlich viele Werte an. Die im nächsten Lemma zu formulierende Aussage präzisiert diesen Sachverhalt und gibt, bei gegebenen Beobachtungen zu den Verteilungsfunktionen F und G , explizit die Punkte an, an denen die Funktionen $\Delta_{m,n}(\mathbf{t})$ und $\bar{\Delta}_{m,n}(\mathbf{t})$ auszuwerten sind.

Seien zunächst wieder $\mathbf{X}_1 = (X_{11}, \dots, X_{1d}), \dots, \mathbf{X}_m = (X_{m1}, \dots, X_{md}), m \in \mathbb{N}$, unabhängig und identisch verteilt mit der Verteilungsfunktion F und $\mathbf{Y}_1 = (Y_{11}, \dots, Y_{1d}), \dots, \mathbf{Y}_n = (Y_{n1}, \dots, Y_{nd}), n \in \mathbb{N}$, unabhängig und identisch verteilt mit der Verteilungsfunktion G . Seien $\mathbf{X}_1, \dots, \mathbf{X}_m, \mathbf{Y}_1, \dots, \mathbf{Y}_n$ unabhängig. Sei $\mathbf{Z}_1 = (Z_{11}, \dots, Z_{1d}), \dots, \mathbf{Z}_N = (Z_{N1}, \dots, Z_{Nd}), N = m + n$, definiert durch $\mathbf{Z}_k = \mathbf{X}_k, k = 1, \dots, m$, und $\mathbf{Z}_k = \mathbf{Y}_{k-m}, k = m + 1, \dots, m + n$.

Ferner vereinbaren wir in diesem Abschnitt für $\mathbf{s} = (s_1, \dots, s_d) \in \mathbb{R}^d, \mathbf{t} = (t_1, \dots, t_d) \in \mathbb{R}^d$ die Schreibweisen

$$[\mathbf{s}, \mathbf{t}] := \prod_{k=1}^d [s_k, t_k], \quad [\mathbf{s}, \mathbf{t}] := \prod_{k=1}^d [s_k, t_k], \quad (\mathbf{s}, \mathbf{t}) := \prod_{k=1}^d (s_k, t_k).$$

Das folgende Resultat liefert die Grundlage für eine einfache Berechnung der Testgrößen.

5.4.1. Lemma. *Seien $\mathbf{Z}_1 = (Z_{11}, \dots, Z_{1d}), \dots, \mathbf{Z}_N = (Z_{N1}, \dots, Z_{Nd}), N = m + n$, gegeben und für $(i_1, \dots, i_d) \in \{1, \dots, m + n\}^d$ sei*

$$\mathbf{V}^{(i_1, \dots, i_d)} := (Z_{(i_1)1}, \dots, Z_{(i_d)d}), \quad (5.4.1)$$

wobei $Z_{(j)k}$ für jedes feste $k = 1, \dots, d$ die j -te Ordnungsstatistik der Z_{1k}, \dots, Z_{Nk} bezeichne, $j = 1, \dots, m + n$. Dann nimmt $\Delta_{m,n}$ die höchstens $(m + n)^d$ verschiedenen Werte

$$\Delta_{m,n}(\mathbf{V}^{(i_1, \dots, i_d)}), \quad (i_1, \dots, i_d) \in \{1, \dots, m + n\}^d,$$

an. Wir setzen ferner für ein fest gewähltes $\varepsilon > 0$ und $(i_1, \dots, i_d) \in \{1, \dots, m+n\}^d$

$$\mathbf{V}_\varepsilon^{(i_1, \dots, i_d)} := (Z_{(i_1)1}^\varepsilon, \dots, Z_{(i_d)d}^\varepsilon),$$

mit $Z_{(1)k}^\varepsilon := Z_{(1)k} - \varepsilon$, $k = 1, \dots, d$, und $Z_{(j)k}^\varepsilon := Z_{(j-1)k}$, $j = 2, \dots, m+n$, $k = 1, \dots, d$.
Dann nimmt $\bar{\Delta}_{m,n}$ die höchstens $(m+n)^d$ verschiedenen Werte

$$\bar{\Delta}_{m,n}(\mathbf{V}_\varepsilon^{(i_1, \dots, i_d)}), \quad (i_1, \dots, i_d) \in \{1, \dots, m+n\}^d,$$

an. Dabei kann $\varepsilon > 0$ beliebig gewählt werden.

Beweis. Die Werte von $\Delta_{m,n}(\mathbf{t})$ ändern sich höchstens dann, wenn sich mindestens einer der Werte $F_m(\mathbf{t})$ oder $G_n(\mathbf{t})$ ändert. Es ist für $\mathbf{t} = (t_1, \dots, t_d)$

$$F_m(\mathbf{t}) = \frac{1}{m} \sum_{k=1}^m \mathbf{1}(X_{k1} \leq t_1, \dots, X_{kd} \leq t_d). \quad (5.4.2)$$

Zunächst ist offenbar

$$F_m(\mathbf{t}) = 0, \quad \forall \mathbf{t} = (t_1, \dots, t_d) : \exists k \in \{1, \dots, d\} \text{ mit } t_k < X_{(1)k}.$$

Wir bezeichnen im Folgenden mit R_{kj} , $k = 1, \dots, m$, $j = 1, \dots, d$, den Rang der j -ten Koordinate X_{kj} von \mathbf{X}_k in X_{1j}, \dots, X_{mj} . Dann erhalten wir direkt aus (5.4.2), dass $F_m(\mathbf{t})$ für ein beliebiges $(i_1, \dots, i_d) \in \{1, \dots, m\}^d$ auf der Rechteckmenge

$$M_{i_1, \dots, i_d} := [X_{i_1 1}, X_{(R_{i_1+1})1}) \times \dots \times [X_{i_d d}, X_{(R_{i_d+1})d}) \quad (5.4.3)$$

konstant gleich $F_m(X_{i_1 1}, \dots, X_{i_d d})$ ist, wobei wir $X_{(R_{kj}+1)j} := \infty$ setzen, falls $R_{kj} = m$ gilt. Es gilt $F_m(\mathbf{t}) > 0$ für ein $\mathbf{t} \in \mathbb{R}^d$ genau dann, wenn ein $k \in \{1, \dots, m\}$ existiert mit $\mathbf{X}_k \leq \mathbf{t}$. Somit gibt es für jedes $\mathbf{t} \in \mathbb{R}^d$ mit $F_m(\mathbf{t}) > 0$ ein $(i_1, \dots, i_d) \in \{1, \dots, m\}^d$ mit $\mathbf{t} \in M_{i_1, \dots, i_d}$. Wir finden daher für jedes $\mathbf{t} \in \mathbb{R}^d$ mit $F_m(\mathbf{t}) > 0$ ein Tupel $(i_1, \dots, i_d) \in \{1, \dots, m\}^d$ mit $F_m(\mathbf{t}) = F_m(X_{i_1 1}, \dots, X_{i_d d})$. Ebenso schließen wir für $G_n(\mathbf{t})$, dass für jedes $\mathbf{t} \in \mathbb{R}^d$ mit $G_n(\mathbf{t}) > 0$ ein Tupel $(i_1, \dots, i_d) \in \{1, \dots, n\}^d$ existiert mit $G_n(\mathbf{t}) = G_n(Y_{i_1 1}, \dots, Y_{i_d d})$.

Aus diesen Vorüberlegungen für F_m und G_n folgt, dass die Differenz $F_m(\mathbf{t}) - G_n(\mathbf{t})$, $\mathbf{t} \in \mathbb{R}^d$, (und daher auch die Funktion $\Delta_{m,n}(\mathbf{t})$, $\mathbf{t} \in \mathbb{R}^d$) auf den Rechteckmengen

$$[\mathbf{V}^{(i_1, \dots, i_d)}, \mathbf{V}^{(i_1+1, \dots, i_d+1)}), \quad (i_1, \dots, i_d) \in \{1, \dots, m+n\}^d,$$

konstant gleich $\Delta_{m,n}(\mathbf{V}^{(i_1, \dots, i_d)})$ ist, wobei wir die Definition von $\mathbf{V}^{(i_1, \dots, i_d)}$ in (5.4.1) auf Indizes $(i_1, \dots, i_d) \in \{1, \dots, m+n+1\}^d$ erweitern, indem wir $Z_{(i_d+1)d} = \infty$ setzen,

falls $i_d = m$ gilt. Also existiert für jedes $\mathbf{t} \in \mathbb{R}^d$ ein $(i_1, \dots, i_d) \in \{1, \dots, m+n\}^d$ mit $\Delta_{m,n}(\mathbf{t}) = \Delta_{m,n}(\mathbf{V}^{(i_1, \dots, i_d)})$, woraus die Behauptung des Lemmas für $\Delta_{m,n}$ folgt.

Wir betrachten nun die Funktion $\mathbf{t} \mapsto \bar{\Delta}_{m,n}(\mathbf{t})$. Aus

$$\bar{F}_m(\mathbf{t}) = \frac{1}{m} \sum_{k=1}^m \mathbf{1}(X_{k1} > t_1, \dots, X_{kd} > t_d)$$

folgt zunächst

$$\bar{F}_m(\mathbf{t}) = 0, \quad \forall \mathbf{t} = (t_1, \dots, t_d) : \exists k \in \{1, \dots, d\} \text{ mit } t_k \geq X_{(m)k},$$

so dass $\bar{\Delta}_{m,n}(\mathbf{V}^{(i_1, \dots, i_d)}) = 0$ gilt, sobald ein $k \in \{1, \dots, d\}$ existiert mit $i_k = m+n$. Mit ähnlichen Überlegungen wie im Fall der Verteilungsfunktionen, macht man sich außerdem klar, dass $\bar{F}_m(\mathbf{t})$ für ein beliebiges $(i_1, \dots, i_d) \in \{1, \dots, m\}^d$ auf der Rechteckmenge

$$M_{i_1, \dots, i_d} = [X_{i_1 1}, X_{(R_{i_1 1}+1)1}) \times \dots \times [X_{i_d d}, X_{(R_{i_d d}+1)d})$$

konstant gleich $\bar{F}_m(X_{i_1 1}, \dots, X_{i_d d})$ ist, wobei wir erneut $X_{(R_{kj}+1)j} := \infty$ setzen, falls $R_{kj} = m$ gilt.

Im Fall der Überlebensfunktion ist $\bar{F}_m(\mathbf{t})$ allerdings im Allgemeinen auch für solche $\mathbf{t} = (t_1, \dots, t_d) \in \mathbb{R}^d$ positiv, für die ein $k \in \{1, \dots, d\}$ mit $t_k < X_{(1)k}$, also $t_k = X_{(1)k} - \varepsilon$ für ein $\varepsilon > 0$, existiert. (Beispielsweise gilt für $\mathbf{t}_0 := (X_{(1)1} - \varepsilon, \dots, X_{(1)d} - \varepsilon)$ mit einem beliebigen $\varepsilon > 0$ offenbar $\bar{F}_m(\mathbf{t}_0) = 1$, aber es ist $\mathbf{t}_0 \notin M_{i_1, \dots, i_d}$ für jedes $(i_1, \dots, i_d) \in \{1, \dots, m\}^d$.)

Für jedes $\mathbf{t} \in \mathbb{R}^d$ mit $t_k < X_{(1)k}$ für mindestens ein $k \in \{1, \dots, d\}$ und jedes feste $\varepsilon > 0$, gibt es dann aber ein Tupel $(i_1, \dots, i_d) \in \{1, \dots, m\}^d$ und einen Punkt $(X_{(i_1)1}^\varepsilon, \dots, X_{(i_d)d}^\varepsilon)$, $X_{(1)k}^\varepsilon := X_{(1)k} - \varepsilon$, $X_{(j)k}^\varepsilon := X_{(j-1)k}$, $j = 2, \dots, m$, $k = 1, \dots, d$, mit

$$\bar{F}_m(\mathbf{t}) = \bar{F}_m(X_{(i_1)1}^\varepsilon, \dots, X_{(i_d)d}^\varepsilon).$$

Analoge Überlegungen für $\bar{G}_n(\mathbf{t})$ liefern dann schließlich die im Lemma formulierte Behauptung für $\bar{\Delta}_{m,n}(\mathbf{t})$. □

5.4.2. Bemerkung. Die Ergebnisse des vorherigen Lemmas erlauben eine einfache Berechnung der Supremumstestgrößen $C_{m,n}^{Kui}$ und $C_{m,n}^{KS}$. Aus der vorliegenden zusammengelegten Stichprobe $\mathbf{Z}_1, \dots, \mathbf{Z}_N$ sind lediglich die Funktionswerte von $\Delta_{m,n}(\mathbf{t})$ in den Punkten

$\mathbf{V}^{(i_1, \dots, i_d)}$ und die Funktionswerte von $\bar{\Delta}_{m,n}(\mathbf{t})$ in den Punkten $\mathbf{V}_\varepsilon^{(i_1, \dots, i_d)}$ für ein beliebiges $\varepsilon > 0$ zu berechnen. Anschließend ergeben sich die Werte der Testgrößen $C_{m,n}^{Kui}$ und $C_{m,n}^{KS}$ als entsprechende Maxima dieser Werte. Bevor wir diesen Sachverhalt im nächsten Korollar formulieren, schließen wir zunächst einige Bemerkungen an.

