

**Friedrich-Alexander-Universität  
Erlangen-Nürnberg**  
Wirtschafts-und Sozialwissenschaftliche Fakultät

Diskussionspapier

82 / 2007

Untersuchung asymptotischer Eigenschaften von  
Schätzern diskreter bivariater Copula Modelle mit  
Kovariablen

Nina Meinel



---

Lehrstuhl für Statistik und Ökonometrie  
Lange Gasse 20 · D-90403 Nürnberg

# Untersuchung asymptotischer Eigenschaften von Schätzern diskreter bivariater Copula Modelle mit Kovariablen

NINA MEINEL

*Friedrich-Alexander-Universität Erlangen-Nürnberg,  
Volkswirtschaftliches Institut, Lehrstuhl für Statistik und Ökonometrie*

*Email: nina.meinel@wiso.uni-erlangen.de*

## SUMMARY

Diskrete Copula Modelle bilden die Abhängigkeiten zwischen multiplen kategorialen Responses sowie die Einflüsse von Kovariablen auf die jeweiligen Responses ab. In einer Simulationsstudie soll das Verhalten von Schätzern diskreter Copula Modelle bei unterschiedlichen Strukturen der Kovariablen untersucht werden. Insbesondere wird Schiefe in der Verteilung der Kovariablen und das Problem der Multikollinearität zwischen den Kovariablengruppen der einzelnen Responses betrachtet.

*Keywords and phrases:* Copula; inference functions of margins; Jackknife; Schiefe; Multikollinearität

## 1 Einleitung

Eine Vielzahl von erhobenen Daten im Bereich der Marktforschung, der Sozialforschung oder im Bereich der Arbeitsmarktforschung beinhalten kategoriale Variablen. Diese können binärer, ungeordneter oder geordneter Natur sein oder es kann sich dabei um so genannte Zähldaten handeln. Ziel und Zweck von Erhebungen sind die Erklärung einer abhängigen Variablen durch ihre Einflussgrößen. Oftmals ist jedoch nicht nur eine abhängige Variable von Interesse, sondern ein ganzes System von Responses, in welchem Interdependenzen auftreten können. Beispielsweise sollen die Kundenzufriedenheit und die Kundenbindung durch bestimmte Kovariablen erklärt werden. Dabei liegt die Vermutung nahe, dass die beiden Variablen nicht unabhängig voneinander sind. Ein weiteres Beispiel wäre die Modellierung von Arbeitsmotivation und Arbeitszufriedenheit. Damit sind Modelle gefordert, welche die Abhängigkeitsstruktur zwischen den Responses und den Einfluss auf die einzelne Response gemeinsam abbilden können. Eine Modellierung von zeitlichen Abhängigkeiten zwischen den Endogenen, wie sie beispielsweise bei Paneldaten auftreten können, ist durchaus möglich, soll in dieser Arbeit aber nicht betrachtet werden.

Copula Modelle bieten eine Möglichkeit, Zusammenhänge zwischen den kategorialen Variablen zu untersuchen und zudem die Einflussfaktoren von Kovariablen auf die jeweilige Responsevariable zu modellieren. Sie stellen eine Erweiterung bereits bekannter univariater Modelle dar, wie z.B. das Logit oder Probit Modell. Die Parameter der Einflussgrößen der einzelnen Responses und die Abhängigkeitsparameter zwischen den Responses können gemeinsam mittels der Maximum-Likelihood-Methode geschätzt werden, welche auf einer

multivariaten Likelihood-Funktion beruht. Eine weitere Möglichkeit der Schätzung ist die Methode der "Inference Functions for Margins". Dieses Verfahren schätzt die Parameter der Einflussgrößen der einzelnen Responses anhand ihrer univariaten Likelihoods. Die Abhängigkeitsparameter werden hingegen mit Hilfe der multivariaten Likelihood-Funktion, gegeben der Werte der univariaten Parameter, ermittelt. Aufgrund des Vorgehens wird diese auch als eine zweistufige Maximum-Likelihood-Schätzung bezeichnet. Simulationsstudien von Xu und Joe (1996)[7] bestätigen die Effizienz der "Inference Functions for Margins"-Methode im Vergleich zum Maximum-Likelihood-Verfahren.

Wenn Kovariablen ähnliche oder sogar gleiche Konstrukte messen, korrelieren sie stark miteinander. Dies wird bezeichnet als ein Multikollinearitätsproblem. Die Konsequenzen für die univariate Schätzung sind ungenaue Schätzergebnisse und hohe Standardfehler (Verbeek S. 42f)[15]. Bei einer multivariaten Betrachtung wird jede Responsevariable  $y_i$  durch eine Kovariablenmatrix  $X_i$  und durch den Zusammenhang zwischen den Responses erklärt. In dieser Arbeit soll der Frage nachgegangen werden, welche Konsequenzen sich für die Schätzer der Maximum-Likelihood- bzw. der "Inference Functions for Margins"-Methode ergeben, wenn die Kovariablen  $X_i$  der Endogenen  $y_i$  mit den Kovariablen  $X_j$  der Endogenen  $y_j$  stark korrelieren? Multikollinearitäten können zum Einen auftreten, wenn unterschiedliche Variablen aus beiden Kovariablen-Gruppen  $X_i$  und  $X_j$  ähnliche Konstrukte messen und damit miteinander korrelieren. Zum Anderen ist der Effekt von Kovariablen nicht nur auf eine einzige Response beschränkt, stattdessen können sie gleichzeitig auf mehrere wirken. Ist dies der Fall, tritt eine perfekte Multikollinearität zwischen  $X_i$  und  $X_j$  auf.

Bei Einstellungsvariablen kann, durch die Tendenz zu bestimmten Antwortkategorien, eine schiefe Verteilung der Kovariablen auftreten. Dabei kann die Schiefe der Kovariablen in gleicher Richtung vorliegen, d.h. alle Kovariablen weisen entweder eine Links- oder eine Rechtsschiefe auf. Sie könnte jedoch auch in unterschiedlicher Richtung vorhanden sein, d.h. die Kovariablen weisen eine Links- und eine Rechtsschiefe auf. Der Einfluss der Schiefe der Kovariablen auf die Eigenschaften der Schätzer der beiden genannten Methoden wurde bisher wenig in der Literatur betrachtet und wird in dieser Arbeit ebenfalls untersucht.

Anhand einer Simulationsstudie soll das Verhalten der Maximum-Likelihood- bzw. das der "Inference Functions for Margins"-Schätzer und deren Effizienz bei Vorliegen von Multikollinearitäten und Schiefe in den Kovariablen genauer betrachtet werden.

Im folgenden Abschnitt wird zunächst ein Überblick über die multivariate Modellierung von diskreten Variablen gegeben. Anschließend erfolgen die Definition der Copula sowie eine Beschreibung der diskreten Copula Modelle und deren Schätzmethode. Die Simulationsstudie, die Kriterien des Effizienzvergleiches zwischen beiden Schätzern und die einzelnen Ergebnisse werden im dritten Abschnitt vorgestellt. Eine Zusammenfassung der Ergebnisse und eine darauf aufbauende Schlussfolgerung folgen im letzten Teil der Arbeit.

## 2 Diskrete Modellierung

### 2.1 Überblick multivariate diskrete Modellierung

Die Literatur bietet eine Vielzahl an Möglichkeiten für die Modellierung von diskreten Responses an. Diese sollten allerdings bestimmte Ansprüche erfüllen. Mindestanforderungen, die an ein Modell für multiple kategoriale Daten gestellt werden, sind:

- Beschreibung der Abhängigkeiten zwischen den Responses,
- Berücksichtigung von Einflüssen durch Kovariablen,
- Ausschöpfung der vollständigen Information der vorliegenden diskreten Skala.

Die Vielzahl an Modellen kann allgemein unterschieden werden in unverbundene, scheinbar unverbundene und interdependente Systeme.

Bei *unverbundenen* Systemen werden gleichzeitig mehrere univariate Modelle geschätzt, wobei keinerlei Abhängigkeiten unterstellt werden. Diese erfüllen die Anforderungen an ein multiples Modell jedoch nicht, da Interdependenzen nicht berücksichtigt werden.

In *scheinbar unverbundenen* Modellen sind die Abhängigkeiten im Gleichungssystem nicht direkt zu erkennen, da eine abhängige Variable nicht als Erklärende einer anderen Responses auftritt. Über die Annahme einer multivariaten Verteilung werden die Interdependenzen in das Modell aufgenommen. Durch die Berücksichtigung von Kovariablen und von unterschiedlichen Skalenniveaus sind die oben genannten Anforderungen erfüllt. Die wohl bekanntesten Vertreter dieser Modellklasse sind das Multivariate Probit Modell von Ashford und Sowden (1970)[1], das "Exchangeable Mixture Modell" von Prentice (1986)[13], das Plackett Modell von Dale (1986)[3] und das Modell von Koehler und Symanowski (1992)[8].

*Interdependente* Systeme sind dadurch gekennzeichnet, dass Abhängigkeiten zwischen einzelnen Responses im Gleichungssystem direkt erkennbar sind. Zum Einen kann eine abhängige Variable als Einflussgröße für eine andere abhängige Variable auftreten. Hierzu zählen unter anderem Strukturgleichungsmodelle für kategoriale Daten von Muthen (1979) und simultane Logit Modelle von Strauss und Schmidt (1975)[14]. Zum Anderen werden bei den saturierten log-linearen Modellen von Cox (1972)[2] die Interdependenzen zwischen den Responses direkt in der Gleichung formuliert. Interdependente Systeme genügen ebenfalls den Anforderungen an eine multivariate Modellierung von diskreten Responses.

### 2.2 Copula

Abhängigkeiten zwischen Responses können unter Annahme einer multivariaten Verteilung modelliert werden. Eine weitere Möglichkeit besteht darin, diese Abhängigkeit mit Hilfe von Copulas darzustellen. Nelsen (1999, S. 5)[11] beschreibt Copulas als "function that join or couple multivariate distribution functions to their one-dimensional marginal distribution functions" und als "distribution function whose one-dimensional marginal functions are uniform". Zunächst erfolgt die Definition einer Copula und wichtiger Theoreme für den zweidimensionalen Fall, um anschließend eine Erweiterung für den multivariaten Fall zugeben.

**Definition 2.1.** Eine zwei-dimensionale Copula  $C$  ist eine Funktion, die auf dem Einheitsquadrat  $\mathbf{I}^2$  mit  $\mathbf{I} = [0, 1]$  definiert ist. Die Copula  $C(u, v)$  besitzt folgende Eigenschaften:

→ für jedes  $u$  und  $v$  in  $\mathbf{I}$  gilt:

$$\begin{aligned} C(u, 0) &= C(0, v) = 0, \\ C(u, 1) &= u, \\ C(1, v) &= v, \end{aligned}$$

→ für jedes  $u_1, u_2, v_1$  und  $v_2$  in  $\mathbf{I}$  mit  $u_1 \leq u_2$  und  $v_1 \leq v_2$  gilt:

$$C(u_2, v_2) - C(u_1, v_2) - C(u_2, v_1) + C(u_1, v_1) \geq 0.$$

Diese Bedingung wird als die "two-increasing" Bedingung bezeichnet.

Weiterhin kann für eine Copula die obere bzw. untere Fréchet-Hoeffding-Grenze mit folgender Ungleichung angegeben werden:

$$\max(u + v - 1, 0) \leq C(u, v) \leq \min(u, v).$$

Die Grenzen selbst stellen wiederum Copulas dar. Die untere Fréchet-Hoeffding-Grenze ist nur für den Fall  $d = 2$  eine Copula. Sie wird in der Literatur als die *Maximumscopula*  $W(u, v) = \max(u + v - 1, 0)$  bezeichnet und gibt die perfekte negative Abhängigkeit an. Die obere Grenze ist die so genannte *Minimumscopula*  $M(u, v) = \min(u, v)$ . Sie entspricht der perfekten positiven Abhängigkeit. Mit Definition 2.1. wird eine Copula als eine Funktion in einem bestimmten Wertebereich verstanden. Die Verbindung einer Copula zu den multivariaten Verteilung wird allerdings nicht widerspiegelt. Das zentrale Theorem der Copula-Theorie stellt der Satz von Sklar (1959) dar, welcher die Beziehung zwischen einer Copula und einer multivariaten Verteilung herstellt.

**Theorem 1.** Ist  $H$  eine Verteilungsfunktion mit den univariaten Randverteilungen  $F$  und  $G$ . Dann existiert eine Copula  $C$ , so dass für alle  $x, y$  in  $\mathbb{R}$  gilt:

$$H(x, y) = C(F(x), G(y)).$$

Falls  $F$  und  $G$  stetig sind, dann ist die Copula eindeutig. Ist dies nicht der Fall, dann ist  $C$  eindeutig auf  $\text{Ran}F \times \text{Ran}G$ .

Das Theorem von Sklar zeigt, dass sich eine mehrdimensionale Verteilungsfunktion aus den Randverteilungen und einer Copula zusammensetzt. Die Aufgabe der Copula ist es, die Abhängigkeitsstruktur zwischen den Zufallsvariablen  $X$  und  $Y$  wiederzugeben. Basierend auf der Wahrscheinlichkeitsintegraltransformation kann nun folgendes Corollary zur Konstruktion einer Copula aufgestellt werden.

*Corollary 2.1.* Sind  $X$  und  $Y$  stetig verteilte Zufallsvariablen mit einer gemeinsamen Verteilungsfunktion  $H(x, y)$  und den univariaten Randverteilungen  $F(x)$  und  $G(y)$ , dann sind  $u = F(x)$  und  $v = G(y)$  gleichverteilt im Bereich  $[0, 1]$ .  $F^{-1}$  und  $G^{-1}$  sind die Quantilsfunktionen der beiden Randverteilung. Es gilt:

$$C(u, v) = H(F^{-1}(u), G^{-1}(v)).$$

Damit kann für bereits bekannte multivariate Verteilungsfunktionen die zugehörige Copula bestimmt werden. Corollary 2.1 stellt eine Möglichkeit dar, um eine Copula zu konstruieren.

Bisher wurden nur zwei-dimensionale Zufallsvariablen betrachtet. Aufbauend auf den bisherigen Erörterungen folgt nun eine Erweiterung für den multivariaten Fall. Für einen  $d$ -dimensionalen Zufallsvariablenvektor  $\mathbf{Y} = (Y_1, \dots, Y_d)'$ , deren  $d$ -dimensionale Verteilungsfunktion mit  $H(y_1, \dots, y_d)$  gegeben ist und deren univariaten Randverteilungen mit  $F_1(y_1), \dots, F_d(y_d)$  bezeichnet werden, gilt Folgendes:

**Definition 2.2.** Eine  $d$ -dimensionale Copula  $C$  ist eine Funktion, die definiert ist auf  $\mathbf{I}^d$  mit  $\mathbf{I} = [0, 1]$ . Die  $d$ -dimensionale Copula  $C(\mathbf{u})$  mit  $\mathbf{u} = (u_1, \dots, u_d)'$  besitzt folgende Eigenschaften:

→ für jedes  $\mathbf{u}$  in  $\mathbf{I}$  gilt:

$$\begin{aligned} C(\mathbf{u}) &= 0, & \text{falls nur ein } u_i = 0, \\ C(\mathbf{u}) &= u_j, & \text{falls alle } \mathbf{u} = 1 \text{ außer } u_j, \end{aligned}$$

→ für jedes  $\mathbf{a}$  und  $\mathbf{b}$  in  $\mathbf{I}^d$  mit  $\mathbf{a} \leq \mathbf{b}$  gilt:

$$V_C([a, b]) \geq 0$$

mit

$$V_C([a, b]) = \sum \text{sgn}(c)C(c).$$

$\text{sgn}(c)$  ist gegeben mit:

$$\text{sgn}(c) = \begin{cases} 1, & \text{falls } c_k = a_k \text{ für eine gerade Zahl von } k, \\ -1, & \text{falls } c_k = a_k \text{ für eine ungerade Zahl von } k. \end{cases}$$

Diese Bedingung wird als die "d-increasing" Bedingung bezeichnet.

**Theorem 2.** Ist  $H$  eine  $d$ -dimensionale Verteilungsfunktion mit den univariaten Randverteilungen  $F_1(y_1), \dots, F_d(y_d)$ , dann existiert eine Copula  $C$ , so dass für alle  $Y$  in  $\mathbb{R}$  gilt:

$$H(y_1, \dots, y_d) = C(F(y_1), \dots, F(y_d)). \quad (2.1)$$

Falls  $F_1(y_1), \dots, F_d(y_d)$  stetig sind, dann ist die Copula eindeutig. Ist dies nicht der Fall, dann ist  $C$  eindeutig auf  $\text{Ran}F_1 \times \text{Ran}F_2 \times \dots \times \text{Ran}F_d$ .

*Corollary 2.2.* Sind  $Y = (Y_1, \dots, Y_d)'$  stetig verteilte Zufallsvariablen mit einer gemeinsamen Verteilungsfunktion  $H(y_1, \dots, y_d)$  und den univariaten Randverteilungen  $F_1(y_1), \dots, F_d(y_d)$ , dann sind  $u_1 = F_1(y_1), \dots, u_d = F_d(y_d)$  gleichverteilt im Bereich  $[0, 1]$ .  $F^{-1}$  ist die jeweilige Quantilsfunktion der univariaten Randverteilung. Es gilt:

$$C(\mathbf{u}) = H(F_1^{-1}(u_1), \dots, F_d^{-1}(u_d)). \quad (2.2)$$

Wie bereits erwähnt kann eine Copula aus einer bereits bekannten Verteilung konstruiert werden. An dieser Stelle kann der Umkehrschluss gebildet werden, dass aus einer bekannten Copula mit gegebenen Rändern eine multivariate Verteilung entwickelt werden kann. Daraus ergibt sich ein weites Einsatzgebiet von Copulafunktion. Eine weitere wichtige Anwendung ist in der skaleninvarianten Modellierung von Zusammenhängen zwischen Variablen zu sehen.

