

Risikotheorie und -management

Pflichtfach
Masterstudium der Finanz- und Versicherungsmathematik
3. Semester

Erstellt nach der Vorlesung von Eranda Dragoti-Çela im Wintersemester 2009

Institut für Optimierung und Diskrete Mathematik
Technische Universität Graz

Inhaltsverzeichnis

1 Risiko und Risikomanagement: Hintergrund und Ziele	5
2 Ermittlung der Gewinn-Verlust-Funktion	8
2.1 Verlustoperatoren	8
2.2 Der allgemeine Fall	9
2.3 Finanzderivate	9
3 Risikomaße	12
3.1 Einige grundlegende Risikomaße	12
3.2 Risikomaße basierend auf der Verlustverteilung	13
3.2.1 Die Standardabweichung ($std(L) := \sqrt{\sigma^2(F_L)}$)	13
3.2.2 Value at Risk ($VaR_\alpha(L)$)	14
3.2.3 Conditional Value at Risk ($CVaR_\alpha(L)$ - oder <i>Expected Shortfall</i> (ES))	16
3.3 Methoden zur Berechnung von VaR und CVaR	18
3.3.1 Der empirische VaR bzw. CVaR	18
3.3.2 Ein nicht-parametrisches Bootstrapping Verfahren zur Ermittlung von Konfidenzintervallen der Schätzer	19
3.3.3 Zusammenfassung des nicht-parametrischen Bootstrapping-Verfahrens zur Berechnung von Konfidenzintervallen	20
3.3.4 Eine approximative Lösung ohne Bootstrapping	20
3.3.5 Historische Simulation	21
3.3.6 Varianz-Kovarianz Methode	21
3.3.7 Monte-Carlo Verfahren	22
4 Extremwerttheorie	24
4.1 Klassische Extremwerttheorie	28
4.1.1 Grenzverteilungen von normierten und zentrierten Summen	28
4.1.2 Grenzverteilungen von normierten und zentrierten Maxima	30
4.2 Graphische Methoden zur Untersuchung des Verteilungsrandes	36
4.2.1 Der Hill Schätzer	38
4.2.2 Die POT Methode (Peaks over Threshold)	40
4.2.3 Schätzer für den Tail und das Quantil der Exzess-Verteilung	42
4.2.4 Wie wird eine hohe Schwelle u gewählt?	43
4.2.5 Die empirische durchschnittliche Exzess-Funktion	44
4.2.6 Schätzung der Parameter γ und β	45

5	Multivariate Verteilungen	49
5.1	Grundlegende Eigenschaften von Zufallsvektoren	49
5.2	Probleme bei der Modellierung der Abhängigkeit zwischen Finanzgrößen mit Hilfe der (multivariaten) Normalverteilung	52
5.3	Abhängigkeitsmaße	53
5.3.1	Lineare Korrelation	53
5.3.2	Rangkorrelation	54
5.3.3	Tail-Abhängigkeit	55
5.4	Multivariate elliptische Verteilungen	56
5.4.1	Die multivariate Normalverteilung	56
5.4.2	Varianz-gemischte Normalverteilungen	57
5.4.3	Sphärische Verteilungen	58
5.4.4	Elliptische Verteilungen	59
5.5	Koherente Risikomaße	63
5.6	Konvexe Risikomaße	64
5.7	Elliptische Verteilungen und Portfoliooptimierung	66
6	Einführung in Copulas	68
6.1	Grundlegende Eigenschaften	68
6.2	Co-Monotonie und Anti-Monotonie	72
6.3	Kendall's Tau und Spearman's Rho	77
6.4	Tail-Abhängigkeit	78
6.5	Elliptische Copulas	80
6.6	Weitere Eigenschaften von Copulas	82
6.7	Archimedische Copulas	85
6.7.1	Bivariate Archimedische Copulas	85
6.7.2	Multivariate Archimedische Copulas	87
6.7.3	Vorteile und Nachteile Archimedischer Copulas	88
6.8	Simulation Archimedischer Copulas	88
6.8.1	Simulation von Gauss'schen Copulas und t-Copulas	88
6.8.2	Simulation von Gumbel und Clayton Copulas	89
6.9	Schätzung von Copulas	91
6.9.1	Die Schätzer $\hat{\theta}$ für C_R^{Ga} , $C_{\nu,R}^t$, C_{θ}^{Cl} und C_{θ}^{Gu}	92
6.9.2	Schätzung von Gauss'schen Copulas und t-Copulas	92
6.10	Die wichtigsten Familien von Copulas im Überblick	93
7	Kreditrisiko	95
7.1	Modelle	95
7.1.1	Ein einfaches Modell	95
7.1.2	Modelle mit latenten Variablen	96
7.1.3	Das univariate KMV Modell (siehe auch www.moodyskmv.com)	96
7.1.4	Das multivariate KMV Modell: Berechnung von multivariaten Default Wahrscheinlichkeiten	99
7.1.5	Credit Metrics	100
7.1.6	CreditRisk ⁺ - Ein Poisson Mixture Modell	104
7.2	Monte Carlo Methoden im Kreditrisiko-Management	107
7.2.1	Grundlagen von „Importance Sampling“	108

7.2.2	Exponential tilting: Bestimmung der IS-Dichte für „light tailed“ Variablen	110
7.2.3	IS im Falle von Wahrscheinlichkeitsmaßen	111

Kapitel 1

Risiko und Risikomanagement: Hintergrund und Ziele

Begriffsherkunft:

- Risicare: Gefahr laufen, wagen
- Resecum: Felsklippe

Risiko: die aus der Unvorhersagbarkeit der Zukunft resultierende Möglichkeit eines Abweichens von (Unternehmens-)Zielen.

Risikomanagement: Systematisches, aktives, zukunfts- und zielorientiertes Denken und Handeln im Umgang mit unternehmerischen Risiken.

Gegenstand des RM:

- Sowohl Einzelrisiken als auch das Gesamtrisiko der Unternehmung
- Berücksichtigung von Risikointerdependenzen
- Integration des RM in die Unternehmenssteuerung

Zentrale Fragen des strategischen RM:

- Welche sind die strategischen Risiken?
- Welche Risiken soll das Unternehmen selbst tragen?
- Welche(r) Instrumente(mix) sollen zur Steuerung der Risiken zum Einsatz kommen?
- Welches Risikodeckungspotential ist erforderlich?
- Welcher risikoadjustierte Erfolgsmaßstab dient als Zielgröße der Unternehmenssteuerung?

Beispiel 1.0.1. Anfangskapital $V_0 = 100$.

Spiel: man verliert oder gewinnt 50 mit einer Wahrscheinlichkeit von jeweils 1/2.

Kapital nach dem Spiel $V_1 = \begin{cases} 150 & \text{mit Wahrsch. } 1/2 \\ 50 & \text{mit Wahrsch. } 1/2 \end{cases}$

◇

Sei $X := V_1 - V_0$ der Gewinn/Verlust. Die Verteilungsfunktion der Zufallsvariable X heißt **Gewinn/Verlust Verteilung (GVV)**.

Die Verteilungsfunktion von $L := V_0 - V_1$ heißt **Verlustverteilung**.

$L \geq 0 \Rightarrow$ Risiko!

Viele Leute hätten lieber keinen Gewinn und keinen Verlust mit Sicherheit, als entweder Gewinn oder Verlust von 50 Einheiten mit Wahrscheinlichkeit von jeweils $1/2 \Rightarrow$ **Risikoaversion!**

Die Entscheidung, ob gespielt wird oder nicht, hängt von der Verlustverteilung ab. Diese ist aber in der Regel unbekannt!

Definition 1.0.2. Ein Risikomaß ρ ist eine Abbildung der Zufallsvariablen in die reellen Zahlen, die jeder Zufallsvariable L eine reelle Zahl $\rho(L) \in \mathbb{R}$ zuordnet.

Beispiel 1.0.3. Mögliche Risikomaße sind:

- Standardabweichung (Nachteil: symmetrisch; bestraft Verluste wie Gewinne gleichermaßen)
- Quantil der Verlustverteilung (z.B. VaR, CVaR. Nachteil: i.A. keine konvexen Funktionen)

◇

Warum Risikomanagement ?

Das Volumen des risikoreichen Handels im Globalen Markt steigt kontinuierlich:

Global OTC Derivatives: Nominalwert in Trillionen von USD¹

Kontrakte	2008	2007	2006	2005	2004	2003	2002
interest rate derivatives	403	382,3	285,7	213,2	164,5	142,3	99,8
credit default swaps	38,6	62,2	34,4	17,1	5,44	3,78	2,15
equity derivatives	8,7	10	7,2	5,6	4,2	3,4	2,5

Beispiele großer Verluste in den Finanzmärkten ²

- Orange County (1994)
- LTCM (1998)
- BAWAG (2006)
- Lehman Brothers (2008)
- Barings Bank (1995)
- Bankgesellschaft Berlin (2001)
- Fannie May and Freddie Mac (2008)
- Hypo Real Estate (2008)

Risikotypen: Für eine Organisation entsteht Risiko durch Ereignisse oder Handlungen, welche die Organisation daran hindern könnten ihre Verpflichtungen zu erfüllen bzw. ihre Strategien durchzuführen.

¹Quelle: ISDA - International Swaps and Derivatives Association, Inc. <http://www.isda.org>

²siehe zB. <http://www.erisk.com>

Finanzielles Risiko:

- Marktrisiko
- Kreditrisiko
- Operationelles Risiko
- Liquiditätsrisiko, Rechtliches Risiko, Rufschädigungsrisiko

Es wird versucht diese Risiken möglichst genau abzuschätzen; dazu wird idealerweise die GVV verwendet.

Regulierung und Aufsicht: Gründung des Basler Ausschusses für Bankenaufsicht in 1974.

- Sicherheitskapital abhängig von der GVV.
- Basler Ausschuss: Vorschläge und Richtlinien über Anforderungen und Methoden zur Berechnung des Sicherheitskapitals.
- International akzeptierte Standards für die Berechnung des Volumens des Sicherheitskapitals, sowie darauf basierende gesetzliche Bestimmungen werden angestrebt.
- Kontrolle durch die Aufsichtsbehörde.

1988 Basel I: Internationale Mindestkapitalanforderungen insbesondere bzgl. Kreditrisiko.

1996 Novelle formuliert standardisierte Modelle für Marktrisiko mit einer Option für größere Banken zur Verwendung von Value at Risk (VaR) Modellen.

2007 Basel II: Mindestkapitalanforderungen (bzgl. Kredit- und Marktrisiko sowie bzgl. operationelle Risiken), aufsichtliche Überprüfungsverfahren, Marktdisziplin³.

2009 ? BASEL III - Verbesserung und Weiterentwicklung von BASEL II im Hinblick auf die Umsetzbarkeit und operationelles Risiko.

³Siehe <http://www.bis.org>

Kapitel 2

Ermittlung der Gewinn-Verlust-Funktion

2.1 Verlustoperatoren

$V(t)$... Wert des PF zum Zeitpunkt t .

Δt ... Zeithorizont.

Verlust: $L_{[t,t+\Delta t]} := -(V(t + \Delta t) - V(t))$.

Diskretisierung der Zeit: $t_n = n\Delta t$, $n = 0, 1, 2, \dots$

$$L_{n+1} := L_{[t_n, t_{n+1}]} = L_{[n\Delta t, (n+1)\Delta t]} = -(V_{n+1} - V_n)$$

wobei $V_n = V(n\Delta t)$.

Beispiel 2.1.1. Ein Aktienportfolio:

Das Portfolio besteht aus α_i Stück von Aktie A_i , $i = 1, 2, \dots, d$.

$S_{n,i}$... Preis von Aktie A_i zum Zeitpunkt t_n .

$$\begin{aligned} V_n &= \sum_{i=1}^d \alpha_i S_{n,i} \\ X_{n+1,i} &:= \ln S_{n+1,i} - \ln S_{n,i} \\ Z_{n,i} &:= \ln S_{n,i} \end{aligned}$$

Seien $w_{n,i} := \alpha_i S_{n,i} / V_n$, $i = 1, 2, \dots, d$, die relativen Portfoliogewichte.

Es gilt:

$$\begin{aligned}
L_{n+1} &:= -(V_{n+1} - V_n) = - \left(\sum_{i=1}^d \alpha_i S_{n+1,i} - \sum_{i=1}^d \alpha_i S_{n,i} \right) \\
&= - \sum_{i=1}^d \alpha_i (\exp(Z_{n+1,i}) - \exp(Z_{n,i})) = - \sum_{i=1}^d \alpha_i \exp(Z_{n,i}) \left[\frac{\exp(Z_{n+1,i})}{\exp(Z_{n,i})} - 1 \right] \\
&= - \sum_{i=1}^d \alpha_i S_{n,i} [\exp(Z_{n+1,i} - Z_{n,i}) - 1] = - \sum_{i=1}^d \alpha_i S_{n,i} [\exp(X_{n+1,i}) - 1] \\
&= -V_n \sum_{i=1}^d w_{n,i} [\exp(X_{n+1,i}) - 1] =: l_n(X_{n+1})
\end{aligned}$$

Mit $e^x = 1 + x + o(x^2) \sim 1 + x$ für $x \ll 1$ erhält man die Linearisierung

$$L_{n+1}^\Delta = -V_n \sum_{i=1}^d w_{n,i} X_{n+1,i} =: l_n^\Delta(X_{n+1}).$$

◇

2.2 Der allgemeine Fall

$V_n = f(t_n, Z_n)$; $Z_n = (Z_{n,1}, \dots, Z_{n,d})$ ist ein Vektor von Risikofaktoren.
Veränderungen der Risikofaktoren: $X_{n+1} = Z_{n+1} - Z_n$

$$L_{n+1} = - (f(t_{n+1}, Z_n + X_{n+1}) - f(t_n, Z_n)) =: l_n(X_{n+1})$$

wobei $l_n(x) := - (f(t_{n+1}, Z_n + x) - f(t_n, Z_n))$ der Verlustoperator ist.

Der linearisierte Verlust ist dann

$$L_{n+1}^\Delta = - \left(f_t(t_n, Z_n) \Delta t + \sum_{i=1}^d f_{z_i}(t_n, Z_n) X_{n+1,i} \right)$$

wobei f_t und f_{z_i} die partiellen Ableitungen von f sind.

Damit ergibt sich der linearisierte Verlustoperator

$$l_n^\Delta(x) = - \left(f_t(t_n, Z_n) \Delta t + \sum_{i=1}^d f_{z_i}(t_n, Z_n) x_i \right).$$

2.3 Finanzderivate

... sind Finanzprodukte oder Kontrakte, die aus einem fundamentalen Basiswert (zB. Aktienpreis, Aktienindex, Zinssatz, Rohstoffpreis) abgeleitet werden.

Definition 2.3.1. Eine *Europäische Call Option (ECO)* auf eine bestimmte Aktie S gibt dem Besitzer das Recht, aber nicht die Pflicht, die Aktie S an einem Tag T um einen Preis K zu kaufen. Die Option wird um einen bestimmten Preis am Tag 0 erworben. Der Wert der ECO zum Zeitpunkt t ist

$$C(t) = \max\{S(t) - K, 0\},$$

wobei $S(t)$ der Preis der Aktie S zum Zeitpunkt t ist.

Definition 2.3.2. Eine *Nullkuponanleihe* mit Laufzeit T ist ein Kontrakt, der dem Besitzer eine Währungseinheit zum Zeitpunkt T bringt.

Definition 2.3.3. Ein *Währungs-Forward* ist ein Kontrakt zwischen zwei Parteien, der dem Käufer das Recht einräumt, eine bestimmte Menge \bar{V} einer fremden Währung zu einem bestimmten Zeitpunkt T und zu einem bestimmten Wechselkurs \bar{e} vom Verkäufer zu erwerben.

Beispiel 2.3.4. Ein Anleihen-Portfolio:

Sei $B(t, T)$ der Preis der Nullkuponanleihe zum Zeitpunkt $t < T$.

Die kontinuierliche Rendite (yield) - $y(t, T) = -\frac{1}{T-t} \ln B(t, T)$ - wird interpretiert als der kontinuierliche Zinssatz, der zum Zeitpunkt t für den gesamten Zeitraum $[t, T]$ vereinbart wurde. Für unterschiedliche Laufzeiten gibt es unterschiedliche Renditen.

Die Renditenkurve (yield curve) zum fixen Zeitpunkt t ist $T \mapsto y(t, T)$.

Das Portfolio besteht aus α_i Stück der Nullkuponanleihe i mit Laufzeit T_i und Preis $B(t, T_i)$, $i = 1, 2, \dots, d$. Der Wert des Portfolios ist

$$V_n = \sum_{i=1}^d \alpha_i B(t_n, T_i) = \sum_{i=1}^d \alpha_i \exp\{-(T_i - t_n)Z_{n,i}\} = f(t_n, Z_n),$$

wobei $Z_{n,i} := y(t_n, T_i)$ die Risikofaktoren sind.

Sei $X_{n+1,i} = Z_{n+1,i} - Z_{n,i}$ die Veränderung der Risikofaktoren, dann gilt:

$$\begin{aligned} l_{[n]}(x) &= - \sum_{i=1}^d \alpha_i B(t_n, T_i) (\exp\{Z_{n,i}\Delta t - (T_i - t_{n+1})x_i\} - 1) \\ L_{n+1}^\Delta &= - \sum_{i=1}^d \alpha_i B(t_n, T_i) (Z_{n,i}\Delta t - (T_i - t_{n+1})X_{n+1,i}). \end{aligned}$$

◇

Beispiel 2.3.5. Ein Währungs-Forward-Portfolio:

Die Partei, die die fremde Währung kauft, hält eine *Long Position*. Die Partei, die verkauft, hält eine *Short Position*.

Eine Long Position über \bar{V} Einheiten in einem Währungs-Forward entspricht einer Long Position über \bar{V} Einheiten in einer fremden Nullkuponanleihe (NCA) und einer Short Position über $\bar{e}\bar{V}$ Einheiten in einer einheimischen Nullkuponanleihe.

Es wird folgende Annahme getroffen: Der Euro-Investor hält eine Long Position in einem USD/EUR Forward über \bar{V} USD.

Sei $B^f(t, T)$ ($B^d(t, T)$) der Preis einer USD- (EUR)-basierten NCA und $e(t)$ der Kassawechselkurs (spot exchange rate) für USD/EUR.

Der Wert der Long Position des Währungs-Forwards zum Zeitpunkt T ist $V_T = \bar{V}(e(T) - \bar{e})$. Die Short Position in der einheimischen NCA kann wie im Beispiel (2.3.4) behandelt werden.

Die Long Position in der fremden NCA:

Risikofaktoren: $Z_n = (\ln e(t_n), y^f(t_n, T))^\top$, wobei $y^f(t_n, T) = -\frac{1}{T-t_n} \ln B^f(t_n, T)$.

Wert der Long Position:

$$\begin{aligned} V_n &= \bar{V} \exp\{Z_{n,1} - (T - t_n)Z_{n,2}\} \\ &= \bar{V} \exp\{\ln e(t_n) - (T - t_n) \left(-\frac{1}{T - t_n} \ln B^f(t_n, T) \right)\} \\ &= \bar{V} \exp\{\ln e(t_n) + \ln B^f(t_n, T)\} = \bar{V} \exp\{\ln(e(t_n)B^f(t_n, T))\} \\ &= \bar{V} e(t_n)B^f(t_n, T). \end{aligned}$$

Der linearisierte Verlust: $L_{n+1}^\Delta = -V_n(Z_{n,2}\Delta t + X_{n+1,1} - (T - t_n)X_{n+1,2})$ ◇

Beispiel 2.3.6. Europäische Call Option (ECO) auf eine Aktie S mit Laufzeit T und Ausübungspreis (Strikepreis) K :

Der Wert der Call Option zum Zeitpunkt T ist $\max\{S_T - K, 0\}$.

Der Preis der ECO zum Zeitpunkt $t < T$ ist $C = C(t, S, r, \sigma)$ (Black-Scholes Modell), wobei t die Zeit, S der Preis der Aktie zum Zeitpunkt t , r der Zinssatz und σ die Volatilität ist.

Die Risikofaktoren sind:

$$\begin{aligned} Z_n &= (\ln S_n, r_n, \sigma_n)^\top \\ X_{n+1} &= (\ln S_{n+1} - \ln S_n, r_{n+1} - r_n, \sigma_{n+1} - \sigma_n)^\top. \end{aligned}$$

Der Wert des Portfolios ist

$$V_n = C(t_n, S_n, r_n, \sigma_n) = f(t_n, \exp(Z_{n,1}), Z_{n,2}, Z_{n,3})$$

und der linearisierte Verlust ist

$$L_{n+1}^\Delta = -(C_t \Delta t + C_S S_n X_{n+1,1} + C_r X_{n+1,2} + C_\sigma X_{n+1,3})$$

mit den Greeks C_t (Theta), C_S (Delta), C_r (Rho) und C_σ (Vega). ◇

Kapitel 3

Risikomaße

Verwendungszweck von Risikomanagement:

- Bestimmung der Mindestkapitalanforderungen:
Kapital, das benötigt wird, um event. Verluste abzudecken.
- Als Management Tool:
zur Bestimmung der Risiken, die unterschiedliche Einheiten einer Firma eingehen dürfen.

3.1 Einige grundlegende Risikomaße

- *Gewichtete Summe der Aktiva* (Assetklassenspezifische Gewichte)
ZB. Basel I (1998):

$$\text{Cooke Ratio} = \frac{\text{Eigenkapital}}{\text{risikogewichtete Summe der Aktiva}} \geq 8\%$$

$$\text{Gewicht} := \begin{cases} 0\% & \text{für Forderungen gegenüber staatlichen Schuldner (OECD-Staaten)} \\ 20\% & \text{für Forderungen gegenüber Kreditinstituten} \\ 50\% & \text{für grundpfandrechtl. gesicherte Realkredite} \\ 100\% & \text{für alle sonstigen Risikoaktiva, d.h. alle Kredite an Unternehmen} \end{cases}$$

Nachteile: Kein Unterschied zwischen Long und Short Positionen, berücksichtigt keine Diversifikationseffekte.

- *Sensitivität gegenüber Risikofaktoren*

Der Portfoliowert zum Zeitpunkt t_n ist $V_n = f(t_n, Z_n)$, wobei Z_n ein Vektor von d Risikofaktoren ist.

Die Sensitivitätskoeffizienten sind $f_{z_i} = \frac{\delta f}{\delta z_i}(t_n, Z_n)$, $1 \leq i \leq d$, z.B.: „The Greeks“ eines Portfolios.

Nachteile: Aggregation zum Risikomaß bei simultanen Veränderungen von mehreren Faktoren schwierig; bei mehreren Märkten ist die Aggregation zum Risikomaß für das Gesamtportfolio schwierig.

- *Szenario basierte Risikomaße*

Sei N die Anzahl möglicher Veränderungen der Risikofaktoren (= Szenarien), X_i der Vektor der Risikofaktorveränderungen, wenn Szenario i , $1 \leq i \leq N$, eintritt.

$\chi = \{X_1, X_2, \dots, X_N\}$ sei die Menge der Szenarien und $l_{[n]}(\cdot)$ der Verlustoperator des Portfolios.

Jedem Szenario wird ein Gewicht w_i , $1 \leq i \leq N$, zugeordnet. Das Portfoliorisiko ist dann

$$\Psi[\chi, w] = \max\{w_1 l_{[n]}(X_1), w_2 l_{[n]}(X_2), \dots, w_N l_{[n]}(X_N)\}$$

Beispiel 3.1.1. SPAN Regeln verwendet von CME (siehe Artzner et al., 1999):

Das Portfolio besteht aus mehreren Einheiten eines Future-Kontrakts und mehreren *Put* bzw. *Call Optionen* desselben Kontrakts mit gleicher Laufzeit.

Berechnung der SPAN Marge:

Szenarien i , $1 \leq i \leq 14$:

Szenarien 1 bis 8		Szenarien 9 bis 14	
Volatilität	Preis der Future	Volatilität	Preis der Future
\nearrow	$\nearrow_{\frac{1}{3}} * Range$	\nearrow	$\searrow_{\frac{1}{3}} * Range$
\searrow	$\nearrow_{\frac{2}{3}} * Range$	\searrow	$\searrow_{\frac{2}{3}} * Range$
	$\nearrow_{\frac{3}{3}} * Range$		$\searrow_{\frac{3}{3}} * Range$
	\rightarrow		

Szenarien i , $i = 15, 16$, stellen extreme Bewegungen nach oben bzw. unten des Futurepreises dar. Die Gewichte sind

$$w_i = \begin{cases} 1 & 1 \leq i \leq 14 \\ 0,35 & 15 \leq i \leq 16 \end{cases}$$

Ein bestimmtes Modell (zB. Black-Scholes) wird verwendet um die Optionspreise in den entsprechenden Szenarien zu generieren. \diamond

3.2 Risikomaße basierend auf der Verlustverteilung

Sei $F_L := F_{L_{n+1}}$ die Verteilung des Verlustes L_{n+1} . Die Parameter von $F_{L_{n+1}}$ werden anhand von historischen Daten entweder direkt oder mit Hilfe der Risikofaktoren geschätzt.

3.2.1 Die Standardabweichung ($std(L) := \sqrt{\sigma^2(F_L)}$)

Sie wird vor allem in der Portfolio-Theorie verwendet.

Nachteile:

- Die Standardabweichung existiert nur für Verteilungen mit $\mathbb{E}(F_L^2) < \infty$, d.h. sie ist nicht einsetzbar bei leptokurtischen („fat tailed“) Verlustverteilungen
- Gewinne und Verluste beeinflussen die Standardabweichung gleichermaßen.

Beispiel 3.2.1. $L_1 \sim N(0, 2)$, $L_2 \sim t_4$ (Student-Verteilung mit 4 Freiheitsgraden).

Es gilt $\sigma^2(L_1) = 2$ und $\sigma^2(L_2) = \frac{m}{m-2} = 2$. Die Verlustwahrscheinlichkeit bei L_2 ist jedoch viel größer als bei L_1 . (Hinweis: Plotte den logarithmischen Quotient $\ln[\mathbb{P}(L_2 > x)/\mathbb{P}(L_1 > x)]$!)

\diamond

3.2.2 Value at Risk ($VaR_\alpha(L)$)

Definition 3.2.2. Sei L eine Verlustfunktion und $\alpha \in (0, 1)$ ein gegebenes Konfidenzniveau. $VaR_\alpha(L)$ ist die kleinste Zahl l , sodass $\mathbb{P}(L > l) \leq 1 - \alpha$ gilt.

$$VaR_\alpha(L) = \inf\{l \in \mathbb{R}: \mathbb{P}(L > l) \leq 1 - \alpha\} = \inf\{l \in \mathbb{R}: 1 - F_L(l) \leq 1 - \alpha\} = \inf\{l \in \mathbb{R}: F_L(l) \geq \alpha\}$$

Ein Vorschlag von der BIS (Bank of International Settlements) für den VaR ist z.B.: $VaR_{0,99}(L)$ über einen Horizont von 10 Tagen als Maß für das Marktrisiko eines Portfolios.

Definition 3.2.3. Sei $F: \mathbb{R} \rightarrow \mathbb{R}$ eine monoton steigende Funktion (d.h. $x \leq y \implies F(x) \leq F(y)$). Die Funktion

$$F^{\leftarrow}: \mathbb{R} \rightarrow \mathbb{R}, y \mapsto \inf\{x \in \mathbb{R}: F(x) \geq y\}$$

heißt *verallgemeinerte inverse Funktion von F* .

Hier gilt $\inf \emptyset = -\infty$.

Falls F streng monoton steigend ist, gilt $F^{-1} = F^{\leftarrow}$.

Beispiel 3.2.4. Sei $F: [0, +\infty) \rightarrow [0, 1]$ mit

$$F(x) = \begin{cases} 1/2 & 0 \leq x < 1 \\ 1 & 1 \leq x. \end{cases}$$

Dann ist die verallgemeinerte inverse Funktion gegeben durch

$$F^{\leftarrow}(y) = \inf\{x \in [0, \infty]: F(x) \geq y\} = \begin{cases} 1 & y > \frac{1}{2} \\ 0 & y \leq \frac{1}{2}. \end{cases}$$

◇

Definition 3.2.5. Sei $F: \mathbb{R} \rightarrow \mathbb{R}$ eine monoton steigende Funktion.

$q_\alpha(F) := \inf\{x \in \mathbb{R}: F(x) \geq \alpha\}$ heißt α -Quantil von F .

Für die Funktion L und ihre Verteilungsfunktion F gilt:

$$VaR_\alpha(L) = q_\alpha(F) = F^{\leftarrow}(\alpha).$$

Beispiel 3.2.6. Sei $L \sim N(\mu, \sigma^2)$. Es gilt

$$VaR_\alpha(L) = \mu + \sigma q_\alpha(\Phi) = \mu + \sigma \Phi^{-1}(\alpha),$$

wobei Φ die Verteilungsfunktion einer standardnormalverteilten Zufallsvariable ist, denn

$$\begin{aligned} VaR_\alpha(L) = \beta &\Rightarrow \mathbb{P}(L \leq VaR_\alpha(L)) = \mathbb{P}(L \leq \mu + \sigma \phi^{-1}(\alpha)) = \\ &\mathbb{P}\left(\underbrace{\frac{L - \mu}{\sigma}}_{=L' \sim N(0,1)} \leq \phi^{-1}(\alpha)\right) = \mathbb{P}(L' \leq \phi^{-1}(\alpha)) = \phi(\phi^{-1}(\alpha)) = \alpha. \end{aligned}$$

$$\Rightarrow \mathbb{P}(L \leq \mu + \sigma \phi^{-1}(\alpha)) = \mathbb{P}(L \leq VaR_\alpha(L)) \Rightarrow VaR_\alpha(L) = \mu + \sigma \phi^{-1}(\alpha),$$

da die Normalverteilung kontinuierlich ist.

◇

Beispiel 3.2.7. Ein Portfolio besteht aus $d = 5$ Stück einer Aktie A . Der heutige Preis von A ist $S_0 = 100$. Die täglichen Log-Renditen sind normalverteilt:

$X_1 = \ln \frac{S_1}{S_0}, X_2 = \ln \frac{S_2}{S_1}, \dots \stackrel{iid}{\sim} N(0, 0.01)$. Sei L_1 der 1-Tages PF-Verlust von heute auf morgen.