- a) Die Menge der Punkte $\{\mathbf{V}^{(i_1, \dots, i_d)}, (i_1, \dots, i_d) \in \{1, \dots, m+n\}^d\}$ ist offensichtlich mit der Menge der Punkte $\{(Z_{i_1 1}, \dots, Z_{i_d d}), (i_1, \dots, i_d) \in \{1, \dots, m+n\}^d\}$ identisch. Die Differenz $\Delta_{m,n}(\mathbf{t})$ muss also höchstens an sämtlichen Punkten ausgewertet werden, die sich im Sinne von Lemma 5.4.1 aus Mischungen der Komponenten der Zufallsvektoren $\mathbf{Z}_1, \dots, \mathbf{Z}_N$ ergeben. Offenbar ist ferner jedes $\mathbf{V}_\varepsilon^{(i_1, \dots, i_d)}$ mit der Eigenschaft, dass $i_k \neq 1$ für alle $k = 1, \dots, d$ gilt, mit einem Punkt $\mathbf{V}^{(j_1, \dots, j_d)}$ identisch. Die Differenz $\bar{\Delta}_{m,n}(\mathbf{t})$ wird also zusätzlich noch an solchen Punkten $\mathbf{t} \in \mathbb{R}^d$ ausgewertet, für die für mindestens eine Komponente t_k , $k = 1, \dots, d$, die Beziehung $t_k < Z_{(1)k}$ besteht.
- b) In der Situation von Lemma 5.4.1 sind die Verteilungsfunktionen F und G stetig und die über die Ordnungsstatistiken definierten Punkte $\mathbf{V}^{(i_1, \dots, i_d)}$ und $\mathbf{V}_\varepsilon^{(i_1, \dots, i_d)}$ somit wohldefiniert. Bei der Durchführung des Bootstrapverfahrens aus Abschnitt 5.3 zur Ermittlung der benötigten kritischen Werte, sind die Beobachtungen zu den Zufallsvektoren $\mathbf{Z}_1, \dots, \mathbf{Z}_N$ im vorherigen Lemma allerdings mit positiver Wahrscheinlichkeit nicht paarweise verschieden.

In diesen Fällen sind die aufgeführten Ordnungsstatistiken und Ränge nicht wohldefiniert und die Definition der Punkte $\mathbf{V}^{(i_1, \dots, i_d)}$ und $\mathbf{V}_\varepsilon^{(i_1, \dots, i_d)}$ in Lemma 5.4.1 in dieser Form nicht sinnvoll. Die Punkte, an denen die Funktionen $\Delta_{m,n}(\mathbf{t})$ und $\bar{\Delta}_{m,n}(\mathbf{t})$ dann ausgewertet werden müssen, ergeben sich aber weiterhin durch Mischung der Koordinaten der Zufallsvektoren $\mathbf{Z}_1, \dots, \mathbf{Z}_N$:

Präziser formuliert, seien $\mathbf{Z}_1 = (Z_{11}, \dots, Z_{1d}), \dots, \mathbf{Z}_N = (Z_{N1}, \dots, Z_{Nd})$, $N = m+n$, gegeben und es seien $\mathbf{U}_1 = (U_{11}, \dots, U_{1d}), \dots, \mathbf{U}_M = (U_{M1}, \dots, U_{Md})$, $M \leq N$, definiert als die paarweise verschiedenen \mathbf{Z}_k . Der Beweis zu Lemma 5.4.1 zeigt, dass im Fall der Funktion $\Delta_{m,n}(\mathbf{t})$ nur die Funktionswerte von $\Delta_{m,n}(\mathbf{t})$ in den Punkten $(Z_{i_1 1}, \dots, Z_{i_d d})$, $(i_1, \dots, i_d) \in \{1, \dots, N\}^d$ relevant sind. Die Menge der Punkte

$$\left\{ (U_{(i_1)1}, \dots, U_{(i_d)d}), (i_1, \dots, i_d) \in \{1, \dots, M\}^d \right\}$$

ist aber identisch mit der Menge der Punkte

$$\left\{ (Z_{i_1 1}, \dots, Z_{i_d d}), (i_1, \dots, i_d) \in \{1, \dots, N\}^d \right\},$$

so dass die Funktion $\Delta_{m,n}(\mathbf{t})$, basierend auf den Zufallsvariablen $\mathbf{Z}_1, \dots, \mathbf{Z}_N$, also keine anderen Werte als

$$\Delta_{m,n}(U_{(i_1)1}, \dots, U_{(i_d)d}), \quad (i_1, \dots, i_d) \in \{1, \dots, M\}^d,$$

annimmt.

Mit dem gleichen Argument ist offensichtlich, dass die Funktion $\bar{\Delta}_{m,n}(\mathbf{t})$, basierend auf den Zufallsvariablen $\mathbf{Z}_1, \dots, \mathbf{Z}_N$, keine anderen Werte als

$$\bar{\Delta}_{m,n}(U_{(i_1)1}^\varepsilon, \dots, U_{(i_d)d}^\varepsilon), \quad (i_1, \dots, i_d) \in \{1, \dots, M\}^d,$$

annimmt, wobei $\varepsilon > 0$ beliebig und $U_{(1)k}^\varepsilon := U_{(1)k} - \varepsilon$, sowie $U_{(j)k}^\varepsilon := U_{(j-1)k}$ für $j = 2, \dots, M$, $k = 1, \dots, d$ ist.

Bei Vorliegen von nicht zwingend paarweise verschiedenen Beobachtungen der Zufallsvektoren $\mathbf{Z}_1, \dots, \mathbf{Z}_N$ sind die folgenden Aussagen dieses Abschnitts daher in der Weise zu verstehen, dass sich die Punkte $\mathbf{V}^{(i_1, \dots, i_d)}$ und $\mathbf{V}_\varepsilon^{(i_1, \dots, i_d)}$ in der oben dargelegten Weise aus den paarweise verschiedenen Zufallsvektoren $\mathbf{U}_1, \dots, \mathbf{U}_M$ ergeben.

Insbesondere reduziert sich in diesen Fällen die Anzahl der zu berechnenden Funktionswerte.

- c) Abschließend bemerken wir, dass die Funktionen $\Delta_{m,n}$ und $\bar{\Delta}_{m,n}$ in vielen Fällen jeweils deutlich weniger als $(m+n)^d$ verschiedene Werte annehmen.

Als Folgerung des letzten Lemmas erhalten wir eine einfache Darstellung der Supremums-testgrößen.

5.4.3. Korollar. Es seien $\mathbf{X}_1 = (X_{11}, \dots, X_{1d}), \dots, \mathbf{X}_m = (X_{m1}, \dots, X_{md})$, $m \in \mathbb{N}$, und $\mathbf{Y}_1 = (Y_{11}, \dots, Y_{1d}), \dots, \mathbf{Y}_n = (Y_{n1}, \dots, Y_{nd})$, $n \in \mathbb{N}$, gegeben. Sei $I := \{1, \dots, m+n\}^d$. Dann gilt für ein beliebiges $\varepsilon > 0$

$$C_{m,n}^{Kui} = \max_{(i_1, \dots, i_d) \in I} \Delta_{m,n}(\mathbf{V}^{(i_1, \dots, i_d)}) + \max_{(i_1, \dots, i_d) \in I} \bar{\Delta}_{m,n}(\mathbf{V}_\varepsilon^{(i_1, \dots, i_d)}),$$

und

$$C_{m,n}^{KS} = \max \left(\max_{(i_1, \dots, i_d) \in I} \Delta_{m,n}(\mathbf{V}^{(i_1, \dots, i_d)}), \max_{(i_1, \dots, i_d) \in I} \bar{\Delta}_{m,n}(\mathbf{V}_\varepsilon^{(i_1, \dots, i_d)}) \right).$$

□

Tabelle 5.4 enthält erste Simulationsergebnisse zur geschätzten Güte des Kolmogorov-Smirnov-Tests $\varphi_{m,n}^{KS}$ im Fall einiger bivariater Verteilungen. Zwecks der besseren Übersicht wurde abkürzend

$$I^{(\alpha)} := \begin{pmatrix} 1 & \alpha \\ \alpha & 1 \end{pmatrix}, \quad \alpha \in (-1, 1),$$

und $I := I^{(0)}$ gesetzt. Mit $N_2(\mathbf{a}, \Sigma)$ bezeichnen wir wie gewohnt die bivariate Normalverteilung mit Erwartungswertvektor \mathbf{a} und Kovarianzmatrix Σ . Ferner bezeichnet $t_n(\mathbf{a}, \Sigma)$, $n \in \mathbb{N}$, die bivariate t -Verteilung mit n Freiheitsgraden im Sinne von Beispiel 4.3.5 c). Das generelle Vorgehen bei der Durchführung der Simulationen wurde bereits in Abschnitt 2.5 dargelegt. Die folgenden Ergebnisse ergaben sich stets für $R = 50$ Bootstrapwiederholungen.

Die Zeilen 1-3 im ersten Block von Tabelle 5.4 repräsentieren Verteilungen bei Gültigkeit der Hypothese (vergleiche Satz 4.3.4 und die Beispiele in 4.3.5), die Zeilen 4-8 repräsentieren die Alternative. Analog repräsentieren Zeilen 1-3 des zweiten Blocks Verteilungen bei Gültigkeit der Hypothese für die t -Verteilung, und die Zeilen 4-8 die Alternative. Es zeigt sich, dass das Niveau bei Gültigkeit der Hypothese in den Fällen $\mathbf{X} \stackrel{v}{=} \mathbf{Y}$ leicht überschritten wird, wohingegen in den Fällen $\mathbf{X} \leq_c \mathbf{Y}$ und $P^{\mathbf{X}} \neq P^{\mathbf{Y}}$ das Niveau auch bei den Simulationen bereits stets eingehalten wird. Bei Gültigkeit der Alternative werden Lagealternativen offenbar leichter aufgedeckt als Skalenalternativen. Der dritte Block repräsentiert Verteilungen bei Gültigkeit der Alternative bei denen nicht $\mathbf{Y} \leq_c \mathbf{X}$ gilt.

Tabelle 5.5 zeigt die Simulationsergebnisse in diesen Fällen für den Kuiper-Test $\varphi_{m,n}^{Kui}$. Auffällig ist, dass $\varphi_{m,n}^{Kui}$ in fast allen Fällen ein konservativeres Verhalten zeigt, im Fall der Lagealternativen aber eine deutlich höhere empirische Güte als $\varphi_{m,n}^{KS}$ aufweist.

Wir wollen nun eine leicht zu implementierende Darstellung der Integral-Testgröße $C_{m,n}^{CvM}$ angeben. Die Tatsache, dass die Funktionen $\Delta_{m,n}(\mathbf{t})$ und $\bar{\Delta}_{m,n}(\mathbf{t})$, wie bereits oben dargelegt, jeweils konstant auf geeignet gewählten und von den Daten abhängenden d -dimensionalen Quadern sind und ferner offenbar

$$\Delta_{m,n}(\mathbf{t}) = 0, \quad \forall \mathbf{t} = (t_1, \dots, t_d) : \exists k \in \{1, \dots, d\} \text{ mit } t_k < Z_{(1)k},$$

und

$$\bar{\Delta}_{m,n}(\mathbf{t}) = 0, \quad \forall \mathbf{t} = (t_1, \dots, t_d) : \exists k \in \{1, \dots, d\} \text{ mit } t_k \geq Z_{(m+n)k}$$

gilt, reduziert die Berechnung der Cramér-von Mises-Testgröße auf die Berechnung von endlichen Summen. Wir illustrieren diesen Sachverhalt zwecks der besseren Übersicht anhand der zweidimensionalen Situation.

Seien $\mathbf{X}_1 = (X_{11}, X_{12}), \dots, \mathbf{X}_m = (X_{m1}, X_{m2}), m \in \mathbb{N}$, unabhängig und identisch verteilt mit Verteilungsfunktion F und $\mathbf{Y}_1 = (Y_{11}, Y_{12}), \dots, \mathbf{Y}_n = (Y_{n1}, Y_{n2}), n \in \mathbb{N}$, unabhängig und identisch verteilt mit Verteilungsfunktion G und $\mathbf{Z}_1 = (Z_{11}, Z_{12}), \dots, \mathbf{Z}_N = (Z_{N1}, Z_{N2}), N = m + n$, seien unabhängig.

Dann ist $\Delta_{m,n}(\mathbf{t})$ jeweils konstant in den Bereichen

$$[\mathbf{V}^{(i,j)}, \mathbf{V}^{(i+1,j+1)}] = [Z_{(i)1}, Z_{(i+1)1}] \times [Z_{(j)2}, Z_{(j+1)2}],$$

$(i, j) \in \{1, \dots, N-1\}^2$, sowie in den Bereichen $[Z_{(m+n)1}, \infty) \times [Z_{(k)2}, Z_{(k+1)2}]$, $k = 1, \dots, N-1$, und $[Z_{(k)1}, Z_{(k+1)1}] \times [Z_{(m+n)2}, \infty)$, $k = 1, \dots, N-1$. Außerdem gilt offenbar $\Delta_{m,n}(\mathbf{V}^{(m+n,m+n)}) = 0$. Betrachten wir die Testgröße

$$C_{m,n}^{CvM} = \int_{\mathbb{R}^d} \left(\Delta_{m,n}^+(\mathbf{t}) + \bar{\Delta}_{m,n}^+(\mathbf{t}) \right) \omega(\mathbf{t}) dt, \quad (5.4.4)$$

und ist Q ein Wahrscheinlichkeitsmaß auf den Borelmengen des \mathbb{R}^2 mit der Lebesgue-Dichte ω , so erhalten wir für das erste Integral in (5.4.4) die folgende handliche Darstellung.