Welchen Vorteil bringt jedoch die Anwendung einer Copula anstelle der multivariaten Verteilung? Die Copula stellt die Abhängigkeiten zwischen den Variablen dar. Sie beinhaltet damit alle Informationen, welche die univariaten Verteilungen zu einer gemeinsamen Verteilung zusammenfügt. Über das Konstrukt der Copula kann somit eine multivariate Verteilung in deren Ränder und in ihre Abhängigkeitsstruktur zerlegt werden. Für die Modellierung von mehreren Responses kann so die Abhängigkeitsstruktur und die Randverteilung separat gebildet werden. Dies erlaubt eine höhere Flexibilität in der Modellierung. Je nach Problemstellung können verschiedene Verteilungen für die Ränder eingesetzt werden, ohne die Abhängigkeitsstruktur zu verändern. Für jede betrachtete Zufallsvariable kann eine andere univariate Verteilung unterstellt werden, oder es können bei gleichbleibenden Rändern verschiedene Assoziationsstrukturen betrachtet werden. Insgesamt bietet das Konstrukt der Copula eine flexiblere Einsatzmöglichkeit im Vergleich zu den multivariaten Verteilungen. Des Weiteren besteht eine Beziehung zwischen bekannten Assoziationskonzepten und Assoziationsmaßen. Dies soll hier nicht näher erläutert werden. Hingewiesen sei auf vertiefende Literatur von Nelsen (1999)[11] und Joe (1997)[4].

Das verschiedene Copulas andere Abhängigkeitsstrukturen wiedergeben und damit unterschiedliche Interpretationen gegeben sind, soll beispielhaft anhand einiger bekannter Copulas kurz veranschaulicht werden. Die Minimums- und Maximumscopula und deren jeweilige Bedeutung wurde bereits im Zusammenhang mit den Fréchet-Hoeffding-Grenzen erwähnt.

**Example 2.1.** Eine weitere bekannte Copula ist die *Unabhängigkeitscopula*, welche definiert ist als:

$$C_I(u_1, \dots, u_d) = \prod_{j=1}^d u_j.$$

In diesem Fall werden die betrachteten Zufallsvariablen als unabhängig vorausgesetzt.

**Example 2.2.** Eine weitere sehr wichtige Copula stellt die *Multivariate Normal Copula* dar. Diese ist gegeben mit:

$$C(u_1, \dots, u_d; \Theta) = \Phi_d(\Phi^{-1}(u_1), \dots, \Phi^{-1}(u_d); \Theta), \quad (2.3)$$

wobei  $\Phi$  die Standardnormalverteilungsfunktion und  $\Phi_d$  die d-variate Normalverteilung mit einem Mittelwertsvektor  $\mu = 0$  und einem Varianzvektor  $\sigma^2 = 1$  ist. Die Konstruktion dieser Copula beruht auf dem Prinzip der Gleichung (2.2). Die Matrix  $\Theta$  fungiert als Assoziationsmaß zwischen den Responses. Die einzelnen Parameter  $\theta_{jk}$  dieser Matrix können als Korrelationskoeffizient von Pearson mit dem entsprechenden Wertebereich von  $[-1,1]$  interpretiert werden. Die perfekte positive Abhängigkeit wird mit  $\theta_{jk} = 1$  erzielt, die perfekte negative bei  $\theta_{jk} = -1$  und Unabhängigkeit liegt vor, wenn  $\theta_{jk} = 0$  ist. Die Erweiterung zu einer Meta-Copula erlaubt eine höhere Flexibilität bei der Wahl der univariaten Verteilung in Kombination mit dem Zusammenhangsmaß  $\Theta$ . Diese kann dargestellt werden als:

$$C(G_1(z_1), \dots, G_d(z_d); \Theta) = \Phi_d(\Phi^{-1}(G_1(z_1)), \dots, \Phi^{-1}(G_d(z_d)); \Theta). \quad (2.4)$$

Für die Wahl von  $G_j(z_j)$  kann jede beliebige Verteilung unterstellt werden, so zum Beispiel die logistische Funktion mit  $G_j(z_j) = \exp(z_j)/(1 + \exp(z_j))$ . Der Einsatz dieser Copula für die Anwendung auf diskrete Daten wurde intensiv in der Arbeit von Xu (1996) untersucht.

**Example 2.3.** Die d-variate *Farlie-Gumbel-Morgenstern (FGM) Copula* ist definiert als:

$$C(u_1, \dots, u_d; \theta) = \prod_{i=1}^d u_i P(u_1, \dots, u_d)$$

mit dem Polynomial  $P(u_1, \dots, u_d) = 1 + \sum_{k=2}^d \sum_{1 \leq j_1 < \dots < j_k \leq d} \theta_{j_1, \dots, j_k} \prod_{i=1}^s (1 - u_i)$ . Die Parameter  $\theta_{j_1, \dots, j_k}$  stellen dabei die Assoziationen zwischen den Responses dar. Diese müssen bestimmte Restriktionen in Abhängigkeit der Dimension  $d$  erfüllen, um zu gewährleisten, dass  $C(u_1, \dots, u_d)$  eine Copula ist. Für den bivariate Fall ergibt sich beispielsweise die folgende Form:

$$C(u, v; \theta) = uv(1 + \theta(1 - u)(1 - v)),$$

wobei  $\theta$  den Zusammenhang zwischen den zwei betrachteten Variablen widerspiegelt. Es gilt für den Copulaparameter  $\theta$ , dass  $|\theta| \leq 1$  ist. Für diesen kann Kendall's tau mit  $2/9\theta$  angegeben werden. Es ergibt sich daraus ein maximaler Wertebereich von  $-2/9$  bis  $2/9$ .

**Example 2.4.** Erwähnt sei an dieser Stelle noch die *Plackett-Copula*, welche sich im bivariaten Fall definiert als:

$$C(u, v; \psi) = \frac{\{1 + (\psi - 1)(u + v) - [(1 + (\psi - 1)(u + v))^2 - 4(\psi - 1)uv]^{1/2}\}}{2(\psi - 1)}. \quad (2.5)$$

Die Plackett-Copula beruht auf der gleichnamigen Verteilung, welche auf Plackett (1969)[12] zurückzuführen ist. Für  $u$  und  $v$  kann jede beliebige Verteilungsfunktion eingesetzt werden. Der Parameter  $\psi$  der Copula stellt den Abhängigkeitsparameter dar und kann als das Odd-Ratio zwischen zwei Variablen interpretiert werden.  $\psi$  nimmt Werte zwischen 0 und  $\infty$  an, wobei mit  $\psi \rightarrow 0$  die Maximumscopula erreicht wird. Bei  $\psi \rightarrow 1$  liegt Unabhängigkeit zwischen den Variablen vor und mit  $\psi \rightarrow \infty$  wird die Minimumscopula erzielt. Eine Erweiterung für den mehrdimensionalen Fall geht auf Molenberghs und Lesaffre (1994)[10] zurück.

Auf eine ausführliche Darstellung wird hier verzichtet, da sie an späterer Stelle nochmals aufgegriffen wird.

Anzumerken ist, dass nicht jede Copula in der Lage ist, den Abhängigkeitsbereich vollständig abzudecken. Es gibt Copulas, die nur positive Zusammenhänge erfassen, wie zum Beispiel die Gumbel-Copula oder die Galambos-Copula. Einige Copulafunktionen können nur einen eingeschränkten Wertebereich wiedergeben, wie bei der bereits genannte FGM-Copula. Abschließend sei für eine tiefergehende Darstellung über die Theorie und Anwendung von Copulas auf die beiden Autoren Nelsen (1999)[11] und Joe (1997)[4] verwiesen.

### 2.3 Diskrete Copula Modelle

Die formale Darstellung des multivariaten diskreten Copula Modells beruht auf der Arbeit von Xu (1996)[16]. Für einen d-dimensionalen diskreten Zufallsvariablenvektor  $\mathbf{Y} = (Y_1, \dots, Y_d)'$  ist die Wahrscheinlichkeitsmassenfunktion mit  $P(y_1, \dots, y_d; \Theta)$  gegeben.  $\Theta$  steht ganz allgemein für einen oder mehrere Abhängigkeitsparameter. Die gemeinsame Verteilungsfunktion von  $\mathbf{Y}$  ist definiert durch:

$$G(y_1, \dots, y_d) = C(G_1(y_1), \dots, G_d(y_d)).$$

$C$  ist dabei eine d-dimensionale Copula und  $G_j(y_j)$  ist die univariate Verteilungsfunktion der diskreten Zufallsvariablen  $Y_j$ . Die Wahrscheinlichkeitsmassenfunktion  $P(y_1, \dots, y_d; \Theta)$  für diskrete Zufallsvariablen, als eine Funktion der Copula, ergibt sich aus:

$$P(y_1, \dots, y_d; \Theta) = \sum_{k_1=1}^2 \dots \sum_{k_d=1}^2 (-1)^{k_1+\dots+k_d} C(x_{1k_1}, \dots, x_{dk_d}) \quad (2.6)$$

mit  $x_{j1} = G_j(y_j - 1)$  und  $x_{j2} = G_j(y_j)$  für  $G_j(y_j - 1) < G_j(y_j)$ . Für die Konkretisierung des allgemeinen Modells für bestimmte zugrunde liegende diskrete Variablen muss zum Einen die Copula und zum Anderen die Randverteilungsfunktionen spezifiziert werden. Anhand der Unterscheidung multivariater Modelle in Abschnitt 2.1 kann das diskrete Copula Modell in die Klasse der scheinbar unverbundenen Systeme eingeordnet werden. Der Zusammenhang zwischen den Responses ist nicht direkt aus dem Gleichungssystem erkennbar, sondern ergibt sich aus der Annahme eine multivariaten Verteilung bzw. einer Copula mit entsprechenden Randverteilungen.

Für die Wahl einer geeigneten Copula sollten im Voraus bestimmte Überlegungen getroffen werden. Die folgenden Eigenschaften dienen der Auswahl möglicher Copulas und sind teilweise aus Joe (1997, S.221)[4] entnommen:

- Geschlossene Form der Copula,
- "reflection symmetry",
- Modellierung eines großen Abhängigkeitsbereiches,

→ Geeignete Interpretation des Copulaparameters,

→ Möglichkeit der Parametrisierung des Copula-Parameters.

Die geschlossene Form der Copula ist aus schätztheoretischem Zweck notwendig, um die Möglichkeit einer Schätzung überhaupt zu gewährleisten. Die Eigenschaft der "reflection symmetry" ist gegeben, falls für  $(U, V) \sim C$ , gleich  $(1 - U, 1 - V) \sim C$  gilt. Dies bedeutet, dass die Richtung der Diskretisierung einer stetigen Variablen  $Z_j$ , also  $Y_j = I(Z_j < z_j)$  oder  $Y_j = I(Z_j \geq z_j)$  keine Auswirkung zeigt. Die drei letzten Eigenschaften erlauben eine höhere Flexibilität bei der Modellbildung und eine geeignete Interpretationsmöglichkeit. Im bivariaten Fall sind folgende Copulas geeignet:

→ *Plackett-Copula*

$$C(u, v; \psi) = \frac{\{1 + (\psi - 1)(u + v) - [(1 + (\psi - 1)(u + v))^2 - 4(\psi - 1)uv]^{1/2}\}}{2(\psi - 1)}$$

mit  $0 < \psi < \infty$ ,  $M$  bei  $\psi \rightarrow 0$ ,  $C_I$  bei  $\psi \rightarrow 0$  und  $W$  bei  $\psi \rightarrow \infty$ ,

→ *Frank-Copula*

$$C(u, v, \delta) = -\delta^{-1} \log \left( \frac{(1 - \exp(-\delta)) - (1 - \exp(-\delta u))(1 - \exp(-\delta v))}{(1 - \exp(-\delta))} \right)$$

mit  $-\infty \leq \delta < \infty$ ,  $M$  bei  $\psi = -\infty$ ,  $C_I$  bei  $\psi = 0$  und  $W$  bei  $\psi \rightarrow \infty$ ,

→ *Bivariate Normal Copula*

$$C(u, v; \theta) = \Phi_2(\Phi^{-1}(u), \Phi^{-1}(v); \theta)$$

mit  $-1 \leq \theta \leq 1$ ,  $M$  bei  $\psi = -1$ ,  $C_I$  bei  $\psi = 0$  und  $W$  bei  $\psi = 1$ .

In  $u$  und  $v$  wird jeweils die Randverteilung  $G_j(y_j)$  der diskreten Variable eingesetzt. Die dargestellte bivariate Normal Copula entspricht dann der definierten Meta-Copula in (2.4). Die Plackett-Copula bzw. die Frank-Copula sind nur für den bivariaten Fall definiert. Eine Erweiterung auf den d-variablen Fall ist mit Hilfe des *Konstruktionsprinzips von Molenberghs und Lesaffre* (1994)[10] möglich. Das Prinzip beruht auf der Bildung von "product ratios". Als ein Beispiel kann die Bildung einer trivariaten Copula angeführt werden. Das dreidimensionale "product ratio"  $\psi_{123}$  ergibt sich folgendermaßen:

$$\psi_{123} = \frac{C_{123}(C_{123} - b_1)(C_{123} - b_2)(C_{123} - b_3)}{(C_{12} - C_{123})(C_{23} - C_{123})(C_{13} - C_{123})(a - C_{123})} \quad (2.7)$$

mit  $a = 1 - F_1 - F_2 - F_3 + C_{12} + C_{23} + C_{13}$ ,  $b_1 = C_{12} + C_{13} - F_1$ ,  $b_2 = C_{12} + C_{23} - F_2$  und  $b_3 = C_{13} + C_{23} - F_3$  unter der Bedingung  $B \leq A$  mit  $B = \max\{0, b_1, b_2, b_3\}$  und  $A = \min\{C_{12}, C_{23}, C_{13}, a\}$ .  $A$  und  $B$  stellen gleichzeitig die Fréchet-Hoeffding-Grenzen dar.

Der Parameter  $\psi_{123}$  ist der Abhängigkeitsparameter von  $C_{123}$  und misst den Zusammenhang zwischen drei Variablen. Die bivariaten Assoziationen werden mittels der eingesetzten zweidimensionalen Copula modelliert. Die Darstellung eines vier-dimensionalen "product ratios" ist beispielsweise in Xu (1996, S.95)[16] zu finden. Um die trivariate Copula zu erhalten, wird der Ausdruck nach  $C_{123}$  aufgelöst. Das sich ergebende Polynom 4. Ordnung:

$$\begin{aligned}
0 = & (\psi_{123} - 1) C_{123}^4 + ((-a - C_{12} - C_{13} - C_{23}) \psi_{123} + b_1 + b_2 + b_3) C_{123}^3 + \\
& (((a + C_{12} + C_{13}) C_{23} + (a + C_{12}) C_{13} + a C_{12}) \psi_{123} + (-b_1 - b_2) b_3 - b_1 b_2) C_{123}^2 + \\
& ((((-a - C_{12}) C_{13} - a C_{12}) C_{23} - a C_{12} C_{13}) \psi_{123} + b_1 b_2 b_3) C_{123} + \psi_{123} a C_{12} C_{13} C_{23}
\end{aligned}$$

kann mit der Ferrari-Methode gelöst werden. Nur eine der möglichen vier Ergebnisse des Gleichungssystems erfüllt die Anforderung, im Intervall von  $[B, A]$  zu liegen. Für die bivariate Copula  $C_{jk}$  kann jeweils die Frank- oder Plackett-Copula unterstellt werden. Der verwendete R-Code zur Lösung des Polynoms 4. Grades unter Verwendung der Plackett-Copula ist im Anhang zu finden. Das graphische Ergebnis der möglichen Lösungen für die trivariate Copula nach Molenberghs und Lesaffre ist in Abbildung (1) und (2) zu finden. In Darstellung (1) ist für die bivariate Copula die Plackett-Copula eingesetzt worden und bei der Graphik (2) die Frank-Copula. In beiden Abbildung ist die vierte Wurzel nicht aufgeführt, da diese jeweils zu keinem Ergebnis führte. Die Lösung der dritten Wurzel liegt bei beiden Graphiken innerhalb der unteren und oberen Fréchet-Hoeffding-Grenze und ist damit die gesuchte trivariate Copula.

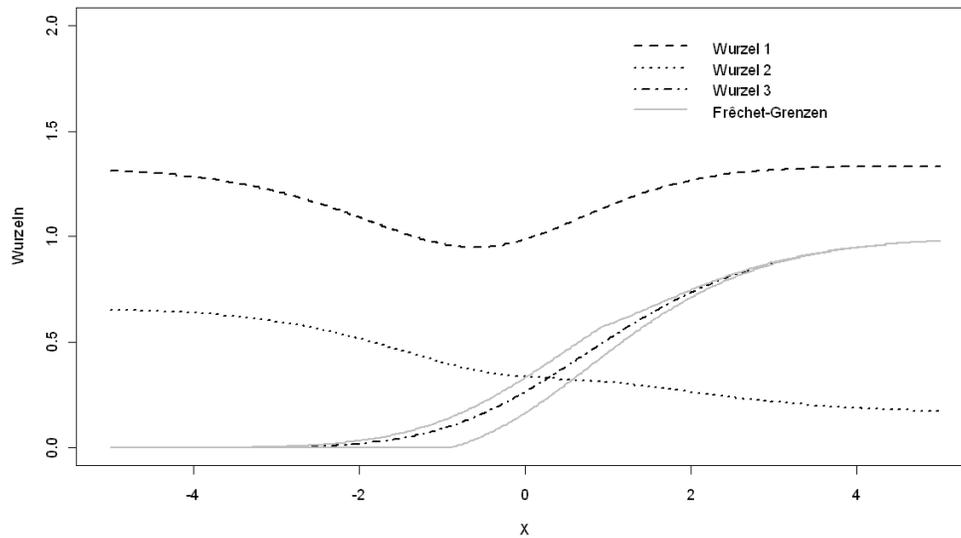


Abbildung 1: Lösungen des Polynom 4. Grades mit der Plackett-Copula

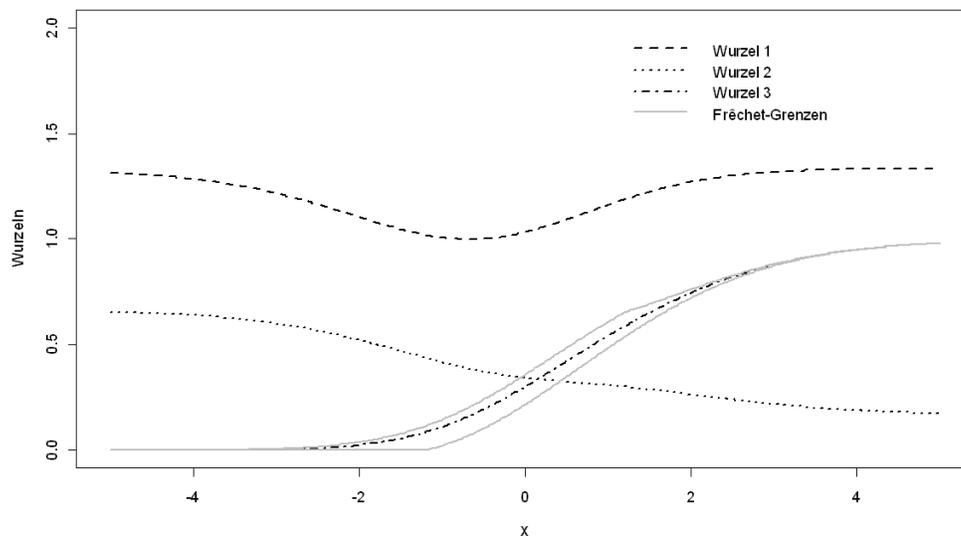


Abbildung 2: Lösungen des Polynom 4. Grades mit der Frank-Copula

Eine weitere Möglichkeit bietet die "Mixture of max-id" Copula von Joe und Hu (1996)[6]. Die Basis der Überlegungen für diese Art von Copula stellen die Abhängigkeitskonzept max-stabil und "max-infinitely divisible" (max-id) dar.