(a) Berechnen Sie $VaR_{0.99}(L_1)$.

$$L_{n+1} = -V_n \sum_{i=1}^d w_{n,i} (\exp(X_{n+1,i}) - 1) \quad X_{n+1,i} = \ln \frac{S_{n+1,i}}{S_{n,i}}$$

$$L_1 = -dS_0 (\exp(X_1) - 1) \quad X_1 = \ln \frac{S_1}{S_0}$$

Für den VaR gilt

$$\begin{aligned} VaR_u(L_1) = F_{L_1}^{-1}(u) = l &\implies F_{L_1}(l) = u = P(L_1 \leq l) \\ \iff P(-500(\exp(X_1) - 1) \leq l) = u &\iff P\left(\exp(X_1) - 1 \geq -\frac{l}{500}\right) = u \\ \iff P\left(\exp(X_1) \geq 1 - \frac{l}{500}\right) = u &\iff P\left(X_1 \geq \ln\left(1 - \frac{l}{500}\right)\right) = u \quad (X_1 \sim N(0, 0.01)) \\ \iff 1 - F_{X_1}\left(\ln\left(1 - \frac{l}{500}\right)\right) = u &\iff 1 - u = F_{X_1}\left(\ln\left(1 - \frac{l}{500}\right)\right) \\ \iff F_{X_1}^{-1}(1 - u) = \ln\left(1 - \frac{l}{500}\right) &\iff 500\left(1 - \exp(F_{X_1}^{-1}(1 - u))\right) = l \\ \iff l = 500\left(1 - \exp(-F_{X_1}^{-1}(u))\right) \end{aligned}$$

Es gilt $F_Z^{-1}(0.99) = 2.3$ und somit $F_{X_1}^{-1}(0.99) = 2.3\sqrt{0.01} = 0.23$. Für $u = 0.99$ folgt insgesamt

$$VaR_{0.99}(L_1) \approx 100.$$

(b) Berechnen Sie $VaR_{0.99}(L_{100})$ und $VaR_{0.99}(L_{100}^\Delta)$, wobei L_{100} der 100-Tage PF-Verlust über einen Zeithorizont von 100 Tagen ausgehend von heute ist. L_{100}^Δ ist die Linearisierung des obigen 100-Tage PF-Verlustes.

$L_{100} = -500\left(\exp(\tilde{X}_{100}) - 1\right)$ mit

$$\tilde{X}_{100} = \ln \frac{S_{100}}{S_0} = \ln \left(\frac{S_1}{S_0} \cdot \frac{S_2}{S_1} \cdots \frac{S_{100}}{S_{99}} \right) = \sum_{i=0}^{99} \ln \frac{S_{i+1}}{S_i} = \sum_{i=1}^{100} X_i \sim N(0, 1).$$

Es folgt somit

$$\begin{aligned} l = VaR_{0.99}(L_{100}) &= 500\left(1 - \exp(-F_{\tilde{X}_{100}}^{-1}(0.99))\right) = 500(1 - \exp(-2.3)) \\ VaR_{0.99}(L_{100}) &\approx 450. \end{aligned}$$

Für die Linearisierung gilt $e^x \approx 1 + x$ für $x \ll 1$, also $L_{100}^\Delta = -500\tilde{X}_{100}$.

$$VaR_{0.99}(L_{100}^\Delta) = 1.150 > 500!! \quad (\text{Anfangskapital})$$

Hier ist die Risikofaktoränderung über 100 Tage nicht infinitesimal klein, die Linearisierung versagt also. \diamond

3.2.3 Conditional Value at Risk ($CVaR_\alpha(L)$ - oder *Expected Shortfall* (**ES**))

Ein Nachteil von VaR ist, dass er keine Auskunft darüber gibt, wie groß der Verlust sein könnte, falls $L \geq VaR_\alpha(L)$.

Definition 3.2.8. Sei α ein vorgegebenes Konfidenzniveau und L eine kontinuierliche Verlustfunktion mit Verteilungsfunktion F_L .

$$CVaR_\alpha(L) := ES_\alpha(L) = \mathbb{E}(L|L \geq VaR_\alpha(L)).$$

Es gilt insbesondere

$$\begin{aligned} CVaR_\alpha(L) &= \mathbb{E}(L|L \geq VaR_\alpha(L)) = \frac{\mathbb{E}(LI_{[q_\alpha(L), \infty)}(L))}{\mathbb{P}(L \geq q_\alpha(L))} \\ &= \frac{1}{1-\alpha} \mathbb{E}(LI_{[q_\alpha(L), \infty)}) = \frac{1}{1-\alpha} \int_{q_\alpha(L)}^{+\infty} l dF_L(l) \end{aligned}$$

wobei I_A die Indikatorfunktion der Menge A ist, also $I_A(x) = \begin{cases} 1 & x \in A \\ 0 & x \notin A \end{cases}$.

Sei F_L die diskrete Verteilungsfunktion einer Verlustverteilung L und α ein vorgegebenes Konfidenzniveau. Der *verallgemeinerte CVaR* wird im diskreten Fall folgendermaßen definiert:

$$GCVaR_\alpha(L) := \frac{1}{1-\alpha} [\mathbb{E}(LI_{[q_\alpha(L), \infty)}) + q_\alpha(L) (1 - \alpha - P(L \geq q_\alpha(L)))] .$$

Lemma 3.2.9. Sei α ein vorgegebenes Konfidenzniveau und L eine kontinuierliche Verlustfunktion mit Verteilungsfunktion F_L . Es gilt

$$CVaR_\alpha(L) = \frac{1}{1-\alpha} \int_\alpha^1 VaR_p(L) dp.$$

Beweis. Für den CVaR gilt $CVaR_\alpha(L) = \frac{1}{1-\alpha} \int_{q_\alpha(L)}^\infty l dF_L(l)$.

Setze $y : [0, 1] \rightarrow \mathbb{R}$, $\alpha \mapsto VaR_\alpha(L) = q_\alpha(L)$ und $y^{\leftarrow} : \mathbb{R} \rightarrow [0, 1]$, $l \mapsto y^{\leftarrow}(l)$. Dann gilt $VaR_{y^{\leftarrow}(l)}(L) = l$.

Man substituiert nun $l = VaR_p(L) = y(p)$ im Integral, was zu $dF_L(l) = dF_L(VaR_p(L)) = dp$ führt, da $F_L(VaR_p(L)) = \mathbb{P}(L \leq VaR_p(L)) = p$ gilt. Die Integrationsgrenzen ändern sich zu

$$\begin{aligned} q_\alpha(L) &\rightsquigarrow \alpha \\ \infty &\rightsquigarrow 1. \end{aligned}$$

Zusammen ergibt das

$$CVaR_\alpha(L) = \frac{1}{1-\alpha} \int_\alpha^1 VaR_p(L) dp.$$

□

Beispiel 3.2.10. (a) Sei $L \sim Exp(\lambda)$. Bestimmen Sie $CVaR_\alpha(L)$.

$$CVaR_\alpha(L) = \frac{1}{1-\alpha} \int_\beta^\infty \lambda l \exp\{-\lambda l\} dl.$$

Mit $\beta = VaR_\alpha(L) = F_L^{-1}(\alpha)$ folgt $\alpha = F_L(\beta) = 1 - \exp\{-\lambda\beta\}$ und somit

$$\beta = -\frac{\ln(1-\alpha)}{\lambda}. \quad (*)$$

Eingesetzt liefert das (für $\lambda l = t$)

$$\begin{aligned} CVaR_\alpha(L) &= \frac{1}{1-\alpha} \int_{\beta\lambda}^\infty t \exp\{-t\} \frac{dt}{\lambda} = \frac{1}{\lambda} \frac{1}{1-\alpha} \left(-e^{-t}(1+t) \right) \Big|_{\beta\lambda}^\infty \\ &= \frac{(1+\beta\lambda)e^{-\beta\lambda}}{\lambda(1-\alpha)} \xrightarrow{(*)} CVaR_\alpha(L) = \lambda^{-1}(1 - \ln(1-\alpha)). \end{aligned}$$

(b) Die Verteilungsfunktion F_L der Verlustfunktion L sei folgendermaßen gegeben:
 $F_L(x) = 1 - (1 + \gamma x)^{-1/\gamma}$ für $x \geq 0$ und $\gamma \in (0, 1)$. Bestimmen Sie $CVaR_\alpha(L)$.

$$\begin{aligned} VaR_\alpha(L) &= \frac{1}{\gamma} \left((1-\alpha)^{-\gamma} - 1 \right) \\ CVaR_\alpha(L) &= \gamma^{-1} \left((1-\alpha)^{-\gamma} (1-\gamma)^{-1} - 1 \right). \end{aligned}$$

◇

Beispiel 3.2.11. Sei $L \sim N(0, 1)$. Seien ϕ und Φ die Verteilungsdichte bzw. -funktion von L . Es gilt $CVaR_\alpha(L) = \frac{\phi(\Phi^{-1}(\alpha))}{1-\alpha}$.

Sei $L' \sim N(\mu, \sigma^2)$. Zeigen Sie, dass $CVaR_\alpha(L') = \mu + \sigma \frac{\phi(\Phi^{-1}(\alpha))}{1-\alpha}$ gilt.

$$\begin{aligned} CVaR_\alpha(L) &= \frac{1}{1-\alpha} \int_{VaR_\alpha(L)}^\infty l d\Phi(l) = \frac{1}{1-\alpha} \int_{\Phi^{-1}(\alpha)}^\infty l \frac{1}{\sqrt{2\pi}} e^{-l^2/2} dl \\ &= \frac{1}{1-\alpha} \frac{1}{\sqrt{2\pi}} \left(-\exp\left(-\frac{l^2}{2}\right) \right) \Big|_{\Phi^{-1}(\alpha)}^\infty = \frac{\phi(\Phi^{-1}(\alpha))}{1-\alpha} \end{aligned}$$

◇

Beispiel 3.2.12. Sei die Verteilungsfunktion von L die Student-t-Verteilung mit $\nu > 1$ Freiheitsgraden. Die Dichtefunktion von L ist

$$g_\nu(x) = \frac{\Gamma((\nu+1)/2)}{\sqrt{\nu\pi}\Gamma(\nu/2)} \left(1 + \frac{x^2}{\nu} \right)^{-(\nu+1)/2}.$$

Zeigen Sie, dass $CVaR_\alpha(L) = \frac{g_\nu(t_\nu^{-1}(\alpha))}{1-\alpha} \left(\frac{\nu + (t_\nu^{-1}(\alpha))^2}{\nu-1} \right)$, wobei t_ν die Verteilungsfunktion von L ist.

◇

3.3 Methoden zur Berechnung von VaR und CVaR

Gegeben ist folgendes Setting:

- Portfoliowert: $V_m = f(t_m, Z_m)$
- Vektor von Risikofaktoren: Z_m
- Verlustfunktion: $L_{m+1} = l_{[m]}(X_{m+1})$
- X_{m+1} , der Vektor der Veränderungen der Risikofaktoren: $X_{m+1} = Z_{m+1} - Z_m$
- Beobachtungen (historische Daten): Z_{m-n+1}, \dots, Z_m .

Wie können diese historischen Daten nun zur Berechnung von $VaR(L_{m+1})$ bzw. $CVaR(L_{m+1})$ verwendet werden?

3.3.1 Der empirische VaR bzw. CVaR

Sei x_1, x_2, \dots, x_n eine Stichprobe der unabhängigen identischverteilten ZV X_1, X_2, \dots, X_n mit Verteilungsfunktion F . (Notation: Die ZV X_1, X_2, \dots, X_n sind i.i.d.)

Die empirische Verteilungsfunktion ist

$$F_n(x) = \frac{1}{n} \sum_{k=1}^n I_{[x_k, +\infty)}(x)$$

mit dem empirischen Quantil $q_\alpha(F_n) = \inf\{x \in \mathbb{R} : F_n(x) \geq \alpha\} = F_n^{\leftarrow}(\alpha)$.

Unter der Annahme $x_1 > x_2 > \dots > x_n$ gilt:

$$q_\alpha(F_n) = x_{[n(1-\alpha)]+1},$$

wobei $[y] = \sup\{n \in \mathbb{N} : n \leq y\}$ für jedes $y \in \mathbb{R}$, denn es gilt

$$\begin{aligned} q_\alpha(F_n) &= \inf\{x \in \mathbb{R} : F_n(x) \geq \alpha\} = x_{\max\{k \in \mathbb{N} : \frac{n-k}{n} \geq \alpha\}+1} \\ &= x_{\max\{k \in \mathbb{N} : 1 - \frac{k}{n} \geq \alpha\}+1} = x_{\max\{k \in \mathbb{N} : k \leq (1-\alpha)n\}+1} \\ &= x_{[n(1-\alpha)]+1}. \end{aligned}$$

$\hat{q}_\alpha(F) := q_\alpha(F_n)$ ist der empirische Schätzer des Quantils $q_\alpha(F)$.

Lemma 3.3.1. *Sei F eine streng monoton steigende Funktion. Dann gilt*

$$\lim_{n \rightarrow \infty} \hat{q}_\alpha(F) = q_\alpha(F), \quad \forall \alpha \in (0, 1),$$

d.h. der Schätzer $\hat{q}_\alpha(F)$ ist konsistent. Der empirische Schätzer des CVaR ist

$$\widehat{CVaR}_\alpha(F) = \frac{\sum_{k=1}^{[n(1-\alpha)]+1} x_k}{[(n(1-\alpha)]+1}.$$

3.3.2 Ein nicht-parametrisches Bootstrapping Verfahren zur Ermittlung von Konfidenzintervallen der Schätzer

Seien die ZV X_1, X_2, \dots, X_n i.i.d. mit Verteilungsfunktion F und sei x_1, x_2, \dots, x_n eine Stichprobe daraus.

Gesucht ist ein Schätzer eines von F abhängigen Parameters θ , z.B. $\theta = q_\alpha(F)$, und das dazugehörige Konfidenzintervall.

Sei $\hat{\theta}(x_1, \dots, x_n)$ ein Schätzer von θ , z.B. $\hat{\theta}(x_1, \dots, x_n) = x_{[(n-1)\alpha]+1, n}$, wobei $x_{1, n} > \dots > x_{n, n}$ die sortierte Stichprobe ist.

Das gesuchte Konfidenzintervall ist ein Intervall (a, b) , $a = a(x_1, \dots, x_n)$ und $b = b(x_1, \dots, x_n)$, sodass $\mathbb{P}(a < \theta < b) = p$, für ein vorgegebenes Konfidenzniveau p .

Fall I: F ist bekannt.

Durch Simulation von F werden N Stichproben (N groß) $\tilde{x}_1^{(i)}, \tilde{x}_2^{(i)}, \dots, \tilde{x}_n^{(i)}$, $1 \leq i \leq N$, erzeugt.

Sei $\tilde{\theta}_i = \hat{\theta}(\tilde{x}_1^{(i)}, \tilde{x}_2^{(i)}, \dots, \tilde{x}_n^{(i)})$, $1 \leq i \leq N$.

Die empirische Verteilungsfunktion von $\hat{\theta}(x_1, x_2, \dots, x_n)$ ist

$$F_N^{\hat{\theta}}(x) := \frac{1}{N} \sum_{i=1}^N I_{[\tilde{\theta}_i, \infty)}(x) \rightarrow F^{\hat{\theta}} \text{ für } N \rightarrow \infty.$$

Das gesuchte Konfidenzintervall ist

$$\left(q_{\frac{1-p}{2}}(F_N^{\hat{\theta}}), q_{\frac{1+p}{2}}(F_N^{\hat{\theta}}) \right),$$

denn für $a := q_{\frac{1-p}{2}}(F_N^{\hat{\theta}})$ und $b := q_{\frac{1+p}{2}}(F_N^{\hat{\theta}})$ gilt

$$\begin{aligned} \mathbb{P}(\theta > b) &= 1 - \frac{1+p}{2} = \frac{1-p}{2} \quad \text{und} \quad \mathbb{P}(\theta < a) = \frac{1-p}{2} \\ \Rightarrow \mathbb{P}(\theta < a \vee \theta > b) &= 1-p \Rightarrow \mathbb{P}(a < \theta < b) = p \end{aligned}$$

Fall II: F ist unbekannt.

Zur Erinnerung: Die empirische Verteilungsfunktion von X_i , $1 \leq i \leq n$, ist

$$F_n(x) = \frac{1}{n} \sum_{i=1}^n I_{[x_i, \infty)}(x).$$

Wenn n groß ist, dann gilt $F_n \approx F$.

Wir können Stichproben aus F_n nehmen, indem wir n Elemente aus $\{x_1, x_2, \dots, x_n\}$ mit Zurücklegen ziehen.

Angenommen es werden N solche Stichproben gezogen: $x_1^{*(i)}, x_2^{*(i)}, \dots, x_n^{*(i)}$, $1 \leq i \leq N$.

Es soll $\theta_i^* = \hat{\theta}(x_1^{*(i)}, x_2^{*(i)}, \dots, x_n^{*(i)})$ berechnet werden.

Die empirische Verteilungsfunktion $F_N^{\theta^*}(x) = \frac{1}{N} \sum_{i=1}^N I_{[\theta_i^*, \infty)}(x)$ approximiert die Verteilungsfunktion $F^{\hat{\theta}}$ von $\hat{\theta}(X_1, X_2, \dots, X_n)$ für $N \rightarrow \infty$.

Für das Konfidenzintervall (a, b) mit $a = q_{(1-p)/2}(F_N^{\theta^*})$, $b = q_{(1+p)/2}(F_N^{\theta^*})$ gilt

$$a = \theta_{[N(1+p)/2]+1, N}^* \quad b = \theta_{[N(1-p)/2]+1, N}^*$$

wobei $\theta_{1, N}^* \geq \theta_{2, N}^* \geq \dots \geq \theta_{N, N}^*$ durch Sortierung von $\theta_1^*, \theta_2^*, \dots, \theta_N^*$ entsteht.

3.3.3 Zusammenfassung des nicht-parametrischen Bootstrapping-Verfahrens zur Berechnung von Konfidenzintervallen

Gegeben: Stichprobe x_1, x_2, \dots, x_n der i.i.d. ZV. X_1, X_2, \dots, X_n mit Verteilungsfunktion F und ein Schätzer $\hat{\theta}(x_1, x_2, \dots, x_n)$ eines unbekanntes Parameters $\theta(F)$.

Gesucht: Ein Konfidenzintervall I_p für θ mit vorgegebenem Konfidenzniveau p .

- Bilde N neue Stichproben $x_1^{*(i)}, x_2^{*(i)}, \dots, x_n^{*(i)}$, $1 \leq i \leq N$, durch ziehen mit Zurücklegen aus $\{x_1, x_2, \dots, x_n\}$.
- Berechne $\theta_i^* = \hat{\theta}(x_1^{*(i)}, x_2^{*(i)}, \dots, x_n^{*(i)})$.
- $I_p = \left(\theta_{[N(1+p)/2]+1, N}^*, \theta_{[N(1-p)/2]+1, N}^* \right)$, wobei $\theta_{1, N}^* \geq \theta_{2, N}^* \geq \dots \geq \theta_{N, N}^*$ durch Sortierung von $\theta_1^*, \theta_2^*, \dots, \theta_N^*$ entsteht.

3.3.4 Eine approximative Lösung ohne Bootstrapping

Gegeben: Eine Stichprobe x_1, x_2, \dots, x_n von ZV X_i , $1 \leq i \leq n$, i.i.d. mit unbekannter kontinuierlicher Verteilungsfunktion F .

Gesucht: Ein Konfidenzintervall (a, b) für $q_\alpha(F)$, $a = a(x_1, x_2, \dots, x_n)$, $b = b(x_1, x_2, \dots, x_n)$, sodass

$$\mathbb{P}(a < q_\alpha(F) < b) = p \text{ und } \mathbb{P}(a \geq q_\alpha(F)) = \mathbb{P}(b \leq q_\alpha(F)) = (1 - p)/2.$$

Wir suchen $i > j$, $i, j \in \{1, 2, \dots, n\}$, und das kleinste $p' > p$, sodass

$$\mathbb{P}(x_{i,n} < q_\alpha(F) < x_{j,n}) = p' \quad (*)$$

und

$$\mathbb{P}(x_{i,n} \geq q_\alpha(F)) \leq (1 - p)/2 \text{ und } \mathbb{P}(x_{j,n} \leq q_\alpha(F)) \leq (1 - p)/2 \quad (**),$$

wobei $x_{1,n} \geq x_{2,n} \geq \dots \geq x_{n,n}$ durch Sortierung von x_1, x_2, \dots, x_n entsteht.

Sei $Y_\alpha = \#\{x_k : x_k > q_\alpha(F)\}$.

Es gilt:

$$\begin{aligned} \mathbb{P}(x_{j,n} \leq q_\alpha(F)) &= \mathbb{P}(x_{j,n} < q_\alpha(F)) = \mathbb{P}(Y_\alpha \leq j - 1) \\ \mathbb{P}(x_{i,n} \geq q_\alpha(F)) &= \mathbb{P}(x_{i,n} > q_\alpha(F)) = 1 - \mathbb{P}(Y_\alpha \leq i - 1) = \mathbb{P}(Y_\alpha \geq i) \end{aligned}$$

$Y_\alpha \sim B(n, 1 - \alpha)$. Berechne $\mathbb{P}(x_{j,n} \leq q_\alpha(F))$ und $\mathbb{P}(x_{i,n} \geq q_\alpha(F))$ für unterschiedliche i und j solange, bis $i, j \in \{1, 2, \dots, n\}$, $i > j$, gefunden werden, die (*) und (**) erfüllen.

Beispiel 3.3.2. Für $n = 10$ und $\alpha = 0.8$ berechne das Konfidenzintervall für $q_{0.8}(F)$ mit Konfidenzniveau $p' \geq 0.75$.

$Y_{0.8} \sim B(10, 0.2) \Rightarrow$

$$\begin{aligned} \mathbb{P}(Y_{0.8} = 0) &= 0.1074 \\ \mathbb{P}(Y_{0.8} = 1) &= 0.2684 \\ \mathbb{P}(Y_{0.8} = 2) &= 0.3020 \\ &\vdots \\ \mathbb{P}(Y_{0.8} = 10) &= 1.02 \cdot 10^{-7} \end{aligned}$$

$$\mathbb{P}(x_{1,10} < q_{0.8}(F)) = \mathbb{P}(Y_{0.8} \leq 0) = 0.1074 \leq \frac{1-p}{2} = 0.125$$

$$\mathbb{P}(x_{4,10} > q_{0.8}(F)) = 1 - \mathbb{P}(Y_{0.8} \leq 3) = 0.12 \leq 0.125$$

$\Rightarrow [x_{4,10}, x_{1,10}]$ und $\mathbb{P}(x_{4,10} \leq q_{0.8}(F) \leq x_{1,10}) \approx 1 - 0.12 - 0.1074 > 0.75$ ◇

3.3.5 Historische Simulation

Seien x_{m-n+1}, \dots, x_m historische Beobachtungen der Veränderungen der Risikofaktoren X_{m-n+1}, \dots, X_m gegeben.

Annahme: Die historischen Verluste sind i.i.d.

Die historischen Verlustwerte $l_k = l_{[m]}(x_{m-k+1})$, $k = 1, 2, \dots, n$, stellen eine Stichprobe der Verlustverteilung dar. Der empirische VaR ist demnach

$$\widehat{VaR} = q_\alpha(\hat{F}_n^L) = l_{[n(1-\alpha)]+1, n}$$

und der empirische CVaR

$$\widehat{CVaR} = \frac{\sum_{i=1}^{[n(1-\alpha)]+1} l_{i,n}}{[n(1-\alpha)]+1},$$

wobei $l_{1,n} \geq l_{2,n} \geq \dots \geq l_{n,n}$ durch die Sortierung von l_i , $1 \leq i \leq n$, entsteht.

Der VaR und CVaR des aggregierten Verlustes über mehrere Tage, zB. 10 Tage, kann mit Hilfe der Verlustwerte

$$l_k^{(10)} = l_{[m]} \left(\sum_{j=1}^{10} x_{m-n+10(k-1)+j} \right) \quad k = 1, \dots, [n/10]$$

geschätzt werden.

Vorteile:

- einfache Implementierung
- berücksichtigt die Abhängigkeitsstruktur zwischen den Komponenten des Vektors der Veränderungen der Risikofaktoren $X_{m-k} = (X_{m-k,1}, \dots, X_{m-k,d})$

Nachteile:

- es sind sehr viele historische Daten notwendig um zuverlässige Schätzer zu bekommen
- Schätzung impliziert, dass der geschätzte Verlust nicht größer als bereits historisch realisierte Verluste sein kann

3.3.6 Varianz-Kovarianz Methode

Die Grundidee hier ist die Verwendung der linearisierten Verlustfunktion

$$L_{m+1}^\Delta = l_m^\Delta(X_{m+1}) = -V \sum_{i=1}^d w_i X_{m+1,i} = -V w^\top X_{m+1},$$

wobei $V := V_m$, $w_i := w_{m,i}$, $w = (w_1, \dots, w_d)^T$, $X_{m+1} = (X_{m+1,1}, X_{m+1,2}, \dots, X_{m+1,d})^T$.
 Unter der Annahme $X_{m+1} \sim N_d(\mu, \Sigma)$ folgt

$$-Vw^\top X_{m+1} \sim N(-Vw^\top \mu, V^2 w^\top \Sigma w).$$

Seien x_{m-n+1}, \dots, x_m historische Beobachtungen der Veränderungen der Risikofaktoren, mit der Annahme sie seien i.i.d. Der Schätzer für μ_i ist

$$\hat{\mu}_i = \frac{1}{n} \sum_{k=1}^n x_{m-k+1,i} \quad i = 1, 2, \dots, d$$

und der Schätzer $\hat{\Sigma} = (\hat{\sigma}_{ij})$ für $\Sigma = (\sigma_{ij})$

$$\hat{\sigma}_{ij} = \frac{1}{n-1} \sum_{k=1}^n (x_{m-k+1,i} - \hat{\mu}_i)(x_{m-k+1,j} - \hat{\mu}_j) \quad i, j = 1, 2, \dots, d.$$

Insgesamt erhält man nun den Schätzer für VaR:

$$\widehat{\text{VaR}}_\alpha(L_{m+1}) = -Vw^\top \hat{\mu} + V\sqrt{w^\top \hat{\Sigma} w} \phi^{-1}(\alpha)$$

(siehe Bsp. (3.2.6)).

Vorteile:

- analytische Lösung
- einfache Implementierung
- keine Simulationen notwendig

Nachteile:

- Linearisierung nicht immer adäquat, nur für einen kurzen Zeithorizont gerechtfertigt (siehe Übung (3.2.7)).
- Annahme der Normalverteilung könnte zur Unterschätzung des Risikos führen und sollte begründet werden (zB. anhand von historischen Daten).

3.3.7 Monte-Carlo Verfahren

Historische Beobachtungen der Risikofaktoren und deren Veränderungen.

↓

Annahme über ein parametrisches Modell für die Verteilungsfunktion der Veränderung der Risikofaktoren X_{m+1} ; zB. gemeinsame Verteilungsfunktion F und Unabhängigkeit von X_{m-n+1}, \dots, X_m .

↓

Schätzung der charakteristischen Parameter dieser Verteilungsfunktion.

Nehme N Stichproben $\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_N$ aus F (N muss groß genug sein).
 Sei $l_k = l_{[m]}(\tilde{x}_k)$, $1 \leq k \leq N$. Die empirische Verteilungsfunktion des Verlustes L_{m+1} ist

$$\hat{F}_N^{L_{m+1}}(x) = \frac{1}{N} \sum_{k=1}^N I_{[l_k, \infty)}(x).$$

Der Schätzer für VaR ist also

$$\widehat{VaR}(L_{m+1}) = q_\alpha \left(\hat{F}_N^{L_{m+1}} \right) = l_{[N(1-\alpha)]+1, N},$$

wobei $l_{1, N} \geq l_{2, N} \geq \dots \geq l_{N, N}$ und der für CVaR

$$\widehat{CVaR}(L_{m+1}) = \frac{\sum_{k=1}^{[N(1-\alpha)]+1} l_{k, N}}{[N(1-\alpha)]+1}.$$

Vorteile:

- Sehr flexibel, es kann jedes Modell verwendet werden, aus dem simuliert werden kann
- Berücksichtigung von zeitlichen Abhängigkeiten zwischen den Veränderungen der Risikofaktoren möglich, etwa durch Verwendung von Zeitreihen

Nachteile:

- Rechenintensiv; große Anzahl von Simulationen notwendig um gute Schätzwerte zu bekommen.

Beispiel 3.3.3. Das Portfolio besteht aus einem Stück einer Aktie mit Preis S_t zum Zeitpunkt t .

Die Veränderungen der Risikofaktoren

$$X_{k+1} = \ln(S_{t_{k+1}}) - \ln(S_{t_k})$$

sind i.i.d. mit Verteilungsfunktion F_θ , wobei θ ein unbekannter Parameter ist.

θ kann mit Hilfe von historischen Daten geschätzt werden (zB. mittels ML Verfahren).

Sei der heutige Preis $S_{t_k} = S$.

Man zeige: Der VaR des Portfolios über $[t_k, t_{k+1}]$ ist folgendermaßen gegeben:

$$VaR_\alpha(L_{t_{k+1}}) = S (1 - \exp\{F_\theta^{\leftarrow}(1 - \alpha)\}).$$

Wenn F_θ kompliziert ist, dann kann die analytische Berechnung von CVaR schwierig sein. Eine Alternative wäre hier die Monte-Carlo Simulation. \diamond

Beispiel 3.3.4. Sei das Portfolio und die Veränderungen der Risikofaktoren X_{k+1} wie im obigen Beispiel.

Ein populäres Modell für die Log-Rendite einer Aktie ist GARCH(1,1) (siehe zB. Alexander 2002):

$$X_{k+1} = \sigma_{k+1} Z_{k+1} \tag{3.1}$$

$$\sigma_{k+1}^2 = a_0 + a_1 X_k^2 + b_1 \sigma_k^2 \tag{3.2}$$

wobei Z_k , $k \in \mathbb{N}$ i.i.d. standard-normalverteilt sind. a_0, a_1 und b_1 sind Parameter, die geschätzt werden.

Es ist einfach aus diesem Modell zu simulieren. Analytische Berechnungen des VaR bzw. CVaR für ein Zeitintervall bestehend aus mehreren Perioden sind hingegen schwierig!

Ausprobieren! \diamond

Kapitel 4

Extremwerttheorie

In diesem Kapitel werden wir oft dieselbe Notation für die Verteilung einer ZV und ihre Verteilungsfunktion verwenden!