5.4.4. Korollar. In der oben beschriebenen Situation gilt

$$\begin{aligned} \int_{\mathbb{R}^d} \Delta_{m,n}^+(\mathbf{t}) \omega(\mathbf{t}) dt &= \sum_{i,j=1}^{N-1} \Delta_{m,n}^+(\mathbf{V}^{(i,j)}) Q \left([\mathbf{V}^{(i,j)}, \mathbf{V}^{(i+1,j+1)}] \right) \\ &\quad + \sum_{i=1}^{N-1} \Delta_{m,n}^+(\mathbf{V}^{(i,m+n)}) Q \left([Z_{(i)1}, Z_{(i+1)1}] \times [Z_{(m+n)2}, \infty) \right) \\ &\quad + \sum_{j=1}^{N-1} \Delta_{m,n}^+(\mathbf{V}^{(m+n,j)}) Q \left([Z_{(m+n)1}, \infty) \times [Z_{(j)2}, Z_{(j+1)2}] \right). \end{aligned}$$

□

Um eine ähnliche Darstellung für den zweiten Integralterm in $C_{m,n}^{CvM}$ zu erhalten, bemerken wir zunächst, dass $\bar{\Delta}_{m,n}(\mathbf{t})$ in den Bereichen $[\mathbf{V}^{(i,j)}, \mathbf{V}^{(i+1,j+1)}]$ mit $(i, j) \in \{1, \dots, N-1\}^2$ konstant gleich $\bar{\Delta}_{m,n}(\mathbf{V}^{(i,j)})$, in den Bereichen der Form $[Z_{(i)1}, Z_{(i+1)1}] \times (-\infty, Z_{(1)2})$, $i = 1, \dots, N-1$, konstant gleich $\bar{\Delta}_{m,n}(\mathbf{V}_\varepsilon^{(i,1)})$ und in den Bereichen der Form $(-\infty, Z_{(1)1}) \times [Z_{(j)2}, Z_{(j+1)2}]$, $j = 1, \dots, N-1$, konstant gleich $\bar{\Delta}_{m,n}(\mathbf{V}_\varepsilon^{(1,j)})$ ist.

5.4.5. Korollar. In der oben beschriebenen Situation gilt

$$\begin{aligned} \int_{\mathbb{R}^d} \bar{\Delta}_{m,n}^+(\mathbf{t}) \omega(\mathbf{t}) d\mathbf{t} &= \sum_{i,j=1}^{N-1} \bar{\Delta}_{m,n}^+(\mathbf{V}^{(i,j)}) Q\left([\mathbf{V}^{(i,j)}, \mathbf{V}^{(i+1,j+1)}]\right) \\ &+ \sum_{i=2}^N \bar{\Delta}_{m,n}^+(\mathbf{V}_\varepsilon^{(i,1)}) Q\left([Z_{(i-1)1}, Z_{(i)1}] \times (-\infty, Z_{(1)2})\right) \\ &+ \sum_{j=2}^N \bar{\Delta}_{m,n}^+(\mathbf{V}_\varepsilon^{(1,j)}) Q\left((-\infty, Z_{(1)1}] \times [Z_{(j-1)2}, Z_{(j)2}]\right), \end{aligned}$$

dabei kann $\varepsilon > 0$ beliebig gewählt werden. \square

Die beiden letzten Korollare ermöglichen eine einfache Implementierung der Cramér-von Mises-Testgröße im bivariaten Fall. Mit höherem formalen Aufwand lassen sich analoge Darstellungen ebenfalls in Dimensionen $d \geq 3$ angeben. Wir verzichten darauf an dieser Stelle und weisen statt dessen auf den Einfluss der Wahl der Gewichtsfunktion ω , bzw. der des zugehörigen Maßes Q , auf die numerischen Berechnungen hin.

Auf die Voraussetzung $\omega(\mathbf{t}) > 0$ für alle $\mathbf{t} \in \mathbb{R}^d$ kann zunächst nicht ohne Weiteres verzichtet werden. In der Tat existiert bei Gültigkeit der Alternative aufgrund der vorausgesetzten Stetigkeit von F und G eine Umgebung $U(\mathbf{t}_0)$ mit $F(\mathbf{t}) > G(\mathbf{t})$, $\mathbf{t} \in U(\mathbf{t}_0)$, oder $\bar{F}(\mathbf{t}) > \bar{G}(\mathbf{t})$, $\mathbf{t} \in U(\mathbf{t}_0)$. Für $U(\mathbf{t}_0) \subset \mathbb{R}^d \setminus \text{supp}(\omega)$ kann das betrachtete Abstandsmaß im Fall der Cramér-von Mises-Testgröße a priori asymptotisch den Wert Null annehmen (siehe 2.4.3, c), für das vergleichbare Problem beim Test auf Laplace-Ordnung).

Eine naheliegende Wahl für ω ist etwa die Dichtefunktion der d -dimensionalen Standard-Normalverteilung ϕ_d . Tabelle 5.6 enthält Simulationsergebnisse für die empirische Güte von $\varphi_{m,n}^{CvM}$ mit $\omega = \phi_d$. Im Vergleich zu den Ergebnissen für $\varphi_{m,n}^{KS}$ und $\varphi_{m,n}^{Kui}$ scheint $\varphi_{m,n}^{CvM}$ bessere Resultate zu liefern. Im Fall der Gleichheit der Verteilungen wird das vorgegebene Testniveau in den Simulationen besser eingehalten, als bei den Supremumstests $\varphi_{m,n}^{KS}$ oder $\varphi_{m,n}^{Kui}$. Ferner zeigt sich für Verteilungen mit $\mathbf{X} \leq_c \mathbf{Y}$ und $P^{\mathbf{X}} \neq P^{\mathbf{Y}}$, dass der Fehler zweiter Art bei $\varphi_{m,n}^{CvM}$ schneller gegen Null zu konvergieren scheint, als bei den anderen Tests. Im Bereich der Alternative liefert $\varphi_{m,n}^{CvM}$ darüber hinaus in den simulierten Fällen eine höhere Güte.

Tabelle 5.3 zeigt allerdings ebenfalls die empirische Güte des Tests $\varphi_{m,n}^{CvM}$ für $\alpha = 0.05$ und $m = n = 20$ anhand eines Beispiels aus dem Bereich der Alternative.

F	G	$P(\varphi^{CvM} = 1)$
$N_2(\mathbf{100}, I^{(0.9)})$	$N_2(\mathbf{100}, I)$	0.000

Tabelle 5.3: Beispiel zum Einfluss der Gewichtsfunktion auf $\varphi_{m,n}^{CvM}$.

Die Ursache für dieses Verhalten scheint offensichtlich: Wegen $\lim_{\|\mathbf{t}\| \rightarrow \infty} \phi_d(\mathbf{t}) = 0$, werden die zu berechnenden Werte $Q([\mathbf{V}^{(i,j)}, \mathbf{V}^{(i+1,j+1)}])$ aus den beiden vorherigen Lemmata im Fall der Dichte der Normalverteilung für betragsmäßig große Beobachtungen $\mathbf{X}_1, \dots, \mathbf{X}_m, \mathbf{Y}_1, \dots, \mathbf{Y}_n$ sehr schnell sehr klein und werden von gängigen Computeralgebra-Systemen nicht mehr von Null unterschieden. Dieser Sachverhalt führt dazu, dass die Hypothese nicht mehr verworfen wird. Um dieses Problem zu umgehen, ist die Transformation der vorliegenden Daten vor der Durchführung des Tests nützlich. Die folgende Aussage ist wegen $F_{\frac{1}{a}\mathbf{X}}(\mathbf{t}) = F_{\mathbf{X}}(a\mathbf{t})$ und $\bar{F}_{\frac{1}{a}\mathbf{X}}(\mathbf{t}) = \bar{F}_{\mathbf{X}}(a\mathbf{t})$ evident.

5.4.6. Lemma. Sei $a > 0$. Dann gilt $\mathbf{X} \leq_c \mathbf{Y} \Leftrightarrow \frac{1}{a}\mathbf{X} \leq_c \frac{1}{a}\mathbf{Y}$. □

Die Daten können also beispielsweise zunächst auf den Bereich $\{\mathbf{t} \in \mathbb{R}^d : \|\mathbf{t}\| \leq 1\}$ transformiert werden, ohne dass das Testproblem verändert wird.

Verteilungen		a)	b)	c)
F	G			
$N_2(\mathbf{0}, I)$	$N_2(\mathbf{0}, I)$	0.074	0.094	0.090
$N_2(\mathbf{0}, I)$	$N_2(\mathbf{0}, I^{(0.5)})$	0.033	0.046	0.044
$N_2(\mathbf{10}, I)$	$N_2(\mathbf{10}, I^{(0.5)})$	0.034	0.044	0.036
$N_2(\mathbf{0}, I^{(0.2)})$	$N_2(\mathbf{0}, I)$	0.098	0.128	0.134
$N_2(\mathbf{0}, I^{(0.5)})$	$N_2(\mathbf{0}, I)$	0.176	0.302	0.227
$N_2(\mathbf{0}, I^{(0.9)})$	$N_2(\mathbf{0}, I)$	0.639	0.824	0.750
$N_2(\mathbf{1}, I)$	$N_2(\mathbf{0}, I)$	0.444	0.603	0.628
$N_2(\mathbf{100}, I^{(0.9)})$	$N_2(\mathbf{100}, I)$	1.000	1.000	1.000
$t_2(\mathbf{0}, I)$	$t_2(\mathbf{0}, I)$	0.076	0.094	0.091
$t_2(\mathbf{0}, I)$	$t_2(\mathbf{0}, I^{(0.5)})$	0.041	0.052	0.043
$t_2(\mathbf{10}, I)$	$t_2(\mathbf{10}, I^{(0.5)})$	0.032	0.045	0.042
$t_2(\mathbf{0}, I^{(0.2)})$	$t_2(\mathbf{0}, I)$	0.107	0.135	0.101
$t_2(\mathbf{0}, I^{(0.5)})$	$t_2(\mathbf{0}, I)$	0.179	0.248	0.186
$t_2(\mathbf{0}, I^{(0.9)})$	$t_2(\mathbf{0}, I)$	0.572	0.754	0.653
$t_2(\mathbf{1}, I)$	$t_2(\mathbf{0}, I)$	0.387	0.474	0.487
$N_2(\mathbf{0}, I)$	$t_2(\mathbf{0}, I)$	0.183	0.247	0.241
$t_1(\mathbf{0}, I)$	$N_2(\mathbf{0}, I)$	0.069	0.091	0.085
$N_2(-\mathbf{1}, I)$	$t_1(\mathbf{0}, I)$	0.836	0.927	0.929

Tabelle 5.4: Empirische Güte von $\varphi_{m,n}^{KS}$ im Fall $\alpha = 0.05$ und a) $m = 20, n = 20$, b) $m = 20, n = 30$, c) $m = 30, n = 20$.

Verteilungen		a)	b)	c)
F	G			
$N_2(\mathbf{0}, I)$	$N_2(\mathbf{0}, I)$	0.058	0.081	0.083
$N_2(\mathbf{0}, I)$	$N_2(\mathbf{0}, I^{(0.5)})$	0.053	0.048	0.052
$N_2(\mathbf{10}, I)$	$N_2(\mathbf{10}, I^{(0.5)})$	0.031	0.041	0.041
$N_2(\mathbf{0}, I^{(0.2)})$	$N_2(\mathbf{0}, I)$	0.092	0.115	0.108
$N_2(\mathbf{0}, I^{(0.5)})$	$N_2(\mathbf{0}, I)$	0.140	0.175	0.158
$N_2(\mathbf{0}, I^{(0.9)})$	$N_2(\mathbf{0}, I)$	0.279	0.404	0.397
$N_2(\mathbf{1}, I)$	$N_2(\mathbf{0}, I)$	0.938	0.979	0.972
$N_2(\mathbf{100}, I^{(0.9)})$	$N_2(\mathbf{100}, I)$	1.000	1.000	1.000
$t_2(\mathbf{0}, I)$	$t_2(\mathbf{0}, I)$	0.074	0.078	0.069
$t_2(\mathbf{0}, I)$	$t_2(\mathbf{0}, I^{(0.5)})$	0.041	0.043	0.045
$t_2(\mathbf{10}, I)$	$t_2(\mathbf{10}, I^{(0.5)})$	0.045	0.048	0.038
$t_2(\mathbf{0}, I^{(0.2)})$	$t_2(\mathbf{0}, I)$	0.080	0.102	0.112
$t_2(\mathbf{0}, I^{(0.5)})$	$t_2(\mathbf{0}, I)$	0.110	0.162	0.146
$t_2(\mathbf{0}, I^{(0.9)})$	$t_2(\mathbf{0}, I)$	0.263	0.385	0.408
$t_2(\mathbf{1}, I)$	$t_2(\mathbf{0}, I)$	0.780	0.842	0.857
$N_2(\mathbf{0}, I)$	$t_2(\mathbf{0}, I)$	0.108	0.152	0.140
$t_1(\mathbf{0}, I)$	$N_2(\mathbf{0}, I)$	0.054	0.049	0.064
$N_2(-\mathbf{1}, I)$	$t_1(\mathbf{0}, I)$	0.896	0.948	0.962

Tabelle 5.5: Empirische Güte von $\varphi_{m,n}^{Kui}$ im Fall $\alpha = 0.05$ und a) $m = 20, n = 20$, b) $m = 20, n = 30$, c) $m = 30, n = 20$.