**Definition 2.3.** Sei  $H_d$  eine  $d$ -variate Verteilungsfunktion, dann ist diese max-stabil, falls zu jedem ( $t > 0$ ) gilt:

$$H_d^t(\mathbf{x}) = H_d(a_{1t} + b_{1t}x_1, \dots, a_{dt} + b_{dt}x_d),$$

für alle  $t \in \mathbb{N}$  existiert ein Vektor  $a_t > 0$  bzw. ein Vektor  $b_t$ .

**Definition 2.4.** Sei  $H_d$  eine  $d$ -variate Verteilungsfunktion, dann ist diese "max-infinitely divisible", falls zu jedem ( $\gamma > 0$ )  $H_d^\gamma$  eine Verteilungsfunktion ist.

Wenn  $H_d$  max-id ist, dann sind für alle  $n \in \mathbb{N}$   $H_d^{1/n}$  Verteilungsfunktionen. Für die notwendige und hinreichende Bedingung gilt:

- $d = 2$   $H_d$  ist max-id, wenn  $H_d$  Total Positiv der Ordnung 2 ( $TP_2$ ) ist,
- $d \geq 2$   $H_d$  ist max-id, wenn  $R_S \geq 0$  für alle  $S$  ist,  
mit  $R = \log(H_d(\mathbf{y}))$ ,  $S = \{1, \dots, d\}$  und  $R_S$  beinhaltet alle partiellen Ableitungen von  $R$  nach  $y_i$ .

Eine allgemeine Konstruktion einer max-id Verteilung ergibt sich über folgenden Ausdruck:

$$F = \int H^\gamma dM(\gamma) = \psi(-\log(H)),$$

wobei  $H$  eine d-dimensionale max-stabile Verteilung,  $M$  die Verteilung einer positiven Zufallsvariablen und  $\psi$  deren Laplace-Transformation darstellt. Für die Bildung der d-variater max-id Copula wird auf eine Mischung des allgemeinen Konstruktionsprinzips zurückgegriffen:

$$\int_0^\infty \prod_{1 \leq j < k \leq d} K_{jk}^\alpha(H_j, H_k) \prod_j^d H_j^{\nu_j \alpha} dM(\alpha) = \psi \left( - \sum_{1 \leq j < k \leq d} \log(K_{jk}(H_j, H_k)) - \sum_j^d \nu_j \log(H_j) \right),$$

mit einer bivariaten max-id Copula  $K_{jk}$  und den Randverteilungen  $H_j$ . Die "Mixture of max-id" Copula ist dann gegeben, wenn  $H_j(u_j) = \exp(-p_j \psi^{-1}(u_j))$  mit  $p_j = (\nu_j + d - 1)^{-1}$  gesetzt wird. Letzteres stellt sicher, dass es sich bei  $H_j(u_j)$  um gleichverteilte Ränder im Bereich  $[0, 1]$  handelt. Setzt man diesen in den oben stehenden Ausdruck ein, erhält man:

$$C(u_1, \dots, u_d) = \psi \left( - \sum_{1 \leq j < k \leq d} \log(K_{jk}(\exp(-p_j \psi^{-1}(u_j)), \exp(-p_k \psi^{-1}(u_k)))) + \sum_j^d \nu_j p_j \psi^{-1}(u_j) \right).$$

Eine geeignete bivariate max-id Copula  $K_{ij}$  ist dann gegeben, wenn sie die  $TP_2$ -Bedingung erfüllt. Von den drei genannten bivariaten Copulas besitzt die Frank- und die Normal Copula diese Eigenschaft. Die Normal Copula ist für den multivariaten Fall definiert, so dass eine Erweiterung nur bezüglich der Frank-Copula sinnvoll ist. Aus der Klasse der Laplace-Transformationen  $\psi(s; \theta)$  können die angewendet werden, bei denen gilt, dass  $-\log(\psi) \in \mathcal{L}_d^*$  ist, mit  $\mathcal{L}_d^* = \{\omega : [0, \infty) \rightarrow [0, \infty) | \omega(0) = 0, \omega(\infty) = \infty, (-1)^{j-1} \omega^{(j)} \geq 0, j = 1, \dots, d\}$ . Die "Mixture of max-id" Copula bietet eine weitere Möglichkeit der flexiblen Modellierung von Abhängigkeiten. Als ein Minimum an Abhängigkeit kann die Laplace-Transformierte  $\psi$  gesehen werden. Die bivariaten Zusammenhänge werden über die eingesetzten Copulas  $K_{jk}$  modelliert und der Parameter  $\nu_j$  misst Asymmetrien in den Verteilungen.

Die beiden vorgestellten Konstruktionsprinzipien können angewendet werden, um die Frank- und Plackett-Copula auf den multivariaten Fall zu erweitern. Die Normal Copula ist für den d-dimensionalen Fall bereits definiert und wird somit direkt angewendet.

Zusammenfassend kann folgende Auswahl an Copulas für den d-variater Fall getroffen werden:

- *Multivariate Frank- bzw. Plackett-Copula nach dem Konstruktionsprinzip von Molenberghs und Lesaffre,*
- *"Mixture of max-id" Copula mit  $K_{jk}$  gleich der Frank-Copula,*
- *Multivariate Normal Copula*

$$C(u_1, \dots, u_d; \Theta) = \Phi_d(\Phi^{-1}(u_1), \dots, \Phi^{-1}(u_d); \Theta)$$

mit  $-1 \leq \theta_{jk} \leq 1$ .

Die Auswahl geeigneter Copulas stellt lediglich eine Empfehlung dar, erhebt jedoch keinen Anspruch auf Vollständigkeit.

Nach der Spezifikation möglicher Copulas wird nun die Konkretisierung der Randverteilungen aufgegriffen. Mit Gleichung (2.6) ist die allgemeine Form der Wahrscheinlichkeitsmassenfunktion  $P(y_1, \dots, y_d; \Theta)$  für diskrete Copula Modell gegeben. Die darin enthaltenen diskreten Randverteilungen  $G_j(y_j)$  sind abhängig vom vorliegenden Skalenniveau. Im Weiteren wird die Spezifizierung von  $G_j(y_j)$  für binäre und ordinale Variablen getrennt betrachtet. Beschränkt man sich in einem *binären* bivariaten Modell auf zwei Responses, dann nehmen die beiden diskreten Zufallsvariablen  $Y_1$  und  $Y_2$  lediglich die Ausprägungen 0 oder 1 an. Im betrachteten Fall können die Auftretenskombinationen (0, 0), (0, 1), (1, 0) oder (1, 1) beobachtet werden, mit den dazugehörigen Wahrscheinlichkeitsmassenfunktionen  $P(0, 0)$ ,  $P(0, 1)$ ,  $P(1, 0)$  und  $P(1, 1)$ . Diese Wahrscheinlichkeitsfunktionen wiederum können als Funktion einer Copula ausgedrückt werden mit:

$$P(y_1, y_2) = C(G_1(y_1), G_2(y_2)) - C(G_1(y_1), G_2(y_2 - 1)) \\ - C(G_1(y_1 - 1), G_2(y_2)) + C(G_1(y_1 - 1), G_2(y_2 - 1)),$$

wobei für  $G_j(y_j)$  eine univariate Verteilungsfunktion eingesetzt wird. Zum Beispiel ergibt sich für  $P(1, 1)$ :

$$P(1, 1) = C(G_1(1), G_2(1)) - C(G_1(1), G_2(0)) - C(G_1(0), G_2(1)) + C(G_1(0), G_2(0)).$$

Für den drei-dimensionalen Fall lässt sich die Wahrscheinlichkeitsmassenfunktionen darstellen mit:

$$P(y_1, y_2, y_3) = C(G_1(y_1), G_2(y_2), G_3(y_3)) - C(G_1(y_1), G_2(y_2 - 1), G_3(y_3)) \\ - C(G_1(y_1), G_2(y_2), G_3(y_3 - 1)) + C(G_1(y_1), G_2(y_2 - 1), G_3(y_3 - 1)) \\ - C(G_1(y_1 - 1), G_2(y_2), G_3(y_3)) - C(G_1(y_1 - 1), G_2(y_2 - 1), G_3(y_3 - 1)) \\ + C(G_1(y_1 - 1), G_2(y_2), G_3(y_3 - 1)) - C(G_1(y_1 - 1), G_2(y_2 - 1), G_3(y_3)).$$

Im binären Fall wird  $G_j(y_j)$  wie folgt definiert:

$$G_j(y_j) = \begin{cases} 0, & y_j < 0 \\ 1 - p_j, & 0 \leq y_j < 1 \\ 1, & y_j \geq 1. \end{cases} \quad (2.8)$$

Der Parameter  $p_j$  gibt die Wahrscheinlichkeit an, dass die Zufallsvariable  $Y_j$  den Wert Eins annimmt, also  $p_j = P(Y_j = 1)$ . Die Spezifizierung dieses Parameters erfolgt nun anhand des latenten diskreten Modells. Dabei wird angenommen, dass die diskrete Variable auf eine nicht beobachtbare stetige Variable zurückzuführen ist. Wenn diese latente Variable einen

bestimmten Schwellenwert überschreitet, so nimmt die diskrete beobachtete Response den Wert Eins an und sonst Null. Formal lässt sich diese Beziehung folgendermaßen beschreiben:

$$Y_j = \begin{cases} 0, & Y_j^* \leq c_j \\ 1, & Y_j^* > c_j. \end{cases} \quad (2.9)$$

Im einfachsten Falle ohne Kovariablen kann die latente Variable  $Y_j^*$  mit dem Fehlerterm  $\epsilon$  gleichgesetzt werden. Für den Fehlerterm wird nun eine stetige Verteilung  $F_j$  angenommen und daraus ergibt sich, dass die Wahrscheinlichkeitsfunktion an der Stelle Eins der Verteilungsfunktion des Fehlerterms entspricht  $p_j = F_j(c_j)$ . Für die gemeinsame Verteilung des Fehlertermvektors  $\epsilon = (\epsilon_1, \dots, \epsilon_d)'$  wird eine d-dimensionale Verteilung unterstellt oder eine Copulafunktion mit entsprechenden Randverteilungen  $F_j$ .

In der Literatur findet sich eine etwas allgemeinere Darstellung dieser Beziehung. Sie ist äquivalent mit  $Y_j = I(Z_j < z_j)$ .  $Z_j$  ist wiederum eine latente Variable, welche eine stetige Verteilung  $F_j$  besitzt und dem Fehlerterm  $\epsilon_j$  der obigen Darstellung entspricht. Für die gemeinsame Verteilung von  $Z = (Z_1, \dots, Z_d)'$  wird im Modell eine Copula-Funktion unterstellt. Es gilt dann für den binären Fall, dass:

$$p_j = F_j(z_j). \quad (2.10)$$

Die Berücksichtigung von Kovariablen erfolgt mittels des Parameters  $z_j$ . Dieser wird als eine Funktion des Kovariablenvektors  $\mathbf{X}_j = \{X_1, \dots, X_p\}$  der  $j$ -ten Response, wie folgt  $z_j = g(\mathbf{X})$ , ausgedrückt. Falls sich  $z_j$  als eine Linearkombination der Kovariablen darstellt, dann gilt:

$$z_j = \alpha_j + X_j \beta_j,$$

mit  $\beta_j$ , dem Vektor der Einflussparameter der Kovariablen und  $\alpha_j$ , dem globalen Achsenabschnitt der  $j$ -ten Response. Damit ergibt sich, dass  $p_j = F_j(\alpha_j + X_j \beta_j)$  ist. Aus Identifikationsgründen wird, wie im univariaten Fall, der Schwellenwertparameter  $c_j$  aus (2.9) gleich Null gesetzt.

Im *ordinalen* Kontext sind jetzt nicht nur zwei, sondern mehrere diskrete Ausprägungen möglich und diese beinhalten zudem eine Ordnungsrelation. Für eine ordinale Response  $y_j$  sind  $m_j$  Kategorien von  $1, \dots, m_j$  zu beobachten. Die Auftretswahrscheinlichkeit der Ausprägung  $s_j$  mit  $s_j = \{1, \dots, m_j\}$  der  $j$ -ten Response wird mit  $p_j^{(s_j)}$  bezeichnet. Die zusätzliche Information wird in der Spezifizierung der Randverteilung  $G_j(y_j)$  berücksichtigt mit:

$$G_j(y_j) = \begin{cases} 0, & y_j < 1 \\ \sum_{s_j=1}^{|y_j|} p_j^{(s_j)}, & 1 \leq y_j < m_j \\ 1, & y_j \geq m_j. \end{cases} \quad (2.11)$$

$|y_j|$  ist die größte ganze Zahl, die kleiner oder gleich der Ausprägung von  $y_j$  ist. Für die Reparameterisierung von  $p_j^{(s_j)}$  wird auf das latente Modell zurückgegriffen. Dieses unterstellt

eine latente Variable  $Y_j^*$ , die repräsentativ für die diskrete Variable  $Y_j$  steht. Die Beziehung (2.9) wird nun folgendermaßen erweitert:

$$Y_j = s_j \iff \alpha_j(s_j - 1) < Y_j^* \leq \alpha_j(s_j). \quad (2.12)$$

Die Beziehung besagt, dass die Variable  $Y_j$  die Kategorie  $s_j$  annimmt, wenn die zugehörige latente Variable  $Y_j^*$  sich im Schwellenwertintervall  $[\alpha_j(s_j - 1), \alpha_j(s_j))$  befindet. Die beiden Schwellenwerte  $\alpha_j(0)$  und  $\alpha_j(m_j)$  werden  $-\infty$  bzw.  $\infty$  gesetzt und es gilt folgende Ordnungsbedingung  $-\infty = \alpha_j(0) < \alpha_j(1) < \dots < \alpha_j(m_j - 1) < \alpha_j(m_j) = \infty$ . Im einfachsten Fall wird die latente Variable  $Y_j^*$  dem Fehlerterm  $\epsilon$  gleichgesetzt. Damit gilt:

$$Y_j = s_j \iff \alpha_j(s_j - 1) < \epsilon \leq \alpha_j(s_j).$$

Wie im binären Modell wird für den Fehlerterm  $\epsilon = (\epsilon_1, \dots, \epsilon_d)'$  eine d-variate Copula mit den univariaten Verteilung  $F_j$  unterstellt. Der gesuchte Parameter  $p_j^{(s_j)}$  kann nun über die Annahme der Fehlertermverteilung durch  $P(\alpha_j(s_j - 1) < \epsilon \leq \alpha_j(s_j))$  ausgedrückt werden und es gilt:

$$p_j^{(s_j)} = F_j(\alpha_j(s_j)) - F_j(\alpha_j(s_j - 1)).$$

In vereinfachter Darstellung kann die Beziehung als  $Y_j = I(z_j(s_j - 1) < Z_j \leq z_j(s_j))$  bezeichnet werden. Die Zufallsvariable  $Z_j$  entspricht wiederum dem Fehlerterm und es werden die selben Annahmen getroffen, wie für  $\epsilon_j$ . Analog ergibt sich für:

$$p_j^{(s_j)} = F(z_j(s_j)) - F(z_j(s_j - 1)). \quad (2.13)$$

Die Berücksichtigung von Kovariablen wird erzielt durch eine Spezifizierung von  $z_j(s_j)$  als Funktion von  $\mathbf{X}_j$ . Bei einer Linearkombination gilt folgender Zusammenhang:

$$z_j(s_j) = \alpha_j(s_j) - X_j \beta_j,$$

wobei der Vektor  $\beta_j$  die Einflussparameter der Kovariablen enthält. Dieser beinhaltet jedoch keinen globalen Achsenabschnitt. Stattdessen wird pro Kategorie ein Achsenabschnitt modelliert, der dem jeweiligen Schwellenwert entspricht. Bei der Schätzung muss die Ordnungsbedingung der Schwellenwerte beachtet werden.

Die Randverteilung  $G_j(y_j)$  ergibt sich im diskreten Modell aus der Summe der univariaten Wahrscheinlichkeitsmassenfunktion  $p_j$ . Auf Basis des latenten Modells erfolgt die Parametrisierung von  $p_j$  im binären Fall mittels der Gleichung (2.10) und von  $p_j^{(s_j)}$  im ordinalen Fall über die Beziehung in (2.13). Die in beiden Gleichungen auftretende Verteilung  $F_j$  muss nun noch konkretisiert werden, indem über die univariaten Verteilung der latenten Variable  $Z_j$  eine Annahme getroffen wird. Dazu wird auf bereits bekannte univariate latente Modelle zurückgegriffen, wie das Logit Modell und das Probit Modell. Darauf basierend können die multivariaten diskreten Copula Modelle in zwei Modellklassen unterschieden werden, dem multivariaten Logit Modell und dem multivariaten Probit Modell.

Im *Multivariaten Logit Modell (MLM)* wird für  $F_j$  eine logistische Verteilung eingesetzt, so dass sich eine natürliche Erweiterung des bekannten univariaten Logit Modells ergibt. Die logistische Verteilung ist definiert als:

$$F(y) = \frac{\exp\left(\frac{y-\mu}{\beta}\right)}{1 + \exp\left(\frac{y-\mu}{\beta}\right)} = \frac{1}{1 + \exp\left(-\frac{y-\mu}{\beta}\right)}$$

mit dem Lageparameter  $\mu$  und dem Skalenparameter  $\beta$ . Die Varianz einer logistisch verteilten Zufallsvariable ergibt sich aus  $Var(Y) = \beta^2\pi^2/3$ . Aus Identifikationsgründen wird an Stelle der logistischen Verteilung deren Standardisierung mit Mittelwert  $\mu = 0$  und Varianz  $\sigma^2 = \pi^2/3$  angenommen.