Bemerkungen:

- $f(x) \sim g(x)$ für $x \rightarrow \infty$ bedeutet $\lim_{x \rightarrow \infty} f(x)/g(x) = 1$
- $\bar{F} := 1 - F$ („right Tail“ einer univariaten Verteilungsfunktion F).

Bezeichnungskonvention: Man sagt eine ZV X hat „fat Tails“ oder ist „heavy tailed“ (h.t.) wenn

$$\lim_{x \rightarrow \infty} \frac{\bar{F}(x)}{e^{-\lambda x}} = \infty \quad \forall \lambda > 0.$$

Auch eine ZV X , für die $\exists k \in \mathbb{N}$, sodass $E(X^k) = \infty$, wird oft *heavy tailed* genannt.

Definition 4.0.5. Eine meßbare Funktion $h: (0, \infty) \rightarrow (0, \infty)$ besitzt eine reguläre Variation mit Index $\rho \in \mathbb{R}$ um ∞ , wenn

$$\lim_{t \rightarrow \infty} \frac{h(tx)}{h(t)} = x^\rho \quad \forall x > 0 \tag{4.1}$$

Bezeichnung: $h \in RV_\rho$.

- Wenn $\rho = 0$, dann heißt es „ h variiert langsam um ∞ “.
- Wenn $h \in RV_\rho$, dann gilt $h(x)/x^\rho \in RV_0$.
- Falls $h \in RV_\rho$, dann $\exists L \in RV_0$ sodass $h(x) = L(x)x^\rho$ ($L(x) = h(x)/x^\rho$).
- Falls $\rho < 0$, dann ist die Konvergenz in (4.1) gleichmäßig in jedem Intervall (b, ∞) für $b > 0$.

Beispiel 4.0.6. 1. Zeigen Sie, dass für die untenstehenden Funktionen L , $L \in RV_0$ gilt.

- (a) $\lim_{x \rightarrow \infty} L(x) = c \in (0, \infty)$
- (b) $L(x) = \ln(1 + x)$
- (c) $L(x) = \ln(1 + \ln(1 + x))$

2. Gilt $f \in RV_0$ für $f(x) = 3 + \sin x$, $f(x) = \ln(e + x) + \sin x$? ◇

Eine Funktion $L \in RV_0$ kann unendlich stark um ∞ variieren im Sinne, dass

$$\liminf_{x \rightarrow \infty} L(x) = 0 \text{ und } \limsup_{x \rightarrow \infty} L(x) = \infty.$$

Ein Beispiel dafür ist $L(x) = \exp\{(\ln(1+x))^2 \cos((\ln(1+x))^{1/2})\}$.

Beispiel 4.0.7. Sei $F(x) = 1 - x^{-\alpha}$, für $x > 1$ und $\alpha > 0$. Dann gilt $\bar{F}(tx)/\bar{F}(t) = x^{-\alpha}$ für $t > 0$. D.h. $\bar{F} \in RV_{-\alpha}$. ◇

Beispiel 4.0.8. Sei $F(x) = \exp\{-x^{-\alpha}\}$ für $x > 0$ und $\alpha > 0$. Es gilt $\lim_{x \rightarrow \infty} \bar{F}(x)/x^{-\alpha} = 1$. Daraus folgt $\bar{F} \in RV_{-\alpha}$, denn

$$\begin{aligned} 1 \leq \frac{\bar{F}(x)}{x^{-\alpha}} &= x^\alpha(1 - F(x)) = x^\alpha \left(1 - \sum_{k=0}^{\infty} \frac{(-x^{-\alpha})^k}{k!} \right) \\ &= -x^\alpha \sum_{k=1}^{\infty} \frac{(-x^{-\alpha})^k}{k!} = 1 + \sum_{k=2}^{\infty} \frac{-x^\alpha (-x^{-\alpha})^k}{k!} \\ &= 1 + \sum_{k=2}^{\infty} \frac{(-x^{-\alpha})^{k-1}}{k!} = 1 + \sum_{k=1}^{\infty} \frac{(-x^{-\alpha})^k}{(k+1)!} \\ &\leq 1 + \sum_{k=1}^{\infty} \frac{(-x^{-\alpha})^k}{k!} = \sum_{k=0}^{\infty} \frac{(-x^{-\alpha})^k}{k!} \\ &= \exp(-x^{-\alpha}) \xrightarrow{x \rightarrow \infty} 1. \end{aligned}$$

Es bleibt somit noch zu zeigen, dass $\lim_{t \rightarrow \infty} \frac{\bar{F}(tx)}{\bar{F}(t)} = x^{-\alpha}$. Das gilt aber, denn

$$\lim_{t \rightarrow \infty} \frac{\bar{F}(tx)/(tx)^{-\alpha}}{\bar{F}(t)/t^{-\alpha}} \cdot \frac{(tx)^{-\alpha}}{t^{-\alpha}} = x^{-\alpha}.$$

◇

Definition 4.0.9. Sei $X > 0$ eine ZV mit Verteilungsfunktion F . Man sagt X hat eine *reguläre Variation*, wenn $\bar{F} \in RV_{-\alpha}$ für ein $\alpha > 0$.

Proposition 4.0.10. Sei $X > 0$ eine ZV mit Verteilungsfunktion F , sodass $\bar{F} \in RV_{-\alpha}$ für ein $\alpha > 0$. Es gilt dann $E(X^\beta) < \infty$ für $\beta < \alpha$ und $E(X^\beta) = \infty$ für $\beta > \alpha$.

Die Umkehrung gilt nicht!

Proposition 4.0.11. (i) Eine meßbare Funktion $h: (0, \infty) \rightarrow (0, \infty)$ hat eine reguläre Variation um $+\infty$, wenn eine Funktion g existiert, sodass

$$\lim_{t \rightarrow +\infty} h(tx)/h(t) = g(x), \quad \forall x > 0.$$

In diesem Fall gilt $g(x) = x^\rho$ für ein $\rho \in \mathbb{R}$ und dann $h \in RV_\rho$.

(ii) Eine monotone Funktion $h: (0, +\infty) \rightarrow (0, +\infty)$ hat eine reguläre Variation um $+\infty$, wenn eine Folge (a_n) von positiven Zahlen und eine Funktion ξ existieren, sodass

$$\lim_{n \rightarrow +\infty} nh(a_n x) = \xi(x), \quad \forall x > 0.$$

In diesem Fall gilt $\xi(x)/\xi(1) = x^\rho$ für ein $\rho \in \mathbb{R}$ und dann $h \in RV_\rho$.

Beweis. Siehe Resnick 1987. □

Beispiel 4.0.12. Seien X_1 und X_2 zwei nichtnegative i.i.d. ZV mit Verteilungsfunktion $F, \bar{F} \in RV_{-\alpha}$ für ein $\alpha > 0$.

Annahme: X_1 (X_2) gibt den Verlust eines Portfolios bestehend aus 1 Stück der Aktie A_1 (A_2) an. Es wird weiters angenommen, dass A_1 und A_2 gleich viel kosten.

Ein Investor hat 2 Stück der Aktie A_1 gekauft. Die Wahrscheinlichkeit, dass sein Verlust größer als l ist, ist durch $\mathbb{P}(2X_1 > l)$ gegeben. Kann der Investor die Verlustwahrscheinlichkeit verringern, indem er auf ein diversifiziertes Portfolio bestehend aus einem Stück der Aktie A_1 und einem Stück der Aktie A_2 wechselt?

Um diese Frage zu beantworten, wird $\frac{\mathbb{P}(X_1 + X_2 > l)}{\mathbb{P}(2X_1 > l)}$ studiert.

Sei $\epsilon > 0$ beliebig:

$$\begin{aligned} \mathbb{P}(X_1 + X_2 > l) &\leq 2\mathbb{P}(X_1 < \epsilon l \wedge X_2 > (1 - \epsilon)l) + \mathbb{P}(X_1 \geq \epsilon l \wedge X_2 \geq \epsilon l) \\ &\leq 2\mathbb{P}(X_1 \geq (1 - \epsilon)l) + \mathbb{P}^2(X_1 \geq \epsilon l) \\ \mathbb{P}(X_1 + X_2 \geq l) &\geq \mathbb{P}(X_1 \geq l \vee X_2 \geq l) = 2\mathbb{P}(X_1 \geq l) - \mathbb{P}(X_1 \geq l \wedge X_2 \geq l) \\ &= 2\mathbb{P}(X_1 \geq l) - \mathbb{P}^2(X_1 \geq l) \end{aligned}$$

Es gilt also

$$\underbrace{\frac{2\mathbb{P}(X_1 > l) - \mathbb{P}^2(X_1 > l)}{\mathbb{P}(2X_1 > l)}}_{h(\alpha, \epsilon, l)} \leq \frac{\mathbb{P}(X_1 + X_2 > l)}{\mathbb{P}(2X_1 > l)} \leq \underbrace{\frac{2\mathbb{P}(X_1 > (1 - \epsilon)l) + \mathbb{P}^2(X_1 > (\epsilon l))}{\mathbb{P}(2X_1 > l)}}_{g(\alpha, \epsilon, l)}$$

mit

$$\begin{aligned}
\lim_{l \rightarrow \infty} h(\alpha, \epsilon, l) &= \lim_{l \rightarrow \infty} \frac{2\mathbb{P}(X_1 > l) - \mathbb{P}^2(X_1 > l)}{\mathbb{P}(X_1 > l/2)} = \lim_{l \rightarrow \infty} \frac{2\mathbb{P}(X_1 > l) \left(1 - \frac{\mathbb{P}(X_1 > l)}{2}\right)}{\mathbb{P}(X_1 > l/2)} \\
&= \lim_{l \rightarrow \infty} \frac{2\mathbb{P}(X_1 > 2 \cdot \frac{l}{2})}{\mathbb{P}(X_1 > l/2)} = 2 \cdot 2^{-\alpha} = 2^{1-\alpha} \\
\lim_{l \rightarrow \infty} g(\alpha, \epsilon, l) &= \lim_{l \rightarrow \infty} \frac{2\mathbb{P}(X_1 > (1-\epsilon)l) \left(1 + \frac{\mathbb{P}^2(X_1 > (\epsilon)l)}{2\mathbb{P}(X_1 > (1-\epsilon)l)}\right)}{\mathbb{P}(2X_1 > l)} \\
&= \lim_{l \rightarrow \infty} \underbrace{\frac{2\mathbb{P}(X_1 > 2(1-\epsilon)\frac{l}{2})}{\mathbb{P}(X_1 > l/2)}}_{\rightarrow 2(2(1-\epsilon))^{-\alpha}} \cdot \underbrace{\left(1 + \frac{\mathbb{P}^2(X_1 > (\epsilon)l)}{2\mathbb{P}(X_1 > (1-\epsilon)l)}\right)}_{\substack{\rightarrow 0 \\ \rightarrow 1}} = 2^{1-\alpha}(1-\epsilon)^{-\alpha}.
\end{aligned}$$

Insgesamt lautet die Ungleichungskette wie folgt:

$$2^{1-\alpha} \leq \lim_{l \rightarrow \infty} \frac{\mathbb{P}(X_1 + X_2 > l)}{\mathbb{P}(2X_1 > l)} \leq 2^{1-\alpha}(1-\epsilon)^{-\alpha} \quad \forall \epsilon \in (0, 1/2).$$

Für $\epsilon \rightarrow 0$ gilt $2^{1-\alpha}(1-\epsilon)^{-\alpha} \rightarrow 2^{1-\alpha}$ und demnach gilt

$$\lim_{l \rightarrow \infty} \frac{\mathbb{P}(X_1 + X_2 > l)}{\mathbb{P}(2X_1 > l)} = 2^{1-\alpha} = \begin{cases} > 1 & \alpha < 1 \\ < 1 & \alpha > 1 \end{cases}.$$

Wenn $\alpha < 1$ ist eine Diversifikation nicht sinnvoll, für $\alpha > 1$ jedoch schon! ◇

Beispiel 4.0.13. Seien X und Y zwei ZV, die die Verluste zweier Geschäftslinien einer Versicherungsgesellschaft darstellen (Brand- bzw. Autoversicherung).

Sei F die Verteilungsfunktion von X für die $\bar{F} \in RV_{-\alpha}$, $\alpha > 0$, gilt. Es gelte weiters $E(Y^k) < \infty$, $\forall k > 0$.

Die Versicherungsgesellschaft möchte $\lim_{x \rightarrow \infty} \mathbb{P}(X > x | X + Y > x)$, d.h. die Wahrscheinlichkeit eines großen Verlustes bei der Brandversicherung, berechnen, vorausgesetzt es gibt einen großen Gesamtverlust der zwei Linien.

Wir betrachten also die bedingte Wahrscheinlichkeit $\mathbb{P}(X > x | X + Y > x)$.

$$\begin{aligned}
\mathbb{P}(X + Y > x) &= \mathbb{P}(X + Y > x \wedge X > (1-\epsilon)x) + \mathbb{P}(X + Y > x \wedge X \leq (1-\epsilon)x) \\
&\leq \mathbb{P}(X + Y > x \wedge X > (1-\epsilon)x) + \mathbb{P}(X + Y > x \wedge Y > \epsilon x) \\
&\leq \mathbb{P}(X > (1-\epsilon)x) + \mathbb{P}(Y > \epsilon x).
\end{aligned}$$

Es gilt

$$1 \leq \frac{\mathbb{P}(X + Y > x)}{\mathbb{P}(X > x)} \leq \frac{\mathbb{P}(X > (1-\epsilon)x) + \mathbb{P}(Y > \epsilon x)}{\mathbb{P}(X > x)}$$

mit

$$\frac{\mathbb{P}(X > (1-\epsilon)x)}{\mathbb{P}(X > x)} \xrightarrow{x \rightarrow \infty} (1-\epsilon)^{-\alpha},$$

$$\mathbb{P}(Y > \epsilon x) = \mathbb{P}(Y^{2\alpha} > (\epsilon x)^{2\alpha}) = \mathbb{P}\left(Y^{2\alpha} > \frac{\mathbb{E}(Y^{2\alpha})}{\mathbb{E}(Y^{2\alpha})}(\epsilon x)^{2\alpha}\right) \stackrel{Markov}{\leq} \frac{\mathbb{E}(Y^{2\alpha})}{(\epsilon x)^{2\alpha}}$$

und

$$\frac{\mathbb{P}(Y > \epsilon x)}{\mathbb{P}(X > x)} \leq \frac{\mathbb{E}(Y^{2\alpha})}{(\epsilon x)^{2\alpha} \mathbb{P}(X > x)} = \frac{\mathbb{E}(Y^{2\alpha})}{(\epsilon x)^{2\alpha} x^{-\alpha} h(x)} = \frac{\mathbb{E}(Y^{2\alpha})}{\underbrace{\epsilon^{2\alpha} x^\alpha h(x)}_{\xrightarrow{x \rightarrow \infty} \infty}} \xrightarrow{x \rightarrow \infty} 0$$

wobei $h \in RV_0$.

Für $x \rightarrow \infty$ folgt

$$1 \leq \frac{\mathbb{P}(X + Y > x)}{\mathbb{P}(X > x)} \leq \underbrace{(1 - \epsilon)^{-\alpha}}_{\xrightarrow{\epsilon \rightarrow 0} 1}$$

und insgesamt für $\epsilon \rightarrow 0$

$$\mathbb{P}(X > x | X + Y > x) \rightarrow 1.$$

◇

4.1 Klassische Extremwerttheorie

4.1.1 Grenzverteilungen von normierten und zentrierten Summen

Seien (X_k) , $k \in \mathbb{N}$, nicht degenerierte i.i.d. ZV mit Verteilungsfunktion F . Für $n \geq 1$ definiere $S_n := \sum_{i=1}^n X_i$.

Frage: Welche sind die möglichen (nicht degenerierten) Grenzverteilungen von normierten und zentrierten S_n ?

Zunächst wird die Grenzverteilung von S_n untersucht:

Für welche nicht degenerierten ZV Z gibt es zwei Zahlenfolgen $a_n > 0$ und $b_n \in \mathbb{R}$, $n \in \mathbb{N}$, sodass $a_n^{-1}(S_n - b_n) \rightarrow Z$ in Verteilung für $n \rightarrow \infty$.

Notation: $\lim_{n \rightarrow \infty} a_n^{-1}(S_n - b_n) = Z$

Definition 4.1.1. Eine ZV X heißt *stabil*, wenn für alle $c_1, c_2 \in \mathbb{R}_+$ und die i.i.d. Kopien X_1 und X_2 von X die Konstanten $a(c_1, c_2) \in \mathbb{R}$ und $b(c_1, c_2) \in \mathbb{R}$ existieren, sodass $c_1 X_1 + c_2 X_2$ und $a(c_1, c_2)X + b(c_1, c_2)$ idente Verteilungsfunktionen haben.

Notation: $c_1 X_1 + c_2 X_2 \stackrel{d}{=} a(c_1, c_2)X + b(c_1, c_2)$

Satz 4.1. Die Klasse von stabilen Verteilungen stimmt mit der Klasse der nicht degenerierten Grenzverteilungen passend normierter und zentrierter Summen von i.i.d. ZV überein.

Beweis. zB. in Rényi, 1962. □

Satz 4.2. Die charakteristische Funktion einer stabilen Verteilung X ist folgendermaßen gegeben:

$$\varphi_X(t) = \mathbb{E}[\exp\{iXt\}] = \exp\{i\gamma t - c|t|^\alpha(1 + i\beta \operatorname{sgn}(t)z(t, \alpha))\} \quad (4.2)$$

wobei $\gamma \in \mathbb{R}$, $c > 0$, $\alpha \in (0, 2]$, $\beta \in [-1, 1]$ und

$$z(t, \alpha) = \begin{cases} \tan(\frac{\pi\alpha}{2}) & \text{wenn } \alpha \neq 1 \\ -\frac{2}{\pi} \ln |t| & \text{wenn } \alpha = 1 \end{cases} .$$

Beweis. Lévy 1954, Gnedenko und Kolmogoroff 1960. □

Definition 4.1.2. Der Parameter α in (4.2) heißt *charakteristischer Exponent*, die dazugehörige Verteilung heißt α -stabil und ihre Verteilungsfunktion wird mit G_α bezeichnet.

Definition 4.1.3. Seien X_k , $k \in \mathbb{N}$, i.i.d. ZV mit Verteilungsfunktion F . Angenommen es existieren die Zahlenfolgen $a_n > 0$ und $b_n \in \mathbb{R}$, $n \in \mathbb{N}$, sodass $\lim_{n \rightarrow \infty} a_n^{-1}(S_n - b_n) = G_\alpha$, für eine α -stabile Verteilung G_α , dann heißt es „ F gehört dem Anziehungsgebiet von G_α “.

Notation: $F \in DA(G_\alpha)$.

Definition 4.1.4. Die ZV Z und \tilde{Z} sind vom selben Typus wenn $\sigma > 0$ und $\mu \in \mathbb{R}$ existieren, sodass $\tilde{Z} \stackrel{d}{=} (Z - \mu)/\sigma$, d.h. $\tilde{F}(x) = F(\mu + \sigma x) \forall x \in \mathbb{R}$, wobei F und \tilde{F} die Verteilungsfunktionen von Z bzw. \tilde{Z} sind.

Satz 4.3. (*Convergence to types theorem*)

Seien $Z, \tilde{Z}, Y_n, n \geq 1$, ZV. Weder Z noch \tilde{Z} sind fast sicher konstant. Seien $a_n, \tilde{a}_n, b_n, \tilde{b}_n \in \mathbb{R}$, $n \in \mathbb{N}$, Zahlenfolgen mit $a_n, \tilde{a}_n > 0$.

(i) Wenn

$$\lim_{n \rightarrow \infty} a_n^{-1}(Y_n - b_n) = Z \text{ und } \lim_{n \rightarrow \infty} \tilde{a}_n^{-1}(Y_n - \tilde{b}_n) = \tilde{Z} \quad (4.3)$$

dann existieren $A > 0$ und $B \in \mathbb{R}$, sodass

$$\lim_{n \rightarrow \infty} \frac{\tilde{a}_n}{a_n} = A \text{ und } \lim_{n \rightarrow \infty} \frac{\tilde{b}_n - b_n}{a_n} = B \quad (4.4)$$

und

$$\tilde{Z} \stackrel{d}{=} (Z - B)/A. \quad (4.5)$$

(ii) Angenommen (4.4) gilt. Dann impliziert jede der zwei Relationen in (4.3) die andere, und (4.5) gilt auch.

Beweis. Siehe Resnick 1987, Prop. 0.2, Seite 7. □

Anmerkung: $X \sim G_2 \iff \varphi_X(t) = \exp\{i\gamma t - \frac{1}{2}t^2(2c)\} \iff X \sim N(\gamma, 2c)$.

Beispiel 4.1.5. Zeigen Sie, dass $F \in DA(G_2) \iff F \in DA(\phi)$, wobei ϕ die Standard-Normalverteilung $N(0, 1)$ ist.

Hinweis: Dazu kann der „Convergence to types“-Satz (4.3) angewendet werden:

$F \in DA(G_2) = DA(N(\gamma, 2c))$. $\exists a_n > 0, b_n, n \in \mathbb{N}$, sodass $\frac{S_n - b_n}{a_n} \xrightarrow{d} G_2$ für $n \rightarrow \infty$. Wähle $\tilde{a}_n = a_n \sqrt{2c} = a_n A$ und $\tilde{b}_n = a_n \gamma + b_n = a_n B + b_n$, dann gilt

$$\lim_{n \rightarrow \infty} \frac{\tilde{a}_n}{a_n} = \sqrt{2c} = A \quad \text{und} \quad \lim_{n \rightarrow \infty} \frac{\tilde{b}_n - b_n}{a_n} = \gamma = B$$

Laut Satz 4.3 Punkt (ii) gilt dann

$$\lim_{n \rightarrow \infty} \frac{S_n - \tilde{b}_n}{\tilde{a}_n} = \tilde{Z} \quad \text{mit} \quad \tilde{Z} \stackrel{d}{=} \frac{G_2 - \gamma}{\sqrt{2c}} \sim N(0, 1).$$

◇

Satz 4.4. (Charakterisierung des Anziehungsgebietes)

(i) Sei ϕ die Verteilungsfunktion der Standardnormalverteilung. Es gilt:

$$F \in DA(\phi) \iff \lim_{x \rightarrow \infty} \frac{x^2 \int_{[-x, x]^C} dF(y)}{\int_{[-x, x]} y^2 dF(y)} = 0,$$

wobei $[-x, x]^C$ das Komplement von $[-x, x]$ in \mathbb{R} ist.

(ii) Für $\alpha \in (0, 2)$ gilt:

$$F \in DA(G_\alpha) \iff F(-x) = \frac{c_1 + o(1)}{x^\alpha} L(x), \quad \bar{F}(x) = \frac{c_2 + o(1)}{x^\alpha} L(x),$$

wobei L eine langsam variierende Funktion ist und $c_1, c_2 \geq 0, c_1 + c_2 > 0$.

Bekannt auch als Satz von Lévy, Feller und Chintschin.

Beweis. in Rényi, 1962. □

Anmerkung: Sei $F \in DA(G_\alpha)$ für $\alpha \in (0, 2)$. Es gilt dann $\mathbb{E}(|X|^\delta) < \infty$ für $\delta < \alpha$ und $\mathbb{E}(|X|^\delta) = \infty$ für $\delta > \alpha$. (Beweis: Hausübung!)

4.1.2 Grenzverteilungen von normierten und zentrierten Maxima

Seien $(X_k), k \in \mathbb{N}$, nicht degenerierte i.i.d. ZV mit Verteilungsfunktion F . Für $n \geq 1$, $M_n := \max\{X_i : 1 \leq i \leq n\}$.

Frage: Welche sind die möglichen (nicht degenerierten) Grenzverteilungen von normierten und zentrierten M_n ?

D.h. wir untersuchen

$$\lim_{n \rightarrow \infty} \mathbb{P}(a_n^{-1}(M_n - b_n) \leq x) = \lim_{n \rightarrow \infty} \mathbb{P}(M_n \leq u_n),$$

wobei $u_n = a_n x + b_n, \forall n \in \mathbb{N}$.

Satz 4.5. (*Poisson Approximation*)

Sei $\tau \in [0, \infty]$ und eine Zahlenfolge $u_n \in \mathbb{R}$ gegeben. Dann gilt:

$$\lim_{n \rightarrow \infty} n\bar{F}(u_n) = \tau \iff \lim_{n \rightarrow \infty} \mathbb{P}(M_n \leq u_n) = \exp\{-\tau\}. \quad (4.6)$$

Beweis. Fall 1: $\tau \in [0, \infty)$.

„ \Rightarrow “: $\lim_{n \rightarrow \infty} n\bar{F}(u_n) = \tau$.

$$\begin{aligned} \mathbb{P}(M_n \leq u_n) &= F^n(u_n) = \exp\{n \ln F(u_n)\} = \exp\{n \ln(1 - \bar{F}(u_n))\} \\ &= \left(1 - \frac{n\bar{F}(u_n)}{n}\right)^n \xrightarrow{n \rightarrow \infty} e^{-\tau} \end{aligned}$$

„ \Leftarrow “: $\lim_{n \rightarrow \infty} \mathbb{P}(M_n \leq u_n) \stackrel{(*)}{=} e^{-\tau}$. Wir zeigen zuerst $\bar{F}(u_n) \xrightarrow{n \rightarrow \infty} 0$.

Ann.: $\exists n_k \in \mathbb{N}, k \in \mathbb{N}, 0 < b < 1 : \bar{F}(u_{n_k}) > b$

$$0 \leq \mathbb{P}(M_{n_k} \leq u_{n_k}) = F^{n_k}(u_{n_k}) = (1 - \bar{F}(u_{n_k}))^{n_k} \leq (1 - b)^{n_k} \xrightarrow{k \rightarrow \infty} 0$$

$\Rightarrow \mathbb{P}(M_{n_k} \leq u_{n_k}) \xrightarrow{k \rightarrow \infty} 0 \not\Leftarrow$ zur Vor.

Aus $(*)$ folgt dann

$$\begin{aligned} \lim_{n \rightarrow \infty} (1 - \bar{F}(u_n))^n &= e^{-\tau} / \ln \\ \ln(1 - \bar{F}(u_n))^n &= n \ln(1 - \bar{F}(u_n)) \sim -n\bar{F}(u_n) \\ \Rightarrow \lim_{n \rightarrow \infty} n\bar{F}(u_n) &= \tau. \end{aligned}$$

Fall 2: $\tau = \infty$.

„ \Rightarrow “: $\lim_{n \rightarrow \infty} n\bar{F}(u_n) = \infty$, „ $\exp(-\infty) = 0$ “.

Für $\mathbb{P}(M_n \leq u_n)$ existiert ein Häufungspunkt α . Falls $\alpha \neq 0$:

$$\lim_{k \rightarrow \infty} \mathbb{P}(M_{n_k} \leq u_{n_k}) = \exp(-\mu)$$

Sei μ so, dass $\exp(-\mu) = \alpha \in (0, 1], \mu \in [0, \infty) \Rightarrow$ Fall 1 $\Rightarrow n_k \bar{F}(u_{n_k}) \rightarrow \mu \in [0, \infty) \not\Leftarrow$.

„ \Leftarrow “: analog. □

Beispiel 4.1.6. Überzeugen Sie Sich mit Hilfe des „convergence to type“ Satzes, dass H und \tilde{H} vom selben Typus sind, wenn $\lim_{n \rightarrow \infty} a_n^{-1}(M_n - b_n) = H$ und $\lim_{n \rightarrow \infty} \tilde{a}_n^{-1}(M_n - \tilde{b}_n) = \tilde{H}$. \diamond

Definition 4.1.7. Eine nicht degenerierte ZV X heißt *max-stabil*, wenn für jedes $n \geq 2$ $\max\{X_1, X_2, \dots, X_n\} \stackrel{d}{=} a_n X + b_n$ für unabhängige Kopien X_1, X_2, \dots, X_n von X und geeignete Konstanten $a_n > 0$ und $b_n \in \mathbb{R}$ gilt.

Satz 4.6. Die Klasse von max-stabilen Verteilungen stimmt mit der Klasse der möglichen nicht degenerierten Grenzverteilungen normierter und zentrierter Maxima von i.i.d. Zufallsvariablen überein.

Beweis. in McNeil, Frey und Embrechts, 2005. □

Satz 4.7. (*Fischer-Tippet Theorem*)

Sei (X_k) eine Folge von i.i.d. ZV. Wenn die Konstanten $a_n, b_n \in \mathbb{R}$, $a_n > 0$, und eine nicht degenerierte Verteilung H existieren, sodass $\lim_{n \rightarrow \infty} a_n^{-1}(M_n - b_n) = H$, dann ist H vom selben Typus wie eine der untenstehenden drei Verteilungen:

$$\begin{array}{ll} \text{Fréchet } \Phi_\alpha(x) & = \begin{cases} 0 & x \leq 0 \\ \exp\{-x^{-\alpha}\} & x > 0 \end{cases} & \alpha > 0 \\ \text{Weibull } \Psi_\alpha(x) & = \begin{cases} \exp\{-(-x)^\alpha\} & x \leq 0 \\ 1 & x > 0 \end{cases} & \alpha > 0 \\ \text{Gumbel } \Lambda(x) & = \exp\{-e^{-x}\} & x \in \mathbb{R} \end{array} .$$

Beweis. Resnick 1987, Seite 9-11. □

Die Verteilungen Φ_α , Ψ_α und Λ heißen *standard Extremwertverteilungen*. Verteilungen, die vom selben Typus wie Φ_α , Ψ_α oder Λ sind, heißen *Extremwertverteilungen*.

Definition 4.1.8. Die ZV X (oder die dazugehörige Verteilung) gehört zum *maximalen Anziugsgebiet* der Extremwertverteilung H wenn die Konstanten $a_n > 0$ und $b_n \in \mathbb{R}$ existieren, sodass $\lim_{n \rightarrow \infty} a_n^{-1}(M_n - b_n) = H$.

Notation: $X \in MDA(H)$ ($F \in MDA(H)$).

Beispiel 4.1.9. $\Phi_\alpha \in MDA(\Phi_\alpha)$: $a_n = n^{1/\alpha}$, $b_n = 0$

$$\begin{aligned} P\left(\frac{M_n - b_n}{a_n} \leq x\right) &= P(M_n \leq n^{1/\alpha}x) = \Phi_\alpha^n(n^{1/\alpha}x) = \exp\left\{-(n^{1/\alpha}x)^{-\alpha}\right\}^n \\ &= \exp\{-n^{-1}x^{-\alpha}\}^n = \exp\left\{-\frac{1}{n}x^{-\alpha}n\right\} = \exp\{-x^{-\alpha}\} = \Phi_\alpha(x) \end{aligned}$$

$\Psi_\alpha(x)$: $a_n = n^{-1/\alpha}$, $b_n = 0$.