Verteilungen		a)	b)	c)
F	G			
$N_2(\mathbf{0}, I)$	$N_2(\mathbf{0}, I)$	0.057	0.057	0.044
$N_2(\mathbf{0}, I)$	$N_2(\mathbf{0}, I^{(0.5)})$	0.007	0.008	0.007
$N_2(\mathbf{10}, I)$	$N_2(\mathbf{10}, I^{(0.5)})$	0.003	0.025	0.000
$N_2(\mathbf{0}, I^{(0.2)})$	$N_2(\mathbf{0}, I)$	0.115	0.136	0.095
$N_2(\mathbf{0}, I^{(0.5)})$	$N_2(\mathbf{0}, I)$	0.252	0.373	0.292
$N_2(\mathbf{0}, I^{(0.9)})$	$N_2(\mathbf{0}, I)$	0.664	0.790	0.753
$N_2(\mathbf{1}, I)$	$N_2(\mathbf{0}, I)$	0.982	0.998	0.991
$N_2(\mathbf{100}, I^{(0.9)})$	$N_2(\mathbf{100}, I)$	0.000	0.000	0.000
$t_2(\mathbf{0}, I)$	$t_2(\mathbf{0}, I)$	0.004	0.064	0.044
$t_2(\mathbf{0}, I)$	$t_2(\mathbf{0}, I^{(0.5)})$	0.005	0.002	0.007
$t_2(\mathbf{10}, I)$	$t_2(\mathbf{10}, I^{(0.5)})$	0.003	0.000	0.003
$t_2(\mathbf{0}, I^{(0.2)})$	$t_2(\mathbf{0}, I)$	0.112	0.142	0.111
$t_2(\mathbf{0}, I^{(0.5)})$	$t_2(\mathbf{0}, I)$	0.262	0.347	0.284
$t_2(\mathbf{0}, I^{(0.9)})$	$t_2(\mathbf{0}, I)$	0.665	0.769	0.717
$t_2(\mathbf{1}, I)$	$t_2(\mathbf{0}, I)$	0.849	0.905	0.892
$N_2(\mathbf{0}, I)$	$t_2(\mathbf{0}, I)$	0.059	0.051	0.064
$t_1(\mathbf{0}, I)$	$N_2(\mathbf{0}, I)$	0.070	0.093	0.066
$N_2(-\mathbf{1}, I)$	$t_1(\mathbf{0}, I)$	0.969	0.955	0.959

Tabelle 5.6: Empirische Güte von $\varphi_{m,n}^{CvM}$ im Fall $\alpha = 0.05$ und a) $m = 20, n = 20$, b) $m = 20, n = 30$, c) $m = 30, n = 20$ sowie ω als Dichtefunktion der d -dimensionalen Standard-Normalverteilung.

Anhang A

Stochastische Integralordnungen

Es sei S ein polnischer Raum und \mathfrak{B} die Borelsche σ -Algebra auf S . Weiter seien $\omega : S \rightarrow [1, \infty)$ eine messbare Abbildung und \mathfrak{B}_ω die Menge aller messbaren Funktionen $f : S \rightarrow \mathbb{R}$ mit der Eigenschaft

$$\|f\|_\omega := \sup_{s \in S} \frac{|f(s)|}{\omega(s)} < \infty. \quad (\text{A.1.1})$$

Ferner bezeichne \mathfrak{P}_ω die Menge der Wahrscheinlichkeitsmaße P auf \mathfrak{B} mit $\int \omega dP < \infty$.

A.1. Definition ([MS02], Definition 2.3.1). Für festes ω sei $\mathcal{F} \subset \mathfrak{B}_\omega$. Dann definieren wir für $P, Q \in \mathfrak{P}_\omega$

$$P \leq_{\mathcal{F}} Q \Leftrightarrow \int f dP \leq \int f dQ \quad \text{für alle } f \in \mathcal{F}, \quad (\text{A.1.2})$$

und nennen $\leq_{\mathcal{F}}$ die von \mathcal{F} generierte *stochastische Integralordnung* auf \mathfrak{P}_ω .

Die Menge \mathcal{F} heißt *Generator* von $\leq_{\mathcal{F}}$. Sind X und Y S -wertige Zufallsgrößen, so setzen wir außerdem $X \leq_{\mathcal{F}} Y \Leftrightarrow P^X \leq_{\mathcal{F}} P^Y$.

Im Spezialfall $(S, \mathfrak{B}) = (\mathbb{R}^d, \mathfrak{B}(\mathbb{R}^d))$ sind in Theorie und Praxis verschiedene stochastische Integralordnungen von Interesse. Für eine ausführliche Behandlung stochastischer Ordnungen und deren Eigenschaften verweisen wir auf die Standardwerke [MS02] und [SS07]. Wir geben in Tabelle A.1 lediglich eine Auflistung der in dieser Arbeit relevanten stochastischen Integralordnungen an.

Die Ordnungen $\leq_{\mathcal{L}}$ und \leq_{mom} betrachten wir dabei für komponentenweise nicht-negative Zufallsvektoren. Wir bezeichnen ferner eine reellwertige Funktion $f : \mathbb{R}^d \rightarrow \mathbb{R}$, $(x_1, \dots, x_d) \mapsto f(x_1, \dots, x_d)$, als *monoton wachsend in x_j* , falls die Funktion $x_j \mapsto$

$f(x_1, \dots, x_d)$ für festgehaltene x_k , $k \neq j$, monoton wachsend ist. Die Funktion f heißt *monoton wachsend* (oder schlicht *wachsend*), falls f monoton wachsend in jeder Komponente ist. Für $\mathbf{t} = (t_1, \dots, t_d) \in \mathbb{R}^d$ setzen wir $\mathbf{1}_{(-\infty, \mathbf{t}]} := \mathbf{1}_{(-\infty, t_1] \times \dots \times (-\infty, t_d]}$ und $\mathbf{1}_{(\mathbf{t}, \infty)} := \mathbf{1}_{(t_1, \infty) \times \dots \times (t_d, \infty)}$.

Eine Funktion $f : \mathbb{R}^d \rightarrow \mathbb{R}$ heißt *supermodular*, falls

$$f(\mathbf{x} \wedge \mathbf{y}) + f(\mathbf{x} \vee \mathbf{y}) \geq f(\mathbf{x}) + f(\mathbf{y}), \quad \mathbf{x}, \mathbf{y} \in \mathbb{R}^d,$$

gilt. Dabei bezeichnet $\mathbf{x} \wedge \mathbf{y}$ das (komponentenweise zu bildende) Minimum von \mathbf{x} und \mathbf{y} und $\mathbf{x} \vee \mathbf{y}$ das (komponentenweise zu bildende) Maximum von \mathbf{x} und \mathbf{y} . Ist f zweimal stetig differenzierbar, so ist f genau dann supermodular, falls

$$\forall \mathbf{x} \in \mathbb{R}^d \forall 1 \leq i < j \leq d : \quad \frac{\partial^2}{\partial x_i \partial x_j} f(\mathbf{x}) \geq 0$$

gilt. (Vergleiche [Bäu97])

Eine Funktion $f : \mathbb{R}^d \rightarrow \mathbb{R}$ heißt *direkt konvex*, falls sie supermodular und - bei jeweils festgehaltenen anderen Komponenten - konvex in jeder Komponente ist.

Schließlich setzen wir

$$\mathcal{M} := \{f_k, k \in \mathbb{N} \mid f_k : \mathbb{R} \rightarrow \mathbb{R}, f_k(x) = x^k\}.$$

Für eine feste Gewichtsfunktion ω und $\mathcal{F} \subset \mathfrak{B}_\omega$ definieren die Integralordnungen aus Tabelle A.1 (partielle) Ordnungen auf der Menge \mathfrak{P}_ω . Im Fall der Ordnung \leq_{icx} ist beispielsweise $\omega(\mathbf{t}) = 1 + \|\mathbf{t}\|$ und somit \mathfrak{P}_ω die Menge aller Wahrscheinlichkeitsmaße mit endlichem Erwartungswert. Vergleiche auch [MS02], Abschnitt 2.6.

Abbildung A.1 veranschaulicht die Beziehungen zwischen den einzelnen Ordnungen. Dabei bedeutet die Notation $\leq_1 \rightarrow \leq_2$, dass $P \leq_1 Q \Rightarrow P \leq_2 Q$ gilt und $\leq_1 \dashrightarrow \leq_2$, dass $P \leq_1 Q \Rightarrow Q \leq_2 P$ gilt. Die Implikation $\leq_c \dashrightarrow \leq_{\mathcal{L}}$ gilt für nicht-negative Zufallsvektoren, die Implikation $\leq_{icx} \rightarrow \leq_{mom}$ für Zufallsvektoren mit Werten in \mathbb{N}_0^d . Beziehungen und Eigenschaften der Ordnungen $\leq_{\mathcal{L}}$ bzw. \leq_c werden ausführlicher in den Kapiteln 1 bzw. 4 behandelt.

Symbol	Generator
\leq_{st}	$\{f : \mathbb{R}^d \rightarrow \mathbb{R}; f \text{ monoton wachsend}\}$
\leq_{icx}	$\{f : \mathbb{R}^d \rightarrow \mathbb{R}; f \text{ wachsend und konvex}\}$
\leq_{icv}	$\{f : \mathbb{R}^d \rightarrow \mathbb{R}; f \text{ wachsend und konkav}\}$
\leq_{sm}	$\{f : \mathbb{R}^d \rightarrow \mathbb{R}; f \text{ supermodular}\}$
\leq_{dir-cx}	$\{f : \mathbb{R}^d \rightarrow \mathbb{R}; f \text{ direkt konvex}\}$
$\leq_{\mathcal{L}}$	$\{f_{\mathbf{t}} : \mathbb{R}^d \rightarrow \mathbb{R}, \mathbf{t} \in \mathbb{R}_{\geq 0}^d; f_{\mathbf{t}}(\mathbf{x}) = -\exp(-\mathbf{t}'\mathbf{x})\}$
\leq_c	$\{\mathbf{1}_{(-\infty, \mathbf{t}]}, \mathbf{t} \in \mathbb{R}^d\} \cup \{\mathbf{1}_{(\mathbf{t}, \infty)}, \mathbf{t} \in \mathbb{R}^d\}$
\leq_{mom}	$\{f : \mathbb{R}^d \rightarrow \mathbb{R}; f(\mathbf{x}) = \prod_{i=1}^d f_i(x_i), f_i \in \mathcal{M}\}$

Tabelle A.1: Generatoren gängiger stochastischer Integralordnungen

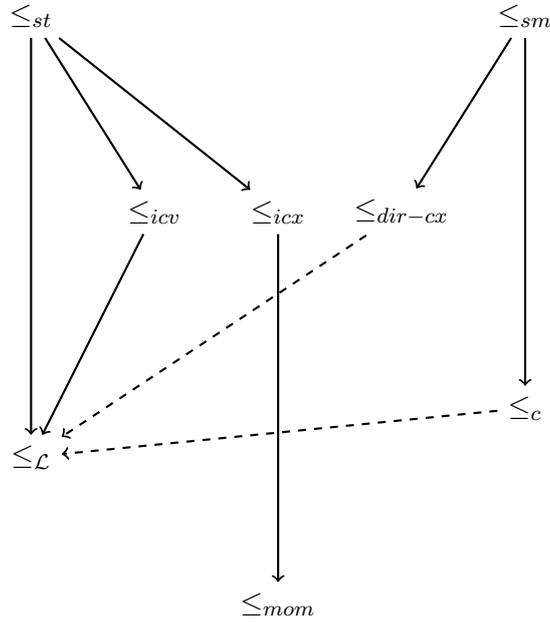


Abbildung A.1: Beziehungen zwischen multivariaten stochastischen Ordnungen.

Anhang B

Hoffmann-Jørgensen-Theorie

B.1 Schwache Konvergenz

Es sei (Ω, \mathcal{A}, P) ein Wahrscheinlichkeitsraum. $X : \Omega \rightarrow \overline{\mathbb{R}}$ eine Abbildung. Es heißt

$$E^*(X) := \inf\{E(Y) \mid Y : (\Omega, \mathcal{A}) \rightarrow (\overline{\mathbb{R}}, \mathfrak{B}(\overline{\mathbb{R}})), Y \geq X, E(Y) \text{ existiert}\}$$

der *äußere Erwartungswert* von X . Für alle $B \subset \Omega$ ist $P^*(B) := \inf\{P(A) : B \subset A \in \mathcal{A}\}$ die *äußere Wahrscheinlichkeit* von B . Ferner definieren wir den *inneren Erwartungswert* bzw. die *innere Wahrscheinlichkeit* als $E_*(X) = -E^*(-X)$ bzw. $P_*(B) = 1 - P^*(B^c)$.

Für eine beliebige Funktion $f : \Omega \rightarrow \overline{\mathbb{R}}$ bezeichne $f^* : \Omega \rightarrow \overline{\mathbb{R}}$ die *kleinste messbare Majorante* von f , also die fast sicher eindeutig bestimmte, messbare Funktion mit den Eigenschaften $f^* \geq f$ f.s. und $f^* \leq g$ f.s. für alle messbaren $g : \Omega \rightarrow \overline{\mathbb{R}}$ mit $g \geq f$ f.s.. Jedes $f : \Omega \rightarrow \overline{\mathbb{R}}$ besitzt eine kleinste messbare Majorante (vgl. [VW96], Lemma 1.2.1). Wir schreiben f^* für jede Version der kleinsten messbaren Majorante.

Im Folgenden sei (\mathbb{D}, ρ_D) ein metrischer Raum, $\mathfrak{B}(\mathbb{D})$ die Borelsche σ -Algebra auf \mathbb{D} , also die von den offenen Teilmengen von \mathbb{D} erzeugte σ -Algebra auf \mathbb{D} . Ein Maß μ auf $\mathfrak{B}(\mathbb{D})$ heißt *Borelmaß*. Ein Wahrscheinlichkeitsmaß P auf $\mathfrak{B}(\mathbb{D})$ heißt *straff*, wenn es zu jedem $\varepsilon > 0$ eine kompakte Menge K gibt, so dass $P(K) \geq 1 - \varepsilon$ gilt. Ferner nennen wir eine Zufallsvariable $X : \Omega \rightarrow \mathbb{D}$ *straff*, falls das Bildmaß P^X *straff* ist. Außerdem nennen wir P *separabel*, falls eine separable Teilmenge $M \in \mathfrak{B}(\mathbb{D})$ mit $P(M) = 1$ existiert. Maßgeblich für die weiteren Begriffe ist die folgende Definition der schwachen Konvergenz.