Das *Multivariate Probit Modell (MPM)* entspricht vom Grundgedanken und der Herleitung her dem *MLM*. Der Unterschied besteht in der Verwendung der Standardnormalverteilung anstelle der standard-logistischen Verteilung. Die Standardnormalverteilung ist definiert als:

$$F(y) = \int_{-\infty}^y \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}t\right) dt.$$

Dieses kann als eine Erweiterung des univariaten Probit Modells auf den d-dimensionalen Fall verstanden werden.

Der Umweg über die latenten Modelle erlaubt unterschiedliche diskrete Skalen korrekt zu modellieren, Kovariablen in das Modell aufzunehmen, den Bezug zu den univariaten diskreten Modellen herzustellen und deren Interpretation zu verwenden. Die so gefundenen Ränder  $G_j(y_j)$  können in den Copulaausdruck (2.6) eingesetzt werden, um die gemeinsame Wahrscheinlichkeit der Responses zu bestimmen. Für die Wahl der geeigneten Copula sei auf die oben stehenden Ausführungen verwiesen. Das *MLM* beinhaltet als Spezialfall bereits genannte Modelle der scheinbar unverbundenen Systeme. Das Plackett-Modell ergibt sich bei Annahme einer Plackett-Copula und das Modell von Koehler und Symanowski bei Annahme einer verallgemeinerten Cook-Johnson Copula. Diese geht auf die beiden Autoren Cook und Johnson (1986) zurück und ist gegeben mit:

$$C(u, v) = (1 + \omega)(u^{-1/\alpha} + v^{-1/\alpha} - 1)^{-\alpha} + \omega(2u^{-1/\alpha} + 2v^{-1/\alpha} - 3)^{-\alpha} - \omega(2u^{-1/\alpha} + v^{-1/\alpha} - 2)^{-\alpha} - \omega(u^{-1/\alpha} + 2v^{-1/\alpha} - 2)^{-\alpha}$$

mit  $\alpha > 0$  und  $-1 \leq \omega \leq 1$ . Diese Copula beinhaltet zwei Parameter, die die Abhängigkeit zwischen zwei Responses steuert, wobei inhaltlich nicht klar ist, wie diese interpretiert werden können. Die Copula wird aus diesen Gründen auch nicht als eine geeignete Copula angesehen. Das *MPM* in Verbindung mit einer Multivariaten Normal Copula ergibt das Modell von Ashford und Sowden. Das diskrete Copula Modell kann damit als Oberklasse der scheinbar unverbundenen Systemen verstanden werden.

## 2.4 Schätzmethoden

Für die Schätzung der Kovariablen- und der Abhängigkeitsparameter wird die Maximum-Likelihood-Methode angewendet. Wie in der Einleitung bereits erwähnt, stehen dafür zwei Ansätze zur Verfügung. Im Folgenden wird die Methode der "Full Maximum-Likelihood" (FML) erläutert und anschließend die der "Inference Functions of Margins" (IFM).

Die *FML-Methode* beruht auf der Maximierung der vollständigen Likelihood-Funktion. Für die Aufstellung der Likelihood-Funktion wird die Verteilung der multivariaten diskreten Response benötigt. Zunächst wird sich auf nur eine Reponse beschränkt und anhand derer die benötigte Likelihood-Funktion hergeleitet und im Anschluss daran erfolgt die Übertragung auf mehrere Responses. Bei Betrachtung nur *einer* Response  $Y_j$  mit  $m_j$  Ausprägungen wird, basierend auf dem Logit bzw. Probit Modell, für das Auftreten einer Ausprägung eine multinomiale Verteilung unterstellt, die folgende Wahrscheinlichkeitsfunktion besitzt:

$$f_{Y_j}(y(1), \dots, y(m_j)) = \frac{n!}{n(1)! \dots n(m_j)!} p(1)^{n(1)} \dots p(m_j)^{n(m_j)},$$

mit  $p(s_j)$  der univariaten Auftrittswahrscheinlichkeit aus (2.10) oder (2.13) und mit  $n(s_j)$  der Anzahl der beobachteten Ausprägungen in der Kategorie  $s_j$ . Folgende Restriktionen müssen dabei beachtet werden:

$$\begin{aligned} \sum_{s=1}^{m_j} p(s_j) &= 1, \\ \sum_{s=1}^{m_j} n(s_j) &= n. \end{aligned} \tag{2.14}$$

Die univariate Likelihood-Funktion  $\mathcal{L}$  bei ungruppierten Daten mit  $(i = 1, \dots, n)$  stellt sich wie folgt dar:

$$\begin{aligned} \mathcal{L}(Y_j; z_j) &= \prod_i^n f_{Y_j}(y_i(1), \dots, y_i(m_j); \theta) \\ &= \prod_{i=1}^n \frac{n!}{n_i(1)! \dots n_i(m_j)!} p_i(1)^{n_i(1)} \dots p_i(m_j)^{n_i(m_j)}. \end{aligned}$$

Dieser Ausdruck läßt sich reduzieren auf:

$$\mathcal{L}(Y_j; z_j) = \prod_{i=1}^n \prod_{s_j=1}^{m_j} p_i(s_j)^{n_i(s_j)}. \tag{2.15}$$

Bei der Betrachtung von *zwei* Responses ergeben sich für deren Ausprägungen verschiedene Möglichkeiten der Kombination. Für das Auftreten der Kombinationen wird, wie im univariaten Fall eine multinomiale Verteilung unterstellt.

Die gemeinsame Wahrscheinlichkeitsfunktion erweitert sich dann zu:

$$f(y_{jk}(1, 1), \dots, y_{jk}(m_j, m_k)) = \frac{n!}{n_{jk}(1,1)! \dots n_{jk}(m_j, m_k)!} p_{jk}(1, 1)^{n_{jk}(1,1)} \dots p_{jk}(m_j, m_k)^{n_{jk}(m_j, m_k)}.$$

Im vorangegangenen Kapitel wurde  $p_{jk} = P(y_j, y_k; \Theta)$  als eine Funktion der Copula in Gleichung (2.6) definiert und gibt die Wahrscheinlichkeit eines gemeinsamen Auftretens für verschiedene Ausprägungskombinationen der beiden Responses  $y_j$  und  $y_k$  an. Die Anzahl der Ausprägungen pro Kombination wird mit  $n_{jk}$  bezeichnet. Die Restriktionen aus (2.14) gelten auch für den zwei-dimensionalen Fall und lauten folgendermassen:

$$\sum_{s_j=1}^{m_j} \sum_{s_k=1}^{m_k} p_{jk}(s_j, s_k) = 1,$$

$$\sum_{s_j=1}^{m_j} \sum_{s_k=1}^{m_k} n_{jk}(s_j, s_k) = n.$$

Die reduzierte Form der bivariaten Likelihood-Funktion kann in Analogie zu (2.15) aufgestellt werden, mit:

$$\mathcal{L}(Y_{jk}; \theta) = \prod_{i=1}^n \prod_{s_j=1}^{m_j} \prod_{s_k=1}^{m_k} p_{ijk}(s_j, s_k)^{n_{ijk}(s_j, s_k)}.$$

Allgemein wird im d-variaten Fall für die möglichen Ausprägungskombinationen der  $d$  Responses die Annahme einer multinomialen Verteilung unterstellt. Die multivariate Likelihood-Funktion in vereinfachter Darstellung läßt sich ableiten als:

$$\mathcal{L}(\mathbf{Y}; \theta) = \prod_{i=1}^n \prod_{s_1=1}^{m_1} \dots \prod_{s_d=1}^{m_d} p_{is_1 \dots s_d}^{n_{is_1 \dots s_d}}. \quad (2.16)$$

Der Vektor  $\theta' = (z'_1, \dots, z'_j, \delta')$  beinhaltet sowohl die zu schätzenden univariaten Parameter  $z_j$  als auch die zu schätzenden Abhängigkeitsparameter  $\delta$ . Bei der FML-Methode wird die logarithmierte multivariate Likelihood-Funktion  $\ell(\mathbf{Y}; \theta)$  maximiert und damit die gesuchten Parameter der Randverteilungen  $G_j$  und der Copula gemeinsam geschätzt. Unter regulären Bedingungen erhält man diese durch Lösung folgender Gleichung:

$$(\partial \ell_{\mathbf{Y}} / \partial z'_1, \dots, \partial \ell_{\mathbf{Y}} / \partial z'_j, \partial \ell_{\mathbf{Y}} / \partial \delta') = 0'.$$

Der FML-Schätzer wird allgemein bezeichnet mit  $\hat{\theta}$ . Für die Berechnung der Varianzen des FML-Schätzer wird die Informationsmatrix  $\mathcal{I}$  benötigt und diese ergibt sich aus dem

Erwartungswert der zweiten Ableitungen. Nach Joe (2005)[5] kann  $\mathcal{I}$  zerlegt werden in:

$$\mathcal{I} = \begin{pmatrix} \mathcal{I}_{11} & \dots & \mathcal{I}_{1d} & \mathcal{I}_{1\delta} \\ \dots & \dots & \dots & \dots \\ \mathcal{I}_{d1} & \dots & \mathcal{I}_{dd} & \mathcal{I}_{d\delta} \\ \mathcal{I}_{\delta 1} & \dots & \mathcal{I}_{\delta d} & \mathcal{I}_{\delta\delta} \end{pmatrix}, \quad (2.17)$$

wobei  $\mathcal{I}_{jk} = -E(\partial^2 \ell_{\mathbf{Y}} / \partial z_j \partial z'_k)$  mit  $1 \leq j \leq k \leq d$  und  $\mathcal{I}_{j\delta} = -E(\partial^2 \ell_{\mathbf{Y}} / \partial z_j \partial \delta')$  ist. Es gilt  $\mathcal{I}_{j\delta} = \mathcal{I}_{\delta j}$ .

Da die multivariate Likelihood-Funktion rechentechnisch schwer zu handhaben ist, wurde von Xu (1996)[16] und von Joe (1997)[4] die *IFM-Methode* als ein alternatives Schätzverfahren vorgeschlagen. Xu selber differenziert dieses nochmals in ein zweistufiges und in ein mehrstufiges Verfahren.

Bei dem zweistufigen Verfahren werden in einem ersten Schritt alle univariaten Parameter  $z_j$  geschätzt anhand der univariaten Likelihood-Funktionen  $\mathcal{L}(Y_j; \theta)$  aus (2.15). In einem zweiten Schritt werden die multivariaten Parameter  $\delta$  ermittelt, gegeben der univariaten Größen auf Basis der multivariaten Likelihood-Funktion  $\mathcal{L}(\mathbf{Y}; \theta)$  aus (2.16).

Das mehrstufigen Verfahren bestimmt im ersten Schritt ebenfalls nur die univariaten Schätzer. Im zweiten Schritt werden die bivariaten Einflussgrößen anhand der bivariaten Likelihood-Funktionen geschätzt. Im Anschluß folgt die Ermittlung der k-variaten Schätzer mit ( $2 < k < d$ ) auf Basis ihrer k-variaten Likelihood-Funktion, bis in einem letzten Schritt die d-variaten Schätzer bestimmt werden. Die geschätzten Parameter werden jeweils in der nachfolgenden Stufe als gegeben vorausgesetzt. Die folgenden Ausführungen beziehen sich auf das zweistufige Verfahren und sind übertragbar auf die mehrstufige Methode.

Wie bereits erwähnt, erfolgt bei dem zweistufigen Verfahren die Schätzung der univariaten Parameter  $z_j$  über die univariaten logarithmierten Likelihood-Funktionen

$$\ell(Y_j; z_j) = \sum_{i=1}^n \sum_{s_j=1}^{m_j} n_i(s_j) \log(p_i(s_j))$$

und die multivariaten Parameter über die d-variate logarithmierte Likelihood-Funktion

$$\ell(\mathbf{Y}; \delta, \tilde{z}_1, \dots, \tilde{z}_d) = \sum_{i=1}^n \sum_{s_1=1}^{m_1} \dots \sum_{s_d=1}^{m_d} n_{i m_1 \dots m_d} \log(p_{i s_1 \dots s_d}),$$

gegeben der univariaten Schätzer  $\tilde{z}_j$ . Das unter regulären Bedingungen zu lösende Maximierungsproblem, stellt sich wie folgt dar:

$$(\partial \ell_{Y_1} / \partial z'_1, \dots, \partial \ell_{Y_j} / \partial z'_j, \partial \ell_{\mathbf{Y}} / \partial \delta') = 0'. \quad (2.18)$$

Der geschätzte Parametervektor der IFM-Methode wird mit  $\tilde{\theta}'$  gekennzeichnet und besitzt folgende Eigenschaften:

**Theorem 3.** Falls  $\tilde{\theta}$  der IFM-Schätzer unter der Bedingung von (2.18) ist, dann ist dieser ein konsistenter Schätzer von  $\theta$ . Weiterhin gilt für den Schätzer mit  $n \rightarrow \infty$  asymptotisch, dass

$$\sqrt{n}(\tilde{\theta} - \theta) \sim N_q(0, J_{\Psi}^{-1}),$$

mit  $q$  der Dimension von  $\theta$  und mit  $J_{\Psi}(\theta) = D'_{\Psi}(\theta)M_{\Psi}^{-1}(\theta)D_{\Psi}(\theta)$ .

*Proof.* Mit  $\Psi$  wird der Vektor der Schätzgleichungen bezeichnet, der sich im vorliegenden Modell beschreiben lässt als:

$$\begin{aligned}\Psi_{Y_j} &= \partial \ell_{Y_j} / \partial z'_j, \\ \Psi_{\mathbf{Y}} &= \partial \ell_{\mathbf{Y}} / \partial \delta',\end{aligned}$$

mit  $j = 1, \dots, d$ . Die Bedingungen für die reguläre Schätzfunktionen gelten ebenfalls für den Vektor  $\Psi$ . Einige wichtige Bedingungen sind:

- $E(\Psi(\theta)) = 0$ ,
- Die partiellen Ableitungen  $\partial \Psi / \partial \theta'$  existieren für fast jedes  $y \in \mathcal{Y}$ ,
- $M_{\Psi}(\theta) = E(\Psi \Psi')$  existiert und ist positiv definit,
- $D_{\Psi}(\theta) = E(\partial \Psi / \partial \theta')$  existiert und ist nicht singular.

Für eine ausführliche Darstellung der regulären Schätzgleichungen sei auf Xu (1996)[16] verwiesen. Die Lösung des Schätzproblems ergibt sich aus:

$$\Psi(\tilde{\theta}) = 0'. \quad (2.19)$$

Nach dem Theorem von Taylor 1. Ordnung kann (2.19) anhand des wahren Parameter  $\theta$  formuliert werden als:

$$\Psi(\tilde{\theta}) = \Psi(\theta) + (\tilde{\theta} - \theta) \left. \frac{\partial \Psi}{\partial \theta'} \right|_{\theta^*} = 0, \quad (2.20)$$

mit einem Wert  $\theta^*$ , der zwischen  $\theta$  und  $\tilde{\theta}$  liegt. Gemäß der Bedingungen der regulären Schätzgleichungen ist  $E(\Psi(\theta)) = 0$  und sind die partiellen Ableitungen  $\partial \Psi / \partial \theta'$  ungleich Null. Da zudem der Ausdruck (2.20) zu Null konvergiert, muss  $(\tilde{\theta} - \theta)$  mit  $n \rightarrow \infty$  nach Null konvergieren. Damit ist der Schätzer  $\tilde{\theta}$  unter den getroffenen Annahmen konsistent. Die Matrix  $H_n(\theta)$  beinhaltet die partiellen Ableitungen des Vektors  $\Psi$  nach den zu schätzenden Parametern  $\theta$  und lässt sich folgendermassen darstellen:

$$H_n(\theta) = \begin{pmatrix} \frac{\partial \Psi_{Y_1}}{\partial z_1} & 0 & \dots & 0 \\ 0 & \frac{\partial \Psi_{Y_2}}{\partial z_2} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ \frac{\partial \Psi_{\mathbf{Y}}}{\partial z_1} & \frac{\partial \Psi_{\mathbf{Y}}}{\partial z_2} & \dots & \frac{\partial \Psi_{\mathbf{Y}}}{\partial \delta} \end{pmatrix}. \quad (2.21)$$

Es gilt weiterhin, dass  $D_\Psi = E(n^{-1}H_n)$  ist. Nach dem Gesetz der großen Zahl konvergiert  $n^{-1}H_n$  mit einer Wahrscheinlichkeit von Eins zu seinem Erwartungswert, so dass gilt:

$$n^{-1}H_n(\tilde{\theta}) \rightarrow D_\Psi(\theta).$$

Die Gleichung (2.20) kann umgeschrieben werden zu:

$$\sqrt{n}(\tilde{\theta} - \theta) = \left[ \frac{1}{n}H_n^* \right]^{-1} \frac{1}{\sqrt{n}}(-\Psi(\theta)),$$

mit  $H_n^*$  an der Stelle von  $\theta^*$ . Es gilt zudem  $n^{-1}H_n^* \rightarrow D_\Psi(\theta)$ . Für  $\Psi(\theta)$  gilt asymptotisch, dass:

$$\frac{1}{\sqrt{n}}\Psi(\theta) \sim N_q(0, M(\theta)).$$

Nach dem Slutsky's Theorem kann für  $\sqrt{n}(\tilde{\theta} - \theta)$  folgende asymptotische Beziehung hergestellt werden:

$$\sqrt{n}(\tilde{\theta} - \theta) \sim N(0, D_\Psi^{-1}(\theta)M_\Psi(\theta)(D_\Psi^{-1}(\theta))').$$

□

Die Matrix  $J_\Psi^{-1} = D_\Psi^{-1}(\theta)M_\Psi(\theta)(D_\Psi^{-1}(\theta))'$  ist die asymptotische Varianz-Kovarianz-Matrix des IFM-Schätzers  $\tilde{\theta}$  und wird bezeichnet als die Inverse der Godambe Informations-Matrix  $J_\Psi$ . Das Theorem (3) und dessen Beweis ist auf die Arbeit von Xu (1996)[16] zurückzuführen. Nach Joe (2005)[5] können die Bestandteile der Varianz-Kovarianz-Matrix mit Hilfe der Matrix in (2.17) dargestellt werden. Durch eine kleine Modifikation erhält man für die Inverse der Gobambe-Informations-Matrix folgenden Ausdruck:

$$J_\Psi^{-1} = (-D_\Psi^{-1}(\theta))M_\Psi(\theta)(-D_\Psi^{-1}(\theta))'.$$

Die Matrix  $D_\Psi$  ist definiert als  $E(\partial\Psi/\partial\theta)$  und  $\mathcal{J}_{jj}$  stellt die Informationsmatrix der  $j$ -ten univariaten Likelihood dar. So kann  $-D_\Psi$  auf Basis von Partitionen formuliert werden:

$$-D_\Psi = \begin{pmatrix} \mathcal{J}_{11} & 0 & \dots & 0 \\ 0 & \mathcal{J}_{22} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ \mathcal{I}_{\delta 1} & \dots & \mathcal{I}_{\delta d} & \mathcal{I}_{\delta\delta} \end{pmatrix}.$$

Die Matrix  $M_\Psi$  beinhaltet die Erwartungswerte von  $\Psi\Psi'$  und kann als Kovarianz zwischen  $\Psi_j$  und  $\Psi_k$  verstanden werden. Sei  $\mathcal{J}_{jk} = Cov(\Psi_{Y_j}, \Psi_{Y_k})$  für  $1 \leq j \leq k \leq m$  und  $\mathcal{J}_{jj}$  bezeichne wiederum die Informationsmatrix der  $j$ -ten univariaten Likelihood, so ist:

$$M_\Psi = \begin{pmatrix} \mathcal{J}_{11} & \dots & \mathcal{J}_{1d} & 0 \\ \dots & \dots & \dots & \dots \\ \mathcal{J}_{d1} & \dots & \mathcal{J}_{dd} & 0 \\ 0 & \dots & 0 & \mathcal{I}_{\delta\delta} \end{pmatrix}.$$

Die Kovarianz  $Cov(\Psi_{Y_j}, \Psi_{\mathbf{Y}})$  zwischen  $\Psi_{Y_j}$  und  $\Psi_{\mathbf{Y}}$  ist Null. Der Beweis dazu befindet sich in Joe (2005). Die Beziehung der Varianz-Kovarianz-Matrix des IFM-Schätzers zu den Informationsmatrizen der univariaten Likelihood-Funktionen und der Informationsmatrix des FML-Schätzer wird bei der Simulation benutzt, um die Godambe-Informations-Matrix zu erzeugen. Dadurch läßt sich die Bildung der rechentechnisch aufwendigen Ableitungen vermeiden.