$\Lambda(x)$: $a_n = 1$, $b_n = \ln n$.

Es gilt

$$X \sim \Phi_\alpha \Rightarrow \begin{cases} -\frac{1}{X} \sim \Psi_\alpha \\ \ln X \sim \Lambda \end{cases} .$$

◇

Satz 4.8. (*Charakterisierung von MDA*)

$F \in MDA(H)$ mit normierenden Konstanten $a_n > 0$ und $b_n \in \mathbb{R}$ dann und nur dann, wenn

$$\lim_{n \rightarrow \infty} n\bar{F}(a_nx + b_n) = -\ln H(x), \forall x \in \mathbb{R}.$$

Für $H(x) = 0$ wird $-\ln H(x)$ durch ∞ ersetzt.

Beweis. Hinweis: Satz 4.8 folgt aus Satz 4.5. □

Es gibt auch Verteilungen, die zu keinem maximalen Anziehungsgebiet einer Extremwertverteilung gehören!

Lemma 4.1.10. (Leadbetter, 1983)

Sei X eine diskrete gzz. ZV mit Verteilungsfunktion F und rechtem Verteilungsrand $x_F = \infty$. Für $\tau \in (0, \infty) \exists u_n \in \mathbb{R}, n \in \mathbb{N}$, sodass

$$\lim_{n \rightarrow \infty} n\bar{F}(u_n) = \tau \iff \lim_{n \rightarrow \infty} \frac{\bar{F}(n)}{\bar{F}(n-1)} = 1.$$

Beispiel 4.1.11. (Die Poisson Verteilung)

Sei $X \sim Poi(\lambda)$. D.h. $\mathbb{P}(X = k) = e^{-\lambda} \lambda^k / k!, k \in \mathbb{N}_0, \lambda > 0$.

Zeigen Sie, dass es keine Extremwertverteilung Z gibt, für die $X \in MDA(Z)$.

Mit Hilfe des Lemmas:

$$\begin{aligned} \bar{F}(n) &= 1 - F(n) = 1 - \mathbb{P}(X \leq n) = 1 - (\mathbb{P}(X \leq n-1) + \mathbb{P}(X = n)) \\ &= 1 - \mathbb{P}(X \leq n-1) - \mathbb{P}(X = n) = 1 - F(n-1) - \mathbb{P}(X = n) \\ &= \bar{F}(n-1) - \mathbb{P}(X = n) \end{aligned}$$

Es gilt somit

$$\begin{aligned} \frac{\bar{F}(n)}{\bar{F}(n-1)} &= \frac{\bar{F}(n-1) - \mathbb{P}(X = n)}{\bar{F}(n-1)} = 1 - \frac{F(n) - F(n-1)}{\bar{F}(n-1)} \\ &= 1 - \frac{\lambda^n}{n!} \left(\sum_{r=n}^{\infty} \frac{\lambda^r}{r!} \right)^{-1} = 1 - \left(1 + \sum_{r=n+1}^{\infty} \frac{n!}{r!} \lambda^{r-n} \right)^{-1} \xrightarrow{n \rightarrow \infty} 0, \end{aligned}$$

denn

$$\sum_{r=n+1}^{\infty} \frac{n!}{r!} \lambda^{r-n} = \sum_{s=1}^{\infty} \frac{\lambda^s}{(n+1) \cdots (n+s)} \leq \sum_{s=1}^{\infty} \left(\frac{\lambda}{n} \right)^s = \frac{\lambda/n}{1 - \lambda/n} \xrightarrow{n \rightarrow \infty} 0 \quad \text{für } \lambda < n.$$

Daraus folgt $\lim_{n \rightarrow \infty} n\bar{F}(u_n) \neq \tau$ und demnach $Poi \notin MDA$ (wegen Satz 4.8). \diamond

Beispiel 4.1.12. (Maxima der Exponentialverteilung)

Sei (X_k) eine Folge von i.i.d. ZV mit Verteilungsfunktion $F, F(x) = 1 - e^{-x}$ für $x \geq 0$. Zeigen Sie, dass $F \in MDA(\Lambda)$ mit normierenden Konstanten $a_n = 1$ und $b_n = \ln n$. \diamond

Beispiel 4.1.13. (Maxima der Cauchy-Verteilung)

Sei (X_k) eine Folge von i.i.d. ZV mit Verteilungsfunktion F und Dichtefunktion $f, f(x) = (\pi(1+x^2))^{-1}$ für $x \in \mathbb{R}$. Zeigen Sie, dass $F \in MDA(\Phi_1)$ mit normierenden Konstanten $a_n = n/\pi$ und $b_n = 0$. \diamond

Definition 4.1.14. (Die Verallgemeinerte Extremwertverteilung)

Die Verteilungsfunktion H_γ sei folgendermaßen gegeben:

$$H_\gamma(x) = \begin{cases} \exp\{-(1+\gamma x)^{-1/\gamma}\} & \text{wenn } \gamma \neq 0 \\ \exp\{-\exp\{-x\}\} & \text{wenn } \gamma = 0, \end{cases}$$

wobei $1 + \gamma x > 0$. D.h. der Definitionsbereich von H_γ ist folgendermaßen gegeben:

$$\begin{aligned} x &> -\gamma^{-1} && \text{wenn } \gamma > 0 \\ x &< -\gamma^{-1} && \text{wenn } \gamma < 0 \\ x &\in \mathbb{R} && \text{wenn } \gamma = 0. \end{aligned}$$

H_γ heißt verallgemeinerte standard Extremwertverteilung.

Satz 4.9. (Charakterisierung von $MDA(H_\gamma)$)

- $F \in MDA(H_\gamma)$ mit $\gamma > 0 \iff F \in MDA(\Phi_\alpha)$ mit $\alpha = 1/\gamma > 0$.
- $F \in MDA(H_0) \iff F \in MDA(\Lambda)$.
- $F \in MDA(H_\gamma)$ mit $\gamma < 0 \iff F \in MDA(\Psi_\alpha)$ mit $\alpha = -1/\gamma > 0$.

Satz 4.10. ($MDA(\Phi_\alpha)$, Gnedenko 1943)

$F \in MDA(\Phi_\alpha)$ ($\alpha > 0$) $\iff \bar{F} \in RV_{-\alpha}$ ($\alpha > 0$).

Wenn $F \in MDA(\Phi_\alpha)$, dann $\lim_{n \rightarrow \infty} a_n^{-1} M_n = \Phi_\alpha$ mit $a_n = F^{\leftarrow}(1 - n^{-1})$.

Beispiel 4.1.15. Zeigen Sie, dass die untenstehenden Verteilungen dem $MDA(\Phi_\alpha)$ angehören und bestimmen Sie die normierenden Konstanten.

- Pareto: $F(x) = 1 - x^{-\alpha}$, $x > 1$, $\alpha > 0$.
- Cauchy: $f(x) = (\pi(1 + x^2))^{-1}$, $x \in \mathbb{R}$.
- Student: $f(x) = \frac{\Gamma((\alpha+1)/2)}{\sqrt{\alpha\pi}\Gamma(\alpha/2)(1+x^2/\alpha)^{(\alpha+1)/2}}$, $\alpha \in \mathbb{N}$, $x \in \mathbb{R}$.
- Loggamma: $f(x) = \frac{\alpha^\beta}{\Gamma(\beta)} (\ln x)^{\beta-1} x^{-\alpha-1}$, $x > 1$, $\alpha, \beta > 0$.

◇

Satz 4.11. ($MDA(\Psi_\alpha)$, Gnedenko 1943)

$F \in MDA(\Psi_\alpha)$ ($\alpha > 0$) $\iff x_F := \sup\{x \in \mathbb{R} : F(x) < 1\} < \infty$ und $\bar{F}(x_F - x^{-1}) \in RV_{-\alpha}$ ($\alpha > 0$).

Wenn $F \in MDA(\Psi_\alpha)$, dann $\lim_{n \rightarrow \infty} a_n^{-1}(M_n - x_F) = \Psi_\alpha$ mit $a_n = x_F - F^{\leftarrow}(1 - n^{-1})$.

Beispiel 4.1.16. Sei $X \sim U(0, 1)$. Es gilt $X \in MDA(\Psi_1)$ mit $a_n = 1/n$, $n \in \mathbb{N}$.

◇

Satz 4.12. ($MDA(\Lambda)$)

Sei F eine Verteilungsfunktion mit rechtem Endpunkt $x_F \leq \infty$.

$F \in MDA(\Lambda)$ dann und nur dann, wenn ein $z < x_F$ existiert, sodass für F folgende Darstellung gilt:

$$\bar{F}(x) = c(x) \exp \left\{ - \int_z^x \frac{g(t)}{a(t)} dt \right\}, \forall x, z < x \leq x_F.$$

Für die Funktionen $c(x)$ und $g(x)$ gilt $\lim_{x \uparrow x_F} c(x) = c > 0$ und $\lim_{t \uparrow x_F} g(t) = 1$, und $a(t)$ ist eine positive absolut stetige Funktion, sodass $\lim_{t \uparrow x_F} a'(t) = 0$.

Satz 4.13. ($MDA(\Lambda)$, alternative Charakterisierung)

Eine Verteilungsfunktion F gehört zu $MDA(\Lambda)$ dann und nur dann, wenn eine positive Funktion \tilde{a} existiert, sodass

$$\lim_{x \uparrow x_F} \frac{\bar{F}(x + u\tilde{a}(x))}{\bar{F}(x)} = e^{-u}, \forall u \in \mathbb{R}$$

Eine mögliche Wahl für \tilde{a} ist $\tilde{a}(x) = a(x)$

$$a(x) = \int_x^{x_F} \frac{\bar{F}(t)}{\bar{F}(x)} dt$$

Die Funktion $a(x)$ heißt durchschnittliche Überschuffunktion (mean excess function):

$$a(x) = E(X - x | X > x), \forall x \leq x_F$$

Einige Verteilungen, die dem $MDA(\Lambda)$ angehören:

- Normal: $f(x) = (2\pi)^{-1/2} \exp\{-x^2/2\}$, $x \in \mathbb{R}$.
- Exponential: $f(x) = \lambda^{-1} \exp\{-\lambda x\}$, $x > 0$, $\lambda > 0$.
- Lognormal: $f(x) = (2\pi x^2)^{-1/2} \exp\{-(\ln x)^2/2\}$, $x > 0$.
- Gamma: $f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} \exp\{-\beta x\}$, $x > 0$, $\alpha, \beta > 0$.

4.2 Graphische Methoden zur Untersuchung des Verteilungsrandes

- Histogramm (Abb. 4.1)

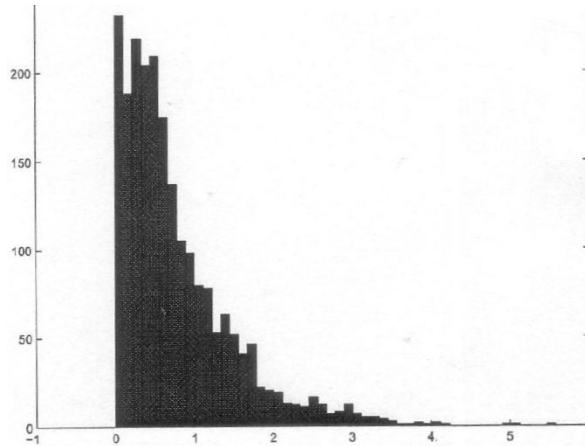


Abbildung 4.1: Histogramm der logarithmierten Schäden eines dänischen Versicherungsunternehmens

- Quantil-Quantil Plots

X_1, X_2, \dots, X_n sind i.i.d. ZV mit einer unbekanntem Verteilung \tilde{F} . Es wird vermutet, dass \tilde{F} am Rand von einer bekannten Verteilung F approximiert wird. Wie kann man diese Vermutung testen?

Sei $X_{n,n} \leq X_{n-1,n} \leq \dots \leq X_{1,n}$ eine sortierte Stichprobe aus X_1, X_2, \dots, X_n .

QQ-Plot: $\{(X_{k,n}, F^{\leftarrow}(\frac{n-k+1}{n+1})) : k = 1, 2, \dots, n\}$.

Bei einer plausiblen Vermutung stellt der QQ-Plot eine einigermaßen lineare Abhängigkeit dar. Diese Eigenschaft bleibt auch dann erhalten, wenn die echte Verteilung und die Referenz-Verteilung nicht übereinstimmen, sondern vom selben Typus sind.

Faustregel: Je größer das Quantil, um so mehr „heavy tailed“ ist die Verteilung!

In den 2 QQ-Plots in Abb. 4.2 sind die logarithmierten Schäden eines dänischen Versicherungsunternehmens einmal gegen eine Exponentialverteilung (links) und einmal gegen die Pareto(1)-Verteilung (rechts) geplottet. Man kann erkennen, dass die Daten *heavier tails* als die Exponentialverteilung und anscheinend eine Pareto(α)-Verteilung haben (aufgrund des fast linearen Verlaufs erkennbar) mit α nahe 1.

Beispiel 4.2.1. QQ-Plot von $F(x) = 1 - e^{-x}$ für $x \geq 0$ (Exponentialverteilung) und $G(x) = 1 - x^{-2}$ für $x \geq 1$ (Paretoverteilung) mit $n = 50, 100, 200, 500$.

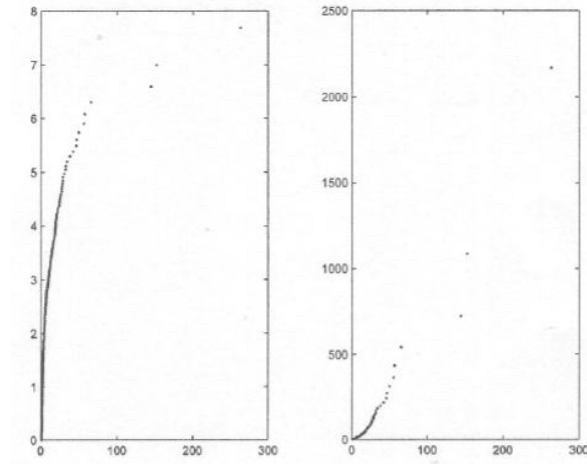
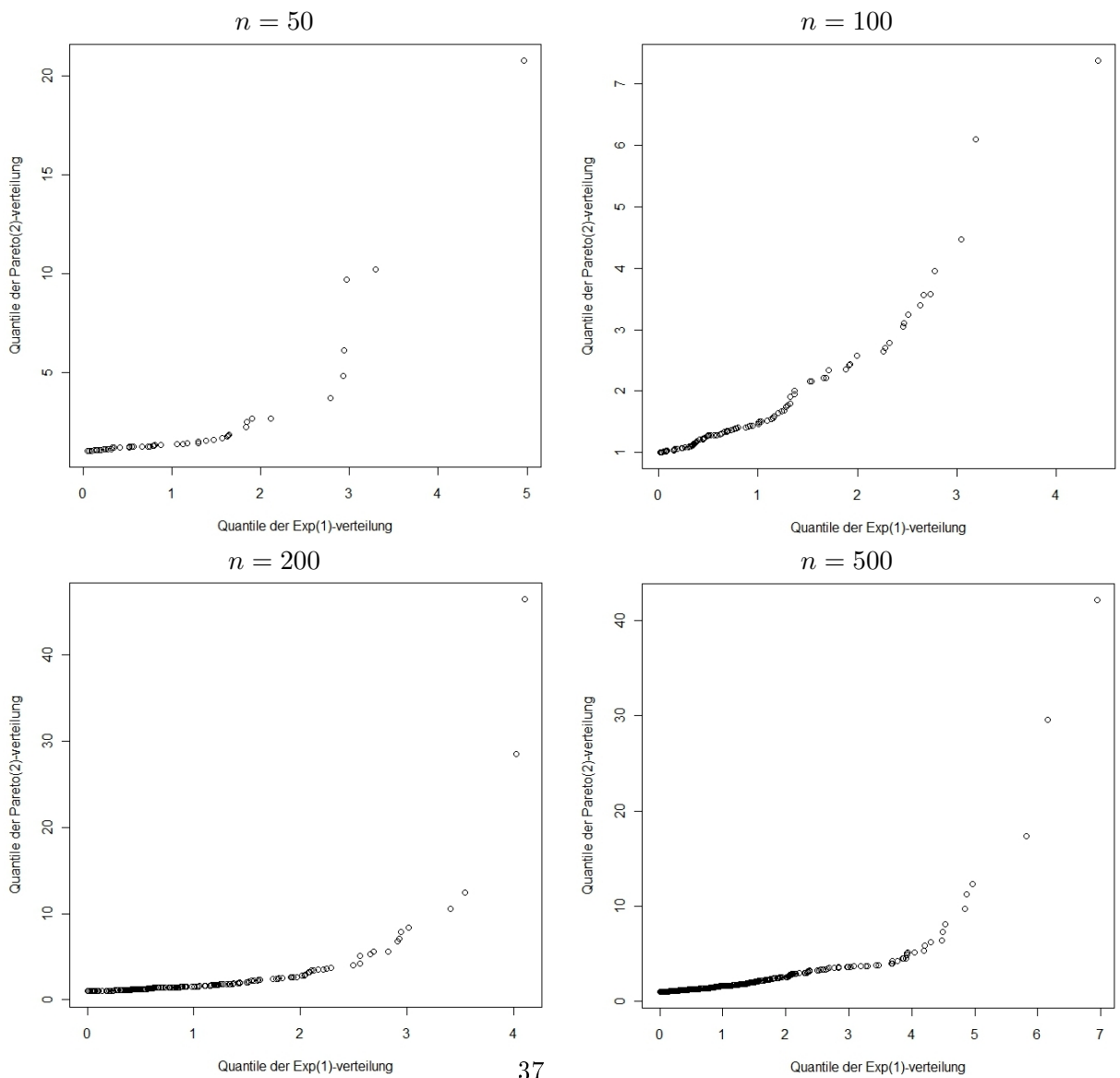


Abbildung 4.2: QQ-Plot der logarithmierten Schäden eines dänischen Versicherungsunternehmens



Man kann deutlich erkennen, dass die Paretoverteilung heavyier tails als die Exponentialverteilung hat. \diamond

4.2.1 Der Hill Schätzer

Seien X_1, X_2, \dots, X_n i.i.d. mit Verteilungsfunktion F , sodass $\bar{F} \in RV_{-\alpha}$, $\alpha > 0$, d.h. $\bar{F}(x) = x^{-\alpha}L(x)$ mit $L \in RV_0$.
Es soll nun α geschätzt werden.

Satz 4.14. (Satz von Karamata)

Sei L eine langsam variierende und lokal beschränkte Funktion auf $[x_0, +\infty)$ für ein $x_0 \in \mathbb{R}$. Dann gilt:

(a) Für $\kappa > -1$:

$$\int_{x_0}^x t^\kappa L(t) dt \sim \frac{1}{\kappa + 1} x^{\kappa+1} L(x) \text{ für } x \rightarrow \infty.$$

(b) Für $\kappa < -1$:

$$\int_x^{+\infty} t^\kappa L(t) dt \sim -\frac{1}{\kappa + 1} x^{\kappa+1} L(x) \text{ für } x \rightarrow \infty.$$

Beweis. in Bingham et al. 1987. \square

Annahme: L ist lokal beschränkt in $[u, +\infty)$.

Aus dem Satz von Karamata folgt:

$$\lim_{u \rightarrow \infty} \mathbb{E}(\ln(X) - \ln(u) | \ln(X) > \ln(u)) = \lim_{u \rightarrow \infty} \frac{1}{\bar{F}(u)} \int_u^\infty (\ln x - \ln u) dF(x) = \alpha^{-1}, \quad (4.7)$$

denn

$$\begin{aligned} \int_u^\infty \underbrace{(\ln x - \ln u)}_t \underbrace{dF(x)}_{v'} &= \underbrace{- (\ln x - \ln u) \bar{F}(x) \Big|_u^\infty}_{=0, \text{ da } \bar{F}(x) \text{ reg. var.}} + \int_u^\infty \frac{\bar{F}(x)}{x} dx \\ &= \int_u^\infty x^{-\alpha} \frac{L(x)}{x} dx = \int_u^\infty x^{-\alpha-1} L(x) dx \\ &\stackrel{\text{Karamata}}{\sim} \frac{1}{\alpha} u^{-\alpha} L(u) \quad \text{für } u \rightarrow \infty \end{aligned}$$

und somit gilt

$$\frac{1}{\bar{F}(u)} \int_u^\infty (\ln u - \ln x) dF(x) \stackrel{u \rightarrow \infty}{\sim} \frac{1}{\bar{F}(u)} \alpha^{-1} \underbrace{u^{-\alpha} L(u)}_{\bar{F}(u)} = \alpha^{-1}.$$

Für die empirische Verteilung $F_n(x) = \frac{1}{n} \sum_{k=1}^n I_{[x_k, \infty)}(x)$ und eine hohe stichprobenabhängige Schwelle $x_{k,n}$, erhält man:

$$\mathbb{E}(\ln(X) - \ln(x_{k,n}) | \ln(X) > \ln(x_{k,n})) \approx$$

$$\frac{1}{\bar{F}_n(x_{k,n})} \int_{x_{k,n}}^{\infty} (\ln x - \ln x_{k,n}) dF_n(x) = \frac{1}{k-1} \sum_{j=1}^{k-1} (\ln x_{j,n} - \ln x_{k,n}).$$

Wenn $k = k(n) \rightarrow \infty$ und $k/n \rightarrow 0$, dann $x_{k,n} \rightarrow \infty$ für $n \rightarrow \infty$, und dann folgt aus (4.7) der Hill-Schätzer $\hat{\alpha}_{k,n}^{(H)}$:

$$\lim_{n \rightarrow \infty} \frac{1}{k-1} \sum_{j=1}^{k-1} (\ln x_{j,n} - \ln x_{k,n}) = \alpha^{-1} \implies \hat{\alpha}_{k,n}^{(H)} = \left(\frac{1}{k} \sum_{j=1}^k (\ln x_{j,n} - \ln x_{k,n}) \right)^{-1}.$$

Wie wird ein passendes k für eine gegebene Stichprobengröße n gewählt?

- k zu klein: hohe Varianz des Schätzers!
- k zu groß: Schätzer basiert auf zentralen Werten der Verteilung \implies Der Schätzer ist verzerrt!

Eine Möglichkeit besteht darin, eine grafische Inspektion des Hill Plots $\left\{ \left(k, \hat{\alpha}_{k,n}^{(H)} \right) : k = 2, \dots, n \right\}$ durchzuführen.

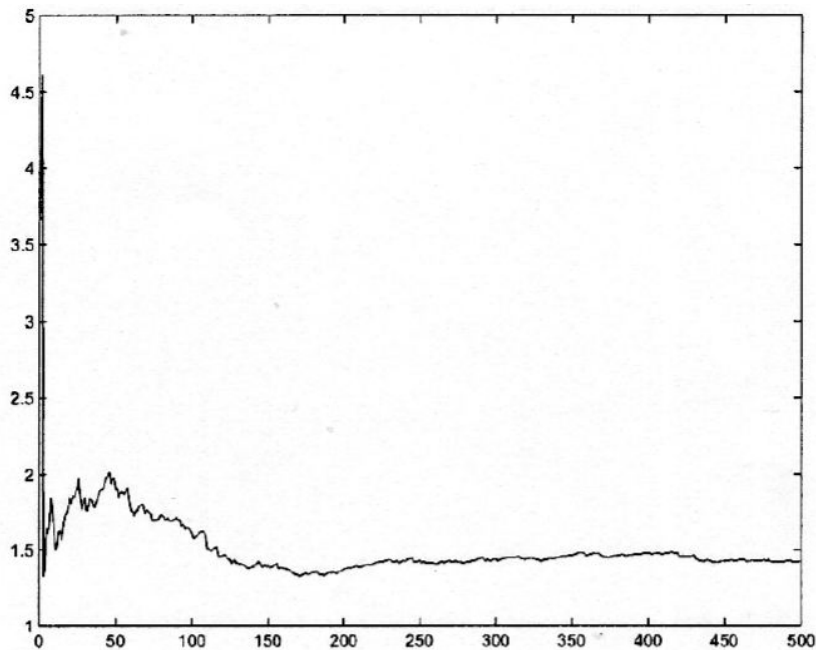


Abbildung 4.3: Hill-Plot der Daten des dänischen Versicherungsunternehmens

Der Hill-Plot in Abb. 4.3 wurde mit den schon oben gesehenen Daten des dänischen Versicherungsunternehmens erzeugt. Eine mögliche Wahl von k ist hier 75, da der Plot ab hier recht stabil aussieht.

Für einen gegebenen Schätzer $\hat{\alpha}_{k,n}^{(H)}$ von α erhält man folgenden Schätzer für die Randverteilung \bar{F} :

$$\hat{\bar{F}}(x) = \frac{k}{n} \left(\frac{x}{x_{k,n}} \right)^{-\hat{\alpha}_{k,n}^{(H)}},$$

denn es gilt für $x/u = x_{k,n}$

$$\begin{aligned}\bar{F}(x) &= \bar{F}\left(\frac{x}{u}\right) \stackrel{u \rightarrow \infty}{\approx} u^{-\alpha} \bar{F}\left(\frac{x}{u}\right) = \left(\frac{x}{x_{k,n}}\right)^{-\alpha} \bar{F}(x_{k,n}) \\ &\approx \left(\frac{x}{x_{k,n}}\right)^{-\alpha} \hat{\bar{F}}(x_{k,n}) = \left(\frac{x}{x_{k,n}}\right)^{-\alpha} \frac{k}{n} \quad \text{für } \frac{k}{n} \rightarrow 0, u \rightarrow \infty\end{aligned}$$

und folgenden Quantil-Schätzer:

$$\hat{q}_p = \hat{F}^{\leftarrow}(p) = \left(\frac{n}{k}(1-p)\right)^{-1/\hat{\alpha}_{k,n}^{(H)}} x_{k,n}.$$

4.2.2 Die POT Methode (Peaks over Threshold)

Definition 4.2.2. (Die verallgemeinerte Pareto Verteilung (GPD))

Die standard GPD G_γ ist definiert durch

$$G_\gamma(x) = \begin{cases} 1 - (1 + \gamma x)^{-1/\gamma} & \text{für } \gamma \neq 0 \\ 1 - \exp\{-x\} & \text{für } \gamma = 0 \end{cases}$$

wobei $x \in D(\gamma)$

$$D(\gamma) = \begin{cases} 0 \leq x < \infty & \text{für } \gamma \geq 0 \\ 0 \leq x \leq -1/\gamma & \text{für } \gamma < 0 \end{cases}$$

oder $G_\gamma(x) = 1 - (1 + \gamma x)^{-1/\gamma}$, $x \in D(\gamma)$ und $G_0 = \lim_{\gamma \rightarrow 0} G_\gamma$.

Sei $\nu \in \mathbb{R}$ und $\beta > 0$. Eine GPD ist durch folgende Verteilungsfunktion gegeben:

$$G_{\gamma,\nu,\beta} = \begin{cases} 1 - (1 + \gamma \frac{x-\nu}{\beta})^{-1/\gamma} & \gamma \neq 0 \\ 1 - \exp\{-\frac{x-\nu}{\beta}\} & \gamma = 0, \end{cases}$$

wobei $x \in D(\gamma, \nu, \beta)$ und

$$D(\gamma, \nu, \beta) = \begin{cases} \nu \leq x < \infty & \text{für } \gamma \geq 0 \\ \nu \leq x \leq \nu - \beta/\gamma & \text{für } \gamma < 0. \end{cases}$$

Satz 4.15. (Charakterisierung von $MDA(H_\gamma)$)

Sei $\gamma \in \mathbb{R}$. Die untenstehenden Aussagen sind äquivalent:

(i) $F \in MDA(H_\gamma)$

(ii) Es existiert eine positive messbare Funktion $a(\cdot)$, sodass für $x \in D(\gamma)$

$$\lim_{u \uparrow x_F} \frac{\bar{F}(u + xa(u))}{\bar{F}(u)} = \bar{G}_\gamma(x).$$

Definition 4.2.3. (Exzess-Verteilung)

Sei X eine ZV mit Verteilungsfunktion F und rechtem Endpunkt x_F . Für $u < x_F$ heißt

$$F_u(x) = P(X - u \leq x | X > u), \quad x \geq 0$$

Exzess-Verteilungsfunktion über die Schwelle u .

Satz 4.16. (Eine weitere Charakterisierung von $MDA(H_\gamma)$)

Sei $\gamma \in \mathbb{R}$. Folgende Aussagen sind äquivalent:

(i) $F \in MDA(H_\gamma)$

(ii) Es existiert eine positive messbare Funktion $\beta(\cdot)$, sodass

$$\lim_{u \uparrow x_F} \sup_{x \in (0, x_F - u)} |F_u(x) - G_{\gamma, 0, \beta(u)}(x)| = 0.$$

Beispiel 4.2.4. $\bar{F} \in RV_{-\alpha}$, $\alpha > 0 \Leftrightarrow F \in MDA(H_{1/\alpha})$.

Da $x_F = \infty$: $\forall x > 0$

$$\lim_{u \rightarrow \infty} \frac{\bar{F}(u + x \frac{u}{\alpha})}{\bar{F}(u)} = \lim_{u \rightarrow \infty} \frac{\bar{F}(u(1 + \frac{x}{\alpha}))}{\bar{F}(u)} = \left(1 + \frac{x}{\alpha}\right)^{-\alpha} = \bar{G}_{1/\alpha}(x)$$

und es gilt

$$\begin{aligned} \frac{\bar{F}(u + xa(u))}{\bar{F}(u)} &= \frac{\mathbb{P}(X > u + xa(u))}{\bar{F}(u)} = \frac{\mathbb{P}(X > u + xa(u), X > u)}{\mathbb{P}(X > u)} \\ &= \mathbb{P}(X > u + xa(u) | X > u) = \bar{F}_u(xa(u)) \\ &\Rightarrow \lim_{u \rightarrow x_F} \bar{F}_u(xa(u)) = \bar{G}_\gamma(x) \end{aligned}$$

◇

Beispiel 4.2.5. $X \sim G_{\gamma, 0, \beta} =: G_{\gamma, \beta}$.

$$G_u(x) = \frac{G_{\gamma, \beta}(x + u) - G_{\gamma, \beta}(u)}{1 - G_{\gamma, \beta}(u)} = \frac{\mathbb{P}(u < X < x + u)}{\mathbb{P}(X > u)} = \mathbb{P}(X - u < x | X > u).$$

Wir zeigen für $0 < \gamma < 1$, $\beta > 0$: $e(u) = \mathbb{E}(X - u | X > u) = \frac{\beta + \gamma u}{1 - \gamma}$.

