

Technische Universität München

ZENTRUM MATHEMATIK

**Lokale Alternativen
in Tests auf Überdispersion
bei zusammengesetzten
Poisson-Verteilungen**

Diplomarbeit

von

Bernhard Niesert

Themensteller: Prof. Dr. Claudia Klüppelberg

Betreuer: Dr. Boris Buchmann

Abgabetermin: 24. Juli 2002

Hiermit erkläre ich, daß ich die Diplomarbeit selbständig angefertigt und nur die angegebenen Quellen verwendet habe.

München, 24. Juli 2002

Inhaltsverzeichnis

1	Einleitung	2
2	Vergleich des Tests mit dem bekannten Dispersionstest	4
2.1	Dispersionstest	4
2.1.1	Verteilungsfunktion des Dispersionstests unter der Hypothese	7
2.1.2	Verteilungsfunktion des Dispersionstests unter der lokalen Alternative	8
2.2	Test mit Hilfe von Kegelapproximationen	16
2.2.1	Verteilungsfunktion des Tests unter der Hypothese	16
2.2.2	Verteilungsfunktion des Tests unter der lokalen Alternative	19
2.3	Vergleich der Tests mittels lokaler Alternativen	22
2.3.1	Berechnung von Maximum und Minimum des Dispersionstests	23
2.3.2	Simulation der Verteilung des Tests mittels Kegelapproximation unter der lokalen Alternative	37
2.3.3	Vergleich der beiden Tests	41
3	Test des Algorithmus zur Bestimmung der Kappa-Werte	46
3.1	Vergleich der Simulation gegen die exakten Werte	46
3.1.1	Berechnung der Kappa-Werte für $S=2$	46
3.1.2	Berechnung der Kappa-Werte für $S=3$	47
3.1.3	Berechnung der Kappa-Werte für $S=4$	52
3.2	Simulationsergebnisse für $S > 4$	65
4	Praktische Implementierung des Tests	70
4.1	Programmierung in <i>matlab</i>	70
4.2	Testen des <i>matlab</i> -Programms	88
A	Algorithmus zur Simulation der lokalen Alternative	90
B	<i>matlab</i> Dateien	102
C	MAPLE Datei zur Bestimmung von κ_4 im Fall $S = 4$	115

Kapitel 1

Einleitung

Einfache und zusammengesetzte Poissonverteilungen spielen heutzutage gerade im Versicherungsbereich eine große Rolle. So dient die diskrete Poissonverteilung in vielen Fällen zur Simulation der Anzahl der Schadenfälle. Betrachtet man nun noch die Verteilung der Schadenhöhe und geht weiter davon aus, daß diese unabhängig von der Schadenanzahl ist, so erhält man für den Gesamtschaden eine zusammengesetzte Poissonverteilung. Dies sei an dieser Stelle nur eine kurze Motivation, wofür gerade in der Praxis in den Versicherungen die einfache und die zusammengesetzte Poissonverteilung benutzt werden. Wer sich für eine detailliertere Beschreibung der Voraussetzungen und Probleme dieser Darstellungen interessiert, kann dies besonders gut in [Ma97] nachlesen.

In dieser Arbeit geht es jedoch vor allem darum, den Test nach Dr. Boris Buchmann aus seiner Doktorarbeit (vgl. [Bu01]) genauer zu betrachten. Dieser Test ist als Alternative zum Dispersionstest gedacht und stellt einen Test dar, der ein Kriterium zur Unterscheidung zwischen einfacher und zusammengesetzter Poissonverteilung liefert. Der Test beruht dabei auf einem Maximum-Likelihood-Quotiententest und verwendet zur Berechnung Kegelprojektionen. Die Herleitung des Tests ist ausführlich in dem zuvor genannten Werk geschildert. Meine Arbeit beruht in weiten Teilen auf den Ausführungen in [Bu01] und verwendet gelegentlich Ergebnisse daraus, ohne dies explizit zu erwähnen. Allerdings werden in dieser Arbeit alle Schreibweisen nochmals eingeführt, so daß der Leser dieser Arbeit auch ohne Kenntnis von [Bu01] folgen kann.

Im zweiten Kapitel geht es als erstes darum, die neue Teststatistik mit dem bekannten Dispersionstest zu vergleichen. Dazu wird der Weg mittels lokaler Alternativen gewählt. Jedoch muß der Ansatz aus [Va00] dazu in eine allgemeinere Form gebracht werden. Die Ergebnisse sind am Ende des zweiten Kapitels zusammengefaßt, wobei jedoch keine allgemeine Aussage getroffen werden kann. Es werden allerdings Tendenzen festgestellt wie z.B., daß der neue Test zum Teil deutlich besser ist, wenn sich die Masse der Basisverteilung auf den ersten Komponenten konzentriert.

Als nächstes wird in Kapitel 3 der Algorithmus zur Bestimmung der κ -Werte aus [Bu01] betrachtet. Für kleine Werte der Abschneidegrenze S werden die mittels dieses Algorithmus bestimmten κ -Werte mit den berechneten Werten verglichen, für große Werte von S wird der Algorithmus lediglich noch auf seine Stabilität getestet. Dabei treten die ersten Instabilitäten erst ab $\lambda \approx 10$ und $S = 7$ auf, also in Bereichen, die für den Test eigentlich nicht mehr von Bedeutung sind.

Das abschließende vierte Kapitel enthält die praktische Implementierung des Tests. Mittels *matlab* wird hierbei eine Routine zur Verfügung gestellt, die entweder feststellt, ob zu einem

gewissen α -Niveau die Hypothese (es handelt sich um eine einfache Poissonverteilung) verworfen werden kann oder den p-Wert des Tests liefert. Anschließend wird diese Routine noch für verschiedene Daten getestet, wie einfach oder zusammengesetzt poissonverteilte Zufallsvariablen (mittels Simulationen erzeugt) oder für die Daten des klassischen Beispiels von Bortkiewicz (Tote in der preussischen Armee, vgl. [Qu87]). In allen Fällen lieferte der Test gute Ergebnisse.

Kapitel 2

Vergleich des Tests mit dem bekannten Dispersionstest

Zu Beginn der Arbeit gilt es, den Test aus [Bu01] auf seine Effizienz im Vergleich zum bisher verwendeten Dispersionstest zu prüfen. Dazu soll der Weg der lokalen Alternative verwendet werden (vgl. auch [Va00]).

Betrachtet man einen Test mit Hypothese $H : \theta \in \Theta_0$ und Alternative $K : \theta \in \Theta_1$, so ist ein Test T_1 (mit Gütefunktion π_n) besser als ein weiterer Test T_2 (mit Gütefunktion $\bar{\pi}_n$) genau dann, wenn gilt:

$$\begin{aligned} \pi_n(\theta) &\leq \bar{\pi}_n(\theta), & \theta \in \Theta_0 \\ \text{und } \pi_n(\theta) &\geq \bar{\pi}_n(\theta), & \theta \in \Theta_1 \end{aligned}$$

Möchte man nun die beiden Tests für $n \rightarrow \infty$ vergleichen, so liegt folgender Ansatz nahe:

Bilde $\pi(\theta) = \lim_{n \rightarrow \infty} \pi_n(\theta)$ für beide Tests und vergleiche die so erhaltenen Grenzwerte für alle zulässigen Werte von θ . Es stellt sich jedoch heraus, daß dieser Ansatz nicht funktioniert. Zwar existiert die Grenzfunktion noch für die meisten Teststatistiken, allerdings hat sie bei konsistenten Tests nach Definition folgende Gestalt:

$$\pi(\theta) = \begin{cases} 0, & \text{falls } \theta \in \Theta_0 \\ 1, & \text{falls } \theta \in \Theta_1 \end{cases}$$

Dies ist auch einleuchtend, da die Teststatistiken für $n \rightarrow \infty$ einen scharf trennenden Test darstellen sollen.

Bei der Methode mit den lokalen Alternativen hingegen nähert man sich der Hypothese aus Richtung der Alternative mit der Rate $O(1/\sqrt{n})$. Auf diesem Weg erhält man für $n \rightarrow \infty$ eine nicht triviale Gütefunktion.

Betrachten wir zuerst die Verteilungsfunktionen der beiden Tests unter der Hypothese und der lokalen Alternative, bevor im Kapitel 2.3 dann auf das weitere Vorgehen eingegangen wird.

2.1 Dispersionstest

Um die Verteilungsfunktionen des Dispersionstests unter der Hypothese und unter der lokalen Alternative zu bestimmen, benötigt man die Waldschen Gleichungen. Da diese meist nur bis

zur Varianz der zusammengesetzten Poissonverteilung berechnet werden, sollen in Satz 2.2 die Gleichungen bis zum vierten zentralen Moment berechnet werden. Dies wird dann zur Bestimmung der Verteilung bei den lokalen Alternativen benötigt. Im folgenden Lemma sei $(X_i)_{i \in \mathbb{N}}$ eine Folge von unabhängigen und identisch verteilten Zufallsvariablen und τ eine von $(X_i)_{i \in \mathbb{N}}$ unabhängige mit Parameter $\lambda > 0$ poissonverteilte Zufallsvariable. $Y := \sum_{k=1}^{\tau} X_k$ ist dann eine zusammengesetzt Poisson-verteilte Zufallsvariable.

Lemma 2.1 (Momentenerzeugende Funktion von Y) *Die momentenerzeugende Funktion sei wie folgt definiert:*

$$\Psi(\theta) = E[e^{\theta Y}]$$

Dann besitzt die zusammengesetzt Poisson-verteilte Zufallsvariable Y die folgende momentenerzeugende Funktion und die dazugehörigen Ableitungen:

$$\Psi(\theta) = e^{\lambda(\Psi_X(\theta)-1)} \quad (2.1)$$

$$\Psi'(\theta) = \Psi(\theta)\lambda\Psi'_X(\theta) \quad (2.2)$$

$$\Psi''(\theta) = \Psi'(\theta)\lambda\Psi'_X(\theta) + \Psi(\theta)\lambda\Psi''_X(\theta) \quad (2.3)$$

$$\Psi'''(\theta) = \Psi''(\theta)\lambda\Psi'_X(\theta) + 2\Psi'(\theta)\lambda\Psi''_X(\theta) + \Psi(\theta)\lambda\Psi'''_X(\theta) \quad (2.4)$$

$$\Psi^{(4)}(\theta) = \Psi'''(\theta)\lambda\Psi'_X(\theta) + 3\Psi''(\theta)\lambda\Psi''_X(\theta) + 3\Psi'(\theta)\lambda\Psi'''_X(\theta) + \Psi(\theta)\lambda\Psi^{(4)}_X(\theta) \quad (2.5)$$

Dabei sei angenommen, daß die entsprechenden Ausdrücke für die Zufallsvariable X existieren.

Beweis:

Setzt man die Definition der Zufallsvariablen Y ein, so ergibt sich für die momentenerzeugende Funktion:

$$\begin{aligned} \Psi(\theta) &= E[e^{\theta Y}] = E[e^{\theta \sum_{k=1}^{\tau} X_k}] = E\left[\sum_{m=0}^{\infty} 1_{\tau=m} \cdot e^{\theta \sum_{i=1}^m X_i}\right] = \\ &\quad \text{(Satz über die majorisierte Konvergenz)} \\ &= \sum_{m=0}^{\infty} E[1_{\tau=m} \cdot e^{\theta \sum_{i=1}^m X_i}] = \\ &\quad \text{(Unabhängigkeit von } \tau \text{ und } (X_i)_{i \in \mathbb{N}}) \\ &= \sum_{m=0}^{\infty} \left(\underbrace{E[1_{\tau=m}]}_{=P(\tau=m)} \cdot \underbrace{E[e^{\theta \sum_{i=1}^m X_i}]}_{=\prod_{i=1}^m E[e^{\theta X_i}]} \right) = \\ &\quad \text{(Unabhängigkeit der } (X_i)) \\ &= \sum_{m=0}^{\infty} \left(e^{-\lambda} \frac{\lambda^m}{m!} \prod_{l=1}^m E[e^{\theta X_l}] \right) = \\ &= e^{-\lambda} \sum_{m=0}^{\infty} \left(\frac{\lambda^m}{m!} (E[e^{\theta X_1}])^m \right) = \\ &= e^{-\lambda} e^{\lambda E[e^{\theta X_1}]} = \\ &= e^{\lambda(\Psi_X(\theta)-1)} \end{aligned}$$

Die Gleichung für $\Psi'(\theta)$ ergibt sich aufgrund der Kettenregel:

$$\Psi'(\theta) = (e^{\lambda(\Psi_X(\theta)-1)})' = \underbrace{e^{\lambda(\Psi_X(\theta)-1)}}_{=\Psi(\theta)} \cdot (\lambda(\Psi_X(\theta) - 1))' = \Psi(\theta)\lambda\Psi'_X(\theta)$$

Die weiteren Gleichungen ergeben sich aus wiederholter Anwendung der Regeln der Differentiation. □

Um nun die Waldschen Gleichungen herzuleiten, benötigt man noch folgenden Zusammenhang zwischen der momentenerzeugenden Funktion und den zentralen Momenten:

$$E[Y^k] = \Psi^{(k)}(0)$$

Damit läßt sich dann der nachfolgende Satz beweisen:

Satz 2.2 (Waldsche Gleichungen)

$$E[Y] = \lambda E[X] \tag{2.6}$$

$$E[Y^2] = \lambda^2 E[X]^2 + \lambda E[X^2] \tag{2.7}$$

$$E[Y^3] = \lambda^3 E[X]^3 + 3\lambda^2 E[X^2]E[X] + \lambda E[X^3] \tag{2.8}$$

$$E[Y^4] = \lambda^4 E[X]^4 + 6\lambda^3 E[X^2]E[X]^2 + 4\lambda^2 E[X^3]E[X] + 3\lambda^2 E[X^2]^2 + \lambda E[X^4] \tag{2.9}$$

Aus (2.6) und (2.7) folgt:

$$\text{Var}(Y) = \lambda E[X^2] \tag{2.10}$$

Beweis:

Zuerst betrachtet man $\Psi(0)$. Aufgrund der Definition und auch mittels der Gleichung $\Psi(0) = E[Y^{(0)}]$ sieht man, daß $\Psi(0) = 1$ gelten muß. Somit ist die Gleichung (2.6) bewiesen. Für die Gleichung (2.7) gilt:

$$E[Y^2] = \Psi''(0) = \Psi'(0)\lambda\Psi'_X(0) + \Psi(0)\lambda\Psi''_X(0) = \lambda E[X]\lambda E[X] + 1\lambda E[X^2]$$

und auch diese Gleichung ist bewiesen. Die restlichen Gleichungen ergeben sich aus den nachfolgenden Rechnungen.

$$\begin{aligned} \Psi'''(0) &= (\lambda^2 E[X]^2 + \lambda E[X^2])\lambda E[X] + 2\lambda^2 E[X]E[X^2] + \lambda E[X^3] = \\ &= \lambda^3 E[X]^3 + 3\lambda^2 E[X^2]E[X] + \lambda E[X^3] \\ \Psi^{(4)}(0) &= (\lambda^3 E[X]^3 + 3\lambda^2 E[X^2]E[X] + \lambda E[X^3])\lambda E[X] + \\ &\quad + 3(\lambda^2 E[X]^2 + \lambda E[X^2])\lambda E[X^2] + 3\lambda E[X]\lambda E[X^3] + \lambda E[X^4] = \\ &= \lambda^4 E[X]^4 + \lambda^3 E[X^2]E[X]^2(3+3) + \lambda^2(E[X^3]E[X] + 3E[X^2]^2 + 3E[X]E[X^3]) + \lambda E[X^4] = \\ &= \lambda^4 E[X]^4 + 6\lambda^3 E[X^2]E[X]^2 + \lambda^2(4E[X^3]E[X] + 3E[X^2]^2) + \lambda E[X^4] \end{aligned}$$

Bleibt noch Gleichung (2.10) zu beweisen. Diese ergibt sich jedoch einfach aus dem Verschiebungssatz:

$$\text{Var}(Y) = E[Y^2] - (E[Y])^2 = \lambda^2 E[X]^2 + \lambda E[X^2] - (\lambda E[X])^2 = \lambda E[X^2]$$

□

2.1.1 Verteilungsfunktion des Dispersionstests unter der Hypothese

Der Grundgedanke beim Dispersionstest ist, die Gleichheit von Varianz und Erwartungswert bei der Poissonverteilung auszunutzen. Dividiert man nun bei einer poissonverteilten Zufallsvariablen die Varianz durch den Erwartungswert, so ergibt sich als Ergebnis die von λ unabhängige Konstante 1. Daher bietet sich beim Dispersionstest folgende Teststatistik an:

$$\tilde{T}_n := \sqrt{n} \cdot \left(\frac{\hat{S}_n}{\bar{Y}_n} - 1 \right)$$

Um später eine standardisierte Verteilung zu erhalten, wählt man jedoch geschickter die leicht veränderte Teststatistik T_n , die wie folgt definiert ist:

$$T_n := \sqrt{\frac{n}{2}} \left(\frac{\hat{S}_n}{\bar{Y}_n} - 1 \right) = \frac{1}{\sqrt{2}} \tilde{T}_n$$

In beiden Fällen bezeichnet \hat{S} die Streuung und \bar{Y} den Mittelwert der statistischen Stichprobe. Um in den weiteren Rechnungen nicht immer die Division durch $\sqrt{2}$ mitschleppen zu müssen, wird zur Analyse die Teststatistik \tilde{T}_n betrachtet und erst am Schluß wieder durch die eigentlich interessante Teststatistik T_n ersetzt. Die Rechnungen für T_n verlaufen bis auf die Konstante analog.

Die Teststatistik \tilde{T}_n läßt sich (analog zum Verschiebungssatz bei der Varianz) umschreiben in $\frac{\hat{S}_n - (\bar{Y}_n)^2}{\bar{Y}_n}$. Dabei bezeichne $\overline{Y_n^2}$ den Mittelwert der Quadrate von Y_n . Somit gilt:

$$\frac{\hat{S}_n}{\bar{Y}_n} = \frac{\overline{Y_n^2} - (\bar{Y}_n)^2}{\bar{Y}_n} = \varphi(\bar{Y}_n, \overline{Y_n^2}) \quad \text{mit } \varphi(x, y) := \frac{y - x^2}{x} \quad \text{definiert auf } (0, \infty) \times (0, \infty)$$

Wir benötigen also die gemeinsame Verteilung von \bar{Y}_n und $\overline{Y_n^2}$, um die Verteilung von \tilde{T}_n bestimmen zu können. Dann liefert der zentrale Grenzwertsatz:

$$\sqrt{n} \left(\begin{array}{c} \bar{Y}_n - E[Y] \\ \overline{Y_n^2} - E[Y^2] \end{array} \right) = \frac{1}{\sqrt{n}} \sum_{k=1}^n \left(\begin{array}{c} Y_k - E[Y] \\ Y_k^2 - E[Y^2] \end{array} \right) \xrightarrow{w} Z$$

wobei

$$Z \sim N_2 \left(0, \underbrace{\begin{pmatrix} \text{Var} Y_k & \text{Cov}(Y_k, Y_k^2) \\ \text{Cov}(Y_k, Y_k^2) & \text{Var}(Y_k^2) \end{pmatrix}}_{:=\Sigma(\lambda)} \right)$$

Als nächstes soll nun die Matrix $\Sigma(\lambda)$ berechnet werden. Für die Varianzen einer einfachen Poissonverteilung gilt nach [Ha67]:

$$\begin{aligned} \text{Var}(Y_k) &= \lambda \\ \text{Var}(Y_k^2) &= E[Y_k^2 - E[Y_k^2]]^2 = E[Y_k^4] - E[Y_k^2]^2 = \\ &= \lambda + 7\lambda^2 + 6\lambda^3 + \lambda^4 - (\lambda + \lambda^2)^2 = \lambda + 6\lambda^2 + 4\lambda^3 \end{aligned}$$

Für die Kovarianz gilt:

$$\text{Cov}(Y_k, Y_k^2) = E[Y_k \cdot Y_k^2] - E[Y_k]E[Y_k^2] = (\lambda + 3\lambda^2 + \lambda^3) - \lambda(\lambda + \lambda^2) = \lambda + 2\lambda^2$$

Zusammen ergibt sich folgender Ausdruck für die Kovarianzmatrix

$$\Sigma(\lambda) = \begin{pmatrix} \lambda & \lambda(1+2\lambda) \\ \lambda(1+2\lambda) & \lambda(1+6\lambda+4\lambda^2) \end{pmatrix} = \lambda \cdot \begin{pmatrix} 1 & 1+2\lambda \\ 1+2\lambda & 1+6\lambda+4\lambda^2 \end{pmatrix}$$

Außerdem wissen wir, daß $\frac{\hat{S}_n}{\bar{Y}_n} \xrightarrow{f.s.} \frac{E[Y^2]-E[Y]^2}{E[Y]}$, da $\bar{Y}_n^2 \xrightarrow{f.s.} E[Y^2]$, $\bar{Y}_n \xrightarrow{f.s.} E[Y] = \lambda > 0$ (falls die Momente existieren) und φ stetig in $(\lambda, \lambda + \lambda^2)$ ist.

Wendet man nun die Delta-Methode auf den Dispersionstest an, so ergibt sich (vergleiche hierzu auch [Va00]):

$$\sqrt{n} \left(\varphi(\bar{Y}_n, \bar{Y}_n^2) - \underbrace{\varphi(\lambda, \lambda + \lambda^2)}_{=1} \right) \xrightarrow{\mathcal{W}} D\varphi_{\lambda, \lambda + \lambda^2}(Z)$$

Bleibt also noch $D\varphi_{\lambda, \lambda + \lambda^2}$ zu bestimmen:

$$D\varphi_{\lambda, \lambda + \lambda^2}(x, y) = \left(\frac{-2x^2 - (y - x^2)}{x^2}, \frac{1}{x} \right)_{\lambda, \lambda + \lambda^2} = \left(\frac{-\lambda - \lambda^2 - \lambda^2}{\lambda^2}, \frac{1}{\lambda} \right) = \left(-\frac{2\lambda + 1}{\lambda}, \frac{1}{\lambda} \right)$$

Somit erhält man:

$$\tilde{T}_n \xrightarrow{\mathcal{W}} \tilde{T} = D\varphi_{\lambda, \lambda + \lambda^2}(Z) \sim N_1(0, \text{Var}(\tilde{T}))$$

Die Varianz läßt sich dabei wie folgt berechnen:

$$\begin{aligned} \text{Var}(\tilde{T}) &= E[\tilde{T} \cdot \tilde{T}^T] - \underbrace{E[\tilde{T}]E[\tilde{T}]^T}_{=0} = E[(D\varphi_{\lambda, \lambda + \lambda^2}(Z))(D\varphi_{\lambda, \lambda + \lambda^2}(Z))^T] = \\ &= D\varphi_{\lambda, \lambda + \lambda^2} \underbrace{E[ZZ^T]}_{\Sigma(\lambda)} D\varphi_{\lambda, \lambda + \lambda^2}^T = \\ &= \frac{1}{\lambda} \cdot \begin{pmatrix} -(2\lambda + 1) & 1 \end{pmatrix} \cdot \lambda \cdot \begin{pmatrix} 1 & 1+2\lambda \\ 1+2\lambda & 1+6\lambda+4\lambda^2 \end{pmatrix} \cdot \frac{1}{\lambda} \cdot \begin{pmatrix} -(2\lambda + 1) \\ 1 \end{pmatrix} = \\ &= \frac{1}{\lambda} \cdot \begin{pmatrix} \underbrace{-(2\lambda + 1) + (2\lambda + 1)}_{=0} & \\ & -(2\lambda + 1)^2 + 1 + 6\lambda + 4\lambda^2 \end{pmatrix} \begin{pmatrix} -(2\lambda + 1) \\ 1 \end{pmatrix} = \\ &= \frac{1}{\lambda} (-1 - 4\lambda - 4\lambda^2 + 1 - 6\lambda + 4\lambda^2) = 2 \end{aligned}$$

Somit ergibt sich für \tilde{T}_n folgende Aussage:

$$\tilde{T}_n \xrightarrow{\mathcal{W}} N(0, 2)$$

Für die Teststatistik T_n gilt aufgrund der Definition $T_n = \frac{1}{\sqrt{2}}\tilde{T}_n$:

$$T_n \xrightarrow{\mathcal{W}} N(0, 1)$$

2.1.2 Verteilungsfunktion des Dispersionstests unter der lokalen Alternative

Jetzt kommt der eigentlich interessantere Teil. Bei den lokalen Alternativen nähern wir uns der Hypothese aus Richtung der Alternative mit der Geschwindigkeit $\frac{1}{\sqrt{n}}$ und beobachten die Güte des Tests. Zur Berechnung der Verteilung der lokalen Alternative benutzen wir den Ansatz nach

Für die Berechnung von $\text{Var}(Y_{n,\cdot}^2)$ benutzen wir den Verschiebungssatz für Varianzen:

$$\text{Var}(Y_{n,\cdot}^2) = E\left[Y_{n,\cdot}^2 - E[Y_{n,\cdot}^2]\right]^2 = E[Y_{n,\cdot}^4] - E[Y_{n,\cdot}^2]^2$$

Es gilt weiter:

$$\begin{aligned} E[Y_{n,\cdot}^2] &= \text{Var}(Y_{n,\cdot}) + E[Y_{n,\cdot}]^2 \stackrel{(2.12)}{=} \lambda + \frac{1}{\sqrt{n}}(h_0 + \lambda \sum_k h_k k^2) + o\left(\frac{1}{\sqrt{n}}\right) + \\ &\quad + \left(\lambda + \frac{1}{\sqrt{n}}h_0 + o\left(\frac{1}{\sqrt{n}}\right)\right)^2 \left(1 + \frac{1}{\sqrt{n}} \sum_k h_k k + o\left(\frac{1}{\sqrt{n}}\right)\right)^2 = \\ &= \lambda + \lambda^2 + \frac{1}{\sqrt{n}}(h_0 + \lambda \sum_k h_k k^2 + 2\lambda^2 \sum_k h_k k + 2\lambda h_0) + o\left(\frac{1}{\sqrt{n}}\right) \\ E[Y_{n,\cdot}^4] &= \lambda_n^4 E[X_{n,\cdot}^4] + 6\lambda_n^3 E[X_{n,\cdot}^2] E[X_{n,\cdot}]^2 + 4\lambda_n^2 E[X_{n,\cdot}^3] E[X_{n,\cdot}] + 3\lambda_n^2 E[X_{n,\cdot}^2]^2 + \lambda_n E[X_{n,\cdot}^4] = \\ &= \left(\lambda + \frac{1}{\sqrt{n}}h_0 + o\left(\frac{1}{\sqrt{n}}\right)\right)^4 \left(1 + \frac{1}{\sqrt{n}} \sum_k h_k k^4 + o\left(\frac{1}{\sqrt{n}}\right)\right) + \\ &\quad + 6 \left(\lambda + \frac{1}{\sqrt{n}}h_0 + o\left(\frac{1}{\sqrt{n}}\right)\right)^3 \left(1 + \frac{1}{\sqrt{n}} \sum_k h_k k^2 + o\left(\frac{1}{\sqrt{n}}\right)\right) \cdot \\ &\quad \cdot \left(1 + \frac{1}{\sqrt{n}} \sum_k h_k k + o\left(\frac{1}{\sqrt{n}}\right)\right)^2 + \\ &\quad + 4 \left(\lambda + \frac{1}{\sqrt{n}}h_0 + o\left(\frac{1}{\sqrt{n}}\right)\right)^2 \left(1 + \frac{1}{\sqrt{n}} \sum_k h_k k^3 + o\left(\frac{1}{\sqrt{n}}\right)\right) \cdot \\ &\quad \cdot \left(1 + \frac{1}{\sqrt{n}} \sum_k h_k k + o\left(\frac{1}{\sqrt{n}}\right)\right)^2 + \\ &\quad + 3 \left(\lambda + \frac{1}{\sqrt{n}}h_0 + o\left(\frac{1}{\sqrt{n}}\right)\right)^2 \left(1 + \frac{1}{\sqrt{n}} \sum_k h_k k^2 + o\left(\frac{1}{\sqrt{n}}\right)\right)^2 + \\ &\quad + \left(\lambda + \frac{1}{\sqrt{n}}h_0 + o\left(\frac{1}{\sqrt{n}}\right)\right) \left(1 + \frac{1}{\sqrt{n}} \sum_k h_k k^4 + o\left(\frac{1}{\sqrt{n}}\right)\right) = \\ &= \lambda^4 + 6\lambda^3 + 7\lambda^2 + \lambda + \\ &\quad + \frac{1}{\sqrt{n}} \left(\lambda^4 \sum_k h_k k^4 + \lambda^3 (4h_0 + 6 \sum_k h_k k + 6 \sum_k h_k k^2) + \right. \\ &\quad \quad \left. + \lambda^2 (18h_0 + 4 \sum_k h_k k + 6 \sum_k h_k k^2 + 4 \sum_k h_k k^3) + \right. \\ &\quad \quad \left. + \lambda (8h_0 + 6h_0 + \sum_k h_k k^4) + h_0 \right) + o\left(\frac{1}{\sqrt{n}}\right) \end{aligned}$$

Damit ergibt sich für die Varianz folgender Ausdruck:

$$\text{Var}(Y_{n,\cdot}^2) = E[Y_{n,\cdot}^4] - E[Y_{n,\cdot}^2]^2 =$$

$$\begin{aligned}
 &= \dots = \\
 &= 4\lambda^3 + 6\lambda^2 + \lambda + \\
 &\quad + \frac{1}{\sqrt{n}} \left(\lambda^4 \left(\sum_k h_k k^4 - 4 \sum_k h_k k \right) + \right. \\
 &\quad \quad + \lambda^3 \left(2 \sum_k h_k k + 4 \sum_k h_k k^2 \right) + \\
 &\quad \quad + \lambda^2 \left(12h_0 + 4 \sum_k h_k k + 4 \sum_k h_k k^2 + 4 \sum_k h_k k^3 \right) + \\
 &\quad \quad \left. + \lambda \left(12h_0 + \sum_k h_k k^4 \right) + h_0 \right) + o\left(\frac{1}{\sqrt{n}}\right) \tag{2.13}
 \end{aligned}$$

Bleibt noch die Kovarianz:

$$\begin{aligned}
 \text{Cov}(Y_n, Y_n^2) &= E[Y_n \cdot Y_n^2] - E[Y_n]E[Y_n^2] = \\
 &\stackrel{(2.4)}{=} \lambda_n E[X_n]E[Y_n^2] + 2\lambda_n^2 E[X_n]E[X_n^2] + \lambda_n E[X_n^3] - \lambda_n E[X_n]E[Y_n^2] = \\
 &= 2\lambda_n^2 E[X_n]E[X_n^2] + \lambda_n E[X_n^3] = \\
 &= 2 \left(\lambda + \frac{1}{\sqrt{n}} h_0 + o\left(\frac{1}{\sqrt{n}}\right) \right)^2 \left(1 + \frac{1}{\sqrt{n}} \sum_k h_k k + o\left(\frac{1}{\sqrt{n}}\right) \right) \cdot \\
 &\quad \cdot \left(1 + \frac{1}{\sqrt{n}} \sum_k h_k k^2 + o\left(\frac{1}{\sqrt{n}}\right) \right) + \\
 &\quad + \left(\lambda + \frac{1}{\sqrt{n}} h_0 + o\left(\frac{1}{\sqrt{n}}\right) \right) \left(1 + \frac{1}{\sqrt{n}} \sum_k h_k k^3 + o\left(\frac{1}{\sqrt{n}}\right) \right) = \\
 &= 2\lambda^2 + \lambda + \frac{1}{\sqrt{n}} \left(2\lambda^2 \left(\sum_k h_k k^2 + \sum_k h_k k \right) + \lambda \left(\sum_k h_k k^3 + 4h_0 \right) + h_0 \right) + \\
 &\quad + o\left(\frac{1}{\sqrt{n}}\right) \tag{2.14}
 \end{aligned}$$

Betrachtet man diese Ergebnisse genauer, so sieht man:

Es existiert ein $N \in \mathbb{N}$, so daß für alle $n \geq N$ gilt:

$$\text{Var}(Y_n) \leq \text{Cov}(Y_n, Y_n^2) \leq \text{Var}(Y_n^2)$$

Setzt man dies wiederum in (2.11) ein, so ergibt sich (für $n \geq N$):

$$\begin{aligned}
 b_n &= \text{Var}(a_1 Y_n + a_2 Y_n^2) = a_1^2 \text{Var}(Y_n) + a_2^2 \text{Var}(Y_n^2) + 2a_1 a_2 \text{Cov}(Y_n, Y_n^2) \geq \\
 &\geq a_1^2 \text{Var}(Y_n) + a_2^2 \text{Var}(Y_n) + 2a_1 a_2 \text{Var}(Y_n) = \\
 &= (a_1 + a_2)^2 \text{Var}(Y_n)
 \end{aligned}$$

Somit hat man für den Fall $a_1 \neq -a_2$ eine von n unabhängige untere Schranke für b_n ($n \geq N$).

Im Falle $a_1 = -a_2 \neq 0$ gilt:

$$b_n = a_1^2 \left(\text{Var}(Y_n) + \text{Var}(Y_n^2) - 2\text{Cov}(Y_n, Y_n^2) \right)$$

Betrachtet man nochmals die Ausdrücke für die entsprechenden Terme, so erkennt man, daß es eine Schranke $M \in \mathbb{N}$ geben muß, so daß es für alle $n \geq M$ eine von n unabhängige untere Schranke für b_n gibt.

An dieser Stelle übernehme ich die Definition des Dreiecksschemas nach [Sc99].

Definition 2.3 (Dreiecksschema) Seien $(\tilde{X}_{m,n}), (1 \leq m \leq k(n), n \in \mathbb{N})$ eine Familie von Zufallsvariablen, wobei die Zufallsvariablen $(\tilde{X}_{m,n}), (1 \leq m \leq k(n), n \in \mathbb{N})$ auf einem Wahrscheinlichkeitsraum $(\Omega_n, \mathcal{F}_n, P_n)$ (also abhängig von $n \in \mathbb{N}$) definiert sind. Sei $S_n = \tilde{X}_{1,n} + \dots + \tilde{X}_{k(n),n}$. $(\tilde{X}_{m,n}), (1 \leq m \leq k(n), n \in \mathbb{N})$ nennt man ein Dreiecksschema, falls

- (i) $E_{P_n}(\tilde{X}_{m,n}) = 0, E_{P_n}(\tilde{X}_{m,n}^2) < \infty$ für $1 \leq m \leq k(n), n \in \mathbb{N}$
- (ii) Für alle $n \in \mathbb{N}$ sind die Zufallsvariablen $\tilde{X}_{1,n}, \dots, \tilde{X}_{k(n),n}$ unabhängig
- (iii) $\text{Var}(S_n) = 1, n \in \mathbb{N}$

Mit dieser Definition liefern die Zufallsvariablen $\tilde{X}_{n,i}$ ein Dreiecksschema, wobei die $\tilde{X}_{n,i}$ wie im nachfolgendem Lemma definiert sind:

Lemma 2.4 (Dreiecksschema) Sei

$$\tilde{X}_{n,i} := \frac{(a_1 Y_{n,i} + a_2 Y_{n,i}^2) - (a_1 E[Y_{n,\cdot}] + a_2 E[Y_{n,\cdot}^2])}{\sqrt{nb_n}} \quad (1 \leq i \leq n)$$

Dann definieren die Zufallsvariablen $\tilde{X}_{n,i}$ ein Dreiecksschema nach Definition (2.3). Dabei sei b_n wie folgt definiert: $b_n := \text{Var}(a_1 Y_{n,i} + a_2 Y_{n,i}^2)$ und $a_1, a_2 \in \mathbb{R}$.

Beweis:

Um das Lemma zu beweisen, müssen die Punkte (i) - (iii) der Definition (2.3) nachgewiesen werden. Punkt (i) ist dabei durch die Zentrierung der Zufallsvariablen garantiert:

$$\begin{aligned} E[\tilde{X}_{n,i}] &= \frac{1}{\sqrt{nb_n}} \left(E \left[a_1 Y_{n,i} + a_2 Y_{n,i}^2 - (a_1 E[Y_{n,\cdot}] + a_2 E[Y_{n,\cdot}^2]) \right] \right) = \\ &= \frac{1}{\sqrt{nb_n}} \left(a_1 E[Y_{n,\cdot}] + a_2 E[Y_{n,\cdot}^2] - a_1 E[Y_{n,\cdot}] - a_2 E[Y_{n,\cdot}^2] \right) = 0 \end{aligned}$$

Die Endlichkeit des zweiten zentralen Moments ist durch die Existenz der Varianz gesichert.

Punkt (ii) ist für die $\tilde{X}_{n,i}$ mit der Unabhängigkeit der $Y_{n,i}$ gesichert.

Bleibt noch (iii) zu zeigen. Dies läßt sich aber ebenfalls leicht nachrechnen. Es gilt:

$$\begin{aligned} \text{Var}(S_n) &= \text{Var} \left(\sum_{i=1}^n \tilde{X}_{n,i} \right) = \\ &\stackrel{\text{Unabhängigkeit}}{=} \sum_{i=1}^n \text{Var}(\tilde{X}_{n,i}) = \\ &= \sum_{i=1}^n \text{Var} \left(\frac{(a_1 Y_{n,i} + a_2 Y_{n,i}^2) - (a_1 E[Y_{n,\cdot}] + a_2 E[Y_{n,\cdot}^2])}{\sqrt{nb_n}} \right) = \\ &= \frac{1}{nb_n} \sum_{i=1}^n \underbrace{\text{Var}(a_1 Y_{n,i} + a_2 Y_{n,i}^2)}_{=b_n} = 1 \end{aligned}$$

Damit sind die Voraussetzungen aus der Definition erfüllt, und man hat gezeigt, daß es sich bei den Zufallsvariablen $\tilde{X}_{n,i}$ um ein Dreiecksschema handelt.

□

Somit ist gezeigt, daß die $(\tilde{X}_{n,i})$ ein Dreiecksschema sind. Dies wiederum ist die Voraussetzung für den nun folgenden Satz, der uns die Grenzverteilung der $(\tilde{X}_{n,i})$ liefert. Dieser ist mit einer kleinen Änderung aus [Sc99] übernommen.

Satz 2.5 (Grenzverteilung eines Dreiecksschemas) Sei $(\tilde{X}_{n,i})$ ein Dreiecksschema derart, daß $(\tilde{X}_{n,i}) \subset L_3(P)$ und

$$\sum_{i=1}^n E_P[|\tilde{X}_{n,i}|^3] \rightarrow 0 \text{ für } n \rightarrow \infty \text{ (Ljapunov-Bedingung),}$$

dann gilt:

Es gibt eine Zufallsvariable Z mit $Z \sim N(0, 1)$, so daß

$$S_n \xrightarrow{W} Z.$$

Nach [Ba91] ist die Ljapunov-Bedingung schon erfüllt, wenn für ein $\delta > 0$ gilt:

$$\sum_{i=1}^n E_P[|\tilde{X}_{n,i}|^{2+\delta}] \rightarrow 0 \text{ für } n \rightarrow \infty$$

Somit muß also noch die Ljapunov-Bedingung nachgewiesen werden. Da in unserem Fall immer $Y_{n,i} \geq 0$ gilt, ist auch $|\tilde{X}_{n,i}| = Y_{n,i}$ erfüllt. Damit ist:

$$\begin{aligned} \sum_{i=1}^n E[|\tilde{X}_{n,i}|^{2+\delta}] &= \sum_{i=1}^n E \left[\left| \frac{1}{\sqrt{nb_n}} \left(a_1 Y_{n,i} + a_2 Y_{n,i}^2 - (a_1 E[Y_{n,\cdot}] + a_2 E[Y_{n,\cdot}^2]) \right) \right|^{2+\delta} \right] \leq \\ &\leq n \frac{4^{2+\delta}}{\sqrt{nb_n}^{2+\delta}} E \left[|a_1 Y_{n,\cdot}|^{2+\delta} + |a_2 Y_{n,\cdot}^2|^{2+\delta} + |a_1 E[Y_{n,\cdot}]|^{2+\delta} + |a_2 E[Y_{n,\cdot}^2]|^{2+\delta} \right] \leq \\ &\leq \frac{(4 \max(|a_1|, |a_2|))^{2+\delta}}{b_n \sqrt{nb_n}^\delta} \left(E[(Y_{n,\cdot})^{2+\delta}] + E[(Y_{n,\cdot}^2)^{2+\delta}] + (E[Y_{n,\cdot}])^{2+\delta} + (E[Y_{n,\cdot}^2])^{2+\delta} \right) \end{aligned} \quad (2.15)$$

Hierzu wurde folgende Ungleichung verwendet, die auch in den nachfolgenden Rechnungen noch des öfteren benutzt wird:

$$\left(\sum_{i=1}^k c_i \right)^d \leq (k \cdot \max_{1 \leq i \leq k} c_i)^d \leq k^d \cdot \max_{1 \leq i \leq k} (c_i^d) \leq k^d \cdot \sum_{i=1}^k c_i^d$$

Betrachtet man nun den für die Ljapunov-Bedingung benötigten Ausdruck, so sieht man, daß der Zähler des ersten Faktors von n unabhängig ist und b_n kann, wie bereits zuvor gezeigt, für $n \rightarrow \infty$ von Null wegbeschränkt werden. Somit geht der erste Faktor für $n \rightarrow \infty$ gegen Null. Damit der gesamte Ausdruck ebenfalls noch mit $n \rightarrow \infty$ gegen Null geht, reicht es zu zeigen, daß

die Summanden des zweiten Faktors beschränkt sind. Nach Definition haben die Zufallsvariablen $(X_{n,l})$ einen endlichen Träger und ihre Verteilungen konvergieren gegen δ_1 . Damit konvergieren alle Momente $E[X_{n,l}^\alpha]$ gegen 1. Aus der Potenzreihendarstellung sieht man, daß alle Momente $E[\tau_n^\alpha]$ der mit der Rate λ_n poissonverteilten Zufallsvariablen τ_n gegen $E[\tau^\alpha]$ konvergieren. Betrachten wir also als nächstes die einzelnen Summanden aus Gleichung (2.15) genauer:

$$\begin{aligned}
 E[(Y_{n,\cdot})^{2+\delta}] &= E\left[\left(\sum_{l=1}^{\tau_n} X_{n,l}\right)^{2+\delta}\right] = E\left[\left(\sum_{k=0}^{\infty} 1_{\tau_n=k} \sum_{l=1}^k X_{n,l}\right)^{2+\delta}\right] = \\
 &= E\left[\sum_{k=0}^{\infty} 1_{\tau_n=k} \left(\sum_{l=1}^k X_{n,l}\right)^{2+\delta}\right] = \\
 &= E\left[\sum_{k=0}^{\infty} 1_{\tau_n=k}\right] \cdot E\left[\left(\sum_{l=1}^k X_{n,l}\right)^{2+\delta}\right] = \\
 &= \sum_{k=0}^{\infty} e^{-\lambda_n} \frac{\lambda_n^k}{k!} \cdot E\left[\left(\sum_{l=1}^k X_{n,l}\right)^{2+\delta}\right] \leq \\
 &\leq \sum_{k=0}^{\infty} e^{-\lambda_n} \frac{\lambda_n^k}{k!} \cdot k^{2+\delta} \cdot E\left[\sum_{l=1}^k X_{n,l}^{2+\delta}\right] = \\
 &= \sum_{k=0}^{\infty} e^{-\lambda_n} \frac{\lambda_n^k}{k!} \cdot k^{3+\delta} \cdot E[X_{n,\cdot}^{2+\delta}] = \\
 &= E[\tau_n^{3+\delta}] \cdot E[X_{n,\cdot}^{2+\delta}]
 \end{aligned}$$

Analog wird bei der Berechnung von $E[(Y_{n,\cdot}^2)^{2+\delta}]$ vorgegangen:

$$\begin{aligned}
 E[(Y_{n,\cdot}^2)^{2+\delta}] &= E\left[\left(\sum_{l=1}^{\tau_n} X_{n,l}\right)^{4+2\delta}\right] = E\left[\left(\sum_{k=0}^{\infty} 1_{\tau_n=k} \sum_{l=1}^k X_{n,l}\right)^{4+2\delta}\right] = \\
 &= E\left[\sum_{k=0}^{\infty} 1_{\tau_n=k} \left(\sum_{l=1}^k X_{n,l}\right)^{4+2\delta}\right] = \\
 &= E\left[\sum_{k=0}^{\infty} 1_{\tau_n=k}\right] \cdot E\left[\left(\sum_{l=1}^k X_{n,l}\right)^{4+2\delta}\right] = \\
 &= \sum_{k=0}^{\infty} e^{-\lambda_n} \frac{\lambda_n^k}{k!} \cdot E\left[\left(\sum_{l=1}^k X_{n,l}\right)^{4+2\delta}\right] \leq \\
 &\leq \sum_{k=0}^{\infty} e^{-\lambda_n} \frac{\lambda_n^k}{k!} \cdot k^{4+2\delta} \cdot E\left[\sum_{l=1}^k X_{n,l}^{4+2\delta}\right] = \\
 &= \sum_{k=0}^{\infty} e^{-\lambda_n} \frac{\lambda_n^k}{k!} \cdot k^{5+2\delta} \cdot E[X_{n,\cdot}^{4+2\delta}] = \\
 &= E[\tau_n^{5+\delta}] \cdot E[X_{n,\cdot}^{4+2\delta}]
 \end{aligned}$$

Den dritten bzw. vierten Summanden erhält man mit den Waldschen Gleichungen:

$$E[(Y_{n,\cdot})]^{2+\delta} = \lambda_n^{2+\delta} E[X_{n,\cdot}]^{2+\delta}$$

$$\begin{aligned} E[(Y_n)^2]^{2+\delta} &= \left(\lambda_n^2 E[X_n]^2 + \lambda_n E[X_n^2] \right)^{2+\delta} \leq \\ &\leq 2^{2+\delta} \left(\lambda_n^{4+2\delta} E[X_n]^{4+2\delta} + \lambda_n^{2+\delta} E[X_n^2]^{2+\delta} \right) \end{aligned}$$

Aufgrund der vor diesen Berechnungen getätigten Überlegungen und der Tatsache, daß für eine Poissonverteilung alle Momente existieren und endlich sind, folgt, daß alle Summanden aus Gleichung (2.15) beschränkt sind. Daher konvergiert der gesamte Ausdruck mit $n \rightarrow \infty$ gegen Null und die Voraussetzungen des Satzes sind erfüllt. Die Zufallsvariable S_n konvergiert also gegen eine Standardnormalverteilung. Damit konvergiert auch $a_1 \bar{Y}_n + a_2 \bar{Y}_n^2$, für alle $a_1, a_2 \in \mathbb{R}$ gegen eine Normalverteilung. Dies wiederum hat zur Folge, daß auch die zweidimensionale gemeinsame Verteilung $\left(\frac{\bar{Y}_n}{\bar{Y}_n^2} \right)$ normalverteilt ist, daß also

$$\sqrt{n} \left(\frac{\bar{Y}_n - E[Y_n]}{\bar{Y}_n^2 - E[Y_n^2]} \right) \xrightarrow{W} Z$$

wobei

$$Z \sim N_2(0, \Sigma(\lambda))$$

Die Struktur der Kovarianzmatrix erhält man durch den Grenzübergang $n \rightarrow \infty$ in den Gleichungen (2.12), (2.13) und (2.14).