Eine weitere Möglichkeit zur Schätzung der Varianz-Kovarianz-Matrix wird von Joe und Xu (1996)[7] vorgeschlagen und besteht in der Anwendung eines Jackknife-Verfahrens. Bei dem "deleted-one" Jackknife wird die Schätzung des IFM-Parameters  $n$ -mal unter Weglassung der jeweils  $i$ -ten Beobachtung durchgeführt. Der Schätzer mit der  $i$ -ten gelöschten Beobachtung wird bezeichnet mit  $\tilde{\theta}^{(i)}$ . Der Jackknife-Schätzer der Varianz-Kovarianz-Matrix ergibt sich, mit:

$$\sum_{i=1}^n = (\tilde{\theta}^{(i)} - \tilde{\theta})'(\tilde{\theta}^{(i)} - \tilde{\theta}),$$

unter der Annahme, dass  $\tilde{\theta}^{(i)}$  bzw.  $\tilde{\theta}$  Zeilenvektoren sind. Für einen sehr großen Stichprobenumfang kann das "deleted-one" Jackknife enorm zeitaufwendig sein, so dass stattdessen ein "deleted- $m$ " Jackknife angewendet werden kann. Dafür werden aus den  $n$  Beobachtungen  $g$  Gruppen vom Umfang  $m$  gebildet, so dass  $n = gm$  gilt. Nun erfolgt die Wiederholung der Schätzung unter Weglassen der  $k$ -ten Gruppe. Der geschätzte Parametervektor wird angegeben mit  $\tilde{\theta}^{(k)}$  und der Schätzer der Varianz-Kovarianz-Matrix unter der gleichen Annahme lautet:

$$\sum_{k=1}^g = (\tilde{\theta}^{(k)} - \tilde{\theta})'(\tilde{\theta}^{(k)} - \tilde{\theta}).$$

Anwendung findet die IFM-Methode zum Einen in der Bereitstellung von Startwerte für die FML-Methode und zum Anderen als alternative Methode an Stelle der FML-Schätzung, wenn diese rechentechnisch zu intensiv ist.

## 3 Simulationsstudie

### 3.1 Ziel und Aufbau der Simulation

Die folgende Studie untersucht das Verhalten der Schätzer von diskreten Copula Modellen bei bestimmten Problemstellungen bezüglich der Kovariablen des Modells. Insbesondere wird das Verhalten bei Vorliegen von Kollinearitäten und von Schiefe in den Kovariablen näher betrachtet. Für die Schätzung der Parameter stehen die beiden erwähnten Methoden aus 2.4 zur Verfügung, deren asymptotische Effizienz geprüft werden soll. Die Simulation beschränkt sich auf die Betrachtung von ordinalen Responsevariablen mit drei Ausprägungen. Für die Simulation der abhängigen Variablen wird auf die Beziehung (2.12) zurückgegriffen

$$Y_j = I(\alpha_j(s_j - 1) - x_j\beta_j < Z_j \leq \alpha_j(s_j) - x_j\beta_j),$$

mit  $\alpha_j(s_j)$  den Schwellenwerten und  $\beta_j$  den Koeffizienten der Kovariablen. Die Simulationsstudie weist folgendes Schema auf:

→ Stichprobenumfang  $n = 100, 500$

(1) Für  $d = 2$ ,  $m_j = 3$ ,  $\alpha_j = (-0.5, 0.5)$ ,  $\beta_j = 0.5$  und  $\psi = 4$  werden folgende Situationen untersucht:

(a)  $X_j = I(U_j \leq 0)$  mit  $U_j \stackrel{\text{iid}}{\sim} U(-1, 1)$ ,

(b)  $X \sim \text{Multinom}(0.35, 0.15, 0.15, 0.35)$  und dem Odd Ratio  $\psi_{x_{i1}; x_{i2}} = 5.4$ ,

(c)  $X_j = I(U_j \leq 0)$  mit  $U_j \stackrel{\text{iid}}{\sim} U(-1, 1)$  und  $X_1 = X_2$ .

→ Stichprobenumfang  $n = 100$

(2) Für  $d = 2$ ,  $m_j = 3$ ,  $\alpha_j = (-0.5, 0.5)$ ,  $\beta_j = 0.5$  und  $\psi = 4$  werden folgende Situationen untersucht:

(a)  $X_j = I(U_j \leq u_j)$  mit  $U_j \stackrel{\text{iid}}{\sim} U(-1, 1)$  und  $-0.8 \leq u_j \leq 0.8$ ,

(b)  $X_j = I(U_j \leq u_j)$  mit  $U_j \stackrel{\text{iid}}{\sim} U(-1, 1)$ ,  $-0.8 \leq u_1 \leq 0.8$  und  $0.8 \geq u_2 \geq -0.8$ .

(3) Für  $d = 2$ ,  $m_j = 3$ ,  $\alpha_j = (-0.5, 0.5)$ ,  $\beta_j = 0.5$  werden folgende Situationen untersucht:

(a)  $X_j = I(U_j \leq 0)$  mit  $U_j \stackrel{\text{iid}}{\sim} U(-1, 1)$  und  $0 < \psi < 4$ .

In den durchgeführten Simulationsstudien liegt die Anzahl der Wiederholung jeweils bei  $k = 1000$ . Die Szenarien (1a), (1b) und (1c) untersuchen verschiedene Konstellationen von Kovariablen der verschiedenen Responses. Simuliert wird ein Szenario mit Unabhängigkeit zwischen den Kovariablengruppen (1a), mit einer positiven Abhängigkeit (1b) und einer perfekten Abhängigkeit (1c) zwischen den Kovariablen der beiden Responses. Der Einfluss von Schiefe in den Kovariablen, auf die zu schätzenden Parameter, wird mit dem Untersuchungsdesign aus (2a) und (2b) modelliert. Betrachtet wird das Verhalten der Schätzer bei dem Vorliegen einer gleichgerichteten Schiefe und einer gegensätzlichen Schiefe. Der Einfluss des Copulaparameters wird mit (3a) analysiert.

In der vorliegenden Simulationsstudie müssen noch die Annahmen über den Zufallsvariablenvektor  $Z' = (Z'_1, \dots, Z'_d)$  getroffen werden, der als der Fehlertermvektor des Modells verstanden werden kann. Für diesen muss zum Einen die Copula und zum Anderen die Randverteilungen festgelegt werden. Im Abschnitt 2.3 wurde bereits eine Auswahl an möglichen Copulas vorgestellt. Bei der Simulationsstudie wird sich auf die Plackett-Copula beschränkt. Das Verhalten der Normal Copula wurde intensiv in Xu (1996)[16] und Zhao (2004)[17] erforscht und soll in diese Arbeit nicht weiter betrachtet werden. Für die "Mixture of max-id" Copula ist ein multivariater Simulationsalgorithmus in der Literatur nicht bekannt. Die Randverteilungen folgen jeweils einer logistischen Verteilung. In der Studie wird damit ein multivariates Logit Modell mit einer Plackett-Copula betrachtet. Die Ergebnisse der Studie können auf die multivariaten Probit Modelle übertragen werden.

Die Plackett-Copula wird verwendet, um die Variable  $Z_j$  in der vorliegenden Studie zu simulieren. Der benötigte Simulationsalgorithmus für eine Plackett-Copula existiert in der Literatur nur für den bivariaten Fall und geht zurück auf Mardia (1967)[9]. Dieser beruht auf der Simulation der Zufallszahlen aus einer bedingten Copula.

### 3.2 Effizienz

Die Simulationsstudie bewertet zum Einen die einzelnen Schätzmethoden und zum Anderen wird ein Vergleich zwischen der FML- und der IFM-Schätzung durchgeführt.

Kriterien für die Einzelbewertung sind der *Bias* und der *Mean Square Error* (MSE) der Schätzer mit:

$$\begin{aligned} BIAS(\hat{\theta}) &= \theta - E(\hat{\theta}) \\ MSE(\hat{\theta}) &= E(\hat{\theta} - \theta)^2. \end{aligned}$$

Anhand dieser Kriterien wird die Güte der Schätzungen simulativ bestimmt.

Bei der Beurteilung der beiden vorgestellten Methoden soll insbesondere die Effizienz des IFM-Verfahrens zur FML-Methode untersucht werden. Der Vergleich basiert auf der Gegenüberstellung der geschätzten Mean Square Errors und den geschätzten Varianzen der Schätzer. Daraus lassen sich drei Kennzahlen ableiten, die bei der Untersuchung angewendet werden. Die *relative Effizienz* des IFM-Schätzer zu dem ML-Schätzer wird ausgedrückt, als das Verhältnis der Root Mean Square Errors zueinander:

$$r = \frac{RMSE(\tilde{\theta})}{RMSE(\hat{\theta})} = \frac{\sqrt{MSE(\tilde{\theta})}}{\sqrt{MSE(\hat{\theta})}},$$

mit dem IFM-Schätzer  $\tilde{\theta}$  und dem FML-Schätzer  $\hat{\theta}$ . Eine weitere Kennzahl ist die *asymptotisch relative Effizienz*, die die geschätzten Varianzen ins Verhältnis setzt. Diese ist gegeben, mit:

$$ARE = \frac{Var(\tilde{\theta})}{Var(\hat{\theta})}.$$

Ein globales Maß ist das *verallgemeinerte Effizienz-Verhältnis*, welches die Summen der Hauptdiagonalen, der zu vergleichenden Varianz-Kovarianz-Matrizen betrachtet:

$$r_g = \frac{tr(J_{\Psi}^{-1})}{tr(\mathcal{I}^{-1})}.$$

Dieses ist eines der Optimalitätskriterien für Schätzfunktionen und wird bezeichnet als die "T-optimality". Weitere Ausführungen zu diesem Thema können in Xu (1996, S.42)[16] nachgelesen werden. Insgesamt kann von einer asymptotischen Effizienz der IFM-Methode zur FML-Methode gesprochen werden, wenn die Kennzahlen jeweils einen Wert nahe Eins aufweisen.

Die Varianz-Kovarianz-Matrix kann bei der IFM-Methode mit Hilfe der Godambe-Informations-Matrix oder dem Jackknife-Verfahren ermittelt werden. In der Studie werden beide Möglichkeiten betrachtet und der FML-Methode gegenübergestellt. Bei dem Jackknife-Verfahren wird das "deleted-one" Jackknife angewendet. In Xu (1996)[16] wurde für einen Stichprobenumfang von 500 bzw. 1000 gezeigt, dass für verschieden hoch gewählte  $m$ 's des "deleted-m" Jackknifes kaum Unterschiede in den Varianzen zu sehen sind. Bei einem Umfang von 100 ist eine Anwendung des "deleted-m" Jackknifes nicht sinnvoll, da der Stichprobenumfang dafür zu gering ist.

### 3.3 Simulationsergebnisse

Die Ergebnisse der durchgeführten Simulationsstudie werden nun im Folgenden vorgestellt. In den Tabellen (1) bis (3) sind die Resultate zu den Szenarien (1a) bis (1c) enthalten. Für die FML- und die IFM-Schätzung ist jeweils der Bias und der MSE (in Klammern) abgebildet. Für die IFM-Schätzung würde die Varianz mittels der Godambe-Informations-Matrix und dem "deleted-one" Jackknife-Verfahren geschätzt. Die Kennzahl  $ARE$  ist wie in (3.2) definiert. Für die Berechnung von  $ARE_1$  wird dafür die geschätzte Godambe-Informations-Matrix und für  $ARE_2$  die geschätzte Varianz des Jackknife-Verfahrens verwendet.

n	Parameter	1			2			
		$\alpha_{11}$	$\alpha_{12}$	$\beta_1$	$\alpha_{21}$	$\alpha_{22}$	$\beta_2$	$\psi$
100	FML	0.018 (0.080)	-0.007 (0.073)	-0.008 (0.129)	0.010 (0.077)	-0.006 (0.077)	0.000 (0.129)	-0.395 (3.341)
	IFM	0.016 (0.087)	-0.011 (0.080)	-0.013 (0.156)	0.015 (0.081)	-0.001 (0.081)	0.009 (0.150)	-0.190 (2.638)
	$r$	1.045	1.044	1.103	1.027	1.022	1.076	0.889
	$ARE_1$	1.083	1.081	1.181	1.083	1.083	1.181	1.003
	$ARE_2$	1.132	1.130	1.241	1.134	1.134	1.240	0.974
	FML	0.000 (0.016)	-0.005 (0.016)	-0.006 (0.025)	-0.003 (0.014)	-0.010 (0.015)	-0.008 (0.023)	-0.087 (0.494)
	IFM	-0.001 (0.018)	-0.006 (0.017)	-0.008 (0.030)	-0.004 (0.015)	-0.011 (0.016)	-0.010 (0.027)	-0.055 (0.477)
	$r$	1.040	1.037	1.097	1.044	1.036	1.077	0.982
	$ARE_1$	1.087	1.085	1.190	1.087	1.085	1.190	0.999
$ARE_2$	1.097	1.095	1.201	1.097	1.094	1.201	1.002	

Tabelle 1: Ergebnis Szenario (1a) mit  $m_j = 3$ ,  $\alpha_j = (-0.5, 0.5)$ ,  $\beta_j = 0.5$ ,  $\psi = 4$  und  $k = 1000$

Für die univariaten Parameter ist in allen drei Szenarien für beide Schätzmethoden ein sehr geringer Bias und MSE zu beobachten. Der geschätzte Copulaparameter  $\psi$  dagegen besitzt einen leichten Bias und einen hohen MSE. Dies ist insbesondere bei einem Stichprobenumfang von  $n = 100$  zu beobachten, relativiert sich jedoch mit steigenden  $n$ . Auffällig ist, dass der Bias und der MSE des IFM-Schätzers von  $\psi$  generell geringer ausfällt. Dementsprechend weist die relative Effizienz  $r$  einen Wert von unter Eins auf. Für kleine Stichproben verstärkt sich bei Vorliegen von Abhängigkeiten zwischen den Kovariablen der Bias der Schätzer, wobei sich der MSE nur leicht erhöht. Die Veränderung im Bias ist allerdings für den FML-Schätzer nur sehr wenig zu beobachten, aber bei dem IFM-Schätzer sehr deutlich ausgeprägt. Dieser Effekt schlägt sich auf die relative Effizienz  $r$  nieder, deren Wert bei zunehmender Abhängigkeit immer größer wird. Bei einem großen Stichprobenumfang von  $n = 500$  ist diese Effekt nicht vorhanden.

n	Parameter	1			2			
		$\alpha_{11}$	$\alpha_{12}$	$\beta_1$	$\alpha_{21}$	$\alpha_{22}$	$\beta_2$	$\psi$
100	FML	0.015 (0.075)	-0.011 (0.083)	-0.016 (0.138)	0.008 (0.080)	-0.015 (0.078)	-0.018 (0.134)	-0.539 (3.443)
	IFM	0.011 (0.080)	-0.015 (0.086)	-0.024 (0.151)	0.002 (0.086)	-0.023 (0.085)	-0.034 (0.157)	-0.354 (2.843)
	$r$	1.030	1.023	1.043	1.035	1.046	1.083	0.909
	$ARE_1$	1.077	1.075	1.162	1.077	1.074	1.162	1.000
	$ARE_2$	1.127	1.123	1.222	1.127	1.125	1.221	1.007
	FML	0.015 (0.014)	-0.011 (0.014)	-0.016 (0.023)	0.008 (0.015)	-0.015 (0.014)	-0.018 (0.024)	-0.043 (0.432)
500	IFM	0.012 (0.016)	0.008 (0.015)	0.009 (0.028)	0.005 (0.016)	0.000 (0.015)	-0.001 (0.028)	-0.014 (0.419)
	$r$	1.040	1.036	1.105	1.044	1.027	1.079	0.984
	$ARE_1$	1.075	1.074	1.160	1.075	1.074	1.160	0.999
	$ARE_2$	1.087	1.085	1.170	1.084	1.082	1.171	0.998

Tabelle 2: Ergebnis Szenario (1b) mit  $m_j = 3$ ,  $\alpha_j = (-0.5, 0.5)$ ,  $\beta_j = 0.5$ ,  $\psi = 4$  und  $k = 1000$