Für $\nu \neq 0$:

$$e_{GPD}(u) = \frac{1}{\bar{G}_{\gamma, \nu, \beta}(u)} \underbrace{\int_u^\infty (X - u) dG_{\gamma, \nu, \beta}(x)}_I$$

mit

$$\begin{aligned} \bar{G}_{\gamma, \nu, \beta}(x) &= \left(1 + \gamma \frac{x - \nu}{\beta}\right)^{-1/\gamma} \\ d\bar{G}_{\gamma, \nu, \beta}(x) &= -\frac{1}{\gamma} \left(1 + \gamma \frac{x - \nu}{\beta}\right)^{-1/\gamma - 1} \frac{\gamma}{\beta}, \quad x \geq \nu. \end{aligned}$$

folgt mit

$$\begin{aligned}
I &= \int_u^\infty (x-u) \frac{1}{\beta} \left(1 + \gamma \frac{x-\nu}{\beta}\right)^{-1/\gamma-1} dx = \int_u^\infty \left(\frac{x-\nu}{\beta} + \frac{\nu}{\beta} - \frac{u}{\beta}\right) \left(1 + \gamma \frac{x-\nu}{\beta}\right)^{-1/\gamma-1} dx \\
&= \beta \left(\int_{\frac{u-\nu}{\beta}}^\infty t(1+\gamma t)^{-1/\gamma-1} dt + \int_{\frac{u-\nu}{\beta}}^\infty (1+\gamma t)^{-1/\gamma-1} \frac{\nu-u}{\beta} dt \right) \\
&= \beta \left(\frac{1}{\gamma} \int_{\frac{u-\nu}{\beta}}^\infty (1+\gamma t)(1+\gamma t)^{-1/\gamma-1} dt - \frac{1}{\gamma} \int_{\frac{u-\nu}{\beta}}^\infty (1+\gamma t)^{-1/\gamma-1} dt \right) + \beta \frac{\nu-u}{\beta} \int_{\frac{u-\nu}{\beta}}^\infty (1+\gamma t)^{-1/\gamma-1} dt \\
&= \beta \left(\frac{1}{\gamma^2} \int_{\frac{u-\nu}{\beta}}^\infty (1+\gamma t)^{-1/\gamma} d(1+\gamma t) + \left(\frac{\nu-u}{\beta} - \frac{1}{\gamma}\right) \int_{\frac{u-\nu}{\beta}}^\infty (1+\gamma t)^{-1/\gamma-1} dt \right),
\end{aligned}$$

dass

$$e_{GPD}(u) = \frac{\beta + \gamma u}{1 - \gamma} - \frac{\nu \gamma}{1 - \gamma}$$

◇

Beispiel 4.2.6. $X \sim \text{Exp}(\lambda)$, $f(x) = \lambda e^{-\lambda x}$ für $x, \lambda > 0$

$$\begin{aligned}
e(u) &= \frac{1}{\bar{F}(u)} \int_u^\infty (y-u) dF(y) = \frac{1}{\bar{F}(u)} \left(\int_u^\infty y dF(y) - \int_u^\infty u dF(y) \right) \\
&= \frac{1}{\bar{F}(u)} \left(\int_u^\infty y \lambda e^{-\lambda y} dy - u \bar{F}(u) \right) = \frac{1}{\bar{F}(u)} \left(u e^{-\lambda u} + \frac{1}{\lambda} (1 - F(u)) - u \bar{F}(u) \right) \\
&= \frac{1}{\bar{F}(u)} \left(u e^{-\lambda u} + \left(\frac{1}{\lambda} - u\right) \bar{F}(u) \right) = u + \frac{1}{\lambda} - u = \frac{1}{\lambda}.
\end{aligned}$$

◇

4.2.3 Schätzer für den Tail und das Quantil der Exzess-Verteilung

Seien X_1, \dots, X_n i.i.d. ZV mit Verteilungsfunktion $F \in MDA(H_\gamma)$ für $\gamma \in \mathbb{R}$.

- Wähle eine hohe Schwelle u (unter Verwendung von geeigneten stat. Verfahren) und berechne

$$N_u = \#\{i \in \{1, 2, \dots, n\}: X_i > u\}$$

- Sei Y_1, Y_2, \dots, Y_{N_u} eine Stichprobe von Exzess-Beobachtungen. Bestimme $\hat{\beta}$ und $\hat{\gamma}$, sodass folgendes gilt:

$$\bar{F}_u(y) \approx \bar{G}_{\hat{\gamma}, 0, \hat{\beta}(u)}(y),$$

wobei $\bar{F}_u(y) = P(X - u > y | X > u)$

- Kombiniere die obigen zwei Schritte um folgende Schätzer zu berechnen:

$$\hat{\bar{F}}(u+y) = \frac{N_u}{n} \left(1 + \hat{\gamma} \frac{y}{\hat{\beta}}\right)^{-1/\hat{\gamma}} \quad (4.8)$$

$$\hat{q}_p = u + \frac{\hat{\beta}}{\hat{\gamma}} \left(\left(\frac{n}{N_u} (1-p)\right)^{-\hat{\gamma}} - 1 \right). \quad (4.9)$$

Beweis. 2. Punkt:

$$\bar{F}_u(y) = \mathbb{P}(X - u > y | X > u) = \frac{\mathbb{P}(X > u + y, X > u)}{\mathbb{P}(X > u)} = \frac{\mathbb{P}(X > u + y)}{\mathbb{P}(X > u)} = \frac{\bar{F}(u + y)}{\bar{F}(u)}$$

$\Rightarrow \bar{F}(u + y) = \bar{F}_u(y)\bar{F}(u)$. Mit dem letzten Satz gilt also $\bar{F}_u(y) \approx \bar{G}_{\hat{\gamma}, 0, \hat{\beta}(u)}(y)$.

3. Punkt: Approximation von $\bar{F}(u)$:

$$\hat{F}(u) = \frac{1}{n} \sum_{i=1}^n I_{\{X_i > u\}} = \frac{N_u}{n}$$

$\hat{F}_u(y) \approx \bar{G}_{\hat{\gamma}, 0, \hat{\beta}} \Rightarrow$

$$\hat{F}(u + y) = \frac{N_u}{n} \left(1 + \hat{\gamma} \frac{y}{\hat{\beta}}\right)^{-1/\hat{\gamma}} \quad \text{falls } \hat{\gamma} \neq 0$$

$$\hat{F}(u + y) = \frac{N_u}{n} \exp\left(-\frac{y}{\hat{\beta}}\right) \quad \text{falls } \hat{\gamma} = 0$$

und für das Quantil $q_p(F)$ gilt:

$q_p(F) = \text{VaR}_p(F) = F^{\leftarrow}(p) \approx \hat{F}^{\leftarrow}(p) \Rightarrow \hat{F}(q_p) = p$. Sei $q_p = u + y$.

$$\hat{F}(u + y) = 1 - \frac{N_u}{n} \left(1 + \hat{\gamma} \frac{y}{\hat{\beta}}\right)^{-1/\hat{\gamma}} = p$$

$$\stackrel{\text{löse nach } y}{\Rightarrow} y = \left(\left((1-p) \frac{n}{N_u} \right)^{-\hat{\gamma}} - 1 \right) \frac{\hat{\beta}}{\hat{\gamma}}$$

$$\Rightarrow \hat{q}_p = u + \frac{\hat{\beta}}{\hat{\gamma}} \left(\left(\frac{n}{N_u} (1-p) \right)^{-\hat{\gamma}} - 1 \right).$$

□

4.2.4 Wie wird eine hohe Schwelle u gewählt?

- u zu groß: Wenige Beobachtungen für die Schätzung von $\hat{\beta}$ und $\hat{\gamma}$.
- u zu klein: die Approximation $\bar{F}_u(y) \approx \bar{G}_{\hat{\gamma}, \hat{\beta}(u)}(y)$ ist nicht gut genug.

Grundidee: Inspektion des Plots der empirischen durchschnittlichen Exzess-Funktion und Auswahl einer Schwelle u_0 , sodass die empirische durchschnittliche Exzess-Funktion für $u > u_0$ annähernd linear ist.

Begründung: die durchschnittliche Exzess-Funktion der $GPD_{\gamma, 0, \beta}$ ist linear!

4.2.5 Die empirische durchschnittliche Exzess-Funktion

Seien X_1, X_2, \dots, X_n i.i.d ZV und sei $N_u = \#\{i: 1 \leq i \leq n, X_i > u\}$ die Anzahl der Überschreitung von u durch X_i .

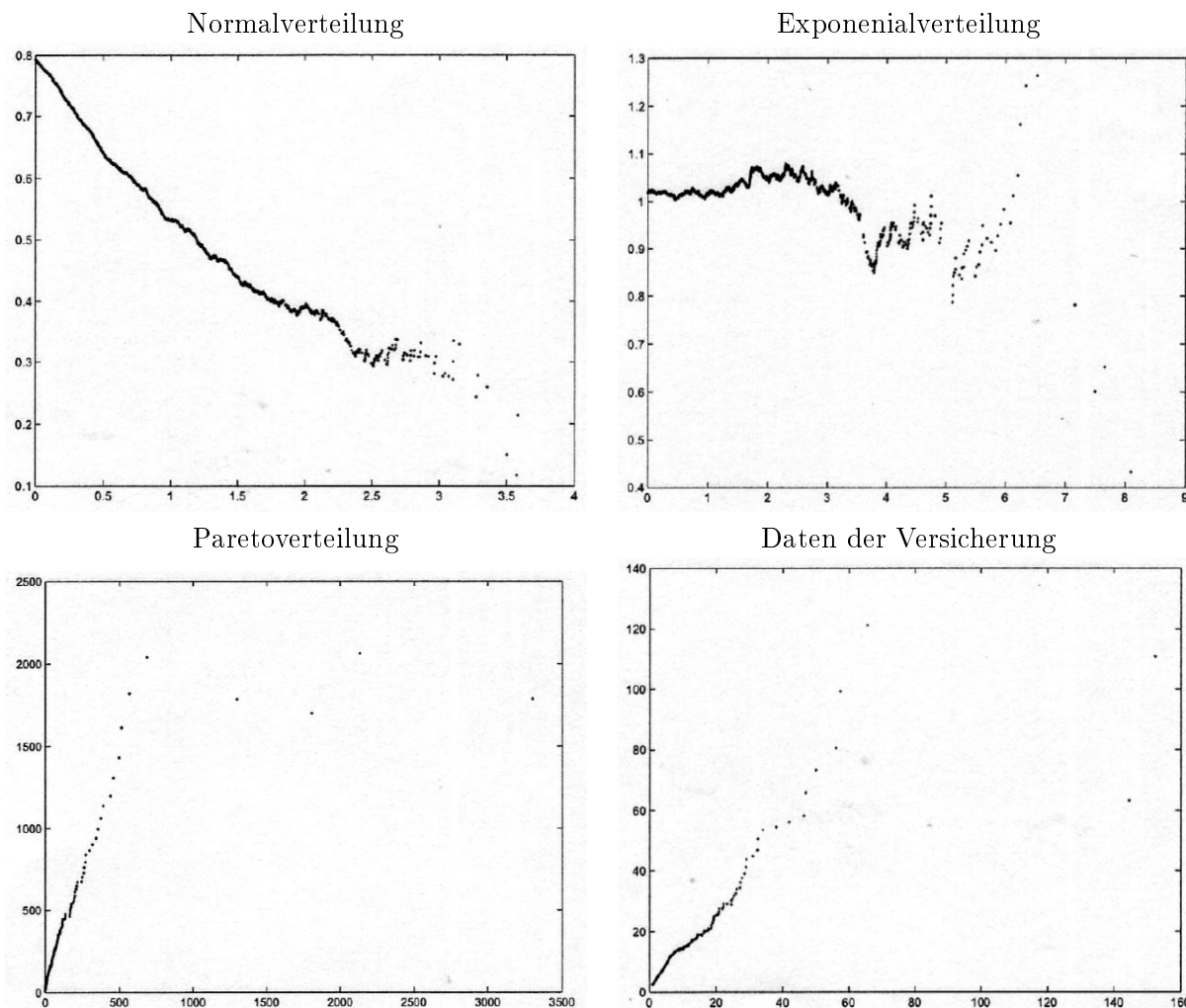
Die empirische durchschnittliche Exzess-Funktion $e_n(u)$ ist gegeben als

$$e_n(u) = \frac{1}{N_u} \sum_{i=1}^n (X_i - u) I\{X_i > u\}.$$

Man betrachtet den Plot der durchschnittlichen Exzess-Funktion: $(X_{k,n}, e_n(X_{k,n}))$ für $k = 1, 2, \dots, n - 1$.

Wenn dieser Plot annähernd linear mit einem positiven Gradienten ist, so wird angenommen, dass die Verteilung einen heavy-tailed Pareto-ähnlichen Tail hat.

Beispiel 4.2.7. Es werden aus der Normal-, Exponential und Paretoverteilung durchschnittliche überschüsse simuliert und mit denen der dänischen Versicherungsgesellschaft verglichen.



Der Plot der Versicherungsdaten sieht annähernd linear aus, was eine Paretoverteilung vermuten lässt. \diamond

4.2.6 Schätzung der Parameter γ und β

Sei u eine gegebene Schwelle und Y_1, Y_2, \dots, Y_{N_u} Beobachtungen der überschüsse $X_i > u$, $1 \leq i \leq n$.

Die Log-Likelihood Funktion ist definiert als

$$\ln L(\gamma, \beta, Y_1, \dots, Y_{N_u}) = -N_u \ln \beta - \left(\frac{1}{\gamma} + 1\right) \sum_{i=1}^{N_u} \ln \left(1 + \frac{\gamma}{\beta} Y_i\right),$$

wobei $Y_i \geq 0$ für $\gamma > 0$ und $0 \leq Y_i \leq -\beta/\gamma$ für $\gamma < 0$.

$L(\gamma, \beta, Y_1, \dots, Y_{N_u})$ ist die bedingte Wahrscheinlichkeit, dass $\bar{F}_u(y) \approx \bar{G}_{\gamma,0,\beta}(y)$ unter der Bedingung, dass die Beobachtungen der Überschüsse Y_1, Y_2, \dots, Y_{N_u} sind.

Für die Ermittlung der Likelihood Funktion siehe Daley, Veve-Jones (2003) and Coles (2001).

Als Schätzer $\hat{\gamma}$ und $\hat{\beta}$ werden jene Werte von γ bzw. β verwendet, die die Log-Likelihood Funktion maximieren (ML-Schätzer).

Die Methode funktioniert gut für $\gamma > -1/2$.

Die ML-Schätzer sind in diesem Fall normal verteilt:

$$(\hat{\gamma} - \gamma, \frac{\hat{\beta}}{\beta} - 1) \sim N_2(0, \Sigma^{-1}/N_u) \text{ wobei } \Sigma^{-1} = \begin{pmatrix} 1 + \gamma & -1 \\ -1 & 2 \end{pmatrix}.$$

Um die Unsicherheit über die einigermaßen willkürliche Auswahl von u zu reduzieren, überprüft man, wie die ML-Schätzer in Abhängigkeit von u variieren.

Weiters wird der Schätzer

$$\hat{F}(u + y) = \frac{N_u}{n} \left(1 + \hat{\gamma} \frac{y}{\hat{\beta}}\right)^{-1/\hat{\gamma}}$$

grafisch dargestellt und inspiziert.

Beispiel 4.2.8. In Abb. 4.4 wurden die ML-Schätzer von γ der Versicherungsdaten als Funktion von N_u geplottet. Die Schätzer sind für $N_u \in (39, 109)$ halbwegs stabil. \diamond

Berechnung von Risikomaßen VaR und CVaR mit Hilfe der POT-Methode

Seien X_1, X_2, \dots, X_n Beobachtungen von i.i.d. ZVen mit unbekannter Verteilungsfunktion F . Direkt aus der POT-Methode erhält man folgende Schätzer für die Randverteilung und das Quantil $q_p = VaR_p(F)$ von F :

$$\hat{F}(u + y) = \frac{N_u}{n} \left(1 + \hat{\gamma} \frac{y}{\hat{\beta}}\right)^{-1/\hat{\gamma}}$$

$$\hat{q}_p = u + \frac{\hat{\beta}}{\hat{\gamma}} \left(\left(\frac{n}{N_u} (1 - p) \right)^{-\hat{\gamma}} - 1 \right).$$

Für $0 < \hat{\gamma} < 1$ zeigen wir, dass $\widehat{CVaR}_p(F) = \hat{q}_p + \frac{\hat{\beta} + \hat{\gamma}(\hat{q}_p - u)}{1 - \hat{\gamma}}$ ein Schätzer von $CVaR_p(F)$ ist.

Beweis. Der Beweis erfolgt mit Hilfe der folgenden 3 Schritte:

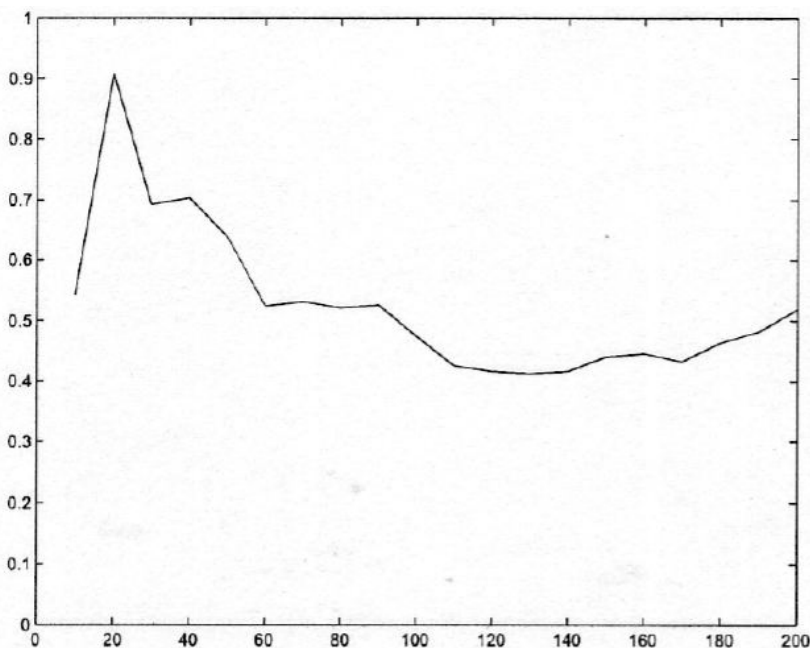


Abbildung 4.4: ML-Schätzer von γ als Funktion von N_u

(1) Sei $Z_1 \sim F$ eine ZV mit Verteilungsfunktion F und seien $u \geq u_0 > 0$ zwei Schwellwerte.

Es gilt $e_{Z_1}(u) = e_{Z_2}(u - u_0)$ wobei $Z_2 \stackrel{d}{=} (Z_1 - u_0) | (Z_1 > u_0)$ und e_{Z_1} und e_{Z_2} die durchschnittlichen Exzessfunktionen von Z_1 bzw. Z_2 sind.

Sei F_2 die Verteilungsfunktion von Z_2 mit

$$\begin{aligned} F_2(x) &= \mathbb{P}(Z_2 < x) = \mathbb{P}(Z_1 - u_0 < x | Z_1 > u_0) = \frac{\mathbb{P}(Z_1 < u_0 + x \wedge Z_1 > u_0)}{\mathbb{P}(Z_1 > u_0)} \\ &= \begin{cases} \frac{\mathbb{P}(u_0 < Z_1 < u_0 + x)}{\mathbb{P}(Z_1 > u_0)} & x > 0 \\ 0 & x \leq 0 \end{cases} \\ &= \begin{cases} \frac{F(u_0 + x) - F(u_0)}{F(u_0)} & x > 0 \\ 0 & x \leq 0 \end{cases} \end{aligned}$$

$$e_{Z_2}(u - u_0) = \frac{1}{\bar{F}_2(u - u_0)} \int_{u - u_0}^{\infty} (y - u + u_0) dF_2(y) \stackrel{zz.}{=} \frac{1}{\bar{F}(u)} \int_u^{\infty} (y - u) dF(y) = e_{Z_1}(u)$$

Für $y > u_0$ gilt $dF_2(y) = \frac{1}{F(u_0)} dF(y + u_0)$ und somit für $u_0 + y = t$:

$$\begin{aligned} \frac{1}{\bar{F}_2(u - u_0)} \int_u^{\infty} (t - u) \frac{dF(t)}{\bar{F}(u_0)} &= \frac{1}{\bar{F}_2(u - u_0) \bar{F}(u_0)} \int_u^{\infty} (t - u) dF(t) \\ &= \frac{1}{\bar{F}(u)} \int_u^{\infty} (t - u) dF(t) = e_{Z_1}(u), \end{aligned}$$

denn

$$\bar{F}_2(u - u_0) = 1 - F_2(u - u_0) = 1 - \frac{F(u) - F(u_0)}{\bar{F}(u_0)} = \frac{\bar{F}(u_0) - F(u) + F(u_0)}{\bar{F}(u_0)} = \frac{1 - F(u)}{\bar{F}(u_0)} = \frac{\bar{F}(u)}{\bar{F}(u_0)}.$$

(2) Sei X eine ZV mit $X \sim GPD_{\gamma,0,\beta}$ und $0 < \gamma < 1$. Es gilt

$$CVaR_p(X) = q_p + \frac{\beta + \gamma q_p}{1 - \gamma},$$

wobei $q_p := VaR_p(X)$ das p -Quantil von X ist.

$$\begin{aligned} CVaR_p(X) &= \frac{1}{1-p} \int_{q_p}^{\infty} x dG_{\gamma,0,\beta}(x) \\ &= \frac{1}{1-p} \left[\int_{q_p}^{\infty} (x - VaR_p(X)) dG_{\gamma,0,\beta}(x) + \underbrace{\int_{q_p}^{\infty} VaR_p(X) dG_{\gamma,0,\beta}(x)}_{=VaR_p(X) \int_{q_p}^{\infty} dG_{\gamma,0,\beta}(x) = VaR_p(X)(1-p)} \right] \\ &= e_{G_{\gamma,0,\beta}}(q_p) + VaR_p(X) = \frac{\beta + \gamma q_p}{1 - \gamma} + q_p \end{aligned}$$

(3) Sei X eine ZV mit $X \sim F$. Die Randverteilung $\bar{F}(x)$ wird durch $\bar{F}(u)\bar{G}_{\gamma,0,\beta}(x-u)$ approximiert. Daraus folgt $F \approx \tilde{F}$ mit $\tilde{F} := 1 - \bar{F}(u)\bar{G}_{\gamma,0,\beta}(x-u)$. Für $q_p > u$ ist der CVaR der Approximation \tilde{F} folgendermaßen gegeben:

$$CVaR_p(\tilde{F}) = q_p + \frac{\beta + \gamma(q_p - u)}{1 - \gamma}$$

$d\tilde{F} = -\bar{F}(u)d\bar{G}_{\gamma,0,\beta}(x-u) = \bar{F}(u)dG_{\gamma,0,\beta}(x-u)$. Damit gilt

$$\begin{aligned} CVaR_p(\tilde{F}) &= \frac{\bar{F}(u)}{\tilde{F}(q_p)} \int_{q_p}^{\infty} t dG_{\gamma,0,\beta}(t-u) \\ &= \frac{\bar{F}(u)}{\tilde{F}(q_p)} \left[\int_{q_p}^{\infty} (t-u) dG_{\gamma,0,\beta}(t-u) + \int_{q_p}^{\infty} u dG_{\gamma,0,\beta}(t-u) \right] \\ &= \frac{\bar{F}(u)}{\underbrace{\tilde{F}(q_p)}_{=\bar{F}(u)\bar{G}_{\gamma,0,\beta}(q_p-u)}} \left[\int_{q_p-u}^{\infty} y dG_{\gamma,0,\beta}(y) + u\bar{G}_{\gamma,0,\beta}(q_p-u) \right] \\ &= \frac{1}{\bar{G}_{\gamma,0,\beta}(q_p-u)} \left[\int_{q_p-u}^{\infty} y dG_{\gamma,0,\beta}(y) + u\bar{G}_{\gamma,0,\beta}(q_p-u) \right] \\ &= CVaR_{p'}(G_{\gamma,0,\beta}) + u \quad \text{mit } p' \text{ so, dass } VaR_{p'}(G_{\gamma,0,\beta}) = q_p - u = q_{p'} \\ &= q_{p'} + \frac{\beta + \gamma q_{p'}}{1 - \gamma} + u = q_p + \frac{\beta + \gamma(q_p - u)}{1 - \gamma} \end{aligned}$$

$$\implies \widehat{CVaR}_p(F) = \hat{q}_p + \frac{\hat{\beta} + \hat{\gamma}(\hat{q}_p - u)}{1 - \hat{\gamma}}. \quad \square$$

Beispiel 4.2.9. In den Abb. 4.5, 4.6 und 4.7 wurden der empirische Tail bzw. seine durch die POT-Methode gefittete Approximation, der 99%-VaR und der 99%-CVaR der Daten des dänischen Versicherungsunternehmens geplottet. \diamond

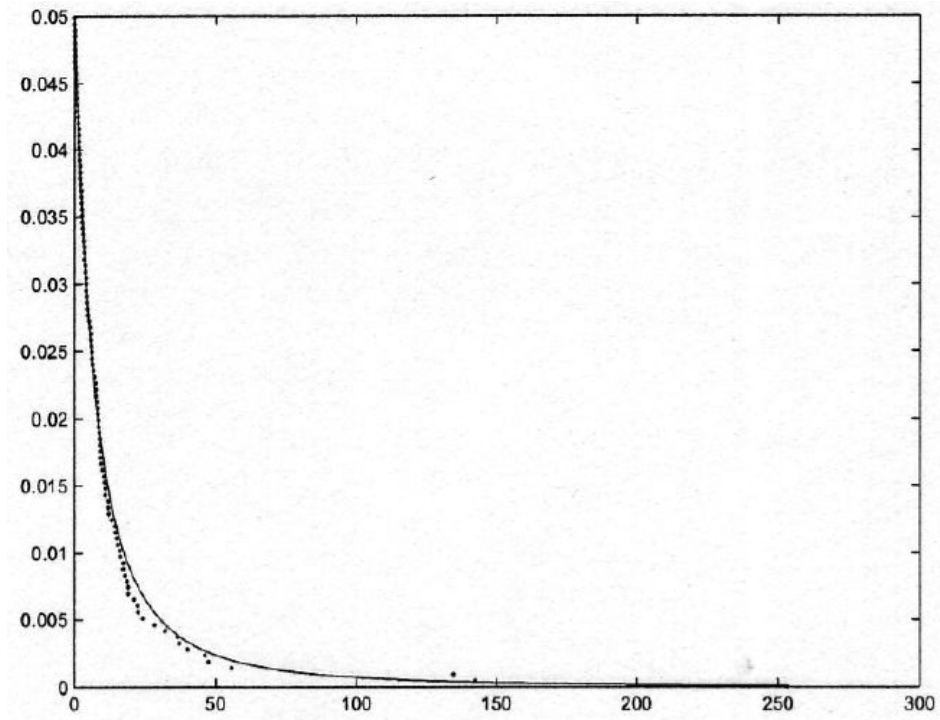


Abbildung 4.5: Empirischer Tail der Daten und die mittels POT gefittete Approximation

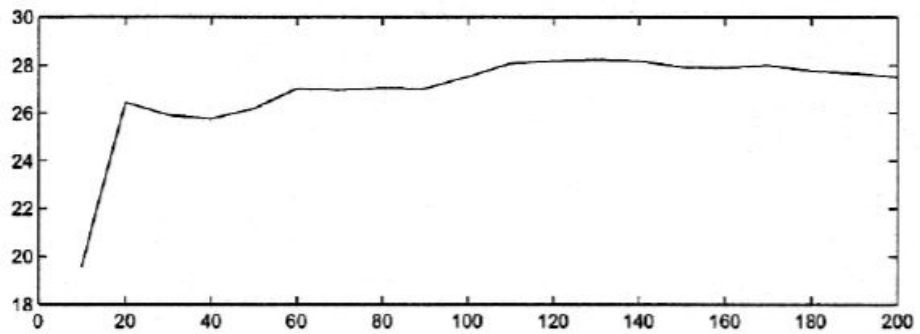


Abbildung 4.6: Geschätzte 99%-Quantile als Funktion von N_u

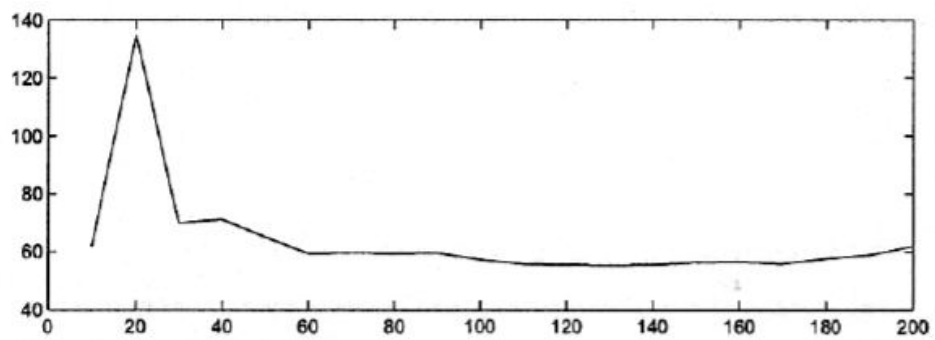


Abbildung 4.7: Geschätzter 99%-CVaR als Funktion von N_u

Kapitel 5

Multivariate Verteilungen

Ziel dieses Kapitels ist es, die Veränderungen der Risikofaktoren zum Zeitpunkt n , $X_n = (X_{n,1}, X_{n,2}, \dots, X_{n,d})$, zu modellieren. Es wird dazu folgende Annahme getroffen: $X_{n,i}$ und $X_{n,j}$ sind abhängig, aber $X_{n,i}$ und $X_{n\pm k,j}$ sind unabhängig für $k \in \mathbb{N}$ ($k \neq 0$), $1 \leq i, j \leq d$.