B.1.1. Definition. Sei $(\Omega_n, \mathcal{A}_n, P_n)_{n \in \mathbb{N}}$ eine Folge von Wahrscheinlichkeitsräumen und $X_n : \Omega_n \rightarrow \mathbb{D}$, $n \in \mathbb{N}$, eine Folge von Abbildungen. Dann heißt $\{X_n\}_{n \in \mathbb{N}}$ *schwach konver-*

gent gegen ein Borelmaß P auf $\mathfrak{B}(\mathbb{D})$, falls

$$E^* f(X_n) \longrightarrow \int f dP, \quad n \rightarrow \infty,$$

für jede stetige und beschränkte Funktion $f : \mathbb{D} \rightarrow \mathbb{R}$ gilt. Wir schreiben dann $X_n \rightsquigarrow P$. Ist P die Verteilung einer Zufallsvariablen X , so schreiben wir auch $X_n \xrightarrow{v} X$.

Mittels des äußeren Erwartungswerts wird auf diese Weise Verteilungskonvergenz auch für solche Abbildungen formuliert, die selbst nicht Borel-messbar sind.

B.1.2. Satz ([VW96], **Theorem 1.3.4**). *Seien (Ω, \mathcal{A}, P) ein Wahrscheinlichkeitsraum, $X_n, X : \Omega \rightarrow \mathbb{D}$, $n \in \mathbb{N}$, X Borel-messbar. Die folgenden Eigenschaften sind äquivalent:*

- a) $X_n \xrightarrow{v} X$;
- b) $\liminf_{n \rightarrow \infty} P_*(X_n \in G) \geq P(X \in G)$ für alle offenen Mengen $G \subset \mathbb{D}$;
- c) $\limsup_{n \rightarrow \infty} P^*(X_n \in G) \leq P(X \in G)$ für alle abgeschlossenen Mengen $G \subset \mathbb{D}$.

B.1.3. Definition. Eine Folge $\{X_n\}_{n \in \mathbb{N}}$ von Abbildungen heißt *asymptotisch messbar*, falls für $n \rightarrow \infty$ und jede stetige und beschränkte Funktion f

$$E^* f(X_n) - E_* f(X_n) \rightarrow 0$$

gilt. Die Folge der Abbildungen $\{X_n\}_{n \in \mathbb{N}}$ heißt *asymptotisch straff*, falls für alle $\epsilon > 0$ eine kompakte Menge K existiert, so dass für jedes $\delta > 0$

$$\liminf P_*(X_n \in K^\delta) \geq 1 - \epsilon$$

ist. Dabei ist $K^\delta = \{x \in \mathbb{D} : \rho_D(x, K) < \delta\}$.

B.1.4. Lemma ([VW96], **Lemma 1.3.8**). *Seien $X_n, X : \Omega \rightarrow \mathbb{D}$ Abbildungen. Dann gilt*

- a) *Ist $X_n \xrightarrow{v} X$, so ist $\{X_n\}_{n \in \mathbb{N}}$ asymptotisch messbar.*
- b) *Sei $X_n \xrightarrow{v} X$. Es ist $\{X_n\}_{n \in \mathbb{N}}$ genau dann asymptotisch straff, wenn der Grenzwert X straff ist.*

B.1.5. Definition. Sei (Ω, \mathcal{A}, P) ein Wahrscheinlichkeitsraum, $X_n, X : \Omega \rightarrow \mathbb{D}$ seien Abbildungen. Dann konvergiert X_n *fast gleichmäßig* gegen X - in Zeichen $X_n \xrightarrow{au} X$ - falls gilt:

$$\forall \varepsilon > 0 \exists A \in \mathcal{A} : P(A) \geq 1 - \varepsilon \text{ und } \sup_{\omega \in A} \rho_D(X_n(\omega), X(\omega)) \rightarrow 0.$$

Außerdem schreiben wir $X_n \xrightarrow{f.s.*} X$, falls $\rho_D(X_n, X)^* \rightarrow 0$ f.s. gilt.

Ist X Borel-messbar, so ist $X_n \xrightarrow{au} X$ äquivalent zu $X_n \xrightarrow{f.s.*} X$ (vgl. [VW96], Lemma 1.9.2).

B.1.6. Satz ([VW96], Theorem 1.10.4). *Seien $(\Omega_n, \mathcal{A}_n, P_n)_{n \in \mathbb{N}}$, (Ω, \mathcal{A}, P) Wahrscheinlichkeitsräume und für $n \in \mathbb{N}$ seien $X_n : \Omega_n \rightarrow \mathbb{D}$, $X : \Omega \rightarrow \mathbb{D}$ Abbildungen. Sei $X_n \xrightarrow{v} X$ und P^X separabel. Dann existieren ein Wahrscheinlichkeitsraum $(\hat{\Omega}, \hat{\mathcal{A}}, \hat{P})$ und Abbildungen $\hat{X}_n, \hat{X} : \hat{\Omega} \rightarrow \mathbb{D}$, $n \in \mathbb{N}$, mit der Eigenschaft*

a) $\hat{X}_n \xrightarrow{au} \hat{X}$,

b) $E^* f(\hat{X}_n) = E^* f(X_n)$ und $E^* f(\hat{X}) = E^* f(X)$ für jedes $n \in \mathbb{N}$ und alle beschränkten Funktionen $f : \mathbb{D} \rightarrow \mathbb{R}$.

Darüber hinaus können die Abbildungen \hat{X}_n bzw. \hat{X} derart gewählt werden, dass sie keine anderen Werte als die Abbildungen X_n bzw. X annehmen. Ist ferner für ein $n \in \mathbb{N}$ die Abbildung X_n messbar, so kann \hat{X}_n messbar gewählt werden und ist die Abbildung X messbar, so kann \hat{X} messbar gewählt werden. (vgl. [VW96], Addendum 1.10.5).

Für zwei Abbildungen $\hat{X} : \hat{\Omega} \rightarrow \mathbb{D}$ und $X : \Omega \rightarrow \mathbb{D}$, die die im letzten Satz genannte Eigenschaft erfüllen, dass $E^* f(\hat{X}) = E^* f(X)$ für alle beschränkten Funktionen $f : \mathbb{D} \rightarrow \mathbb{R}$ zutrifft, verwenden wir auch die Schreibweise $\hat{X} \sim X$.

B.2 Empirische Prozesse in $l^\infty(T)$

Sei T eine Menge, $l^\infty(T)$ der Raum der beschränkten und reellen Funktionen auf T . Dann ist $(l^\infty(T), \rho)$ mit $\rho(x, y) := \|x - y\|_\infty$ ein metrischer Raum. Das für uns wichtigste Beispiel von Abbildungen mit Werten in $l^\infty(T)$ ist das des empirischen Prozesses:

Sind X_1, X_2, \dots unabhängige und identisch verteilte Zufallsvariablen mit Werten in einem Messraum (R, \mathfrak{G}) , $P = P^{X_1}$, $P_n = \frac{1}{n} \sum_{k=1}^n \delta_{X_k}$, und ist \mathcal{F} eine Menge messbarer und bezüglich P quadratintegrierbarer Funktionen $f : R \rightarrow \mathbb{R}$, so heißt

$$(\mathbb{F}_n f, f \in \mathcal{F}) := (\sqrt{n}(P_n - P)f, f \in \mathcal{F})$$

der durch \mathcal{F} indizierte empirische Prozess der X_1, \dots, X_n . Dabei setzen wir $Pf := \int f dP$. Ist $\mathcal{F} = \{f_t; t \in T\}$ so sprechen wir auch vom durch T indizierten empirischen Prozess. Die endlichdimensionalen Randverteilungen von \mathbb{F}_n konvergieren nach dem mehrdimensionalen zentralen Grenzwertsatz gegen eine Normalverteilung. Genauer gilt für beliebige $f_1, \dots, f_k \in \mathcal{F}$

$$(\mathbb{F}_n f_1, \dots, \mathbb{F}_n f_k) \xrightarrow{v} N_k(0, \Sigma), \quad n \rightarrow \infty,$$

mit $\Sigma = (\Sigma_{ij})_{1 \leq i, j \leq k}$, $\Sigma_{ij} = P((f_i - Pf_i)(f_j - Pf_j))$. Nach dem Existenzsatz von Kolmogorov existiert ein zentrierter Gaußprozess \mathbb{G} der die Verteilungen $N_k(0, \Sigma)$ mit $\Sigma = (\Sigma_{ij})_{1 \leq i, j \leq k}$, $\Sigma_{ij} = P((f_i - Pf_i)(f_j - Pf_j))$, als k -dimensionale Randverteilungen hat. Unter der Bedingung

$$\sup_{f \in \mathcal{F}} |f - Pf| < \infty$$

kann die Pfadmenge von \mathbb{F}_n und \mathbb{G} als Teilmenge von $l^\infty(\mathcal{F})$ aufgefasst werden. Das Grenzwverhalten von Abbildungen mit Werten in $l^\infty(T)$ ist daher von besonderem Interesse.

B.2.1. Satz ([VW96], Theorem 1.5.4). Sei $X_n : \Omega_n \rightarrow l^\infty(T)$, $n \in \mathbb{N}$. Dann konvergiert $\{X_n\}_{n \in \mathbb{N}}$ genau dann schwach gegen einen straffen Grenzwert, wenn X_n asymptotisch straff ist und $(X_n(t_1), \dots, X_n(t_k))$ für sämtliche Auswahlen $\{t_1, \dots, t_k\} \subset T$, $k \in \mathbb{N}$, schwach konvergiert.

Ist $\{X_n\}_{n \in \mathbb{N}}$ asymptotisch straff und konvergiert $(X_n(t_1), \dots, X_n(t_k))$ für sämtliche Auswahlen $\{t_1, \dots, t_k\} \subset T$, $k \in \mathbb{N}$, schwach gegen die Randverteilungen $(X(t_1), \dots, X(t_k))$ eines stochastischen Prozesses $X = (X(t), t \in T)$, so existiert eine Version von X mit beschränkten Pfaden und $X_n \xrightarrow{v} X$.

B.2.2. Lemma ([VW96], Lemma 1.5.3). Seien X und Y straffe Borel-messbare Abbildungen in $l^\infty(T)$. Dann sind die Maße P^X und P^Y genau dann gleich, wenn alle endlichdimensionalen Randverteilungen von X und Y gleich sind.

B.2.3. Definition. Es sei ρ eine Semimetrik auf T .

- a) Die Menge T heißt *total beschränkt* bezüglich ρ , falls es für jedes $\varepsilon > 0$ endlich viele Elemente $t_1, \dots, t_n \in T$ gibt mit der Eigenschaft, dass $T = \bigcup_{k=1}^n B(t_k, \varepsilon)$ gilt, wobei für $k = 1, \dots, n$ $B(t_k, \varepsilon) = \{t \in T : \rho(t_k, t) < \varepsilon\}$ die ε -Umgebung von t_k ist.

b) Eine Folge $X_n : \Omega_n \rightarrow l^\infty(T)$, $n \in \mathbb{N}$, heißt *asymptotisch gleichgradig gleichmäßig ρ -stetig in Wahrscheinlichkeit*, falls für alle $\varepsilon, \eta > 0$ ein $\delta > 0$ existiert mit

$$\limsup_{n \rightarrow \infty} P^* \left(\sup_{\rho(s,t) < \delta} |X_n(s) - X_n(t)| > \varepsilon \right) < \eta.$$

B.2.4. Satz ([VW96], Satz 1.5.7). *Eine Folge $X_n : \Omega_n \rightarrow l^\infty(T)$ ist genau dann asymptotisch straff, wenn $X_n(t)$ asymptotisch straff in \mathbb{R} für jedes $t \in T$ ist und eine Semimetrik ρ existiert, so dass (T, ρ) total beschränkt ist und X_n asymptotisch gleichmäßig gleichgradig ρ -stetig in Wahrscheinlichkeit ist.*

Das folgende Resultat charakterisiert die Straffheit von *Gaußprozessen* mit Pfaden in $l^\infty(T)$, die häufig als Grenzwerte von Folgen von Abbildungen mit Werten in $l^\infty(T)$ auftreten. Ein stochastischer Prozess $(X(t), t \in T)$ heißt bekanntlich Gaußprozess, falls alle endlichdimensionalen Randverteilungen $(X(t_1), \dots, X(t_k))$, $t_k \in T$, $k \in \mathbb{N}$, k -dimensionale Normalverteilungen in \mathbb{R}^k sind.

B.2.5. Satz ([VW96], 1.5.10). *Ein Borel-messbarer Gaußprozess X mit Pfaden in $l^\infty(T)$ ist genau dann straff, falls für*

$$\rho_p(s, t) := (E|X(s) - X(t)|^p)^{\frac{1}{\max(1,p)}}, \quad s, t \in T, 0 < p < \infty,$$

gilt: (T, ρ_p) ist total beschränkt und die Pfade von X sind fast sicher gleichmäßig ρ_p -stetig für ein (und somit für alle) $0 < p < \infty$.

B.2.1 VC- und Donsker-Klassen

Die Bemerkungen des letzten Abschnittes und die Aussagen in Satz B.2.1 zeigen, dass für die Existenz eines straffen und Borel-messbaren Prozesses \mathbb{G} mit $\mathbb{F}_n \xrightarrow{v} \mathbb{G}$ die asymptotische Straffheit von $(\mathbb{F}_n f, f \in \mathcal{F})$ entscheidend ist. Für empirische Prozesse hängt diese Eigenschaft eng mit der Funktionenklasse \mathcal{F} zusammen. Funktionenklassen, für die ein solches \mathbb{G} existiert nennen wir *Donsker*, oder präziser *P -Donsker* um die Abhängigkeit vom zugrunde liegenden Wahrscheinlichkeitsmaß zu unterstreichen. Ist \mathcal{F} Donsker bezüglich jedem P , so heißt \mathcal{F} gleichmäßig Donsker.