Da wir jedoch an der gemeinsamen Verteilung von $\frac{1}{\sqrt{n}} \sum_{k=1}^n \begin{pmatrix} Y_{n,k} - E[Y] \\ Y_{n,k}^2 - E[Y^2] \end{pmatrix}$ interessiert sind, muß man zusätzlich noch folgende Rechnung tätigen:

$$\frac{1}{\sqrt{n}} \sum_{k=1}^n \begin{pmatrix} Y_{n,k} - E[Y_n] \\ Y_{n,k}^2 - E[Y_n^2] \end{pmatrix} = \frac{1}{\sqrt{n}} \sum_{k=1}^n \begin{pmatrix} Y_{n,k} - E[Y] \\ Y_{n,k}^2 - E[Y^2] \end{pmatrix} + \sqrt{n} \begin{pmatrix} E[Y] - E[Y_n] \\ E[Y^2] - E[Y_n^2] \end{pmatrix}$$

Weiter weiß man:

$$\begin{aligned} E[Y] &= \lambda \\ E[Y^2] &= \lambda^2 + \lambda \\ E[Y_n] &= \lambda + \frac{1}{\sqrt{n}}(h_0 + \lambda \sum_k h_k k) + o\left(\frac{1}{\sqrt{n}}\right) \\ E[Y_n^2] &= \lambda^2 + \lambda + \frac{1}{\sqrt{n}}(h_0 + \lambda \sum_k h_k k^2) + o\left(\frac{1}{\sqrt{n}}\right) \end{aligned}$$

Und somit erhält man für den gesuchten Ausdruck die folgende Gleichung:

$$\begin{aligned} &\frac{1}{\sqrt{n}} \sum_{k=1}^n \begin{pmatrix} Y_{n,k} - E[Y] \\ Y_{n,k}^2 - E[Y^2] \end{pmatrix} = \\ &= \frac{1}{\sqrt{n}} \sum_{k=1}^n \begin{pmatrix} Y_{n,k} - E[Y_n] \\ Y_{n,k}^2 - E[Y_n^2] \end{pmatrix} - \sqrt{n} \begin{pmatrix} \lambda - \lambda - \frac{1}{\sqrt{n}}(h_0 + \lambda \sum_k h_k k) + o\left(\frac{1}{\sqrt{n}}\right) \\ \lambda^2 + \lambda - \lambda^2 + \lambda + \frac{1}{\sqrt{n}}(h_0 + \lambda \sum_k h_k k^2) + o\left(\frac{1}{\sqrt{n}}\right) \end{pmatrix} = \\ &= \underbrace{\frac{1}{\sqrt{n}} \sum_{k=1}^n \begin{pmatrix} Y_{n,k} - E[Y_n] \\ Y_{n,k}^2 - E[Y_n^2] \end{pmatrix}}_{\rightarrow N_2(0, \Sigma(\lambda))} + \underbrace{\begin{pmatrix} h_0 + \lambda \sum_k h_k k \\ h_0 + \lambda \sum_k h_k k^2 \end{pmatrix}}_{=: \tilde{H}} + \underbrace{o(1)}_{\rightarrow 0} \end{aligned}$$

Daraus kann man nun die Verteilung ableiten:

$$\sqrt{n} \begin{pmatrix} \overline{Y_n} - E[Y] \\ \overline{Y_n^2} - E[Y^2] \end{pmatrix} \xrightarrow{\mathcal{W}} \tilde{Z} \sim N_2 \left(\tilde{H}, \Sigma(\lambda) \right)$$

Um nun noch die Verteilung des Schätzers unter der lokalen Alternative zu erhalten, wendet man auf die gemeinsame Verteilung die Funktion $\varphi(x, y)$ an, die wie im vorhergehenden Abschnitt wie folgt definiert ist:

$$\varphi(x, y) := \frac{y - x^2}{x}$$

Wendet man darauf, ebenfalls analog zum vorigen Abschnitt, die Delta-Methode an (vgl. [Va00]), so ergibt sich:

$$\begin{aligned} & \sqrt{n} \left(\varphi(\overline{Y_n}, \overline{Y_n^2}) - \varphi(\lambda, \lambda^2 + \lambda) \right) \xrightarrow{\mathcal{W}} D\varphi_{\lambda, \lambda^2 + \lambda}(\tilde{Z}) \\ \Rightarrow T_n := & \sqrt{\frac{n}{2}} \left(\frac{\overline{Y_n^2} - (\overline{Y_n})^2}{(\overline{Y_n})^2} - 1 \right) \xrightarrow{\mathcal{W}} \frac{1}{\sqrt{2}} D\varphi_{\lambda, \lambda^2 + \lambda}(\tilde{Z}) =: T \end{aligned}$$

Weiter gilt:

$$\begin{aligned} E[T] &= \frac{1}{\sqrt{2}} D\varphi_{\lambda, \lambda^2 + \lambda} E[\tilde{Z}] = \frac{1}{\sqrt{2}} \left(-\frac{2\lambda + 1}{\lambda}, \frac{1}{\lambda} \right) \begin{pmatrix} h_0 + \lambda \sum_k h_k k \\ h_0 + \lambda \sum_k h_k k^2 \end{pmatrix} = \\ &= \frac{1}{\sqrt{2}} \left(-2h_0 + \sum_k h_k k^2 - (2\lambda + 1) \sum_k h_k k \right) =: H \\ \text{Var}(T) &= 1 \end{aligned}$$

Somit gilt für die Teststatistik T_n unter der lokalen Alternative:

$$T_n \xrightarrow{\mathcal{W}} N(H, 1)$$

2.2 Test mit Hilfe von Kegelapproximationen

2.2.1 Verteilungsfunktion des Tests unter der Hypothese

Dieser Abschnitt wurde bereits ausführlich in [Bu01] behandelt. An dieser Stelle sollen nur noch einmal die wichtigsten Ergebnisse aufgezeigt werden. Die Beweise und Motivationen sind dem oben genannten Werk zu entnehmen.

In diesem Abschnitt soll zur Vereinfachung der Schreibweise der Satz von Skohorod angewendet werden (vgl. [Po84]). Damit gilt: Für die Zufallsvariable q_n und die standardnormalverteilte Zufallsvariable Z gelte:

$$\sqrt{n}(q_n - q_0) \xrightarrow{\mathcal{W}} H + Z$$

Dann existiert ein Wahrscheinlichkeitsraum mit den Zufallsvariablen q'_n und Z' derart, daß:

$$\begin{aligned} q'_n &\stackrel{\mathcal{D}}{=} q_n \text{ und } Z' \stackrel{\mathcal{D}}{=} Z \text{ sowie} \\ \sqrt{n}(q'_n - q_0) &\stackrel{f.s.}{\rightarrow} H + Z' \end{aligned}$$

Da wir lediglich an den Verteilungen interessiert sind, spielt es für uns keine Rolle, ob die Zufallsvariablen q_n und Z oder q'_n und Z' benutzt werden. Allerdings kann bei der fast sicheren Konvergenz analytisch gerechnet werden, was die Handhabung enorm erleichtert. Daher wird im Folgenden mit den Zufallsvariablen q'_n und Z' gerechnet, jedoch aus Gründen der Übersichtlichkeit die Bezeichnung ohne Strich (also q_n und Z) benutzt.

Zurück zum eigentlichen Inhalt dieses Abschnitts. Dafür müssen jedoch zu Beginn noch einige Abkürzungen eingeführt werden. Sei $T_k(x) := x_k, k \in \mathbb{N}_0$ die Projektion auf die k -te Komponente und $T_0^\perp(x) := (0, x_1, x_2, \dots)$. Dabei sei T_k eine Abbildung von $\ell^1 := \ell^1_{\mathbb{R}}(\mathbb{N}_0)$ nach \mathbb{R} und T_0^\perp eine Abbildung von ℓ^1 nach ℓ^1 . Damit läßt sich die Abbildung $\Psi : \ell^1 \rightarrow \ell^1$ definieren:

$$\Psi(x) := e^{(T_0, x)(T_0^\perp x - \delta_0)}$$

Diese Abbildung dient dazu, aus den Werten von (λ, p) mittels Panjerrekursion

$$\begin{aligned} q_0 &= e^{-\lambda} \\ q_k &= \frac{\lambda}{k} \sum_{l=1}^k l p_l q_{k-l} \end{aligned}$$

die Werte von q zu erhalten. Da wir jedoch im weiteren den abgeschnittenen Fall betrachten, muß noch die folgende Abbildung definiert werden:

$$\begin{aligned} \Psi_k^S(x) &:= T_k \Psi(x), \quad k = 0, 1, \dots, S \\ \Psi_{S+1}^S(x) &:= 1 - \sum_{l=0}^S \Psi_l^S(x) \end{aligned}$$

Damit läßt sich nun eine Teststatistik wie folgt definieren:

$$T_1^n := 2n \log \frac{\max_{(\lambda, p)^T \in K_1} \prod_{i=1}^n \Psi_{z_i}^S(\lambda, p)}{\max_{\lambda \in [0, \infty]} \prod_{i=1}^n \Psi_{z_i}^S(\lambda, 1, 0, \dots, 0)}$$

Es handelt sich hierbei, wie man leicht sehen kann, um einen Maximum-Likelihood-Quotiententest. Somit muß K_1 auch für den Zulässigkeitsbereich der Alternative stehen:

$$K_1 := \left\{ (\lambda, p_1, \dots, p_S)^T \in \mathbb{R}^{S+1} : \lambda \in [0, \infty], \sum_{l=1}^S p_l \leq 1, p_l \geq 0, l = 1, \dots, S \right\}$$

Zur weiteren Analyse wird noch der Begriff der Kegelprojektion benötigt.

Eine abgeschlossene konvexe Menge $C \in V$ (V der d -dimensionale euklidische Raum) sei dabei genau dann ein Kegel, wenn für alle $x \in C$ und alle $\alpha \geq 0$ auch $\alpha x \in C$. In dieser Arbeit sind aber nur solche Kegel von Interesse, die von einer Menge $\{\nu_1, \nu_2, \dots, \nu_m\} \in V$ erzeugt werden. Dann sei der von diesen Vektoren erzeugte Kegel wie folgt definiert:

$$\mathcal{C}(\nu_1, \nu_2, \dots, \nu_m) := \left\{ \sum_{k=1}^m \alpha_k \nu_k : \alpha_1, \alpha_2, \dots, \alpha_m \geq 0 \right\}$$

Da nun der Begriff des Kegels definiert ist, kann nun auch die Kegelprojektion definiert werden. Es handelt sich dabei um die Abbildung eines Punktes $x \in V$ auf einen Punkt $y \in C$ mit der Eigenschaft (es gibt dabei aufgrund der Konvexität von C keine Eindeutigkeitsprobleme):

$$\|x - y\| = \inf_{z \in C} \|x - z\|$$

Dafür verwenden wir folgende Notation:

$$\pi(x|C) := y$$

Falls C ein Kegel ist, so ist die nachfolgende Bedingung notwendig und hinreichend dafür, daß $y \in V$ die Kegelprojektion $\pi(x|C)$ ist:

$$(NSC) \quad y \in C \quad \text{und} \quad \langle x - y, z - y \rangle \leq 0 \quad \forall z \in C$$

Anschaulich bedeutet dies, daß es sich bei dem Winkel zwischen dem Vektor $x - y$ und allen Vektoren $z - y$ ($\forall z \in C$) um einen stumpfen Winkel handelt.

Bleibt noch eine letzte Vereinbarung: Die Notationen $\pi_D(x|C)$, $\|\cdot\|_D$ usw. bedeuten immer, daß das innere Produkt bezüglich der Matrix D zu nehmen ist, d.h.: $\langle x, y \rangle_D := x^T D y$.

Mit diesen Vereinbarungen kann dann folgende Aussage gezeigt werden (die genauen Ausführungen sind dem oben genannten Werk zu entnehmen oder können im nachfolgenden Abschnitt für die lokale Alternative in ähnlicher Form nachgelesen werden):

$$T_n^1 \xrightarrow{D} \|\pi_D(Z|V) - \pi_D(Z|C_q)\|_D^2$$

wobei

$$\begin{aligned} Z &\sim N_{S+2} \left(0, (q_i \delta_{ij} - q_i q_j)_{\substack{0 \leq i, j \leq S+1 \\ 0 \leq i, j \leq S+1}} \right) \\ D &:= \text{diag}(1/q_0, \dots, 1/q_S) \\ C_q &:= (\Psi^S)'_{\lambda, p} \mathcal{C}(\Delta_S - (\lambda, p)) \\ \Delta_S &:= \left\{ (\lambda, p_1, p_2, \dots, p_S)^T \in \mathbb{R}^{S+1} : \lambda > 0, \sum_{l=1}^S p_l \leq 1, p_l \geq 0, l = 1, \dots, S \right\} \\ V &:= (\Psi^S)'_{(\lambda, 1, 0, \dots, 0)} \text{lin}\{e_0\} \end{aligned}$$

Dies führt dann nach einigen weiteren Umformungen zur folgenden Aussage:

$$\|\pi_D(Z|V) - \pi_D(Z|C_q)\|_D^2 \sim \sum_{i=0}^S \kappa_i \chi_{S-i}^2$$

D. h. im Grenzfall $n \rightarrow \infty$ erhält man für T_n^1 eine Mischung aus Chi-Quadrat-Verteilungen. Weitere Rechnungen, die an dieser Stelle nicht durchgeführt werden sollen (sie sind erneut bei [Bu01] zu finden), ergeben den folgenden Ausdruck für die Mischungskoeffizienten κ_i :

$$\kappa_i = \sum_{\#I=i} P \left(\tilde{Z} \in \mathcal{C}(\nu_i | i \in I) \oplus (\text{lin}\{\nu_i : i \in I\})^{\perp_D} \cap \{x \in \mathbb{R}^{S+1} : \langle \nu_i, x \rangle_D < 0; i \in I^C\} \right)$$

Betrachtet man sich die Ausdrücke genauer, so erkennt man darin den Zusammenhang zu den Kegeln: Es handelt sich bei den κ_i um die Wahrscheinlichkeit in einem durch die Vektoren ν_i auf eine bestimmte Art aufgespannten Kegel zu landen. Diese Kegel spannen zusammen den gesamten Raum auf, so daß sich die κ_i zu 1 aufaddieren. Des weiteren weiß man aufgrund der Tatsache, daß die κ_i Wahrscheinlichkeiten sind, daß alle κ_i zwischen Null und Eins liegen müssen.

2.2.2 Verteilungsfunktion des Tests unter der lokalen Alternative

Erneut geht es bei der lokalen Alternative darum, die Verteilung der Teststatistik zu berechnen, falls man sich mit der Geschwindigkeit $\frac{1}{\sqrt{n}}$ der Hypothese nähert. Grundlage ist dabei Kapitel 3.5 aus [Bu01], worin der Beweis für die Verteilung unter der Hypothese geführt ist.

Bezeichne in diesem Kapitel \hat{q}^n die relativen Frequenzen unter der lokalen Alternative und

$$\begin{aligned}\hat{q}_H^n &= \operatorname{argmax}_{q \in \Psi^S([0, \infty] \times \{(1, 0, \dots, 0)\})} L(q | \hat{q}^n) \\ \hat{q}_K^n &= \operatorname{argmax}_{q \in \Psi^S(K_1)} L(q | \hat{q}^n)\end{aligned}$$

worin die Funktion $L(\cdot | y)$ wie folgt definiert ist:

$$L(x | y) := \sum_{l=1}^d y_l \log x_l$$

Mithilfe dieser Funktion wird auch noch die Funktion Θ definiert:

$$\Theta(r) := \left\{ P \in K : L(P | r) = \max_{Q \in K} L(Q | r) \right\}$$

Betrachtet man nun den Beweis von Theorem 3.7 aus [Bu01], so erkennt man, daß wiederum alles auf rein analytischen Betrachtungen basiert und somit die gemeinsame Verteilung ohne Probleme angegeben werden kann. Auch zeigt sich, daß die Rechnungen in den Beweisen aus Kapitel 3.3 in [Bu01] ebenso für die lokalen Alternativen durchgeführt werden können. Einzig bei Theorem 3.7 [Bu01] muß man etwas aufpassen. Dies soll an dieser Stelle auch vorgerechnet werden, da dies die entscheidenden Veränderungen in der Verteilung der Teststatistik liefert. Um den Vergleich zu Theorem 3.7 [Bu01] zu erhalten, wird dieses Theorem auch hier in der etwas allgemeineren Form und mit den bei [Bu01] verwendeten Bezeichnungen bewiesen. Daher sei

$$\hat{r}_n := r + \frac{1}{\sqrt{n}} \Gamma + o\left(\frac{1}{\sqrt{n}}\right)$$

mit Γ einer beliebigen Zufallsvariablen, die die Richtung angibt, aus der man sich r nähert.

Um den nachfolgenden Satz verstehen zu können, müssen jedoch erneut erst einige Bezeichnungen eingeführt werden. So nehmen wir an, daß K in einer Umgebung von $P \in K$ durch einen Kegel approximiert werden kann, und zwar in folgendem Sinne:

Sei $O_1 \subset \mathbb{R}^{d-1}$ eine offene konvexe Menge und $O_2 \subset P + \{x \in \mathbb{R}^s : \sum_{l=1}^d x_l = 0\}$ eine offene Menge, die nach Definition den Punkt $P \in K$ enthält. Sei weiter $\Phi : O_1 \rightarrow O_2$ ein C^2 -Diffeomorphismus und $C \in \mathbb{R}^{d-1}$ ein Kegel. Falls

$$\begin{aligned}\Phi(0) &= P \\ \Phi(C \cap O_1) &= K \cap O_2\end{aligned}$$

gilt, so sagen wir, daß K durch einen Kegel in P approximiert wird mit dem Repräsentations-triple (O_1, O_2, Φ, C) .

Dann gilt das folgende Lemma:

Lemma 2.6 *K lasse sich durch einen Kegel in R mit dem Repräsentationstriple (O_1, O_2, Φ, C) approximieren. Sei Z ein $d-1$ -dimensionale Gaußsche Zufallsvariable mit Erwartungswert Null und Kovarianzmatrix I^{-1} , wobei I^{-1} die Inverse der Fisher-Information*

$$I = (\Phi'_0)^T \text{diag}(1/r_0, \dots, 1/r_d) \Phi'_0$$

bezeichne. Dann gilt:

$$\sqrt{n} \Phi^{-1}(\Theta(\hat{r}^n)) \rightarrow \pi_I \left((\Phi'_0)^{-1} \Gamma \middle| C \right)$$

Beweis:

Sei $\gamma := (\Phi'_0)^{-1} \Gamma$.

Analog zum Beweis von Theorem 3.7 aus [Bu01] muß dann noch gezeigt werden, daß

$$\langle \gamma - \eta, c - \eta \rangle_I \leq 0, \quad \forall c \in C$$

Dabei bezeichne η den Grenzwert der Folge $\eta_n := \sqrt{n} \Phi^{-1}(\Theta(\hat{r}^n))$. Wähle nun $c \in C$ beliebig und definiere $c_n := \frac{1}{\sqrt{n}} c$. Wenn n groß genug wird, kann erreicht werden, daß c_n beliebig klein wird und somit $c_n \in C \cap O_1$ gilt. In diesem Fall ist auch

$$DL(\Phi(\cdot)|\hat{r}^n)_{\underbrace{\Phi^{-1}(\Theta(\hat{r}^n))}_{= \frac{1}{\sqrt{n}} \eta_n}} \left(c_n - \Phi^{-1}(\Theta(\hat{r}^n)) \right) \leq 0, \quad (2.16)$$

da dies die Richtungsableitung von $L(\Phi(\cdot)|\hat{r}^n)$ an der Stelle $\Phi^{-1}(\Theta(\hat{r}^n)) = \frac{1}{\sqrt{n}} \eta_n$ in Richtung $(c_n - \Phi^{-1}(\Theta(\hat{r}^n)))$ ist und sich somit auch schreiben läßt als:

$$\lim_{\alpha \searrow 0} \frac{1}{\alpha} \left\{ L \left(\Phi(\Phi^{-1}(\Theta(\hat{r}^n)) + \alpha(c_n - \Phi^{-1}(\Theta(\hat{r}^n)))) | \hat{r}^n \right) - \underbrace{L \left(\Phi(\Phi^{-1}(\Theta(\hat{r}^n))) | \hat{r}^n \right)}_{\text{Maximum}} \right\} \leq 0$$

Darin ist der zweite Term das Maximum der Funktion $L(\cdot|\hat{r}^n)$ und somit ist dieser Ausdruck immer kleiner gleich Null. Daher gilt dies auch für Gleichung (2.16).

Somit gilt weiter:

$$\begin{aligned} 0 &\leq DL(\Phi(\cdot)|\hat{r}^n)_{\frac{1}{\sqrt{n}} \eta_n} \left(\frac{1}{\sqrt{n}} c - \frac{1}{\sqrt{n}} \eta_n \right) \\ \Leftrightarrow 0 &\leq DL(\Phi(\cdot)|\hat{r}^n)_{\frac{1}{\sqrt{n}} \eta_n} \left(c - \eta_n \right) = \\ &= DL(\Phi(\cdot)|\hat{r}^n)_{\Phi^{-1}(\hat{r}^n)} \left(c - \eta_n \right) + \left(\frac{1}{\sqrt{n}} \eta_n - \Phi^{-1}(\hat{r}^n) \right)^T \underbrace{D^2 L(\Phi(\cdot)|\hat{r}^n)_{\xi_n}}_{:= I_n} \left(c - \eta_n \right) \\ &\quad \text{mit } \xi_n \in \text{conv} \left(\frac{1}{\sqrt{n}} \eta_n, \Phi^{-1}(\hat{r}^n) \right) \end{aligned}$$

Der Ausdruck $DL(\Phi(\cdot)|\hat{r}^n)_{\Phi^{-1}(\hat{r}^n)}$ ist jedoch gleich Null, da die Funktion $L(\Phi(\cdot)|\hat{r}^n)$ im Punkt $\Phi^{-1}(\hat{r}^n)$ ihr Maximum annimmt. Somit ist auch

$$0 \leq \left(\frac{1}{\sqrt{n}} \eta_n - \Phi^{-1}(\hat{r}^n) \right)^T I_n \left(c - \eta_n \right)$$

Betrachten wir als nächstes den Ausdruck I_n . Da ξ_n in der konvexen Hülle von $\frac{1}{\sqrt{n}}\eta_n$ und $\Phi^{-1}(\hat{r}^n)$ liegt und sowohl $\frac{1}{\sqrt{n}}\eta_n \xrightarrow{n \rightarrow \infty} 0$ als auch $\Phi^{-1}(\hat{r}^n) \xrightarrow{n \rightarrow \infty} 0$ gilt, muß auch $\xi_n \xrightarrow{n \rightarrow \infty} 0$ erfüllt sein. Dann ist aber auch

$$D^2L(\Phi(\cdot)|\hat{r}^n)_0 = -(\Phi')_0^T \text{diag} \left(\frac{1}{r_1^n}, \dots, \frac{1}{r_d^n} \right) (\Phi')_0 + \underbrace{(1, 1, \dots, 1)\Phi''_0}_{=0} = -I$$

Da die obige Ungleichung auch für den Grenzfalle $n \rightarrow \infty$ erfüllt ist, muß auch folgendes gelten (zusätzlich wurde der erste Faktor noch mit \sqrt{n} aufgeblasen):

$$0 \leq -(\eta_n - \sqrt{n}\Phi^{-1}(\hat{r}^n))^T I (c - \eta_n)$$

Schreibt man nun noch den Term $\sqrt{n}\Phi^{-1}(\hat{r}^n)$ mittels der Taylorentwicklung um und benutzt die Tatsachen, daß $\Phi^{-1}(r)$ gleich Null ist und $\eta_n \rightarrow \eta$ nach Definition der Folge η_n , so erhält man die Ungleichung:

$$0 \geq ((\Phi'_0)^{-1}\Gamma - \eta)^T I (c - \eta) = \langle \gamma - \eta, c - \eta \rangle_I$$

□

Somit kann man auch die für die weitere Berechnung der Verteilung unter der lokalen Alternative notwendige gemeinsame Verteilung berechnen. Es gilt dann:

$$\sqrt{n} \left(\begin{pmatrix} \hat{q}^n \\ \hat{q}_K^n \\ \hat{q}_H^n \end{pmatrix} - \begin{pmatrix} q \\ q \\ q \end{pmatrix} \right) \xrightarrow{\mathcal{D}} \begin{pmatrix} Z + \mu \\ \pi_D(Z + \mu | C_q) \\ \pi_D(Z + \mu | V) \end{pmatrix} \quad (2.17)$$

wobei

$$\mu := (\Psi^S)'_{(\lambda, 1, 0, \dots, 0)} \begin{pmatrix} h_0 \\ \mathbf{h} \end{pmatrix} \quad (2.18)$$

Man kann analog zur Rechnung unter der Hypothesenannahme wie folgt weiterrechnen: Sei dazu die Teststatistik definiert als:

$$T_\infty^n := 2n \log \frac{\max_{(\lambda, p)^T \in K_1} \prod_{i=1}^n \Psi_{z_i}^S(\lambda, p)}{\max_{\lambda \in [0, \infty]} \prod_{i=1}^n \Psi_{z_i}^S(\lambda, 1, 0, \dots, 0)}$$

Dies läßt sich mit den Definitionen von \hat{q}_H^n und \hat{q}_K^n wie folgt umformen:

$$T_\infty^n = -2n (L(\hat{q}_H^n | \hat{q}^n) - L(\hat{q}_K^n | \hat{q}^n))$$

Dieser Ausdruck wird nun mittels Taylorentwicklung im Punkt \hat{q}^n umgeschrieben zu:

$$\begin{aligned} T_\infty^n &= -2n \left(L(\hat{q}_H^n | \hat{q}^n) - L(\hat{q}^n | \hat{q}^n) + L(\hat{q}^n | \hat{q}^n) - L(\hat{q}_K^n | \hat{q}^n) \right) = \\ &= -n \left((\hat{q}_H^n - \hat{q}^n)^T C_n^1 (\hat{q}_H^n - \hat{q}^n) - (\hat{q}_K^n - \hat{q}^n)^T C_n^2 (\hat{q}_K^n - \hat{q}^n) \right) \end{aligned} \quad (2.19)$$

Dies gilt für zwei Matrizen C_n^1 und C_n^2 , die noch näher bestimmt werden müssen. Bei der Taylorentwicklung in obiger Gleichung fehlt der Term erster Ordnung. Dabei würde es sich jedoch um die Richtungsableitung der Funktion $L(\cdot|\hat{q}^n)$ handeln. Diese verschwindet jedoch entlang sämtlicher Richtungen, da \hat{q}^n genau das Maximum der Funktion $L(\cdot|\hat{q}^n)$ ist. Betrachten wir also als nächstes die Matrizen C_n^1 bzw. C_n^2 . Aus der Taylorentwicklung folgt, daß diese Matrizen sich darstellen lassen in der Form:

$$C_n^i = D^2L(\cdot|\hat{q}^n)_{\theta_n^i} \quad i \in \{1, 2\} \quad \text{mit} \quad \theta_n^1 \in \text{conv}\{\hat{q}^n, \hat{q}_H^n\} \\ \theta_n^2 \in \text{conv}\{\hat{q}^n, \hat{q}_K^n\}$$

Aber es gilt, daß mit $n \rightarrow \infty$ auch $\hat{q}_H^n \rightarrow q$, $\hat{q}_K^n \rightarrow q$ und $\hat{q}^n \rightarrow q$ f.s. Daher muß auch $\theta_n^i \rightarrow q$ und, da die Log-Likelihoodfunktion glatt genug nahe q ist, auch $C_n^i \rightarrow -D$ gelten.

Mit der in Gleichung (2.17) hergeleiteten gemeinsamen Verteilung gilt nun für T_∞^n aus Gleichung (2.19) folgende Grenzaussage:

$$T_\infty^n \xrightarrow{\mathcal{D}} \|\pi_D(Z + \mu|V) - (Z + \mu)\|_D^2 - \|\pi_D(Z + \mu|C_q) - (Z + \mu)\|_D^2$$

Zur weiteren Rechnung wird nun die für die Kegelprojektion notwendige und hinreichende Bedingung (NSC) benötigt. Wendet man diese Bedingung auf den Fall der lokalen Alternative an, so ergibt sich:

Da $V = -V \subset C_q$ gilt, erhalten wir für alle ν in V :

$$\langle \nu - \pi_D(Z + \mu|C_q), (Z + \mu) - \pi_D(Z + \mu|C_q) \rangle_D \leq 0$$

Ersetzt man in obiger Ungleichung ν durch $-\nu$, so erhält man die genau entgegengesetzte Ungleichung. Diese beiden Ungleichungen zusammen ergeben dann die nachfolgende Gleichung:

$$\langle \nu - \pi_D(Z + \mu|C_q), (Z + \mu) - \pi_D(Z + \mu|C_q) \rangle_D = 0$$

Dies gilt für alle $\nu \in V$ und somit auch für das spezielle Element $\pi_D(Z + \mu|V)$. Zusammen mit dem Satz von Pythagoras ergibt sich:

$$\|(Z + \mu) - \pi_D(Z + \mu|C_q)\|_D^2 + \|\pi_D(Z + \mu|V) - \pi_D(Z + \mu|C_q)\|_D^2 = \|(Z + \mu) - \pi_D(Z + \mu|V)\|_D^2$$

Und daher:

$$T_\infty^n \xrightarrow{\mathcal{D}} \|\pi_D(Z + \mu|V) - \pi_D(Z + \mu|C_q)\|_D^2$$

2.3 Vergleich der Tests mittels lokaler Alternativen

Nachdem nun die jeweiligen Verteilungen berechnet wurden, wollen wir ermitteln, welches der bessere Test ist. Da bei van der Vaart [Va00] lediglich der Fall einer eindimensionalen Hypothese behandelt wird, muß man für unseren Fall eine allgemeinere Vergleichsmethode finden. Bei van der Vaart wird dabei folgender Test betrachtet: $H_0 : \theta = 0$ gegen $H_1 : \theta > 0$. Zum Vergleich zweier Tests definiert er dann die lokale Grenzfunktion der Gütefunktion wie folgt:

$$\pi(h) = \lim_{n \rightarrow \infty} \pi_n \left(\frac{h}{\sqrt{n}} \right), \quad h \geq 0$$

Dieser Grenzwert existiert normalerweise und kann wie folgt bestimmt werden:

Sei dazu angenommen, daß die Teststatistik T_n für alle Folgen $\theta_n = h/\sqrt{n}$ asymptotisch normalverteilt ist,

$$\frac{\sqrt{n}(T_n - \mu(\theta_n))}{\sigma(\theta_n)} \xrightarrow{\theta_n} N(0, 1)$$

Dabei stehen $\mu(\theta)$ und $\sigma^2(\theta)$ oft für den Erwartungswert und die Varianz von T_n , allerdings muß dies nicht der Fall sein. Mit dieser Bedingung ist nun der Test, der die Hypothese ablehnt für den Fall, daß $\sqrt{n}(T_n - \mu(0))$ den Wert $\sigma(0)z_\alpha$ übersteigt, asymptotisch vom Level α (z_α ist dabei das α -Quantil der Standardnormalverteilung). Die Gütefunktion kann somit geschrieben werden als

$$\pi_n(\theta_n) = P_{\theta_n} \left(\sqrt{n}(T_n - \mu(\theta_n)) > \sigma(0)z_\alpha - \sqrt{n}(\mu(\theta_n) - \mu(0)) \right)$$

Für $\theta_n = h/\sqrt{n}$ konvergiert die Folge $\sqrt{n}(\mu(\theta_n) - \mu(0))$ nach $h\mu'(0)$ (falls μ differenzierbar in Null). Falls auch $\sigma(\theta_n) \rightarrow \sigma(0)$, dann gilt:

$$\pi_n \left(\frac{h}{\sqrt{n}} \right) \rightarrow 1 - \Phi \left(z_\alpha - h \frac{\mu'(0)}{\sigma(0)} \right)$$

Der Ausdruck $\mu'(0)/\sigma(0)$ wir dabei als „slope“ bezeichnet.

Betrachtet man sich jedoch die Vorgehensweise bei van der Vaart genauer, so erkennt man, daß der dort eingeführte Begriff „slope“ ein Maß für die Steigung der Gütefunktion ist. Dabei gilt: Je größer der „slope“, desto steiler ist die Gütefunktion auf Seiten der Alternative und desto schärfer trennt der Test. Somit steht ein größerer „slope“ für einen besseren Test. Dieses Prinzip kann auch in den mehrdimensionalen Fall übertragen werden, indem man analog zu [Va00] die Richtungsableitungen betrachtet.

Da es sich jedoch als schwer erwiesen hat, die Verteilung von $\|\pi_D(Z + \mu|V) - \pi_D(Z + \mu|C_q)\|_D^2$ zu bestimmen, sollen hier nicht die jeweiligen „slopes“ verglichen werden. Wir werden für den Dispersionstest die Richtungsableitungen bestimmen, die den Test möglichst scharf bzw. schwach trennen. Diese Richtungsvektoren seien mit Maximum bzw. Minimum des Dispersionstests bezeichnet.

2.3.1 Berechnung von Maximum und Minimum des Dispersionstests

Unter der lokalen Alternative gilt:

$$T_n \xrightarrow{\mathcal{W}} N(H, 1)$$

wo

$$H = \frac{1}{\sqrt{2}} \left(-2h_0 + \sum_k h_k k^2 - (2\lambda + 1) \sum_k h_k k \right)$$

Zur vereinfachten Schreibweise werden noch folgende Abkürzungen eingeführt:

$$\bar{\gamma} := \begin{pmatrix} \gamma_0 \\ \gamma \\ \gamma \end{pmatrix} := \begin{pmatrix} \gamma_0 \\ \gamma_1 \\ \vdots \\ \gamma_S \end{pmatrix}$$

$$h := \begin{pmatrix} h_0 \\ h_1 \\ \vdots \\ h_S \end{pmatrix}$$

mit

$$\begin{aligned} \gamma_0 &= -\frac{2}{\sqrt{2}} \\ \gamma_k &= \frac{1}{\sqrt{2}}(k^2 - (2\lambda + 1)k) \quad (1 \leq k \leq n). \end{aligned}$$

Mit diesen Definitionen erhält man:

$$H = \bar{\gamma}^T h$$

Somit erhält man als Gütefunktion auf der Alternative folgenden Ausdruck:

$$1 - \Phi(z_\alpha - \bar{\gamma}^T h)$$

Die Richtungsableitung dieser Funktion ist:

$$(\nabla(1 - \Phi(z_\alpha - \bar{\gamma}^T h))|_{h=0})^T \begin{pmatrix} e_0 \\ e \end{pmatrix} = \varphi(z_\alpha) \cdot \bar{\gamma}^T \begin{pmatrix} e_0 \\ e \end{pmatrix}$$

wo $\bar{e} := \begin{pmatrix} e_0 \\ e \end{pmatrix}$ ein zulässiger Richtungsvektor der Länge 1 ist. Betrachten wir die Eigenschaften, die $\begin{pmatrix} e_0 \\ e \end{pmatrix}$ erfüllen muß, um zulässig zu sein: Da es sich bei der Richtung, aus der man sich der Hypothese nähert (also $\delta_1 + e$), um ein Wahrscheinlichkeitsmaß handeln muß, gelten folgende Bedingungen:

$$\begin{aligned} \sum_{i=0}^S e_i^2 &= 1 \quad (\text{Einheitsvektor}) \\ \sum_{i=1}^S e_i &= 0 \\ e_1 &\leq 0 \\ e_i &\geq 0 \quad (2 \leq i \leq S) \end{aligned}$$

Die Menge aller Vektoren, die diese Bedingungen erfüllen, sei mit \bar{E} bezeichnet (beginnt die Summe in der ersten Bedingung erst bei $i = 1$, so sei die Menge der Vektoren mit E bezeichnet). Um nun die Richtungen zu erhalten, für die der Test besonders gut (bzw. schlecht) trennt, gilt es (da $\varphi(z_\alpha)$ konstant bleibt), den Ausdruck $\bar{\gamma}^T \begin{pmatrix} e_0 \\ e \end{pmatrix}$ zu maximieren (bzw. minimieren).

Maximum- und Minimumberechnung unter $e_0 = 0$

Setzen wird zuerst einmal die Komponente $e_0 = 0$, da sie - im Gegensatz zu den anderen Komponenten - die Richtung beschreibt, aus der man sich λ nähert.

Somit ergibt sich folgendes Problem:

$$\begin{aligned} & \text{Maximiere (Minimiere) } \gamma^T e \\ & e \in E \end{aligned}$$

Betrachten wir die Nebenbedingungen etwas genauer, so sieht man, daß durch die Bedingung $\sum_{i=1}^S e_i = 0$ eine Hyperebene beschrieben wird. Daher kann die Menge der dafür zulässigen Vektoren durch einen $(S - 1)$ -dimensionalen Raum, der von $(S - 1)$ linear unabhängigen Vektoren aufgespannt wird, beschrieben werden. Um eine Orthonormalbasis zu erhalten, müssen diese Vektoren zueinander senkrecht und normiert gewählt werden. Wählt man die Vektoren darüberhinaus auch noch senkrecht zum Vektor $(1, \dots, 1)^T$, so liegen alle Vektoren des durch diese linear unabhängigen Vektoren aufgespannten Raumes in obiger Hyperebene.

Solche Vektoren sind z.B.:

$$\frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \frac{1}{\sqrt{6}} \begin{pmatrix} -1 \\ -1 \\ 2 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \frac{1}{\sqrt{12}} \begin{pmatrix} -1 \\ -1 \\ -1 \\ 3 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \frac{1}{\sqrt{20}} \begin{pmatrix} -1 \\ -1 \\ -1 \\ -1 \\ 4 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \dots$$

Die Vektoren seien mit $z_1, z_2, z_3, \dots, z_S$ bezeichnet.

Aufgrund der Bedingung, daß e ein Einheitsvektor sein soll, ergibt sich eine lineare Zielfunktion unter einer quadratischen Nebenbedingung. Dies ist im Allgemeinen nicht leicht zu lösen. Daher wird hier versucht, das Problem mittels einer Darstellung über Kugelkoordinaten zu behandeln. Denn die Einschränkung auf die Menge der Einheitsvektoren läßt sich auf diese Weise leichter darstellen. Bei der Verwendung der Einheitsvektoren sieht die Darstellung wie folgt aus:

$$\begin{pmatrix} \sin(\theta_{S-2}) \cdots \sin(\theta_1) \\ \sin(\theta_{S-2}) \cdots \sin(\theta_2) \cos(\theta_1) \\ \vdots \\ \sin(\theta_{S-2}) \sin(\theta_{S-3}) \cos(\theta_{S-4}) \\ \sin(\theta_{S-2}) \cos(\theta_{S-3}) \\ \cos(\theta_{S-2}) \end{pmatrix}$$

Da hier jedoch nicht die Einheitsvektoren benutzt werden, sieht die Kugeldarstellung wie folgt aus:

$$z_1 \sin(\theta_{n-2}) \cdots \sin(\theta_1) + z_2 \sin(\theta_{n-2}) \cdots \sin(\theta_2) \cos(\theta_1) + \dots + z_{S-1} \cos(\theta_{S-2})$$

Somit sind alle Vektoren, die sich durch so eine Linearkombination darstellen lassen, genau diejenigen, die die ersten beiden Bedingungen an den Richtungsvektor erfüllen. Bleiben also nur noch die Vorzeichenbedingungen. Dabei kann man erkennen, daß aufgrund der Wahl der Vektoren z_1, \dots, z_S nur positive Werte für die Linearkombinationskoeffizienten zulässig sind. Die weiteren Bedingungen an die Variablen $\theta_1, \dots, \theta_{S-2}$ erhält man, indem man die entsprechenden

Zeilen aus obiger Kugeldarstellung Null setzt. Somit ist das Optimierungsproblem zu einem um zwei Dimensionen verkleinerten Problem mit Nebenbedingungen geworden.

Da das allgemeine Problem jedoch immer noch sehr komplex ist, sollen an dieser Stelle nur die Fälle $S = 3$ und $S = 4$ explizit behandelt werden. Danach soll versucht werden, aus dem Verhalten in diesen niedrigdimensionalen Fällen Rückschlüsse auf den allgemeinen Fall zu tätigen.

Fall $S = 3$:

Betrachten wir zuerst die Orthonormalbasis in diesem Fall. Sie wird gebildet von den folgenden Vektoren:

$$x = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \quad \text{und} \quad y = \frac{1}{\sqrt{6}} \begin{pmatrix} -1 \\ -1 \\ 2 \end{pmatrix}$$

Somit läßt sich die Kugel wie folgt darstellen:

$$\psi(\varphi) := \frac{1}{\sqrt{2}} \cos(\varphi) \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} + \frac{1}{\sqrt{6}} \sin(\varphi) \begin{pmatrix} -1 \\ -1 \\ 2 \end{pmatrix} \quad (2.20)$$

Der Vektor γ besitzt folgende Struktur:

$$\gamma = \frac{1}{\sqrt{2}} \begin{pmatrix} -2\lambda \\ 2 - 4\lambda \\ 6 - 6\lambda \end{pmatrix}$$

Daraus ergibt sich folgende Zielfunktion:

$$\begin{aligned} f(\varphi) = \psi(\varphi)^T \gamma &= \cos(\varphi) \left[\frac{1}{\sqrt{2}} \frac{1}{\sqrt{2}} \left((-1)(-2\lambda) + 1(2 - 4\lambda) + 0(6 - 6\lambda) \right) \right] + \\ &+ \sin(\varphi) \left[\frac{1}{\sqrt{6}} \frac{1}{\sqrt{2}} \left((-1)(-2\lambda) + (-1)(2 - 4\lambda) + 2(6 - 6\lambda) \right) \right] = \\ &= (1 - \lambda) \cos(\varphi) + \frac{1}{\sqrt{3}} (5 - 3\lambda) \sin(\varphi) \end{aligned}$$

Aus Gleichung (2.20) ergibt sich, daß φ in nachfolgendem Intervall liegen muß, um die Vorzeichenbeschränkung aus den Nebenbedingungen zu erfüllen:

$$\varphi \in [0^\circ, 60^\circ]$$

Berechnen wir weiter die ersten beiden Ableitungen der Funktion f :

$$\begin{aligned} f' &= -(1 - \lambda) \sin(\varphi) + \frac{1}{\sqrt{3}} (5 - 3\lambda) \cos(\varphi) \\ f'' &= -(1 - \lambda) \cos(\varphi) - \frac{1}{\sqrt{3}} (5 - 3\lambda) \sin(\varphi) \end{aligned}$$

Damit ergibt sich als Extremum von f ($\lambda \neq 1$):

$$\begin{aligned} 0 &\stackrel{!}{=} f'(\varphi) = (\lambda - 1) \sin(\varphi) + \frac{1}{\sqrt{3}} (5 - 3\lambda) \cos(\varphi) \\ &\Leftrightarrow \tan(\varphi) = \frac{\sin(\varphi)}{\cos(\varphi)} = \frac{5 - 3\lambda}{\sqrt{3}(1 - \lambda)} \\ &\Leftrightarrow \varphi = \arctan \left(\frac{5 - 3\lambda}{\sqrt{3}(1 - \lambda)} \right) \end{aligned}$$

Setzt man dies in die zweite Ableitung ein, so erhält man:

$$\begin{aligned}
 f''(\varphi) &= -(1-\lambda)\cos(\varphi) - \frac{1}{\sqrt{3}}(5-3\lambda)\sin(\varphi) \stackrel{\sin(\varphi)=\cos(\varphi)\frac{5-3\lambda}{\sqrt{3}(1-\lambda)}}{=} \\
 &= (\lambda-1)\cos(\varphi) - \frac{1}{\sqrt{3}}(5-3\lambda)\cos(\varphi)\frac{5-3\lambda}{\sqrt{3}(1-\lambda)} = \\
 &= \frac{\cos(\varphi)}{\lambda-1} \underbrace{\left((\lambda-1)^2 + \frac{1}{3}(5-3\lambda)^2 \right)}_{>0} = \frac{c}{\lambda-1} \quad \text{mit } c > 0
 \end{aligned}$$

Somit handelt es sich bei dem Extremum für $\lambda > 1$ um ein Minimum, für $\lambda < 1$ um ein Maximum.

Bleibt nur noch zu überprüfen, ob das Extremum auch im Intervall $[0^\circ, 60^\circ]$ liegt. Dies ist jedoch durch eine einfache Rechnung möglich. Insgesamt ergibt sich folgender Sachverhalt:

λ	$(0; \frac{5}{3}]$	$[\frac{5}{3}; 2)$	2	$(2, \infty)$
Max	60°	60°	0° und 60°	0°
Min	0°	$\arctan\left(\frac{5-3\lambda}{\sqrt{3}(1-\lambda)}\right)$	$30^\circ = \arctan\left(\frac{5-3\cdot 2}{\sqrt{3}(1-2)}\right)$	$\arctan\left(\frac{5-3\lambda}{\sqrt{3}(1-\lambda)}\right)$

Setzt man die Winkel in die Kugeldarstellung (2.20) ein, so ergeben sich folgende Vektoren:

$$\begin{aligned}
 60^\circ &\hat{=} \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \\
 0^\circ &\hat{=} \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}
 \end{aligned}$$

Auffällig an diesem Ergebnis ist dabei der Sprung des Maximums für $\lambda = 2$. Während für $\lambda < 2$ der Vektor $\frac{1}{\sqrt{2}}(-1, 0, 1)^T$ das Maximum bildet, springt das Maximum für $\lambda > 2$ auf den Vektor $\frac{1}{\sqrt{2}}(-1, 1, 0)^T$, der zudem noch für $\lambda < \frac{5}{3}$ das Minimum darstellt. Beim Minimum hingegen gibt es keine Unstetigkeitsstellen, für $\lambda \rightarrow \infty$ wird das Minimum jedoch im Vektor $\frac{1}{\sqrt{2}}(-1, 0, 1)^T$ angenommen, der wiederum für $\lambda < \frac{5}{3}$ das Maximum der Funktion ist.

Fall S = 4:

Betrachten wir auch hier zuerst die Vektoren der Orthonormalbasis. Diese sind:

$$x = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad y = \frac{1}{\sqrt{6}} \begin{pmatrix} -1 \\ -1 \\ 2 \\ 0 \end{pmatrix} \quad \text{und} \quad z = \frac{1}{\sqrt{12}} \begin{pmatrix} -1 \\ -1 \\ -1 \\ 3 \end{pmatrix}$$

Damit hat die Einheitskugel die folgende Gestalt:

$$\psi(\varphi, \theta) = x \cos(\varphi) \sin(\theta) + y \sin(\varphi) \sin(\theta) + z \cos(\theta)$$

Durch Nullsetzen der entsprechenden Zeilen dieser Kugeldarstellung erhält man die Bedingungen, die (φ, θ) erfüllen müssen, um eine zulässige Lösung darzustellen. Es ergibt sich:

- $\varphi \in [0, 60^\circ]$
- $\theta \in [\arctan\left(\frac{1}{\sqrt{2}}\right), 90^\circ]$
- $2\sqrt{2} \tan(\theta) \sin(\varphi) \geq 1$
- $\left(\sqrt{6} \cos(\varphi) - \sqrt{2} \sin(\varphi)\right) \tan(\theta) \geq 1$

Berechnen wir als nächstes die Zielfunktion:

$$\begin{aligned}
 f(\varphi, \theta) &= \psi(\varphi, \theta)^T \gamma = \\
 &= \cos(\varphi) \sin(\theta) \left[\frac{1}{\sqrt{2}} \frac{1}{\sqrt{2}} \left((-1)(-2\lambda) + 1(2 - 4\lambda) \right) \right] + \\
 &\quad + \sin(\varphi) \sin(\theta) \left[\frac{1}{\sqrt{2}} \frac{1}{\sqrt{6}} \left((-1)(-2\lambda) + (-1)(2 - 4\lambda) + 2(6 - 6\lambda) \right) \right] + \\
 &\quad + \cos(\theta) \left[\frac{1}{\sqrt{2}} \frac{1}{\sqrt{12}} \left((-1)(-2\lambda) + (-1)(2 - 4\lambda) + (-1)(6 - 6\lambda) + 3(12 - 8\lambda) \right) \right] = \\
 &= (1 - \lambda) \cos(\varphi) \sin(\theta) + \frac{1}{\sqrt{3}}(5 - 3\lambda) \sin(\varphi) \sin(\theta) + \frac{1}{\sqrt{6}}(14 - 6\lambda) \cos(\theta)
 \end{aligned}$$

Um ein Extremum dieser Funktion zu erhalten, muß der Gradient gleich Null gesetzt werden. Dazu berechnen wir die partiellen Ableitungen:

$$\begin{aligned}
 \frac{\delta}{\delta \varphi} f(\varphi, \theta) &= f'(\varphi) \sin(\theta) \\
 \frac{\delta}{\delta \theta} f(\varphi, \theta) &= \cos(\theta) \left[(1 - \lambda) \cos(\varphi) + \frac{1}{\sqrt{3}}(5 - 3\lambda) \sin(\varphi) \right] - \frac{1}{\sqrt{6}}(14 - 6\lambda) \sin(\theta)
 \end{aligned}$$

Setzt man diese Ausdrücke gleich Null, so ergibt sich bei der partiellen Ableitung nach φ der gleiche Ausdruck wie im Fall $S = 3$, für die partielle Ableitung nach θ gilt (für $\lambda \neq \frac{7}{3}$):

$$\tan(\theta) = \frac{(1 - \lambda) \cos(\varphi) + \frac{1}{\sqrt{3}}(5 - 3\lambda) \sin(\varphi)}{\frac{1}{\sqrt{6}}(14 - 6\lambda)}$$

Setzt man darin noch die Bedingung ein, daß auch die partielle Ableitung nach φ gleich Null sein soll (und einen für φ zulässigen Punkt liefert), so erhält man:

$$\begin{aligned}
 \tan(\theta) &= \frac{\sqrt{6} \cos(\varphi)}{(14 - 6\lambda)(1 - \lambda)} \left(4\lambda^2 - 12\lambda + \frac{28}{3} \right) = \\
 &= \frac{\sqrt{6}}{6\lambda - 14} \sqrt{4\lambda^2 - 12\lambda + \frac{28}{3}}
 \end{aligned}$$

Insgesamt ergibt sich als Extrempunkt für $\lambda > \frac{7}{3}$ das Paar:

$$(\varphi_0, \theta_0) = \left(\arctan\left(\frac{5 - 3\lambda}{\sqrt{3}(1 - \lambda)}\right), \arctan\left(\frac{\sqrt{6\lambda^2 - 18\lambda + 14}}{3\lambda - 7}\right) \right) \quad (2.21)$$

Für $\lambda \leq \frac{7}{3}$ weiß man, daß es entweder kein Extremum gibt oder daß der Punkt mit verschwindenden Gradienten sicher außerhalb der zulässigen Menge liegt. Als nächstes wird überprüft, ob das Extremum auch ein zulässiger Wert für (φ, θ) ist. Dazu betrachtet man z.B. die Randbedingung

$$\left(\sqrt{6} \cos(\varphi) - \sqrt{2} \sin(\varphi) \right) \tan(\theta) \geq 1$$

und berechnet, für welche Werte von λ diese unter den Extremumsbedingungen erfüllt ist:

$$\begin{aligned} & \left(\sqrt{6} \cos(\varphi) - \sqrt{2} \sin(\varphi) \right) \tan(\theta) = \\ & = \left(\sqrt{6} \cos(\varphi) - \sqrt{2} \frac{5-3\lambda}{\sqrt{3}(1-\lambda)} \cos(\varphi) \right) \frac{\sqrt{6}}{6\lambda-14} \sqrt{4\lambda^2-12\lambda+\frac{28}{3}} = \\ & = \frac{2}{3\lambda-7} \geq 1 \quad \text{für } \lambda \in \left(\frac{7}{3}, 3 \right] \end{aligned}$$

Die andere Randbedingung ist immer erfüllt. Somit hat man für $\lambda \in (\frac{7}{3}, 3]$ einen zulässigen Extrempunkt. Um noch zu erfahren, ob es sich dabei um ein Maximum, Minimum oder einen Sattelpunkt handelt, muß die Hessematrix in diesem Punkt berechnet werden. Mittels MAPLE wurde nachgerechnet, daß es sich um ein Minimum handelt.