5.1 Grundlegende Eigenschaften von Zufallsvektoren

Ein d -dimensionaler Zufallsvektor $X = (X_1, X_2, \dots, X_d)^\top$ wird durch die Verteilungsfunktion F spezifiziert

$$F(x) = F(x_1, x_2, \dots, x_d) = \mathbb{P}(X_1 \leq x_1, X_2 \leq x_2, \dots, X_d \leq x_d) = \mathbb{P}(X \leq x).$$

Die i -te Randverteilung F_i von F ist die Verteilungsfunktion von X_i und ist folgendermaßen gegeben:

$$F_i(x_i) = \mathbb{P}(X_i \leq x_i) = F(\infty, \dots, \infty, x_i, \infty, \dots, \infty).$$

Die Verteilungsfunktion F ist stetig, wenn es eine nicht negative Funktion $f \geq 0$ gibt, sodass

$$F(x_1, x_2, \dots, x_d) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} \dots \int_{-\infty}^{x_d} f(u_1, u_2, \dots, u_d) du_1 du_2 \dots du_d.$$

f ist in diesem Fall die Dichte von F .

Die Komponenten von X sind unabhängig dann und nur dann, wenn

$$F(x) = \prod_{i=1}^d F_i(x_i)$$

oder, wenn die Dichten f und f_i , $1 \leq i \leq d$, existieren, dann sind die Komponenten von X d.u.n.d. unabhängig, wenn

$$f(x) = \prod_{i=1}^d f_i(x_i).$$

Ein Zufallsvektor wird durch seine charakteristische Funktion $\phi_X(t)$ eindeutig spezifiziert:

$$\phi_X(t) := \mathbb{E}(\exp\{it^\top X\}), t \in \mathbb{R}^d$$

Beispiel 5.1.1. Für die multivariate Normalverteilung mit Mittelwert μ und Kovarianzmatrix Σ sind die Dichtefunktion f bzw. die charakteristische Funktion ϕ_X folgendermaßen gegeben ($|\Sigma| = |\text{Det}(\Sigma)|$):

$$f(x) = \frac{1}{\sqrt{(2\pi)^d |\Sigma|}} \exp \left\{ -\frac{1}{2} (x - \mu)^\top \Sigma^{-1} (x - \mu) \right\}, x \in \mathbb{R}^d$$

$$\phi_X(t) = \exp \left\{ it^\top \mu - \frac{1}{2} t^\top \Sigma t \right\}, t \in \mathbb{R}^d.$$

◇

Wenn $\mathbb{E}(X_k^2) < \infty$ für alle k , dann ist die Kovarianzmatrix eines Zufallsvektors folgendermaßen gegeben:

$$\text{Cov}(X) = \mathbb{E}((X - \mathbb{E}(X))(X - \mathbb{E}(X))^\top).$$

Anmerkung: Für einen n -dimensionalen Zufallsvektor X , eine konstante Matrix $B \in \mathbb{R}^{n \times n}$ und einen konstanten Vektor $b \in \mathbb{R}^n$ gelten folgende Gleichungen:

$$\mathbb{E}(BX + b) = B\mathbb{E}(X) + b \quad \text{Cov}(BX + b) = B\text{Cov}(X)B^\top.$$

Beispiel 5.1.2. (Portfolio Optimierung, Markowitz-Modell)

Ein Investor möchte in d (risikoreiche) Assets investieren, sodass der erwartete Portfolioertrag μ_p beträgt und die Varianz des Ertrags minimiert wird.

Sei $X = (X_1, X_2, \dots, X_d)^\top$ der Zufallsvektor der Asset>Returns mit $\mathbb{E}(X) = \mu$ und $\text{Cov}(X) = \Sigma$. Die Gewichte des Minimum-Varianz Portfolios sind als Lösung des folgenden quadratischen Optimierungsproblems gegeben:

$$\begin{aligned} \min_w \quad & w^\top \Sigma w \\ \text{sodass} \quad & \\ & w^\top \mu = \mu_p \\ & \sum_{i=1}^d |w_i| = 1 \end{aligned} \tag{5.1}$$

(siehe zB. Campbell et al. (1997)).

Die Restriktion 5.1 erlaubt dem Portfolioverwalter sogenanntes *short selling*. Die Investition von Kapital 1 in long und short Positionen kann auf Folgende Art veranschaulicht werden:

Zeit	long	short
	Investor Markt	Investor Broker Markt
t	$\xrightarrow{w_i \text{ (Geld)}}$ $\xleftarrow{w_i/p_i \text{ (Anleihe)}}$	$\xrightarrow{w_i \text{ (Geld)}}$ $\xleftarrow{w_i/p_i \text{ (Anleihe)}}$ $\xrightarrow{w_i/p_i \text{ (Anleihe)}}$ $\xleftarrow{w_i \text{ (Geld)}}$
t+1	$\xrightarrow{w_i/p_i \text{ (Anleihe)}}$ $\xleftarrow{p'_i w_i/p_i \text{ (Geld)}}$	$\xrightarrow{p'_i w_i/p_i \text{ (Geld)}}$ $\xleftarrow{w_i/p_i \text{ (Anleihe)}}$ $\xrightarrow{w_i/p_i \text{ (Anleihe)}}$ $\xleftarrow{w_i(1 - \alpha_i) \text{ (Geld)}}$

α_i ist eine an den Broker zu bezahlende Margin. Im Markowitz-Modell wird α_i gleich 0 gesetzt. Zum Zeitpunkt $t + 1$ hat der Investor in der long Position $p'_i \frac{w_i}{p_i}$ Geldeinheiten, in der short Position hingegeben

$$w_i - p'_i \frac{w_i}{p_i} + w_i(1 - \alpha_i) = w_i \left(\frac{p_i - p'_i}{p_i} + 1 - \alpha_i \right)$$

Geldeinheiten zur Verfügung. In Summe ergibt das für das gesamte Portfolio

$$\underbrace{\sum_{i \in L} p'_i \frac{w_i}{p_i}}_{\text{Gewinn long}} + \underbrace{\sum_{i \in S} w_i \left(\frac{p_i - p'_i}{p_i} + 1 - \alpha_i \right)}_{\text{Gewinn short}} - \underbrace{\sum_{i \in L \cup S} w_i}_{\text{Gesamtinvestition}} = \sum_{i \in L} w_i \left(\frac{p'_i}{p_i} - 1 \right) + \sum_{i \in S} w_i \left(\frac{p_i - p'_i}{p_i} - \alpha_i \right)$$

und da bei Markowitz die Marge α_i gleich 0 gesetzt wird folgt

$$\sum_{i \in L} w_i \underbrace{\frac{p'_i - p_i}{p_i}}_{x_i} - \sum_{i \in S} w_i \underbrace{\frac{p'_i - p_i}{p_i}}_{x_i}$$

Der Gesamtportfolioertrag ist also

$$P_R = \sum_{i \in L} w_i x_i - \sum_{i \in S} w_i x_i \quad \text{mit} \quad w'_i = \begin{cases} w_i & i \in L \\ -w_i & i \in S \end{cases}$$

und damit

$$P_R = \sum_{i=1}^d w'_i x_i \quad \text{mit} \quad \sum_{i=1}^d |w'_i| = 1.$$

◇

5.2 Probleme bei der Modellierung der Abhängigkeit zwischen Finanzgrößen mit Hilfe der (multivariaten) Normalverteilung

- Finanzgrößen haben i.A. heavier Tails als die Normalverteilung.
- Die Zusammenhänge bei größeren Verlusten sind i.A. stärker als bei „normalen“ Werten. Diese Art von Zusammenhängen kann mit der multivariaten Normalverteilung nicht modelliert werden.

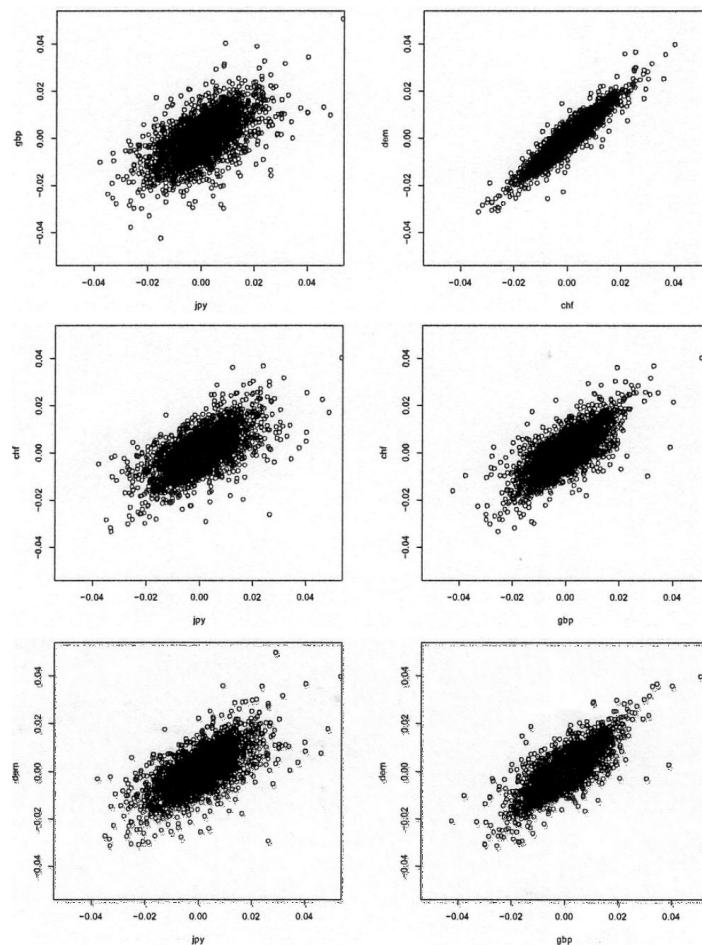


Abbildung 5.1: Log>Returns von Wechselkursen gegen den US-Dollar

Die echten Daten in Abb. 5.1 sind heavier tailed als die simulierten aus Abb. 5.2, was an der „Wolke“ um den Scatterplot ersichtlich wird. Gibt es viele „Ausreißer“ links unten deutet dies auf eine *untere Tailabhängigkeit* hin, sind derartige Ausreißer rechts oben zu finden, lässt das auf eine *obere Tailabhängigkeit* schließen (vgl. später).

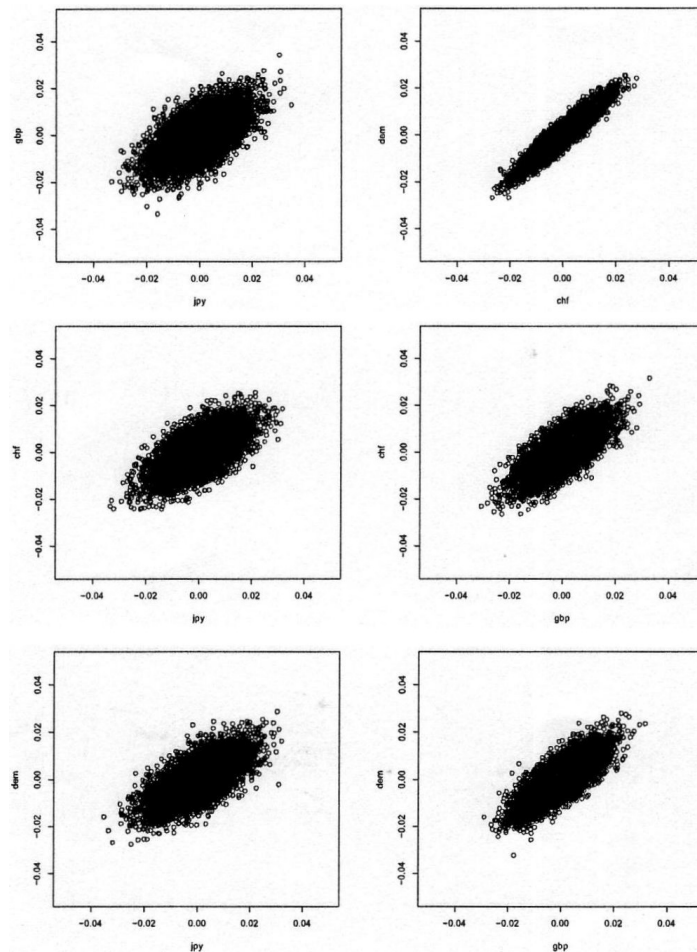


Abbildung 5.2: Mit einer bivariaten Normalverteilung simulierte Wechselkurse (mit geschätztem Mittelwert und Kovarianzmatrix)

5.3 Abhängigkeitsmaße

Seien X_1 und X_2 zwei Zufallsvariablen. Es gibt einige skalare Maße für die Abhängigkeit zwischen X_1 und X_2 .

5.3.1 Lineare Korrelation

Annahme: $Var(X_1), Var(X_2) \in (0, \infty)$. Der Koeffizient der linearen Korrelation $\rho_L(X_1, X_2)$ ist folgendermaßen gegeben:

- $\rho_L(X_1, X_2) = \frac{Cov(X_1, X_2)}{\sqrt{Var(X_1)Var(X_2)}}$
- X_1 und X_2 sind unabhängig $\Rightarrow \rho_L(X_1, X_2) = 0$
- $\rho_L(X_1, X_2) = 0$ impliziert nicht, dass X_1 und X_2 unabhängig sind

Beispiel 5.3.1. Sei $X_1 \sim N(0, 1)$ und $X_2 = X_1^2$. Es gilt $\rho_L(X_1, X_2) = 0$, aber X_1 und X_2 sind klarerweise abhängig. \diamond

Weiters gilt:

$$|\rho_L(X_1, X_2)| = 1 \Leftrightarrow \exists \alpha, \beta \in \mathbb{R}, \beta \neq 0, \text{ sodass } X_2 \stackrel{d}{=} \alpha + \beta X_1$$

$$\text{und } \text{sgn}(\beta) = \text{sgn}(\rho_L(X_1, X_2))$$

Der lineare Korrelationskoeffizient ist eine Invariante unter streng monoton steigenden linearen Transformationen. D.h. für zwei Zufallsvariablen X_1 und X_2 und reelle Konstante $\alpha_1, \alpha_2, \beta_1, \beta_2 \in \mathbb{R}, \beta_1 > 0$ und $\beta_2 > 0$ gilt:

$$\rho_L(\alpha_1 + \beta_1 X_1, \alpha_2 + \beta_2 X_2) = \rho_L(X_1, X_2).$$

Der lineare Korrelationskoeffizient ist jedoch keine Invariante unter streng monoton steigenden nicht-linearen Transformationen.

Beispiel 5.3.2. Seien $X \sim \text{Exp}(\lambda)$, $X_2 = X_1$, und T_1, T_2 zwei streng monoton steigende Transformationen: $T_1(X_1) = X_1$ und $T_2(X_1) = X_1^2$. Dann gilt:

$$\rho_L(X_1, X_1) = 1 \text{ und } \rho_L(T_1(X_1), T_2(X_1)) = \frac{2}{\sqrt{5}}.$$

◇

5.3.2 Rangkorrelation

Die Koeffizienten der Rangkorrelation (Spearman's Rho und Kendall's Tau) sind Maße für die Übereinstimmung von bivariaten Zufallsvektoren.

Seien (x_1, x_2) und $(\tilde{x}_1, \tilde{x}_2)$ zwei Punkte in \mathbb{R}^2 . Die zwei Punkte heißen *übereinstimmend*, wenn $(x_1 - \tilde{x}_1)(x_2 - \tilde{x}_2) > 0$, und *nicht übereinstimmend*, wenn $(x_1 - \tilde{x}_1)(x_2 - \tilde{x}_2) < 0$.

Seien (X_1, X_2) und $(\tilde{X}_1, \tilde{X}_2)$ zwei unabhängige Zufallsvektoren mit gemeinsamer bivariater Verteilung. Kendall's Tau ρ_τ ist definiert als

$$\rho_\tau(X_1, X_2) = \mathbb{P}\left((X_1 - \tilde{X}_1)(X_2 - \tilde{X}_2) > 0\right) - \mathbb{P}\left((X_1 - \tilde{X}_1)(X_2 - \tilde{X}_2) < 0\right).$$

Sei (\hat{X}_1, \hat{X}_2) ein dritter von (X_1, X_2) und $(\tilde{X}_1, \tilde{X}_2)$ unabhängiger Zufallsvektor mit derselben Verteilung wie (X_1, X_2) und $(\tilde{X}_1, \tilde{X}_2)$. Spearman's Rho ρ_S ist definiert als

$$\rho_S(X_1, X_2) = 3 \left\{ \mathbb{P}\left((X_1 - \tilde{X}_1)(X_2 - \hat{X}_2) > 0\right) - \mathbb{P}\left((X_1 - \tilde{X}_1)(X_2 - \hat{X}_2) < 0\right) \right\}.$$

Einige Eigenschaften von ρ_τ und ρ_S :

- $\rho_\tau(X_1, X_2) \in [-1, 1]$ und $\rho_S(X_1, X_2) \in [-1, 1]$.
- Wenn X_1 und X_2 unabhängig sind, dann gilt $\rho_\tau(X_1, X_2) = \rho_S(X_1, X_2) = 0$. Die Umkehrung gilt i.A. nicht.
- Sei $T: \mathbb{R} \rightarrow \mathbb{R}$ eine streng monoton steigende Funktion, dann gilt:

$$\rho_\tau(T(X_1), T(X_2)) = \rho_\tau(X_1, X_2)$$

$$\rho_S(T(X_1), T(X_2)) = \rho_S(X_1, X_2)$$

Beweis. 1) und 2) sind trivial. Beweis von 3) erfolgt mit Hilfe von Copulas (später). □

5.3.3 Tail-Abhängigkeit

Definition 5.3.3. Sei $(X_1, X_2)^\top$ ein Zufallsvektor mit Randverteilungen F_1 und F_2 . Der *Koeffizient der oberen Tail-Abhängigkeit* von $(X_1, X_2)^\top$ wird folgendermaßen definiert:

$$\lambda_U(X_1, X_2) = \lim_{u \rightarrow 1^-} \mathbb{P}(X_2 > F_2^{\leftarrow}(u) | X_1 > F_1^{\leftarrow}(u)),$$

vorausgesetzt der Limes existiert.

Der *Koeffizient der unteren Tail-Abhängigkeit* von $(X_1, X_2)^\top$ wird folgendermaßen definiert:

$$\lambda_L(X_1, X_2) = \lim_{u \rightarrow 0^+} \mathbb{P}(X_2 \leq F_2^{\leftarrow}(u) | X_1 \leq F_1^{\leftarrow}(u)),$$

vorausgesetzt der Limes existiert.

Wenn $\lambda_U > 0$ ($\lambda_L > 0$), heißt es $(X_1, X_2)^\top$ hat eine obere (untere) Tail-Abhängigkeit.

(Siehe Joe 1997, Schmidt und Stadtmüller 2002)

Beispiel 5.3.4. Sei $X_1 \sim \text{Exp}(\lambda)$ und $X_2 = X_1^2$. Bestimmen Sie $\lambda_U(X_1, X_2)$, $\lambda_L(X_1, X_2)$ und zeigen Sie, dass $(X_1, X_2)^\top$ eine obere und eine untere Tail-Abhängigkeit hat. Berechnen Sie auch den linearen Korrelationskoeffizienten $\rho_L(X_1, X_2)$.

$$F_1(t) = \begin{cases} 1 - e^{-\lambda t} & t > 0 \\ 0 & t \leq 0 \end{cases} \quad F_1^{\leftarrow}(u) = \begin{cases} \frac{\ln(1-u)}{-\lambda} & u \in [0, 1) \\ \infty & u = 1 \end{cases}$$

$$F_2(t) = \begin{cases} \mathbb{P}(X_1^2 \leq t) = \int_0^{\sqrt{t}} \lambda e^{-\lambda x} dx = 1 - e^{-\lambda\sqrt{t}} & t > 0 \\ 0 & t \leq 0 \end{cases} \quad F_2^{\leftarrow}(u) = \begin{cases} \frac{\ln^2(1-u)}{\lambda^2} & u \in [0, 1) \\ \infty & u = 1 \end{cases}$$

$$\lambda_U(X_1, X_2) = \lim_{u \rightarrow 1^-} \mathbb{P}(X_2 > F_2^{\leftarrow}(u) | X_1 > F_1^{\leftarrow}(u)) = \lim_{u \rightarrow 1^-} \mathbb{P}\left(X_1^2 > \frac{\ln^2(1-u)}{\lambda^2} | X_1 > \frac{\ln(1-u)}{-\lambda}\right) = 1$$

$$\lambda_L(X_1, X_2) = \lim_{u \rightarrow 0^+} \mathbb{P}(X_2 \leq F_2^{\leftarrow}(u) | X_1 \leq F_1^{\leftarrow}(u)) = \lim_{u \rightarrow 0^+} \mathbb{P}\left(X_1^2 \leq \frac{\ln^2(1-u)}{\lambda^2} | X_1 \leq \frac{\ln(1-u)}{-\lambda}\right) = 1$$

\Rightarrow sehr starke Tail-Abhängigkeit, trotz $\rho_L = 2/\sqrt{5}$.

Explizite Berechnung von $\rho_L(X_1, X_2) = \frac{\text{Cov}(X_1, X_2)}{\sqrt{\text{Var}(X_1)\text{Var}(X_2)}}$:

$$\text{Var}(X_1) = \frac{1}{\lambda^2} \quad \text{Var}(X_2) = \int_0^\infty (X_2 - \mathbb{E}(X_2))^2 \lambda e^{-\lambda X_1} dX_1 = \int_0^\infty \left(x^2 - \frac{2}{\lambda}\right)^2 \lambda e^{-\lambda x} dx = \dots = \frac{20}{\lambda^4}$$

wobei $\mathbb{E}(X_2) = \int_0^\infty x^2 \lambda e^{-\lambda x} dx = \frac{2}{\lambda^2}$ und

$$\text{Cov}(X_1, X_2) = \int_0^\infty \underbrace{(X_1 - \mathbb{E}(X_1))}_{1/\lambda} \underbrace{(X_1^2 - \mathbb{E}(X_1^2))}_{2/\lambda^2} \lambda e^{-\lambda X_1} dX_1 = \frac{4}{\lambda^3}.$$

Insgesamt gilt also

$$\rho_L(X_1, X_2) = \frac{4/\lambda^3}{\sqrt{1/\lambda^2 \cdot 20/\lambda^4}} = \frac{4\sqrt{\lambda^6}}{2\lambda^3\sqrt{5}} = \frac{2}{\sqrt{5}}.$$

◇

Beispiel 5.3.5. Graphische Darstellung zweier bivariater Verteilungen mit linearer Korrelation $\rho_L = 0.8$ und standard-normalverteilten Randverteilungen.

Im linken Plot von Abb. 5.3 ist eine obere und untere Tail-Abhängigkeit von $\lambda_U = \lambda_L = 0.49$ zu erkennen, im rechten keine ($\lambda_U = \lambda_L = 0$). \diamond

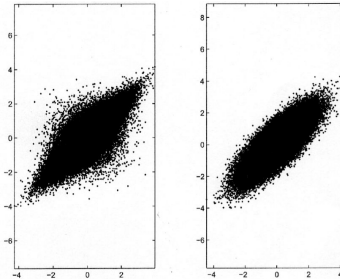


Abbildung 5.3: Zwei bivariate Verteilungen mit linearer Korrelation $\rho_L = 0.8$ und standard-normalverteilten Randverteilungen

5.4 Multivariate elliptische Verteilungen

5.4.1 Die multivariate Normalverteilung

Definition 5.4.1. Der Zufallsvektor $(X_1, X_2, \dots, X_d)^\top$ hat eine *multivariate Normalverteilung* (oder eine *multivariate Gauss'sche Verteilung*), wenn $X \stackrel{d}{=} \mu + AZ$, wobei $Z = (Z_1, Z_2, \dots, Z_k)^\top$ ein Vektor von i.i.d. normalverteilten ZV ($Z_i \sim N(0, 1), \forall i = 1, 2, \dots, k$), $A \in \mathbb{R}^{d \times k}$ eine konstante Matrix und $\mu \in \mathbb{R}^d$ ein konstanter Vektor ist.

Für so einen Zufallsvektor X gilt: $\mathbb{E}(X) = \mu$, $Cov(X) = \Sigma = AA^\top$ (Σ positiv semidefinit).

Notation: $X \sim N_d(\mu, \Sigma)$.

Satz 5.1. (Multivariate Normalverteilung: äquiv. Definitionen)

1. $X \sim N_d(\mu, \Sigma)$ für einen Vektor $\mu \in \mathbb{R}^d$ und eine positiv semidefinite Matrix $\Sigma \in \mathbb{R}^{d \times d}$ dann und nur dann, wenn $\forall a \in \mathbb{R}^d, a = (a_1, a_2, \dots, a_d)^\top$, die Zufallsvariable $a^\top X$ normalverteilt ist.
2. Ein Zufallsvektor $X \in \mathbb{R}^d$ ist multivariat normalverteilt dann und nur dann, wenn seine charakteristische Funktion folgendermaßen gegeben ist:

$$\phi_X(t) = \mathbb{E}(\exp\{it^\top X\}) = \exp\left\{it^\top \mu - \frac{1}{2}t^\top \Sigma t\right\}$$

für einen Vektor $\mu \in \mathbb{R}^d$ und eine positiv semidefinite Matrix $\Sigma \in \mathbb{R}^{d \times d}$.

3. Ein Zufallsvektor $X \in \mathbb{R}^d$ mit $\mathbb{E}(X) = \mu$ und $Cov(X) = \Sigma$, wobei die Determinante von Σ positiv ist ($\det(\Sigma) := |\Sigma| > 0$), ist normalverteilt, d.h. $X \sim N_d(\mu, \Sigma)$, dann und nur dann, wenn seine Dichtefunktion folgendermaßen gegeben ist

$$f_X(x) = \frac{1}{\sqrt{(2\pi)^d |\Sigma|}} \exp\left\{-\frac{(x - \mu)^\top \Sigma^{-1} (x - \mu)}{2}\right\}.$$

Beweis. (siehe zB. Gut 1995)

□

Satz 5.2. (Eigenschaften der multivariaten Normalverteilung)

Für $X \sim N_d(\mu, \Sigma)$ gilt:

Lineare Kombinationen:

Für $B \in \mathbb{R}^{k \times d}$ und $b \in \mathbb{R}^k$. Es gilt dann $BX + b \in N_k(B\mu + b, B\Sigma B^\top)$.

Randverteilungen:

Setze $X^\top = (X^{(1)\top}, X^{(2)\top})$ für $X^{(1)\top} = (X_1, X_2, \dots, X_k)^\top$ und $X^{(2)\top} = (X_{k+1}, X_{k+2}, \dots, X_d)^\top$ und analog

$$\mu^\top = (\mu^{(1)\top}, \mu^{(2)\top}) \quad \text{und} \quad \Sigma = \begin{pmatrix} \Sigma^{(1,1)} & \Sigma^{(1,2)} \\ \Sigma^{(2,1)} & \Sigma^{(2,2)} \end{pmatrix}.$$

Es gilt dann $X^{(1)} \sim N_k(\mu^{(1)}, \Sigma^{(1,1)})$ und $X^{(2)} \sim N_{d-k}(\mu^{(2)}, \Sigma^{(2,2)})$.

Bedingte Verteilungen:

Wenn Σ regulär, dann ist auch der bedingte Zufallsvektor $X^{(2)} | X^{(1)} = x^{(1)}$ multivariat normalverteilt

$$X^{(2)} | X^{(1)} = x^{(1)} \sim N_{d-k}(\mu^{(2,1)}, \Sigma^{(22,1)}) \quad \text{wobei}$$

$$\mu^{(2,1)} = \mu^{(2)} + \Sigma^{(2,1)} (\Sigma^{(1,1)})^{-1} (x^{(1)} - \mu^{(1)}) \quad \text{und}$$

$$\Sigma^{(22,1)} = \Sigma^{(2,2)} - \Sigma^{(2,1)} (\Sigma^{(1,1)})^{-1} \Sigma^{(1,2)}.$$

Quadratische Formen:

Wenn Σ regulär, dann gilt $D^2 = (X - \mu)^\top \Sigma^{-1} (X - \mu) \sim \chi_d^2$.

Die Zufallsvariable D heißt Mahalanobis Distanz.

Faltung:

Seien $X \sim N_d(\mu, \Sigma)$ und $Y \sim N_d(\tilde{\mu}, \tilde{\Sigma})$ zwei unabhängige Zufallsvektoren. Es gilt dann $X + Y \sim N_d(\mu + \tilde{\mu}, \Sigma + \tilde{\Sigma})$.

5.4.2 Varianz-gemischte Normalverteilungen

Definition 5.4.2. Ein Zufallsvektor $X \in \mathbb{R}^d$ hat eine *multivariate Varianz-gemischte Normalverteilung*, wenn $X \stackrel{d}{=} \mu + WAZ$, wobei: $Z \sim N_k(0, I)$, $W \geq 0$ eine von Z unabhängige positive Zufallsvariable, $\mu \in \mathbb{R}^d$ ein konstanter Vektor, $A \in \mathbb{R}^{d \times k}$ eine konstante Matrix und I die Einheitsmatrix ist.

Unter der Bedingung $W = w$ ist X normalverteilt: $X \sim N_d(\mu, w^2\Sigma)$, wobei $\Sigma = AA^\top$.

$$\mathbb{E}(X) = \mu \quad \text{und} \quad \text{Cov}(X) = \mathbb{E}(W^2 A Z Z^\top A^\top) = \mathbb{E}(W^2) \Sigma \quad \text{falls} \quad \mathbb{E}(W^2) < \infty.$$

Beispiel 5.4.3. Die multivariate t_α -Verteilung:

Sei $Y \sim IG(\alpha, \beta)$ (Inverse Gamma-Verteilung) mit Dichtefunktion:

$$f_{\alpha, \beta}(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{-(\alpha+1)} \exp(-\beta/x) \quad x > 0, \alpha > 0, \beta > 0.$$

Dann gilt:

$$\mathbb{E}(Y) = \frac{\beta}{\alpha - 1} \text{ für } \alpha > 1, \text{ Var}(Y) = \frac{\beta^2}{(\alpha - 1)^2(\alpha - 2)} \text{ für } \alpha > 2$$

Sei $W^2 \sim IG(\alpha/2, \alpha/2)$. Dann ist die Verteilung von $X = \mu + WAZ$ eine multivariate t_α -Verteilung mit α Freiheitsgraden: $X \sim t_d(\alpha, \mu, \Sigma)$.

$$\text{Cov}(X) = \mathbb{E}(W^2)\Sigma = \frac{\alpha}{\alpha - 2}\Sigma$$

◇

5.4.3 Sphärische Verteilungen

Definition 5.4.4. Ein Zufallsvektor $X = (X_1, X_2, \dots, X_d)^\top$ hat eine *sphärische Verteilung*, wenn für jede orthogonale Matrix $U \in \mathbb{R}^{d \times d}$ die Gleichung $UX \stackrel{d}{=} X$ gilt.