Die Größe der Funktionenklasse \mathcal{F} ist bestimmend für die asymptotische Straffheit des empirischen Prozesses. Eine Möglichkeit die Größe von Funktionenklassen zu messen sind sogenannte Überdeckungszahlen. Dazu definieren wir zunächst eine *Envelope von \mathcal{F}*

als eine Funktion F mit der Eigenschaft, dass für alle x und alle $f \in \mathcal{F}$ die Beziehung $F(x) \geq |f(x)|$ besteht. Wegen $|f(x)| \leq F(x) := \sup_{f \in \mathcal{F}} f(x)$ für alle x existiert immer eine Envelope von \mathcal{F} .

B.2.1.1. Definition. Sei $(\mathcal{F}, \|\cdot\|)$ ein normierter Raum, $\varepsilon > 0$ und F eine Envelope von \mathcal{F} . Die *Überdeckungszahl* $N(\varepsilon, \mathcal{F}, \|\cdot\|)$ ist die kleinste Anzahl von offenen ε -Umgebungen mit denen sich \mathcal{F} überdecken lässt, d.h. es ist

$$N(\varepsilon, \mathcal{F}, \|\cdot\|) = \inf \left\{ k \in \mathbb{N} \mid \exists f_1, \dots, f_k : R \rightarrow \mathbb{R}, \|f_j\| < \infty, \mathcal{F} \subset \bigcup_{j=1}^k \{g : \|f_j - g\| < \varepsilon\} \right\}$$

Es sei Q ein Wahrscheinlichkeitsmaß auf $\mathcal{B}(\mathbb{R})$ und für $r \geq 1$ sei $L_r(Q)(f) := \|f\|_{Q,r} = (\int |f|^r dQ)^{\frac{1}{r}}$. Die *Entropiezahl* von \mathcal{F} bezüglich L_r ist definiert als

$$\sup_Q \log N(\varepsilon \|F\|_{Q,r}, \mathcal{F}, L_r(Q)),$$

wobei das Supremum über alle Wahrscheinlichkeitsmaße Q mit $0 < QF^r < \infty$ gebildet wird.

B.2.1.2. Definition ([VW96], Definition 2.3.3). Sei \mathcal{F} eine Klasse messbarer reeller Funktionen definiert auf einem Wahrscheinlichkeitsraum (Ω, \mathcal{A}, P) . Für ein festes $n \in \mathbb{N}$ sei $(\Omega^n, \overline{\mathcal{A}^n}, \overline{P^n})$ die P^n -Vervollständigung des Produktraums $(\Omega^n, \mathcal{A}^n, P^n)$ und für $(e_1, \dots, e_n) \in \mathbb{R}^n$ sei

$$H_{e_1, \dots, e_n} : (\Omega^n, \overline{\mathcal{A}^n}, \overline{P^n}) \rightarrow \mathbb{R}$$

$$(x_1, \dots, x_n) \mapsto \sup_{f \in \mathcal{F}} \left| \sum_{i=1}^n e_i f(x_i) \right|.$$

Dann heißt \mathcal{F} P -messbare Klasse, falls H_{e_1, \dots, e_n} für jedes $n \in \mathbb{N}$ und alle $(e_1, \dots, e_n) \in \mathbb{R}^n$ $(\overline{\mathcal{A}^n}, \mathcal{B}(\mathbb{R}))$ -messbar ist.

Für unsere Zwecke häufig ausreichend ist die Tatsache, dass eine Klasse \mathcal{F} messbarer reeller Funktionen dann eine für jedes P P -messbare Klasse ist, falls eine abzählbare Teilmenge $\mathcal{G} \subset \mathcal{F}$ existiert mit

$$\forall f \in \mathcal{F} \exists \{g_n\}_{n \in \mathbb{N}} \subset \mathcal{G} \forall x \in R : \lim_{n \rightarrow \infty} g_n(x) = f(x).$$

Siehe dazu [VW96], 2.3.4. Insbesondere trifft das Letztgenannte etwa auf Klassen von (auf kompakten metrischen Räumen definierten) stetigen und reellwertigen Funktionen zu.

B.2.1.3. Satz ([VW96], Theorem 2.5.2). Sei \mathcal{F} eine Klasse messbarer Funktionen. Es seien $\{f - g : f, g \in \mathcal{F}, \|f - g\|_{P,2} < \delta\}$ und $\{(f - g)^2 : f, g \in \mathcal{F}\}$ P -messbare Klassen für alle $\delta > 0$. Gilt für eine bezüglich P quadratintegrierbare Envelope F von \mathcal{F}

$$\int_0^\infty \sup_Q \sqrt{\log N(\varepsilon \|F\|_{Q,2}, \mathcal{F}, L_2(Q))} d\varepsilon < \infty, \quad (\text{B.2.1})$$

so ist \mathcal{F} P -Donsker.

Die Bedingung (B.2.1) ist insbesondere für die kombinatorisch motivierten sogenannten VC-Funktionenklassen erfüllt:

B.2.1.4. Definition. Sei M eine Menge und $\mathfrak{C} \subset \mathfrak{P}(M)$. Für $x_1, \dots, x_n \in M$ sei

$$\Delta_n(\mathfrak{C}, x_1, \dots, x_n) = \#\{C \cap \{x_1, \dots, x_n\} : C \in \mathfrak{C}\}$$

und

$$V(\mathfrak{C}) = \inf\{n \in \mathbb{N} : \max_{x_1, \dots, x_n} \Delta_n(\mathfrak{C}, x_1, \dots, x_n) < 2^n\}, \quad \inf_\emptyset := \infty.$$

Ist $V(\mathfrak{C}) < \infty$, so heißt \mathfrak{C} eine *VC-Klasse* vom Index $V(\mathfrak{C})$.

Sei \mathcal{F} eine Menge messbarer Funktionen $f : M \rightarrow \mathbb{R}$. Dann heißt \mathcal{F} *VC-Funktionenklasse*, falls die Menge der Subgraphen $\{(x, t) \in M \times \mathbb{R} : t < f(x)\}, f \in \mathcal{F}$ eine VC-Klasse in $M \times \mathbb{R}$ bilden.

B.2.1.5. Satz ([VW96], Theorem 2.6.7). Sei \mathcal{F} eine VC-Funktionenklasse mit messbarer Envelope F . Dann erfüllt \mathcal{F} die Bedingung (B.2.1) für jedes Wahrscheinlichkeitsmaß Q mit $\|F\|_{Q,2} > 0$.

Der folgende Satz ergibt sich aus [VW96], Theorem 2.4.3, und der zugehörigen Bemerkung auf Seite 125 in [VW96].

B.2.1.6. Satz. Sei P ein Wahrscheinlichkeitsmaß und P_n das zugehörige empirische Maß. Sei \mathcal{F} eine P -messbare VC-Funktionenklasse mit messbarer Envelope F . Es gelte $PF < \infty$. Dann gilt

$$\left(\sup_{f \in \mathcal{F}} |(P_n - P)f| \right)^* \rightarrow 0 \quad \text{fast sicher.}$$

B.2.2 Konvergenzverhalten bei Dreiecksschemata

Für jedes $n \in \mathbb{N}$ seien X_{n1}, \dots, X_{nn} unabhängige und identisch verteilte Zufallsvariablen je mit der Verteilung P_n und Werten in (R, \mathcal{S}) . Sei $\mathbb{P}_n = \frac{1}{n} \sum_{k=1}^n \delta_{X_{nk}}$ das empirische Maß zu X_{n1}, \dots, X_{nn} . Wir definieren für $f, g \in \mathcal{F}$ $\rho_P(f, g) := \rho_P(f - g)$ mit

$$\rho_P(f) := (P(f - Pf)^2)^{\frac{1}{2}}.$$

Es ist ρ_P eine Semimetrik auf \mathcal{F} .

B.2.2.1. Satz ([VW96], Theorem 2.8.9). *Sei \mathcal{F} eine Klasse messbarer Funktionen. Seien $\{f - g : f, g \in \mathcal{F}, \|f - g\|_{P,2} < \delta\}$ und $\{(f - g)^2 : f, g \in \mathcal{F}\}$ für jedes $n \in \mathbb{N}$ P_n -messbare Klassen für alle $\delta > 0$. Sei F eine bezüglich jedem P_n quadratintegrierbare Envelope von \mathcal{F} und sei $P_n F^2 = \mathcal{O}(1)$. Es gelte (B.2.1),*

$$\forall \varepsilon > 0 : \quad \limsup_{n \rightarrow \infty} P_n F^2 \{F \geq \varepsilon \sqrt{n}\} = 0,$$

und

$$\sup_{f, g \in \mathcal{F}} |\rho_{P_n}(f, g) - \rho_{P_0}(f, g)| \rightarrow 0, \quad n \rightarrow \infty. \quad (\text{B.2.2})$$

für ein Wahrscheinlichkeitsmaß P_0 auf \mathcal{S} . Dann gilt

$$(\sqrt{n}(P_n - P_n)f, f \in \mathcal{F}) \xrightarrow{v} (\mathbb{G}f, f \in \mathcal{F}) \quad \text{in } l^\infty(\mathcal{F}),$$

mit einem zentrierten Gaußprozess \mathbb{G} mit Kovarianzfunktion $\text{Cov}(\mathbb{G}f, \mathbb{G}g) = P_0fg - P_0fP_0g, f, g \in \mathcal{F}$.

B.2.2.2. Bemerkung. Im Fall, dass $\sup_{f \in \mathcal{F}} |P_n f| < \infty$ für alle $n \in \mathbb{N}$ und $\sup_{f \in \mathcal{F}} |P_0 f| < \infty$ ist, lässt sich die Bedingung (B.2.2) im obigen Satz durch die Bedingung

$$\sup_{f, g \in \mathcal{F}} |L_2(P_n)(f - g) - L_2(P_0)(f - g)| \rightarrow 0, \quad n \rightarrow \infty,$$

ersetzen. Vergleiche dazu [VW96], Problem 2.1.1-2.1.3.

Anhang C

Benachbarte Alternativen

Es sei (R_n, \mathfrak{S}_n) , $n \in \mathbb{N}$, eine Folge von Messräumen und für jedes $n \in \mathbb{N}$ seien P_n und Q_n Wahrscheinlichkeitsmaße auf \mathfrak{S}_n .

C.1. Definition. Die Folge $\{Q_n\}_{n \in \mathbb{N}}$ heißt *benachbart* zu $\{P_n\}_{n \in \mathbb{N}}$, falls für jede Folge $\{A_n\}_{n \in \mathbb{N}}$ von Ereignissen mit $A_k \in \mathfrak{S}_k$, $k \in \mathbb{N}$, gilt:

$$\lim_{n \rightarrow \infty} P_n(A_n) = 0 \quad \Rightarrow \quad \lim_{n \rightarrow \infty} Q_n(A_n) = 0.$$

Die Folgen $\{P_n\}_{n \in \mathbb{N}}$ und $\{Q_n\}_{n \in \mathbb{N}}$ heißen (*beidseitig*) *benachbart*, falls $\{Q_n\}_{n \in \mathbb{N}}$ benachbart zu $\{P_n\}_{n \in \mathbb{N}}$ und $\{P_n\}_{n \in \mathbb{N}}$ benachbart zu $\{Q_n\}_{n \in \mathbb{N}}$ ist.

Wir nehmen im Folgenden an, dass die Maße Q_n und P_n für jedes $n \in \mathbb{N}$ von einem σ -endlichen Maß μ_n dominiert werden (z.B. $\mu_n = P_n + Q_n$, $n \in \mathbb{N}$). Es seien für $n \in \mathbb{N}$ jeweils $q_n := \left(\frac{dQ_n}{d\mu_n}\right)$ und $p_n := \left(\frac{dP_n}{d\mu_n}\right)$ die μ_n -Dichten von Q_n bzw. P_n und weiter sei für alle $n \in \mathbb{N}$

$$\Lambda_n := \log(L_n), \quad L_n := \begin{cases} q_n/p_n, & p_n > 0, \\ 1, & q_n = p_n = 0, \\ \infty, & q_n > 0, p_n = 0. \end{cases}$$

Dann ist Λ_n der Log-Likelihoodquotient von P_n und Q_n .

Im Zusammenhang mit dem zentralen Grenzwertsatz liefern die folgenden *Le Cam Lemmata* häufig nützliche Hilfsmittel um die Benachbarkeit von Maßen zu verifizieren:

C.2. Lemma ([WMF95], Korollar 6.124). *Gilt in der oben beschriebenen Situation für $\infty > \sigma^2 > 0$*

$$\Lambda_n \xrightarrow{v} N\left(-\frac{1}{2}\sigma^2, \sigma^2\right), \quad n \rightarrow \infty,$$

unter P_n , so ist $\{Q_n\}_{n \in \mathbb{N}}$ benachbart zu $\{P_n\}_{n \in \mathbb{N}}$. Außerdem gilt dann

$$\Lambda_n \xrightarrow{v} N\left(\frac{1}{2}\sigma^2, \sigma^2\right), \quad n \rightarrow \infty,$$

unter Q_n und $\{P_n\}_{n \in \mathbb{N}}$ ist benachbart zu $\{Q_n\}_{n \in \mathbb{N}}$.

C.3. Satz ([VW96], Theorem 3.10.7). Für jedes $n \in \mathbb{N}$ seien P_n und Q_n Wahrscheinlichkeitsmaße auf \mathfrak{S}_n und $X_n : R_n \rightarrow \mathbb{D}$ eine Abbildung mit Werten im metrischen Raum (\mathbb{D}, ρ) . Es sei $\{Q_n\}_{n \in \mathbb{N}}$ benachbart zu $\{P_n\}_{n \in \mathbb{N}}$ und es gelte für eine Zufallsvariable (X, Y) mit Werten in $\mathbb{D} \times \mathbb{R}$

$$\left(X_n, \frac{dQ_n}{dP_n}\right) \xrightarrow{v} (X, Y), \quad n \rightarrow \infty,$$

unter $\{P_n\}_{n \in \mathbb{N}}$. Dann gilt $P^{X_n} \rightsquigarrow L$ unter $\{Q_n\}_{n \in \mathbb{N}}$, mit einem Borelmaß L auf \mathbb{D} , definiert durch $L(B) = E(\mathbf{1}_B(X)Y)$ für alle $B \in \mathcal{B}(\mathbb{D})$. Ist X straff oder separabel, so auch L .