Für die anderen Werte von λ weiß man, daß sowohl das Minimum als auch das Maximum auf dem Rand des Zulässigkeitsbereichs liegen müssen. Berechnet man die Extremstellen auf den Rändern, so erhält man letztendlich die folgende Tabelle (die Abkürzung von $(\varphi_0|\theta_0)$ nach Gleichung (2.21)):

λ	$(0, \frac{5}{3}]$	$[\frac{5}{3}, \frac{7}{3}]$	$[\frac{7}{3}, \frac{5}{2})$
<i>Max</i> $(\varphi \theta)$	$(30^\circ \arctan(\frac{1}{\sqrt{2}}))$	$(30^\circ \arctan(\frac{1}{\sqrt{2}}))$	$(30^\circ \arctan(\frac{1}{\sqrt{2}}))$
<i>Min</i> $(\varphi \theta)$	$(0^\circ 90^\circ)$	$(\arctan(\frac{5-3\lambda}{\sqrt{3}(1-\lambda)}) 90^\circ)$	$(\varphi_0 \theta_0)$

λ	$\frac{5}{2}$	$(\frac{5}{2}, 3]$	$[3, \infty)$	' ∞ '
<i>Max</i> $(\varphi \theta)$	$(30^\circ \arctan(\frac{1}{\sqrt{2}}))$ und $(0^\circ 90^\circ)$	$(0^\circ 90^\circ)$	$(0^\circ 90^\circ)$	$(0^\circ 90^\circ)$
<i>Min</i> $(\varphi \theta)$	$(\arctan(\frac{5\sqrt{3}}{9}) \arctan(\sqrt{26}))$	$(\varphi_0 \theta_0)$	$(m_1 m_2)$	$(\arctan(\frac{7\sqrt{3}}{15}) \arctan(\frac{\sqrt{62}}{8}))$

wo m_1 das eindeutige Minimum (für $\varphi \in [30^\circ, 60^\circ]$) der Funktion $\frac{(15-7\lambda)\cos(\varphi)+\sqrt{3}(\lambda-3)\sin(\varphi)}{\sqrt{1+(\sqrt{6}\cos(\varphi)-\sqrt{2}\sin(\varphi))^2}}$ und m_2 die Lösung der Gleichung $\left(\sqrt{6} \cos(m_1) - \sqrt{2} \sin(m_1) \right) \tan(m_2) = 1$ ist. Mittels MAPLE

wurden diese Werte noch berechnet. Dabei ergab sich:

$$m_1 = \arccos\left(\frac{5\lambda - 9}{2} \sqrt{\frac{3}{81 - 99\lambda + 31\lambda^2}}\right)$$

$$m_2 = \arctan\left(\frac{\sqrt{162 - 198\lambda + 62\lambda^2}}{8\lambda - 18}\right)$$

Betrachtet man auch in diesem Fall die Winkel von Maximum und Minimum für $\lambda \in (0, \frac{5}{3}]$, so ergeben sich folgende zugehörige Richtungsvektoren:

$$(\varphi|\theta) = (30^\circ|\arctan(\frac{1}{\sqrt{2}})) \cong \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

$$(\varphi|\theta) = (0^\circ|90^\circ) \cong \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \end{pmatrix}$$

Auch hier läßt sich analog zum Fall $S = 3$ feststellen, daß der Vektor $\frac{1}{\sqrt{2}}(-1, 1, 0, 0)^T$, der für $\lambda < \frac{5}{3}$ noch das Minimum darstellt, für $\lambda \geq \frac{5}{3}$ zum Maximum wird.

Hieran sieht man schon, wie schnell die Komplexität mit der Dimension wächst. Daher wurde das Verhalten der Extrema für höhere Dimensionen nicht mehr berechnet, sondern zu Beginn des Abschnitts lediglich der allgemeine Weg erklärt, wie die Extrema zu berechnen sind. Betrachtet man an dieser Stelle noch einmal das Optimierungsproblems, so sieht man, daß darin sogar der Fall $\lambda = 0$ zulässig ist. Als Grenzverhalten für $\lambda \rightarrow 0$ wurden daher für höhere Dimensionen noch die Lösungen für $\lambda = 0$ ermittelt. Dies geschah u.a. unter Zuhilfenahme des Optimierungstools von *matlab*. Betrachtet man dabei für $S \in \{5, 6, \dots, 9\}$ die Ergebnisse, so sieht man, daß dort nur Masse in den jeweils letzten beiden Komponenten vorhanden ist. Unter dieser Annahme konnte dann mittels Lagrange-Multiplikatoren die genaue Lösung bestimmt werden.

Es ergeben sich letztendlich folgende Richtungsvektoren:

Maximum:

$S = 5$	$S = 6$	$S = 7$	$S = 8$	$S = 9$
$\frac{1}{\sqrt{114}} \begin{pmatrix} -8 \\ 0 \\ 0 \\ 1 \\ 7 \end{pmatrix}$	$\frac{1}{\sqrt{42}} \begin{pmatrix} -5 \\ 0 \\ 0 \\ 0 \\ 1 \\ 4 \end{pmatrix}$	$\frac{1}{\sqrt{26}} \begin{pmatrix} -4 \\ 0 \\ \vdots \\ 0 \\ 1 \\ 3 \end{pmatrix}$	$\frac{1}{\sqrt{78}} \begin{pmatrix} -7 \\ 0 \\ \vdots \\ 0 \\ 2 \\ 5 \end{pmatrix}$	$\frac{1}{\sqrt{402}} \begin{pmatrix} -16 \\ 0 \\ \vdots \\ 0 \\ 5 \\ 11 \end{pmatrix}$

$S = 10$	$S = 15$	$S = 20$
$\begin{pmatrix} -0.8089 \\ 0 \\ \vdots \\ 0 \\ 0.0223 \\ 0.2598 \\ 0.5269 \end{pmatrix}$	$\begin{pmatrix} -0.8418 \\ 0 \\ \vdots \\ 0 \\ 0.1168 \\ 0.2765 \\ 0.4486 \end{pmatrix}$	$\begin{pmatrix} -0.8741 \\ 0 \\ \vdots \\ 0 \\ 0.0783 \\ 0.1690 \\ 0.2639 \\ 0.3628 \end{pmatrix}$

Minimum:

Für die Richtungsvektoren des Minimierungsproblems kann gezeigt werden, daß für alle S der Lösungsvektor für $\lambda \leq 1$ die folgende Struktur besitzt:

$$\frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Dies ergibt sich, wenn man sich das Minimierungsproblem genauer betrachtet. So kann z.B. festgestellt werden, daß die erste Komponente von γ immer negativ ist; im Fall $\lambda \leq \frac{1}{2}$ gilt darüberhinaus auch noch $\gamma_k \geq 0$ ($k = 2, 3, \dots, S$), während im Fall $\frac{1}{2} \leq \lambda \leq 1$ folgendes gilt: $\gamma_2 \leq 0$, $\gamma_k \geq 0$ ($k = 3, 4, \dots, S$).

Im Gegensatz zum Maximierungsproblem bleibt die Masse also auf die ersten beiden Komponenten konzentriert.

Maximum- und Minimumberechnung unter Berücksichtigung der Komponente e_0

Bis jetzt wurde bei allen Berechnungen davon ausgegangen, daß die Komponente e_0 gleich Null ist. Im weiteren soll nun erreicht werden, daß auch diese Komponente zum Optimierungsproblem hinzugenommen wird. Es geht also darum, folgendes Problem zu lösen:

$$\begin{aligned} \text{Maximiere (Minimiere)} \quad & \bar{\gamma}^T \begin{pmatrix} e_0 \\ e \end{pmatrix} \\ & \begin{pmatrix} e_0 \\ e \end{pmatrix} \in \bar{E} \end{aligned}$$

Allerdings sollten dazu auch die Ergebnisse, die bereits im vorangehenden Abschnitt berechnet wurden, wieder verwendet werden. Somit wird ein Ansatz benötigt, der aus der Optimallösung für den Fall $e_0 = 0$ den Fall der insgesamt optimalen Lösung herleitet. Dazu dient folgender Satz:

Satz 2.7 Sei Ext Optimallösung des Optimierungsproblems unter der Voraussetzung $e_0 = 0$ und M ein beliebiger zulässiger Vektor sowie $p \in \mathbb{R}$ beliebig. Dann gilt:

$$\bar{Ext} := \frac{1}{\sqrt{p^2+1}} \begin{pmatrix} p \\ Ext \end{pmatrix} \text{ ist eine bessere Lösung als } \bar{M} := \frac{1}{\sqrt{p^2+1}} \begin{pmatrix} p \\ M \end{pmatrix}$$

Dieser Satz hat zur Folge, daß die Optimallösung des obigen Problems nur unter den Vektoren $\frac{1}{\sqrt{p^2+1}} \begin{pmatrix} p \\ Ext \end{pmatrix}$ zu suchen ist, die Lösung sich somit ergibt als:

$$\text{Maximiere (Minimiere)} \quad \frac{1}{\sqrt{p^2+1}} \begin{pmatrix} p \\ Ext \end{pmatrix}$$

Beweis:

Um die Aussage des Satzes zu beweisen, muß die Zielfunktion der beiden Vektoren verglichen werden. Dabei gilt:

$$\begin{aligned} \langle \bar{\gamma}, \overline{Ext} \rangle &= \frac{1}{\sqrt{p^2+1}} \left\langle \begin{pmatrix} -\sqrt{2} \\ \gamma \end{pmatrix}, \begin{pmatrix} p \\ Ext \end{pmatrix} \right\rangle = \frac{1}{\sqrt{p^2+1}} (-\sqrt{2}p + \langle \gamma, Ext \rangle) \\ \langle \bar{\gamma}, \overline{M} \rangle &= \frac{1}{\sqrt{p^2+1}} \left\langle \begin{pmatrix} -\sqrt{2} \\ \gamma \end{pmatrix}, \begin{pmatrix} p \\ M \end{pmatrix} \right\rangle = \frac{1}{\sqrt{p^2+1}} (-\sqrt{2}p + \langle \gamma, M \rangle) \end{aligned}$$

Da es sich bei dem Ausdruck $\langle \gamma, Ext \rangle$ um die Lösung des Optimierungsproblem unter $e_0 = 0$ handelt, ist auch der gesamte Ausdruck mit $\langle \gamma, Ext \rangle$ maximal (bzw. minimal).

□

Betrachten wir zunächst den Fall des Minimierens. Definiere dazu:

$$f(p) := \frac{1}{\sqrt{p^2+1}} (-\sqrt{2}p + \langle \gamma, Min \rangle)$$

Von dieser Funktion ist das Minimum über p zu suchen. Bildet man von $f(p)$ die erste Ableitung und setzt diese gleich Null, so ergibt sich:

$$p_0 = -\frac{\sqrt{2}}{\langle \gamma, Min \rangle} \quad (\langle \gamma, Min \rangle \neq 0)$$

Um nun bestimmen zu können, ob es sich um ein Maximum oder ein Minimum handelt, wird die zweite Ableitung an dieser Stelle ausgewertet. Dies liefert:

$$\text{sgn} \left[f'' \left(-\frac{\sqrt{2}}{\langle \gamma, Min \rangle} \right) \right] = -\text{sgn}[\langle \gamma, Min \rangle]$$

Somit handelt es sich für $\langle \gamma, Min \rangle < 0$ um ein Minimum, für $\langle \gamma, Min \rangle > 0$ um ein Maximum. Betrachtet man noch das Grenzverhalten der Funktion $f(p)$ für $p \rightarrow \pm\infty$, so ergibt alles zusammen:

$$p_0 = \text{argmin} f(p) = \begin{cases} -\frac{\sqrt{2}}{\langle \gamma, Min \rangle} & : \langle \gamma, Min \rangle < 0 \\ \infty & : \langle \gamma, Min \rangle \geq 0 \end{cases}$$

Wir suchen jedoch den Vektor, der das Optimierungsproblem löst. Für ihn gilt:

$$\overline{Min} = \begin{cases} \frac{1}{\sqrt{2+\langle \gamma, Min \rangle^2}} \begin{pmatrix} \sqrt{2} \\ -Min \cdot \langle \gamma, Min \rangle \end{pmatrix} & : \langle \gamma, Min \rangle < 0 \\ (1, 0, \dots, 0)^T & : \langle \gamma, Min \rangle \geq 0 \end{cases}$$

Eine analoge Rechnung ergibt sich auch für das Maximierungsproblem. Diesmal gilt es, die folgende Funktion zu maximieren:

$$f(p) := \frac{1}{\sqrt{p^2 + 1}}(-\sqrt{2}p + \langle \gamma, Max \rangle)$$

Somit gilt weiter:

$$p_0 = \operatorname{argmax} f(p) = \begin{cases} -\frac{\sqrt{2}}{\langle \gamma, Max \rangle} & : \langle \gamma, Max \rangle > 0 \\ -\infty & : \langle \gamma, Max \rangle \leq 0 \end{cases}$$

Und für den maximalen Vektor \overline{Max} gilt:

$$\overline{Max} = \begin{cases} \frac{1}{\sqrt{2 + \langle \gamma, Max \rangle^2}} \begin{pmatrix} -\sqrt{2} \\ Max \cdot \langle \gamma, Max \rangle \end{pmatrix} & : \langle \gamma, Max \rangle > 0 \\ (-1, 0, \dots, 0)^T & : \langle \gamma, Max \rangle \leq 0 \end{cases}$$

Diese allgemeinen Formeln sollen nun für die Fälle $S = 3$ und $S = 4$ explizit durchgerechnet werden.

Fall $S = 3$:

Betrachten wir zuerst den Fall des Minimierens. Wie aus den vorangehenden Überlegungen ersichtlich ist, muß das Skalarprodukt $\langle \gamma, Min \rangle$ berechnet werden. Dabei hat γ die folgende Gestalt:

$$\gamma = \frac{1}{\sqrt{2}} \begin{pmatrix} -2\lambda \\ 2 - 4\lambda \\ 6 - 6\lambda \end{pmatrix}$$

Weiter sind dann die beiden Fälle $\lambda \in (0, \frac{5}{3}]$ und $\lambda \in (\frac{5}{3}, \infty)$ zu unterscheiden, da sich dort das Minimum anders berechnet. Sei zuerst $\lambda \in (0, \frac{5}{3}]$. Damit ist

$$Min = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}$$

und das Skalarprodukt ist gegeben durch:

$$\langle \gamma, Min \rangle = \frac{1}{2} \left(-2\lambda \cdot (-1) + (2 - 4\lambda) \cdot 1 + (6 - 6\lambda) \cdot 0 \right) = 1 - \lambda$$

Im Fall $\lambda \in (\frac{5}{3}, \infty)$ hat das Minimum die folgende Gestalt

$$Min = \arctan \left(\frac{5 - 3\lambda}{\sqrt{3}(1 - \lambda)} \right)$$

und für das Skalarprodukt gilt:

$$\langle \gamma, Min \rangle = -2\sqrt{\frac{3\lambda^2 - 9\lambda + 7}{3}} < 0$$

Als nächstes betrachten wir das Maximum:

Erneut benötigen wir eine Fallunterscheidung. Diesmal muß jedoch unterschieden werden, ob $\lambda \in (0, 2)$ oder $\lambda \in [2, \infty)$.

Sei zunächst $\lambda \in (0, 2)$ und somit

$$Max = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$$

Dann gilt für das Skalarprodukt:

$$\langle \gamma, Max \rangle = \frac{1}{2} \left(-2\lambda \cdot (-1) + (2 - 4\lambda) \cdot 0 + (6 - 6\lambda) \cdot 1 \right) = 3 - 2\lambda$$

Im Fall $\lambda \in [2, \infty)$ gilt für das Maximum:

$$Max = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}$$

Damit berechnet sich das Skalarprodukt wie folgt:

$$\langle \gamma, Max \rangle = \frac{1}{2} \left(-2\lambda \cdot (-1) + (2 - 4\lambda) \cdot 1 + (6 - 6\lambda) \cdot 0 \right) = 1 - \lambda$$

Insgesamt ergibt sich für den Fall $S = 3$ die nachfolgende Tabelle:

λ	$(0, 1]$	$(1, \frac{3}{2})$
\overline{Min}	$(1, 0, 0, 0)^T$	$\frac{1}{\sqrt{2+(1-\lambda)^2}} \left(\sqrt{2}, \frac{1-\lambda}{\sqrt{2}}, \frac{\lambda-1}{\sqrt{2}}, 0 \right)^T$
\overline{Max}	$\frac{1}{\sqrt{2+(3-2\lambda)^2}} \left(-\sqrt{2}, \frac{2\lambda-3}{\sqrt{2}}, 0, \frac{3-2\lambda}{\sqrt{2}} \right)^T$	$\frac{1}{\sqrt{2+(3-2\lambda)^2}} \left(-\sqrt{2}, \frac{2\lambda-3}{\sqrt{2}}, 0, \frac{3-2\lambda}{\sqrt{2}} \right)^T$
λ	$[\frac{3}{2}, \frac{5}{3}]$	$[\frac{5}{3}, \infty)$
\overline{Min}	$\frac{1}{\sqrt{2+(1-\lambda)^2}} \left(\sqrt{2}, \frac{1-\lambda}{\sqrt{2}}, \frac{\lambda-1}{\sqrt{2}}, 0 \right)^T$	$\frac{1}{\sqrt{2\lambda^2-6\lambda+\frac{17}{3}}} \left(1, \frac{4-3\lambda}{3}, \frac{1}{3}, \frac{3\lambda-5}{3} \right)^T$
\overline{Max}	$(-1, 0, 0, 0)^T$	$(-1, 0, 0, 0)^T$

Fall $S = 4$:

In diesem Fall hat der Vektor γ die folgende Gestalt:

$$\gamma = \frac{1}{\sqrt{2}} \begin{pmatrix} -2\lambda \\ 2 - 4\lambda \\ 6 - 6\lambda \\ 12 - 8\lambda \end{pmatrix}$$

Auch hier soll zunächst der Fall des Minimierens betrachtet werden. Somit benötigen wir, erneut mit einer Fallunterscheidung nach λ , das Skalarprodukt $\langle \gamma, Min \rangle$.

Betrachten wir zunächst den Fall $\lambda \in (0, \frac{7}{3}]$. Aus der Minimallösung für $S = 4$ sieht man, daß für diese λ -Werte der Winkel θ den Wert 90° annimmt. Dies hat zur Folge, daß die letzte Komponente des Lösungsvektors Null ist, wodurch man \overline{Min} einfach dadurch erhält, daß man zur entsprechenden Lösung aus dem Fall $S = 3$ eine Nullkomponente hinzufügt.

Betrachten wir als nächstes den Fall $\lambda \in [\frac{7}{3}, 3]$. Mittels MAPLE wurde hierfür das Skalarprodukt $\langle \gamma, Min \rangle$ berechnet. Dabei ergab sich der folgende Ausdruck:

$$\langle \gamma, Min \rangle = -\sqrt{2(5\lambda^2 - 20\lambda + 21)}$$

Man sieht sofort, daß dieser Ausdruck stets negativ ist. Als nächstes wird der Wert des Minimums für $e_0 = 0$ aus der Kugeldarstellung umgerechnet in:

$$Min = \frac{1}{2\sqrt{5\lambda^2 - 20\lambda + 21}} \begin{pmatrix} 3\lambda - 5 \\ 3 - \lambda \\ \lambda - 1 \\ 3\lambda - 7 \end{pmatrix}$$

Dieser Ausdruck wird später noch zur Berechnung von \overline{Min} benötigt.

Bleibt noch der Fall $\lambda > 3$ zu betrachten. In diesem Fall hat man:

$$\langle \gamma, Min \rangle = -\frac{2}{3}\sqrt{3(7\lambda^2 - 27\lambda + 27)}$$

Auch dieser Ausdruck ist wieder negativ. Daher berechnen wir auch hier das Minimum für $e_0 = 0$ aus der Kugeldarstellung:

$$Min = \frac{1}{\sqrt{6(7\lambda^2 - 27\lambda + 27)}} \begin{pmatrix} 5\lambda - 9 \\ 0 \\ \lambda \\ 4\lambda - 9 \end{pmatrix}$$

Somit lassen sich für das Minimierungsproblem alle Richtungsvektoren bestimmen und wir können uns nun dem Maximierungsproblem zuwenden. Wie sich zeigen wird, ist dieses um einiges einfacher zu handhaben.

Auch hier ist natürlich wieder eine Fallunterscheidung nach λ durchzuführen. Jedoch sind dabei nur die beiden Fälle $\lambda \in (0, \frac{5}{2}]$ und $\lambda \in [\frac{5}{2}, \infty)$ zu unterscheiden.

Sei zunächst $\lambda \in (0, \frac{5}{2}]$. Dann ist das Maximum gegeben durch:

$$Max = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

Dies liefert weiter:

$$\langle \gamma, Max \rangle = \frac{1}{2} \left(2\lambda + (12 - 8\lambda) \right) = 6 - 3\lambda$$

Somit erhält man für $\lambda \in [2, \frac{5}{2}]$:

$$\overline{Max} = (-1, 0, 0, 0, 0)^T$$

Im Fall $\lambda \in [0, 2)$ ist \overline{Max} gegeben durch den Ausdruck:

$$\overline{Max} = \frac{1}{\sqrt{2 + (6 - 3\lambda)^2}} \begin{pmatrix} -\sqrt{2} \\ \frac{1}{\sqrt{2}}(3\lambda - 6) \\ 0 \\ 0 \\ \frac{1}{\sqrt{2}}(6 - 3\lambda) \end{pmatrix}$$

Bleibt noch $\lambda \in [\frac{5}{2}, \infty)$ zu betrachten. Allerdings hat hier wiederum Max in der letzten Komponente den Wert Null und \overline{Max} ergibt sich aus dem Wert von $S = 3$ durch Anfügen einer Nullkomponente.

Somit sind alle Fälle für $S = 4$ abgehandelt und es ergibt sich letztendlich folgende Tabelle:

λ	$(0, 1]$	$(1, \frac{5}{3}]$
\overline{Min}	$(1, 0, 0, 0, 0)^T$	$\frac{1}{\sqrt{2+(1-\lambda)^2}} \left(\sqrt{2}, \frac{1-\lambda}{\sqrt{2}}, \frac{\lambda-1}{\sqrt{2}}, 0, 0 \right)^T$
\overline{Max}	$\frac{1}{\sqrt{2+(6-3\lambda)^2}} \left(-\sqrt{2}, \frac{3\lambda-6}{\sqrt{2}}, 0, 0, \frac{6-3\lambda}{\sqrt{2}} \right)^T$	$\frac{1}{\sqrt{2+(6-3\lambda)^2}} \left(-\sqrt{2}, \frac{3\lambda-6}{\sqrt{2}}, 0, 0, \frac{6-3\lambda}{\sqrt{2}} \right)^T$
λ	$(\frac{5}{3}, 2)$	$[2, \frac{7}{3}]$
\overline{Min}	$\frac{1}{\sqrt{2\lambda^2-6\lambda+\frac{17}{3}}} \left(1, \frac{4-3\lambda}{3}, \frac{1}{3}, \frac{3\lambda-5}{3}, 0 \right)^T$	$\frac{1}{\sqrt{2\lambda^2-6\lambda+\frac{17}{3}}} \left(1, \frac{4-3\lambda}{3}, \frac{1}{3}, \frac{3\lambda-5}{3}, 0 \right)^T$
\overline{Max}	$\frac{1}{\sqrt{2+(6-3\lambda)^2}} \left(-\sqrt{2}, \frac{3\lambda-6}{\sqrt{2}}, 0, 0, \frac{6-3\lambda}{\sqrt{2}} \right)^T$	$(-1, 0, 0, 0, 0)^T$

λ	$[\frac{7}{3}, 3]$	$[3, \infty)$
\overline{Min}	$\frac{1}{2\sqrt{5\lambda^2-20\lambda+22}} \left(2, 5 - 3\lambda, 3 - \lambda, \lambda - 1, 3\lambda - 7 \right)^T$	$\frac{1}{\sqrt{42\lambda^2-162\lambda+171}} \left(3, 9 - 5\lambda, 0, \lambda, 4\lambda - 9 \right)^T$
\overline{Max}	$\left(-1, 0, 0, 0, 0 \right)^T$	$\left(-1, 0, 0, 0, 0 \right)^T$

An dieser Tabelle ist erneut eine Ähnlichkeit von Minimum für $\lambda < 1$ und Maximum für $\lambda > 2$ festzustellen. Hierbei ist lediglich das Vorzeichen der ersten Komponente vertauscht. Weitere Ergebnisse sind im Vergleich in Kapitel 2.3.3 dargestellt.

2.3.2 Simulation der Verteilung des Tests mittels Kegelapproximation unter der lokalen Alternative

In Kapitel 2.1.2 wurde bereits gezeigt, daß der Test mittels Kegelapproximationen die folgende Verteilung unter der lokalen Alternative besitzt:

$$T_\infty^n \xrightarrow{\mathcal{D}} \|\pi_D(Z + \mu|V) - \pi_D(Z + \mu|C_q)\|_D^2$$

wo

$$\mu = (\Psi^S)'_{(\lambda, 1, 0, \dots, 0)} \begin{pmatrix} h_0 \\ \mathbf{h} \end{pmatrix}$$

Diesen Ausdruck gilt es nun für die im vorangehenden Abschnitt bestimmten Extremalrichtungen zu berechnen. Da es sich bei den zu projizierenden Vektoren nicht mehr um zentrierte Normalverteilungen handelt, kann der Algorithmus, der im Fall der Hypothese verwendet worden ist, nicht ohne weiteres übernommen werden.

Jedoch kann obiger Ausdruck analog zum Hypothesenfall wie folgt umgewandelt werden:

$$\begin{aligned} & \|\pi_D(Z + \mu|V) - \pi_D(Z + \mu|C_q)\|_D^2 = \\ & = \|T^{-1}\pi_{T^{-T}DT^{-1}}(TZ + T\mu|TV) - T^{-1}\pi_{T^{-T}DT^{-1}}(TZ + T\mu|TC_q)\|_D^2 = \\ & = \|\pi_{\tilde{D}}(\tilde{Z} + T\mu|\tilde{V}) - \pi_{\tilde{D}}(\tilde{Z} + T\mu|\tilde{C})\|_{\tilde{D}}^2 \end{aligned} \quad (2.22)$$

wo

$$\begin{aligned} T & := (\text{id}_{(S+1) \times (S+1)}, 0) \\ \tilde{Z} & := TZ \sim N_{S+1}(\mathbf{0}, \tilde{D}^{-1}) \\ \tilde{D} & := T^{-T}DT^{-1} \\ \tilde{C} & := TC_q = \mathcal{C}(\nu_1, \dots, \nu_S)^{* \tilde{D}} \\ & \quad \nu_i \text{ die Spalten von } \mathcal{N} := \tilde{D}^{-1}A^{-T}N \end{aligned}$$

Die hierin vorkommenden Matrizen A und N sind wie folgt definiert:

$$N := \begin{pmatrix} 0 & & & & 0 \\ 1 & & & & 0 \\ 1 & & & & \\ \vdots & \text{diag}(-1, \dots, -1)_{(S-1) \times (S-1)} & & & \\ 1 & & & & \end{pmatrix} \in \mathbb{R}^{(S+1) \times S}$$

$$A := \begin{pmatrix} -e^{-\lambda} & 0 & & \dots & 0 \\ e^{-\lambda} \lambda \left(\frac{1}{\lambda} - 1\right) & \lambda e^{-\lambda} & \ddots & & 0 \\ e^{-\lambda} \frac{\lambda^2}{2!} \left(\frac{2}{\lambda} - 1\right) & \lambda^2 e^{-\lambda} & \ddots & & 0 \\ e^{-\lambda} \frac{\lambda^3}{3!} \left(\frac{3}{\lambda} - 1\right) & \frac{\lambda^3}{2!} e^{-\lambda} & \ddots & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & 0 \\ e^{-\lambda} \frac{\lambda^S}{S!} \left(\frac{S}{\lambda} - 1\right) & \frac{\lambda^S}{(S-1)!} e^{-\lambda} & \frac{\lambda^{S-1}}{(S-2)!} e^{-\lambda} & \dots & \lambda e^{-\lambda} \end{pmatrix} \in \mathbb{R}^{(S+1) \times (S+1)}$$

Als nächstes soll nun versucht werden, die Differenz in diesem Ausdruck zu einem Term zusammenzufassen. Dazu wird eine Aussage über die Linearität bei Kegelprojektionen unter gewissen Voraussetzungen benötigt. Um dies zu zeigen, wird zuerst der folgende Satz benötigt:

Satz 2.8 *Sei C ein beliebiger Kegel und $V \subset C$ ein beliebiger Vektorraum. Dann gilt: Falls $y \in V^\perp$, dann ist $\pi(y|C) \in V^\perp$*

Beweis:

Wenn $v \in V$, dann ist aufgrund der Definition eines Vektorraums auch $-v \in V$. Somit liegen beide Vektoren auch in C , und es gilt:

$$\begin{aligned}
 0 &\geq \langle v - \pi(y|C), y - \pi(y|C) \rangle = \langle v, y - \pi(y|C) \rangle - \langle \pi(y|C), y - \pi(y|C) \rangle = \\
 &= \langle v, y - \pi(y|C) \rangle - \underbrace{\langle \pi(y|C), \pi(y|C^*) \rangle}_{=0, \text{ da } \perp \text{ (vgl. [Bu01])}} \\
 0 &\geq \langle -v - \pi(y|C), y - \pi(y|C) \rangle = \langle -v, y - \pi(y|C) \rangle - \langle \pi(y|C), y - \pi(y|C) \rangle = \\
 &= \langle -v, y - \pi(y|C) \rangle - \underbrace{\langle \pi(y|C), \pi(y|C^*) \rangle}_{=0, \text{ da } \perp \text{ (vgl. [Bu01])}}
 \end{aligned}$$

Somit gilt:

$$0 = \langle v, y - \pi(y|C) \rangle = \underbrace{\langle v, y \rangle}_{=0} - \langle v, \pi(y|C) \rangle$$

Somit ist $\pi(y|C)$ senkrecht zu allen Vektoren $v \in V$ und es gilt die Behauptung. □

Mit Hilfe dieses Satzes kann nun der Satz, der Aufschluß über die Linearität bei Kegelprojektionen gibt, bewiesen werden.

Satz 2.9 Sei $V \subset C$ ein Vektorraum. Dann gilt:

$$\pi(x|C) = \pi(\pi(x|V)|C) + \pi(\pi(X|V^\perp)|C)$$

Beweis:

Um den Satz zu beweisen, muß gezeigt werden, daß:

$$\langle c - \pi(\pi(x|V)|C) - \pi(\pi(X|V^\perp)|C), x - \pi(\pi(x|V)|C) - \pi(\pi(X|V^\perp)|C) \rangle \leq 0 \quad \forall c \in C$$

Dazu formulieren wir den Ausdruck um und erhalten:

$$\begin{aligned} & \langle c - \pi(\pi(x|V)|C) - \pi(\pi(X|V^\perp)|C), x - \underbrace{\pi(\pi(x|V)|C)}_{=\pi(x|V)} - \pi(\pi(X|V^\perp)|C) \rangle = \\ & \quad \underbrace{\langle c, \pi(x|V^\perp) - \pi(\pi(X|V^\perp)|C) \rangle}_{=\pi(x|V^\perp)} - \langle \underbrace{\pi(\pi(x|V)|C) - \pi(\pi(X|V^\perp)|C)}_{=\pi(x|V^\perp)}, \underbrace{\pi(x|V^\perp) - \pi(\pi(X|V^\perp)|C)}_{=\pi(\pi(x|V^\perp)|C^*)} \rangle = \\ & \quad \underbrace{\langle c, \pi(x|V^\perp) - \pi(\pi(X|V^\perp)|C) \rangle}_{=0} - \langle \underbrace{\pi(\pi(X|V^\perp)|C)}_{=\pi(X|V^\perp) - \pi(\pi(X|V^\perp)|C)}, \underbrace{\pi(\pi(X|V^\perp)|C^*)}_{=\pi(\pi(X|V^\perp)|C)} \rangle = \\ & \quad \langle c - \pi(\pi(X|V^\perp)|C), \pi(x|V^\perp) - \pi(\pi(X|V^\perp)|C) \rangle = \\ & \quad \langle c - \pi(p|C), p - \pi(p|C) \rangle \quad \text{wo} \quad p := \pi(X|V^\perp) \end{aligned}$$

Bei dem letzten Ausdruck der obigen Umformulierungen handelt es sich genau um die Bedingung (NSC), so daß dieser Ausdruck kleiner gleich Null sein muß und der Satz bewiesen ist. □

Nun kann die Gleichung (2.22) weiter umgeformt werden. Mit der Definition

$$\tilde{X} := \tilde{Z} + \tilde{\mu} := \tilde{Z} + T\mu$$

gilt:

$$\begin{aligned} & \|\pi_{\tilde{D}}(\tilde{Z} + T\mu|\tilde{V}) - \pi_{\tilde{D}}(\tilde{Z} + T\mu|\tilde{C})\|_{\tilde{D}}^2 = \\ & = \|\pi_{\tilde{D}}(\tilde{X}|\tilde{V}) - \pi_{\tilde{D}}(\tilde{X}|\tilde{C})\|_{\tilde{D}}^2 = \\ & = \|\pi_{\tilde{D}}(\pi_{\tilde{D}}(\tilde{X}|\tilde{V}) + \pi_{\tilde{D}}(\tilde{X}|\tilde{V}^\perp)|\tilde{V}) - \pi_{\tilde{D}}(\pi_{\tilde{D}}(\tilde{X}|\tilde{V}) + \pi_{\tilde{D}}(\tilde{X}|\tilde{V}^\perp)|\tilde{C})\|_{\tilde{D}}^2 = \\ & = \|\underbrace{\pi_{\tilde{D}}(\pi_{\tilde{D}}(\tilde{X}|\tilde{V})|\tilde{V})}_{=\pi_{\tilde{D}}(\tilde{X}|\tilde{V})} + \underbrace{\pi_{\tilde{D}}(\pi_{\tilde{D}}(\tilde{X}|\tilde{V}^\perp)|\tilde{V})}_{=0} - \underbrace{\pi_{\tilde{D}}(\pi_{\tilde{D}}(\tilde{X}|\tilde{V})|\tilde{C})}_{=\pi_{\tilde{D}}(\tilde{X}|\tilde{V})} - \pi_{\tilde{D}}(\pi_{\tilde{D}}(\tilde{X}|\tilde{V}^\perp)|\tilde{C})\|_{\tilde{D}}^2 = \\ & = \|\pi_{\tilde{D}}(\pi_{\tilde{D}}(\tilde{X}|\tilde{V}^\perp)|\tilde{C})\|_{\tilde{D}}^2 \end{aligned}$$

Zur weiteren Rechnung wird die Matrix \overline{Q} eingeführt, die wie folgt definiert ist:

$$\overline{Q} := N^T A^{-1} \tilde{D}^{-1} A^{-T} N$$

Um den vorangehenden Ausdruck nun weiter vereinfachen zu können, wird die Umkehrabbildung der durch die Matrix N erzeugten Abbildung benötigt. Da die Matrix N nicht quadratisch ist, kann dies nicht einfach mittels der Inversen geschehen. Betrachtet man jedoch die durch N erzeugte Abbildung genauer, so sieht man, daß, obwohl die Abbildung von \mathbb{R}^S nach \mathbb{R}^{S+1} geht, keine Dimensionsänderung vorliegt, da die Bildmenge von N nur ein S -dimensionaler Teilraum des \mathbb{R}^{S+1} ist. Somit kann aufgrund der speziellen Struktur der Matrix N eine eindeutige Umkehrabbildung definiert werden. Diese sei mit N^{-1} bezeichnet, wodurch sich der oben berechnete Ausdruck weiter umformen läßt.

$$\begin{aligned}
 & \left\| \pi_{\tilde{D}} \left(\pi_{\tilde{D}}(\tilde{X} | \tilde{V}^\perp) \middle| \tilde{C} \right) \right\|_{\tilde{D}}^2 = \\
 &= \left\| \pi_{A^{-T} N \bar{Q} \bar{Q}^{-1} N^T A^{-1}} \left((N^T A^{-1})^{-1} \bar{Q} \bar{Q}^{-1} N^T A^{-1} \pi_{\tilde{D}}(\tilde{X} | \tilde{V}^\perp) \middle| (N^T A^{-1})^{-1} \bar{Q} \bar{Q}^{-1} N^T A^{-1} \tilde{C} \right) \right\|_{\tilde{D}}^2 = \\
 &= \left\| \pi_{\bar{Q}} \left(\bar{Q}^{-1} N^T A^{-1} \pi_{\tilde{D}}(\tilde{X} | \tilde{V}^\perp) \middle| \underbrace{\bar{Q} N^T A^{-1} \tilde{C}}_{=: \bar{C}} \right) \right\|_{\bar{Q}}^2 = \\
 &= \left\| \pi_{\bar{Q}} \left(\bar{Q}^{-1} N^T A^{-1} (\tilde{X} - A e_0 (e_0^T A^T A e_0)^{-1} e_0^T A^T \tilde{D} \tilde{X}) \middle| \bar{C} \right) \right\|_{\bar{Q}}^2 = \\
 &= \left\| \pi_{\bar{Q}} \left(\bar{Q}^{-1} N^T A^{-1} \tilde{X} - \underbrace{\bar{Q}^{-1} N^T A^{-1} A e_0 (e_0^T A^T A e_0)^{-1} e_0^T A^T \tilde{D} \tilde{X}}_{=I} \middle| \bar{C} \right) \right\|_{\bar{Q}}^2 = \\
 &= \left\| \pi_{\bar{Q}} \left(\bar{Q}^{-1} N^T A^{-1} \tilde{X} \middle| \bar{C} \right) \right\|_{\bar{Q}}^2 = \\
 &= \|\pi_{\bar{Q}}(\bar{X} | \bar{C})\|_{\bar{Q}}^2 = \pi_{\bar{Q}}(\bar{X} | \bar{C})^T \cdot \bar{Q} \cdot \pi_{\bar{Q}}(\bar{X} | \bar{C}) = (\bar{X} - \pi_{\bar{Q}}(\bar{X} | \bar{C}^* \bar{Q}))^T \cdot \bar{Q} \cdot \pi_{\bar{Q}}(\bar{X} | \bar{C}) = \\
 &= \bar{X}^T \cdot \bar{Q} \cdot \pi_{\bar{Q}}(\bar{X} | \bar{C})
 \end{aligned}$$

wo

$$\begin{aligned}
 \bar{X} &:= \bar{Q}^{-1} N^T A^{-1} \tilde{X} = \bar{Z} + \bar{Q}^{-1} N^T A^{-1} T \mu \\
 \bar{Z} &:= \bar{Q}^{-1} N^T A^{-1} \tilde{Z}
 \end{aligned}$$

Für den Kegel \bar{C} gilt:

$$\begin{aligned}
 \bar{C} &= \bar{Q}^{-1} N^T A^{-1} A \cdot \{x \in \mathbb{R}^{S+1} : N^T x \leq 0\} = \bar{Q}^{-1} N^T \cdot \{x \in \mathbb{R}^{S+1} : N^T x \leq 0\} = \\
 &= \bar{Q}^{-1} \cdot \{x \in \mathbb{R}^S : x \leq 0\} = \{x \in \mathbb{R}^S : I \cdot \bar{Q} \cdot x \leq 0\} \\
 \Rightarrow \bar{C} &= \mathcal{C}(e_1, \dots, e_s)^* \bar{Q}
 \end{aligned}$$

Dabei bezeichne C^* den dualen Kegel von C . Dieser besteht aus allen Vektoren, die zwischen sich selbst mit allen Vektoren aus C einen stumpfen Winkel besitzen, d.h. es gilt:

$$C^* := \{x \in V : \langle x, c \rangle \leq 0 \quad \forall c \in C\}$$

Die Matrix \bar{Q} hinter dem Stern bedeutet, wie bereits schon zuvor erwähnt, daß das innere Produkt bezüglich der Matrix \bar{Q} zu bilden ist.

Bleibt also noch die Projektion von \bar{X} auf den Kegel \bar{C} zu bestimmen. Dazu wird mittels des Algorithmus von [Bu01] die Indexmenge I bestimmt, die angibt, wo der Vektor relativ zum Kegel liegt. Dies bedeutet, daß man jenes I benötigt, welches die nachfolgende Bedingung erfüllt (dabei bezeichne E_I die $\mathbb{R}^{S \times \#I}$ -Matrix mit den Spalten $e_i (i \in I)$, wo e_i den i -ten Einheitsvektor bezeichnet).

$$\begin{aligned} \bar{X} &\in \mathcal{C}(e_i | i \in I) \oplus \left(\text{lin}\{e_i : i \in I\}^{\perp \bar{Q}} \cap \mathcal{C}^{*\bar{Q}}(e_i | i \in I^C) \right) \\ \Leftrightarrow \quad &\pi_{\bar{Q}}(\bar{X} | \text{lin}\{e_i, i \in I\}) \in \mathcal{C}(e_i | i \in I) \\ &\wedge \pi_{\bar{Q}}(\bar{X} | \text{lin}\{e_i, i \in I\}^{\perp \bar{Q}}) \in \mathcal{C}^{*\bar{Q}}(e_i | i \in I^C) \\ \Leftrightarrow \quad &(E_I^T \bar{Q} E_I)^{-1} E_I^T \bar{Q} \cdot \bar{X} \geq 0 \\ &\wedge E_{I^C}^T \bar{Q} \cdot (\bar{X} - E_I (E_I^T \bar{Q} E_I)^{-1} E_I^T \bar{Q} \cdot \bar{X}) \leq 0 \end{aligned}$$

Somit erhält man denselben Ausdruck wie in [Bu01], lediglich der Vektor \bar{Z} muß durch den Vektor \bar{X} ersetzt werden. Um nun die Projektion berechnen zu können, muß nur noch die folgende Gleichung angewendet werden:

$$\pi_{\bar{Q}}(\bar{X} | \bar{C}) = \pi_{\bar{Q}}(\bar{X} | \text{lin}\{e_i, i \in I\}^{\perp \bar{Q}}) = \bar{X} - E_I (E_I^T \bar{Q} E_I)^{-1} E_I^T \bar{Q} \cdot \bar{X}$$

Diese erste Gleichheit ist dabei eine Folgerung, die man aus der Indexmenge I gewinnt. Da man mittels I weiß, wo \bar{X} relativ zum Kegel liegt, kann man die Kegelpjektion durch eine Projektion auf einen linearen Teilraum ersetzen. Somit erhält man, wenn man alle Schritte zusammenfaßt, die folgende Gleichung:

$$\|\pi_D(Z + \mu | V) - \pi_D(Z + \mu | C_q)\|_D^2 = \bar{X}^T (\bar{Q} - \bar{Q} E_I (E_I^T \bar{Q} E_I)^{-1} E_I^T \bar{Q}) \bar{X}$$

2.3.3 Vergleich der beiden Tests

In diesem Abschnitt sollen nun die Ergebnisse des Vergleichs zusammengefaßt werden. Wie schon in den vorangehenden Abschnitten erläutert, wurde dabei wie folgt vorgegangen:

Zu den in Kapitel 2.3.1 bestimmten Richtungen wurde der Fehler 2. Art bestimmt. Dazu verwendete man die Grenzverteilung der lokalen Alternative, die sich für $n \rightarrow \infty$ ergab. Somit erhielt man für den Fall des Dispersionstests den maximalen und den minimalen Fehler 2. Art, je nachdem aus welcher Richtung man sich der Hypothese nähert. Dann wurde der Fehler 2. Art für den Test mittels Kegelpjektionen bestimmt. In diesem Fall waren dies zwar nicht mehr die Maxima bzw. Minima, jedoch konnte auf diesem Wege immerhin ein Vergleich erzielt werden. In nachfolgenden Tabellen sind die Ergebnisse zusammengefaßt. Dabei wurde die Rechnung für die Fälle $S = 3$ und $S = 4$ durchgeführt, unterschieden nach dem 95% und 99% Quantil. Die Auswertung geschah an den folgenden Werten: $\lambda \in \{\frac{1}{2}, \frac{3}{4}, 1, \frac{3}{2}, 2, 3\}$

In der Tabelle bedeuten leere Felder, daß der Fehler 2. Art den Wert 99,99% übersteigt. Des weiteren steht α_{Max} für den minimalen Fehler 2. Art (da er sich aus dem Maximum der Richtungsableitung ergibt), α_{Min} für den maximalen Fehler 2. Art. In den Zeilen β_{Max} bzw. β_{Min} stehen die dazugehörigen Fehler 2. Art des Tests mittels Kegelpjektionen (diese sind aber,

wie oben beschrieben, nicht die Extremwerte). Die Ergebnisse für den Test mittels Kegelapproximationen beruhen auf dem Simulationsalgorithmus, der im Anhang zu finden ist. Dabei wurden 100.000 Simulationen gerechnet.

Betrachtet man sich die Ergebnisse genauer, so sieht man, daß der Test mittels Kegelapproximationen in den meisten Fällen besser abschneidet. So ist der Dispersionstest nur in den Fällen besser, in denen h_0 beliebig gewählt wurde und gleichzeitig λ sehr groß ist, und dann auch nur im scharf trennenden Fall. In den anderen Fällen zeigt jedoch der Test mittels Kegelapproximationen ein besseres Ergebnis. Gerade im scharf trennenden Fall ist dies jedoch erstaunlich, da es sich beim Dispersionstest um ein Extremum handelt, der Kegelapproximationstest jedoch noch nicht zwingend seinen besten Wert erreicht. Im Falle der lokalen Alternativen, bei denen man sich aus Richtungen nähert, die den (für den Dispersionstest) am schwächsten trennenden Test liefern, ist in allen beobachteten Fällen der Kegelapproximationstest besser. Allerdings bleibt hier die Frage, ob es schon der schlechteste Fall für diesen Test ist, oder ob es noch Richtungen gibt, die noch schwächere Ergebnisse liefern.