Satz 5.3. Die folgenden Aussagen sind äquivalent.

1. Der Zufallsvektor $X \in \mathbb{R}^d$ hat eine sphärische Verteilung.
2. Es existiert eine Funktion $\psi: \mathbb{R} \rightarrow \mathbb{R}$, sodass die charakteristische Funktion von X folgendermaßen gegeben ist:

$$\phi_X(t) = \psi(t^\top t) = \psi(t_1^2 + t_2^2 + \dots + t_d^2)$$

3. Für jeden Vektor $a \in \mathbb{R}^d$ gilt $a^\top X \stackrel{d}{=} \|a\|X_1$ wobei $\|a\|^2 = a_1^2 + a_2^2 + \dots + a_d^2$.
4. X lässt sich als $X \stackrel{d}{=} RS$ repräsentieren, wobei der Zufallsvektor $S \in \mathbb{R}^d$ gleichverteilt auf der Einheitskugel \mathcal{S}^{d-1} , $\mathcal{S}^{d-1} = \{x \in \mathbb{R}^d: \|x\| = 1\}$, ist und $R \geq 0$ eine von S unabhängige ZV ist.

Notation: $X \sim S_d(\psi)$

Beispiel 5.4.5. Normalverteilungen sind sphärische Verteilungen.

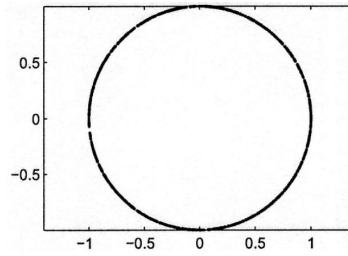
Sei $X \sim N_d(0, I)$. Dann ist $X \sim S_d(\psi)$ mit $\psi = \exp(-x/2)$.

Tatsächlich gilt: $\phi_X(t) = \exp\{it^\top 0 - \frac{1}{2}t^\top I t\} = \exp\{-t^\top t/2\} = \psi(t^\top t)$.

Sei $X = RS$ die stochastische Darstellung von $X \sim N_d(0, I)$. Es gilt $\|X\|^2 \stackrel{d}{=} R^2 \sim \chi_d^2$. ◇

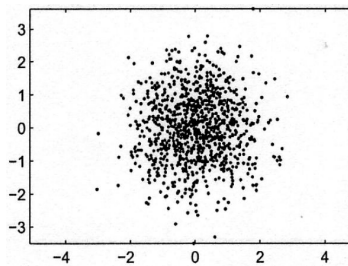
Simulation einer sphärischen Verteilung:

- (i) Simuliere s aus einem gleichmäßig verteilten Zufallsvektor in \mathcal{S}^{d-1} (zB. indem y aus einer multivariaten Standardnormalverteilung $Y \sim N_d(0, I)$ simuliert und $s = y/\|y\|$ gesetzt wird).



(ii) Simuliere r aus R .

(iii) Setze $x = rs$.



5.4.4 Elliptische Verteilungen

Definition 5.4.6. Ein Zufallsvektor $X \in \mathbb{R}^d$ hat eine *elliptische Verteilung*, wenn $X \stackrel{d}{=} \mu + AY$, wobei $Y \sim S_k(\psi)$, $\mu \in \mathbb{R}^d$ ein konstanter Vektor und $A \in \mathbb{R}^{d \times k}$ eine konstante Matrix ist.

Die charakteristische Funktion ist

$$\begin{aligned} \phi_X(t) &= \mathbb{E}(\exp\{it^\top X\}) = \mathbb{E}(\exp\{it^\top(\mu + AY)\}) = \exp\{it^\top \mu\} \mathbb{E}(\exp\{i(A^\top t)^\top Y\}) \\ &= \exp\{it^\top \mu\} \psi(t^\top \Sigma t), \end{aligned}$$

wobei $\Sigma = AA^\top$.

Notation: $X \sim E_d(\mu, \Sigma, \psi)$.

μ heißt Positionsparameter (*location parameter*), Σ heißt Dispersionsparameter (*dispersion parameter*) und ψ heißt charakteristischer Generator der elliptischen Verteilung.

Falls $A \in \mathbb{R}^{d \times d}$ regulär, dann gilt folgende Relation zwischen elliptischen und sphärischen Verteilungen:

$$X \sim E_d(\mu, \Sigma, \psi) \Leftrightarrow A^{-1}(X - \mu) \sim S_d(\psi), \quad A \in \mathbb{R}^{d \times d}, AA^\top = \Sigma.$$

Satz 5.4. (*Stochastische Darstellung der elliptischen Verteilung*)

Sei $X \in \mathbb{R}^d$ ein d -dimensionaler Zufallsvektor.

$X \sim E_d(\mu, \Sigma, \psi)$ dann und nur dann, wenn $X \stackrel{d}{=} \mu + RAS$, wobei $S \in \mathbb{R}^k$ ein auf der Einheitskugel \mathcal{S}^{k-1} gleichverteilter Zufallsvektor, $R \geq 0$ eine von S unabhängige Zufallsvariable, $A \in \mathbb{R}^{d \times k}$ eine konstante Matrix ($\Sigma = AA^\top$) und $\mu \in \mathbb{R}^d$ ein konstanter Vektor ist.

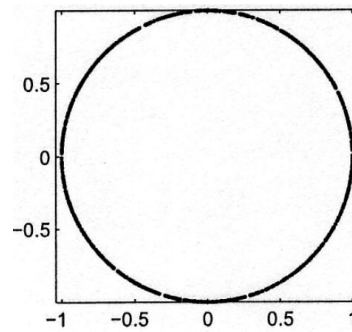
Beweis.

$$X = \mu + A \underbrace{Y}_{S_k(\psi)} = \mu + A \underbrace{R}_{\geq 0} \underbrace{S}_{\sim U(S^{k-1})} = \mu + RAS$$

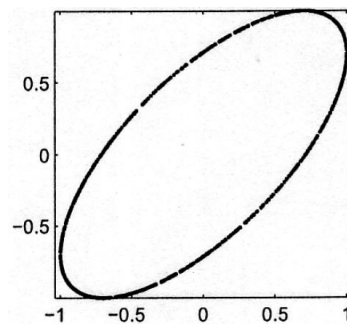
□

Simulation einer elliptischen Verteilung:

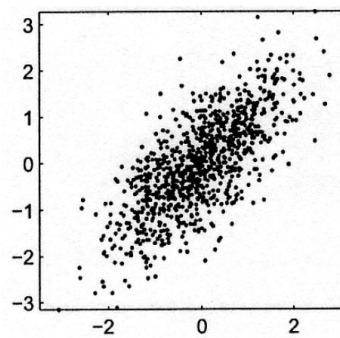
- (i) Simuliere s aus einem gleichmäßig verteilten Zufallsvektor in \mathcal{S}^{d-1} (zB. indem y aus einer multivariaten Standardnormalverteilung $Y \sim N_d(0, I)$ simuliert und $s = y/\|y\|$ gesetzt wird).



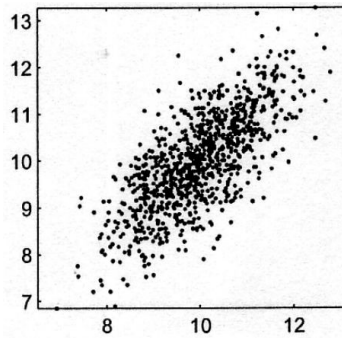
- (ii) Multipliziere jedes s_i mit A .



- (iii) Simuliere r aus R und multipliziere es mit As .



- (iv) Setze $x = \mu + rAs$.



Beispiel 5.4.7. Multivariate Normalverteilung:

Sei $X \sim N_d(\mu, \Sigma)$. Es existiert eine Matrix $A \in \mathbb{R}^{d \times k}$, sodass $X \stackrel{d}{=} \mu + AZ$, wobei $Z \in N_k(0, I)$ und $AA^\top = \Sigma$. Weiters gilt $Z = RS$, wobei S ein gleichmäßig verteilter Zufallsvektor in \mathcal{S}^{k-1} ist und $R^2 \sim \chi_k^2$. Daraus folgt $X \stackrel{d}{=} \mu + RAS$ und daher $X \sim E_d(\mu, \Sigma, \psi)$ mit $\psi(x) = \exp\{-x/2\}$. \diamond

Beispiel 5.4.8. Multivariate Varianz-gemischte Normalverteilung:

Sei $Z \sim N_d(0, I)$ ein normalverteilter Zufallsvektor. Z ist sphärisch-verteilt mit stochastischer Darstellung $Z \stackrel{d}{=} VS$, wobei $V^2 = \|Z\|^2 \sim \chi_d^2$. Sei $X = \mu + WAZ$ eine Varianz-gemischte Normalverteilung. Dann gilt $X \stackrel{d}{=} \mu + VWAS$ wobei $V^2 \sim \chi_d^2$ und VW eine nicht-negative von S unabhängige ZV ist. D.h., X ist elliptisch verteilt mit $R = VW$. \diamond

Satz 5.5. (*Eigenschaften der elliptischen Verteilung*)

Sei $X \sim E_k(\mu, \Sigma, \psi)$. X hat folgende Eigenschaften:

Lineare Kombinationen:

Für $B \in \mathbb{R}^{k \times d}$ und $b \in \mathbb{R}^k$ gilt:

$$BX + b \in E_k(B\mu + b, B\Sigma B^\top, \psi).$$

Randverteilungen:

Setze $X^\top = (X^{(1)\top}, X^{(2)\top})$ für $X^{(1)\top} = (X_1, X_2, \dots, X_n)^\top$ und $X^{(2)\top} = (X_{n+1}, X_{n+2}, \dots, X_k)^\top$ und analog $\mu^\top = (\mu^{(1)\top}, \mu^{(2)\top})$ sowie $\Sigma = \begin{pmatrix} \Sigma^{(1,1)} & \Sigma^{(1,2)} \\ \Sigma^{(2,1)} & \Sigma^{(2,2)} \end{pmatrix}$.
Es gilt dann $X_1 \sim E_n(\mu^{(1)}, \Sigma^{(1,1)}, \psi)$ und $X_2 \sim E_{k-n}(\mu^{(2)}, \Sigma^{(2,2)}, \psi)$.

Bedingte Verteilungen:

Wenn Σ regulär, dann ist auch die bedingte Verteilung $X^{(2)} | X^{(1)} = x^{(1)}$ elliptisch verteilt:

$$X^{(2)} | X^{(1)} = x^{(1)} \sim E_{k-n}(\mu^{(2,1)}, \Sigma^{(22,1)}, \tilde{\psi})$$

wobei

$$\mu^{(2,1)} = \mu^{(2)} + \Sigma^{(2,1)} \left(\Sigma^{(1,1)} \right)^{-1} \left(x^{(1)} - \mu^{(1)} \right)$$

und

$$\Sigma^{(22,1)} = \Sigma^{(2,2)} - \Sigma^{(2,1)} \left(\Sigma^{(1,1)} \right)^{-1} \Sigma^{(1,2)}.$$

Typischerweise sind $\tilde{\psi}$ und ψ unterschiedlich (siehe Fang, Katz und Ng 1987).

Quadratische Formen:

Wenn Σ regulär ist, dann gilt

$$D^2 = (X - \mu)^\top \Sigma^{-1} (X - \mu) \sim R^2.$$

wobei R die nicht-negative ZV aus der stochastischen Darstellung $Y = RS$ der sphärischen Verteilung Y mit $S \sim U(S^{(d-1)})$ und $X = \mu + AY$ ist. Die Zufallsvariable D heißt Mahalanobis Distanz.

Faltung:

Seien $X \sim E_k(\mu, \Sigma, \psi)$ und $Y \sim E_k(\tilde{\mu}, \Sigma, \tilde{\psi})$ zwei unabhängige Zufallsvektoren. Es gilt dann $X + Y \sim E_k(\mu + \tilde{\mu}, \Sigma, \bar{\psi})$ wobei $\bar{\psi} = \psi\tilde{\psi}$.

Achtung: Σ muss i.A. dieselbe für X und Y sein.

Anmerkung: Aus $X \sim E_k(\mu, I_k, \psi)$ folgt nicht, dass die Komponenten von X unabhängig sind. Die Komponenten von X sind dann und nur dann unabhängig, wenn X multivariat normalverteilt mit der Einheitsmatrix als Kovarianzmatrix ist.

Beweis. Lineare Kombinationen:

$$\begin{aligned}\phi_{BX+b} &= \mathbb{E}(\exp\{it^\top(BX + b)\}) = \mathbb{E}(\exp\{it^\top b\} \exp\{it^\top BX\}) \\ &= \exp\{it^\top b\} \mathbb{E}(\exp\{i(B^\top t)^\top X\}) = \exp\{it^\top b\} \underbrace{\exp\{i(B^\top t)^\top \mu\} \psi(t^\top B \Sigma B^\top t)}_{\phi_X(B^\top t)} \\ &= \exp\{it^\top(b + B\mu)\} \psi(t^\top B \Sigma B^\top t)\end{aligned}$$

$$\Rightarrow BX + b \sim E_k(B\mu + b, B\Sigma B^\top, \psi).$$

Faltung: $X \sim E_k(\mu, \Sigma, \psi)$, $Y \sim E_k(\tilde{\mu}, \Sigma, \tilde{\psi})$.

$$\begin{aligned}\phi_{X+Y}(t) &= \mathbb{E}(\exp\{it^\top(X + Y)\}) = \mathbb{E}(\exp\{it^\top X\} \exp\{it^\top Y\}) \\ &\stackrel{u.a.}{=} \mathbb{E}(\exp\{it^\top X\}) \mathbb{E}(\exp\{it^\top Y\}) = \exp\{it^\top \mu\} \psi(t^\top \Sigma t) \exp\{it^\top \tilde{\mu}\} \tilde{\psi}(t^\top \Sigma t) \\ &= \exp\{it^\top(\mu + \tilde{\mu})\} \psi \tilde{\psi}(t^\top \Sigma t)\end{aligned}$$

$$\Rightarrow X + Y \sim E_k(\mu + \tilde{\mu}, \Sigma, \psi \tilde{\psi}). \quad \square$$

5.5 Koherente Risikomaße

Sei (Ω, \mathcal{F}, P) ein Wahrscheinlichkeitsraum mit Ereignismenge Ω , Ereignisalgebra \mathcal{F} und Wahrscheinlichkeitsmaß P .

Sei $L^{(0)}(\Omega, \mathcal{F}, P)$ die Menge aller Zufallsgrößen aus (Ω, \mathcal{F}) , die fast sicher endlich sind.

Sei $M \subseteq L^{(0)}$ die Menge, die die Zufallsvariablen der Verluste beinhaltet.

Sei $\rho: M \rightarrow \mathbb{R}$ ein Risikomaß in M .

Definition 5.5.1. Ein Risikomaß ρ , das folgende Eigenschaften besitzt, heißt *koherent* auf M :

(C1) Invarianz bzgl. Translation:

$$\rho(X + r) = \rho(X) + r, \text{ für jede Konstante } r \text{ und jedes } X \in M.$$

(C2) Subadditivität:

$$\forall X_1, X_2 \in M \text{ gilt } \rho(X_1 + X_2) \leq \rho(X_1) + \rho(X_2).$$

(C3) Positive Homogenität:

$$\rho(\lambda X) = \lambda \rho(X), \forall \lambda \geq 0, \forall X \in M.$$

(C4) Monotonie:

$$\forall X_1, X_2 \in M \text{ gilt } X_1 \stackrel{f.s.}{\leq} X_2 \implies \rho(X_1) \leq \rho(X_2).$$

Anmerkungen:

- ad C(2): Diese Eigenschaft macht unter der Annahme „Diversifikation mindert das Risiko“ Sinn.
- ad C(3): Die positive Homogenität macht wirtschaftlich nur für $\lambda \in \mathbb{N}$ Sinn, $\lambda \in \mathbb{R}$ hat technische Gründe.

5.6 Konvexe Risikomaße

Betrachte die Eigenschaft

(C5) Konvexität:
 $\forall X_1, X_2 \in M, \forall \lambda \in [0, 1]$ gilt

$$\rho(\lambda X_1 + (1 - \lambda)X_2) \leq \lambda \rho(X_1) + (1 - \lambda)\rho(X_2).$$

(C5) ist schwächer als (C2) und (C3), d.h. (C2) und (C3) zusammen implizieren (C5), aber nicht umgekehrt, denn

$$\rho(\lambda X_1 + (1 - \lambda)X_2) \stackrel{(C2)}{\leq} \rho(\lambda X_1) + \rho((1 - \lambda)X_2) \stackrel{(C3)}{=} \lambda \rho(X_1) + (1 - \lambda)\rho(X_2) \quad \forall \lambda \in [0, 1].$$

Definition 5.6.1. Ein Risikomaß ρ , das die Eigenschaften (C1),(C4) und (C5) besitzt, heißt *konvex* auf M .

Beispiel 5.6.2. VaR ist nicht kohärent:

Sei das Wahrscheinlichkeitsmaß \mathbb{P} durch eine beliebige kontinuierliche oder diskrete Wahrscheinlichkeitsverteilung F definiert.

$VaR_\alpha(F) = F^{\leftarrow}(\alpha)$ besitzt die Eigenschaften (C1), (C3) und (C4).

(C4): $F_2(x) = \mathbb{P}(X_2 \leq x) \leq \mathbb{P}(X_1 \leq x) = F_1(x)$, weil $X_1 \stackrel{f.s.}{\leq} X_2 \Rightarrow (X_2 \leq x \stackrel{f.s.}{\Rightarrow} X_1 \leq x)$ (vgl. Abb. 5.4).

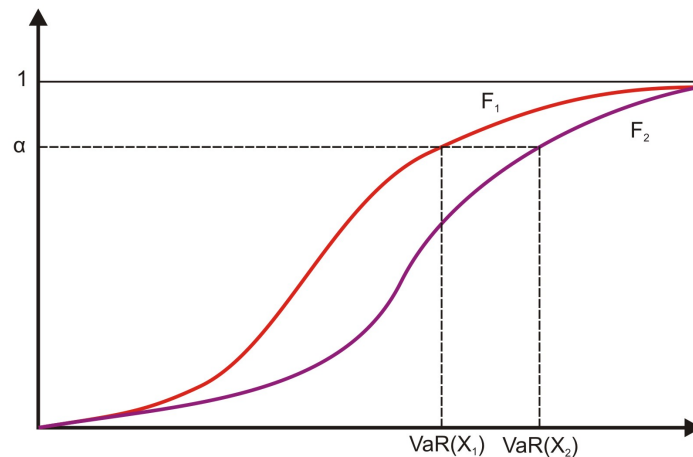


Abbildung 5.4: Graphische Veranschaulichung von Eigenschaft (C4) für VaR_α

Sei das Wahrscheinlichkeitsmaß \mathbb{P} durch die Binomialverteilung $B(p, n)$ für $n \in \mathbb{N}$, $p \in (0, 1)$, definiert.

$VaR_\alpha(B(p, n))$ ist nicht subadditiv!

Um dies zu zeigen, wird der VaR der Verluste eines Bond-Portfolios berechnet, das aus 100 Bonds besteht die unabhängig voneinander mit Wahrscheinlichkeit $p = 0.02$ defaultieren. Beobachten Sie, dass dieser Wert größer als das Hundertfache des VaR des Verlustes eines einzigen Bonds ist:

$$\begin{aligned} P_i^t &= 100 \quad \forall 1 \leq i \leq 100 \quad (\text{Preis zum Zeitpunkt } t). \\ P_i^{t+1} &= \begin{cases} 105 & \text{wenn kein Default} \\ 0 & \text{sonst} \end{cases} \quad \forall i. \\ L_i &= \begin{cases} -5 & \text{kein Default} \\ 100 & \text{Default} \end{cases} \quad \dots \text{Bondverlust.} \\ Y_i &= \begin{cases} 1 & \text{Bond } i \text{ defaultiert in } [t, t+1] \\ 0 & \text{sonst} \end{cases} \end{aligned}$$

Der Verlust von Bond i ist demnach $L_i = 100Y_i - 5(1 - Y_i) \forall i$ und es gilt, dass die L_i i.i.d. sind.

Portfolio A bestehe aus 100 Stück von Bond 1, Portfolio B aus je 1 Stück von Bond i , $1 \leq i \leq 100$.

Die entsprechenden Verluste sind $L_A = 100L_1$ und $L_B = \sum_{i=1}^{100} L_i = 105 \sum_{i=1}^{100} Y_i - 500$.

Es gilt also

$$\begin{aligned} VaR_{0.95}(L_A) &= 100 \underbrace{VaR_{0.95}(L_1)}_{-5} = -500 \\ VaR_{0.95}(L_B) &= 105 VaR_{0.95}\left(\underbrace{\sum_{i=1}^{100} Y_i}_{\sim Bin(100, 0.02)}\right) - 500 = 105 \cdot 5 - 500 = 25 \end{aligned}$$

weil

$$\begin{aligned} \mathbb{P}\left(\sum_{i=1}^{100} Y_i \leq 5\right) &\approx 0.984 \geq 0.95 \\ \mathbb{P}\left(\sum_{i=1}^{100} Y_i \leq 4\right) &\approx 0.949 < 0.95 \end{aligned}$$

Das ergibt einen Widerspruch zur Subadditivität, da der VaR im nicht-diversifizierten Fall beträchtlich kleiner als im diversifizierten Fall ist. \diamond

Satz 5.6. Sei (Ω, \mathcal{F}, P) ein Wahrscheinlichkeitsraum und $M \subseteq L^{(0)}(\Omega, \mathcal{F}, P)$ die Menge aller in (Ω, \mathcal{F}, P) definierten Zufallsvariablen mit einer kontinuierlichen Wahrscheinlichkeitsverteilung F .

$CVaR_\alpha$ ist ein kohärentes Risikomaß in M , $\forall \alpha \in (0, 1)$.

Beweis. (C1): $CVaR_\alpha(L + l) = CVaR_\alpha(L) + l$, $l \in \mathbb{R}$, denn

$$CVaR_\alpha(L) = \frac{1}{1-\alpha} \int_\alpha^1 VaR_u(L) du \quad \text{und} \quad VaR_\alpha(L + l) = VaR_\alpha(L) + l.$$

$$(C3): CVaR_\alpha(\lambda L) = \frac{1}{1-\alpha} \int_\alpha^1 VaR_u(\lambda L) du = \frac{\lambda}{1-\alpha} \int_\alpha^1 VaR_u(L) du = \lambda CVaR_\alpha(L) \quad \forall \lambda \geq 0.$$

$$(C4): L_1 \leq L_2 \Rightarrow VaR_\alpha(L_1) \leq VaR_\alpha(L_2) \quad \forall \alpha$$

$$\Rightarrow CVaR_\alpha(L_1) = \frac{1}{1-\alpha} \int_\alpha^1 VaR_u(L_1) du \leq \frac{1}{1-\alpha} \int_\alpha^1 VaR_u(L_2) du = CVaR_\alpha(L_2).$$

(C2): bleibt offen. □

5.7 Elliptische Verteilungen und Portfoliooptimierung

Es werden d Aktien und die Klasse \mathcal{P} aller Portfolios bestehend aus diesen Aktien betrachtet. Jedes (long-short) Portfolio aus \mathcal{P} ist eindeutig durch den Gewichtsvektor $w = (w_i) \in \mathbb{R}^d$ definiert. Daher gilt:

$$\mathcal{P} = \left\{ w = (w_i) \in \mathbb{R}^d : \sum_{i=1}^d |w_i| = 1 \right\}$$

Sei $X = (X_1, X_2, \dots, X_d)$ ein Zufallsvektor, der die Rendite der d Aktien darstellt. Sei $\mathbb{E}(X) = \mu$.

Die Portfoliorendite ist die Zufallsvariable $Z(w) = \sum_{i=1}^d w_i X_i$ und dementsprechend ist die erwartete Portfoliorendite: $\mathbb{E}(Z(w)) = w^\top \mu$.

Es gelte $X \sim E_d(\mu, \Sigma, \psi)$ mit $\mathbb{E}(X_k^2) < \infty$ und $\Sigma = Cov(X)$. Sei \mathcal{P}_m die Klasse jener Portfolios aus \mathcal{P} , sodass $\mathbb{E}(Z(w)) = m$, $m \in \mathbb{R}$, $m > 0$.

$$\mathcal{P}_m = \left\{ w = (w_i) \in \mathbb{R}^d, \sum_{i=1}^d |w_i| = 1, w^\top \mu = m \right\}$$

Das **Mean-Variance Portfolio-Optimierungsmodell** nach Markowitz (1952, 1987) hat die Form

$$\min_{w \in \mathcal{P}_m} Var(Z(w)) \tag{5.2}$$

Sei ρ ein Risikomaß. Das **Mean- ρ Portfolio-Optimierungsmodell** lautet:

$$\min_{w \in \mathcal{P}_m} \rho(Z(w)) \tag{5.3}$$

Sei $\rho = VaR_\alpha$, $\alpha \in (0, 1)$. Das **Mean-VaR Portfolio-Optimierungsmodell** lautet:

$$\min_{w \in \mathcal{P}_m} VaR_\alpha(Z(w)) \tag{5.4}$$

Frage: Wie hängen die Probleme (5.2) und (5.3) (insbesondere (5.4)) zusammen?

Satz 5.7. Sei M die Menge der erwarteten Renditen der Portfolii aus \mathcal{P} . Die Risikofaktoren, d.h. die Renditen der einzelnen Aktien, seien elliptisch verteilt, $X = (X_1, X_2, \dots, X_d) \sim E_d(\mu, \Sigma, \psi)$ für gegebene $\mu \in \mathbb{R}^d$, $\sigma \in \mathbb{R}^{d \times d}$ und $\psi: \mathbb{R} \rightarrow \mathbb{R}$. VaR_α ist kohärent auf M für jedes $\alpha \in (0.5, 1)$.

Beweis. Zu zeigen ist die Subadditivität von VaR_α . Seien dazu L_1 und L_2 die Renditen zweier Portfolii aus \mathcal{P} .

z.z.: $VaR_\alpha(L_1 + L_2) \leq VaR_\alpha(L_1) + VaR_\alpha(L_2), \forall \alpha > 0.5$.

$\forall L \in M$ gilt $L = w^\top X$. Da $X \sim E_d(\mu, \Sigma, \psi)$ gilt außerdem $X = \mu + AY$, wobei $AA^\top = \Sigma$ und $Y \in S_k(\psi)$, $A \in \mathbb{R}^{d \times k}$, $\mu \in \mathbb{R}^d$.

$$L = w^\top (\mu + AY) = (w^\top A)Y + w^\top \mu \stackrel{\text{Satz 5.3}}{\stackrel{d}{=} \|w^\top A\| Y_1 + w^\top \mu}$$

$$\Rightarrow VaR_\alpha(L) = VaR_\alpha(\|w^\top A\| Y_1 + w^\top \mu) = \|w^\top A\| VaR_\alpha(Y_1) + w^\top \mu.$$

Seien $L_1 = w_1^\top X$ und $L_2 = w_2^\top X$.

$$\begin{aligned} VaR_\alpha(L_1 + L_2) &= VaR_\alpha((w_1^\top + w_2^\top)X) = \|(w_1^\top + w_2^\top)A\| VaR_\alpha(Y_1) + (w_1^\top + w_2^\top)\mu \\ &\stackrel{(*)}{\leq} \|w_1^\top A\| VaR_\alpha(Y_1) + w_1^\top \mu + \|w_2^\top A\| VaR_\alpha(Y_1) + w_2^\top \mu \\ &= VaR_\alpha(L_1) + VaR_\alpha(L_2). \end{aligned}$$

Ad (*): Hier braucht man, dass $\alpha > 0.5$ ist, da Y_1 sphärisch verteilt und deswegen symmetrisch verteilt ist. Somit ist $VaR_\alpha(Y_1) \geq 0$ für $\alpha > 0.5$. \square

Satz 5.8. (Embrechts et al., 2002)

Sei $X = (X_1, X_2, \dots, X_d) = \mu + AY$ elliptisch verteilt mit $\mu \in \mathbb{R}^d$, $A \in \mathbb{R}^{d \times k}$, und einem sphärisch verteilten Zufallsvektor $Y \sim S_k(\psi)$ und $\mathbb{E}(X_k^2) < \infty, \forall k$. Sei ρ ein Risikomaß, das die Eigenschaften (C1) und (C3) besitzt, und $\rho(Y_1) \geq 0$ erfüllt, wobei Y_1 die erste Komponente des sphärisch verteilten Zufallsvektors Y ist.

Es gilt dann:

$$\arg \min\{\rho(Z(w)) : w \in \mathcal{P}_m\} = \arg \min\{Var(Z(w)) : w \in \mathcal{P}_m\}$$

Beweis.

$$\underbrace{Z(w)}_{PF\text{-Rendite}} = w^\top X \stackrel{d}{=} w^\top (AY + \mu) \stackrel{d}{=} \|w^\top A\| Y_1 + \underbrace{w^\top \mu}_{=: m} \quad \forall w \in \mathcal{P}_m$$

Sei ρ ein Risikomaß, das die Eigenschaften (C1) und (C3) erfüllt und außerdem

$$\frac{\rho(Y_1)}{\sqrt{Var(Y_1)}} =: k > 0, \text{ da } \rho(Y_1) > 0 \text{ laut Ann. } \Rightarrow \rho(Y_1) = k\sqrt{Var(Y_1)}$$

Es gilt

$$\begin{aligned} \rho(Z(w)) &= \rho(\|w^\top A\| Y_1 + m) \stackrel{(C1)}{=} \rho(\|w^\top A\| Y_1) + m \stackrel{(C3)}{=} \|w^\top A\| \rho(Y_1) + m \\ &= \|w^\top A\| k\sqrt{Var(Y_1)} + m = k\sqrt{Var(\underbrace{\|w^\top A\| Y_1}_{\stackrel{d}{=} (w^\top A)Y})} + m = k\sqrt{Var(\underbrace{(w^\top A)Y + m}_{\stackrel{d}{=} w^\top X})} + m \\ &= k\sqrt{Var(Z(w))} + m \\ &\Rightarrow \min_{w \in \mathcal{P}_m} \rho(Z(w)) = k\sqrt{\min_{w \in \mathcal{P}_m} Var(Z(w))} + m \\ &\Rightarrow \arg \min\{\rho(Z(w)) : w \in \mathcal{P}_m\} = \arg \min\{Var(Z(w)) : w \in \mathcal{P}_m\} \end{aligned}$$

\square

Kapitel 6

Einführung in Copulas

6.1 Grundlegende Eigenschaften

Definition 6.1.1. Eine d -dimensionale Copula ist eine Verteilungsfunktion auf $[0, 1]^d$, deren Randverteilungen jeweils standardgleichverteilt auf $[0, 1]$ sind.