Vergleicht man die Werte der Effizienz-Kennzahlen, so liegen diese alle Nahe dem Wert Eins. Das Szenario (1c) mit der höchsten simulierten Korrelation erzielt dabei das beste Ergebnis. Über alle drei Simulationen ist die Effizienz der asymptotischen Varianz beruhend auf der Godambe-Informations-Matrix ( $ARE_1$ ) für beide betrachteten Stichprobenumfänge gegeben. Für das Jackknife-Verfahren ( $ARE_2$ ) kann eine Überschätzung der Varianzen beobachtet werden. Die Effekte sind insbesondere bei einem kleinen Stichprobenumfang vorhanden,

reduzieren sich bei einem größeren Umfang. Auffallend in den Szenarien (1a) und (1b) sind die Werte der asymptotisch relativen Effizienz ( $ARE_1$ ,  $ARE_2$ ) für die  $b_j$ -Parameter, den Effektgrößen der Kovariablen. Diese sind vergleichsmäßig hoch, auch für  $n = 500$ .

n	Parameter	1			2			
		$\alpha_{11}$	$\alpha_{12}$	$\beta_1$	$\alpha_{21}$	$\alpha_{22}$	$\beta_2$	$\psi$
100	FML	0.010 (0.081)	0.000 (0.085)	0.005 (0.159)	0.023 (0.082)	0.005 (0.082)	0.008 (0.149)	-0.459 (3.107)
	IFM	0.011 (0.084)	0.001 (0.087)	0.008 (0.163)	0.023 (0.084)	0.005 (0.084)	0.007 (0.152)	-0.420 (2.972)
	$r$	1.018	1.015	1.013	1.010	1.011	1.010	0.978
	$ARE_1$	1.013	1.013	1.014	1.013	1.013	1.014	0.999
	$ARE_2$	1.061	1.062	1.066	1.060	1.059	1.065	1.089
	FML	-0.008 (0.015)	-0.008 (0.016)	-0.010 (0.025)	0.002 (0.017)	-0.001 (0.016)	0.001 (0.030)	-0.078 (0.438)
500	IFM	-0.009 (0.016)	-0.009 (0.016)	-0.011 (0.026)	0.003 (0.017)	0.000 (0.016)	0.002 (0.031)	-0.071 (0.435)
	$r$	1.011	1.010	1.008	1.012	1.012	1.013	0.996
	$ARE_1$	1.014	1.014	1.015	1.014	1.014	1.015	0.998
	$ARE_2$	1.024	1.024	1.025	1.024	1.023	1.025	1.015
	FML	-0.008 (0.015)	-0.008 (0.016)	-0.010 (0.025)	0.002 (0.017)	-0.001 (0.016)	0.001 (0.030)	-0.078 (0.438)
	IFM	-0.009 (0.016)	-0.009 (0.016)	-0.011 (0.026)	0.003 (0.017)	0.000 (0.016)	0.002 (0.031)	-0.071 (0.435)

Tabelle 3: Ergebnis Szenario (1c) mit  $m_j = 3$ ,  $\alpha_j = (-0.5, 0.5)$ ,  $\beta_j = 0.5$ ,  $\psi = 4$  und  $k = 1000$

Zusammenfassend für das Szenario (1) hat die Abhängigkeit zwischen den Kovariablen lediglich bei einem kleinen Stichprobenumfang einen Einfluss auf die IFM-Schätzung. Eine Verzerrung in der Schätzung betrifft nur den Copulaparameter und nicht die Parameter der Randverteilung. Bei der FML-Schätzung ist der erhöhte Bias bzw. MSE für den Abhängigkeitsparameter schon bei unabhängigen Kovariablen vorhanden, so dass hier die IFM-Schätzung der FML-Methode überlegen ist. Auch bei Abhängigkeiten zwischen den Kovariablen ist die Effizienz der IFM-Schätzung gegeben, insbesondere für die Schwellenwerte und den Abhängigkeitsparameter. Die Varianz für die Effektparameter der Kovariablen wird leicht überschätzt. Bei einem kleinen Stichprobenumfang ist allerdings von dem "deleted-one" Jackknife abzuraten. Die Varianz-Kovarianz-Matrix sollte stattdessen auf Basis der Godambe-Informations-Matrix erzeugt werden.

Die Simulation des zweiten Szenarios befasst sich mit dem Einfluss der Schiefe auf die Schätzmethodik. Die beiden durchgeführten Szenarien (2a) und (2b) führen zu ähnlichen Ergebnissen, so dass die genauere Auswertung der Ergebnisse sich lediglich auf das Szenario (2a) bezieht. Eine graphische Darstellung der Resultate des Szenarios mit gleichgerichteter Schiefe (2a) ist den Abbildungen (3) bis (10) zu finden. Die Abbildungen (3) bis (9) enthalten jeweils drei Graphiken für jeden einzelnen geschätzten Parameter. In der ersten Graphik ist der Bias für die FML- und IFM-Methode gegenübergestellt, die zweite bildet den MSE ab und die letzte Graphik zeigt die asymptotisch relative Effizienz (*ARE*) für die Godambe-Informationen-Matrix und für das "deleted-one" Jackknife.  $u$  gibt dabei den Grad der Schiefe an und Symmetrie liegt vor, bei  $u = 0$ . Die Rechtsschiefe nimmt zu, mit einem steigenden  $u$  und eine Linksschiefe wird erzeugt, bei einem sinkenden  $u$ .

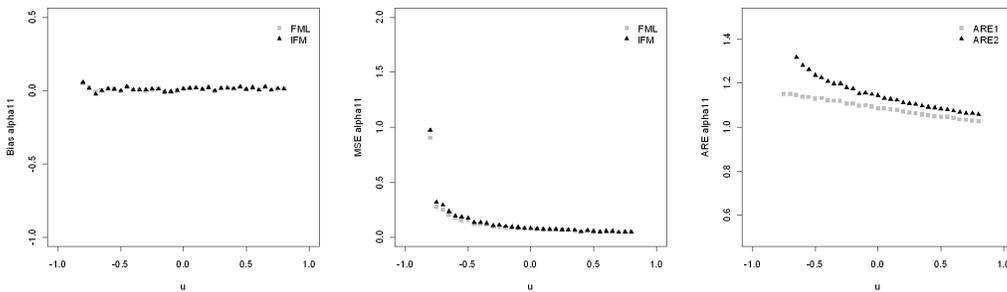


Abbildung 3: Ergebnis Szenario (2a) für den Parameter  $\alpha_{11}$  mit  $m_j = 3$ ,  $\alpha_j = (-0.5, 0.5)$ ,  $\beta_j = 0.5$ ,  $\psi = 4$ ,  $n = 100$ ,  $k = 1000$  und  $-0.8 \leq u_j \leq 0.8$

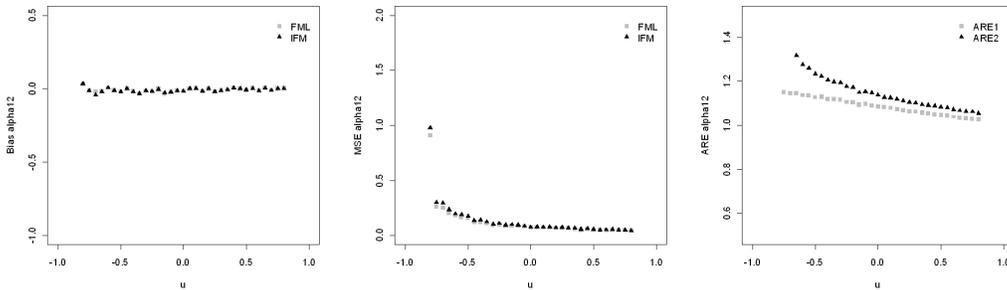


Abbildung 4: Ergebnis Szenario (2a) für den Parameter  $\alpha_{12}$  mit  $m_j = 3$ ,  $\alpha_j = (-0.5, 0.5)$ ,  $\beta_j = 0.5$ ,  $\psi = 4$ ,  $n = 100$ ,  $k = 1000$  und  $-0.8 \leq u_j \leq 0.8$

Für die Parameter der univariaten Randverteilungen ist der Bias über den betrachteten Wertebereich von  $u$  bei beiden Schätzmethoden nahe Null. Der Bias für den Abhängigkeitsparameter (Abbildung (9)) ist leicht erhöht und weist über den Schiefeparameter  $u$  Schwankungen auf, die jedoch unsystematisch sind. Wie im Szenario (1) ist für den IFM-Schätzer eine geringere Verzerrung zu beobachten.

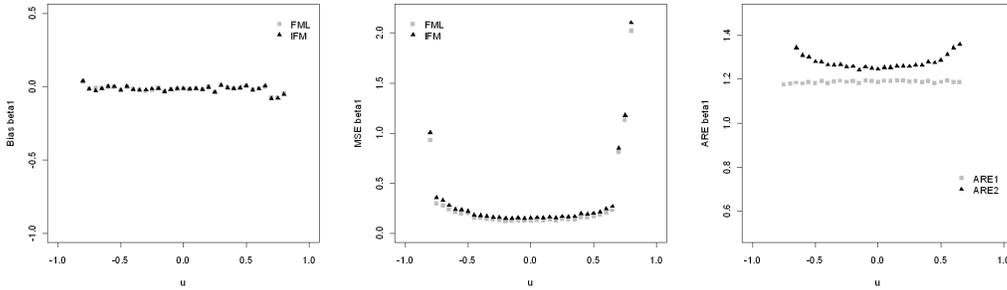


Abbildung 5: Ergebnis Szenario (2a) für den Parameter  $\beta_1$  mit  $m_j = 3$ ,  $\alpha_j = (-0.5, 0.5)$ ,  $\beta_j = 0.5$ ,  $\psi = 4$ ,  $n = 100$ ,  $k = 1000$  und  $-0.8 \leq u_j \leq 0.8$

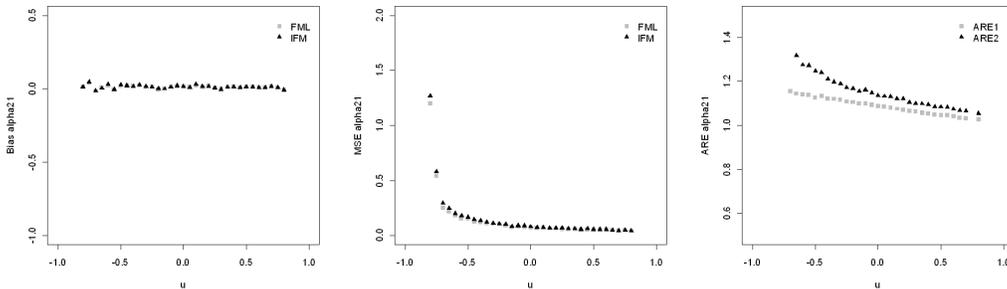


Abbildung 6: Ergebnis Szenario (2a) für den Parameter  $\alpha_{21}$  mit  $m_j = 3$ ,  $\alpha_j = (-0.5, 0.5)$ ,  $\beta_j = 0.5$ ,  $\psi = 4$ ,  $n = 100$ ,  $k = 1000$  und  $-0.8 \leq u_j \leq 0.8$

Der MSE wird für die Parameter getrennt interpretiert. Der MSE der Schwellenwerte wird nur von einer sehr starken Linksschiefe beeinflusst und sinkt mit einem zunehmenden  $u$  gegen Null. Bei den beiden Effektparametern der Kovariablen verhält sich der MSE anders. So kann hier ein erhöhter MSE bei Vorliegen einer starken Linksschiefe und bei einer starken Rechtsschiefe beobachtet werden. Je schwächer der Einfluss der Schiefe ist, desto geringer ist der Wert des MSE. Bei einer symmetrischen Verteilung der Kovariablen fällt dieser am geringsten aus. Diese Effekte treten bei beiden betrachteten Schätzmethoden auf. Bezüglich

des Copulaparameters tritt ein ähnliches Ergebnis wie beim Szenario (1) ein. Der MSE ist hier am höchsten ausgeprägt und schwankt über den Bereich von  $u$ . Der MSE der IFM-Schätzung ist jeweils geringer als der der FML-Methode.

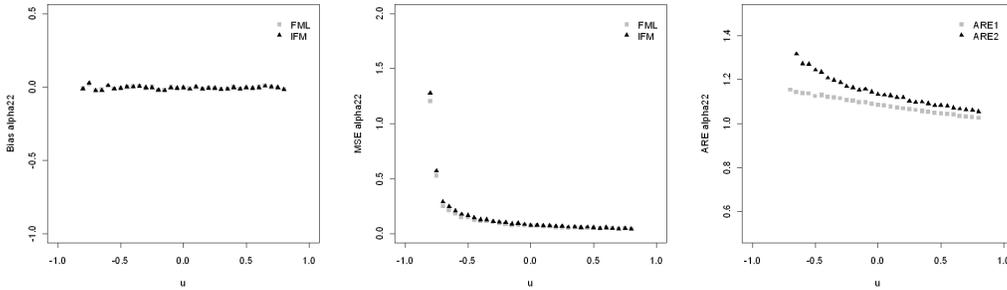


Abbildung 7: Ergebnis Szenario (2a) für den Parameter  $\alpha_{22}$  mit  $m_j = 3$ ,  $\alpha_j = (-0.5, 0.5)$ ,  $\beta_j = 0.5$ ,  $\psi = 4$ ,  $n = 100$ ,  $k = 1000$  und  $-0.8 \leq u_j \leq 0.8$

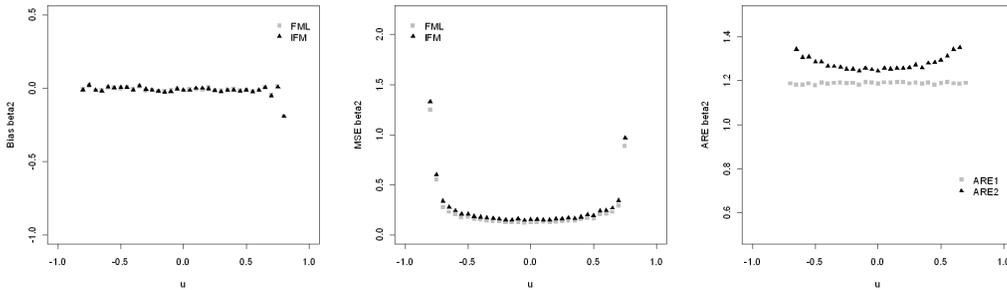


Abbildung 8: Ergebnis Szenario (2a) für den Parameter  $\beta_2$  mit  $m_j = 3$ ,  $\alpha_j = (-0.5, 0.5)$ ,  $\beta_j = 0.5$ ,  $\psi = 4$ ,  $n = 100$ ,  $k = 1000$  und  $-0.8 \leq u_j \leq 0.8$

Eine Gegenüberstellung der Kennzahlen  $ARE_1$  und  $ARE_2$  ist in der dritten Graphik in den jeweiligen Abbildungen zu finden. Ein Wert nahe Eins weist auf eine gute Effizienz der verwendeten Varianz-Kovarianz-Matrix gegenüber der FML-Methode hin. Über alle Abbildungen ist die Kennzahl nur für den Wertebereich von  $[0.5, 1.5]$  eingezeichnet. Werte außerhalb des Wertebereiches stellen eine sehr hohe Ineffizienz der IFM-Schätzungen dar. Eine unvollständige Abbildung der Kennzahlen über  $u$  ist auf eine hohe Ineffizienz in den jeweiligen Bereichen zurückzuführen. Allgemein ist die Godambe-Informations-Matrix über alle Parameter effizienter als das "deleted-one" Jackknife-Verfahren, da das  $ARE_1$  immer kleinere

Werte annimmt als das  $ARE_2$ . Für die Schwellenwerte können wieder gleiche Schlussfolgerungen gezogen werden. So sinkt der  $ARE$ -Wert bei einem steigenden Wert von  $u$ . Bei einer sehr stark ausgeprägten Linksschiefe in den Kovariablen ist die IFM-Schätzung nicht mehr effizient, dies gilt insbesondere für das "deleted-one" Jackknife.

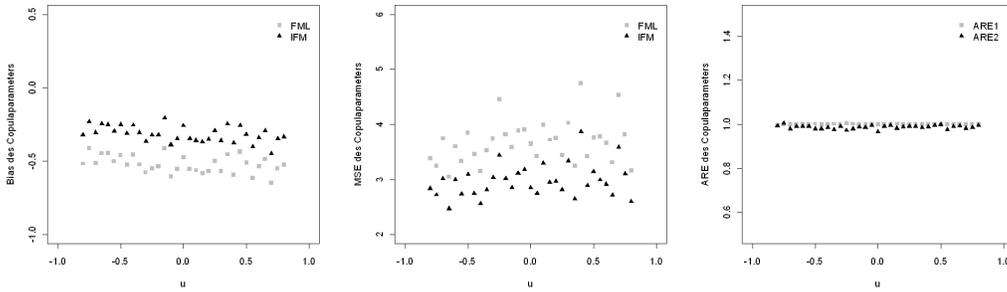


Abbildung 9: Ergebnis Szenario (2a) für den Parameter  $\psi$  mit  $m_j = 3$ ,  $\alpha_j = (-0.5, 0.5)$ ,  $\beta_j = 0.5$ ,  $\psi = 4$ ,  $n = 100$ ,  $k = 1000$  und  $-0.8 \leq u_i \leq 0.8$

Für die Parameter  $\beta_1$  und  $\beta_2$  ist ein ähnliches Bild zu beobachten, wie in Szenario (1). Auch hier weist die asymptotisch relative Effizienz den höchsten Wert auf, gesehen zu den anderen geschätzten Parametern. Die Effizienz der Godambe-Informationsmatrix wird von der Schiefe der Kovariablen nicht beeinflusst. Die Kennzahl  $ARE_1$  ist nahezu konstant. Die Werte des Jackknife-Verfahren besitzen die höchsten Ausprägungen und werden äußerst ineffizient an den Rändern des Wertebereiches von  $u$ . Für den Copulaparameter dagegen ist eine gute Effizienz der Schätzung über alle  $u$  gegeben, da bei beiden Kennzahlen einen Wert nahe Eins zu beobachten ist.