Bemerkenswert ist noch das gegenläufige Verhalten im schwach trennenden Fall. Hier steigt der Fehler 2. Art beim Dispersionstest mit wachsendem λ , während er beim Kegelapproximationstest fällt. Im schwach trennenden Fall hingegen ist bei beiden Tests ein Wachsen des Fehlers 2. Art mit wachsendem λ festzustellen.

Betrachtet man noch die Aussagen, die bereits weiter zuvor in diesem Kapitel gezeigt wurden, so hat man dort gesehen, daß das Minimum bei kleinen λ -Werten für solche Richtungen angenommen wird, bei denen sich die Masse auf die ersten Komponenten konzentriert. Für $\lambda \leq 1$ konnte sogar gezeigt werden, daß die gesamte Masse der Richtung, aus der man sich nähert, auf die ersten beiden Komponenten konzentriert ist. Betrachtet man sich nun die nachfolgenden Tabellen, so erkennt man, daß in diesen Fällen der Test mittels Kegelapproximation (oft sogar weit) überlegen ist.

An dieser Stelle sei noch auf ein Ergebnis aus [Po66] hingewiesen. Darin wird gezeigt, daß der Dispersionstest unter allen Tests ohne (lokalen) Bias gegen die Alternative einer Negativen Binomialverteilung der lokal beste Test ist. Dies spielt für unseren Fall insofern eine Rolle, da die Negative Binomialverteilung und die zusammengesetzte Poisson-Verteilung eine identische Verteilung liefern, falls folgende Bedingungen erfüllt sind:

$$\lambda = -r \log(q)$$

$$p_k = \frac{r(1-q)^k}{\lambda k}$$

Dabei stehen r und q für die Parameter der Negativen Binomialverteilung $NB(r, q)$, die Parameter λ und p_k seien wie gewohnt die Parameter der zusammengesetzten Poisson-Verteilung. Diese Aussage läßt sich durch Koeffizientenvergleich der erzeugenden Funktionen beweisen. Diese lautet für die Negative Binomialverteilung:

$$g_1(t) = \left(\frac{q}{1 - (1-q)t} \right)^r$$

Für die zusammengesetzte Poisson-Verteilung ergibt sich:

$$g_2(t) = \exp \left(\lambda \left(\sum_{k=1}^{\infty} p_k t^k - 1 \right) \right)$$

Diese beiden Ausdrücke setzt man nun noch gleich und erhält:

$$\begin{aligned}
 g_1(t) &\stackrel{!}{=} g_2(t) \\
 \left(\frac{q}{1 - (1 - q)t} \right)^r &= \exp \left(\lambda \left(\sum_{k=1}^{\infty} p_k t^k - 1 \right) \right) \\
 r \log(q) - r \log(1 - (1 - q)t) &= \lambda \left(\sum_{k=1}^{\infty} p_k t^k - 1 \right) \\
 r \log(q) + r \sum_{k=1}^{\infty} \frac{(1 - q)^k}{k} t^k &= -\lambda + \lambda \sum_{k=1}^{\infty} p_k t^k
 \end{aligned}$$

Der oben bereits angesprochene Koeffizientenvergleich liefert sodann das bereits genannte Ergebnis. In der Rechnung wurde dabei der Logarithmus in eine Potenzreihe umgewandelt. Dies gilt aber nur für den Fall $-\frac{1}{1-q} \leq t \leq \frac{1}{1-q}$. Allerdings stellt dies kein Problem dar, da sowieso nur Werte von t nahe der Null von Interesse sind. Ist also eine solche Darstellung möglich, ist der Dispersionstest auch dem Test mittels Kegelapproximationen überlegen.

95%-Quantil $S = 3, h_0$ beliebig

λ	0,5	0,75	1	1,5	2	3
α_{Max}	21,1%	33,9%	46,5%	59,1%	59,1%	59,1%
β_{Max}	12,3%	28,5%	50,4%	90,1%	89,9%	89,6%
α_{Min}	99,9%	99,9%	99,9%	99,9%	99,97%	
β_{Min}	91,4%	90,9%	90,5%	85,6%	70,2%	66,1%

 $S = 3, h_0 = 0$

λ	0,5	0,75	1	1,5	2	3
α_{Max}	36,1%	55,8%	74,0%	95,0%	99,6%	
β_{Max}	3,8%	8,7%	14,5%	28,1%	74,8%	78,9%
α_{Min}	87,4%	91,9%	95,0%	98,4%	99,7%	
β_{Min}	73,8%	72,7%	72,2%	73,0%	52,1%	62,6%

 $S = 4, h_0$ beliebig

λ	0,5	0,75	1	1,5	2	3
α_{Max}	0,1%	0,9%	4,8%	33,9%	59,1%	59,1%
β_{Max}	0%	0%	0%	1,3%	83,4%	83,4%
α_{Min}	99,9%	99,9%	99,9%	99,9%	99,97%	
β_{Min}	85,3%	84,4%	83,9%	78,1%	53,2%	17,6%

 $S = 4, h_0 = 0$

λ	0,5	0,75	1	1,5	2	3
α_{Max}	0,2%	1,8%	7,8%	55,8%	95,0%	
β_{Max}	0%	0%	0%	0,1%	0,6%	64,4%
α_{Min}	87,4%	91,9%	95,0%	98,4%	99,7%	
β_{Min}	67,1%	69,8%	63,4%	62,5%	30,4%	13,6%

99%-Quantil

$S = 3, h_0$ beliebig

λ	0,5	0,75	1	1,5	2	3
α_{Max}	45,6%	60,4%	72,3%	81,9%	81,9%	81,9%
β_{Max}	40,0%	55,1%	76,6%	97,8%	97,7%	97,6%
α_{Min}						
β_{Min}	98,2%	98,1%	97,4%	96,4%	89,4%	87,4%

$S = 3, h_0 = 0$

λ	0,5	0,75	1	1,5	2	3
α_{Max}	62,7%	79,6%	90,7%	99%	99,6%	
β_{Max}	13,7%	25,2%	36,0%	54,9%	91,8%	93,6%
α_{Min}	96,5%	98,1%	99%	99,8%	99,98%	
β_{Min}	91,1%	90,6%	90,4%	90,8%	78,3%	85,4%

$S = 4, h_0$ beliebig

λ	0,5	0,75	1	1,5	2	3
α_{Max}	0,8%	4,7%	16,1%	60,4%	81,9%	81,9%
β_{Max}	0%	0%	0%	6,20%	95,7%	95,6%
α_{Min}						
β_{Min}	96,5%	96,1%	95,9%	93,7%	80,0%	43,0%

$S = 4, h_0 = 0$

λ	0,5	0,75	1	1,5	2	3
α_{Max}	1,5%	7,7%	25,0%	79,5%	99,0%	
β_{Max}	0%	0%	0%	3,7%	95,7%	87,2%
α_{Min}	96,5%	98,1%	99%	99,8%	99,98%	
β_{Min}	87,8%	86,6%	86,0%	85,7%	59,1%	36,3%

Kapitel 3

Test des Algorithmus zur Bestimmung der Kappa-Werte

Dieses Kapitel befaßt sich nun mit dem Algorithmus zur Bestimmung der Kappa-Werte aus [Bu01]. Dabei kommt es im Wesentlichen auf zwei Punkte an: Zuerst soll festgestellt werden, ob die Simulation bei einer gewissen Anzahl von Simulationen auch die richtigen Kappa-Werte liefert. Dazu werden die genauen Kappa-Werte berechnet und zusammen mit den simulierten in ein gemeinsames Diagramm geplottet. Da zur Berechnung der Kappa-Werte jedoch die Wahrscheinlichkeiten einer S-dimensionalen Normalverteilung bestimmt werden müssen, werden die Kappa-Werte nur für S=2, S=3 und S=4 berechnet. Als zweites gilt es zu überprüfen, ob der Algorithmus auch für extreme Lambdabereiche und für große Werte von S noch sinnvoll arbeitet. Dazu muß unter anderem festgestellt werden, ob das heuristische Verfahren zur Bestimmung der Indexmenge I endlich ist und inwiefern sich numerische Instabilitäten ergeben.

3.1 Vergleich der Simulation gegen die exakten Werte

3.1.1 Berechnung der Kappa-Werte für S=2

Im Fall S=2 müssen die Werte von κ_0 , κ_1 und κ_2 berechnet werden. Nach [Bu01] handelt es sich dabei um die Koeffizienten der folgenden zu den Kappa-Werten gehörenden Kegel:

$$\begin{aligned}\kappa_2 & : \mathcal{C}(e_1, e_2) \\ \kappa_1 & : \mathcal{C}(e_1) \oplus \left(\{e_1\}^{\perp q} \cap \mathcal{C}^{*\bar{Q}}(e_2) \right), \quad \mathcal{C}(e_2) \oplus \left(\{e_2\}^{\perp q} \cap \mathcal{C}^{*\bar{Q}}(e_1) \right) \\ \kappa_0 & : \mathcal{C}^{*\bar{Q}}(e_1, e_2)\end{aligned}$$

Da die Kegel für alle Werte von S, und somit auch hier, den gesamten Raum aufspannen, gilt natürlich:

$$\sum_{i=0}^S \kappa_i = 1$$

Daher genügt es hier, zwei der drei Kappa-Werte zu berechnen, da sich der dritte aus der obigen Gleichung berechnen läßt.

Betrachten wir als erstes κ_1 . Dabei ergibt sich nach [Bu01], daß $\kappa_1 = 1/2$ gilt.

Die Berechnung von κ_2 ist schon etwas aufwendiger, aber auch hier sei auf obige Quelle verwiesen. Letztendlich ergibt sich:

$$\begin{aligned} \kappa_2 &= \frac{1}{2\pi} \arccos\left(\frac{\bar{Q}_{1,2}}{\bar{Q}_{1,1}\bar{Q}_{2,2}}\right) = \frac{1}{2\pi} \arccos(-\rho_{1,2}) = \frac{1}{2\pi} \arccos\left(\frac{z(\lambda)}{n(\lambda)}\right) \\ &\text{wobei} \\ z(\lambda) &:= \frac{1}{4}\lambda e^\lambda(2 + \lambda) \\ n^2(\lambda) &:= \frac{1}{16}e^{2\lambda}\lambda^4 + \frac{1}{4}e^{2\lambda}\lambda^3 + \frac{1}{4}e^{2\lambda}\lambda^2 + \frac{1}{4}e^{2\lambda}\lambda - \frac{1}{2}e^{2\lambda} + \\ &\quad + \frac{1}{2}\frac{e^{2\lambda}}{\lambda} - \frac{1}{4}e^\lambda + \frac{1}{2}\frac{e^{2\lambda}}{\lambda^2} - \frac{e^\lambda}{\lambda} - \frac{1}{2}\frac{e^\lambda}{\lambda^2} \\ &\text{und } \rho_{12} \text{ der zu der Zufallsvariable } \bar{Z} \text{ gehörende Korrelationskoeffizient ist} \end{aligned}$$

Für κ_0 gilt: $\kappa_0 = 1 - \kappa_1 - \kappa_2$.

Somit können nun die tatsächlichen Kappawerte bestimmt und mit den simulierten Ergebnissen verglichen werden. Für diese Arbeit wird dazu für λ der Bereich von $(0, 2)$, $(0, 5)$ und $(0, 10)$ betrachtet. Dabei wird das jeweilige Intervall gleichmäßig zerlegt und an jeweils 100 Stützstellen werden die Kappawerte bestimmt. Die Simulation arbeitet mit 100000 Simulationen. Das Ergebnis sieht man in Abbildung 3.1 auf der nächsten Seite.

3.1.2 Berechnung der Kappa-Werte für S=3

Dieser Fall erfordert schon etwas mehr Berechnungen, denn aufgrund der Dimensionserhöhung vergrößert sich sowohl die Anzahl der Kappa-Werte als auch die Anzahl der Kegel. Darüberhinaus muß hier auch mit einer dreidimensionalen Normalverteilung gerechnet werden. An dieser Stelle wird etwas genauer in die Berechnung eingegangen, da in dem Buch von [Bu01] hauptsächlich die Theorie bearbeitet wurde, hier aber auch die Umsetzung in den Algorithmus behandelt wird.

Zu Beginn betrachten wir erneut die zu den κ_i gehörenden Kegel:

$$\begin{aligned} \kappa_3 &: \mathcal{C}(e_1, e_2, e_3) \\ \kappa_2 &: \mathcal{C}(e_k, e_l) \oplus \left(\text{lin}\{e_k, e_l\}^{\perp_{\bar{Q}}} \cap \mathcal{C}^{*\bar{Q}}(e_m) \right), \quad \{k, l, m\} = \{1, 2, 3\} \\ \kappa_1 &: \mathcal{C}(e_k) \oplus \left(\text{lin}\{e_k\}^{\perp_{\bar{Q}}} \cap \mathcal{C}^{*\bar{Q}}(e_l, e_m) \right), \quad \{k, l, m\} = \{1, 2, 3\} \\ \kappa_0 &: \mathcal{C}^{*\bar{Q}}(e_1, e_2, e_3) \end{aligned}$$

Betrachten wir zunächst κ_3 . Dazu muß die Wahrscheinlichkeit $P(\bar{Z} \in \mathcal{C}(e_1, e_2, e_3)) = P(\bar{Z} \geq 0)$ berechnet werden. Dies geschieht mittels der Flächenberechnung eines sphärischen Dreiecks auf der Einheitskugel im \mathbb{R}^3 relativ zur Gesamtoberfläche der Einheitskugel. Damit ergibt sich für κ_3 folgende Formel:

$$\kappa_3 = \frac{1}{4\pi} \cdot \frac{1}{2} \sum_{\{k,l,m\}} \arccos\left(\frac{\bar{Q}_{l,m}\bar{Q}_{k,k} - \bar{Q}_{k,l}\bar{Q}_{k,m}}{\sqrt{(\bar{Q}_{k,k}\bar{Q}_{m,m} - \bar{Q}_{k,m}^2)(\bar{Q}_{k,k}\bar{Q}_{l,l} - \bar{Q}_{k,l}^2)}}\right) =$$

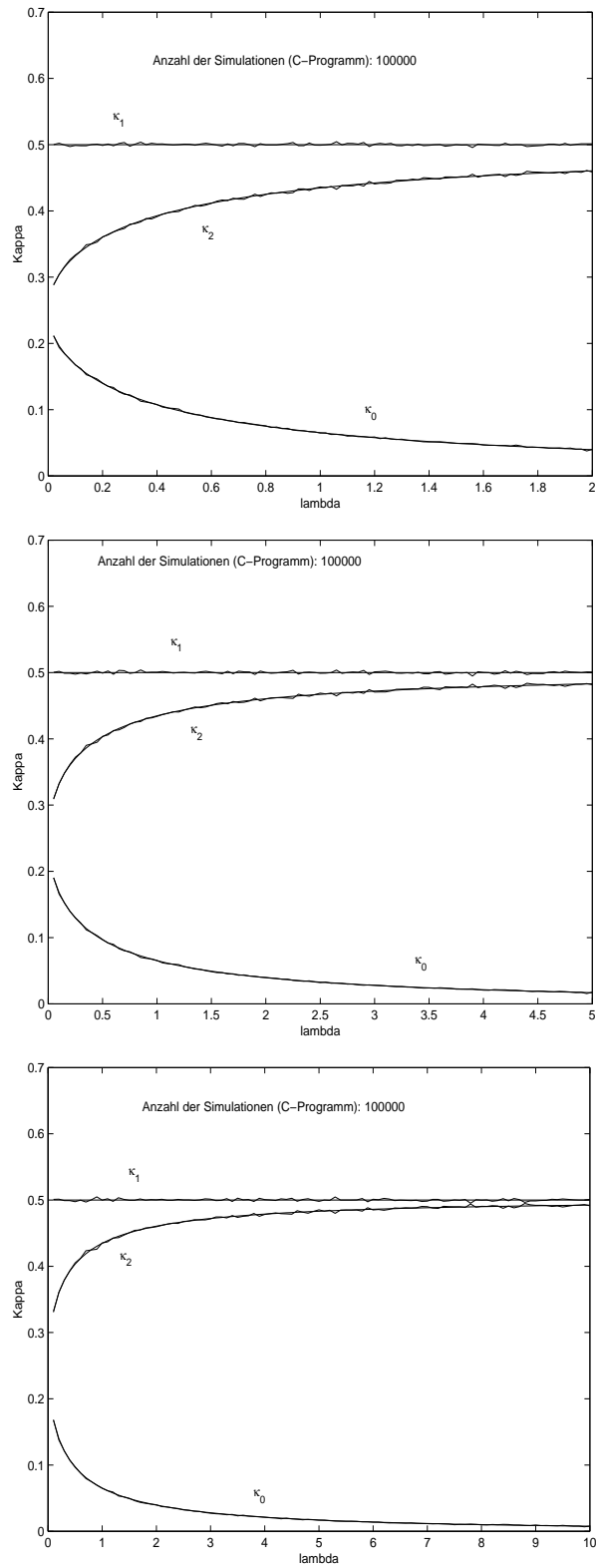


Abbildung 3.1: Vergleich der Simulation mit den berechneten Werten für $S=2$

$$= \frac{\arccos(-\rho_{1,2}) + \arccos(-\rho_{1,3}) + \arccos(-\rho_{2,3}) - \pi}{4\pi}$$

Bei der zweiten Formel handelt es sich um die Berechnung nach [P154], auf deren Herleitung noch genauer im nächsten Abschnitt eingegangen wird. In der Formel stehen $\rho_{1,2}$, $\rho_{1,3}$ und $\rho_{2,3}$ für die Korrelationskoeffizienten der Zufallsvariable \bar{Z} .

Als nächstes soll nun κ_2 berechnet werden. Sei hierzu $k = 1$, $l = 2$ und $m = 3$. Die anderen Fälle lassen sich analog berechnen. Somit ist folgende Wahrscheinlichkeit zu bestimmen:

$$P\left(\bar{Z} \in \mathcal{C}(e_1, e_2) \oplus \left(\text{lin}\{e_1, e_2\}^{\perp_Q} \cap \mathcal{C}^{*\bar{Q}}(e_3)\right)\right)$$

Aufgrund der direkten Summe und der Unabhängigkeit der zugrundeliegenden Normalverteilung läßt sich die Wahrscheinlichkeit zerlegen in das folgende Produkt:

$$P\left(\pi_{\bar{Q}}(\bar{Z} | \text{lin}\{e_1, e_2\}) \in \mathcal{C}(e_1, e_2)\right) \times P\left(\pi_{\bar{Q}}(\bar{Z} | \text{lin}\{e_1, e_2\})^{\perp_Q} \in \mathcal{C}^{*\bar{Q}}(e_3)\right)$$

Beim zweiten Faktor wird \bar{Z} auf das orthogonale Komplement einer zweidimensionalen Fläche projiziert und ist damit eine eindimensionale zentrierte Normalverteilung. Somit gilt:

$$P\left(\pi_{\bar{Q}}(\bar{Z} | \text{lin}\{e_1, e_2\})^{\perp_Q} \in \mathcal{C}^{*\bar{Q}}(e_3)\right) = \frac{1}{2}$$

Für den ersten Faktor soll die weitere Berechnung über einen Satz für den allgemeinen Fall berechnet werden:

Satz 3.1 *Unter den Bezeichnungen dieses Kapitels gilt:*

$$P\left(\pi_{\bar{Q}}(\bar{Z} | \text{lin}\{e_i, i \in I\}) \geq 0\right) = P(Y \geq 0)$$

$$\text{wobei } Y \sim N_{\#I}(0, (Q^*)^{-1}) \text{ und } Q^* = \begin{pmatrix} \bar{Q}_{i_1, i_1} & \cdots & \bar{Q}_{i_1, i_{\#I}} \\ \vdots & & \vdots \\ \bar{Q}_{i_{\#I}, i_1} & \cdots & \bar{Q}_{i_{\#I}, i_{\#I}} \end{pmatrix}.$$

Beweis:

Aufgrund der Projektion gilt:

$$\begin{aligned} P\left(\pi_{\bar{Q}}(\bar{Z} | \text{lin}\{e_i, i \in I\}) \geq 0\right) &= P\left(E_I(E_I^T \bar{Q} E_I)^{-1} E_I^T \bar{Q} \bar{Z} \geq 0\right) = \\ &= P\left(\bar{Y} \geq 0\right) \text{ mit } \bar{Y} := E_I(E_I^T \bar{Q} E_I)^{-1} E_I^T \bar{Q} \bar{Z} \end{aligned}$$

Aufgrund der Projektion auf die lineare Hülle von $(e_i, i \in I)$ sind die Komponenten von \bar{Y} , auf die nicht projiziert wird, gleich Null und die Zufallsvariable \bar{Y} ist in diesen Komponenten entartet, d.h. es gilt:

$$\bar{Y} = \begin{pmatrix} \bar{Y}_1 \\ \vdots \\ \bar{Y}_S \end{pmatrix} \text{ mit } \bar{Y}_i = 0 \text{ falls } i \notin I.$$

Faßt man nun die Komponenten, die nicht Null sind, zu einer neuen Zufallsvariablen Y zusammen, so gilt (da die Nullkomponenten natürlich auch die Bedingung größer gleich Null erfüllen):

$$\begin{aligned} P\left(\pi_{\bar{Q}}(\bar{Z} | \text{lin}\{e_i, i \in I\}) \geq 0\right) &= P\left(\bar{Y} \geq 0\right) = P\left(Y \geq 0\right) \\ \text{mit } Y &:= \left((E_I(E_I^T \bar{Q} E_I)^{-1} E_I^T \bar{Q} \bar{Z})_i\right) \quad (i \in I) \end{aligned}$$

Dabei handelt es sich bei Y um eine $\#I$ -dimensionale Normalverteilung, von der noch die Kovarianzmatrix bestimmt werden muß. Berechnen wir zuerst die Kovarianzmatrix von \bar{Y} :

$$\begin{aligned} E[\bar{Y}\bar{Y}^T] &= E_I(E_I^T \bar{Q} E_I)^{-1} E_I^T \bar{Q} \cdot \underbrace{E[\bar{Z}\bar{Z}^T]}_{=\bar{Q}^{-1}} \cdot \underbrace{\bar{Q}^T}_{=\bar{Q}} E_I \underbrace{(E_I^T \bar{Q} E_I)^{-T}}_{=(E_I^T \bar{Q} E_I)^{-1}} E_I^T = \\ &\quad \underbrace{\hspace{10em}}_{=I} \\ &\quad \underbrace{\hspace{10em}}_{=E_I^T \bar{Q} E_I} \\ &= E_I(E_I^T \bar{Q} E_I)^{-1} E_I^T \end{aligned}$$

Sei $I = \{i_1, \dots, i_{\#I}\}$, so gilt aufgrund der Matrixmultiplikation:

$$E_I^T \bar{Q} E_I = \begin{pmatrix} \bar{Q}_{i_1, i_1} & \cdots & \bar{Q}_{i_1, i_{\#I}} \\ \vdots & & \vdots \\ \bar{Q}_{i_{\#I}, i_1} & \cdots & \bar{Q}_{i_{\#I}, i_{\#I}} \end{pmatrix} =: \tilde{Q}$$

Da $E[\bar{Y}\bar{Y}^T] = E_I(E_I^T \bar{Q} E_I)^{-1} E_I^T = E_I(\tilde{Q})^{-1} E_I^T$ gilt, folgt aufgrund der speziellen Matrixstruktur von E_I :

$$\begin{cases} \text{Falls } k, l \notin I : E[\bar{Y}\bar{Y}^T]_{k,l} = 0 \\ \text{Falls } k, l \in I \Rightarrow \exists p, q : i_p = k, i_q = l \text{ und } (E[\bar{Y}\bar{Y}^T])_{k,l} = (E[\bar{Y}\bar{Y}^T])_{i_p, i_q} = (\tilde{Q}^{-1})_{p,q} \end{cases}$$

$$\Rightarrow (Q^*) = [E(YY^T)]^{-1} = [(\tilde{Q}^{-1})]^{-1} = \tilde{Q} = \begin{pmatrix} \bar{Q}_{i_1, i_1} & \cdots & \bar{Q}_{i_1, i_{\#I}} \\ \vdots & & \vdots \\ \bar{Q}_{i_{\#I}, i_1} & \cdots & \bar{Q}_{i_{\#I}, i_{\#I}} \end{pmatrix}$$

□

Wendet man diesen Satz nun auf dieses konkrete Problem an, so ergibt sich:

$$P(\pi_{\bar{Q}}(\bar{Z} | \text{lin}\{e_1, e_2\}) \geq 0) = P(Y \geq 0)$$

mit

$$Y \sim N_2(0, (Q^*)^{-1}) \text{ wobei } Q^* := \begin{pmatrix} \bar{Q}_{1,1} & \bar{Q}_{1,2} \\ \bar{Q}_{1,2} & \bar{Q}_{2,2} \end{pmatrix}$$

Setzt man nun wieder die Formel für die zweidimensionale Normalverteilung an, so ergibt sich:

$$P(\pi_{\bar{Q}}(\bar{Z} | \text{lin}\{e_1, e_2\}) \geq 0) = P(Y \geq 0) = \frac{1}{2\pi} \arccos \left(\frac{\bar{Q}_{1,2}}{\sqrt{\bar{Q}_{1,1}\bar{Q}_{2,2}}} \right)$$

Alles zusammen liefert:

$$P\left(\bar{Z} \in \mathcal{C}(e_1, e_2) \oplus \left(\text{lin}\{e_1, e_2\}^{\perp \bar{Q}} \cap \mathcal{C}^{*\bar{Q}}(e_3)\right)\right) = P(Y \geq 0) = \frac{1}{2\pi} \arccos \left(\frac{\bar{Q}_{1,2}}{\sqrt{\bar{Q}_{1,1}\bar{Q}_{2,2}}} \right) \times \frac{1}{2}$$

Da sich κ_2 aus 3 disjunkten Kegeln (den Permutationen der Menge $\{1, 2, 3\}$ unter Beachtung der Tatsache, daß je zwei dieser 6 Kegel identisch sind) zusammensetzt, ergibt sich κ_2 als Summe von 3 Wahrscheinlichkeiten:

$$\kappa_2 = \frac{1}{4\pi} \left(\arccos \left(\frac{\bar{Q}_{1,2}}{\sqrt{\bar{Q}_{1,1}\bar{Q}_{2,2}}} \right) + \arccos \left(\frac{\bar{Q}_{2,3}}{\sqrt{\bar{Q}_{2,2}\bar{Q}_{3,3}}} \right) + \arccos \left(\frac{\bar{Q}_{1,3}}{\sqrt{\bar{Q}_{3,3}\bar{Q}_{1,1}}} \right) \right)$$

Bleibt noch κ_1 zu berechnen. Hier gilt (wiederum nur für den Fall $k = 1, l = 2$ und $m = 3$ vorgerechnet, der Rest ergibt sich analog):

$$\begin{aligned} & P \left(\bar{Z} \in \mathcal{C}(e_1) \oplus \left(\text{lin}\{e_1\}^{\perp_{\bar{Q}}} \cap \mathcal{C}^{*\bar{Q}}(e_2, e_3) \right) \right) = \\ & = P \left(\pi_{\bar{Q}}(\bar{Z} | \text{lin}\{e_1\}) \in \mathcal{C}(e_1) \right) \times P \left(\pi_{\bar{Q}}(\bar{Z} | \text{lin}\{e_1\})^{\perp_{\bar{Q}}} \in \mathcal{C}^{*\bar{Q}}(e_2, e_3) \right) \end{aligned}$$

Der erste Faktor gleicht von der Berechnung her einem Fall, der bereits zuvor berechnet wurde. Erneut wird \bar{Z} auf eine eindimensionale Gerade projiziert und somit ist die Wahrscheinlichkeit, daß die so entstehende Zufallsvariable positiv ist, gleich $\frac{1}{2}$.

Der zweite Faktor bedarf schon einer genaueren Betrachtung. Es gilt jedoch:

$$\pi_{\bar{Q}}(\bar{Z} | \text{lin}\{e_1\})^{\perp_{\bar{Q}}} \in \mathcal{C}^{*\bar{Q}}(e_2, e_3) \Leftrightarrow Y := E_{\{2,3\}}^T \bar{Q} (e_1 (e_1^T \bar{Q} e_1)^{-1} e_1^T \bar{Q} \bar{Z} - \bar{Z}) \geq 0$$

Dazu wird die Kovarianzmatrix von Y benötigt. Es kann leicht nachgerechnet werden, daß sie folgende Struktur besitzt:

$$E[YY^T] = E_{\{2,3\}}^T \bar{Q} E_{\{2,3\}} - \frac{1}{\|e_1\|_{\bar{Q}}^2} E_{\{2,3\}}^T \bar{Q} e_1 e_1^T \bar{Q} E_{\{2,3\}}$$

Da man für die Formel im zweidimensionalen Fall die Inverse der Kovarianzmatrix benötigt, muß diese auch noch berechnet werden. Die Berechnung der Inversen wurde mittels MAPLE durchgeführt. Dabei sei c eine von der Matrix \bar{Q} abhängige Konstante:

$$(E[YY^T])^{-1} = c \cdot \underbrace{\begin{pmatrix} \bar{Q}_{3,3}\bar{Q}_{1,1} - \bar{Q}_{1,3}^2 & \bar{Q}_{1,2}\bar{Q}_{1,3} - \bar{Q}_{2,3}\bar{Q}_{1,1} \\ \bar{Q}_{1,2}\bar{Q}_{1,3} - \bar{Q}_{2,3}\bar{Q}_{1,1} & \bar{Q}_{2,2}\bar{Q}_{1,1} - \bar{Q}_{1,2}^2 \end{pmatrix}}_{:=\tilde{Q}}$$

Damit ergibt sich:

$$\begin{aligned} & P \left(\pi_{\bar{Q}}(\bar{Z} | \text{lin}\{e_1\})^{\perp_{\bar{Q}}} \in \mathcal{C}^{*\bar{Q}}(e_2, e_3) \right) = P(Y \geq 0) = \frac{1}{2\pi} \arccos \left(\frac{c\tilde{Q}_{1,2}}{\sqrt{c\tilde{Q}_{1,1}c\tilde{Q}_{2,2}}} \right) = \\ & = \frac{1}{2\pi} \arccos \left(\frac{\bar{Q}_{1,2}\bar{Q}_{1,3} - \bar{Q}_{2,3}\bar{Q}_{1,1}}{\sqrt{(\bar{Q}_{3,3}\bar{Q}_{1,1} - \bar{Q}_{1,3}^2) \cdot (\bar{Q}_{2,2}\bar{Q}_{1,1} - \bar{Q}_{1,2}^2)}} \right) \end{aligned}$$

κ_1 ergibt sich auch diesmal wieder als Summe der Wahrscheinlichkeiten der disjunkten Kegel. Dazu muß die obige Rechnung noch für $\{k = 2, l = 3, m = 1\}$ und $\{k = 3, l = 1, m = 2\}$ analog

durchgeführt werden, so daß sich letztendlich folgender Ausdruck für κ_2 ergibt:

$$\begin{aligned} \kappa_2 = & \frac{1}{4\pi} \arccos \left(\frac{\bar{Q}_{1,2}\bar{Q}_{1,3} - \bar{Q}_{2,3}\bar{Q}_{1,1}}{\sqrt{(\bar{Q}_{3,3}\bar{Q}_{1,1} - \bar{Q}_{1,3}^2) \cdot (\bar{Q}_{2,2}\bar{Q}_{1,1} - \bar{Q}_{1,2}^2)}} \right) + \\ & + \frac{1}{4\pi} \arccos \left(\frac{\bar{Q}_{2,3}\bar{Q}_{2,1} - \bar{Q}_{3,1}\bar{Q}_{2,2}}{\sqrt{(\bar{Q}_{1,1}\bar{Q}_{2,2} - \bar{Q}_{2,1}^2) \cdot (\bar{Q}_{3,3}\bar{Q}_{2,2} - \bar{Q}_{2,3}^2)}} \right) + \\ & + \frac{1}{4\pi} \arccos \left(\frac{\bar{Q}_{3,1}\bar{Q}_{3,2} - \bar{Q}_{1,2}\bar{Q}_{3,3}}{\sqrt{(\bar{Q}_{2,2}\bar{Q}_{3,3} - \bar{Q}_{3,2}^2) \cdot (\bar{Q}_{1,1}\bar{Q}_{3,3} - \bar{Q}_{3,1}^2)}} \right) \end{aligned}$$

Bleibt nur noch κ_0 zu berechnen. Dies jedoch geschieht einfach mittels: $\kappa_0 = 1 - \kappa_1 - \kappa_2 - \kappa_3$. Auch hier wurden erneut die berechneten Kappawerte mit den Werten aus der Simulation verglichen. Die Abbildung 3.2 auf der nächsten Seite wurde wiederum mit 100000 Simulationen durchgeführt.

3.1.3 Berechnung der Kappa-Werte für S=4

Im Fall S=4 gibt es zur Berechnung der Kappa-Werte keine einfachen Formeln mehr, gerade zur Berechnung von κ_4 bedarf es schon eines enormen Aufwands. Hierbei ist die Wahrscheinlichkeit zu bestimmen, daß eine zentrierte vierdimensionale Normalverteilung größer gleich Null ist. Hierfür gibt es für spezielle Verteilungen in der Literatur einige explizite Formeln. So behandelt Coxeter in [Co35] den Fall, daß die Korrelationsmatrix eine Tridiagonalform besitzt (d.h. es gilt: $\rho_{1,3}$, $\rho_{1,4}$ und $\rho_{2,4}$ sind alle Null). Ruben [Ru54] hingegen betrachtet eine vierdimensionale Normalverteilung, bei der alle Korrelationskoeffizienten identisch sind. Im allgemeinen Fall, der hier vorliegt, muß dem Weg von Plackett [Pl54] gefolgt werden, der auf einer Dimensionsreduktion basiert. Dabei zeigt sich schon der Rechenaufwand, der für höhere Dimensionen zu betreiben ist. Dies ist auch der Grund, warum es nicht mehr viel Sinn macht, die genauen Kappa-Werte ab S=4 zu berechnen.

Bevor wir jedoch die einzelnen Kappa-Werte berechnen wollen, betrachten wir zuerst noch die dazugehörigen Kegel:

$$\begin{aligned} \kappa_4 : & \mathcal{C}(e_1, e_2, e_3, e_4) \\ \kappa_3 : & \mathcal{C}(e_k, e_l, e_m) \oplus \left(\text{lin}\{e_k, e_l, e_m\}^{\perp \bar{Q}} \cap \mathcal{C}^{*\bar{Q}}(e_n) \right), \quad \{k, l, m, n\} = \{1, 2, 3, 4\} \\ \kappa_2 : & \mathcal{C}(e_k, e_l) \oplus \left(\text{lin}\{e_k, e_l\}^{\perp \bar{Q}} \cap \mathcal{C}^{*\bar{Q}}(e_m, e_n) \right), \quad \{k, l, m, n\} = \{1, 2, 3, 4\} \\ \kappa_1 : & \mathcal{C}(e_k) \oplus \left(\text{lin}\{e_k\}^{\perp \bar{Q}} \cap \mathcal{C}^{*\bar{Q}}(e_l, e_m, e_n) \right), \quad \{k, l, m, n\} = \{1, 2, 3, 4\} \\ \kappa_0 : & \mathcal{C}^{*\bar{Q}}(e_1, e_2, e_3, e_4) \end{aligned}$$

Betrachten wir zuerst κ_1 . Zu berechnen ist dazu folgender Ausdruck:

$$P \left(\bar{Z} \in \mathcal{C}(e_k) \oplus \left(\text{lin}\{e_k\}^{\perp \bar{Q}} \cap \mathcal{C}^{*\bar{Q}}(e_l, e_m, e_n) \right) \right)$$

Dabei läßt sich die Wahrscheinlichkeit erneut aufgrund der Unabhängigkeit der zugrundeliegenden Normalverteilung wie folgt zerlegen:

$$P \left(\pi_{\bar{Q}}(\bar{Z} | \text{lin}\{e_k\}) \in \mathcal{C}(e_k) \right) \times P \left(\pi_{\bar{Q}}(\bar{Z} | \text{lin}\{e_k\}^{\perp \bar{Q}}) \in \mathcal{C}^{*\bar{Q}}(e_l, e_m, e_n) \right)$$

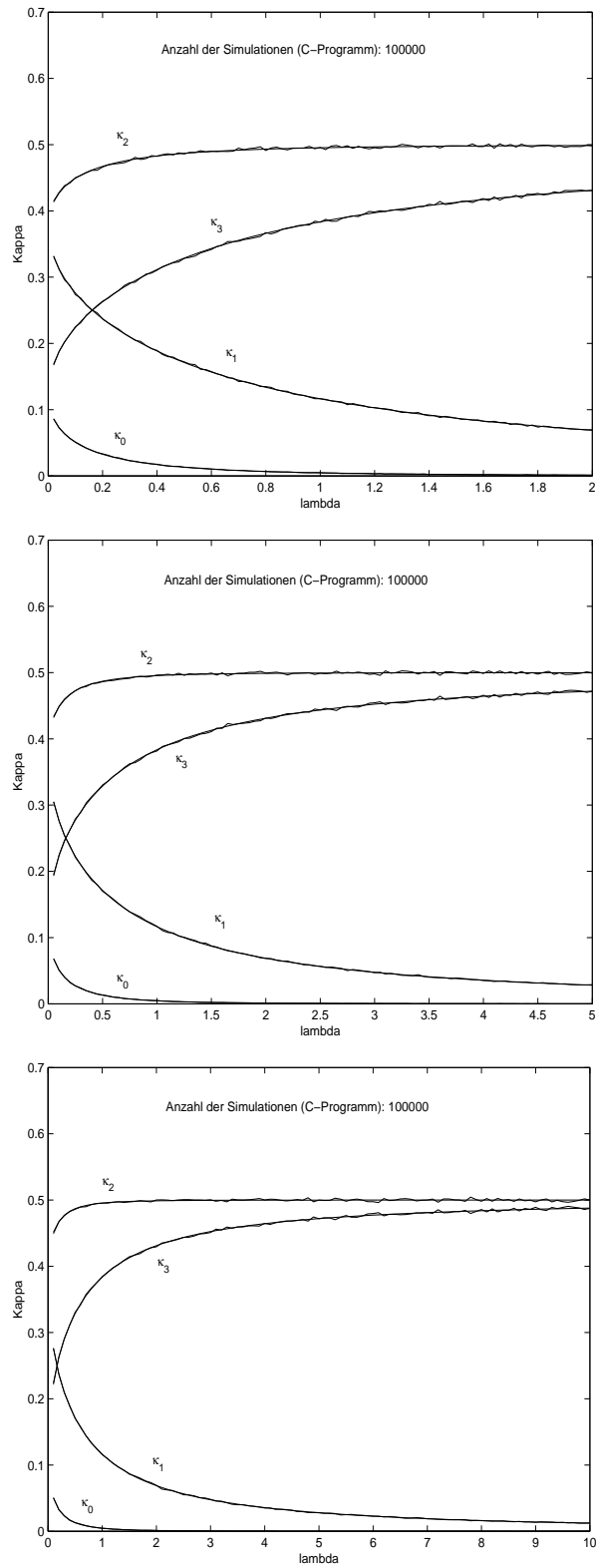


Abbildung 3.2: Vergleich der Simulation mit den berechneten Werten für $S=3$

Hier sei die weitere Rechnung nur an der Permutation $\{k = 1, l = 2, m = 3, n = 4\}$ vorgeführt. Die anderen Fälle lassen sich genauso berechnen. Dann gilt:

$$P\left(\pi_{\bar{Q}}(\bar{Z}|\text{lin}\{e_1\}^{\perp\bar{Q}} \in \mathcal{C}^{*\bar{Q}}(e_2, e_3, e_4))\right) = P\left(\underbrace{E_{\{2,3,4\}}^T \bar{Q} (e_1 (e_1^T \bar{Q} e_1)^{-1} e_1^T \bar{Q} - I)}_{:=Y} \bar{Z} \geq 0\right) = P(\tilde{Z} \geq 0)$$

$$\text{mit } \tilde{Z} := Y\bar{Z} \text{ und } \tilde{Z} \sim N_3\left(0, \underbrace{(YQ^{-1}Y^T)}_{:=Q_3}\right).$$

Somit hat man das Problem auf eine dreidimensionale Normalverteilung reduziert und kann die Wahrscheinlichkeit wie im Fall S=3 berechnen. Dazu muß allerdings zuerst die Matrix Q_3 berechnet und dann invertiert werden. Um nun κ_1 berechnen zu können, benötigt man noch den ersten Teil des Produkts; dieser beträgt jedoch erneut genau $\frac{1}{2}$.

Als nächstes soll nun κ_2 berechnet werden. Dazu wird die direkte Summe wie in den Beispielen zuvor zerlegt:

$$P\left(\bar{Z} \in \mathcal{C}(e_k, e_l) \oplus \left(\text{lin}\{e_k, e_l\}^{\perp\bar{Q}} \cap \mathcal{C}^{*\bar{Q}}(e_m, e_n)\right)\right) = \\ P\left(\pi_{\bar{Q}}(\bar{Z}|\text{lin}\{e_k, e_l\}) \in \mathcal{C}(e_k, e_l)\right) \times P\left(\pi_{\bar{Q}}(\bar{Z}|\text{lin}\{e_k, e_l\}^{\perp\bar{Q}}) \in \mathcal{C}^{*\bar{Q}}(e_m, e_n)\right)$$

Zur Berechnung des ersten Faktors greifen wir erneut auf Satz 3.1 zurück. Damit läßt sich die Wahrscheinlichkeit umformulieren in

$$P\left(\pi_{\bar{Q}}(\bar{Z}|\text{lin}\{e_k, e_l\}) \in \mathcal{C}(e_k, e_l)\right) = P(\tilde{Z} \geq 0) \text{ mit } \tilde{Z} \sim N_2\left(0, \begin{pmatrix} Q_{k,k} & Q_{k,l} \\ Q_{l,k} & Q_{l,l} \end{pmatrix}\right)$$

und es gilt:

$$P\left(\pi_{\bar{Q}}(\bar{Z}|\text{lin}\{e_k, e_l\}) \in \mathcal{C}(e_k, e_l)\right) = \frac{1}{2\pi} \arccos\left(\frac{Q_{k,l}}{\sqrt{Q_{k,k}Q_{l,l}}}\right)$$

Für den zweiten Faktor setzt man wie schon bei der Berechnung von κ_1 die Formel für die Projektion an. Dann gilt:

$$P\left(\pi_{\bar{Q}}(\bar{Z}|\text{lin}\{e_k, e_l\}^{\perp\bar{Q}}) \in \mathcal{C}^{*\bar{Q}}(e_m, e_n)\right) = \\ = P\left(\underbrace{E_{\{m,n\}}^T \bar{Q} \left(E_{\{k,l\}} (E_{\{k,l\}}^T \bar{Q} E_{\{k,l\}})^{-1} E_{\{k,l\}}^T \bar{Q} - I\right)}_{:=\tilde{Y}(k,l)} \bar{Z} \geq 0\right) = \\ = P(\tilde{Y}(k, l, m, n)\bar{Z} \geq 0) \text{ mit } \tilde{Y}(k, l, m, n)\bar{Z} \sim N_2(0, \tilde{Q}(k, l, m, n)^{-1}) \\ \text{und } \tilde{Q}(k, l, m, n) = \begin{pmatrix} \tilde{Q}(k, l, m, n)_{1,1} & \tilde{Q}(k, l, m, n)_{1,2} \\ \tilde{Q}(k, l, m, n)_{2,1} & \tilde{Q}(k, l, m, n)_{2,2} \end{pmatrix}$$

Die Matrix \tilde{Q} erhält man dabei durch einfache Matrixmultiplikation und aus den Sätzen über die Kovarianzmatrix bei Transformationen. Sie hat folgende Gestalt:

$$\tilde{Q}(k, l, m, n) = \text{Cov}(\tilde{Y}(k, l, m, n)\bar{Z}) = \tilde{Y}(k, l, m, n)\text{Cov}(\bar{Z})\tilde{Y}(k, l, m, n)^T \\ = \tilde{Y}(k, l, m, n)\bar{Q}^{-1}\tilde{Y}(k, l, m, n)^T = \frac{1}{Q_{k,k}Q_{l,l} - Q_{k,l}^2} \cdot \begin{pmatrix} A & B \\ B & D \end{pmatrix}$$

wobei

$$\begin{aligned}
 A &:= \bar{Q}_{k,k}\bar{Q}_{l,l}\bar{Q}_{m,m} - \bar{Q}_{m,m}\bar{Q}_{k,l}^2 - \bar{Q}_{k,m}\bar{Q}_{l,l}^2 - \bar{Q}_{l,m}^2\bar{Q}_{k,k} + 2\bar{Q}_{l,m}\bar{Q}_{k,l}\bar{Q}_{k,m} \\
 B &:= \bar{Q}_{m,n}\bar{Q}_{k,k}\bar{Q}_{l,l} - \bar{Q}_{m,n}\bar{Q}_{k,l}^2 - \bar{Q}_{k,m}\bar{Q}_{k,n}\bar{Q}_{l,l} - \bar{Q}_{l,m}^2\bar{Q}_{k,k}\bar{Q}_{l,n} + \bar{Q}_{l,m}\bar{Q}_{k,l}\bar{Q}_{k,n} + \bar{Q}_{k,l}\bar{Q}_{k,m}\bar{Q}_{l,n} \\
 D &:= \bar{Q}_{k,k}\bar{Q}_{l,l}\bar{Q}_{n,n} - \bar{Q}_{k,k}\bar{Q}_{l,n}^2 - \bar{Q}_{k,n}^2\bar{Q}_{l,l} - \bar{Q}_{n,n}^2\bar{Q}_{k,l} + 2\bar{Q}_{l,n}\bar{Q}_{k,l}\bar{Q}_{k,n}
 \end{aligned}$$

Somit befindet man sich erneut im zweidimensionalen Fall und die Wahrscheinlichkeit läßt sich wie folgt berechnen:

$$P\left(\pi_{\bar{Q}}(\bar{Z}|\text{lin}\{e_k, e_l\}^{\perp\bar{Q}}) \in \mathcal{C}^{*\bar{Q}}(e_m, e_n)\right) = \frac{1}{2\pi} \arccos\left(\frac{\tilde{Q}_{1,2}}{\sqrt{\tilde{Q}_{1,1}\tilde{Q}_{2,2}}}\right)$$

Somit ergibt sich κ_2 als Produkt der beiden Faktoren:

$$\kappa_2 = \frac{1}{16\pi^2} \sum_{\{k,l,m,n\}=\{1,2,3,4\}} \arccos\left(\frac{Q_{k,l}}{Q_{k,k}Q_{l,l}}\right) \arccos\left(\frac{\tilde{Q}(k,l,m,n)_{1,2}}{\tilde{Q}(k,l,m,n)_{1,1}\tilde{Q}(k,l,m,n)_{2,2}}\right)$$

Der Vorfaktor entsteht dadurch, daß hier je 4 Permutationen von $\{1,2,3,4\}$ denselben Kegel liefern und daher nur einmal gezählt werden dürfen, in obiger Summation aber alle 4 Kegel mit ihren Wahrscheinlichkeiten addiert werden. Somit muß das Ergebnis am Ende noch durch 4 geteilt werden. Zur Berechnung von κ_3 zerlegen wir die zugrundeliegenden Kegel wieder in die Teile, die durch die direkte Summe bestimmt werden:

$$\begin{aligned}
 &P\left(\bar{Z} \in \mathcal{C}(e_k, e_l, e_m) \oplus \left(\text{lin}\{e_k, e_l, e_m\}^{\perp\bar{Q}} \cap \mathcal{C}^{*\bar{Q}}(e_n)\right)\right) = \\
 &= P\left(\pi_{\bar{Q}}(\bar{Z}|\text{lin}\{e_k, e_l, e_m\}) \in \mathcal{C}(e_k, e_l, e_m)\right) \times P\left(\pi_{\bar{Q}}(\bar{Z}|\text{lin}\{e_k, e_l, e_m\}^{\perp\bar{Q}}) \in \mathcal{C}^{*\bar{Q}}(e_n)\right)
 \end{aligned}$$

Der erste Faktor wird mittels Satz 3.1 berechnet:

$$\begin{aligned}
 &P\left(\pi_{\bar{Q}}(\bar{Z}|\text{lin}\{e_k, e_l, e_m\}) \in \mathcal{C}(e_k, e_l, e_m)\right) = P\left(\tilde{Z} \geq 0\right) \\
 &\text{mit } \tilde{Z} \sim N_3(0, \tilde{Q}^{-1}) \text{ wobei } \tilde{Q} = \begin{pmatrix} \bar{Q}_{k,k} & \bar{Q}_{k,l} & \bar{Q}_{k,m} \\ \bar{Q}_{k,l} & \bar{Q}_{l,l} & \bar{Q}_{l,m} \\ \bar{Q}_{k,m} & \bar{Q}_{l,m} & \bar{Q}_{m,m} \end{pmatrix}
 \end{aligned}$$

Somit läßt sich der erste Faktor als dreidimensionale Normalverteilung leicht berechnen:

$$\begin{aligned}
 &P\left(\pi_{\bar{Q}}(\bar{Z}|\text{lin}\{e_k, e_l, e_m\}) \in \mathcal{C}(e_k, e_l, e_m)\right) = \\
 &= \frac{1}{8\pi} \sum_{\{k,l,m\}=\{1,2,3\}} \arccos\left(\frac{\bar{Q}_{l,m}\bar{Q}_{k,k} - \bar{Q}_{k,l}\bar{Q}_{k,m}}{\sqrt{(\bar{Q}_{k,k}\bar{Q}_{m,m} - \bar{Q}_{k,m}^2) \cdot (\bar{Q}_{k,k}\bar{Q}_{l,l} - \bar{Q}_{k,l}^2)}}\right) - \frac{1}{4}
 \end{aligned}$$

Beim zweiten Faktor wird die vierdimensionale Normalverteilung auf das orthogonale Komplement eines dreidimensionalen Raums projiziert. Im \mathbb{R}^4 handelt es sich somit um eine Gerade und die gesuchte Wahrscheinlichkeit beträgt somit $\frac{1}{2}$. Alles zusammen liefert:

$$\kappa_2 = \frac{1}{2} \left(\frac{1}{8\pi} \sum_{\{k,l,m\} \in \{1,2,3,4\}} \arccos\left(\frac{\bar{Q}_{l,m}\bar{Q}_{k,k} - \bar{Q}_{k,l}\bar{Q}_{k,m}}{\sqrt{(\bar{Q}_{k,k}\bar{Q}_{m,m} - \bar{Q}_{k,m}^2) \cdot (\bar{Q}_{k,k}\bar{Q}_{l,l} - \bar{Q}_{k,l}^2)}}\right) - 1 \right)$$

Bleibt nur noch $\kappa_4 = P(\bar{Z} \in \mathcal{C}(e_1, e_2, e_3, e_4)) = P(\bar{Z} \geq 0)$ zu berechnen.