Ist speziell $X_n : (R_n, \mathfrak{S}_n) \rightarrow (\mathbb{R}^d, \mathcal{B}^d)$ und gilt unter P_n

$$(X_n, \Lambda_n) \xrightarrow{v} N_{d+1} \left(\begin{pmatrix} a \\ -\frac{1}{2}\sigma^2 \end{pmatrix}, \begin{pmatrix} \Sigma & c \\ c' & \sigma^2 \end{pmatrix} \right), \quad n \rightarrow \infty,$$

für geeignete Vektoren $a, c \in \mathbb{R}^d$, eine symmetrische und positiv definite Matrix $\Sigma \in \mathbb{R}^{d \times d}$ und $\sigma^2 > 0$, so gilt unter Q_n

$$X_n \xrightarrow{v} N_d(a + c, \Sigma), \quad n \rightarrow \infty.$$

Anhang D

Verschiedenes

D.1 Gaußmaße

D.1.1. Satz ([Tsi75]). *Es sei $\{X_n\}_{n \in \mathbb{N}}$ eine Folge von Zufallsvariablen mit $X_k \sim N(\mu_k, \sigma_k^2)$, wobei $\mu_k \geq 0$ für alle $k \in \mathbb{N}$ und $\sigma_k^2 \neq 0$ für mindestens ein $k \in \mathbb{N}$ gelte. Dann ist die Funktion $F(a) := P(\sup_{k \in \mathbb{N}} X_k < a)$, $a \in \mathbb{R}$, stetig, außer eventuell im Punkt $a_0 := \inf\{a : F(a) > 0\}$. Die Funktion F ist absolut stetig auf (a_0, ∞) und für die auf (a_0, ∞) fast überall definierte Ableitung F' gilt*

$$F'(x) \geq \frac{1}{c} \Phi' \left(\frac{x - c_1}{c} \right), \quad x \in (a_0, \infty),$$

fast überall, mit Φ' als Dichte der Standardnormalverteilung und $c^2 := \sup_{k \in \mathbb{N}} \text{Var}(X_k)$, sowie $c_1 := \lim_{p \uparrow 1} (F^{-1}(p) - c\Phi^{-1}(p)) \in [0, \infty)$.

Für die im Folgenden benötigten Definitionen des *abstrakten Wienertraums* und des *Wienermaßes*, sowie der in diesem Kontext verwendeten Begriffe *Inklusionsabbildung* und *kanonisches Gaußmaß* siehe z.B. [Kuo70]. Wir führen hier aus Gründen der Übersichtlichkeit nur die verwendeten Resultate auf:

D.1.2. Satz ([BM86], Proposition 2). *Es sei (i, H, B) ein abstrakter Wienertraum. Sei P das kanonische Gaußmaß auf B . Dann hat die Verteilungsfunktion der auf B definierten Normabbildung $x \mapsto \|x\|$ unter P auf $[0, \infty)$ eine Dichte und eine auf $[0, \infty)$ streng monoton wachsende Verteilungsfunktion.*

D.1.3. Satz ([Kue70]). *Sei B ein separabler Banachraum und P ein Gaußmaß auf B mit $P(M) > 0$ für alle offenen Mengen $\emptyset \neq M \subset B$. Dann existiert ein reeller separabler Hilbertraum H , so dass (i, H, B) , $i : H \rightarrow B$ die Inklusionsabbildung, ein abstrakter Wienertraum ist und P ein Wienermaß auf B ist.*

Ist die Bedingung $P(M) > 0$ für alle offenen Mengen $\emptyset \neq M \subset B$ im vorherigen Satz nicht erfüllt, ist folgendes Resultat nützlich (siehe [Sat69], Theorem 2):

Ist $B_0 \subset B$ der Träger von P , also der kleinste abgeschlossene lineare Teilraum von B mit $P(B_0) = 1$, so ist $P(M) > 0$ für alle offenen Mengen $\emptyset \neq M \subset B_0$. Die Aussage des Satzes trifft dann für das auf die Borelsche σ -Algebra von B_0 eingeschränkte Gaußmaß P zu, welches dann ein Gaußmaß auf B_0 ist.

D.2 Separable Prozesse

Es sei (Ω, \mathcal{A}, P) ein Wahrscheinlichkeitsraum. Weiter sei S ein kompakter metrischer Raum, \mathfrak{S} die Borelsche σ -Algebra auf S und T Teilmenge eines separablen metrischen Raums.

D.2.1. Definition ([AG75], 4.1.2). *Es sei $X = (X(t), t \in T)$ ein stochastischer Prozess, $X(t) : (\Omega, \mathcal{A}) \rightarrow (S, \mathfrak{S})$, $t \in T$. Dann heißt X separabler stochastischer Prozess, falls es eine abzählbare dichte Teilmenge $T_0 \subset T$ und eine P -Nullmenge $N \in \mathcal{A}$ gibt, so dass für alle $\omega \notin N$ und alle $t \in T$ eine Folge $\{t_k\}_{k \in \mathbb{N}} \subset T_0$ existiert mit $t_k \rightarrow t$ und $X(t_k, \omega) \rightarrow X(t, \omega)$ für $k \rightarrow \infty$.*

D.2.2. Satz ([AG75], Theorem 4.1.5). *Es sei $X = (X(t), t \in T)$ ein stochastischer Prozess, $X(t) : (\Omega, \mathcal{A}) \rightarrow (S, \mathfrak{S})$, $t \in T$. Sei $N \in \mathcal{A}$ eine P -Nullmenge mit der Eigenschaft, dass für alle $\omega \notin N$ die Abbildung $t \mapsto X(\omega, t)$ stetig ist. Dann ist X ein separabler stochastischer Prozess.*

D.3 Hilfssätze

D.3.1. Satz. *Sei $f : \mathbb{R}^d \rightarrow \mathbb{R}$ eine analytische Funktion, $f \not\equiv 0$. Sei $A = \{\mathbf{t} \in \mathbb{R}^d : f(\mathbf{t}) = 0\}$. Dann hat A das Maß Null bezüglich des d -dimensionalen Lebesgue-Borelschen-Maßes λ^d , es gilt also $\lambda^d(A) = 0$.*

Beweis. Für führen den Beweis per Induktion über d . Sei $d = 1$. Angenommen A ist überabzählbar. Dann hat A als überabzählbare Teilmenge der reellen Zahlen einen reellen Häufungspunkt. Daraus ergibt sich bereits $f \equiv 0$, im Widerspruch zur Voraussetzung. Also ist A abzählbar und somit $\lambda^1(A) = 0$. Sei nun die Behauptung für $d \in \mathbb{N}$ bereits gezeigt und sei

$$\begin{aligned} f : \mathbb{R} \times \mathbb{R}^d &\rightarrow \mathbb{R} \\ (x, \mathbf{t}) &\mapsto f(x, \mathbf{t}) \end{aligned}$$

analytisch und nicht identisch Null. Wir setzen

$$A_x := \{\mathbf{t} \in \mathbb{R}^d : f(x, \mathbf{t}) = 0\}, \quad x \in \mathbb{R}.$$

Da f analytisch ist, ist auch die Funktion $f(x, \cdot) : \mathbb{R}^d \rightarrow \mathbb{R}$ für jedes feste $x \in \mathbb{R}$ analytisch und hat die Nullstellenmenge A_x . Es folgt

$$\lambda^d(A_x) = 0 \quad \forall x \in \mathbb{R} : f(x, \cdot) \not\equiv 0.$$

Sei nun $N := \{x \in \mathbb{R} : f(x, \cdot) \equiv 0\}$. Wir zeigen $\lambda(N) = 0$: Wegen $f \not\equiv 0$ existiert zunächst ein $\mathbf{t}_0 \in \mathbb{R}^d$ mit $f(\cdot, \mathbf{t}_0) \not\equiv 0$. Es ist $f(\cdot, \mathbf{t}_0)$ analytisch mit

$$\forall x \in N : f(x, \mathbf{t}_0) = 0,$$

also ist N Teilmenge der Nullstellenmenge der analytischen und nicht identisch verschwindenden Funktion $f(\cdot, \mathbf{t}_0)$. Aus dem Induktionsanfang folgt $\lambda(N) = 0$. Schließlich erhalten wir mit dem Satz von Fubini

$$\begin{aligned} \lambda^{d+1}(A) &= \int_{\mathbb{R}} \lambda^d(A_x) d\lambda(x) \\ &= \int_N \lambda^d(A_x) d\lambda(x) + \int_{N^c} \lambda^d(A_x) d\lambda(x) \\ &= 0, \end{aligned}$$

was zu zeigen war. □

Wir bemerken, dass die gleichen Argumente des vorherigen Satzes auch auf reellwertige Funktionen angewendet werden können, die analytisch auf einer offenen Teilmenge $D \subset \mathbb{R}^d$ sind. Eine ähnliche Aussage wie die des letzten Satzes, allerdings für analytische Funktionen $f : D \rightarrow \mathbb{C}$, $D \subset \mathbb{C}^d$ offen, findet sich in [GR65], Chapter I, Corollary 10.

Der dortige Beweis kann allerdings nicht ohne Weiteres auf die hier vorliegende Situation übertragen werden.

D.3.2. Satz. *Es sei (K, ρ) ein kompakter metrischer Raum, $\{a_n\}_{n \in \mathbb{N}} \subset \mathbb{R}_{>0}$ eine Folge positiver reeller Zahlen mit $\lim_{n \rightarrow \infty} a_n = \infty$. Seien $f_n : K \rightarrow \mathbb{R}$ eine Folge von Funktionen, $f, h : K \rightarrow \mathbb{R}$ stetig, $h \leq 0$, mit $\lim_{n \rightarrow \infty} \sup_{t \in K} |f_n(t) - f(t)| = 0$. Sei $A := \{t \in K; h(t) = 0\} \neq \emptyset$. Dann gilt*

$$\lim_{n \rightarrow \infty} \sup_{t \in K} [f_n(t) + a_n h(t)] = \sup_{t \in A} f(t).$$

Beweis. Sei $\varepsilon > 0$ gegeben. Wegen der gleichmäßigen Konvergenz der f_n gegen f finden wir dann ein $n_0 \in \mathbb{N}$ derart, dass $-\varepsilon + f(t) \leq f_n(t) \leq f(t) + \varepsilon$ für alle $t \in K$, $n \geq n_0$, gilt. Daher erhalten wir

$$\begin{aligned} \liminf_{n \rightarrow \infty} \sup_{t \in K} [f_n(t) + a_n h(t)] &\geq \liminf_{n \rightarrow \infty} \sup_{t \in K} [f(t) + a_n h(t)] - \varepsilon \\ &\geq \liminf_{n \rightarrow \infty} \sup_{t \in A} [f(t) + a_n h(t)] - \varepsilon \\ &= \sup_{t \in A} f(t) - \varepsilon, \end{aligned}$$

und somit

$$\liminf_{n \rightarrow \infty} \sup_{t \in K} [f_n(t) + a_n h(t)] \geq \sup_{t \in A} f(t).$$

Zum Beweis der Gleichheit definieren wir zunächst für ein $t \in K$ und Teilmengen $\emptyset \neq B \subset K$ den Abstand $\text{dist}(t, B) := \inf\{\rho(t, s); s \in B\}$. Die Abbildung $t \mapsto \text{dist}(t, B)$ ist gleichmäßig stetig auf K und daher sind die Mengen

$$\begin{aligned} A_\delta &:= \{t \in K; \text{dist}(t, A) \leq \delta\} \\ B_\delta &:= \{t \in K; \text{dist}(t, A) \geq \delta\} \end{aligned}$$

für alle $\delta > 0$ abgeschlossene und somit kompakte Teilmengen von K . Stetige Funktionen nehmen auf kompakten Mengen ihr Maximum an. Wir erhalten deswegen für jedes $\delta > 0$, hinreichend große $n \in \mathbb{N}$ und mit $\sup_\emptyset := -\infty$

$$\begin{aligned} \sup_{t \in K} [f_n(t) + a_n h(t)] &\leq \max \left(\sup_{t \in A_\delta} [f(t) + a_n h(t)], \sup_{t \in B_\delta} [f(t) + a_n h(t)] \right) + \varepsilon \\ &\leq \max \left(\sup_{t \in A_\delta} f(t), \sup_{t \in B_\delta} [f(t) + a_n h(t)] \right) + \varepsilon \\ &\leq \max \left(\sup_{t \in A_\delta} f(t), \sup_{t \in B_\delta} f(t) + a_n \sup_{t \in B_\delta} h(t) \right) + \varepsilon, \end{aligned}$$

wobei $\sup_{t \in B_\delta} h(t) < 0$ ist. Es folgt zusammen mit der Stetigkeit von f auf K

$$\limsup_{n \rightarrow \infty} \sup_{t \in K} [f_n(t) + a_n h(t)] \leq \sup_{t \in A_\delta} f(t) + \varepsilon \searrow \sup_{t \in A} f(t) + \varepsilon, \quad \text{für } \delta \rightarrow 0.$$