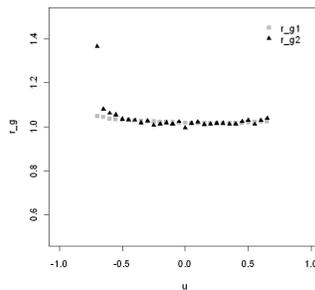


Abbildung 10: Verallgemeinerte Effizienzverhältnis  $r_g$  Szenario (2a) mit  $m_j = 3$ ,  $\alpha_j = (-0.5, 0.5)$ ,  $\beta_j = 0.5$ ,  $\psi = 4$ ,  $n = 100$ ,  $k = 1000$  und  $-0.8 \leq u_j \leq 0.8$

Eine globale Aussage kann anhand des verallgemeinerten Effizienzverhältnisses  $r_g$  in Abbildung (10) getroffen werden. An dieser Stelle wird die Kennzahl angegeben unter Verwendung der Godambe-Informations-Matrix ( $r_{g1}$ ) bzw. des Jackknifes ( $r_{g2}$ ). Insgesamt zeigt das globale Maß  $r_{g2}$  einen ähnlich guten Verlauf wie für  $r_{g1}$  auf. Bei einer starken Linksschiefe kann dies nicht mehr bestätigt werden, hier weist das verallgemeinerte Effizienzverhältnis für das Jackknife auf Ineffizienzen hin.

Im Szenario (2b) wird eine entgegengerichtete Schiefe betrachtet.  $u_1$  wird dabei wieder im Bereich  $[-0.8, 0, 8]$  modelliert und  $u_2$  nimmt den Wert  $-u_1$  an. Die graphischen Abbildungen der Ergebnisse befinden sich im Anhang und entsprechen den Resultaten des Szenarios (2a).

Für das Szenario (2) kann insgesamt festgehalten werden, dass eine Schiefe in den Kovariablen keine Auswirkung auf das Schätzergebnis zeigt. Trotzdem ist die Verwendung der IFM-Methode der FML-Methode vorzuziehen, da diese einen geringeren Bias bezüglich des Abhängigkeitsparameters aufweist. Die Varianz der Schätzer wird bei starker Schiefe insbesondere bei der Anwendung des "deleted-one" Jackknifes überschätzt. Dieses Ergebnis ist unter anderem bedingt durch den gewählten Stichprobenumfang von  $n = 100$ . Der beobachtete Effekt kann sich bei einem erhöhten  $n$  relativieren, so dass der Einsatz des Jackknife-Verfahrens in diesem Fall unbedenklich wäre.

Das Szenario (3) untersucht das Verhalten der beiden Schätzer bei unterschiedlicher Höhe des Abhängigkeitsparameters  $\psi$ . Dieser ist durch die Verwendung der Plackett-Copula im Bereich  $[0, \infty)$  definiert. Für die Interpretation des Parameters sei auf das Beispiel 2.4. verwiesen. Der betrachtete Wertebereich der Simulationsstudie beschränkt sich auf den Bereich von  $0 < \psi \leq 4$ . Die Kovariablen der Responses wurden aus einer symmetrischen Verteilung unabhängig voneinander simuliert. Die Ergebnisse des Szenarios sind in den Abbildungen (11) bis (18) festgehalten. Die ersten sieben Darstellung zeigen jeweils den Bias, den MSE und den ARE der einzelnen geschätzten Parameter. Die letzte Graphik beinhaltet eine Zusammenfassung des Effizienzverhaltens über alle Parameter.

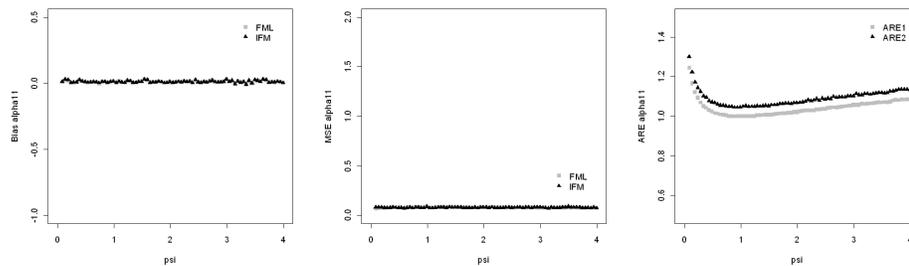


Abbildung 11: Ergebnis Szenario (3) für den Parameter  $\alpha_{11}$  mit  $m_j = 3$ ,  $\alpha_j = (-0.5, 0.5)$ ,  $\beta_j = 0.5$ ,  $n = 100$ ,  $k = 1000$  und  $0 < \psi \leq 4$

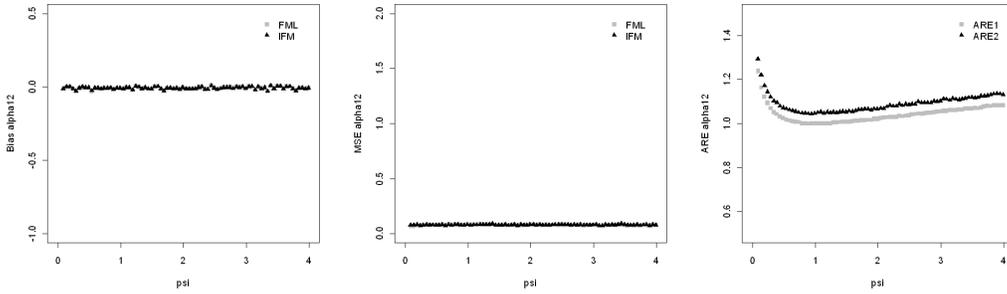


Abbildung 12: Ergebnis Szenario (3) für den Parameter  $\alpha_{12}$  mit  $m_j = 3$ ,  $\alpha_j = (-0.5, 0.5)$ ,  $\beta_j = 0.5$ ,  $n = 100$ ,  $k = 1000$  und  $0 < \psi \leq 4$

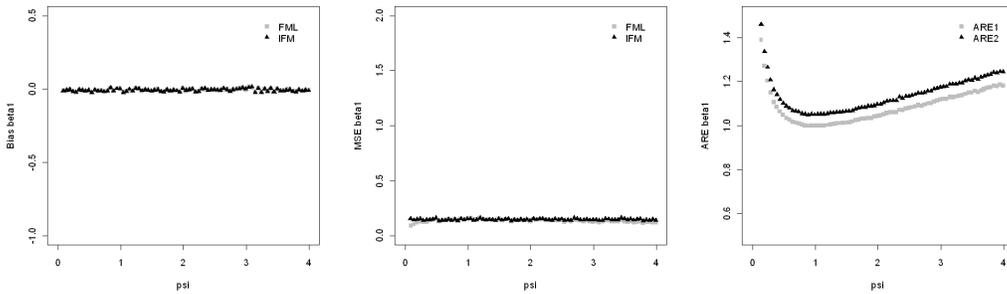


Abbildung 13: Ergebnis Szenario (3) für den Parameter  $\beta_1$  mit  $m_j = 3$ ,  $\alpha_j = (-0.5, 0.5)$ ,  $\beta_j = 0.5$ ,  $n = 100$ ,  $k = 1000$  und  $0 < \psi \leq 4$

In Bezug auf die Schwellenwert- und die Kovariablenparameter ist zu beobachten, dass der Grad der modellierten Abhängigkeit keinerlei Einfluss auf den MSE und den Bias der Schätzer haben. Allerdings sind Einflüsse auf die Effizienz des IFM-Schätzers im Vergleich zum FML-Schätzer zu sehen. Bei Vorliegen von starken Abhängigkeiten werden die Varianzen der IFM-Schätzer gegenüber dem FML-Schätzern überschätzt. Dies ist in den erhöhten Werten der *ARE*-Kennzahl zu sehen. Für die Kovariablen zeigt sich, dass sich, bei erhöhten Abhängigkeiten zwischen den Responses, die Effizienz der IFM-Schätzung stärker verschlechtert, als bei den Schwellenwertparametern. Dieser Effekt könnte schon in den vorherigen Szenarien mit  $\psi = 4$  beobachtet werden. Ebenfalls kann die verstärkte Überschätzung des Jackknife-Verfahrens (siehe  $ARE_2$ ) festgestellt werden. Die Ergebnisse für die Kovariablen und Schwellenwertparameter aus den ersten beiden Szenarien können bestätigt werden.

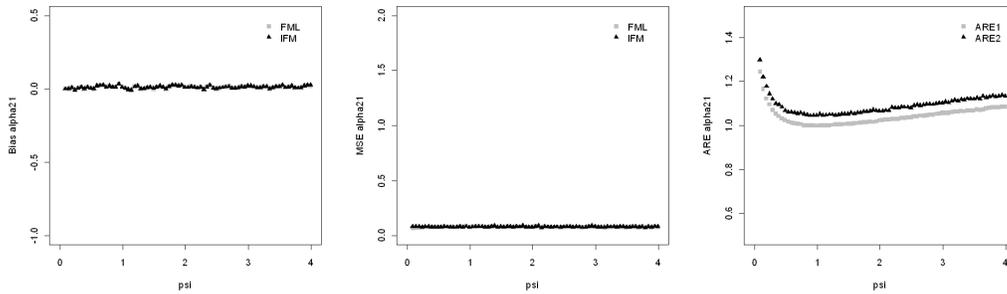


Abbildung 14: Ergebnis Szenario (3) für den Parameter  $\alpha_{21}$  mit  $m_j = 3$ ,  $\alpha_j = (-0.5, 0.5)$ ,  $\beta_j = 0.5$ ,  $n = 100$ ,  $k = 1000$  und  $0 < \psi \leq 4$

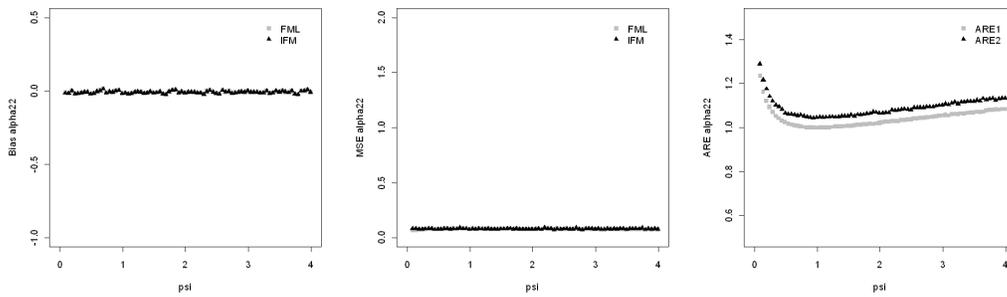


Abbildung 15: Ergebnis Szenario (3) für den Parameter  $\alpha_{22}$  mit  $m_j = 3$ ,  $\alpha_j = (-0.5, 0.5)$ ,  $\beta_j = 0.5$ ,  $n = 100$ ,  $k = 1000$  und  $0 < \psi \leq 4$

Die Darstellung (17) zeigt die Resultate für den Copulaparameter. Der Bias und der MSE des Schätzer erhöhen sich mit einer Erhöhung von  $\psi$ . Dies bedeutet, dass eine Unterschätzung des Copulaparameters nur bei einer positiven Abhängigkeit auftritt. Auch hier ist die Verzerrung des IFM-Schätzers geringer als die des FML-Schätzers. Bei dem Parameter  $\psi$  spricht man von einer positiven Abhängigkeit, wenn dieser Werte größer als Eins annimmt. Da der Parameter keinen eingeschränkten Wertebereich bezüglich des positiven Astes besitzt, ist es schwierig, diesen zu interpretieren. Eine Hilfestellung dazu gibt die Abbildung (19). Diese zeigt die Veränderung des Korrelationskoeffizienten von Pearson  $\rho$  bei einer Veränderung des Copulaparameters  $\psi$  für zwei verschiedene Wertebereiche. Die Darstellungen zeigen, dass eine geringe Veränderung in  $\psi$  (z.B. um 0.5) eine minimale Veränderung im Wert der Korrelation nach Pearson nach sich zieht und sich damit wenig in der Interpretation verändert. Dieser Effekt ist ab einem  $\psi > 4$  zu beobachten. Für die Schätzung bedeutet dies, dass zwar ein Bias für den Abhängigkeitsparameter vorhanden ist mit einem steigenden  $\psi$ , allerdings

auf die Interpretation ein geringer Einfluß besteht.

Die *ARE*-Kennzahlen weisen auf eine Ineffizienz des IFM-Schätzers bei starken negativen Zusammenhängen zwischen den Responses hin. Das Jackknife-Verfahren weist eine ähnlich gute Effizienz auf, wie die geschätzte Varianz des IFM-Schätzers mit der Hilfe der Godambe-Informations-Matrix.

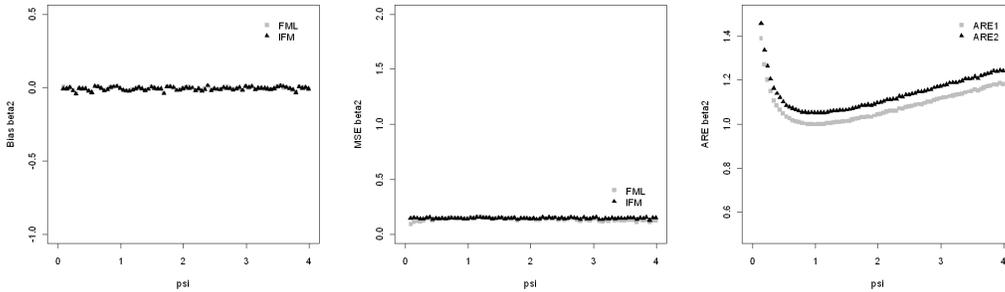


Abbildung 16: Ergebnis Szenario (3) für den Parameter  $\beta_2$  mit  $m_j = 3$ ,  $\alpha_j = (-0.5, 0.5)$ ,  $\beta_j = 0.5$ ,  $n = 100$ ,  $k = 1000$  und  $0 < \psi \leq 4$

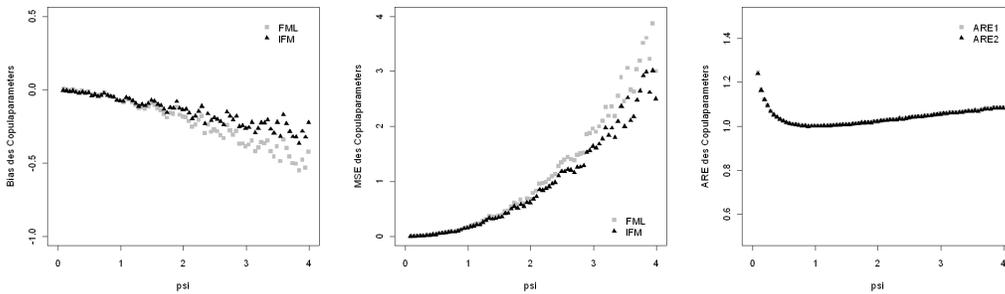


Abbildung 17: Ergebnis Szenario (3) für den Parameter  $\psi$  mit  $m_j = 3$ ,  $\alpha_j = (-0.5, 0.5)$ ,  $\beta_j = 0.5$ ,  $n = 100$ ,  $k = 1000$  und  $0 < \psi \leq 4$

Die ganzheitliche Betrachtung der Effizienz über alle Parameter erfolgt anhand der Abbildung (18). Insgesamt ist eine Überschätzung der Varianzen des IFM-Schätzers bei stärkeren Abhängigkeiten zwischen den abhängigen Variablen zu beobachten. Hierbei ist deutlich zu erkennen, dass das Jackknife-Verfahren die geschätzten Varianzen der IFM-Schätzer stärker überschätzt als die ermittelten Varianzen anhand der Godambe-Informations-Matrix.

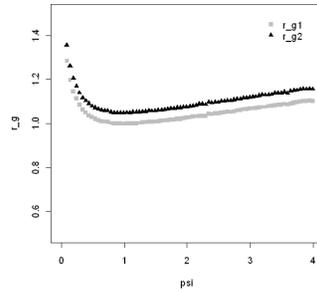


Abbildung 18: Verallgemeinerte Effizienzverhältnis  $r_g$  Szenario (3) mit  $m_j = 3$ ,  $\alpha_j = (-0.5, 0.5)$ ,  $\beta_j = 0.5$ ,  $n = 100$ ,  $k = 1000$  und  $0 < \psi \leq 4$

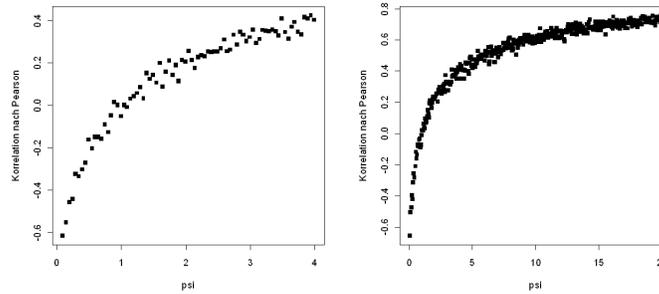


Abbildung 19: Plackett-Copula mit logistischen Rändern:  $\rho$  vs.  $\psi$  bei  $n=1000$

Das Szenario (3) zeigt, dass die Stärke der Abhängigkeit zwischen den Responses keinen Einfluss auf die Schätzer der univariaten Modelle ausübt. Für den Schätzer des Copula-parameter selber ist jedoch ein Bias vorhanden, der für den IFM-Schätzer wieder geringer ausfällt. Die Varianzen der IFM-Schätzer werden bei Abhängigkeiten leicht überschätzt, wobei der Effekt bei dem Jackknife-Verfahren verstärkt auftritt. Allerdings muss beachtet werden, dass die Ergebnisse auf einem Umfang von  $n = 100$  beruhen. Wie im Szenario (1) gezeigt, relativiert sich der Effekt auf die Effizienz bei einer Erhöhung des Stichprobenumfangs. Von dem Jackknife-Verfahren ist demnach nur bei einem kleinen Stichprobenumfang abzuraten.

## 4 Zusammenfassung

Die drei Szenarien haben den Einfluss der Multikollinearität, der Schiefe in den Kovariablen und der Stärke der Abhängigkeit zwischen den Responses untersucht.

Eine Verzerrung in den univariaten Schätzern ist bei keinem Szenario zu beobachten. Der Bias des Copulaparameters steigt leicht bei Kollinearitäten in den Kovariablen und bei einer Erhöhung des Zusammenhangs zwischen den Responses. Allerdings ist diese Erhöhung des Bias nur geringfügig und hat keinen Einfluß auf die Interpretation der Stärke des Einflusses. Der IFM-Schätzer weist in allen Szenarien für den Copulaparameter jedoch den geringeren Bias auf im Vergleich zu dem FML-Schätzer.

Die Effizienz des IFM-Schätzers des Schwellenwertparameters ist nur bei Vorliegen sehr linkschiefer Kovariablen und bei starken Abhängigkeiten zwischen den Responses nicht gegeben. Die Schätzung der Varianz des Copulaparameters mittels des Jackknifes oder der Godambe-Informations-Matrix ist sehr effizient bei allen Szenarien. Die Varianz des IFM-Schätzers für den Kovariablenparameter wird bei Abhängigkeiten zwischen den Responses überschätzt. Dieser Effekt ist im Szenario (1) und (2) ebenfalls zu beobachten, da hier ein  $\psi = 4$  unterstellt wurde. Alle drei Szenarien weisen darauf hin, dass von der Verwendung des Jackknife-Verfahrens zu Schätzung der Varianzen bei einem kleinen Stichprobenumfang abzuraten ist.