Wie schon zu Beginn dieses Abschnitts erklärt, werden wir hierbei dem Weg von Plackett [P154] folgen. Da jedoch einige Ungenauigkeiten in diesem Weg enthalten sind, werde ich an einigen Stellen abweichen. Der Grundgedanke bleibt dennoch unverändert.

Es gelten folgende Bezeichnungen: $f_n(\mathbf{x}; \mathbf{0}, \Sigma)$ sei die Dichte einer zentrierten n -dimensionalen

Normalverteilung mit Kovarianzmatrix $\Sigma = \begin{pmatrix} \sigma_{1,1} & \cdots & \sigma_{1,n} \\ \vdots & \ddots & \vdots \\ \sigma_{1,n} & \cdots & \sigma_{n,n} \end{pmatrix}$, $F_n(\mathbf{0}; \mathbf{0}, \Sigma)$ die zu f gehören-

de Verteilungsfunktion an der Stelle $\mathbf{0}$. Außerdem seien $\rho_{i,j}$ ($1 \leq i, j \leq n$) die Elemente der Korrelationsmatrix. Sei weiter $\mathbf{X} = (X_1, X_2, \dots, X_n)^T$ verteilt gemäß dieser Verteilungsfunktion. Für die weitere Rechnung kann o.B.d.A. angenommen werden, daß $\text{Var}(X_i) = 1$ ($1 \leq i \leq n$) ist. Der Ausgangspunkt unserer Rechnungen ist die folgende Gleichung:

$$\frac{\partial}{\partial \sigma_{i,j}} f(\mathbf{x}, \mathbf{0}, \Sigma) = \frac{\partial^2}{\partial x_i \partial x_j} f(\mathbf{x}, \mathbf{0}, \Sigma), \quad (i \neq j) \quad (3.1)$$

Um dies zu beweisen, benutzen wir die folgende Darstellung der Dichte mittels einer Transformation ihrer charakteristischen Funktion:

$$f_n(\mathbf{x}; \mathbf{0}, \Sigma) = (2\pi)^{-n} \int \int \cdots \int \exp(-it'\mathbf{x} - \frac{1}{2}\mathbf{t}'\Sigma\mathbf{t}) dt_1 dt_2 \dots dt_n$$

Somit ergeben beide Seiten von (3.1) den Ausdruck

$$-(2\pi)^{-n} \int \int \cdots \int t_i t_j \exp(-it'\mathbf{x} - \frac{1}{2}\mathbf{t}'\Sigma\mathbf{t}) dt_1 dt_2 \dots dt_n$$

und die Gleichung (3.1) ist bewiesen.

Da wir angenommen haben, daß $\sigma_{i,i} = 1$ ($1 \leq i \leq n$) ist, gilt hier speziell $\sigma_{i,j} = \rho_{i,j}$ ($1 \leq i, j \leq n$) und somit gilt in diesem Fall auch die Gleichung:

$$\frac{\partial}{\partial \rho_{i,j}} f(\mathbf{x}, \mathbf{0}, \Sigma) = \frac{\partial^2}{\partial x_i \partial x_j} f(\mathbf{x}, \mathbf{0}, \Sigma), \quad (i \neq j) \quad (3.2)$$

Da sich die Verteilungsfunktion in der Form

$$F_n(\mathbf{0}; \mathbf{0}, \Sigma) = \int_{-\infty}^0 \int_{-\infty}^0 \int_{-\infty}^0 \cdots \int_{-\infty}^0 f_n(\mathbf{x}; \mathbf{0}, \Sigma) \prod_{i=1}^n dx_i \quad (3.3)$$

schreiben läßt und man die Integration und Differentiation vertauschen darf, gilt auch die folgende Gleichung:

$$\begin{aligned} \frac{\partial}{\partial \rho_{n-1,n}} F_n(\mathbf{0}; \mathbf{0}, \Sigma) &\stackrel{(3.3)}{=} \int_{-\infty}^0 \int_{-\infty}^0 \cdots \left(\int_{-\infty}^0 \int_{-\infty}^0 \frac{\partial}{\partial \rho_{n-1,n}} f_n(\mathbf{x}; \mathbf{0}, \Sigma) dx_{n-1} dx_n \right) \prod_{i=1}^{n-2} dx_i = \\ &\stackrel{(3.2)}{=} \int_{-\infty}^0 \int_{-\infty}^0 \cdots \int_{-\infty}^0 \left(\int_{-\infty}^0 \int_{-\infty}^0 \frac{\partial^2}{\partial x_{n-1} \partial x_n} f_n(\mathbf{x}; \mathbf{0}, \Sigma) dx_{n-1} dx_n \right) \prod_{i=1}^{n-2} dx_i = \\ &= \int_{-\infty}^0 \int_{-\infty}^0 \cdots \int_{-\infty}^0 f_n(\mathbf{0}, \mathbf{0}, \mathbf{x}_1, \dots, \mathbf{x}_{n-2}; \mathbf{0}, \Sigma) \prod_{i=1}^{n-2} dx_i \end{aligned} \quad (3.4)$$

Um dies in eine für die spätere Rechnung bessere Form zu bringen, benutzen wir hier die Darstellung mittels der bedingten Dichte. Dazu partitionieren wir den Vektor \mathbf{x} wie folgt:

$$\begin{aligned}\mathbf{x}'_1 &:= (x_1, \dots, x_{n-2}) \\ \mathbf{x}'_2 &:= (x_{n-1}, x_n)\end{aligned}$$

und

$$\Sigma = \begin{pmatrix} \Sigma_{1,1} & \Sigma_{1,2} \\ \Sigma_{2,1} & \Sigma_{2,2} \end{pmatrix}$$

wo $\Sigma_{2,2}$ eine 2×2 Matrix ist. Damit läßt sich Gleichung (3.4) umformulieren in:

$$\frac{\partial}{\partial \sigma_{n-1,n}} F_n(\mathbf{0}; \mathbf{0}, \Sigma) = f_2(0, 0; \mathbf{0}, \Sigma_{2,2}) \int_{-\infty}^0 \cdots \int_{-\infty}^0 f_{1|2}(\mathbf{x}_2; \mathbf{0}, \Sigma_{11.2}) \prod_{i=1}^{n-2} dx_i \quad (3.5)$$

wo f_2 die marginale Dichte von $(\mathbf{X}_{n-1}, \mathbf{X}_n)'$ ausgewertet an der Stelle $(0, 0)'$ und $f_{1|2}$ die bedingte Dichte von $(\mathbf{X}_1, \dots, \mathbf{X}_{n-2})'$ gegeben $(\mathbf{X}_{n-1}, \mathbf{X}_n)' = (0, 0)'$ ist. Dabei wird $\Sigma_{11.2}$ wie folgt definiert:

$$\Sigma_{11.2} := \Sigma_{1,1} - \Sigma_{1,2} \Sigma_{2,2}^{-1} \Sigma_{2,1}$$

Somit hat man einen Ausdruck für die partielle Ableitung der Verteilungsfunktion nach $\rho_{n-1,n}$. Als nächstes betrachten wir die zur Verteilung $F_n(0; 0, \Sigma)$ gehörende Korrelationsmatrix. Diese sei mit \mathbf{R} bezeichnet. Ihre Elemente seien $\rho_{i,j}$. Durch die $\binom{n}{2}$ Nebendiagonalelemente ist in einem Raum der Dimension $\binom{n}{2}$ ein Punkt P mit den Koordinaten $\rho_{i,j}$ definiert. Da wir die Verteilungsfunktion an der Stelle $\mathbf{0}$ auswerten wollen, hängt der Wert von $F_n(\mathbf{0}; \mathbf{0}, \Sigma)$ nur von Σ bzw. von \mathbf{R} ab. Daher bezeichne $F_n(P)$ den Wert der Verteilungsfunktion einer zentrierten Normalverteilung mit der zum Punkt P gehörenden Korrelationsmatrix ausgewertet an der Stelle $\mathbf{0}$. Nehmen wir nun an, K sei ein Punkt mit den Koordinaten $\kappa_{i,j}$, für den $F_n(K)$ offensichtlich oder zumindest leicht zu berechnen sei. Weiter seien $\lambda_{i,j}$ die Koordinaten des Punktes L , der die Strecke KP im Verhältnis $t : 1 - t$ teilt. Dann läßt sich L als Linearkombination von K und P schreiben. Somit läßt sich folgende Kurve definieren:

$$\lambda_{i,j}(t) = t\rho_{i,j} + (1-t)\kappa_{i,j}$$

Bezeichne $\mathbf{R}(t)$ die Matrix mit den Nebendiagonaleinträgen $\{\lambda_{i,j}\}$ und 1 auf der Diagonale. Dann gehört $\mathbf{R}(0)$ zum Punkt K und $\mathbf{R}(1)$ zum Punkt P . Daher ist auch

$$\mathbf{R}(t) = t\mathbf{R}(1) + (1-t)\mathbf{R}(0)$$

positiv definit für $0 < t \leq 1$. Daher ist Gleichung (3.4) für jeden Zwischenpunkt der Strecke KP anwendbar. Mit dem Hauptsatz der Integration und Differentiation folgt:

$$F_n(P) = F_n(K) + \sum_{i < j} \int_{\kappa_{i,j}}^{\rho_{i,j}} \frac{\partial F_n}{\partial \lambda_{i,j}}(L) d\lambda_{i,j} \quad (3.6)$$

Als nächstes muß ein geeigneter Punkt K bestimmt werden.

Dazu soll $\mathbf{R}(0)$ als singuläre Matrix gewählt werden. Um aber nicht zuviele Summanden in (3.6)

zu erhalten, möchte man möglichst wenig, am besten nur einen Eintrag der Matrix verändern. Um dies zu erreichen, geht man wie folgt vor: Bezeichne $\mathbf{R}_{1,1}$ die ersten $(n-1)$ Spalten und Zeilen von \mathbf{R} . Die restliche Matrix sei wie immer zerlegt in $\mathbf{R}_{1,2}$, $\mathbf{R}_{2,1}$ und $\mathbf{R}_{2,2} = 1$. Da für uns der Eintrag $\rho_{n-1,n}$ eine besondere Rolle spielt, wird der Vektor $\mathbf{R}_{1,2}$ nochmals aufgespalten:

$$\mathbf{R}_{1,2} = \begin{pmatrix} \bar{\mathbf{R}}_{1,2} \\ \rho_{n-1,n} \end{pmatrix}$$

Damit hat die Matrix \mathbf{R} die folgende Gestalt:

$$\mathbf{R} = \begin{pmatrix} \mathbf{R}_{1,1} & \bar{\mathbf{R}}_{1,2} \\ \bar{\mathbf{R}}_{1,2}^T & \rho_{n-1,n} & 1 \end{pmatrix}$$

In dieser Matrix soll nun der Eintrag $\rho_{n-1,n}$ durch $\rho_{n-1,n}^*$ derart ersetzt werden, daß die so entstehende Matrix \mathbf{R}^* singulär wird. Zur Berechnung von $\rho_{n-1,n}^*$ dient der folgende Satz:

Satz 3.2 (Determinantenberechnung bei spezieller Matrixstruktur) *Sei \mathbf{R} wie oben in Teilmatrizen zerlegt. Dann gilt:*

$$|\mathbf{R}| = |\mathbf{R}_{1,1}| \cdot (1 - \mathbf{R}_{1,2}^T \mathbf{R}_{1,1}^{-1} \mathbf{R}_{1,2})$$

Beweis:

Zuerst zerlegt man \mathbf{R} mittels einer QR-Zerlegung in das folgende Matrixprodukt:

$$\mathbf{R} = \begin{pmatrix} \mathbf{R}_{1,1} & \mathbf{R}_{1,2} \\ \mathbf{R}_{1,2}^T & 1 \end{pmatrix} = \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{R}_{1,2}^T \mathbf{R}_{1,1}^{-1} & 1 \end{pmatrix} \begin{pmatrix} \mathbf{R}_{1,1} & \mathbf{R}_{1,2} \\ \mathbf{0} & 1 - \mathbf{R}_{1,2}^T \mathbf{R}_{1,1}^{-1} \mathbf{R}_{1,2} \end{pmatrix}$$

Die Determinante eines Matrixprodukts läßt sich als Produkt der Determinanten berechnen (falls wie hier beide Matrizen quadratisch sind), und somit gilt:

$$\begin{aligned} |\mathbf{R}| &= \underbrace{\left| \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{R}_{1,2}^T \mathbf{R}_{1,1}^{-1} & 1 \end{pmatrix} \right|}_{=1} \cdot \left| \begin{pmatrix} \mathbf{R}_{1,1} & \mathbf{R}_{1,2} \\ \mathbf{0} & 1 - \mathbf{R}_{1,2}^T \mathbf{R}_{1,1}^{-1} \mathbf{R}_{1,2} \end{pmatrix} \right| = \\ &= |\mathbf{R}_{1,1}| \cdot (1 - \mathbf{R}_{1,2}^T \mathbf{R}_{1,1}^{-1} \mathbf{R}_{1,2}) \end{aligned}$$

□

Wendet man diesen Satz auf obiges Problem an, so sieht man:

$\rho_{n-1,n}^*$ muß, um eine singuläre Matrix zu erhalten, derart gewählt werden, daß die Gleichung

$$\begin{pmatrix} \bar{\mathbf{R}}_{1,2} \\ \rho_{n-1,n}^* \end{pmatrix}^T \mathbf{R}_{1,1}^{-1} \begin{pmatrix} \bar{\mathbf{R}}_{1,2} \\ \rho_{n-1,n}^* \end{pmatrix} = 1$$

erfüllt ist. Schreibt man die obige Gleichung um, so ergibt sich:

$$\begin{aligned} &\left(\begin{pmatrix} \bar{\mathbf{R}}_{1,2} \\ \rho_{n-1,n}^* \end{pmatrix} - \begin{pmatrix} \bar{\mathbf{R}}_{1,2} \\ \rho_{n-1,n} \end{pmatrix} \right)^T \mathbf{R}_{1,1}^{-1} \left(\begin{pmatrix} \bar{\mathbf{R}}_{1,2} \\ \rho_{n-1,n}^* \end{pmatrix} - \begin{pmatrix} \bar{\mathbf{R}}_{1,2} \\ \rho_{n-1,n} \end{pmatrix} \right) + \\ &+ 2 \left(\begin{pmatrix} \bar{\mathbf{R}}_{1,2} \\ \rho_{n-1,n}^* \end{pmatrix} - \begin{pmatrix} \bar{\mathbf{R}}_{1,2} \\ \rho_{n-1,n} \end{pmatrix} \right)^T \mathbf{R}_{1,1}^{-1} \begin{pmatrix} \bar{\mathbf{R}}_{1,2} \\ \rho_{n-1,n} \end{pmatrix} - \left(1 - \begin{pmatrix} \bar{\mathbf{R}}_{1,2} \\ \rho_{n-1,n} \end{pmatrix}^T \mathbf{R}_{1,1}^{-1} \begin{pmatrix} \bar{\mathbf{R}}_{1,2} \\ \rho_{n-1,n} \end{pmatrix} \right) = 0 \end{aligned}$$

Faßt man noch die Terme zusammen, so ergibt sich:

$$\begin{aligned} & \begin{pmatrix} 0 \\ \rho_{n-1,n}^* - \rho_{n-1,n} \end{pmatrix}^T \mathbf{R}_{1,1}^{-1} \begin{pmatrix} 0 \\ \rho_{n-1,n}^* - \rho_{n-1,n} \end{pmatrix} + 2 \begin{pmatrix} 0 \\ \rho_{n-1,n}^* - \rho_{n-1,n} \end{pmatrix}^T \mathbf{R}_{1,1}^{-1} \begin{pmatrix} \bar{\mathbf{R}}_{1,2} \\ \rho_{n-1,n} \end{pmatrix} - \\ & - \left(1 - \begin{pmatrix} \bar{\mathbf{R}}_{1,2} \\ \rho_{n-1,n} \end{pmatrix}^T \mathbf{R}_{1,1}^{-1} \begin{pmatrix} \bar{\mathbf{R}}_{1,2} \\ \rho_{n-1,n} \end{pmatrix} \right) = 0 \end{aligned} \quad (3.7)$$

Bei diesem Ausdruck handelt es sich um eine quadratische Gleichung in $\rho_{n-1,n}^*$. Um nun die Anzahl der Lösungen bestimmen zu können, muß das Vorzeichen der Diskriminante bestimmt werden. Dies geschieht im folgenden Satz:

Satz 3.3 *Die Diskriminante der quadratischen Gleichung (3.7) ist unter den Voraussetzungen dieses Abschnitts stets positiv.*

Beweis:

Die Diskriminante D der allgemeinen quadratischen Gleichung $ax^2 + bx + c = 0$ ist wie folgt definiert: $D = b^2 - 4ac$

Dabei ist zu beachten, daß eine Verschiebung der Variablen $\rho_{n-1,n}^*$ um die Konstante $\rho_{n-1,n}$ keine Veränderung der Diskriminante bewirkt. Angewendet auf den konkreten Fall sind daher a , b und c dabei wie folgt definiert:

$$\begin{aligned} a & := (\mathbf{R}_{1,1}^{-1})_{n-1,n-1} \\ b & := 2 \left(\mathbf{R}_{1,1}^{-1} \begin{pmatrix} \bar{\mathbf{R}}_{1,2} \\ \rho_{n-1,n} \end{pmatrix} \right)_{n-1} \\ c & := - \left(1 - \begin{pmatrix} \bar{\mathbf{R}}_{1,2} \\ \rho_{n-1,n} \end{pmatrix}^T \mathbf{R}_{1,1}^{-1} \begin{pmatrix} \bar{\mathbf{R}}_{1,2} \\ \rho_{n-1,n} \end{pmatrix} \right) \end{aligned}$$

Betrachten wir zuerst den Ausdruck $a = (\mathbf{R}_{1,1}^{-1})_{n-1,n-1}$. Es handelt sich dabei um ein Diagonalelement der Matrix $\mathbf{R}_{1,1}^{-1}$. Wir wissen, daß die Matrix \mathbf{R} eine Korrelationsmatrix ist und somit positiv definit ist. Nach Definition der positiven Definitheit muß auch jede Hauptuntermatrix positiv definit sein. Daher ist auch $\mathbf{R}_{1,1}$ als eine solche Hauptuntermatrix positiv definit und mit dieser Matrix auch ihre Inverse $\mathbf{R}_{1,1}^{-1}$. Da es sich bei dem Ausdruck $a = (\mathbf{R}_{1,1}^{-1})_{n-1,n-1}$ somit um ein Diagonalelement einer positiv definiten Matrix handelt, ist $a = (\mathbf{R}_{1,1}^{-1})_{n-1,n-1} > 0$ garantiert.

Als nächstes soll der Ausdruck $c = - \left(1 - \begin{pmatrix} \bar{\mathbf{R}}_{1,2} \\ \rho_{n-1,n} \end{pmatrix}^T \mathbf{R}_{1,1}^{-1} \begin{pmatrix} \bar{\mathbf{R}}_{1,2} \\ \rho_{n-1,n} \end{pmatrix} \right)$ untersucht werden.

Nach Satz (3.2) gilt: $|\mathbf{R}| = -|\mathbf{R}_{1,1}| \cdot c$. Wie oben schon gezeigt, gilt aber $|\mathbf{R}| > 0$ sowie $|\mathbf{R}_{1,1}| > 0$.

Damit muß $c = - \left(1 - \begin{pmatrix} \bar{\mathbf{R}}_{1,2} \\ \rho_{n-1,n} \end{pmatrix}^T \mathbf{R}_{1,1}^{-1} \begin{pmatrix} \bar{\mathbf{R}}_{1,2} \\ \rho_{n-1,n} \end{pmatrix} \right) < 0$ sein.

Der Ausdruck für b spielt keine größere Rolle, da er sowieso nur als Quadrat in der Diskriminante auftaucht und somit $b^2 \geq 0$ gilt.

Insgesamt ergibt sich für die Diskriminante D die folgende Ungleichung:

$$\begin{aligned}
 D &= b^2 - 4ac = \\
 &= \underbrace{\left[2 \left(\mathbf{R}_{1,1}^{-1} \begin{pmatrix} \bar{\mathbf{R}}_{1,2} \\ \rho_{n-1,n} \end{pmatrix} \right)_{n-1} \right]^2}_{\geq 0} - 4 \underbrace{(\mathbf{R}_{1,1}^{-1})_{n-1,n-1}}_{> 0} \underbrace{\left(\begin{pmatrix} \bar{\mathbf{R}}_{1,2} \\ \rho_{n-1,n} \end{pmatrix}^T \mathbf{R}_{1,1}^{-1} \begin{pmatrix} \bar{\mathbf{R}}_{1,2} \\ \rho_{n-1,n} \end{pmatrix} - 1 \right)}_{< 0}
 \end{aligned}$$

Damit ist die Diskriminante der quadratischen Gleichung stets positiv. □

Mit Hilfe dieses Satzes wurde gezeigt, daß die quadratische Gleichung eine positive Diskriminante besitzt und daher immer zwei Lösungen für $\rho_{n-1,n}^*$ liefert. Bis jetzt wurde aber noch nicht gezeigt, daß diese Lösungen auch Korrelationskoeffizienten sind. Dazu fehlt noch der Beweis, daß $|\rho_{n-1,n}^*| \leq 1$ gilt. Dies soll aber auf konstruktivem Wege gezeigt werden. Dazu definieren wir die Zufallsvariable X_n^* wie folgt:

$$X_n^* := \psi_1 X_1 + \psi_2 X_2 + \dots + \psi_{n-1} X_{n-1}$$

wo

$$\mathbf{R}_{1,1} \psi = \begin{pmatrix} \bar{\mathbf{R}}_{1,2} \\ \rho_{n-1,n}^* \end{pmatrix}$$

Die Korrelationsmatrix der so erhaltenen Zufallsvariablen $\mathbf{X}^* = (X_1, X_2, \dots, X_n^*)^T$ kann leicht bestimmt werden. Die ersten $(n-1)$ Einträge der Zufallsvariablen \mathbf{X}^* wurden nicht verändert und somit ändern sich auch deren Korrelationskoeffizienten nicht. Einzig die letzte Zeile bzw. Spalte der Korrelationsmatrix ändert sich. Wie man leicht aus der Konstruktion von X_n^* einsehen kann, ergibt sich für die letzte Spalte genau der Vektor $\begin{pmatrix} \bar{\mathbf{R}}_{1,2} \\ \rho_{n-1,n}^* \end{pmatrix}$. Damit wurde auch gleichzeitig gezeigt, daß $\rho_{n-1,n}^*$ ein Korrelationskoeffizient ist und somit auch die Bedingung $|\rho_{n-1,n}^*| \leq 1$ erfüllt.

Bis jetzt wurde alles für den n -dimensionalen Fall gezeigt. Da aber für uns nur der Fall $n = 4$ interessant ist, sei jetzt also $n = 4$ vorausgesetzt. Damit ergeben sich einige Vereinfachungen, die die weiteren Rechnungen erleichtern.

So vereinfacht sich die quadratische Gleichung (3.7) und $\rho_{n-1,1}^*$ ergibt sich als Lösung von (verifiziert mittels MAPLE):

$$\begin{aligned}
 (1 - \rho_{1,2}^2) \rho_{3,4}^* &= \rho_{1,4}(\rho_{1,3} - \rho_{1,2}\rho_{2,3}) + \rho_{2,4}(\rho_{2,3} - \rho_{1,2}\rho_{1,3}) \\
 &\pm \{(1 - \rho_{1,2}^2 - \rho_{1,3}^2 - \rho_{2,3}^2 + 2\rho_{1,2}\rho_{1,3}\rho_{2,3})(1 - \rho_{1,2}^2 - \rho_{1,4}^2 - \rho_{2,4}^2 + 2\rho_{1,2}\rho_{1,4}\rho_{2,4})\}^{\frac{1}{2}}
 \end{aligned}$$

Die Koeffizienten von X_4^* erhält man durch Lösen des Gleichungssystems:

$$\begin{aligned}
 \psi_1 + \rho_{1,2}\psi_2 + \rho_{1,3}\psi_3 &= \rho_{1,4} \\
 \rho_{1,2}\psi_1 + \psi_2 + \rho_{2,3}\psi_3 &= \rho_{2,4} \\
 \rho_{1,3}\psi_1 + \rho_{2,3}\psi_2 + \psi_3 &= \rho_{3,4}^*
 \end{aligned}$$

und sie werden überprüft mittels der Bedingung:

$$\rho_{1,4}\psi_1 + \rho_{2,4}\psi_2 + \rho_{3,4}^*\psi_3 = 1$$

Um nun die Wahrscheinlichkeit $F_4(K)$ berechnen zu können, muß man noch die Vorzeichen von ψ_i betrachten. Dabei kommt es vor allem auf die Anzahl der positiven bzw. negativen Komponenten an. Zur weiteren Rechnung bezeichne S_1 die Menge der X_1 , die größer Null sind (S_2 und S_3 analog) und T die Menge mit X_4^* größer als Null. Ein Strich über einer Menge kennzeichnet dabei das Komplement. Zur Berechnung der Wahrscheinlichkeit wird dabei folgende Rechenregel angewendet: $P(A \cap B) = P(A) - P(A \cap \bar{B})$.

Es gilt somit:

- $\psi_1, \psi_2, \psi_3 > 0 \Rightarrow S_1 \cap S_2 \cap S_3 \cap \bar{T} = \emptyset$
 $\Rightarrow P(X_1 > 0, X_2 > 0, X_3 > 0, X_4^* > 0) =$
 $= P(X_1 > 0, X_2 > 0, X_3 > 0) - \underbrace{P(X_1 > 0, X_2 > 0, X_3 > 0, X_4^* < 0)}_{=0} =$
 $= P(X_1 > 0, X_2 > 0, X_3 > 0)$
- $\psi_1, \psi_2 > 0$ und $\psi_3 < 0 \Rightarrow \bar{S}_1 \cap \bar{S}_2 \cap S_3 \cap T = \emptyset$
 $\Rightarrow P(S_1 \cap S_2 \cap S_3 \cap T) = P(S_1 \cap S_3 \cap T) + P(S_2 \cap S_3 \cap T) - P(S_3 \cap T)$
 $\Rightarrow P(X_1 > 0, X_2 > 0, X_3 > 0, X_4^* > 0) =$
 $= P(X_1 > 0, X_3 > 0, X_4^* > 0) + P(X_2 > 0, X_3 > 0, X_4^* > 0) - P(X_3 > 0, X_4^* > 0)$
- $\psi_1 > 0$ und $\psi_2, \psi_3 < 0 \Rightarrow \bar{S}_1 \cap S_2 \cap S_3 \cap T = \emptyset$
 $\Rightarrow P(X_1 > 0, X_2 > 0, X_3 > 0, X_4^* > 0) =$
 $= P(X_2 > 0, X_3 > 0, X_4^* > 0) - \underbrace{P(X_1 > 0, X_2 > 0, X_3 > 0, X_4^* > 0)}_{=0} =$
 $= P(X_2 > 0, X_3 > 0, X_4^* > 0)$
- $\psi_1, \psi_2, \psi_3 < 0 \Rightarrow S_1 S_2 S_3 T = \emptyset$
 $\Rightarrow P(X_1 > 0, X_2 > 0, X_3 > 0, X_4^* > 0) = 0$

Bei anderer Reihenfolge der Vorzeichen ergibt sich eine analoge Rechnung. Dies soll aber hier nicht vorgerechnet werden, die Implementierung wird jedoch aus dem Algorithmus im Anhang ersichtlich. Somit kann $F_4(K)$ mit den Formeln für den niedrigdimensionalen Fall berechnet werden.

Wenden wir uns, nachdem $F_4(K)$ berechnet wurde, wieder der Formel (3.6) zu. Durch die geeignete Wahl von K hat sich die Summe über die Integrale auf lediglich einen Summanden reduziert und es gilt somit:

$$F_4(P) = F_4(K) + \int_{\rho_{3,4}^*}^{\rho_{3,4}} \frac{\partial F_4}{\partial \lambda_{3,4}}(L) d\lambda_{3,4} \quad (3.8)$$

Setzt man hier nun die bedingte Dichte ein, so ergibt sich folgender Ausdruck:

$$F_4(P) = F_4(K) + \int_{\rho_{3,4}^*}^{\rho_{3,4}} \left(\frac{1}{2\pi} (1 - \lambda_{3,4}^2)^{-\frac{1}{2}} \times \int_{-\infty}^0 \int_{-\infty}^0 f_{1|2}(x_3, x_4; \mathbf{0}, \boldsymbol{\Sigma}_{11.2}) dx_3 dx_4 \right) d\lambda_{3,4} \quad (3.9)$$

Bei der bedingten Dichte handelt es sich um eine zweidimensionale Normalverteilung und die Berechnung des Doppelintegrals entspricht der Wahrscheinlichkeit, daß beide Komponenten positiv sind. Somit ist die Rechnung analog zum Fall $S=2$ und man benötigt nur noch den Korrelationskoeffizienten. Dieser soll mittels des folgenden Satzes berechnet werden. Wie sich zeigen wird, hängt der gesuchte Korrelationskoeffizient auch noch von der Integrationsvariablen $\lambda_{3,4}$ ab.

Satz 3.4 (Bedingte Verteilungen) Sei $(x_1, x_2, x_3, x_4)^T \sim N(0, \Sigma)$ mit

$$\Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} & \sigma_{14} \\ \sigma_{12} & \sigma_{22} & \sigma_{23} & \sigma_{24} \\ \sigma_{13} & \sigma_{23} & \sigma_{33} & \sigma_{34} \\ \sigma_{14} & \sigma_{24} & \sigma_{34} & \sigma_{44} \end{pmatrix} \text{ und dazugehöriger Korrelationsmatrix } \rho = \begin{pmatrix} 1 & \rho_{12} & \rho_{13} & \rho_{14} \\ \rho_{12} & 1 & \rho_{23} & \rho_{24} \\ \rho_{13} & \rho_{23} & 1 & \rho_{34} \\ \rho_{14} & \rho_{24} & \rho_{34} & 1 \end{pmatrix}.$$

Sei weiter ρ_{12}^* der Korrelationskoeffizient der zweidimensionalen Normalverteilung von $(x_1, x_2)^T$ gegeben $(x_3, x_4)^T = (0, 0)^T$. Dann gilt:

Die bedingte Normalverteilung ist zentriert und

$$\rho_{12}^* = \begin{vmatrix} \rho_{12} & \rho_{13} & \rho_{14} \\ \rho_{23} & 1 & \rho_{34} \\ \rho_{24} & \rho_{34} & 1 \end{vmatrix} \cdot \left(\begin{vmatrix} 1 & \rho_{23} & \rho_{24} \\ \rho_{23} & 1 & \rho_{34} \\ \rho_{24} & \rho_{34} & 1 \end{vmatrix} \cdot \begin{vmatrix} 1 & \rho_{13} & \rho_{14} \\ \rho_{13} & 1 & \rho_{34} \\ \rho_{14} & \rho_{34} & 1 \end{vmatrix} \right)^{-1/2}$$

Beweis:

Nach Theorem 3.3.4 aus [To90] gilt:

$\mathbf{X} \sim N_n(\mu, \Sigma) \Rightarrow$ Für jedes beliebige, aber fest gewählte, $k < n$ ist die bedingte Dichte von \mathbf{X}_1 gegeben $\mathbf{X}_2 = \mathbf{x}_2$ normalverteilt mit $N_k(\mu_{1.2}, \Sigma_{11.2})$ wobei gilt:

$$\begin{aligned} \mu_{1.2} &:= \mu_1 + \Sigma_{12} \Sigma_{22}^{-1} (\mathbf{x}_2 - \mu_2) \\ \Sigma_{11.2} &:= \Sigma_{11} - \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21} \end{aligned}$$

Angewendet auf den Satz folgt daher mit $\mu = 0, n = 4, k = 2$ und $\mathbf{X}_2 = 0$:

$$\begin{aligned} \mu_{1.2} &= 0 + \Sigma_{12} \Sigma_{22}^{-1} (0 - 0) = 0 \\ \Sigma_{11.2} &= \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{pmatrix} - \begin{pmatrix} \sigma_{13} & \sigma_{14} \\ \sigma_{23} & \sigma_{24} \end{pmatrix} \begin{pmatrix} \sigma_{33} & \sigma_{34} \\ \sigma_{34} & \sigma_{44} \end{pmatrix}^{-1} \begin{pmatrix} \sigma_{13} & \sigma_{23} \\ \sigma_{14} & \sigma_{24} \end{pmatrix} = \\ &= \frac{1}{\sigma_{33}\sigma_{44} - \sigma_{34}^2} \cdot \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B} & \mathbf{C} \end{pmatrix} \end{aligned}$$

mit

$$\begin{aligned} \mathbf{A} &:= \sigma_{11}\sigma_{33}\sigma_{44} - \sigma_{11}\sigma_{34}^2 - \sigma_{13}^2\sigma_{44} + 2\sigma_{13}\sigma_{14}\sigma_{34} - \sigma_{14}^2\sigma_{33} \\ &= \begin{vmatrix} \sigma_{11} & \sigma_{13} & \sigma_{14} \\ \sigma_{13} & \sigma_{33} & \sigma_{34} \\ \sigma_{14} & \sigma_{34} & \sigma_{44} \end{vmatrix} \\ \mathbf{B} &:= \sigma_{12}\sigma_{33}\sigma_{44} - \sigma_{12}\sigma_{34}^2 - \sigma_{23}\sigma_{13}\sigma_{44} + \sigma_{23}\sigma_{14}\sigma_{34} + \sigma_{24}\sigma_{13}\sigma_{34} - \sigma_{24}\sigma_{14}\sigma_{33} \\ &= \begin{vmatrix} \sigma_{12} & \sigma_{13} & \sigma_{14} \\ \sigma_{23} & \sigma_{33} & \sigma_{34} \\ \sigma_{24} & \sigma_{34} & \sigma_{44} \end{vmatrix} \\ \mathbf{C} &:= \sigma_{22}\sigma_{33}\sigma_{44} - \sigma_{22}\sigma_{34}^2 - \sigma_{23}^2\sigma_{44} + 2\sigma_{23}\sigma_{24}\sigma_{34} - \sigma_{24}^2\sigma_{33} \\ &= \begin{vmatrix} \sigma_{22} & \sigma_{23} & \sigma_{24} \\ \sigma_{23} & \sigma_{33} & \sigma_{34} \\ \sigma_{24} & \sigma_{34} & \sigma_{44} \end{vmatrix} \end{aligned}$$

Damit gilt:

$$\rho_{12}^* = \frac{(\Sigma_{11 \cdot 2})_{1,2}}{\sqrt{(\Sigma_{11 \cdot 2})_{1,1} \cdot (\Sigma_{11 \cdot 2})_{2,2}}} = \frac{\frac{B}{\sigma_{33}\sigma_{44} - \sigma_{34}^2}}{\sqrt{\frac{A}{\sigma_{33}\sigma_{44} - \sigma_{34}^2} \cdot \frac{C}{\sigma_{33}\sigma_{44} - \sigma_{34}^2}}} = \frac{B}{\sqrt{A \cdot C}}$$

Der Satz ist damit schon fast bewiesen. Zum Schluß müssen nur noch die Kovarianzeinträge durch die Korrelationseinträge ersetzt werden. Hierzu benutzt man die Sätze über die Determinanten, im Speziellen:

$$\begin{aligned} |\lambda \mathbf{a} \ \mathbf{E}| &= \lambda \cdot |\mathbf{a} \ \mathbf{E}| \\ \begin{vmatrix} \lambda \mathbf{a} \\ \mathbf{E} \end{vmatrix} &= \lambda \cdot \begin{vmatrix} \mathbf{a} \\ \mathbf{E} \end{vmatrix} \end{aligned}$$

mit \mathbf{a} beliebiger Zeilen- bzw. Spaltenvektor und \mathbf{E} beliebige (von den Dimensionen passende) Matrix.

Damit gilt für \mathbf{A} :

$$\begin{aligned} \mathbf{A} &= \begin{vmatrix} \sigma_{11} & \sigma_{13} & \sigma_{14} \\ \sigma_{13} & \sigma_{33} & \sigma_{34} \\ \sigma_{14} & \sigma_{34} & \sigma_{44} \end{vmatrix} = \begin{vmatrix} 1 \cdot \sqrt{\sigma_{11} \cdot \sigma_{11}} & \rho_{13} \cdot \sqrt{\sigma_{33} \cdot \sigma_{11}} & \rho_{14} \cdot \sqrt{\sigma_{11} \cdot \sigma_{44}} \\ \rho_{13} \cdot \sqrt{\sigma_{11} \cdot \sigma_{33}} & 1 \cdot \sqrt{\sigma_{33} \cdot \sigma_{33}} & \rho_{34} \cdot \sqrt{\sigma_{33} \cdot \sigma_{44}} \\ \rho_{14} \cdot \sqrt{\sigma_{11} \cdot \sigma_{44}} & \rho_{34} \cdot \sqrt{\sigma_{33} \cdot \sigma_{44}} & 1 \cdot \sqrt{\sigma_{44} \cdot \sigma_{44}} \end{vmatrix} = \\ \text{spaltenweise} &= \frac{1}{\sqrt{\sigma_{11}}} \cdot \frac{1}{\sqrt{\sigma_{33}}} \cdot \frac{1}{\sqrt{\sigma_{44}}} \cdot \begin{vmatrix} 1 \cdot \sqrt{\sigma_{11}} & \rho_{13} \cdot \sqrt{\sigma_{11}} & \rho_{14} \cdot \sqrt{\sigma_{11}} \\ \rho_{13} \cdot \sqrt{\sigma_{33}} & 1 \cdot \sqrt{\sigma_{33}} & \rho_{34} \cdot \sqrt{\sigma_{33}} \\ \rho_{14} \cdot \sqrt{\sigma_{44}} & \rho_{34} \cdot \sqrt{\sigma_{44}} & 1 \cdot \sqrt{\sigma_{44}} \end{vmatrix} = \\ \text{zeilenweise} &= \frac{1}{\sigma_{11}} \cdot \frac{1}{\sigma_{33}} \cdot \frac{1}{\sigma_{44}} \cdot \begin{vmatrix} 1 & \rho_{13} & \rho_{14} \\ \rho_{13} & 1 & \rho_{34} \\ \rho_{14} & \rho_{34} & 1 \end{vmatrix} \end{aligned}$$

Analog folgt:

$$\begin{aligned} \mathbf{B} &= \frac{1}{\sqrt{\sigma_{11}}} \cdot \frac{1}{\sqrt{\sigma_{22}}} \cdot \frac{1}{\sigma_{33}} \cdot \frac{1}{\sigma_{44}} \cdot \begin{vmatrix} \rho_{12} & \rho_{13} & \rho_{14} \\ \rho_{23} & 1 & \rho_{34} \\ \rho_{24} & \rho_{34} & 1 \end{vmatrix} \\ \mathbf{C} &= \frac{1}{\sigma_{22}} \cdot \frac{1}{\sigma_{33}} \cdot \frac{1}{\sigma_{44}} \cdot \begin{vmatrix} 1 & \rho_{23} & \rho_{24} \\ \rho_{23} & 1 & \rho_{34} \\ \rho_{24} & \rho_{34} & 1 \end{vmatrix} \end{aligned}$$

Dies führt letztendlich zur Gleichung:

$$\rho_{12}^* = \frac{B}{\sqrt{A \cdot C}} = \frac{\frac{1}{\sqrt{\sigma_{11}}} \cdot \frac{1}{\sqrt{\sigma_{22}}} \cdot \frac{1}{\sigma_{33}} \cdot \frac{1}{\sigma_{44}} \cdot \begin{vmatrix} \rho_{12} & \rho_{13} & \rho_{14} \\ \rho_{23} & 1 & \rho_{34} \\ \rho_{24} & \rho_{34} & 1 \end{vmatrix}}{\sqrt{\frac{1}{\sigma_{11}} \cdot \frac{1}{\sigma_{33}} \cdot \frac{1}{\sigma_{44}}} \cdot \begin{vmatrix} 1 & \rho_{13} & \rho_{14} \\ \rho_{13} & 1 & \rho_{34} \\ \rho_{14} & \rho_{34} & 1 \end{vmatrix}^{1/2} \cdot \sqrt{\frac{1}{\sigma_{22}} \cdot \frac{1}{\sigma_{33}} \cdot \frac{1}{\sigma_{44}}} \cdot \begin{vmatrix} 1 & \rho_{23} & \rho_{24} \\ \rho_{23} & 1 & \rho_{34} \\ \rho_{24} & \rho_{34} & 1 \end{vmatrix}^{1/2}} =$$

$$= \frac{\begin{vmatrix} \rho_{12} & \rho_{13} & \rho_{14} \\ \rho_{23} & 1 & \rho_{34} \\ \rho_{24} & \rho_{34} & 1 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_{13} & \rho_{14} \\ \rho_{13} & 1 & \rho_{34} \\ \rho_{14} & \rho_{34} & 1 \end{vmatrix}^{1/2} \cdot \begin{vmatrix} 1 & \rho_{23} & \rho_{24} \\ \rho_{23} & 1 & \rho_{34} \\ \rho_{24} & \rho_{34} & 1 \end{vmatrix}^{1/2}}$$

□

Somit könnte man das Integral schon berechnen, jedoch kann der Ausdruck mittels einer Substitution noch vereinfacht werden. Man setzt in Gleichung (3.9)

$$\lambda_{3,4} := \cos(\theta)$$

und erhält:

$$\begin{aligned} F_4(P) &= F_4(K) + \int_{\rho_{3,4}^*}^{\rho_{3,4}} \left(\frac{1}{2\pi} (1 - \lambda_{3,4}^2)^{-\frac{1}{2}} \times \underbrace{\int_{-\infty}^0 \int_{-\infty}^0 f_{1|2}(x_3, x_4; \mathbf{0}, \Sigma_{11 \cdot 2}) dx_3 dx_4}_{:=g(\lambda_{3,4})} \right) d\lambda_{3,4} = \\ &= F_4(K) + \frac{1}{2\pi} \int_{\arccos(\rho_{3,4}^*)}^{\arccos(\rho_{3,4})} \frac{(-\sin(\theta)) \cdot g(\cos(\theta))}{\sqrt{1 - \cos^2(\theta)}} d\theta = \\ &= F_4(K) - \frac{1}{2\pi} \int_{\arccos(\rho_{3,4}^*)}^{\arccos(\rho_{3,4})} g(\cos(\theta)) d\theta \end{aligned}$$

Setzt man nun die Formel für den zweidimensionalen Fall sowie das Ergebnis aus dem zuvor bewiesenen Satz ein, so ergibt sich letztendlich:

$$F_4(P) = F_4(K) - \frac{1}{4\pi^2} \int_{\arccos(\rho_{3,4}^*)}^{\arccos(\rho_{3,4})} \arccos \left(\frac{\begin{vmatrix} \rho_{12} & \rho_{13} & \rho_{14} \\ \rho_{23} & 1 & \cos(\theta) \\ \rho_{24} & \cos(\theta) & 1 \end{vmatrix}}{\left(\begin{vmatrix} 1 & \rho_{23} & \rho_{24} \\ \rho_{23} & 1 & \cos(\theta) \\ \rho_{24} & \cos(\theta) & 1 \end{vmatrix} \cdot \begin{vmatrix} 1 & \rho_{13} & \rho_{14} \\ \rho_{13} & 1 & \cos(\theta) \\ \rho_{14} & \cos(\theta) & 1 \end{vmatrix} \right)^{1/2}} \right) d\theta$$

In Abbildung 3.3 ist das Ergebnis für $\lambda \in [0, 2]$ und für $\lambda \in [0, 5]$ zu sehen. In beiden Fällen wurde die additive Lösung der quadratischen Gleichung (3.7) gewählt.