Da $\varepsilon > 0$ beliebig war, folgt

$$\limsup_{n \rightarrow \infty} \sup_{t \in K} [f_n(t) + a_n h(t)] \leq \sup_{t \in A} f(t)$$

und somit die Behauptung. \square

D.3.3. Lemma. Sei $(B, \|\cdot\|)$ ein separabler Banachraum. Sei $(B^*, \|\cdot\|^*)$ der zugehörige Dualraum, versehen mit der üblichen Operatornorm. Dann existiert eine abzählbare Teilmenge $B_0^* \subset \{f \in B^* : \|f\|^* \leq 1\} := K$, so dass für alle $x \in B$ gilt

$$\|x\| = \sup_{f \in B_0^*} f(x).$$

Beweis. Der Satz von Hahn-Banach ([Rud91], 3.3) liefert für jedes $x \in B$ die Darstellung $\|x\| = \sup_{f \in K} f(x)$. Für ein festes $x \in B$ sei $M_x^* \subset K$ eine abzählbare Teilmenge mit der Eigenschaft, dass

$$\|x\| = \sup_{f \in M_x^*} f(x)$$

gilt. Sei $B_0 \subset B$ eine abzählbare dichte Teilmenge. Dann ist $B_0^* := \bigcup_{x \in B_0} M_x^*$ ebenfalls abzählbar mit der Eigenschaft

$$\forall x \in B_0 : \quad \|x\| = \sup_{f \in B_0^*} f(x).$$

Sei $x \in B$ beliebig. Offenbar ist $\sup_{f \in K} f(x) \geq \sup_{f \in B_0^*} f(x)$. Andererseits existiert für jedes $\varepsilon > 0$ ein $x_0 \in B_0$ mit $\|x - x_0\| < \frac{\varepsilon}{2}$. Es folgt

$$\begin{aligned} \|x\| &\leq \|x_0\| + \frac{\varepsilon}{2} = \sup_{f \in B_0^*} f(x_0) + \frac{\varepsilon}{2} \\ &\leq \sup_{f \in B_0^*} \left(f(x) + \frac{\varepsilon}{2} \right) + \frac{\varepsilon}{2} \\ &= \sup_{f \in B_0^*} f(x) + \varepsilon, \end{aligned}$$

wobei die letzte Abschätzung daraus folgt, dass jedes $f \in B_0^*$ ein stetiges und lineares Funktional auf B mit $|f(x)| \leq \|x\|$ ist. Da $\varepsilon > 0$ beliebig war, erhalten wir somit für jedes $x \in B$

$$\sup_{f \in B_0^*} f(x) = \sup_{f \in K} f(x).$$

\square

D.3.4. Satz. *Es sei (\mathcal{F}, ρ) ein totalbeschränkter metrischer Raum, $C^u(\mathcal{F}) := \{f : \mathcal{F} \rightarrow \mathbb{R} \mid f \text{ gleichmäßig stetig}\}$. Dann ist $(C^u(\mathcal{F}), \|\cdot\|_\infty)$ ein separabler Banachraum.*

Beweis. Es sei $\overline{\mathcal{F}}$ die bis auf Isometrie eindeutig bestimmte Vervollständigung des metrischen Raums \mathcal{F} . Dann ist $\overline{\mathcal{F}}$ nach Konstruktion vollständig und \mathcal{F} eine total beschränkte, dichte Teilmenge in $\overline{\mathcal{F}}$, also ist insbesondere $\overline{\mathcal{F}}$ selbst total beschränkt und somit kompakt. Sei nun $f \in C^u(\mathcal{F})$ beliebig. Nach [Die71], 3.15.6, existiert für jedes $f \in C^u(\mathcal{F})$ eine eindeutig bestimmte, (gleichmäßig) stetige Fortsetzung von f auf $\overline{\mathcal{F}}$, also ein $f^* \in C(\overline{\mathcal{F}})$ mit $f^*|_{\mathcal{F}} = f$. Es ist $(C(\overline{\mathcal{F}}), \|\cdot\|_\infty)$ separabel und vollständig. Ist dann $A \subset C(\overline{\mathcal{F}})$ eine abzählbar dichte Teilmenge, so ist $A|_{\mathcal{F}} := \{f|_{\mathcal{F}} : f \in A\}$ eine abzählbar dichte Teilmenge in $C^u(\mathcal{F})$ und somit $C^u(\mathcal{F})$ ebenfalls separabel.

Bleibt die Vollständigkeit von $C^u(\mathcal{F})$ zu zeigen: Sei dazu $\{f_n\}_{n \in \mathbb{N}} \subset C^u(\mathcal{F})$ eine Cauchyfolge. Dann gilt

$$\forall \varepsilon > 0 \exists n(\varepsilon) \in \mathbb{N} \forall k, l \geq n(\varepsilon) : \|f_l - f_k\|_\infty \leq \varepsilon.$$

Wir betrachten die Folge der Funktionen $\{f_n^*\}_{n \in \mathbb{N}} \subset C^u(\overline{\mathcal{F}})$, wobei für jedes $n \in \mathbb{N}$ f_n^* die eindeutig bestimmte Fortsetzung von f_n auf $\overline{\mathcal{F}}$ sei. Da jedes f_n^* gleichmäßig stetig und $\mathcal{F} \subset \overline{\mathcal{F}}$ dicht ist, gilt für alle $k, l \in \mathbb{N}$

$$\sup_{x \in \overline{\mathcal{F}}} |f_k^*(x) - f_l^*(x)| = \sup_{x \in \mathcal{F}} |f_k^*(x) - f_l^*(x)| = \sup_{x \in \mathcal{F}} |f_k(x) - f_l(x)|,$$

und daher, dass die Folge $\{f_n^*\}_{n \in \mathbb{N}} \subset C^u(\overline{\mathcal{F}})$ eine Cauchyfolge ist. Aufgrund der Vollständigkeit von $(C(\overline{\mathcal{F}}), \|\cdot\|_\infty)$ existiert ein $f^* \in C(\overline{\mathcal{F}})$ mit $\|f_n^* - f^*\|_\infty \rightarrow 0, n \rightarrow \infty$. Es folgt $\|f_n^*|_{\mathcal{F}} - f^*|_{\mathcal{F}}\|_\infty = \|f_n - f^*|_{\mathcal{F}}\|_\infty \rightarrow 0$. Folglich konvergiert die Cauchyfolge $\{f_n\}_{n \in \mathbb{N}}$ gleichmäßig gegen die Restriktion $f^*|_{\mathcal{F}}$. Es ist f^* gleichmäßig stetig und somit auch $f^*|_{\mathcal{F}}$ gleichmäßig stetig auf \mathcal{F} □

Literaturverzeichnis

- [Adl90] R. J. Adler. *An Introduction to Continuity, Extrema, and Related Topics for General Gaussian Processes*. Lecture Notes-Monograph Series. 1990.
- [AG75] R. B. Ash and M. F. Gardner. *Topics in stochastic processes*. 1975.
- [AKJ91] A. Alzaid, Kim, and F. J.S., Proschan. Laplace orderings and its applications. *Journal of Applied Probability*, 28:116–130, 1991.
- [Bah60a] R.R. Bahadur. On the asymptotic efficiency of tests and estimates. *Sankhya*, 22:229–252, 1960.
- [Bah60b] R.R. Bahadur. Stochastic comparison of tests. *Ann. Math. Statist.*, 31:276–295, 1960.
- [Bah67] R.R. Bahadur. Rates of convergence of estimates and test statistics. *Ann. Math. Statist.*, 38:303–324, 1967.
- [Ben07] C. J. Bennett. Consistent tests for completely monotone stochastic dominance. unpublished, 2007.
- [Ber91] R. Bergmann. Stochastic orders and their application to a unified approach to various concepts of dependence and association. In K. Mosler and M. Scarsini (Eds.). *Stochastic Orders and Decision under Risk*, IMS Lecture Notes in Mathematics 19, 1991.
- [BG09] L. Baringhaus and R. Grübel. Nonparametric two-sample tests for increasing convex order. *Bernoulli*, 15:99–123, 2009.
- [Bil68] P. Billingsley. *Convergence of Probability Measures*. 1968.
- [BM86] R. Beran and P. W. Millar. Confidence sets for a multivariate distribution. *Ann. Statist.*, 14:431–442, 1986.

- [Bäu97] N. Bäuerle. Inequalities for stochastic models via supermodular orderings. *Comment. Math. Univ. Carolin.*, 32:181–201, 1997.
- [CDS04] A. C. Cebrián, M. Denuit, and O. Scaillet. Testing for concordance ordering. *Astin Bull.*, 34:151–173, 2004.
- [Den01] M. Denuit. Laplace transform order of actuarial quantities. *Insurance: Mathematics and Economics*, 29:83–102, 2001.
- [Den05] M. Denuit. *Actuarial Theory for dependent risks*. 2005.
- [DG96] J. Dhaene and M.J. Goovaerts. Dependency of risks and stop-loss order. *ASTIN Bulletin*, 26:201–212, 1996.
- [Die71] J. Dieudonné. *Grundzüge der modernen Analysis*. 1971.
- [EMS02] P. Embrechts, A. McNeil, and D. Straumann. *Risk Management: Value at Risk and beyond*. Cambridge: Cambridge University Press, 2002.
- [FA90] K.-T. Fang and T.-W. Anderson. *Statistical Inference in Elliptically Contoured and Related Distributions*. 1990.
- [FC84] K.T. Fang and H.F. Chen. Relationships among classes of spherical matrix distributions. *Acta Math. Appl. Sinica*, 1984.
- [FKN87] K.-T. Fang, S. Kotz, and K.-W. Ng. *Symmetric Multivariate and Related Distributions*. Chapman and Hall, 1987.
- [GR65] R. C. Gunning and H. Rossi. *Analytic Functions of Several Complex Variables*. 1965.
- [Joe90] H. Joe. Multivariate concordance. *Journal of Multivariate Analysis*, 35:12–30, 1990.
- [KBJ00] S. Kotz, N. Balakrishnan, and N. L. Johnson. *Continuous Multivariate Distributions*. 2000.
- [Kel70] D. Kelker. Distribution theory of spherical distributions and location-scale parameter generalisations. *Sankhya*, 32:419–430, 1970.
- [KK95] W.C.M. Kallenberg and A.J. Koning. On Wieand’s theorem. *Statist. Probab. Lett.*, 25:121–132, 1995.

-
- [Kle83] B. Klefsjö. A useful ageing property based on the Laplace transform. *Journal of Applied Probability*, 20:615–626, 1983.
- [KS87] G. Kimeldorf and A. R. Sampson. A framework for positive dependence. *Ann. Inst. Statist. Math.*, 41:31–45, 1987.
- [Kue70] J. Kuelbs. Gaussian measure on a Banach space. *J. Func. Anal.*, 5:354–367, 1970.
- [Kuo70] H.-H. Kuo. *Gaussian Measures in Banach spaces*. Lecture Notes in Mathematics. 1970.
- [Leh86] E.L. Lehmann. *Testing statistical hypotheses*. Second edition, 1986.
- [MS02] A. Müller and D. Stoyan. *Comparison Methods for Stochastic Models and Risks*. 2002.
- [Mül97] A. Müller. Stop-loss order for portfolios of dependent risks. *Insurance Math. Econom.*, 21:219–223, 1997.
- [PS70] G. Pólya and G. Szegő. *Problems and Theorems in Analysis II*. 1970.
- [Rud91] W. Rudin. *Functional Analysis*. 1991.
- [Sat69] Hiroshi Sato. Gaussian measures on a Banach space and abstract Wiener measure. *Nagoya Math. J.*, 36:65–81, 1969.
- [Sca00] M. Scarsini. Some remarks on the supermodular order. *Journal of Multivariate Analysis*, 73:107–119, 2000.
- [SS07] M. Shaked and J. G. Shanthikumar. *Stochastic Orders*. Springer Verlag, 2007.
- [Sto83] D. Stoyan. *Comparison Models For Queues and Other Stochastic Models*. 1983.
- [Tah09] F. Taheri. *Empirical Hankel transform and statistical goodness-of-fit tests for exponential distributions*. PhD thesis, 2009.
- [Tay83] J. M. Taylor. Comparison of certain distribution functions. *Operationsforsch. Statist. Ser. Statist.*, 14:397–408, 1983.
- [Ton80] Y. L. Tong. *Probability inequalities in multivariate distributions*. New York: Academic Press, 1980.

- [Tsi75] V. S. Tsirelson. The density of the distribution of the maximum of a Gaussian process. *Theory. Probab. Appl.*, 20:847–856, 1975.
- [Vak79] N. N. Vakhania. Correspondence between Gaussian measures and the Gaussian process. *Mat. Zametki*, 26:293–297, 1979.
- [VW96] van der Vaart, A. W. and J. A. Wellner. *Weak Convergence and Empirical Processes*. 1996.
- [WMF95] H. Witting and U. Müller-Funk. *Mathematische Statistik II*. 1995.
- [YO69] T. Yanagimoto and M. Okamoto. Partial orderings of permutations and monotonicity of a rank correlation statistic. *Ann. Inst. Statist. Math.*, 21:489–506, 1969.

Lebenslauf

Persönliche Informationen

Name Dennis Behnke
Geburtsdatum 10. Mai 1984
Geburtsort Burgwedel
Familienstand ledig

Schulischer Werdegang

1990-1994 **Grundschule**, *Grundschule Elze*, Elze-Bennemühlen.
1994-1996 **Orientierungsstufe**, *OS Mellendorf*, Mellendorf.
1996-2003 **Gymnasium**, *Gymnasium Mellendorf*, Mellendorf.
Abschluss: Abitur im Juli 2003

Studium

2003-2008 **Studium der Mathematik**, *Leibniz Universität Hannover*, Hannover.
Abschluss: Dipl.-Math. am 02. September 2008

Promotionsstudium

seit 01.10.2008 **Promotionsstudent**, *Leibniz Universität Hannover*, Hannover.
seit 01.01.2009 **Wissenschaftlicher Mitarbeiter**, *Institut für Mathematische Stochastik*, Leibniz Universität Hannover.

Voßstraße 21 – 30161 Hannover

📞 0151 41930428 • 📠 0511 48893177
• ✉ dbehnke@stochastik.uni-hannover.de