## 5 Anhang

R-Code zur Lösung des Polynoms 4. Grades nach der Methode von Ferrari zur Konstruktion einer trivariaten Copula nach dem Prinzip von Molenberghs und Lesaffre.

```
A=expression((psi-1))
B=expression((((-a-C12-C13-C23)*psi+b1+b2+b3)))
C1=expression((((a+C12+C13)*C23+(a+C12)*C13+a*C12)*psi+(-b1-b2)*b3-b1*b2))
D1=expression(((((-a-C12)*C13-a*C12)*C23-a*C12*C13)*psi+b1*b2*b3))
E=expression((psi*C12*C13*C23*a))

alpha=paste("((-3*",B,"^2)/(8*",A,"^2)+",C1,"/",A,")")
beta=paste("(",B,"^3/(8*",A,"^3)-(",B,"*",C1,")/(2*",A,"^2)+",D1,"/",A,")")
gamma1=paste("((-3*",B,"^4)/(256*",A,"^4)+(",C1,"*",B,"^2)/(16*",A,"^3)-(",B,"*",D1,")/(4*",A,"^2)+",E,"/",A,")")

P=paste("(-",alpha,"^2/12-",gamma1,")")
Q=paste("(-",alpha,"^3/108+(",alpha,"*",gamma1,")/3-",beta,"^2/8)")
R=paste("(",Q,"/2+sqrt(",Q,"^2/4+",P,"^3/27)")")
U=paste("(",R,"^(1/3)")")
y=paste("(-5/6*",alpha,"-",U,"+",P,"/(3*",U,")")")

W=paste("(sqrt(",alpha,"+2*",y,")")")

## x1:=-B/(4*A)+(W+sqrt(-(3*alpha+2*y+(2*beta)/W)))/2;
x1=paste("-",B,"/(4*",A,")+(",W,"+sqrt(-(3*",alpha,"+2*",y,"+(2*",beta,")/",W,")"))/2")
cat(x1,file="Root1.Rdmped")
Root1=parse(file = "Root1.Rdmped")
unlink("Root1.Rdmped")

## x2:=-B/(4*A)+(-W+sqrt(-(3*alpha+2*y-(2*beta)/W)))/2;
x2=paste("-",B,"/(4*",A,")+(",W,"+sqrt(-(3*",alpha,"+2*",y,"-(2*",beta,")/",W,")"))/2")
cat(W,file="Root2.Rdmped")
Root2=parse(file = "Root2.Rdmped")
unlink("Root2.Rdmped")

## x3:=-B/(4*A)+(W-sqrt(-(3*alpha+2*y+(2*beta)/W)))/2;
x3=paste("-",B,"/(4*",A,")+(",W,"-sqrt(-(3*",alpha,"+2*",y,"+(2*",beta,")/",W,")"))/2")
cat(x3,file="Root3.Rdmped")
Root3=parse(file = "Root3.Rdmped")
unlink("Root3.Rdmped")

## x4:=-B/(4*A)+(-W-sqrt(-(3*alpha+2*y-(2*beta)/W)))/2;
x4=paste("-",B,"/(4*",A,")+(",W,"-sqrt(-(3*",alpha,"+2*",y,"-(2*",beta,")/",W,")"))/2")
cat(x4,file="Root4.Rdmped")
Root4=parse(file = "Root4.Rdmped")
unlink("Root4.Rdmped")
```

```

### Evaluation der Loesungen im Interval von [-5;5]
X=seq(-5,5,0.01) F1=plogis(X) F2=plogis(X) F3=plogis(X)

psi12=4
psi13=4
psi23=4
psi=4

C12=(1+(psi12-1)*(F1+F2))-sqrt(((1+(psi12-1)*(F1+F2))^2-4*F1*F2*psi12*(psi12-1)))/(2*(psi12-1))
C13=(1+(psi13-1)*(F1+F3))-sqrt(((1+(psi13-1)*(F1+F3))^2-4*F1*F3*psi13*(psi13-1)))/(2*(psi13-1))
C23=(1+(psi23-1)*(F2+F3))-sqrt(((1+(psi23-1)*(F2+F3))^2-4*F2*F3*psi23*(psi23-1)))/(2*(psi23-1))
b1=C12+C13-F1
b2=C12+C23-F2
b3=C13+C23-F3
a=1-F1-F2-F3+C12+C13+C23

R1=eval(Root1) R2=eval(Root2) R3=eval(Root3) R4=eval(Root4)

### Grapische Darstellung
mini=c()
maxi=c()
for(i in 1:length(a))
{
mini[i]=max(0, b1[i], b2[i], b3[i])
maxi[i]=min(C12[i],C23[i],C13[i], a[i])
}

plot(X,R1,ylim=c(0,2),xlim=c(-5,5),type="l",lty=2,lwd=2,ylab="Wurzeln")
points(X,R2,type = "l",lty=3,lwd=2)
points(X,R3,type="l",lty=4,lwd=2)
points(X,mini,type="l",col="grey",lwd=2,lty=1)
points(X,maxi,type="l",col="grey",lwd=2,lty=1)
legend(c(2,1.5),c("Wurzel1", "Wurzel2", "Wurzel3", "Frechet-Grenzen"),
col=c("black", "black", "black", "grey"),lty=c(2,3,4,1),lwd=2,bty="n")

```

Graphische Darstellung der Ergebnisse des Simulationsszenarios (2b)

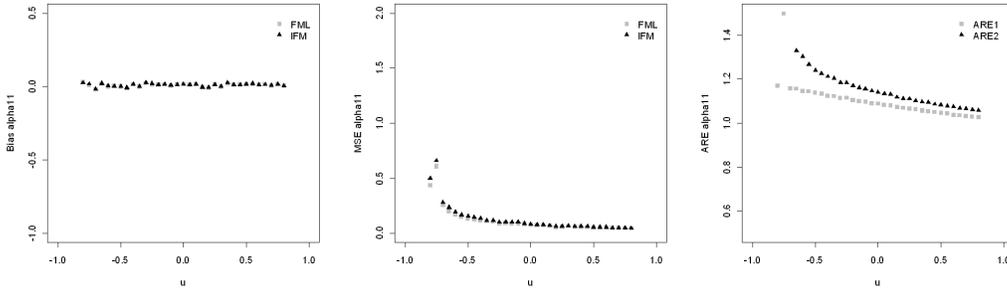


Abbildung 20: Ergebnis Szenario (2b) für den Parameter  $\alpha_{11}$  mit  $m_j = 3$ ,  $\alpha_j = (-0.5, 0.5)$ ,  $\beta_j = 0.5$ ,  $\psi = 4$ ,  $n = 100$ ,  $k = 1000$  und  $-0.8 \leq u_j \leq 0.8$

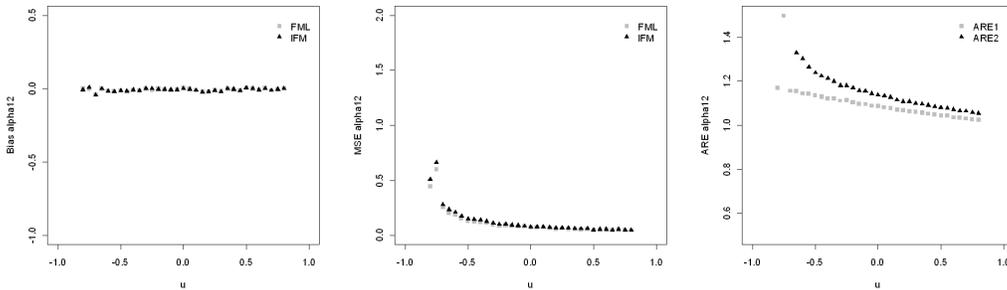


Abbildung 21: Ergebnis Szenario (2b) für den Parameter  $\alpha_{12}$  mit  $m_j = 3$ ,  $\alpha_j = (-0.5, 0.5)$ ,  $\beta_j = 0.5$ ,  $\psi = 4$ ,  $n = 100$ ,  $k = 1000$  und  $-0.8 \leq u_j \leq 0.8$

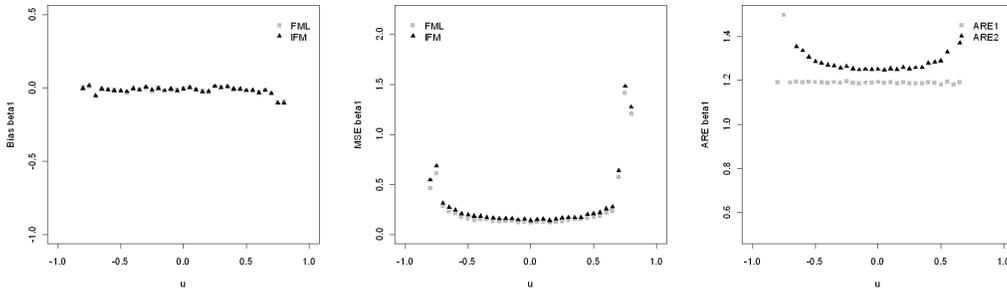


Abbildung 22: Ergebnis Szenario (2b) für den Parameter  $\beta_1$  mit  $m_j = 3$ ,  $\alpha_j = (-0.5, 0.5)$ ,  $\beta_j = 0.5$ ,  $\psi = 4$ ,  $n = 100$ ,  $k = 1000$  und  $-0.8 \leq u_j \leq 0.8$

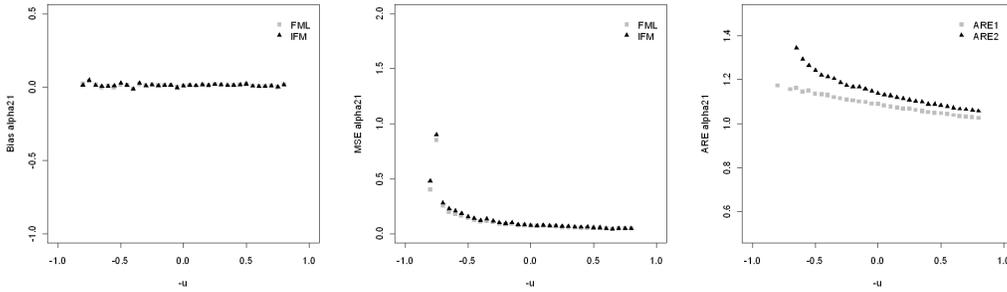


Abbildung 23: Ergebnis Szenario (2b) für den Parameter  $\alpha_{21}$  mit  $m_j = 3$ ,  $\alpha_j = (-0.5, 0.5)$ ,  $\beta_j = 0.5$ ,  $\psi = 4$ ,  $n = 100$ ,  $k = 1000$  und  $-0.8 \leq u_j \leq 0.8$

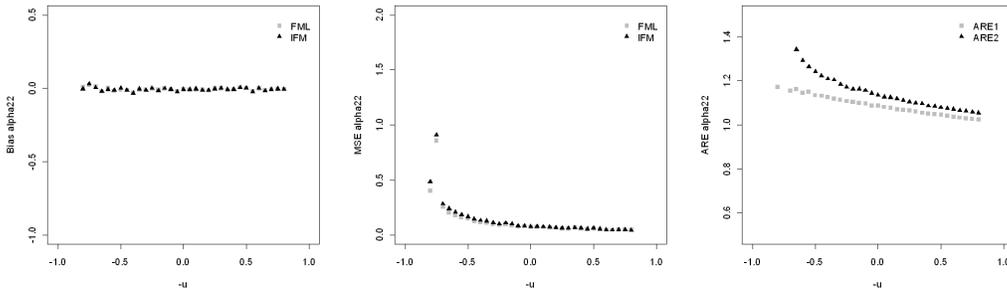


Abbildung 24: Ergebnis Szenario (2b) für den Parameter  $\alpha_{22}$  mit  $m_j = 3$ ,  $\alpha_j = (-0.5, 0.5)$ ,  $\beta_j = 0.5$ ,  $\psi = 4$ ,  $n = 100$ ,  $k = 1000$  und  $-0.8 \leq u_j \leq 0.8$

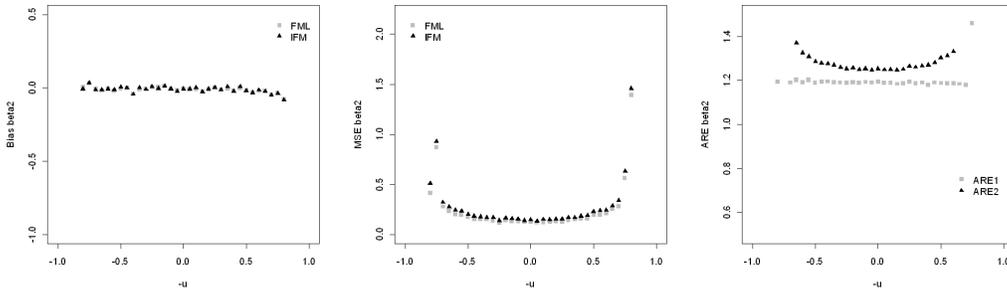


Abbildung 25: Ergebnis Szenario (2b) für den Parameter  $\beta_2$  mit  $m_j = 3$ ,  $\alpha_j = (-0.5, 0.5)$ ,  $\beta_j = 0.5$ ,  $\psi = 4$ ,  $n = 100$ ,  $k = 1000$  und  $-0.8 \leq u_j \leq 0.8$

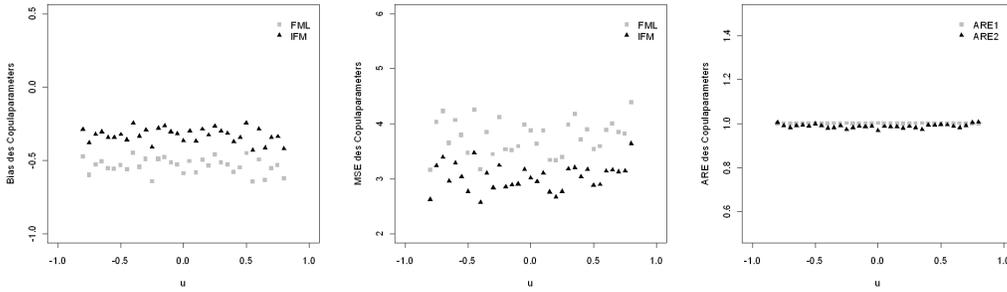


Abbildung 26: Ergebnis Szenario (2b) für den Parameter  $\psi$  mit  $m_j = 3$ ,  $\alpha_j = (-0.5, 0.5)$ ,  $\beta_j = 0.5$ ,  $\psi = 4$ ,  $n = 100$ ,  $k = 1000$  und  $-0.8 \leq u_i \leq 0.8$

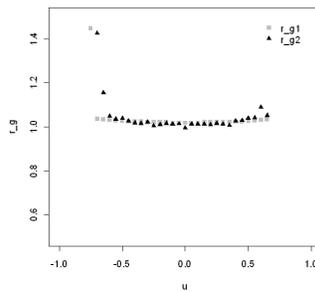


Abbildung 27: Verallgemeinerte Effizienzverhältnis  $r_g$  Szenario (2b) mit  $m_j = 3$ ,  $\alpha_j = (-0.5, 0.5)$ ,  $\beta_j = 0.5$ ,  $\psi = 4$ ,  $n = 100$ ,  $k = 1000$  und  $-0.8 \leq u_j \leq 0.8$

## Literatur

- [1] ASHFORD, J.R.; SOWDEN, T.: *Multi-Variate Probit Analysis*. Biometrics, 26:535–546, 1970.
- [2] COX, R.L.: *The analysis of multivariate binary data*. Applied Statistics, 21:113–120, 1972.
- [3] DALE, J.R.: *Global Cross-Ratio Models for Bivariate, Discrete, Ordered Responses*. Biometrics, 42:909–917, 1986.
- [4] JOE, H.: *Multivariate Models and Dependence Concepts*. Chapman and Hall, London, 1997.
- [5] JOE, H.: *Asymptotic efficiency of the two-stage estimation method for copula-based models*. Journal of Multivariate Analysis, 94:401–419, 2005.
- [6] JOE, H.; HU T.: *Multivariate Distributions from Mixtures of Max-Infinitely Divisible Distributions*. Journal of Multivariate Analysis, 57:240–265, 1996.
- [7] JOE, H; XU, J.: *The Estimation Method of Inference Functions for Margins for Multivariate Models*. Technical Report 166, Department of Statistics, University of British Columbia, 1996.
- [8] KOEHLER, J.K.; SYMANOWSKI, J.T.: *Applications To Ordered Categorical Data*. In: BALAKRISHNAN, N. (Herausgeber): *Handbook of the logistic distribution*, Seiten 495–512. American Marketing Association, Marcel Dekker Inc., 1992.
- [9] MARDIA, K.V.: *Some contributions to contingency-type bivariate distributions*. Biometrika, 54:235–249, 1967.
- [10] MOLENBERGHS, G.; LESAFFRE, E.: *Marginal Modelling of Correlated Ordinal Data Using a Plackett Distribution*. Journal of American Statistical Association, 89:633–644, 1994.
- [11] NELSEN, R.B.: *An introduction to copulas*. Springer Verlag, New York, 1999.
- [12] PLACKETT, R.L.: *A class of bivariate Distributions*. Journal of American Statistical Association, 60:516–522, 1965.
- [13] PRENTICE, R.L.: *Binary Regression using an extended beta-binomial distribution, with discussion of correlation induced by covariate measurement errors*. Journal of American Statistical Association, 81:321–327, 1986.
- [14] SCHMIDT, P.; STRAUSS, R.P.: *Estimation of models with jointly dependent qualitative variables: A simultaneous logit approach*. Econometrica, 43:745–755, 1975.
- [15] VERBEEK, M.: *A Guide to Modern Econometrics*. John Wiley & Sons, Ltd, New York/Chichester/Weinheim/Brisbane/Singapore/Toronto, 2004.

- [16] XU, J.J.: *Statistical Modelling and Inferences for Multivariate and Longitudinal Discrete Response Data*. Doktorarbeit, Department of Statistics, University of British Columbia, 1996.
- [17] ZHAO, Y.: *Statistical Modelling and Inferences for Discrete and Censored Familial Data*. Doktorarbeit, Department of Statistics, University of British Columbia, 2004.