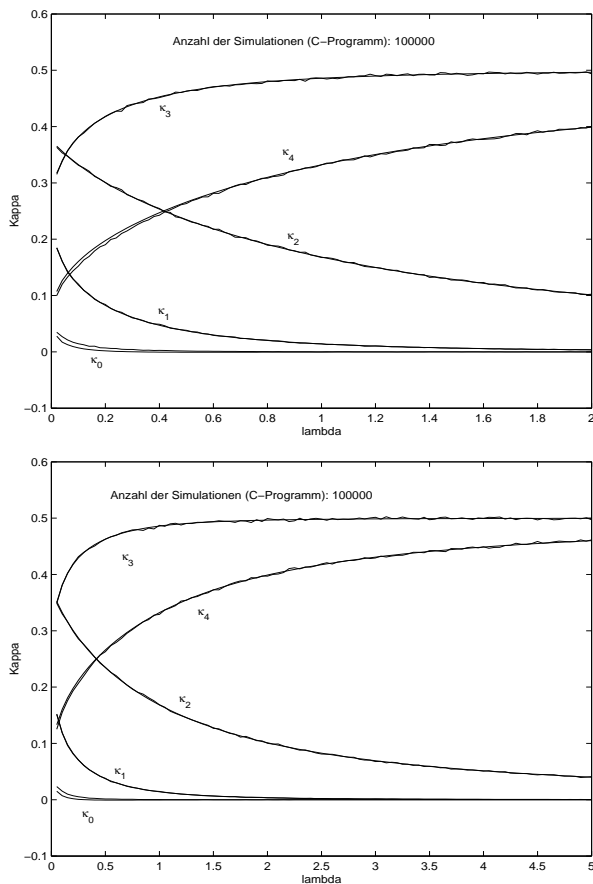
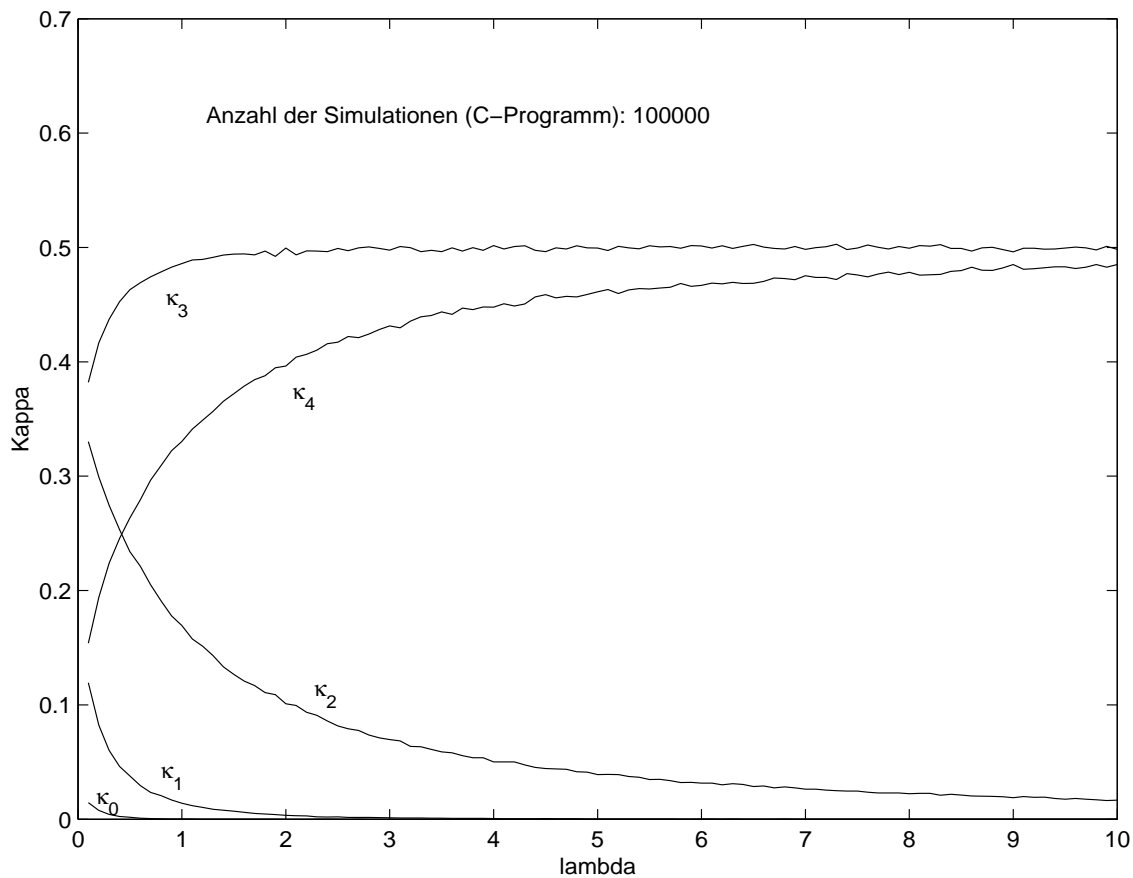


Abbildung 3.3: Vergleich der Simulation mit den berechneten Werten für $S=4$

Für $\lambda \in [0, 10]$ wurde kein Vergleich mehr geplottet. Der Algorithmus aus *matlab*, der zur Berechnung der Lösung diente, zeigte hier bereits einige Instabilitäten. Daher sind in Abbildung 3.4 lediglich die Simulationen für $\lambda \in [0, 10]$ geplottet.

3.2 Simulationsergebnisse für $S > 4$

Wie schon zu Beginn dieses Kapitels angedeutet, werden die Kappawerte für $S > 4$ aufgrund des enormen Rechenaufwands nicht mehr berechnet. An dieser Stelle werden lediglich noch die Simulationsergebnisse angeführt. Für $S = 5$ und $S = 6$ zeigt sich, daß zumindest im Bereich $\lambda \in (0, 10)$ noch keine Instabilitäten im Algorithmus auftreten (siehe Abbildungen 3.5 bzw. 3.6 auf den nächsten Seiten). Auch für $S = 7$ liefert der Algorithmus für $\lambda \in (0, 9)$ noch vernünftige Ergebnisse. Jedoch muß man hier bereits bei der unteren Schranke aufpassen, daß man nicht zu nahe an die Null herankommt, und ab $\lambda = 9$ gibt es ebenfalls Probleme. Dies äußert sich darin, daß der Algorithmus, der zur Findung der Indexmenge I verwendet wird, anfängt zu kreiseln und somit keine Lösung findet. Das Ergebnis dieser Simulation findet sich in Abbildung 3.7.

Abbildung 3.4: Simulation der Kappawerte für $S=4$

Insgesamt ist bei allen Kurven aufgefallen, daß sich anscheinend die κ -Werte, die zur χ_0^2 - und zur χ_1^2 -Verteilung gehören, asymptotisch gegen $\frac{1}{2}$ bewegen, während sich die restlichen κ -Werte gegen Null bewegen. Allerdings konnte dafür weder eine analytische Begründung gefunden werden noch eine, die dieses Phänomen aufgrund der Anschauung erklären könnte.

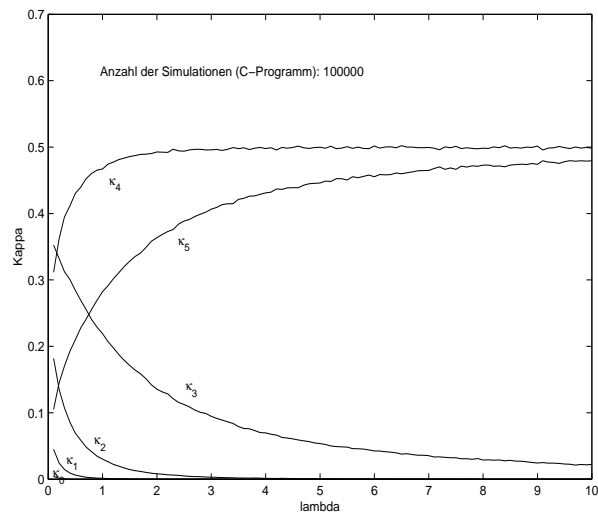
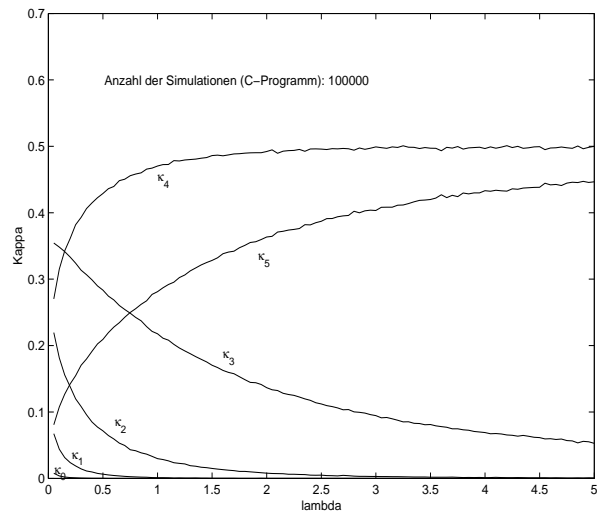
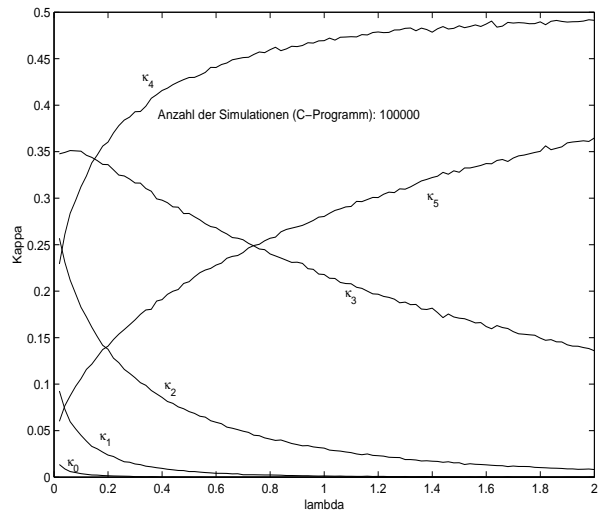


Abbildung 3.5: Simulation der Kappawerte für $S=5$

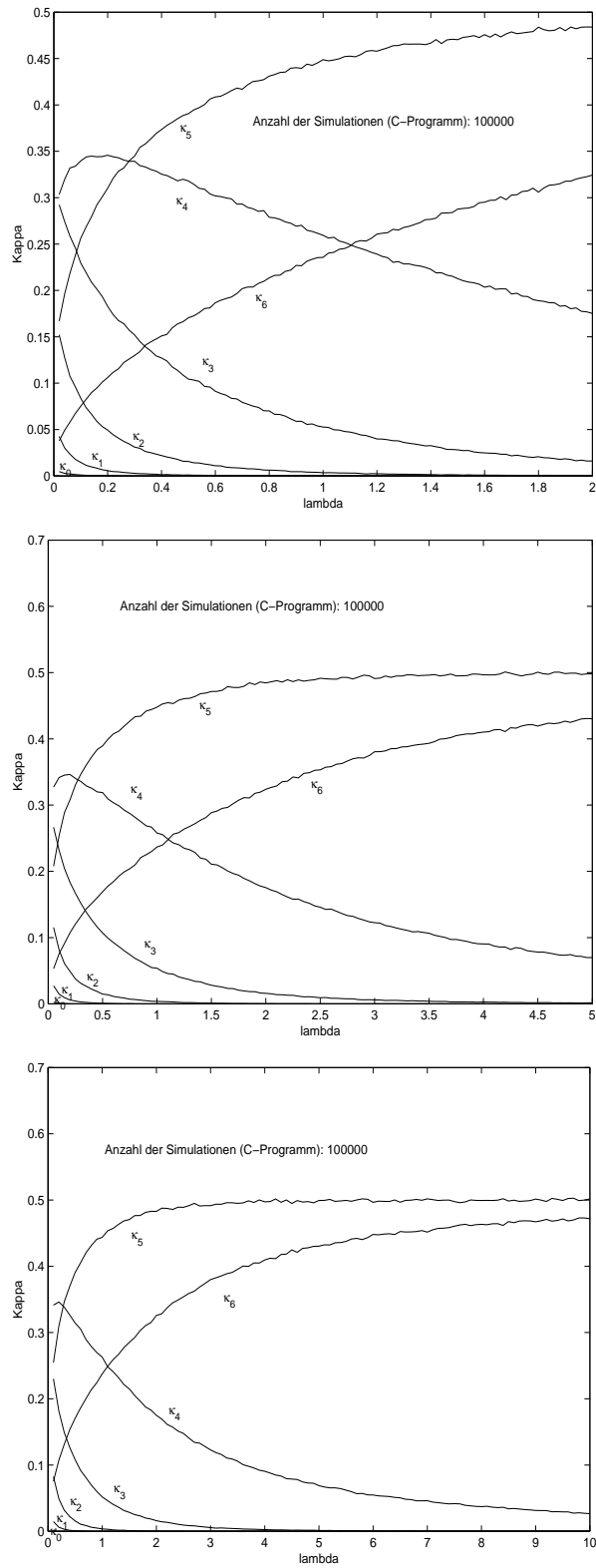


Abbildung 3.6: Simulation der Kappawerte für $S=6$

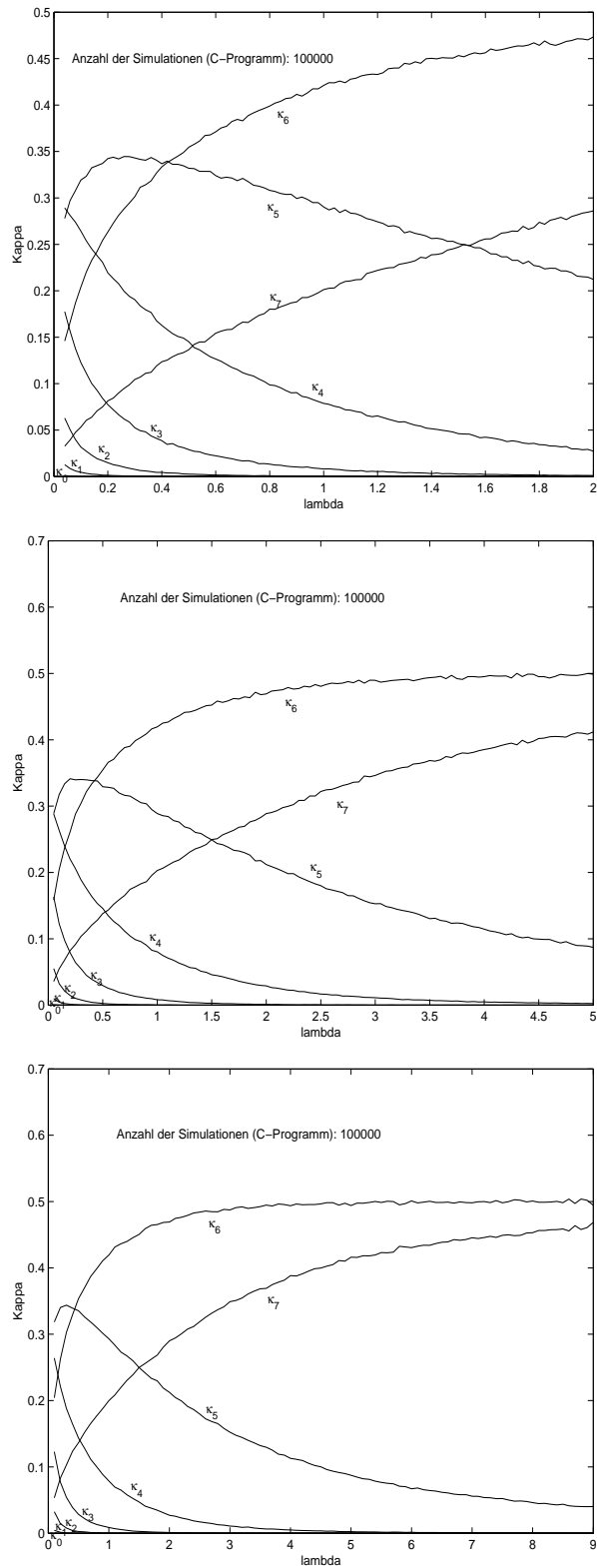


Abbildung 3.7: Simulation der Kappawerte für $S=7$

Kapitel 4

Praktische Implementierung des Tests

4.1 Programmierung in *matlab*

Dieses Kapitel stellt nun einen Algorithmus zur Verfügung, der den Test selbständig durchführt. Der Algorithmus dazu ist in *matlab* implementiert. Sollte jedoch der Anwender die Testgröße selber bestimmen wollen, so ist zu Beginn dieses Kapitels eine Zusammenfassung der kritischen Werte für $S = 2$ bis $S = 7$ gegeben. Dies ist jeweils für das 95% Quantil geschehen. Die Skala für λ wurde dabei in den jeweils drei zusammengehörenden Bildern unterteilt in die Intervalle $\lambda \in [0, 2]$, $\lambda \in [0, 5]$ und $\lambda \in [0, 10]$ (bzw. $\lambda \in [0, 9]$ für $S = 7$).

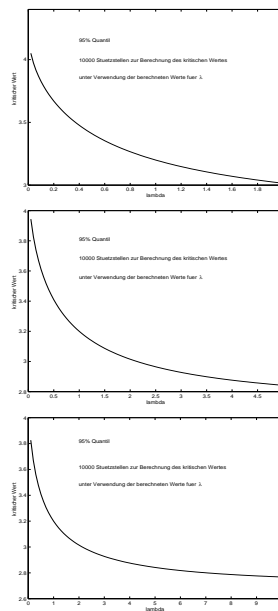


Abbildung 4.1: Abhängigkeit des kritischen Wertes von λ im Fall $S = 2$

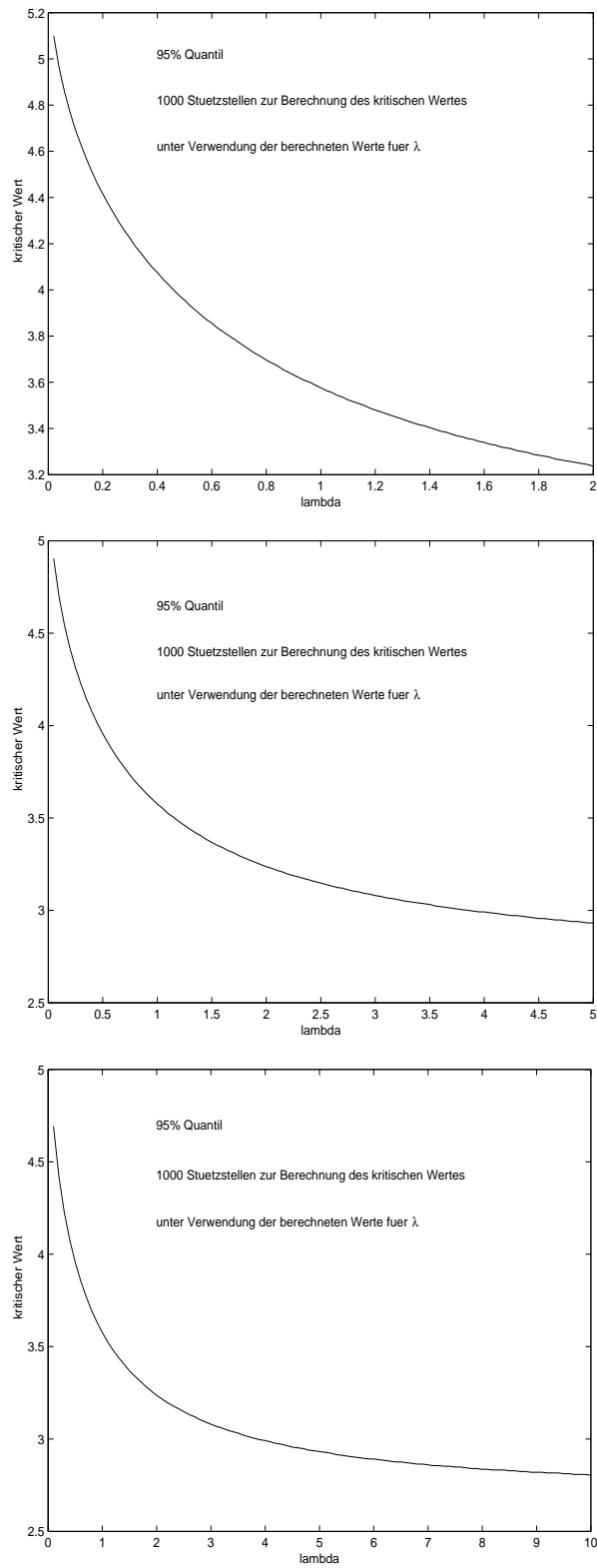


Abbildung 4.2: Abhängigkeit des kritischen Wertes von λ im Fall $S = 3$

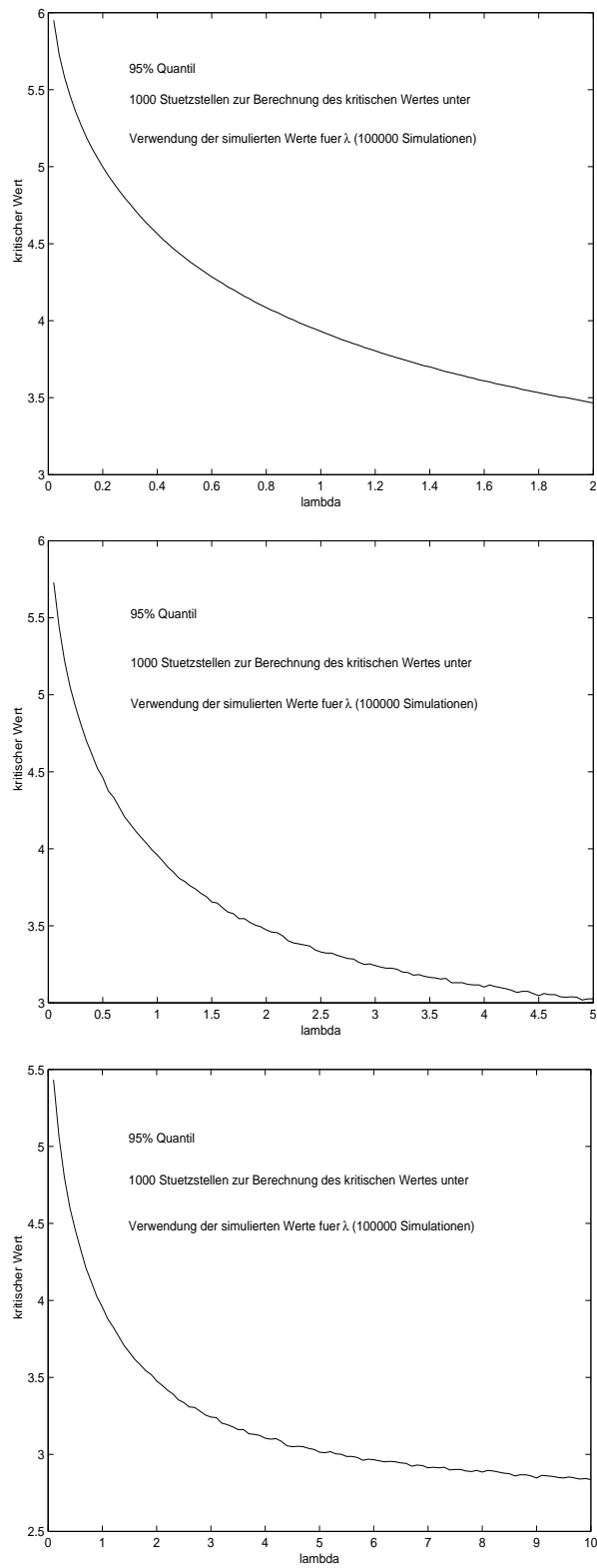


Abbildung 4.3: Abhängigkeit des kritischen Wertes von λ im Fall $S = 4$

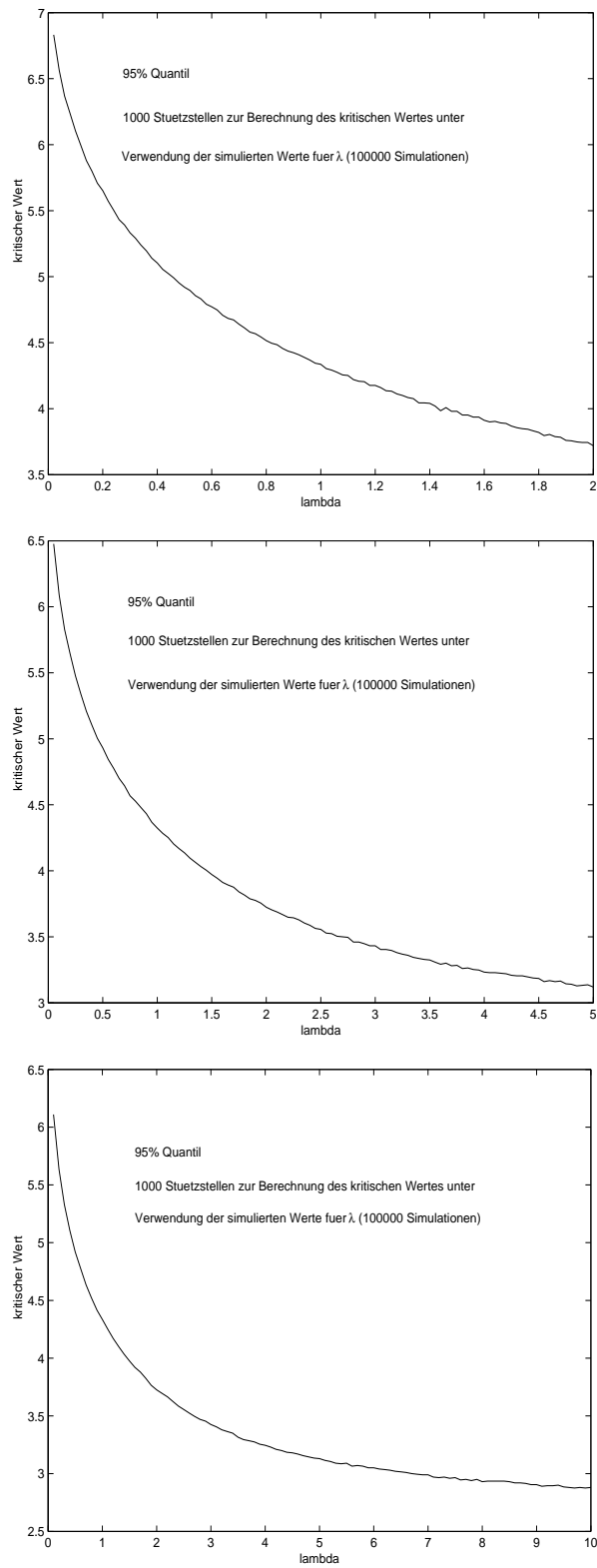


Abbildung 4.4: Abhängigkeit des kritischen Wertes von λ im Fall $S = 5$

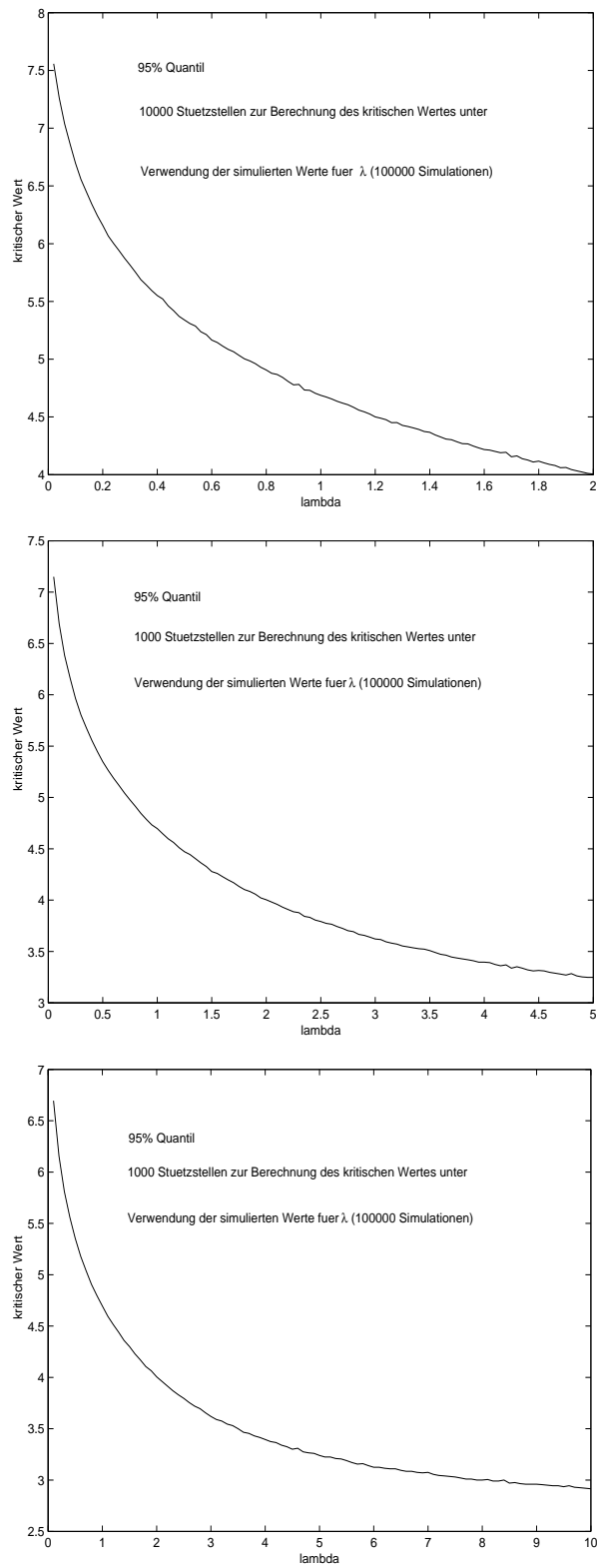


Abbildung 4.5: Abhängigkeit des kritischen Wertes von λ im Fall $S = 6$

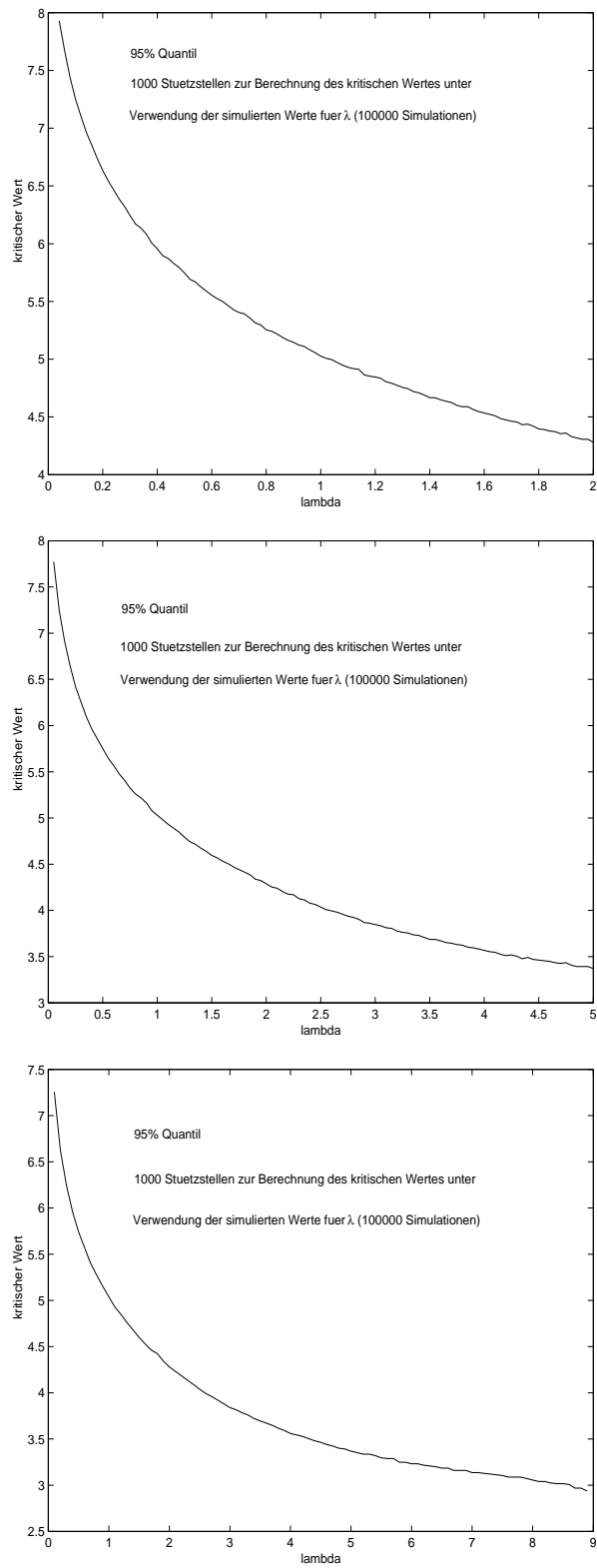


Abbildung 4.6: Abhängigkeit des kritischen Wertes von λ im Fall $S = 7$

Als nächstes wird nun ein Algorithmus zur Bestimmung der Testgröße benötigt. Dies bedeutet, daß der folgende Ausdruck bestimmt werden muß:

$$\begin{aligned}
T_1^n &= 2n \log \frac{\max_{(\lambda,p)^T \in K_1} \prod_{i=1}^n \Psi_{z_i}^S(\lambda,p)}{\max_{\lambda \in [0,\infty]} \prod_{i=1}^n \Psi_{z_i}^S(\lambda,1,0,\dots,0)} = \\
&= 2n \left[\log \left(\max_{(\lambda,p)^T \in K_1} \prod_{i=1}^n \Psi_{z_i}^S(\lambda,p) \right) - \log \left(\max_{\lambda \in [0,\infty]} \prod_{i=1}^n \Psi_{z_i}^S(\lambda,1,0,\dots,0) \right) \right] \quad (4.1)
\end{aligned}$$

Dieser Ausdruck soll nun mit Hilfe der Newton-Raphson Methode bestimmt werden. Dieser iterative Ansatz beginnt mit einem zulässigen Punkt $\theta_j \in \Theta$ und berechnet eine Approximation zweiter Ordnung der zu maximierenden Funktion $\Lambda(\theta)$:

$$\begin{aligned}
\Lambda(\theta) &\approx \Lambda(\theta_j) + D\Lambda_{\theta_j}(\theta - \theta_j) + \frac{1}{2}(\theta - \theta_j)^T D^2\Lambda_{\theta_j}(\theta - \theta_j) = \\
&= b_j + D\Lambda_{\theta_j}\theta - \theta_j^T D^2\Lambda_{\theta_j}\theta + \frac{1}{2}\theta^T D^2\Lambda_{\theta_j}\theta = \\
&= b_j + (\text{grad}\Lambda_{\theta_j} - D^2\Lambda_{\theta_j}\theta_j)^T\theta + \frac{1}{2}\theta^T D^2\Lambda_{\theta_j}\theta
\end{aligned}$$

wo b_j nicht von θ abhängt. Damit ergibt sich der neue Parameter aus:

$$\theta_{j+1} = \arg \max_{\theta \in \Theta} \left\{ (\text{grad}\Lambda_{\theta_j} - D^2\Lambda_{\theta_j}\theta_j)^T\theta + \frac{1}{2}\theta^T D^2\Lambda_{\theta_j}\theta \right\}$$

Betrachtet man nochmals die Gleichung (4.1), so sieht man, daß der Logarithmus einer zu maximierenden Funktion zu bestimmen ist. Da der Logarithmus eine streng monoton wachsende Funktion ist, kann das Maximum mit dem Logarithmus vertauscht werden.

Um die weitere Rechnung übersichtlich zu gestalten, werden die nachfolgenden Abkürzungen eingeführt:

$$\begin{aligned}
\theta_0 &:= \lambda \\
\theta_k &:= p_k \quad (1 \leq k \leq S) \\
\theta &:= (\theta_0, \theta_1, \dots, \theta_S)^T \\
a &:= 1 - \sum_{k=0}^S \Psi_k^S(\theta) \\
b_k &:= \frac{q_{n,k}}{\Psi_k^S(\theta)} - \frac{q_{n,S+1}}{a} \\
r_0 &:= 0 \\
s_0 &:= 0 \\
r_k(\theta) &:= \sum_{l=1}^k \theta_l \Psi_{k-l}^S(\theta) \quad (1 \leq k \leq S) \\
s_k(\theta) &:= \sum_{l=1}^k \theta_l r_{k-l}(\theta) \quad (1 \leq k \leq S)
\end{aligned}$$

Untersuchen wir zunächst den ersten Ausdruck aus Gleichung (4.1). Für ihn gilt:

$$\begin{aligned}\Lambda &:= \log \prod_{i=1}^n \Psi_{z_i}^S(\theta) = \log \prod_{k=0}^{S+1} (\Psi_k^S(\theta))^{q_{n,k}} = \\ &= \sum_{k=0}^{S+1} \log(\Psi_k^S(\theta))^{q_{n,k}} = \sum_{k=0}^{S+1} q_{n,k} \log(\Psi_k^S(\theta))\end{aligned}$$

Dabei steht $q_{n,k}$ für die absolute Häufigkeit. Als nächstes ist der Gradient und die Hessematrix von Λ zu bestimmen:

$$\begin{aligned}\partial_i \Lambda &= \sum_{k=0}^{S+1} q_{n,k} \frac{\partial_i \Psi_k^S(\theta)}{\Psi_k^S(\theta)} = \sum_{k=0}^S q_{n,k} \frac{\partial_i \Psi_k^S(\theta)}{\Psi_k^S(\theta)} + q_{n,S+1} \frac{\partial_i (1 - \sum_{k=0}^S \Psi_k^S(\theta))}{1 - \sum_{k=0}^S \Psi_k^S(\theta)} = \\ &= \sum_{k=0}^S \partial_i \Psi_k^S(\theta) \left(\frac{q_{n,k}}{\Psi_k^S(\theta)} - \frac{q_{n,S+1}}{1 - \sum_{k=0}^S \Psi_k^S(\theta)} \right) = \\ &= \sum_{k=0}^S b_k \cdot \partial_i \Psi_k^S(\theta)\end{aligned}$$

Bereits hier benötigt man die partiellen Ableitungen der Funktion $\Psi_k^S(\theta)$. Da man für die Hessematrix auch die partiellen Ableitungen der Ordnung zwei benötigt, werden auch diese schon jetzt bestimmt. Aufgrund der Definition $\Psi_k = T_k \exp_*(\theta_0(\sum_{l \geq 1} \theta_l \delta_l - \delta_0))$ und der Definition $\Psi_k^S = \Psi_k$ für $k \leq S$, erhält man (T_k ist dabei erneut die kanonische Projektion auf die k -te Komponente) mittels einer formalen Differentiation:

$$\begin{aligned}\partial_i \Psi_k^S &= \partial_i \Psi_k = T_k \partial_{\theta_i} \exp_* \left(\theta_0 \left(\sum_{l \geq 1} \theta_l \delta_l - \delta_0 \right) \right) = \\ &= \theta_0 T_k \exp_* \left(\theta_0 \left(\sum_{l \geq 1} \theta_l \delta_l - \delta_0 \right) \right) * \delta_0 = \theta_0 \Psi_{k-i} = \theta_0 \Psi_{k-i}^S \\ &\quad (i \geq 1, \quad k \leq S)\end{aligned}$$

Dies liefert die folgenden partiellen Ableitungen:

$$\begin{aligned}\partial_0 \Psi_k^S &= \sum_{l=1}^k \theta_l \Psi_{k-l}^S - \Psi_k^S = r_k - \Psi_k^S, \\ \partial_j \Psi_k^S &= \theta_0 \Psi_{k-j}^S && 0 < j \leq k, \\ \partial_j \Psi_k^S &= 0 && j > k, \\ \partial_0^2 \Psi_k^S &= \sum_{l=1}^k \theta_l \sum_{m=1}^{k-l} \theta_m \Psi_{k-l-m}^S - 2 \sum_{l=1}^k \theta_l \Psi_{k-l}^S + \Psi_k^S \\ &= s_k - 2r_k + \Psi_k^S, \\ \partial_0 \partial_j \Psi_k^S &= \theta_0 \sum_{l=1}^{k-j} \theta_l \Psi_{k-j-l}^S + (1 - \theta_0) \Psi_{k-j}^S \\ &= \theta_0 r_{k-j} + (1 - \theta_0) \Psi_{k-j}^S && 0 < j \leq k, \\ \partial_i \partial_j \Psi_k^S &= \theta_0^2 \Psi_{k-i-j}^S && 1 \leq i, j, i+j \leq k, \\ \partial_i \partial_j \Psi_k^S &= 0 && k < i+j\end{aligned}$$

Dies wird nun in den zuvor berechneten Ausdruck für den Gradienten eingesetzt. Somit erhält man:

$$\begin{aligned}\partial_0\Lambda &= \sum_{k=0}^S b_k \cdot (r_k - \Psi_k^S(\theta)) \\ \partial_i\Lambda &= \theta_0 \sum_{k=i}^S b_k \cdot \Psi_{k-i}^S(\theta) \quad 0 < i \leq S\end{aligned}$$

Als nächstes muß nun die Hessematrix berechnet werden. Dabei liefert die formale Ableitung den folgenden Ausdruck:

$$\begin{aligned}\partial_i\partial_j\Lambda &= \sum_{k=0}^S b_k \cdot \partial_i\partial_j\Psi_k^S(\theta) - \sum_{k=0}^S \partial_i\Psi_k^S(\theta) \left(\frac{q_{n,k} \cdot \partial_j\Psi_k^S(\theta)}{(\Psi_k^S(\theta))^2} + \frac{q_{n,S+1} \sum_{k=0}^S \partial_j\Psi_k^S(\theta)}{a^2} \right) = \\ &= \sum_{k=0}^S b_k \cdot \partial_i\partial_j\Psi_k^S(\theta) - \sum_{k=0}^S q_{n,k} \frac{\partial_i\Psi_k^S(\theta) \cdot \partial_j\Psi_k^S(\theta)}{(\Psi_k^S(\theta))^2} - \left(\sum_{k=0}^S \partial_i\Psi_k^S(\theta) \right) \cdot \left(\sum_{k=0}^S \partial_j\Psi_k^S(\theta) \right) \cdot \frac{q_{n,S+1}}{a^2}\end{aligned}$$

Erneut werden noch die partiellen Ableitungen der Funktion Ψ_k^S eingesetzt und man erhält

$$\begin{aligned}\partial_0^2\Lambda &= \sum_{k=0}^S b_k \cdot \left(s_k - 2r_k + \Psi_k^S(\theta) \right) - \sum_{k=0}^S q_{n,k} \left(\frac{r_k - \Psi_k^S(\theta)}{\Psi_k^S(\theta)} \right)^2 - \left(\sum_{k=0}^S r_k - \Psi_k^S(\theta) \right)^2 \cdot \frac{q_{n,S+1}}{a^2} = \\ &= \sum_{k=0}^S b_k \cdot \left(s_k - 2r_k + \Psi_k^S(\theta) \right) - \sum_{k=0}^S q_{n,k} \left(\frac{r_k}{\Psi_k^S(\theta)} - 1 \right)^2 - \left(\sum_{k=0}^S r_k - \Psi_k^S(\theta) \right)^2 \cdot \frac{q_{n,S+1}}{a^2} = \\ &= \sum_{k=0}^S q_{n,k} \frac{s_k}{\Psi_k^S(\theta)} - 2 \sum_{k=0}^S q_{n,k} \frac{r_k}{\Psi_k^S(\theta)} + \sum_{k=0}^S q_{n,k} - \frac{q_{n,S+1}}{a} \sum_{k=0}^S s_k + 2 \frac{q_{n,S+1}}{a} \sum_{k=0}^S r_k - \\ &\quad - \frac{q_{n,S+1}}{a} \sum_{k=0}^S \Psi_k^S(\theta) - \sum_{k=0}^S q_{n,k} \left(\frac{r_k}{\Psi_k^S(\theta)} \right)^2 + 2 \sum_{k=0}^S q_{n,k} \frac{r_k}{\Psi_k^S(\theta)} - \sum_{k=0}^S q_{n,k} - \\ &\quad - \frac{q_{n,S+1}}{a^2} \left(\sum_{k=0}^S r_k \right)^2 + 2 \frac{q_{n,S+1}}{a^2} \sum_{k=0}^S r_k \sum_{k=0}^S \Psi_k^S(\theta) - \frac{q_{n,S+1}}{a^2} \left(\sum_{k=0}^S \Psi_k^S(\theta) \right)^2 = \\ &= - \sum_{k=0}^S q_{n,k} \left(\frac{r_k}{\Psi_k^S(\theta)} \right)^2 + \sum_{k=0}^S q_{n,k} \frac{s_k}{\Psi_k^S(\theta)} + \\ &\quad + \frac{q_{n,S+1}}{a^2} \left[\left(1 - \sum_{k=0}^S \Psi_k^S(\theta) \right) \cdot \left(- \sum_{k=0}^S s_k + 2 \sum_{k=0}^S r_k - \sum_{k=0}^S \Psi_k^S(\theta) \right) - \right. \\ &\quad \left. - \left(\sum_{k=0}^S r_k \right)^2 + 2 \sum_{k=0}^S r_k \sum_{k=0}^S \Psi_k^S(\theta) - \left(\sum_{k=0}^S \Psi_k^S(\theta) \right)^2 \right] = \\ &= - \sum_{k=0}^S q_{n,k} \left(\frac{r_k}{\Psi_k^S(\theta)} \right)^2 + \sum_{k=0}^S q_{n,k} \frac{s_k}{\Psi_k^S(\theta)} +\end{aligned}$$

$$\begin{aligned}
& + \frac{q_{n,S+1}}{a^2} \left[\left(1 - \sum_{k=0}^S \Psi_k^S(\theta) \right) \cdot \sum_{k=0}^S (-s_k) - \left(\sum_{k=0}^S r_k \right)^2 + 2 \sum_{k=0}^S r_k - 1 + 1 - \sum_{k=0}^S \Psi_k^S(\theta) \right] = \\
& = - \sum_{k=0}^S q_{n,k} \left(\frac{r_k}{\Psi_k^S(\theta)} \right)^2 + \sum_{k=0}^S q_{n,k} \frac{s_k}{\Psi_k^S(\theta)} - \\
& \quad - \frac{q_{n,S+1}}{a^2} \cdot \left(1 - \sum_{k=0}^S r_k \right)^2 + \frac{q_{n,S+1}}{a} \cdot \left(1 - \sum_{k=0}^S s_k \right) \\
\partial_0 \partial_j \Lambda & = \sum_{k=j}^S b_k \cdot \left(\theta_0 r_{k-j} + (1 - \theta_0) \Psi_{k-j}^S(\theta) \right) - \theta_0 \sum_{k=j}^S q_{n,k} \frac{\left(r_k - \Psi_k^S(\theta) \right) \cdot \Psi_{k-j}^S(\theta)}{\left(\Psi_k^S(\theta) \right)^2} - \\
& \quad - \theta_0 \cdot \left(\sum_{k=0}^S r_k - \Psi_k^S(\theta) \right) \cdot \left(\sum_{k=j}^S \Psi_{k-j}^S(\theta) \right) \cdot \frac{q_{n,S+1}}{a^2} = \\
& = \theta_0 \cdot \sum_{k=j}^S q_{n,k} \frac{r_{k-j}}{\Psi_k^S(\theta)} + (1 - \theta_0) \cdot \sum_{k=j}^S q_{k,n} \frac{\Psi_{k-j}^S(\theta)}{\Psi_k^S(\theta)} - \theta_0 \cdot \frac{q_{n,S+1}}{a} \cdot \sum_{k=j}^S r_{k-j} - \\
& \quad - (1 - \theta_0) \frac{q_{n,S+1}}{a} \cdot \sum_{k=j}^S \Psi_{k-j}^S(\theta) - \theta_0 \cdot \sum_{k=j}^S q_{n,k} \frac{r_k \Psi_{k-j}^S(\theta)}{\left(\Psi_k^S(\theta) \right)^2} + \theta_0 \cdot \sum_{k=j}^S q_{n,k} \frac{\Psi_{k-j}^S(\theta)}{\Psi_k^S(\theta)} - \\
& \quad - \theta_0 \cdot \frac{q_{n,S+1}}{a^2} \cdot \left(\sum_{k=0}^S r_k \sum_{k=j}^S \Psi_{k-j}^S(\theta) - \sum_{k=0}^S \Psi_k^S(\theta) \sum_{k=j}^S \Psi_{k-j}^S(\theta) \right) = \\
& = -\theta_0 \cdot \sum_{k=j}^S q_{n,k} \frac{r_k \Psi_{k-j}^S(\theta)}{\left(\Psi_k^S(\theta) \right)^2} + \sum_{k=j}^S q_{k,n} \frac{\Psi_{k-j}^S(\theta)}{\Psi_k^S(\theta)} + \theta_0 \cdot \sum_{k=j}^S q_{n,k} \frac{r_{k-j}}{\Psi_k^S(\theta)} + \\
& \quad + \frac{q_{n,S+1}}{a^2} \left[\left(1 - \Psi_k^S(\theta) \right) \cdot \left(-\theta_0 \sum_{k=j}^S r_{k-j} - (1 - \theta_0) \sum_{k=j}^S \Psi_{k-j}^S(\theta) \right) - \right. \\
& \quad \quad \left. - \theta_0 \sum_{k=0}^S r_k \sum_{k=j}^S \Psi_{k-j}^S(\theta) + \theta_0 \sum_{k=0}^S \Psi_k^S(\theta) \sum_{k=j}^S \Psi_{k-j}^S(\theta) \right] = \\
& = -\theta_0 \cdot \sum_{k=j}^S q_{n,k} \frac{r_k \Psi_{k-j}^S(\theta)}{\left(\Psi_k^S(\theta) \right)^2} + \sum_{k=j}^S q_{k,n} \frac{\Psi_{k-j}^S(\theta)}{\Psi_k^S(\theta)} + \theta_0 \cdot \sum_{k=j}^S q_{n,k} \frac{r_{k-j}}{\Psi_k^S(\theta)} + \\
& \quad + \frac{q_{n,S+1}}{a^2} \left[-\theta_0 \cdot \left(1 - \sum_{k=0}^S \Psi_k(\theta) \right) \sum_{k=j}^S r_{k-j} + \right. \\
& \quad \quad \left. + \sum_{k=j}^S \Psi_{k-j}(\theta) \cdot \left(-\theta_0 \sum_{k=0}^S r_k + \theta_0 - 1 + \sum_{k=0}^S \Psi_k(\theta) \right) \right] = \\
& = -\theta_0 \cdot \sum_{k=j}^S q_{n,k} \frac{r_k \Psi_{k-j}^S(\theta)}{\left(\Psi_k^S(\theta) \right)^2} + \sum_{k=j}^S q_{k,n} \frac{\Psi_{k-j}^S(\theta)}{\Psi_k^S(\theta)} + \theta_0 \cdot \sum_{k=j}^S q_{n,k} \frac{r_{k-j}}{\Psi_k^S(\theta)} -
\end{aligned}$$

$$\begin{aligned}
& - \left(-\theta_0 \cdot \frac{q_{n,S+1}}{a^2} \left(1 - \sum_{k=0}^S r_k\right) \sum_{k=j}^S \Psi_{k-j} + \frac{q_{n,S+1}}{a} \sum_{k=j}^S \Psi_{k-j}(\theta) + \theta_0 \cdot \frac{q_{n,S+1}}{a} \sum_{k=j}^S r_{k-j} \right) \\
\partial_i \partial_j \Lambda &= \theta_0^2 \sum_{k=i+j}^S b_k \cdot \Psi_{k-i-j}^S(\theta) - \theta_0^2 \sum_{k=i \vee j}^S q_{n,k} \frac{\Psi_{k-j}^S(\theta) \Psi_{k-i}^S(\theta)}{(\Psi_k^S(\theta))^2} - \\
& - \theta_0^2 \left(\sum_{k=i}^S \Psi_{k-i}^S(\theta) \right) \cdot \left(\sum_{k=j}^S \Psi_{k-j}^S(\theta) \right) \cdot \frac{q_{n,S+1}}{a^2} = \\
&= \theta_0^2 \cdot \sum_{k=i+j}^S q_{n,k} \frac{\Psi_{k-i-j}^S(\theta)}{\Psi_k^S(\theta)} - \theta_0^2 \cdot \sum_{k=i \vee j}^S q_{n,k} \frac{\Psi_{k-i}^S(\theta) \Psi_{k-j}^S(\theta)}{(\Psi_k^S(\theta))^2} - \\
& - \theta_0^2 \cdot \frac{q_{n,S+1}}{a} \cdot \sum_{k=i+j}^S \Psi_{k-i-j}^S(\theta) - \theta_0^2 \cdot \frac{q_{n,S+1}}{a^2} \cdot \sum_{k=i}^S \Psi_{k-i}^S(\theta) \cdot \sum_{k=j}^S \Psi_{k-j}^S(\theta)
\end{aligned}$$

Als nächstes betrachten wird den zweiten Ausdruck aus Gleichung (4.1). Auch hier sollen zuerst die nachfolgenden Abkürzungen eingeführt werden, um den Ausdruck übersichtlich zu halten:

$$\begin{aligned}
c_1 &:= \sum_{k=0}^S q_{n,k} \cdot k \\
c_2 &:= \sum_{k=0}^S q_{n,k} \cdot \log(k!) \\
c_3(\lambda) &:= e^\lambda - \sum_{k=0}^S \frac{\lambda^k}{k!}
\end{aligned}$$

Somit gilt für den gesuchten Ausdruck:

$$\begin{aligned}
& \log \prod_{i=1}^n \Psi_{Z_i}^S(\theta) = \\
&= \sum_{k=0}^{S+1} q_{n,k} \cdot \log(\Psi_k^S(\theta)) = \sum_{k=0}^S \left(q_{n,k} \cdot \log\left(e^{-\lambda} \frac{\lambda^k}{k!}\right) \right) + q_{n,S+1} \cdot \log\left(1 - \sum_{k=0}^S \Psi_k^S(\theta)\right) = \\
&= \sum_{k=0}^S q_{n,k} \cdot \underbrace{\log(e^{-\lambda})}_{=-\lambda} + \sum_{k=0}^S q_{n,k} \cdot \log(\lambda^k) - \underbrace{\sum_{k=0}^S q_{n,k} \log(k!)}_{=c_2} + q_{n,S+1} \cdot \log\left(1 - \sum_{k=0}^S \Psi_k^S(\theta)\right) = \\
&= -\lambda(n - q_{n,S+1}) + \log(\lambda) \underbrace{\sum_{k=0}^S (q_{n,k} \cdot k)}_{=c_1} - c_2 + q_{n,S+1} \cdot \log\left(1 - \sum_{k=0}^S \Psi_k^S(\theta)\right) = \\
&= -\lambda \cdot n + c_1 \cdot \log(\lambda) - c_2 + q_{n,S+1} \left(\log\left(1 - \sum_{k=0}^S e^{-\lambda} \frac{\lambda^k}{k!}\right) + \log(e^\lambda) \right) =
\end{aligned}$$

$$\begin{aligned}
&= -\lambda \cdot n + c_1 \cdot \log(\lambda) - c_2 + q_{n,S+1} \cdot \log \left(\underbrace{e^\lambda - \sum_{k=0}^S \frac{\lambda^k}{k!}}_{=c_3(\lambda)} \right) = \\
&= -\lambda \cdot n + c_1 \cdot \log(\lambda) - c_2 + q_{n,S+1} \log(c_3(\lambda))
\end{aligned}$$

Für die Ableitung der Funktion gilt:

$$\begin{aligned}
&\frac{d}{d\lambda} \left(-\lambda \cdot n + c_1 \cdot \log(\lambda) - c_2 + q_{n,S+1} \log(c_3(\lambda)) \right) = \\
&= -n + \frac{c_1}{\lambda} + q_{n,S+1} \frac{e^\lambda - \sum_{k=0}^{S-1} \frac{\lambda^k}{k!}}{e^\lambda - \sum_{k=0}^S \frac{\lambda^k}{k!}} = \\
&= -n + \frac{c_1}{\lambda} + q_{n,S+1} + \frac{q_{n,S+1}}{S!} \cdot \frac{\lambda^S}{e^\lambda - \sum_{k=0}^S \frac{\lambda^k}{k!}} = \\
&= q_{n,S+1} - n + \frac{c_1}{\lambda} + \frac{q_{n,S+1}}{S!} \cdot \frac{\lambda^S}{c_3(\lambda)}
\end{aligned}$$

Für die zweite Ableitung ergibt sich:

$$\begin{aligned}
&\frac{d}{d\lambda} \left(q_{n,S+1} - n + \frac{c_1}{\lambda} + \frac{q_{n,S+1}}{S!} \cdot \frac{\lambda^S}{c_3(\lambda)} \right) = \\
&= -\frac{c_1}{\lambda^2} + \frac{q_{n,S+1}}{S!} \cdot \frac{S \cdot c_3(\lambda) \cdot \lambda^{S-1} - \lambda^S (e^\lambda - \sum_{k=0}^{S-1} \frac{\lambda^k}{k!})}{c_3(\lambda)^2} = \\
&= -\frac{c_1}{\lambda^2} + \frac{q_{n,S+1} \cdot \lambda^{S-1}}{c_3(\lambda)^2 \cdot S!} \left(c_3 \cdot S - c_3 \cdot \lambda - \frac{\lambda^{S+1}}{S!} \right) = \\
&= -\frac{c_1}{\lambda^2} + \frac{q_{n,k} \cdot (S - \lambda) \cdot \lambda^{S-1}}{c_3(\lambda) \cdot S!} - q_{n,S+1} \left(\frac{\lambda^S}{c_3(\lambda) \cdot S!} \right)^2
\end{aligned}$$

Die so berechneten Ausdrücke werden nun in *matlab* benutzt, um mittels des dort implementierten Algorithmus *quadprog* iterativ die Lösung des Maximierungsproblems zu finden. Der zugehörige Algorithmus ist im Anhang zu finden.

Im letzten Schritt muß nun noch die Teststatistik mit dem kritischen Wert verglichen werden, um eine Aussage zu erhalten, ob die Hypothese zu einem bestimmten α -Niveau abgelehnt werden kann. Dabei kann der Anwender zwischen zwei Möglichkeiten wählen. Er kann erstens ein α vorgeben und dann den Algorithmus bestimmen lassen, ob zu diesem Quantil abgelehnt wird, oder sich zweitens den p -Wert für seine Testgröße bestimmen lassen.

Dazu wird jedoch noch die Verteilungsfunktion unter der Hypothese benötigt. Diese erhält man aus den κ -Werten, die im Algorithmus in [Bu01] bestimmt wurden. Es gilt schließlich:

$$T_n^1 \xrightarrow{D} \|\pi_D(Z|V) - \pi_D(Z|C_q)\|_D^2 \sim \sum_{i=0}^S \kappa_i \chi_{S-i}^2$$

Da allerdings nicht bei jedem Aufruf der Funktion die κ -Werte neu simuliert werden sollen, wurden im Intervall von $[0,2]$, $[0,5]$ bzw. $[0,10]$ die Kurven der jeweiligen Kappawerte durch eine Ausgleichsrechnung angenähert. Dabei wurde ein Polynom 8. Grades benutzt. Dies wurde für

$S \in \{2, 3, \dots, 7\}$ durchgeführt. Nachfolgend sind die entsprechenden Plots. Darin wurden, soweit im vorangehenden Kapitel bestimmt, die berechneten κ -Werte benutzt, ansonsten wurden erneut die κ -Werte benutzt, die der Simulationsalgorithmus bei 100000 Simulationen liefert. Aus den Graphiken erkennt man, daß die Ausgleichspolynome nur unmerkliche Abweichungen aufweisen. Im Anhang sind die entsprechend umgesetzten Algorithmen zu finden.

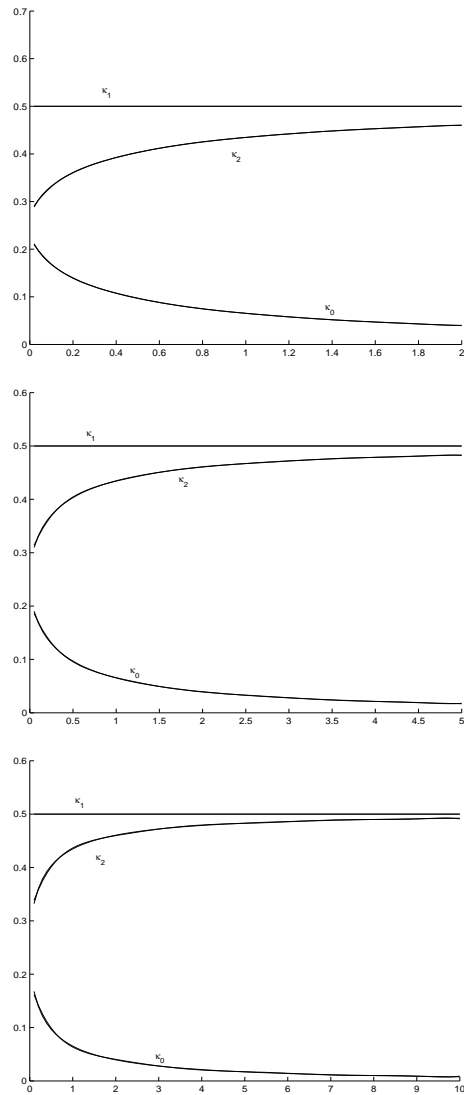


Abbildung 4.7: Ausgleichspolynom 8. Grades für $S=2$

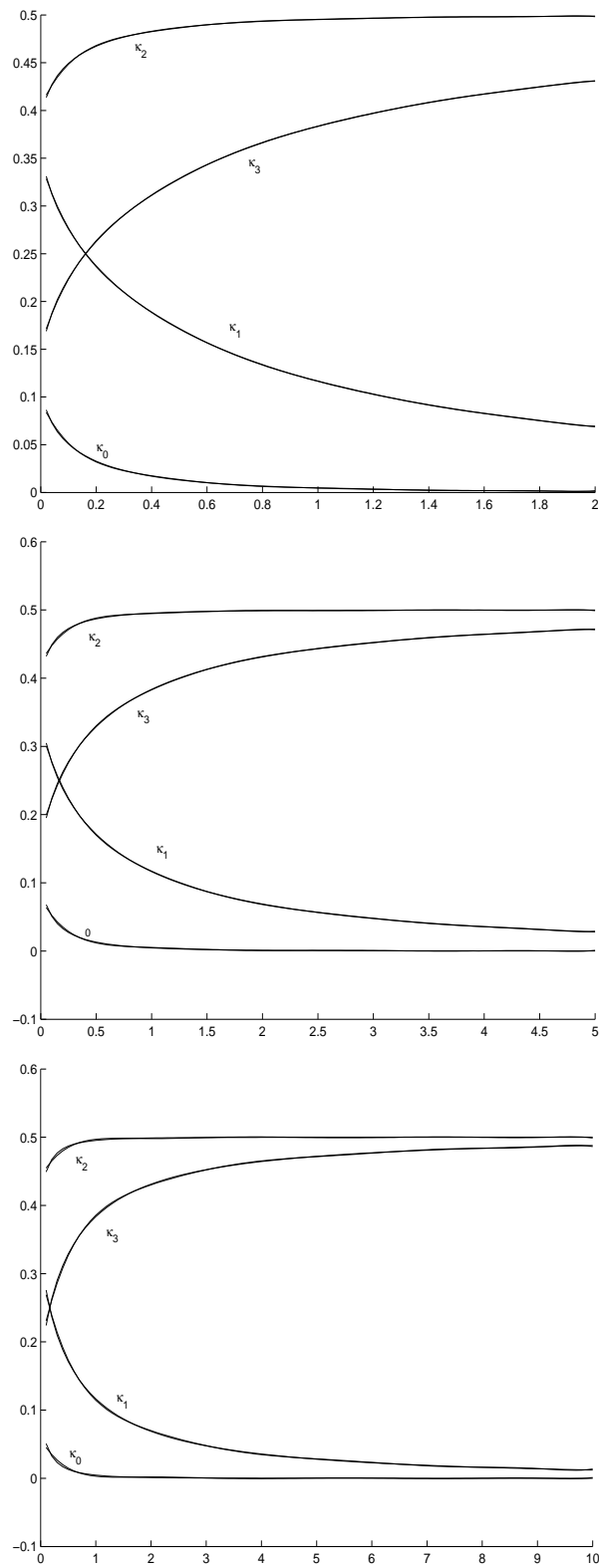


Abbildung 4.8: Ausgleichspynom 8. Grades für $S=3$

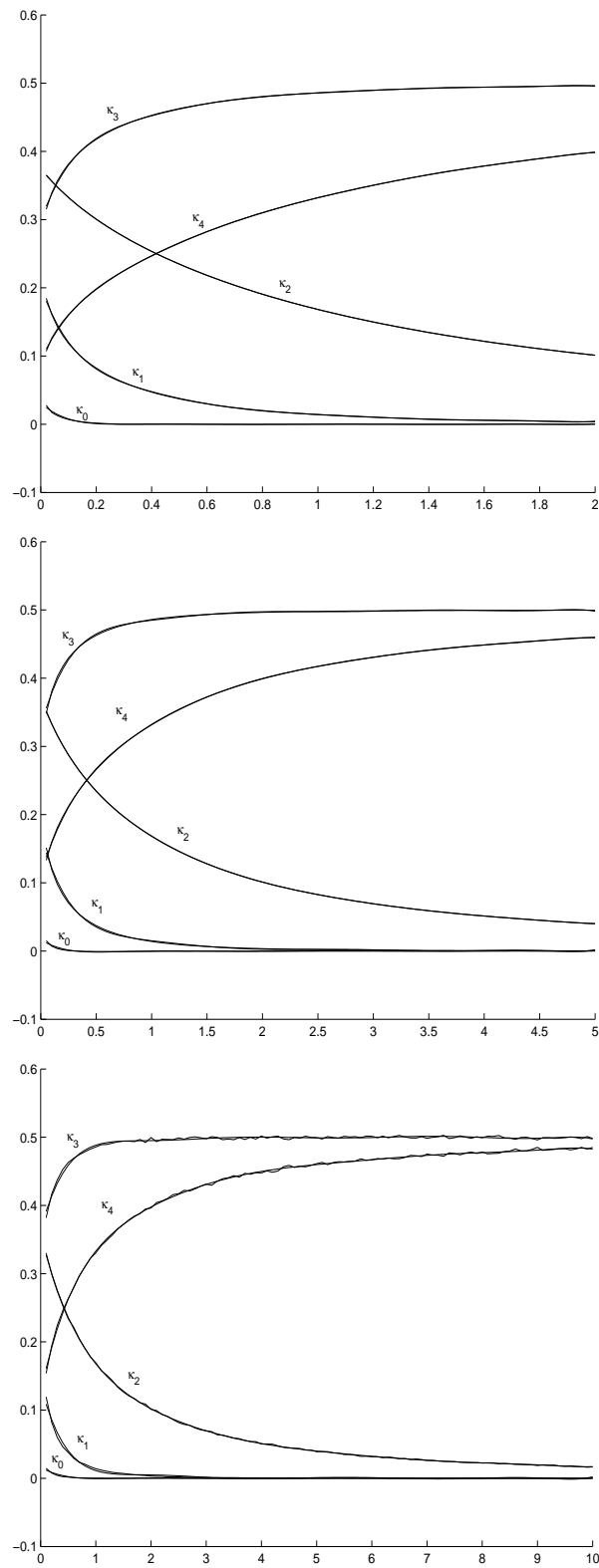


Abbildung 4.9: Ausgleichspynom 8. Grades für $S=4$

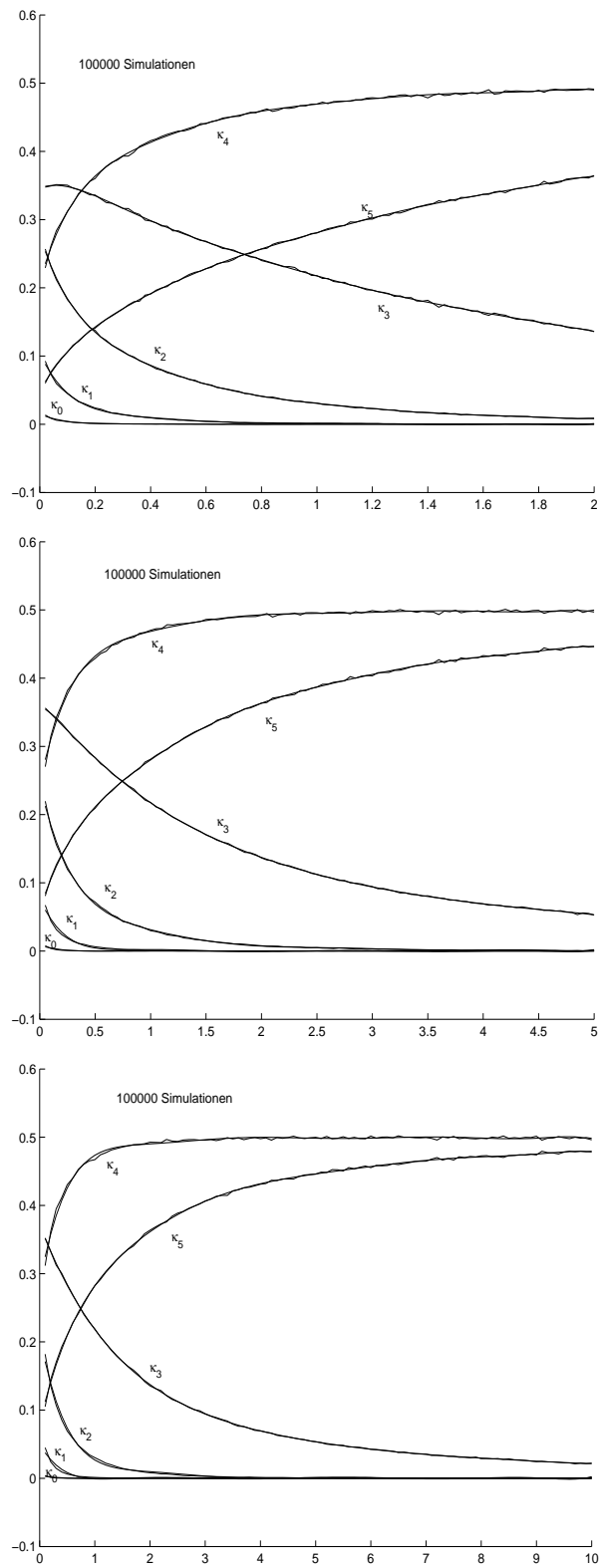


Abbildung 4.10: Ausgleichspolynom 8. Grades für $S=5$

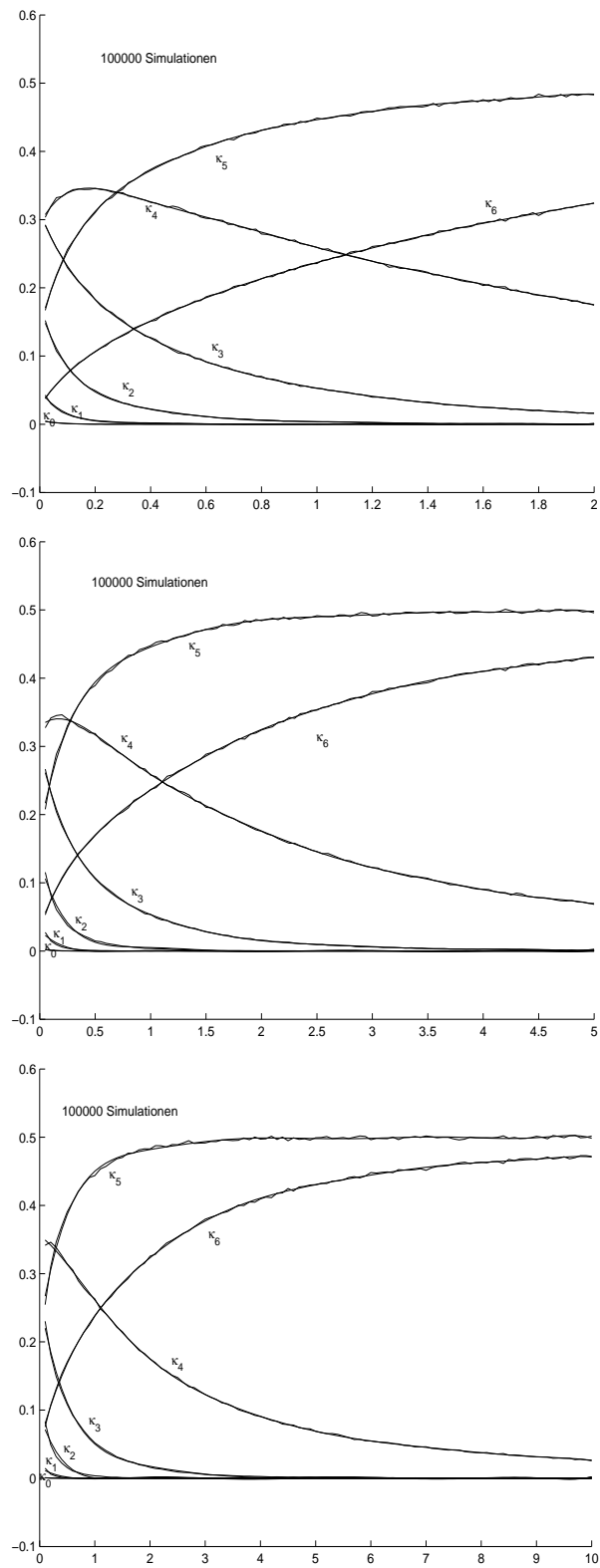


Abbildung 4.11: Ausgleichspolynom 8. Grades für $S=6$

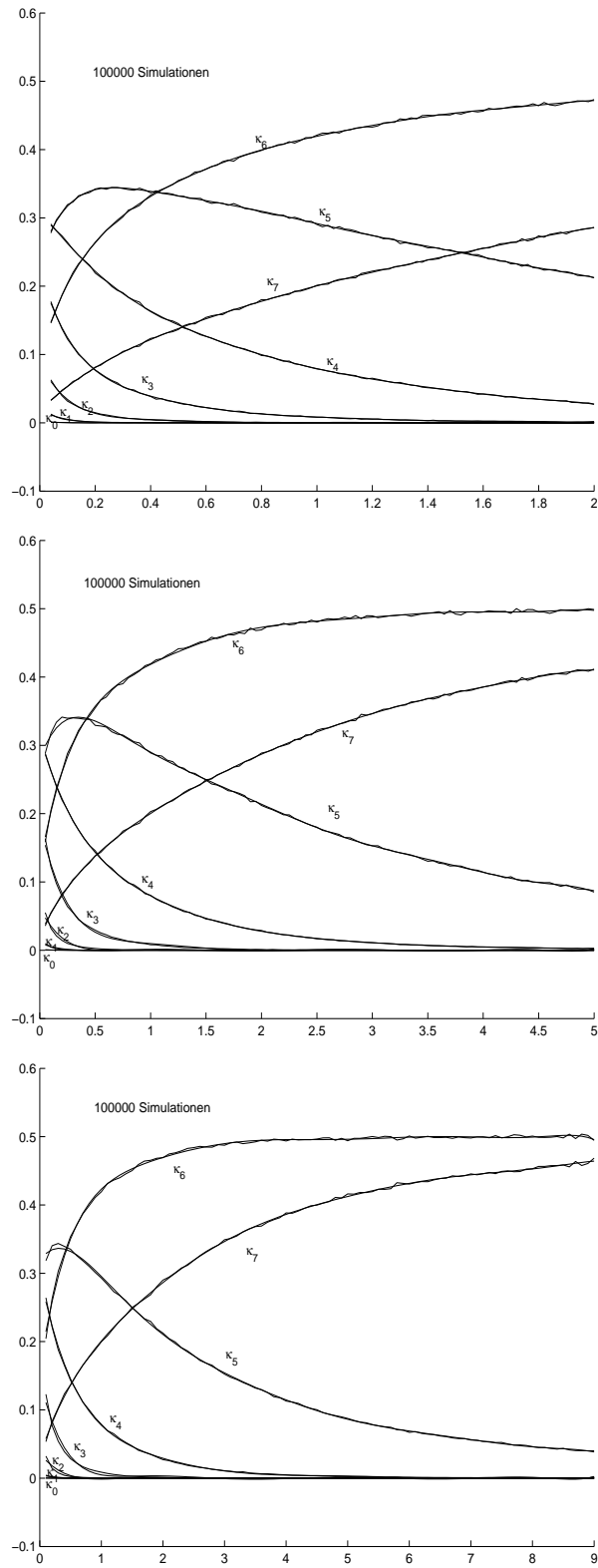


Abbildung 4.12: Ausgleichspolynom 8. Grades für $S=7$

4.2 Testen des *matlab*-Programms

Zum Abschluß der Arbeit sollte das im vorangehenden Abschnitt erklärte Programm noch getestet werden. Dazu wurden in *matlab* poissonverteilte und zusammengesetzt poissonverteilte Zufallsvariablen simuliert. Des weiteren wurden Daten geprüft, von denen man bereits mit mehreren Tests die Verteilung bestimmt hatte.

Betrachten wir zuerst die Ergebnisse, die der Algorithmus bei zusammengesetzt poissonverteilten Zufallsvariablen liefert. Hierbei wurden jeweils 50.000 Zufallsvariablen simuliert. Die genauen Werte sowie die von der *matlab*-Routine gelieferten Werte stehen in nachfolgender Tabelle:

	λ	p_1	p_2	p_3	p_4	p_5	p-Wert
simulierte Werte	1.4	0.7500	0.1250	0.0750	0.0250	0.0250	
<i>matlab</i> -Ergebnisse ($S = 5$)	1.4020	0.7522	0.1157	0.0841	0.0226	0.0246	$1.2351 \cdot 10^{-4}$
<i>matlab</i> -Ergebnisse ($S = 4$)	1.4020	0.7522	0.1157	0.0841	0.0226		0
<i>matlab</i> -Ergebnisse ($S = 3$)	1.4020	0.7522	0.1157	0.0841			$6.8421 \cdot 10^{-6}$

	λ	p_1	p_2	p_3	p_4	p_5	p-Wert
simulierte Werte	2.2	0.8000	0.1000	0.0500	0.0250	0.0250	
<i>matlab</i> -Ergebnisse ($S = 5$)	2.2041	0.8022	0.1005	0.0354	0.0489	0.0	$2.03 \cdot 10^{-6}$
<i>matlab</i> -Ergebnisse ($S = 4$)	2.2046	0.8019	0.1032	0.0273	0.0625		$1.006 \cdot 10^{-7}$
<i>matlab</i> -Ergebnisse ($S = 3$)	2.2046	0.8019	0.1032	0.0273			0

	λ	p_1	p_2	p_3	p_4	p_5	p-Wert
simulierte Werte	1.8	0.8500	0.1000	0.0250	0.0250	0.0000	
<i>matlab</i> -Ergebnisse ($S = 5$)	1,8041	0.8452	0.1082	0.0263	0.0128	0.0075	$6.5772 \cdot 10^{-7}$
<i>matlab</i> -Ergebnisse ($S = 4$)	1,8041	0.8453	0.1076	0.0287	0.0074		$5.4685 \cdot 10^{-8}$
<i>matlab</i> -Ergebnisse ($S = 3$)	1,8041	0.8453	0.1076	0.0287			0

Als nächstes wurden in *matlab* jeweils 10.000 poissonverteilte Zufallsvariablen simuliert. Die Ergebnisse, die die *matlab*-Routine hierbei liefert, sind wiederum in nachfolgenden Tabellen zusammengefaßt:

	λ	p_1	p_2	p_3	p_4	p_5	p-Wert
simulierte Werte	1.2	1.000					
<i>matlab</i> -Ergebnisse ($S = 5$)	1.2172	1.000					0.6959
<i>matlab</i> -Ergebnisse ($S = 4$)	1.2179	1.000					0.6482
<i>matlab</i> -Ergebnisse ($S = 3$)	1.2187	1.000					0.6018

	λ	p_1	p_2	p_3	p_4	p_5	p-Wert
simulierte Werte	0.8	1.000					
<i>matlab</i> -Ergebnisse ($S = 5$)	0.8004	1.000					0.7429
<i>matlab</i> -Ergebnisse ($S = 4$)	0.8005	1.000					0.6897
<i>matlab</i> -Ergebnisse ($S = 3$)	0.8005	1.000					0.6334

	λ	p_1	p_2	p_3	p_4	p_5	p-Wert
simulierte Werte	2.4	1.0000					
<i>matlab</i> -Ergebnisse ($S = 5$)	2.4036	1.0000					0.6170
<i>matlab</i> -Ergebnisse ($S = 4$)	2.4040	1.0000					0.5024
<i>matlab</i> -Ergebnisse ($S = 3$)	2.4090	0.9997	0.0003				0.0033

In allen obigen Fällen wurden bei der Newton-Raphson Methode für die Approximation der Maximallösung 5 Iterationen zur Bestimmung des maximalen λ -Wertes und 10 Iterationen zur Bestimmung des Maximums von $\log \left(\max_{(\lambda,p)^T \in K_1} \prod_{i=1}^n \Psi_{z_i}^S(\lambda,p) \right)$ benutzt.

Zum Abschluß wurde die *matlab*-Routine noch auf das klassische Testbeispiel von Bortkiewicz angewendet. Diese Daten beschreiben die Anzahl der Soldaten der preußischen Armee, die aufgrund von Pferdetritten starben. Die Daten beruhen auf Beobachtungen in 10 Korps über einen Zeitraum von 20 Jahren (vgl. auch [Qu87] für weitere Informationen). In nachfolgender Tabelle sind die Observationen zusammengefaßt:

Anzahl der Toten	0	1	2	3	4	≥ 5
absolute Häufigkeit	109	65	22	3	1	0
relative Häufigkeit	0.545	0.325	0.110	0.015	0.005	0

Die *matlab*-Routine lieferte folgendes Ergebnis:

	λ	p_1	p_2	p_3	p_4	p_5	p-Wert
<i>matlab</i> -Ergebnisse ($S = 5$)	0.61	1.000					0.7703
<i>matlab</i> -Ergebnisse ($S = 4$)	0.61	1.000					0.7161
<i>matlab</i> -Ergebnisse ($S = 3$)	0.6068	0.9975					$4.3487 \cdot 10^{-6}$

Zum Vergleich: Die klassische Poissonapproximation liefert folgendes Ergebnis:

$$\lambda = 0.6100$$

Anhang A

Algorithmus zur Simulation der lokalen Alternative

Dieses c-Programm simuliert die Verteilung der in Kapitel 2.3.2 bestimmten Zufallsgröße:

$$\|\pi_D(Z + \mu|V) - \pi_D(Z + \mu|C_q)\|_D^2 = \bar{X}^T (\bar{Q} - \bar{Q}E_I(E_I^T\bar{Q}E_I)^{-1}E_I^T\bar{Q})\bar{X}$$

Dabei wurde der in [Bu01] beschriebene Algorithmus in weiten Teilen übernommen. Gravierende Unterschiede wurden meist auskommentiert.

```
#include <stdio.h>
#include <math.h>
#include <stdlib.h>
#define S 3
#define zweihochSminus2 8-2
#define NSAMPLE 100000
#define DIMMAX 50*/

double lambda,kr;
double Q[S+1][S+1],QIminus[S+1][S+1],Lminus[S+1][S+1];
double h[S+1],Qminus[S+1][S+1],hq[S+1];
double DI[zweihochSminus2+1][S+1][S+1],DI2[zweihochSminus2+1][S+1][S+1];
int Imenge[S+1],DIberechnet[zweihochSminus2+1],DIberechnet2[zweihochSminus2+1];

void hquer();
void initQLDIber();
void newDecider(int index, int Ber);
int IndexI();
double unif(void);
double gauss(void);
int card(int index);
```

```

main()
{int k,i,j,imin,samplei,algorfails,CardI,ergpos,erneg;
double Y[S+1],Z[S+1],kappa[S+1],halbraum[zweihochSminus2+3];
double Min,summand,sumplus,summinus,Norm;
int fail,failzaehler;

failzaehler=0;
ergpos=0;erneg=0,Norm=0;
printf("\n\n");
printf("Intensitaet lambda =");scanf("%lf",&lambda);
printf("kritischer Wert=");scanf("%lf",&kr);
for(k=0;k<=S;k=k+1)
{
printf("Komponente %i des Richtungsvektors=",k);scanf("%lf",&h[k]);
}
printf("Abschneideparamter S=%i\n",S);
printf("Sampleumfang S=%i",NSAMPLE);

initQLDIber();
hquer();
for(k=0;k<=zweihochSminus2+2;k=k+1)
{halbraum[k]=0;DIberechnet[k]=0;DIberechnet2[k]=0;}
for(i=0;i<=S;i=i+1){kappa[i]=0;}

for(samplei=1;samplei<=NSAMPLE;samplei=samplei+1)
{
fail=0;

Norm=0;
for(i=1;i<=S;i=i+1){Z[i]=gauss();}
/*Berechnung von Y=L^{-T}Z\sim N(0,Q^{-1})*/
for(i=1;i<=S;i=i+1)
{sumplus=0;summinus=0;
for(j=i;j<=S;j=j+1)
{summand=Lminus[j][i]*Z[j];
if(summand>=0){sumplus=sumplus+summand;}
else{summinus=summinus+summand;}}
Y[i]=sumplus+summinus+hq[i];
}
}

```

```

    Min=Y[1];imin=1;
    for(i=2;i<=S;i=i+1){if(Y[i]<=Y[imin]){Min=Y[i];imin=i;}}
    if(Min>=0)
{
halbraum[zweihochSminus2+1]=halbraum[zweihochSminus2+1]+1;
kappa[S]=kappa[S]+1;ergneg=ergneg+1;
}
else
{imin=1;
for(i=1;i<=S;i=i+1){Imenge[i]=0;}
sumplus=0;summinus=0;
for(j=1;j<=S;j=j+1)
{summand=-Q[1][j]*Y[j];
if(summand>=0){sumplus=sumplus+summand;}
else{summinus=summinus+summand;}}
Min=summinus+sumplus;
for(i=2;i<=S;i=i+1)
{sumplus=0;summinus=0;
for(j=1;j<=S;j=j+1)
{summand=-Q[i][j]*Y[j];
if(summand>=0){sumplus=sumplus+summand;}
else{summinus=summinus+summand;}}

if(sumplus+summinus<=Min){Min=sumplus+summinus;imin=i;}}
if(Min>=0)
{
halbraum[0]=halbraum[0]+1;
kappa[0]=kappa[0]+1;k=0;
}
else
{Imenge[imin]=1;algorfails=0;CardI=1;
do
{k=IndexI();algorfails=algorfails+1;
if(DIberechnet[k]==0)
{
newDecider(k,0);
DIberechnet[k]=1;
}
imin=1;

sumplus=0;summinus=0;
for(j=1;j<=S;j=j+1)
{summand=DI[k][1][j]*Y[j];
if(summand>=0){sumplus=sumplus+summand;}
else{summinus=summinus+summand;}}
Min=summinus+sumplus;

```



```

        for(i=2;i<=S;i=i+1)
            {sumplus=0;summinus=0;
            for(j=1;j<=S;j=j+1)
                {summand=DI[k][i][j]*Y[j];
                if(summand>=0){sumplus=sumplus+summand;}
                else{summinus=summinus+summand;}}
            if(sumplus+summinus<=Min)
{Min=sumplus+summinus;imin=i;}}
        if(Min<0)
            {if(Imenge[imin]>0){Imenge[imin]=0;CardI=CardI-1;}
            else{Imenge[imin]=1;CardI=CardI+1;}}
            while(Min<0 && algorfails<zweihochSminus2+1);
            if(Min<0)
{
halbraum[zweihochSminus2+2]=halbraum[zweihochSminus2+2]+1;
fail=1;failzaehler=failzaehler+1;
        }
        else{halbraum[k]=halbraum[k]+1;kappa[CardI]=kappa[CardI]+1;}
    }
/* Bis hierher wurde der Algorithmus mit wenigen Aenderungen uebernommen. */
/* Hier wird nun mittels der Indexmenge I die Projektion berechnet. */

    if(fail==0)
    {
        summand=0;sumplus=0;summinus=0;
if(DIberechnet2[k]==0){newDecider(k,1);DIberechnet2[k]=1;}

sumplus=0;summinus=0;
for(i=1;i<=S;i=i+1)
{
if(k==0){summand=Q[i][i]*Y[i]*Y[i];}
else{summand=DI2[k][i][i]*Y[i]*Y[i];}
if(summand>=0){sumplus=sumplus+summand;}
else{summinus=summinus+summand;}

}
Norm=Norm+sumplus+summinus;
summand=0;sumplus=0;summinus=0;
for(i=1;i<=S-1;i=i+1)
{
for(j=i+1;j<=S;j=j+1)
{
if(k==0){summand=Q[i][j]*Y[i]*Y[j];}
else{summand=DI2[k][i][j]*Y[i]*Y[j];}
if(summand>=0){sumplus=sumplus+summand;}

```

```

else{summinus=summinus+summand;}
}
}
Norm=Norm+2*(sumplus+summinus);
if(Norm>=kr){ergpos=ergpos+1;}
else{ergneg=ergneg+1;}
}
    }}
printf("\n \n \n");
printf("\n Anzahl der Fehlversuche: %i",failzaehler);
printf("\n Anzahl der positiven Ergebnisse: %i", ergpos);
printf("\n Anzahl der negativen Ergebnisse: %i", ergneg);
printf("\n Wahrscheinlichkeit des Fehlers 2. Art:
        %lf", 1.*ergneg/(ergneg+ergpos));
printf("\n");

}/*Ende main*/

/* In der Funktion hquer wird der Ausdruck          */
/*  $\overline{Q}^{-1}N^T A^{-1}$  berechnet          */

void hquer()
{
int i,j,k;
double summand,sumplus,summinus;

for(i=1;i<=S;i=i+1)
{
for(j=i;j<=S;j=j+1)
{
summand=0;sumplus=0;summinus=0;
for(k=j;k<=S;k=k+1)
{
summand=Lminus[k][j]*Lminus[k][i];
if(summand>=0){sumplus=sumplus+summand;}
else{summinus=summinus+summand;}
}
Qminus[i][j]=sumplus+summinus;
Qminus[j][i]=Qminus[i][j];
}
}
}

```

```

summand=0;sumplus=0;summinus=0;
for(i=1;i<=S;i=i+1)
{
sumplus=0;summinus=0;
summand=Qminus[i][1]*h[1];
if(summand>=0){sumplus=sumplus+summand;}
else{summinus=summinus+summand;}
for(j=2;j<=S;j=j+1)
{
summand=(Qminus[i][1]-Qminus[i][j])*h[j];
if(summand>=0){sumplus=sumplus+summand;}
else{summinus=summinus+summand;}
}
hq[i]=sumplus+summinus;
}
summand=0;sumplus=0;summinus=0;

} /*Ende von hquer*/

void initQLDIber()
{
double Delta[S+1][S+1];
double q[S+1],qminus[S+1],Deltasum[S+1];
double normtildeq,sumplus,summinus,summand;
int i,j,k;

for(k=1;k<=zweihochSminus2;k=k+1){DIberechnet[k]=0;}

q[0]=exp(-lambda);
for(i=1; i<=S; i=i+1){q[i]=q[i-1]*lambda/i;}
qminus[0]=exp(lambda);
for(i=1; i<=S; i=i+1){qminus[i]=qminus[i-1]*(-lambda)/i;}

sumplus=0;summinus=0;
for(k=1;k<=S;k=k+2){summinus=summinus+qminus[k];}
for(k=2;k<=S;k=k+2){sumplus=sumplus+qminus[k];}
normtildeq=sumplus+summinus;

/*Berechnung von Delta*/

for(i=0;i<=S;i=i+1)
{for(j=1;j<=i;j=j+1)
{sumplus=0;summinus=0;

```

```

        for(k=0;k<=j;k=k+1)
            {summand=qminus[i-k]*qminus[j-k]*q[k];
if(summand>=0)
    {sumplus=sumplus+summand;}
else
    {summinus=summinus+summand;}}
        Delta[i][j]=sumplus+summinus;
        Delta[j][i]=sumplus+summinus;
    }}

for(i=1;i<=S;i=i+1)
    {sumplus=0;summinus=0;
        for(k=1;k<=S;k=k+1)
            {summand=Delta[i][k];
if(summand>=0)
    {sumplus=sumplus+summand;}
else
    {summinus=summinus+summand;}}
        Deltasum[i]=(sumplus+summinus);
    }

/* Berechnung von
/*  $Q=f^T A^{-1}(\text{diag}(q_0, \dots, q_S) - q \otimes q) A^{-T} f$ 
*/

sumplus=0;summinus=0;
for(k=1;k<=S;k=k+1)
    {summand=Deltasum[k];
        if(summand>=0) {sumplus=sumplus+summand;}
        else {summinus=summinus+summand;}}

Q[1][1]=(sumplus+summinus+2*normtildeq*exp(lambda)-1)/(lambda*lambda);

for(i=2;i<=S;i=i+1)
    {Q[1][i]=-(Deltasum[i]+qminus[i])/(lambda*lambda);Q[i][1]=Q[1][i];
    }

for(i=2;i<=S;i=i+1)
    {for(j=i;j<=S;j=j+1)
        {Q[i][j]=Delta[i][j]/(lambda*lambda);Q[j][i]=Q[i][j];}
    }

/* Cholesky-Zerlegung von
*/

```

```

/*  $Q=f^T A^{-1}(\text{diag}(q_0, \dots, q_S)-q \otimes q) A^{-T} f$  */
for(i=1; i<=S; i=i+1)
  {sumplus=0;
   for(k=1; k<=i-1; k=k+1)
     {sumplus=sumplus+Delta[i][k]*Delta[i][k];}
   Delta[i][i]=sqrt(Q[i][i]-sumplus);
   for(j=i+1; j<= S; j=j+1)
     {sumplus=0; summinus=0;
      for(k=1; k<=i-1; k=k+1)
        {summand=Delta[i][k]*Delta[j][k];
         if(summand>=0)
           {sumplus=sumplus+summand;}
         else
           {summinus=summinus+summand;}}
      Delta[j][i]=(Q[i][j]-summinus-sumplus)/Delta[i][i];}
  }

/*Berechnung der Inversen  $L^{-1}$  der Choleskyzerlegung von  $Q^*$ */

for(i=1; i<=S; i=i+1){Lminus[i][i]=1/Delta[i][i];}
for(i=2; i<=S; i=i+1)
  {for(j=i-1; j>=1; j=j-1)
   {sumplus=0; summinus=0;
    for(k=j+1; k<=i; k=k+1)
      {summand=Lminus[i][k]*Delta[k][j];
       if(summand>=0)
         {sumplus=sumplus+summand;}
       else
         {summinus=summinus+summand;}}
    Lminus[i][j]=-(sumplus+summinus)/Delta[j][j];}}

}/*Ende von initQL*/

/* Die Funktion wurde dahingehend abgeändert, dass noch ein */
/* zweiter Index aufgenommen wurde. Dieser dient zur */
/* Unterscheidung, welche der aehnlich zu berechnenden Matrixen */
/* (Entscheidungsmatrix oder Matrix zur Berechnung der Norm) */
/* bestimmt werden soll */

void newDecider(int index, int Ber)
{
double sumplus, summinus, summand;
double LI[S+1][S+1], LIminus[S+1][S+1], QDI[S+1][S+1];

```

```

int cardImenge,k,k1,l,i,j;

/*Berechnung von  $E_I(E_I^T Q E_I)^{-1} E_I^T$ */

/*Berechnung von  $Q_I=(E_I^T Q E_I)$ */

i=0;cardImenge=0;
for(k=1;k<=S;k=k+1)
  {if(Imenge[k]>0)
{cardImenge=cardImenge+1;i=i+1;j=i-1;
      for(k1=k;k1<=S;k1=k1+1)
        {
          if(Imenge[k1]>0)
{j=j+1;QIminus[i][j]=Q[k][k1];
QIminus[j][i]=Q[k][k1];}
        }
      }
}

/*Cholesky-Zerlegung von QIminus*/

for(i=1;i<=cardImenge;i=i+1)
  {sumplus=0;
  for(k=1;k<=i-1;k=k+1)
    {sumplus=sumplus+LI[i][k]*LI[i][k];}
  LI[i][i]=sqrt(QIminus[i][i]-sumplus);
  for(j=i+1;j<= cardImenge;j=j+1)
    {sumplus=0;summinus=0;
    for(k=1;k<=i-1;k=k+1)
      {summand=LI[i][k]*LI[j][k];
      if(summand>=0)
        {sumplus=sumplus+summand;}
      else
        {summinus=summinus+summand;}}
    LI[j][i]=(QIminus[i][j]-summinus-sumplus)/LI[i][i];}
  }

/*Invertierung von QIminus*/

for(i=1;i<=cardImenge;i=i+1){LIminus[i][i]=1/LI[i][i];}
for(i=2;i<=cardImenge;i=i+1) {for(j=i-1;j>=1;j=j-1)
  {sumplus=0;summinus=0;
  for(k=j+1;k<=i;k=k+1)
    {summand=LIminus[i][k]*LI[k][j];
    if(summand>=0)

```

```

    {sumplus=sumplus+summand;}
      else
    {summinus=summinus+summand;}}
    LIminus[i][j]=-(sumplus+summinus)/LI[j][j];}}

i=0;
for(k=1;k<=S;k=k+1)
  {if(Imenge[k]>0){
      i=i+1;j=i-1;
      for(k1=k;k1<=S;k1=k1+1)
        {if(Imenge[k1]>0)
            {j=j+1;
              sumplus=0;summinus=0;
              for(l=j;l<=cardImenge;l=l+1)
                {summand=LIminus[l][i]*LIminus[l][j];
                  if(summand>=0)
                    {sumplus=sumplus+summand;}
                    else {summinus=summinus+summand;}}
                QIminus[k][k1]=sumplus+summinus;
                QIminus[k1][k]=QIminus[k][k1];}
              else{QIminus[k][k1]=0;QIminus[k1][k]=0;}
            }}
      else{for(k1=1;k1<=S;k1=k1+1){QIminus[k][k1]=0;QIminus[k1][k]=0;}}}

/*Berechnung der Entscheidungsmatrix */
/*DI=E_I(E_I^TQE_I)E_I^TQ+E_I^C(E_I^C^TQ-E_I^C^TE_I(E_I^TQE_I)QE_I^TQ) */

for(i=1;i<=S;i=i+1)
  {for(j=1;j<=S;j=j+1)
      {sumplus=0;summinus=0;
        for(k=1;k<=S+1;k=k+1)
          {summand=QIminus[i][k]*Q[k][j];
            if(summand>=0){sumplus=sumplus+summand;}
            else{summinus=summinus+summand;}}
          DI[index][i][j]=sumplus+summinus;}}

for(i=1;i<=S;i=i+1)
  {for(j=1;j<=S;j=j+1)
      {sumplus=0;summinus=0;
        for(k=1;k<=S+1;k=k+1)
          {summand=Q[i][k]*DI[index][k][j];
            if(summand>=0){sumplus=sumplus+summand;}
            else{summinus=summinus+summand;}}
    }

```

```

    QDI[i][j]=sumplus+summinus;}]

/* Hier erfolgt die unterschiedliche Berechnung */

if(Ber==0)
{
for(i=1;i<=S;i=i+1)
{if(Imenge[i]==0)
    {for(j=1;j<=S;j=j+1)
        {if(Imenge[j]==0){DI[index][i][j]=-(Q[i][j]-QDI[i][j]);}}}
    }
else
    {
for(i=1;i<=S;i=i+1)
    {
for(j=1;j<=S;j=j+1)
    {
DI2[index][i][j]=Q[i][j]-QDI[i][j];
    }
    }
    }
}

}/*Ende newDecider*/

int IndexI()
{int ergebnis,k,zweihochk;

zweihochk=1;ergebnis=0;
for(k=1;k<=S;k=k+1)
{ergebnis=ergebnis+zweihochk*Imenge[k];zweihochk=2*zweihochk;}
return(ergebnis);}

double gauss(void){
static int i = 0; /* 1 if value in stock */
static double x1, x2, y1, y2;

if (i == 0){
x1 = unif();
x2 = unif();
y1 = sqrt(-2 * log(x1));
y2 = 2 * M_PI * x2;
x1 = y1 * sin(y2);
x2 = y1 * cos(y2);
}
}

```



```

i = 1;
return x1;
}
else{
i = 0;
return x2;
}
}
}
/* returns U(0,1)-variates, Marsaglia-Zaman algorithm */
double unif(void){
static unsigned long x[] =
{1276610355UL, 4193469394UL, 2057566612UL, 1886580328UL, 1694206606UL,
 2633431637UL, 1265626433UL,  885029446UL, 3417643270UL, 3311627661UL,
 2615330922UL, 2585171253UL, 2061319010UL,  76799462UL,  217610450UL,
 1970157156UL, 3650280925UL, 3031778051UL, 3936002891UL, 1455404536UL,
 3581605850UL,  978584193UL, 1392725752UL,  424558724UL,  718634923UL,
 2602380921UL, 1073859225UL, 2260449986UL,  437368889UL,  111202475UL,
  430748330UL,  860297108UL,  469595518UL, 2956147077UL, 2998566928UL,
 3679001976UL, 1174826611UL, 3589929608UL, 2670654217UL,  999890898UL,
 3874011621UL, 3680146780UL, 3569051095UL };
static int r = 0, s = 21, carry = 0;

if (r > 42) r -= 43;
if (x[s] >= x[r] + carry){
x[r] = x[s] - x[r] - carry;
carry = 0;
}
else{
x[r] = (4294967291UL - x[r] - carry) + x[s];
carry = 1;
}
if (++s > 42) s -= 43;
return (((double) x[r++] + 0.5) / 4294967291.0);
}

```

Anhang B

matlab Dateien

Die nachfolgenden Dateien sind diejenigen *matlab* Dateien, die zur Durchführung des Tests benötigt werden. Es handelt sich dabei um die nachfolgenden sechs Funktionen, die alle im Laufe der Testdurchführung aufgerufen werden.

Zum Ausführen des Tests ist der Vektor, der die Daten enthält, im selben Verzeichnis wie die *matlab*-Funktionen abzuspeichern. Der Test ist dann mittels des Befehls 'Auswertung(S,Quantil)' auszuführen. Dabei steht 'S' für den Abschneideparameter und 'Quantil' für das Quantil, für das der Test durchgeführt werden soll (bzw. den p-Wert).

Weitere Änderungen, die am Test vorgenommen werden können, sind den einzelnen Funktionen zu entnehmen.

```
function [x_zusammen]=Auswertung(S,Quantil)
% Eingabe von Quantil=1 bewirkt Berechnung des p-Wertes
% Quantil z.B als 0.95 oder 0.99 eingeben

% Die nachfolgenden Daten werden fuer die Funktion quantilber2 benoetigt
linkeSeite=0;
rechteSeite=8;
Feinheit=10000;

% Anzahl der Wiederholungen unter einfacher Poissonverteilung
Anzahl1 = 5;

% Anzahl der Wiederholungen unter zusammengesetzter Poissonverteilung
Anzahl2 = 10;

% Bestimmen der Testgroesse
[Ergebnis,x_einfach,x_zusammen]=Test(S,Anzahl1,Anzahl2);
lambda=x_einfach;
if (Ergebnis<0) Ergebnis=0; end %Rundungsungenauigkeiten beseitigen
```

```

Daten der Ausgleichspolynome
%lambda in (0,2)
Koeff2=zeros(3,6,7,0);
Koeff(1,2,1,:)= [ 9.2691110e-02, -8.2017175e-01, 3.0356867e+00, -6.1084756e+00,
7.2876891e+00, -5.3100975e+00, 2.4024584e+00, -7.3736621e-01, 2.2296032e-01];
Koeff(1,2,2,:)= [ -9.1621151e-15, 7.3752785e-14, -2.3529831e-13, 3.7519723e-13,
-3.0813140e-13, 1.1907352e-13, -1.6658052e-14, 1.1463631e-15, 5.0000000e-01];
Koeff(1,2,3,:)= [ -9.2691106e-02, 8.2017170e-01, -3.0356865e+00, 6.1084751e+00,
-7.2876884e+00, 5.3100970e+00, -2.4024582e+00, 7.3736619e-01, 2.7703968e-01];

Koeff(1,3,1,:)= [1.0143449e-01, -8.9668471e-01, 3.3137531e+00, -6.6495073e+00,
7.8880765e+00, -5.6677090e+00, 2.4557216e+00, -6.3596766e-01, 9.5641937e-02];
Koeff(1,3,2,:)= [ 1.0350242e-01, -9.1551299e-01, 3.3867893e+00, -6.8097956e+00,
8.1173925e+00, -5.9183721e+00, 2.7132364e+00, -9.0620002e-01, 3.4553417e-01];
Koeff(1,3,3,:)= [ -1.0143446e-01, 8.9668440e-01, -3.3137521e+00, 6.6495055e+00,
-7.8880748e+00, 5.6677080e+00, -2.4557214e+00, 6.3596765e-01, 4.0435806e-01];
Koeff(1,3,4,:)= [ -1.0350243e-01, 9.1551313e-01, -3.3867899e+00, 6.8097966e+00,
-8.1173936e+00, 5.9183727e+00, -2.7132366e+00, 9.0620004e-01, 1.5446583e-01];

Koeff(1,4,1,:)= [ 8.9668730e-02, -7.8967736e-01, 2.9011076e+00, -5.7638647e+00,
6.7132939e+00, -4.6456943e+00, 1.8452694e+00, -3.8146937e-01, 3.1501293e-02];
Koeff(1,4,2,:)= [ 1.7201568e-01, -1.5206278e+00, 5.6199533e+00, -1.1280479e+01,
1.3395504e+01, -9.6606314e+00, 4.2444708e+00, -1.1579392e+00, 2.0207210e-01];
Koeff(1,4,3,:)= [ 1.5441433e-02, -1.4135809e-01, 5.4742758e-01, -1.1729627e+00,
1.5354903e+00, -1.3039228e+00, 8.0155041e-01, -4.8781380e-01, 3.7446646e-01];
Koeff(1,4,4,:)= [ -1.7201568e-01, 1.5206278e+00, -5.6199533e+00, 1.1280479e+01,
-1.3395504e+01, 9.6606314e+00, -4.2444708e+00, 1.1579392e+00, 2.9792790e-01];
Koeff(1,4,5,:)= [ -1.0930929e-01, 9.6381439e-01, -3.5499099e+00, 7.0914657e+00,
-8.3626120e+00, 5.9759070e+00, -2.6379746e+00, 8.6645899e-01, 9.4155589e-02];

Koeff(1,5,1,:)= [4.6446907e-02, -4.0543339e-01, 1.4731793e+00, -2.8867515e+00,
3.3044228e+00, -2.2385347e+00, 8.6851170e-01, -1.7688010e-01, 1.5136007e-02];
Koeff(1,5,2,:)= [ 1.5908285e-01, -1.4051180e+00, 5.1851315e+00, -1.0376928e+01,
1.2245231e+01, -8.6983330e+00, 3.6632950e+00, -8.7515130e-01, 1.0421443e-01];
Koeff(1,5,3,:)= [ 1.5421716e-01, -1.3723226e+00, 5.1136800e+00, -1.0373828e+01,
1.2500006e+01, -9.2190231e+00, 4.2236164e+00, -1.2730872e+00, 2.7731927e-01];
Koeff(1,5,4,:)= [ -5.9577226e-02, 5.6365146e-01, -2.2290798e+00, 4.7644639e+00,
-5.9425588e+00, 4.3465164e+00, -1.7271109e+00, 1.5487175e-01, 3.4630672e-01];
Koeff(1,5,5,:)= [ -1.9718551e-01, 1.7353262e+00, -6.3900198e+00, 1.2800677e+01,
-1.5215403e+01, 1.1041673e+01, -4.9397890e+00, 1.4254891e+00, 2.0869414e-01];
Koeff(1,5,6,:)= [ -1.0298418e-01, 8.8389637e-01, -3.1528913e+00, 6.0723660e+00,
-6.8916977e+00, 4.7677017e+00, -2.0885232e+00, 7.4475777e-01, 4.8329435e-02];

Koeff(1,6,1,:)= [ 1.8373279e-02, -1.6076847e-01, 5.8566933e-01, -1.1504878e+00,
1.3192174e+00, -8.9295164e-01, 3.4358708e-01, -6.7923829e-02, 5.3153585e-03];
Koeff(1,6,2,:)= [ 1.2051019e-01, -1.0595387e+00, 3.8849218e+00, -7.7018192e+00,
8.9534376e+00, -6.1959620e+00, 2.4792880e+00, -5.2965999e-01, 4.9092030e-02];
Koeff(1,6,3,:)= [ 1.9582454e-01, -1.7395979e+00, 6.4660075e+00, -1.3063729e+01,
1.5620770e+01, -1.1320511e+01, 4.9333245e+00, -1.2596785e+00, 1.7149071e-01];
Koeff(1,6,4,:)= [ 7.4047803e-02, -6.6669029e-01, 2.5217802e+00, -5.2266940e+00,
6.5194308e+00, -5.1147294e+00, 2.6412370e+00, -1.0060009e+00, 3.1062013e-01];
Koeff(1,6,5,:)= [ -1.7616406e-01, 1.5756163e+00, -5.8856753e+00, 1.1906329e+01,
-1.4153783e+01, 1.0052527e+01, -4.1362909e+00, 7.8326100e-01, 2.9329411e-01];
Koeff(1,6,6,:)= [ -1.8269184e-01, 1.6078790e+00, -5.9328086e+00, 1.1944144e+01,
-1.4332698e+01, 1.0575728e+01, -4.8716914e+00, 1.4961336e+00, 1.4257155e-01];
Koeff(1,6,7,:)= [ -4.9899921e-02, 4.4310014e-01, -1.6398949e+00, 3.2922567e+00,
-3.9263747e+00, 2.8958992e+00, -1.3894542e+00, 5.8386851e-01, 2.7616110e-02];

Koeff(1,7,1,:)= [ 5.5341988e-03, -4.8584369e-02, 1.7761731e-01, -3.5018623e-01,
4.0289693e-01, -2.7327862e-01, 1.0496311e-01, -2.0493893e-02, 1.5392941e-03];
Koeff(1,7,2,:)= [ 4.7592315e-02, -4.2102900e-01, 1.5548464e+00, -3.1079095e+00,
3.6459840e+00, -2.5456974e+00, 1.0236190e+00, -2.1603537e-01, 1.8693100e-02];
Koeff(1,7,3,:)= [ 1.3634719e-01, -1.2154820e+00, 4.5356576e+00, -9.2006638e+00,
1.1033926e+01, -7.9802244e+00, 3.4108573e+00, -8.0784593e-01, 8.7900485e-02];
Koeff(1,7,4,:)= [ 1.5720762e-01, -1.4167823e+00, 5.3646018e+00, -1.1106275e+01,
1.3725905e+01, -1.0417702e+01, 4.8594358e+00, -1.3815599e+00, 2.2355888e-01];
Koeff(1,7,5,:)= [ -6.8921839e-02, 5.6568219e-01, -1.8942597e+00, 3.2977904e+00,
-3.0931220e+00, 1.3212930e+00, 1.4802971e-01, -5.0820104e-01, 3.1101452e-01];
Koeff(1,7,6,:)= [ -1.7923720e-01, 1.6371472e+00, -6.2635683e+00, 1.3032769e+01,
-1.6035873e+01, 1.1904870e+01, -5.2266017e+00, 1.1813531e+00, 2.4081161e-01];
Koeff(1,7,7,:)= [ -1.1234731e-01, 1.0174835e+00, -3.8713778e+00, 8.0681665e+00,
-1.0095567e+01, 7.8729983e+00, -3.9277796e+00, 1.3699879e+00, 9.8018223e-02];
Koeff(1,7,8,:)= [ 1.3825032e-02, -1.1843517e-01, 3.9648264e-01, -6.3369111e-01,
4.1585063e-01, 1.1774124e-01, -3.9252367e-01, 3.8279514e-01, 1.8463895e-02];

%lambda in (0,5)
Koeff(2,2,1,:)= [ 1.0545631e-04, -2.3325633e-03, 2.1574105e-02, -1.0839085e-01,
3.2214881e-01, -5.8077203e-01, 6.3458055e-01, -4.2713257e-01, 2.0602746e-01];
Koeff(2,2,2,:)= [ 8.2786713e-19, -2.1693239e-17, 2.2942607e-16, -1.2401796e-15,
3.5994709e-15, -5.3912557e-15, 3.5728493e-15, -4.5854526e-16, 5.0000000e-01];
Koeff(2,2,3,:)= [ -1.0545633e-04, 2.3325637e-03, -2.1574108e-02, 1.0839086e-01,
-3.2214883e-01, 5.8077204e-01, -6.3458054e-01, 4.2713256e-01, 2.9397254e-01];

Koeff(2,3,1,:)= [ 1.0450314e-04, -2.2997100e-03, 2.1116572e-02, -1.0493616e-01,
3.0631943e-01, -5.3422355e-01, 5.4328861e-01, -3.0159851e-01, 7.7442614e-02];
Koeff(2,3,2,:)= [ 1.1508882e-04, -2.5488940e-03, 2.3627510e-02, -1.1919715e-01,
3.5720323e-01, -6.5580950e-01, 7.4894906e-01, -5.6206828e-01, 3.2677835e-01];

```

```

Koeff(2,3,3,:)= [-1.0450322e-04, 2.2997116e-03, -2.1116586e-02, 1.0493623e-01,
-3.0631961e-01, 5.3422382e-01, -5.4328883e-01, 3.0159858e-01, 4.2255738e-01];
Koeff(2,3,4,:)= [-1.1508883e-04, -2.3627510e-02, 1.1919715e-01,
-3.5720322e-01, 6.5580948e-01, -7.4894904e-01, 5.6206828e-01, 1.7322165e-01];

Koeff(2,4,1,:)= [-1.0176297e-04, 2.2401098e-03, -2.0598815e-02, 1.0280543e-01,
-3.0380073e-01, 5.4953291e-01, -6.2900370e-01, 5.1744052e-01, 1.1304843e-01];
Koeff(2,4,2,:)= [5.5389574e-05, -1.2045007e-03, 1.0885322e-02, -5.2899828e-02,
1.4938606e-01, -2.4702610e-01, 2.2847311e-01, -1.0443232e-01, 1.6687542e-02];
Koeff(2,4,3,:)= [1.8077878e-04, -3.9878467e-03, 3.6744750e-02, -1.8356352e-01,
5.4045381e-01, -9.5692437e-01, 1.0020332e+00, -5.9105360e-01, 1.7122410e-01];
Koeff(2,4,4,:)= [4.6373391e-05, -1.0356091e-03, 9.7134936e-03, -4.9905604e-02,
1.5441467e-01, -3.0250681e-01, 4.0053059e-01, -4.1300821e-01, 3.7026403e-01];
Koeff(2,4,5,:)= [-1.8077878e-04, 3.9878467e-03, -3.6744750e-02, 1.8356352e-01,
-5.4045381e-01, 9.5692437e-01, -1.0020332e+00, 5.9105360e-01, 3.2877590e-01];

Koeff(2,5,1,:)= [2.6345175e-05, -5.7145402e-04, 5.1482540e-03, -2.4922301e-02,
7.0042274e-02, -1.1516987e-01, 1.0597346e-01, -4.8620564e-02, 8.3076614e-03];
Koeff(2,5,2,:)= [1.5058434e-04, -3.2974238e-03, 3.0077112e-02, -1.4807275e-01,
4.2624911e-01, -7.2681652e-01, 7.0982726e-01, -3.6258727e-01, 7.6615858e-02];
Koeff(2,5,3,:)= [1.8028841e-04, -3.9927468e-03, 3.6992181e-02, -1.8629654e-01,
5.555841e-01, -1.0058569e+00, 1.1000170e+00, -7.1073776e-01, 2.4541717e-01];
Koeff(2,5,4,:)= [-2.4711898e-05, 5.4049128e-04, -4.9435004e-03, 2.4396547e-02,
-6.9283088e-02, 1.0691730e-01, -4.6111226e-02, -1.5820375e-01, 3.6420033e-01];
Koeff(2,5,5,:)= [-2.0858948e-04, 4.6223139e-03, -4.2828163e-02, 2.1550192e-01,
-6.4087444e-01, 1.152237e+00, -1.2398834e+00, 7.7532079e-01, 2.4470595e-01];
Koeff(2,5,6,:)= [-1.2391655e-04, 2.6988194e-03, -2.4445884e-02, 1.1939312e-01,
-3.4169226e-01, 5.8870223e-01, -6.2982308e-01, 5.0482856e-01, 6.0753035e-02];

Koeff(2,6,1,:)= [8.9703761e-06, -1.9399183e-04, 1.7408479e-03, -8.3829068e-03,
2.3384520e-02, -3.8024077e-02, 3.4361356e-02, -1.5262216e-02, 2.4407413e-03];
Koeff(2,6,2,:)= [8.4492504e-05, -1.8393125e-03, 1.6646172e-02, -8.1065749e-02,
2.2968900e-01, -3.8208404e-01, 3.5782333e-01, -1.6898144e-01, 3.0400512e-02];
Koeff(2,6,3,:)= [2.1748764e-04, -4.7736018e-03, 4.3679687e-02, -2.1600547e-01,
6.2603561e-01, -1.0794565e+00, 1.0758353e+00, -5.7255723e-01, 1.3210299e-01];
Koeff(2,6,4,:)= [1.4806226e-04, -3.2940904e-03, 3.0698483e-02, -1.5582471e-01,
4.7014181e-01, -8.6850248e-01, 9.9030441e-01, -7.0456135e-01, 2.9433442e-01];
Koeff(2,6,5,:)= [-1.4288251e-04, 3.1114476e-03, -2.8187667e-02, 1.3762734e-01,
-3.9211303e-01, 6.5737606e-01, -6.0228284e-01, 1.5476835e-01, 3.2871905e-01];
Koeff(2,6,6,:)= [-2.3813257e-04, 5.2381321e-03, -4.8161057e-02, 2.4029299e-01,
-7.0782676e-01, 1.2604644e+00, -1.3525890e+00, 8.7206682e-01, 1.7640883e-01];
Koeff(2,6,7,:)= [-7.7997701e-05, 1.7514168e-03, -1.6416466e-02, 8.3358502e-02,
-2.4931114e-01, 4.5022664e-01, -5.0345250e-01, 4.3452708e-01, 3.5593461e-02];

Koeff(2,7,1,:)= [2.6746899e-06, -5.7823627e-05, 5.1864177e-04, -2.4955198e-03,
6.9523589e-03, -1.1279124e-02, 1.0148785e-02, -4.4669458e-03, 6.9873336e-04];
Koeff(2,7,2,:)= [3.5836969e-05, -7.7703730e-04, 6.9958761e-03, -3.3830117e-02,
9.4899779e-02, -1.5551184e-01, 1.4214175e-01, -6.4285129e-02, 1.0611044e-02];
Koeff(2,7,3,:)= [1.5504025e-04, -3.3823765e-03, 3.0698116e-02, -1.5007440e-01,
4.2755161e-01, -7.1713576e-01, 6.8072908e-01, -3.2935608e-01, 6.2172307e-02];
Koeff(2,7,4,:)= [2.4791096e-04, -5.4610686e-03, 5.0212265e-02, -2.4998876e-01,
7.3174692e-01, -1.2816624e+00, 1.3124314e+00, -7.3524600e-01, 1.8751451e-01];
Koeff(2,7,5,:)= [7.5211016e-05, -1.7037696e-03, 1.6247373e-02, -8.5010876e-02,
2.6747896e-01, -5.2610056e-01, 6.6615604e-01, -5.7100343e-01, 3.1361145e-01];
Koeff(2,7,6,:)= [-2.5883030e-04, 5.6237423e-03, -5.0803014e-02, 2.4713294e-01,
-7.0113824e-01, 1.1750158e+00, -1.1138110e+00, 4.4851799e-01, 2.7964296e-01];
Koeff(2,7,7,:)= [-2.1781395e-04, 4.8840534e-03, -4.5739562e-02, 2.3220176e-01,
-6.9494713e-01, 1.2551151e+00, -1.3662186e+00, 9.0982580e-01, 1.2368496e-01];
Koeff(2,7,8,:)= [-4.0029626e-05, 8.7427984e-04, -8.1296968e-03, 4.2064982e-02,
-1.3254426e-01, 2.6155878e-01, -3.3157738e-01, 3.4601380e-01, 2.2064040e-02];

%lambda in (0,10)
Koeff(3,2,1,:)= [6.0559551e-07, -2.6701567e-05, 4.9160778e-04, -4.9047064e-03, 2.8817011e-02,
-1.0173472e-01, 2.1280699e-01, -2.5736156e-01, 1.8539938e-01];
Koeff(3,2,2,:)= [3.2338560e-21, -1.6947843e-19, 3.5847824e-18, -3.8755614e-17, 2.2496693e-16,
-6.7390696e-16, 8.9321234e-16, -2.2927263e-16, 5.0000000e-01];
Koeff(3,2,3,:)= [-6.0559535e-07, 2.6701561e-05, -4.9160767e-04, 4.9047054e-03,
-2.8817006e-02, 1.0173471e-01, -2.1280697e-01, 2.5736155e-01, 3.1460062e-01];

Koeff(3,3,1,:)= [4.7414767e-07, -2.0738373e-05, 3.7765443e-04, -3.7090391e-03,
2.1272690e-02, -7.2106604e-02, 1.3941354e-01, -1.3999375e-01, 5.7809565e-02];
Koeff(3,3,2,:)= [7.2243269e-07, -3.2031890e-05, 5.9428715e-04, -5.9947077e-03,
3.5813546e-02, -1.2990040e-01, 2.8481501e-01, -3.7476710e-01, 3.0396632e-01];
Koeff(3,3,3,:)= [-4.7414780e-07, 2.0738378e-05, -3.7765452e-04, 3.7090400e-03,
-2.1272695e-02, 7.2106619e-02, -1.3941357e-01, 1.3999377e-01, 4.4219043e-01];
Koeff(3,3,4,:)= [-7.2243261e-07, 3.2031886e-05, -5.9428709e-04, 5.9947072e-03,
-3.5813543e-02, 1.2990039e-01, -2.8481500e-01, 3.7476709e-01, 1.9603368e-01];

Koeff(3,4,1,:)= [-5.8914868e-07, 2.6164170e-05, -4.8692704e-04, 4.9387008e-03,
-2.9806648e-02, 1.1045529e-01, -2.5472635e-01, 3.7712825e-01, 1.2551618e-01];
Koeff(3,4,2,:)= [1.9739810e-07, -8.5779251e-06, 1.5486573e-04, -1.5029429e-03,
8.4718106e-03, -2.7953784e-02, 5.1636691e-02, -4.7540886e-01, 1.6238696e-02];
Koeff(3,4,3,:)= [9.5413108e-07, -4.1831773e-05, 7.6429136e-04, -7.5421166e-03,
4.3571238e-02, -1.4943512e-01, 2.9489978e-01, -3.0768009e-01, 1.3663423e-01];
Koeff(3,4,4,:)= [3.3702565e-07, -1.5246440e-05, 2.9051119e-04, -3.0424922e-03,
1.9223916e-02, -7.6197455e-02, 1.9355396e-01, -3.2381904e-01, 3.5793868e-01];

```

```

Koeff(3,4,5,:)= [ -9.0043743e-07 , 3.9535569e-05 , -7.2349092e-04 , 7.1555739e-03 ,
                 -4.1493712e-02 , 1.4322151e-01 , -2.8548736e-01 , 3.0198229e-01 , 3.6366151e-01];

Koeff(3,5,1,:)= [6.0744170e-08, -2.6238607e-06, 4.7012242e-05, -4.5173395e-04,
                 2.5121975e-03, -8.1310969e-03, 1.4584878e-02, -1.2785608e-02, 3.9830958e-03];
Koeff(3,5,2,:)= [ 5.6101883e-07, -2.4410531e-05, 4.4148017e-04, -4.2951225e-03,
                 2.4299838e-02, -8.0644327e-02, 1.5043184e-01, -1.4104493e-01, 5.0027469e-02];
Koeff(3,5,3,:)= [ 1.0932695e-06, -4.8258009e-05, 8.8947452e-04, -8.8813638e-03,
                 5.2162640e-02, -1.8334881e-01, 3.7628734e-01, -4.1985999e-01, 2.0964131e-01];
Koeff(3,5,4,:)= [ 1.5316921e-08, -6.0421751e-07, 1.1620738e-05, -1.5947969e-04,
                 1.6973151e-03, -1.2957376e-02, 6.5492532e-02, -2.0777473e-01, 3.7194324e-01];
Koeff(3,5,5,:)= [ -1.1492425e-06, 5.0578720e-05, -9.2994956e-04, 9.2687672e-03,
                 -5.4381085e-02, 1.9105202e-01, -3.9173672e-01, 4.3537715e-01, 2.8482643e-01];
Koeff(3,5,6,:)= [ -5.8119709e-07, 2.5321860e-05, -4.5970985e-04, 4.5196128e-03,
                 -2.6294442e-02, 9.4039499e-02, -2.1507367e-01, 3.4609612e-01, 7.9577218e-02];

Koeff(3,6,1,:)= [ 1.6029521e-08, -6.9002683e-07, 1.2309788e-05, -1.1761573e-04,
                 6.4909966e-04, -2.0781559e-03, 3.6664075e-03, -3.1264057e-03, 9.2373764e-04];
Koeff(3,6,2,:)= [ 2.1736129e-07, -9.4103435e-06, 1.6909255e-04, -1.6308819e-03,
                 9.1156977e-03, -2.9716254e-02, 5.3880180e-02, -4.8073970e-02, 1.5467971e-02];
Koeff(3,6,3,:)= [ 9.0693220e-07, -3.9594851e-05, 7.1924900e-04, -7.039053e-03,
                 4.0155452e-02, -1.3491150e-01, 2.5658797e-01, -2.4866730e-01, 9.3735139e-02];
Koeff(3,6,4,:)= [ 9.6594447e-07, -4.2785161e-05, 7.9269437e-04, -7.9800168e-03,
                 4.7514665e-02, -1.7112906e-01, 3.6767883e-01, -4.4803769e-01, 2.6146350e-01];
Koeff(3,6,5,:)= [ -4.1860081e-07, 1.8358273e-05, -3.3473004e-04, 3.2761086e-03,
                 -1.8450171e-02, 5.8303870e-02, -8.1844847e-02, -5.5095213e-02, 3.5551502e-01];
Koeff(3,6,6,:)= [ -1.1804009e-06, 5.1962537e-05, -9.5614084e-04, 9.5496186e-03,
                 -5.6308177e-02, 2.0011834e-01, -4.2144394e-01, 4.9787934e-01, 2.2126360e-01];
Koeff(3,6,7,:)= [ -5.0472193e-07, 2.2065535e-05, -4.0107150e-04, 3.9309416e-03,
                 -2.2630057e-02, 7.9303114e-02, -1.7839253e-01, 3.0505318e-01, 5.1640567e-02];

Koeff(3,7,1,:)= [ 9.7131458e-09, -3.7649322e-07, 6.0477484e-06, -5.2030510e-05,
                 2.5855042e-04, -7.4529820e-04, 1.1837215e-03, -9.0832229e-04, 2.4120535e-04];
Koeff(3,7,2,:)= [ 1.6440517e-07, -6.3970209e-06, 1.0325493e-04, -8.9390332e-04,
                 4.4794311e-03, -1.3065719e-02, 2.1121352e-02, -1.6680749e-02, 4.6709620e-03];
Koeff(3,7,3,:)= [ 1.0446168e-06, -4.0861395e-05, 6.6402833e-04, -5.8003828e-03,
                 2.9426971e-02, -8.7381743e-02, 1.4522848e-01, -1.2022237e-01, 3.6831283e-02];
Koeff(3,7,4,:)= [ 2.6061091e-06, -1.0285699e-04, 1.6913288e-03, -1.5014341e-02,
                 7.7956103e-02, -2.3979446e-01, 4.2228647e-01, -3.8783435e-01, 1.4559239e-01];
Koeff(3,7,5,:)= [ 1.5765621e-06, -6.3035570e-05, 1.0562022e-03, -9.6503447e-03,
                 5.2527335e-02, -1.7561044e-01, 3.6215287e-01, -4.5329414e-01, 3.0043899e-01];
Koeff(3,7,6,:)= [ -1.6183111e-06, 6.4884288e-05, -1.0845701e-03, 9.7906864e-03,
                 -5.1626982e-02, 1.5966838e-01, -2.6545757e-01, 1.2512510e-01, 3.1890716e-01];
Koeff(3,7,7,:)= [ -2.7861463e-06, 1.0953817e-04, -1.7989306e-03, 1.6025854e-02,
                 -8.4300606e-02, 2.6824551e-01, -5.1330432e-01, 5.7650578e-01, 1.6164769e-01];
Koeff(3,7,8,:)= [ -9.4581899e-07, 3.7403399e-05, -6.1421774e-04, 5.4292408e-03,
                 -2.8058740e-02, 8.7207868e-02, -1.7151258e-01, 2.7646531e-01, 3.1785399e-02];

```

```

% Fallunterscheidung, ob
% lambda \in [0,2] lambda \in [2,5] oder lambda \in [5,10]
if (lambda<2)
dec=1;
else
if (lambda<5)
dec=2;
else
dec=3;
end
end

% Berechnung des Ausgleichspolynoms
for l=1:S+1
kappa(l)=Koeff(dec,S,l,1);
for k=1:8
kappa(l)=kappa(l).*lambda+Koeff(dec,S,l,1+k);
end
end

if(Quantil==1)

```

```
% p-Wert bestimmen
y=0*Ergebnis+kappa(S+1);
for i=1:S
y=y+kappa(i)*chi2cdf(Ergebnis,S+1-i);
end
if (y>1) y=1; end
y=1-y;
display('Der p-Wert betraegt:');
y
else
erg=quantilber2(kappa,[linkeSeite,rechteSeite],Feinheit,Quantil);
if (Ergebnis>erg)
display('Hypothese verwerfen');
else
display('Hypothese nicht verwerfen');
end
end
end
```

```
function [erg,x2,x] = Test(S,Anzahl1,Anzahl2)

% Der Vektor mit den Auspraegungen der Zufallsvariable Z_I
% sei bereits gespeichert in der Datei 'vektor'
% Diese Funktion bestimmt die Testgroesse

vektor = load('vektor');
y = zeros(1,S+2);
n = size(vektor);
n = n(2);

% Bestimmen der absoluten Haeufigkeiten aus den Daten
lam=0;
for k = 1:n
for i=0:S
if (vektor(k) == i) y(i+1) = y(i+1) + 1;
end
end
if (vektor(k) > S) y(S+2) = y(S+2) + 1;
end
lam= lam+ vektor(k);
end
lam=lam/n;

% Beginn fuer einfache Poissonverteilung
A2 = [];
b2 = [];
Aeq2 = [];
beq2 = [];
ub2 = [];
lb2 = 0;

x2 = lam;
% Als Startwert der Iteration wird der Mittelwert genommen

for i=1:Anzahl1
[f,H] = myfun2(x2,y,S,n);
x2 = quadprog(-H,-f,A2,b2,Aeq2,beq2,lb2,ub2,x2,optimset('LargeScale','off'));
end

% Berechnen des Funktionswertes an der Stelle x2
erg2 = myfun2(x2,y,S,n);

% Beginn fuer zusammengesetzte Poissonverteilung
A = [0,ones(1,S)];
b =1;
```

```
Aeq=[];
beq=[];
lb=zeros(S+1,1);
ub=[];

jump=ones(S,1)*1/S;
x=[x2;jump];
% Als Startwert fuer lambda dient der
% Optimalwert der einfachen Poissonverteilung

for i=1:Anzahl2
    % In einigen Extremfaellen wird lambda als Optimalwert = 0
    % dies wird mit der nachfolgenden Zeile verhindert
    if (x(1) < 0.001) x(1) =0.1;end

% Beginn der eigentlichen Iteration
[f,H] = myfun(x,y,S,n);
[x,fval] =quadprog(-H,-f,A,b,Aeq,beq,lb,ub,x,optimset('LargeScale','off'));
end

% Berechnen des Funktionswertes an der Stelle x
erg1 = myfun(x,y,S,n);

% Gesamtergebnis
erg = 2 *n * (erg1 - erg2);
```



```

function [f,H]=myfun(x,y,S,n)
% Diese Funktion bestimmt den Funktionswert sowie
% Ableitung und Hessematrix

%Berechnung von \Psi_k^S
Psi(1)=exp(-x(1));
sum=Psi(1);
for k=1:S
sum2=0;
for i=1:k
sum2=sum2+i*x(i+1)*Psi(k-i+1);
end
Psi(k+1) = x(1)/k*sum2;
sum=sum+Psi(k+1);
end
a=1-sum;
Psi(S+2)=a;

% Wird die Funktion nur mit einer Ausgabe aufgerufen,
% liefert sie als Ausgabe den Funktionswert
if (nargout==1)
f=0;
for k=0:S+1
f=f+y(k+1)*log(Psi(k+1));
end
end % Ende if(nargout==1)

% Bei Aufruf der Funktion mit zwei Ausgabeparametern
% wird der veraenderte Gradient und die Hessematrix bestimmt
if (nargout==2)
% Gradient
% Berechnung der Abkuerzungen
r(1)=0;
s(1)=0;

for k=0:S
b(k+1) = y(k+1) /Psi(k+1) - y(S+2)/a;
end

for k=1:S
sum=0;
for l=1:k
sum=sum+x(l+1) *Psi(k-l+1);
end
r(k+1)=sum;
end

```

```

for k=1:S
sum=0;
for l=1:k
sum=sum+x(l+1) *r(k-l+1);
end
s(k+1)=sum;
end

% rMinus steht fuer  $1-\sum_{k=0}^s r_k$ 
% sMinus analog
sMinus=0;
rMinus=0;
for k=0:S
sMinus = sMinus +s(k+1);
rMinus = rMinus +r(k+1);
end
rMinus=1-rMinus;
sMinus=1-sMinus;
sum=0;
for k=0:S
sum=sum+b(k+1)*(r(k+1) - Psi(k+1));
end
grad(1)=sum;

for i=1:S
sum=0;
for k=i:S
sum=sum+b(k+1)*Psi(k-i+1) ;
end
grad(i+1)=x(1)*sum;
end

% Hessematrix
sum1=0;
sum2=0;
for k=0:S
sum1 = sum1 + y(k+1)*(r(k+1)/Psi(k+1))^2;
sum2 = sum2 + y(k+1)*s(k+1)/Psi(k+1);
end
Hess(1,1) = -sum1+sum2 -y(S+2)/a^2*rMinus^2+y(S+2)/a*sMinus;
for j=1:S
sum1=0;
sum2=0;
sum3=0;
sum4=0;

```

```

sum5=0;
for k=j:S
sum1 = sum1 + y(k+1)*r(k+1)*Psi(k+1-j)/Psi(k+1)^2;
sum2 = sum2 + y(k+1)*Psi(k+1-j)/Psi(k+1);
sum3 = sum3 + y(k+1)*r(k+1-j)/Psi(k+1);
sum4 = sum4 + Psi(k+1-j);
sum5 = sum5 + r(k+1-j);
end
Hess(1,j+1)=-x(1)*sum1+sum2+x(1)*sum3-(-x(1)*y(S+2)/a^2*rMinus*sum4+y(S+2)
/a*sum4+x(1)*y(S+2)/a*sum5);

Hess(j+1,1)=Hess(1,j+1);
sum(j)=sum4;
end
for i=1:S
for j = i:S
sum1=0;
sum2=0;
sum3=0;
for k=(i+j):S
sum1 = sum1 + y(k+1)*Psi(k+1-i-j)/Psi(k+1);
sum2 = sum2 + Psi(k+1-i-j);
end
for k=j:S %j ist das Maximum (j >= i)
sum3 =sum3 + y(k+1)*Psi(k+1-j)*Psi(k+1-i)/Psi(k+1)^2;
end
Hess(i+1,j+1)=x(1)^2*sum1 - x(1)^2 *sum3 -x(1)^2*y(S+2)
/a*sum2-x(1)^2*sum(i)*sum(j)*y(S+2)/a^2;

Hess(j+1,i+1)=Hess(i+1,j+1);
end
end
f=(grad'-Hess*x);
H=Hess;

end % Ende von if nargout == 2

```

```

function [f,H] = myfun2(x,y,S,n)

c1=0;
c2=0;
c3=0;
% cumprod liefert die Fakultät
fak=[1,cumprod(1:S)];

% Berechnung der Konstanten
for k=0:S
c1 = c1 + y(k+1) * k;
c2 = c2 + y(k+1) *log(fak(k+1));
c3 = c3 + x^k / fak(k+1);
end
c3 = exp(x) - c3;

% Wird die Funktion nur mit einem Ausgabeparameter
% aufgerufen, so liefert sie den Funktionswert
if (nargout==1)
f = -x*n +c1 *log(x) - c2 + y(S+2) *log(c3);
end % Ende von if (nargout==1)

% Wird die Funktion mit zwei Ausgabeparametern
% aufgerufen, so liefert sie die erste und zweite Ableitung
if (nargout==2)
grad = y(S+2) - n +c1 / x+ (y(S+2) * x^S) / (fak(S+1) * c3);
H=-c1/x^2 + y(S+2)*(S-x)*x^(S-1)/(c3*fak(S+1)) -
      y(S+2)*(x^S/(fak(S+1)*c3))^2;
f=(grad'-H*x);
end % Ende von if (nargout==2)

```

```

function [quantil,prozent] = quantilber2(kappa,bereich,feinheit,percentage)

% Diese Funktion liefert folgende Ergebnisse:
% quantil   percentage-Quantil der Verteilungsfunktion (=kritischer Wert)
% prozent   genauer Wert (aufgrund der Diskretisierung von perc. abweichend)
%           des zu kritischer Wert gehoerenden Quantils

% Als Eingabe benoetigt die Funktion:
% kappa     enthaelt die Kappawerte von kappa0 bis kappaS
% bereich   ist der Bereich in dem gesucht wird
% feinheit  gibt an, wieviele Stuetzstellen genommen werden
% percentage ist das zu berechnende Quantil

dim=length(kappa);
a=bereich(1);
b=bereich(2);
c=feinheit;
Schrittweite=(b-a)/feinheit;

%Funktionswerte berechnen
x=a:Schrittweite:b;
y=0*x+kappa(dim);
for i=1:dim-1
y=y+kappa(i)*chi2cdf(x,dim-i);
end

% Quantilbestimmung
% Fehler, falls bereich zu klein gewaehlt wurde
if y(feinheit+1)<=percentage
display('Rechte Grenze zu niedrig gewaehlt.
                                               Der groesste Funktionswert betraegt:')
display(y(feinheit+1))
quantil=0;
prozent=0;
elseif y(1)>=percentage
display('Linke Seite zu hoch gewaehlt.
                                               Der kleinste Funktionswert betraegt:')
display(y(1))
quantil=0;
prozent=0;
else
% Die Funktion abgleich bestimmt mittels Bisektion das gewuenschte Quantil
[p,prozent]=abgleich(y,percentage);
quantil=a+Schrittweite*(p-1);
end

```

```
function [Reihe,vektorWert]=abgleich(vektor,Wert)

% Diese Funktion berechnet das Element des Vektors, der als erster den
% kritischen Wert (hier mit Wert bezeichnet) ueberschreitet.
% Alg. gilt nur fuer monoton wachsende Folgen (Verteilungsfunktionen)

% Reihe gibt an, das wievielte Element der krit. Wert in der Reihenfolge ist
% vektorWert gibt den Wert dieses Elements aus

% Zur Berechnung wird das Bisektionsverfahren verwendet

niedrigerWert=1;
hoherWert = length(vektor);

while hoherWert~=niedrigerWert
if hoherWert-niedrigerWert==1
vektorWert=vektor(hoherWert);
Reihe=hoherWert;
break;
else
Mitte=round(0.5*(hoherWert+niedrigerWert));
if vektor(Mitte)>Wert
hoherWert=Mitte;
elseif vektor(Mitte)<Wert
niedrigerWert=Mitte;
else
vektorWert=vektor(Mitte);
Reihe=Mitte;
break;
end
end
vektorWert=vektor(Mitte);
Reihe=Mitte;
end
```

Anhang C

MAPLE Datei zur Bestimmung von κ_4 im Fall $S = 4$

```
with(LinearAlgebra):
z:=Matrix(10):
for i from 0 by 1 to 9 do for k from 0 by 2 to 8 do
x:=(1+5*i+k/2)/50.:
Aminus1:=Transpose(
  < <-exp(x),0,0,0,0>|
  <exp(x)/x*(1-x),exp(x)/x,0,0,0>|
  <exp(x)*x/2,-exp(x),exp(x)/x,0,0>|
  <-exp(x)*x^2/6,exp(x)*x/2,-exp(x),exp(x)/x,0>|
  <exp(x)*x^3/24,-exp(x)*x^2/6,exp(x)*x/2,-exp(x),exp(x)/x>>):
N:=Transpose(<<0,0,0,0>|<1,0,0,0>|<1,-1,0,0>|<1,0,-1,0>|<1,0,0,-1>>):
qt:=<exp(-x)|x*exp(-x)|1/2*exp(-x)*x^2|1/6*exp(-x)*x^3|1/24*exp(-x)*x^4>:
q:=Transpose(qt):
qq:=q.qt:
Dminus1:=DiagonalMatrix(q)-qq:
Q:=simplify(Transpose(N).Aminus1.Dminus1.Transpose(Aminus1).N):
Q:=MatrixInverse(Q):
Q[1,2]:=Q[1,2]/sqrt(Q[1,1]*Q[2,2]):
Q[1,3]:=Q[1,3]/sqrt(Q[1,1]*Q[3,3]):
Q[1,4]:=Q[1,4]/sqrt(Q[1,1]*Q[4,4]):
Q[2,3]:=Q[2,3]/sqrt(Q[2,2]*Q[3,3]):
Q[2,4]:=Q[2,4]/sqrt(Q[2,2]*Q[4,4]):
Q[3,4]:=Q[3,4]/sqrt(Q[3,3]*Q[4,4]):
Q[2,1]:=Q[1,2]:
Q[3,1]:=Q[1,3]:
Q[4,1]:=Q[1,4]:
Q[3,2]:=Q[2,3]:
Q[4,2]:=Q[2,4]:
Q[4,3]:=Q[3,4]:
Q[1,1]:=1:
```

```

Q[2,2]:=1:
Q[3,3]:=1:
Q[4,4]:=1:
Korr:=Q:
erg1:=Korr[1,4]*(Korr[1,3]-Korr[1,2]*Korr[2,3])+
      Korr[2,4]*(Korr[2,3]-Korr[1,2]*Korr[1,3]):
erg2:=sqrt((1-Korr[1,2]^2-Korr[1,3]^2-Korr[2,3]^2+
            2*Korr[1,2]*Korr[1,3]*Korr[2,3])*(1-Korr[1,2]^2-Korr[1,4]^2-
            Korr[2,4]^2+2*Korr[1,2]*Korr[1,4]*Korr[2,4])):
Rho34stern:=(erg1-erg2)/(1-Korr[1,2]^2):
C:=< <Korr[1,1],Korr[1,2],Korr[1,3]>|
     <Korr[2,1],Korr[2,2],Korr[2,3]>|
     <Korr[3,1],Korr[3,2],Korr[3,3]>>:
C:=MatrixInverse(C):
Rho:=<Korr[1,4],Korr[2,4],Rho34stern>:
PSI:=C.Rho:
check:=Transpose(Rho).Ps:
Kegel:=0:
positiv:=0:
negativ:=0:
p:=<0,0,0>:
n:=<0,0,0>:
for j from 1 to 3 do
if (PSI[j]<0)
then negativ:=negativ+1:
n[negativ]:=j
else
positiv:=positiv+1:
p[positiv]:=j:
end if:
end do:

Mat21:=< <Korr[1,2],Korr[1,3],Korr[1,4]>|
        <Korr[3,2],Korr[3,3],cos(theta)>|
        <Korr[4,2],cos(theta),Korr[4,4]>>:
Mat22:=< <Korr[2,2],Korr[2,3],Korr[2,4]>|
        <Korr[3,2],Korr[3,3],cos(theta)>|
        <Korr[4,2],cos(theta),Korr[4,4]>>:
Mat23:=< <Korr[1,1],Korr[1,3],Korr[1,4]>|
        <Korr[3,1],Korr[3,3],cos(theta)>|
        <Korr[4,1],cos(theta),Korr[4,4]>>:
sigma12:=Determinant(Mat21)/sqrt(Determinant(Mat22)*Determinant(Mat23)):
Integral:=-1/4/Pi/Pi*int(arccos(sigma12),
                       theta=arccos(Rho34stern)..arccos(Korr[3,4])):

```



```
Korr[3,4]:=Rho34stern:
if (negativ=0) then
  Kegel:=1/4/Pi*(arccos(-Korr[1,2])+
    arccos(-Korr[1,3])+arccos(-Korr[2,3])-Pi):
end if:

if (negativ=1) then
  Kegel:=1/4/Pi*(arccos(-Korr[p[1],n[1]])+arccos(-Korr[p[1],4])+
    arccos(-Korr[p[2],n[1]])+arccos(-Korr[p[2],4])-2*Pi):
end if:

if (negativ=2) then
  Kegel:= 1/4/Pi*(arccos(-Korr[n[1],4])+
    arccos(-Korr[n[2],4])+arccos(-Korr[n[1],n[2]])-Pi):
end if:

Endergebnis:=evalf(Kegel+Integral);
z[i+1,k+1]:=x:
z[i+1,k+2]:=Endergebnis;
end do:
end do:
z;
```

Literaturverzeichnis

- [Ba91] Bauer, Heinz , *Wahrscheinlichkeitstheorie*, de Gruyter, 1991
- [Bu01] Buchmann, Boris, *Decompounding: An Estimation problem for the Compound Poisson Distribution*, Doctoral thesis, Universität Hannover, 2001
- [Co35] Coxeter, H. S. M., *The functions of Schläfli and Lobatschewsky*, Quart. J. Math 6, 13-20
- [Ha67] Haight, Frank A., *Handbook of the Poisson distribution*, John Wiley & Sons Inc, 1967
- [Ma97] Mack, Thomas, *Schadenversicherungsmathematik*, Versicherungswirtschaft e.V. Karlsruhe, 1997
- [Pl54] Plackett, R.L., *A reduction formula for normal multivariate integrals*, Biometrika, 41, p. 352- 360, 1954
- [Po84] Pollard, David, *Convergence of Stochastic Processes*, Springer Series in Statistics, 1984
- [Po66] Potthoff, Richard F. und Whittinghill, Maurice, *Testing for homogeneity, II. The Poisson distribution*, Biometrika, 53, p. 183 - 190, 1966
- [Qu87] Quine, M.P., Seneta, E., *Bortkiewicz Data and the Law of Small Numbers*, International Statistical Review 55, no.2, p. 173-181, 1987
- [Ru54] Ruben, H., *On the moments of order statistics in samples from normal population*, Biometrika, 41, 200-27
- [Sc99] Schlüchtermann, Georg, Vorlesungsskript: *Preisbildung von Derivaten*, LMU München, 1999 (www.mathematik.uni-muenchen.de/~schluech/vorl.html)
- [To90] Tong, Y. L., *The Multivariate Normal Distribution*, Springer- Verlag, 1990
- [Va00] van der Vaart, Aad W.; *Asymptotic Statistics*, Cambridge University Press; 2000