Oder äquivalent:

Eine Copula C ist eine Funktion $C: [0, 1]^d \rightarrow [0, 1]$, die folgende Eigenschaften hat:

1. $C(u_1, u_2, \dots, u_d)$ ist mon. steigend in jeder Variable u_i , $1 \leq i \leq d$.
2. $C(1, 1, \dots, 1, u_k, 1, \dots, 1) = u_k$ für jedes $k \in \{1, \dots, d\}$, $u_k \in [0, 1]$.
3. Folgende Ungleichung (sogenannte *Rechtecksungleichung*) gilt für alle (a_1, a_2, \dots, a_d) , $(b_1, b_2, \dots, b_d) \in [0, 1]^d$ mit $a_k \leq b_k$, $\forall k \in \{1, 2, \dots, d\}$:

$$\sum_{k_1=1}^2 \dots \sum_{k_d=1}^2 (-1)^{k_1+k_2+\dots+k_d} C(u_{1k_1}, u_{2k_2}, \dots, u_{dk_d}) \geq 0,$$

wobei $u_{j1} = a_j$ und $u_{j2} = b_j$.

Anmerkung: Für $2 \leq k \leq d$ sind die k -dimensionalen Randverteilungen einer d -dimensionalen Copula wieder Copulas, k -dimensionale Copulas.

Lemma 6.1.2. Sei $h: \mathbb{R} \rightarrow \mathbb{R}$ eine monoton steigende Funktion und $h^\leftarrow: \mathbb{R} \rightarrow \mathbb{R}$ die verallgemeinerte inverse Funktion von h . Es gelten dann folgende Aussagen:

1. h ist stetig $\iff h^\leftarrow$ ist streng monoton steigend.
2. h ist streng monoton steigend $\iff h^\leftarrow$ ist stetig.
3. $h^\leftarrow(h(x)) \leq x$.
4. h ist streng monoton steigend $\implies h^\leftarrow(h(x)) = x$.
5. h ist stetig $\implies h(h^\leftarrow(y)) = y$.

Lemma 6.1.3. Sei X eine Zufallsvariable mit stetiger Verteilungsfunktion F . Es gilt: $\mathbb{P}(X = x: F^{\leftarrow}(F(x)) = x) = 1$, d.h. $F^{\leftarrow}(F(x)) \stackrel{f.s.}{=} x$.

Satz 6.1. Sei G eine Verteilungsfunktion in \mathbb{R} .

1. Quantil-Transformation:

Wenn $U \sim U(0, 1)$ (standard Gleichverteilung), dann gilt $\mathbb{P}(G^{\leftarrow}(U) \leq x) = G(x)$.

2. Wahrscheinlichkeits-Transformation:

Sei Y eine Zufallsvariable mit stetiger Verteilungsfunktion G . Es gilt $G(Y) \sim U(0, 1)$.

Beweis. 1) $G(x) \geq y \Leftrightarrow G^{\leftarrow}(y) \leq x$ (G bel. Verteilungsfunktion).

„ \Rightarrow “: $G^{\leftarrow}(y) = \inf\{z: G(z) \geq y\}$.

$$G(x) \geq y \Rightarrow x \in \{z: G(z) \geq y\} \Rightarrow G^{\leftarrow}(y) = \inf\{z: G(z) \geq y\} \leq x.$$

„ \Leftarrow “: $G^{\leftarrow}(y) \leq x \Rightarrow \inf\{z: G(z) \geq y\} \leq x$.

$x > G^{\leftarrow}(y)$: $\Rightarrow \exists x_1: G^{\leftarrow}(y) \leq x_1 < x$ sodass $G(x_1) \geq y \Rightarrow G(x) \geq G(x_1) \geq y$.

$x = G^{\leftarrow}(y)$: Ang. $G(x) < y$. G ist rechtsstetig $\Rightarrow \exists \delta > 0 \forall z \in (x, x + \delta): G(z) < y$. \nexists zu $x = \inf\{z: G(z) \geq y\}$.

Somit gilt:

$$\mathbb{P}(G^{\leftarrow}(U) \leq x) \stackrel{\text{Äquivalenz}}{=} \mathbb{P}(U \leq G(x)) \stackrel{U \sim U(0,1)}{=} G(x).$$

$$2) \mathbb{P}(G(Y) \leq u) \stackrel{(*)}{=} \mathbb{P}(G^{\leftarrow}(G(Y)) \leq G^{\leftarrow}(u)) \stackrel{(**)}{=} \mathbb{P}(Y \leq G^{\leftarrow}(u)) = G(G^{\leftarrow}(u)) \stackrel{(***)}{=} u$$

$\Rightarrow G(Y) \sim U(0, 1)$,

(*) G^{\leftarrow} streng monoton steigend (Lemma 6.1.2, Punkt 1).

(**) Lemma 6.1.3.

(***) Lemma 6.1.2, Punkt 5. □

Satz 6.2. (Sklar, 1959)

Sei $F: \mathbb{R}^d \rightarrow [0, 1]$ eine Gesamtverteilungsfunktion mit Randverteilungsfunktionen F_1, \dots, F_d . Es existiert eine Copula C , sodass für alle $x_1, x_2, \dots, x_d \in \bar{\mathbb{R}} = [-\infty, \infty]$

$$F(x_1, x_2, \dots, x_d) = C(F_1(x_1), F_2(x_2), \dots, F_d(x_d)). \quad (6.1)$$

Wenn F_1, \dots, F_d stetig, dann ist C eindeutig.

Vice-versa, sei C eine Copula und F_1, \dots, F_d Verteilungsfunktionen. Dann ist die Funktion F aus (6.1) eine Gesamtverteilungsfunktion mit Randverteilungsfunktionen F_1, \dots, F_d .

C aus (6.1) heißt Copula von F . Für einen Zufallsvektor $X \in \mathbb{R}^d$ mit Gesamtverteilungsfunktion F heißt C auch Copula von X .

Beweis. Der Beweis wird nur für stetige F_i , $1 \leq i \leq d$, betrachtet.

Sei $(x_1, \dots, x_d) \in \bar{\mathbb{R}}^d$ und F die Verteilungsfunktion von (X_1, \dots, X_d) . Weiters sei C die

Gesamtverteilungsfunktion des Zufallsvektors (U_1, \dots, U_d) und $U_i = F_i(X_i) \sim U(0, 1), 1 \leq i \leq d$. Dann gilt

$$\begin{aligned} F(x_1, \dots, x_d) &= \mathbb{P}(X_1 \leq x_1, \dots, X_d \leq x_d) = \mathbb{P}(\underbrace{F_1(X_1)}_{U_1} \leq F_1(x_1), \dots, \underbrace{F_d(X_d)}_{U_d} \leq F_d(x_d)) \\ &= C(F_1(x_1), \dots, F_d(x_d)) \end{aligned}$$

C ist eine Copula, denn $\forall u \in [0, 1]^d \exists x \in \bar{\mathbb{R}}^d : (u_1, \dots, u_d) = (F_1(x_1), \dots, F_d(x_d))$ und somit $C: [0, 1]^d \rightarrow [0, 1]$.

Damit ist die Existenz bewiesen.

Die Eindeutigkeit folgt aus $x_i = F_i^{\leftarrow}(u_i) \forall 1 \leq i \leq d$, denn (vgl. Korollar 6.1.4)

$$C(u_1, \dots, u_d) = F(F_1^{\leftarrow}(u_1), \dots, F_d^{\leftarrow}(u_d)).$$

Vice versa: Konstruiere einen Zufallsvektor mit Verteilungsfunktionen F_i . Sei U ein Zufallsvektor mit Gesamtverteilungsfunktion C . Betrachte den Zufallsvektor $(\underbrace{F_1^{\leftarrow}(U_1)}_{X_1}, \dots, \underbrace{F_d^{\leftarrow}(U_d)}_{X_d})$.

$$\begin{aligned} F(x_1, \dots, x_d) &= \mathbb{P}(X_1 \leq x_1, \dots, X_d \leq x_d) = \mathbb{P}(F_1^{\leftarrow}(U_1) \leq x_1, \dots, F_d^{\leftarrow}(U_d) \leq x_d) \\ &= \mathbb{P}(U_1 \leq F_1(x_1), \dots, U_d \leq F_d(x_d)) = C(F_1(x_1), \dots, F_d(x_d)). \end{aligned}$$

□

Korollar 6.1.4. Sei F eine Gesamtverteilungsfunktion mit stetigen Randverteilungsfunktionen F_1, \dots, F_d . Die eindeutige Copula von F ist folgendermaßen gegeben:

$$C(u_1, u_2, \dots, u_d) = F(F_1^{\leftarrow}(u_1), F_2^{\leftarrow}(u_2), \dots, F_d^{\leftarrow}(u_d)). \quad (6.2)$$

Beweis. F_i stetig $\forall i \Rightarrow C(u_1, \dots, u_d) = F(F_1^{\leftarrow}(u_1), \dots, F_d^{\leftarrow}(u_d))$.

Sei $X = (X_1, \dots, X_d)$ ein Zufallsvektor mit Verteilungsfunktion F und Randverteilungsfunktionen F_i . Setze $U_i = F_i(X_i) \sim U(0, 1)$. Betrachte (U_1, \dots, U_d) . Die dazugehörige Verteilungsfunktion C ist eine Copula und erfüllt 6.2. □

Satz 6.3. (Copula-Invarianz bzgl. streng monotonen Transformationen)

Sei $X = (X_1, X_2, \dots, X_d)^\top$ ein Zufallsvektor mit stetigen Randverteilungen F_1, F_2, \dots, F_d und Copula C . Seien T_1, T_2, \dots, T_d streng monoton steigende Funktionen in \mathbb{R} . Dann ist C auch eine Copula von $(T_1(X_1), T_2(X_2), \dots, T_d(X_d))^\top$.

Beweis. Sei \tilde{F}_i die Verteilungsfunktion von $T_i(X_i)$.

$$\tilde{F}_i(x) = \mathbb{P}(T_i(X_i) \leq x) = \mathbb{P}(X_i \leq T_i^{-1}(x)) = F_i(T_i^{-1}(x))$$

$\Rightarrow \tilde{F}_i = F_i \circ T_i^{-1} \forall 1 \leq i \leq d$.

T_i ist streng monoton steigend $\Leftrightarrow T_i^{-1}$ ist stetig (Lemma 6.1.2, Punkt 2), F_i ist stetig $\Rightarrow \tilde{F}_i$

ist stetig \Rightarrow es existiert eine eindeutige Copula \tilde{C} für $(T_1(X_1), \dots, T_d(X_d))$.

Mit Satz 6.2 folgt

$$\begin{aligned}\tilde{C}(\tilde{F}_1(x_1), \dots, \tilde{F}_d(x_d)) &= \mathbb{P}(T_1(X_1) \leq x_1, \dots, T_d(X_d) \leq x_d) = \mathbb{P}(X_1 \leq T_1^{-1}(x_1), \dots, X_d \leq T_d^{-1}(x_d)) \\ &= C(F_1(T_1^{-1}(x_1)), \dots, F_d(T_d^{-1}(x_d))) = C(\tilde{F}_1(x_1), \dots, \tilde{F}_d(x_d))\end{aligned}$$

und somit die Gleichheit von C und \tilde{C} in $[\tilde{F}_i(\mathbb{R})]^d = [0, 1]^d$ (\tilde{F}_i ist stetige Verteilungsfunktion). \square

Beispiel 6.1.5. Sei $X \sim N_d(0, \Sigma)$, wobei $\Sigma = R$ die Korrelationsmatrix von X ist. Seien ϕ_R und ϕ die Verteilungsfunktionen von X bzw. X_1 . Die Copula von X ist die so genannte *Gauss'sche Copula* C_R^{Ga} :

$$C_R^{Ga}(u_1, u_2, \dots, u_d) = \phi_R(\phi^{-1}(u_1), \phi^{-1}(u_2), \dots, \phi^{-1}(u_d)).$$

C_R^{Ga} ist auch die Copula jeder nicht degenerierten Normalverteilung $N_d(0, \Sigma)$ mit Korrelationsmatrix R .

Für $d = 2$ und $\rho = R_{12} \in (-1, 1)$ gilt:

$$C_R^{Ga}(u_1, u_2) = \int_{-\infty}^{\phi^{-1}(u_1)} \int_{-\infty}^{\phi^{-1}(u_2)} \frac{1}{2\pi(1-\rho^2)^{1/2}} \exp\left\{-\frac{(x_1^2 - 2\rho x_1 x_2 + x_2^2)}{2(1-\rho^2)}\right\} dx_1 dx_2$$

\diamond

Satz 6.4. (*Fréchet Schranken*)

Für jede Copula gilt

$$\max\left\{\sum_{k=1}^d u_k - d + 1, 0\right\} \leq C(u_1, u_2, \dots, u_d) \leq \min\{u_1, u_2, \dots, u_d\}.$$

Notation: Untere Schranke =: W_d und obere Schranke =: M_d , für $d \geq 2$. Für $d = 2$ setzen wir $M := M_2$, $W := W_2$.

Beweis. • obere Schranke:

$$C(u_1, \dots, u_d) = \mathbb{P}(U_1 \leq u_1, \dots, U_d \leq u_d) = \mathbb{P}\left(\bigcap_{i=1}^d \{U_i \leq u_i\}\right) \leq \mathbb{P}(U_i \leq u_i) = u_i \quad \forall i.$$

• untere Schranke:

$$\begin{aligned}C(u_1, \dots, u_d) &= \mathbb{P}\left(\bigcap_{i=1}^d \{U_i \leq u_i\}\right) = 1 - \underbrace{\mathbb{P}\left(\bigcup_{i=1}^d \{U_i > u_i\}\right)}_{\leq \sum_{i=1}^d \mathbb{P}(U_i > u_i)} \\ &\geq 1 - \sum_{i=1}^d \mathbb{P}(U_i > u_i) = 1 - \sum_{i=1}^d (1 - u_i) = 1 - d + \sum_{i=1}^d u_i.\end{aligned}$$

Andererseits ist $C \geq 0$, da Wahrscheinlichkeitsfunktion $\Rightarrow C(u_1, \dots, u_d) \geq \max\{1 - d + \sum_{i=1}^d u_i, 0\}$. \square

Anmerkung: Ein analoges Ergebnis wie im Satz 6.4 gilt für allgemeine multivariate Verteilungen F mit Randverteilungen F_i , $1 \leq i \leq d$:

$$\max \left\{ \sum_{k=1}^d F_k(x_k) - d + 1, 0 \right\} \leq F(x_1, x_2, \dots, x_d) \leq \min\{F_1(x_1), F_2(x_2), \dots, F_d(x_d)\}.$$

Beispiel 6.1.6. Zeigen Sie, dass die Fréchet untere Schranke W_d für $d \geq 3$ keine Copula ist. Hinweis: Verwenden Sie die Mengenfunktion Q ,

$$Q([a_1, b_1] \times [a_2, b_2] \times \dots \times [a_d, b_d]) = \sum_{k_1=1}^2 \sum_{k_2=1}^2 \dots \sum_{k_d=1}^2 (-1)^{k_1+k_2+\dots+k_d} W_d(u_{1k_1}, u_{2k_2}, \dots, u_{dk_d}),$$

wobei $(a_1, a_2, \dots, a_d), (b_1, b_2, \dots, b_d) \in [0, 1]^d$ mit $a_k \leq b_k$ und $u_{j1} = a_j$ und $u_{j2} = b_j$ für $j \in \{1, 2, \dots, d\}$. \diamond

Satz 6.5. Für jedes $d \geq 3$ und jedes $u \in [0, 1]^d$, existiert eine Copula $C_{d,u}$, sodass $C_{d,u}(u) = W_d(u)$.

Anmerkung: Für jedes $d \geq 2$ ist die Fréchet obere Schranke M_d eine Copula.

Beispiel 6.1.7. M und W sind Copulas.

Sei X eine Zufallsvariable mit Verteilungsfunktion F_X . Seien $Y = T(X)$ und $Z = S(X)$ zwei Zufallsvariablen, wobei T und S zwei streng monotone Funktionen, T steigend und S fallend, sind. Zeigen Sie, dass M die Copula von $(X, T(X))^\top$ und W die Copula von $(X, S(X))^\top$ ist.

• T streng monoton steigende Funktion:

$$\begin{aligned} C(u_1, u_2) &\stackrel{\text{Korollar}}{=} F(F_X^{\leftarrow}(u_1), F_Y^{\leftarrow}(u_2)) = \mathbb{P}(X \leq F_X^{\leftarrow}(u_1), Y \leq F_Y^{\leftarrow}(u_2)) \\ &= \mathbb{P}(F_X(X) \leq u_1, \underbrace{F_Y(Y)}_{(*)} \leq u_2) = \mathbb{P}(F_X(X) \leq u_1, F_X \circ T^{-1}(T(X)) \leq u_2) \\ &= \mathbb{P}(F_X(X) \leq u_1, F_X(X) \leq u_2) = \min\{u_1, u_2\}, \end{aligned}$$

denn $F_X(X) \sim U(0, 1)$ und für $(*)$ gilt, dass $Y = T(X)$ und $F_Y = F_X \circ T^{-1}$, denn

$$\mathbb{P}(Y \leq a) = \mathbb{P}(T(X) \leq a) = \mathbb{P}(X \leq T^{-1}(a)) = F_X(T^{-1}(a)).$$

• S streng monoton fallende Funktion: analog. \diamond

6.2 Co-Monotonie und Anti-Monotonie

Definition 6.2.1. X_1 und X_2 heißen *co-monoton*, wenn M eine Copula von $(X_1, X_2)^\top$ ist. X_1 und X_2 heißen *anti-monoton*, wenn W eine Copula von $(X_1, X_2)^\top$ ist.

Satz 6.6. *Angenommen eine Copula von $(X_1, X_2)^\top$ ist W oder M . Es existieren dann zwei monotone Funktionen $\alpha, \beta: \mathbb{R} \rightarrow \mathbb{R}$ und eine Zufallsvariable Z , sodass*

$$(X_1, X_2) \stackrel{d}{=} (\alpha(Z), \beta(Z)).$$

Falls M die Copula von $(X_1, X_2)^\top$ ist, dann sind α und β monoton steigend, falls W die Copula von $(X_1, X_2)^\top$ ist, dann ist α monoton steigend und β monoton fallend.

Wenn die Randverteilungen F_1 und F_2 von $(X_1, X_2)^\top$ stetig sind, dann gilt:

$$\begin{aligned} C = W &\iff X_2 = T(X_1) \text{ fast sicher, } T = F_2^{\leftarrow} \circ (1 - F_1) \text{ monoton fallend} \\ C = M &\iff X_2 = T(X_1) \text{ fast sicher, } T = F_2^{\leftarrow} \circ F_1 \text{ monoton steigend.} \end{aligned}$$

Beweis. In McNeil et al., 2005. □

Satz 6.7. *Sei $(X_1, X_2)^\top$ ein Zufallsvektor mit Randverteilungsfunktionen F_1, F_2 und einer nicht spezifizierten Abhängigkeitsstruktur.*

Sei $\text{Var}(X_1), \text{Var}(X_2) \in (0, \infty)$. Dann gilt:

1. *Die Menge der möglichen linearen Korrelationen von X_1 und X_2 ist ein abgeschlossenes Intervall $[\rho_{L,\min}; \rho_{L,\max}]$ mit $0 \in (\rho_{L,\min}; \rho_{L,\max})$.*
2. *Die minimale lineare Korrelation wird dann und nur dann erreicht, wenn X_1 und X_2 anti-monoton sind. Die maximale lineare Korrelation wird dann und nur dann erreicht, wenn X_1 und X_2 co-monoton sind.*

Im Beweis wird die Höfding'sche Gleichung verwendet:

Lemma 6.2.2. *(Die Höfding'sche Gleichung)*

Sei $(X_1, X_2)^\top$ ein Zufallsvektor mit Gesamtverteilung F und Randverteilungen F_1, F_2 . Wenn $\text{Cov}(X_1, X_2) < \infty$, dann gilt:

$$\text{Cov}(X_1, X_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (F(x_1, x_2) - F_1(x_1)F_2(x_2)) dx_1 dx_2.$$

Beweis. In McNeil et al., 2005. □

Beweis. (Satz)

$$\begin{aligned}
\rho_L(X_1, X_2) &= \frac{\text{Cov}(X_1, X_2)}{\sqrt{\text{Var}(X_1)\text{Var}(X_2)}} \\
&= \frac{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (F(x_1, x_2) - F_1(x_1)F_2(x_2)) dx_1 dx_2}{\sqrt{\text{Var}(X_1)\text{Var}(X_2)}} \\
&\stackrel{\text{Sklar}}{=} \frac{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (C(F_1(x_1), F_2(x_2)) - F_1(x_1)F_2(x_2)) dx_1 dx_2}{\sqrt{\text{Var}(X_1)\text{Var}(X_2)}}.
\end{aligned}$$

Mit den Fréchet-Schranken folgt:

$$\underbrace{\frac{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (\max\{F_1(x_1) + F_2(x_2) - 1, 0\} - F_1(x_1)F_2(x_2)) dx_1 dx_2}{\sqrt{\text{Var}(X_1)\text{Var}(X_2)}}}_{\rho_{L,\min}} \leq \rho_L(X_1, X_2)$$

und

$$\rho_L(X_1, X_2) \leq \underbrace{\frac{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (\min\{F_1(x_1), F_2(x_2)\} - F_1(x_1)F_2(x_2)) dx_1 dx_2}{\sqrt{\text{Var}(X_1)\text{Var}(X_2)}}}_{\rho_{L,\max}}.$$

- $\rho_{L,\max}$ wird für $C = M$ (Co-Monotonie) angenommen und es gilt $\rho_{L,\max} \geq 0$, denn

$$\min\{F_1(x_1), F_2(x_2)\} \geq F_1(x_1)F_2(x_2).$$

- $\rho_{L,\min}$ wird für $C = W$ (Anti-Monotonie) angenommen und es gilt $\rho_{L,\min} \leq 0$, denn wird das Maximum in 0 angenommen, folgt die Behauptung trivialerweise, wird jedoch das Maximum in $F_1(x_1) + F_2(x_2) - 1$ angenommen, so bleibt zu zeigen, dass $F_1(x_1) + F_2(x_2) - 1 \leq F_1(x_1)F_2(x_2)$ gilt:

$$\begin{aligned}
F_1(x_1) + F_2(x_2) - 1 - F_1(x_1)F_2(x_2) &\leq 0 \\
F_1(x_1)(1 - F_2(x_2)) + F_2(x_2) - 1 &\leq 0 \\
\underbrace{(1 - F_1(x_1))}_{\geq 0} \underbrace{(F_2(x_2) - 1)}_{\leq 0} &\leq 0.
\end{aligned}$$

- Konvexe Kombinationen zweier Copulas sind wieder Copulas:

$$\forall \alpha \in [\rho_{L,\min}, \rho_{L,\max}] \quad \exists C_\alpha : \frac{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (C_\alpha(F_1(x_1), F_2(x_2)) - F_1(x_1)F_2(x_2)) dx_1 dx_2}{\sqrt{\text{Var}(X_1)\text{Var}(X_2)}} = \alpha$$

$C_\alpha = \lambda W + (1 - \lambda)M$ für ein λ , sodass $\alpha = \lambda\rho_{L,\min} + (1 - \lambda)\rho_{L,\max}$, und C_α ist eine Copula. \square

Beispiel 6.2.3. Sei $X_1 \sim LN(0, 1)$ und $X_2 \sim LN(0, \sigma^2)$, $\sigma > 0$. Bestimmen Sie $\rho_{L,\min}(X_1, X_2)$ und $\rho_{L,\max}(X_1, X_2)$:

Sei $Z \sim N(0, 1)$, dann gilt $X_1 \stackrel{d}{=} e^Z$ und $X_2 \stackrel{d}{=} e^{\sigma Z} = (e^Z)^\sigma = X_1^\sigma$. X_2 ist also eine monoton steigende Transformation von $X_1 \Rightarrow$ Co-Monotonie!

$$\Rightarrow \rho_{L,\max} = \rho_L(e^Z, e^{\sigma Z}) = \frac{e^\sigma - 1}{\sqrt{(e - 1)(e^{\sigma^2} - 1)}}.$$

Ist jedoch $X_2 \stackrel{d}{=} e^{-\sigma Z} = (e^Z)^{-\sigma} = X_1^{-\sigma}$, dann ist X_2 eine monoton fallende Transformation von $X_1 \Rightarrow$ Anti-Monotonie!

$$\Rightarrow \rho_{L,\min} = \rho_L(e^Z, e^{-\sigma Z}) = \frac{e^{-\sigma} - 1}{\sqrt{(e-1)(e^{\sigma^2} - 1)}}.$$

In Abb. 6.1 sind die obere bzw. untere Schranke $\rho_{L,\max}$ und $\rho_{L,\min}$ in Abhängigkeit von σ aufgetragen. \diamond

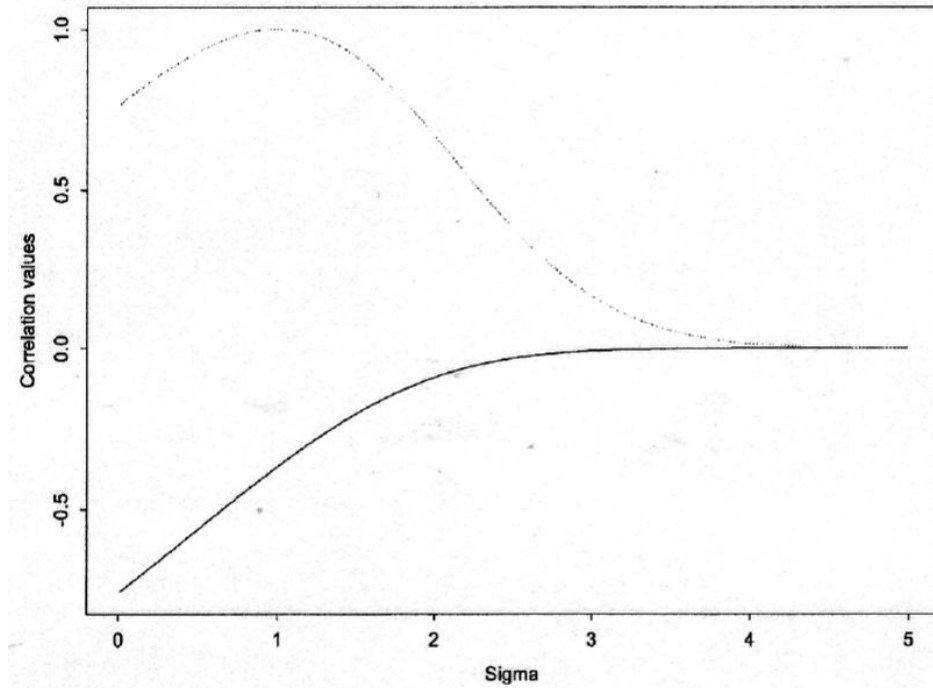


Abbildung 6.1: $\rho_{L,\max}$ und $\rho_{L,\min}$ in Abhängigkeit von σ

Beispiel 6.2.4. Betrachten Sie zwei ZV Z_1 und Z_2 , die die Verluste zweier Portfolii darstellen. Sei $Z_1 \sim N(0, 1)$, $Z_2 \sim N(0, 1)$ und $\rho_L(Z_1, Z_2) = 0$.

Geben Sie zwei Zufallsvektoren $(X_1, X_2)^\top$ und $(Y_1, Y_2)^\top$ mit unterschiedlichen Gesamtverteilungsfunktionen an, die die obigen Annahmen erfüllen, d.h. $X_1, X_2, Y_1, Y_2 \sim N(0, 1)$ und $\rho_L(X_1, X_2) = 0$, $\rho_L(Y_1, Y_2) = 0$.

Bestimmen Sie die Quantile $F_{X_1+X_2}^{\leftarrow}(\alpha)$ und $F_{Y_1+Y_2}^{\leftarrow}(\alpha)$ der Gesamtverluste und vergleichen Sie diese miteinander.

$$(X_1, X_2) \sim N_2(0, I_2), Y_1 \sim N(0, 1) \text{ und } Y_2 = VY_1, \text{ wobei } V = \begin{cases} 1 & \text{mit W! } 0.5 \\ -1 & \text{mit W! } 0.5 \end{cases}.$$

z.z.: $Y_2 \sim N(0, 1)$.

$$\begin{aligned} \mathbb{P}(Y_2 \leq x) &= \mathbb{P}(VY_1 \leq x) = \mathbb{P}(VY_1 \leq x|V=1) \underbrace{\mathbb{P}(V=1)}_{0.5} + \mathbb{P}(VY_1 \leq x|V=-1) \underbrace{\mathbb{P}(V=-1)}_{0.5} \\ &= 0.5(\mathbb{P}(Y_1 \leq x) + \underbrace{\mathbb{P}(-Y_1 \leq x)}_{\mathbb{P}(Y_1 \leq x)}) = \mathbb{P}(Y_1 \leq x) \end{aligned}$$

z.z.: $\rho_L(Y_1, Y_2) = 0$.

$$\begin{aligned} \rho_L(Y_1, Y_2) &= \rho_L(Y_1, VY_1) = \mathbb{E}[(Y_1 - \underbrace{\mathbb{E}(Y_1)}_{=0})(VY_1 - \underbrace{\mathbb{E}(VY_1)}_{=0})] = \mathbb{E}(Y_1 \cdot VY_1) \\ &= \mathbb{E}(Y_1^2) \underbrace{\mathbb{P}(V=1)}_{0.5} + \mathbb{E}(-Y_1^2) \underbrace{\mathbb{P}(V=-1)}_{0.5} = 0.5 \cdot 1 - 0.5 \cdot 1 = 0. \end{aligned}$$

$X_1 + X_2 \sim N(0, 2)$:

$$\mathbb{P}(X_1 + X_2 \geq \alpha) = \mathbb{P}(X_1 \geq \frac{\alpha}{\sqrt{2}}) = \bar{\phi}(\frac{\alpha}{\sqrt{2}}) \quad \forall \alpha$$

$$\Leftrightarrow \text{VaR}_\alpha(X_1 + X_2) = \phi_{X_1+X_2}^{-1}(\alpha) = \sqrt{2}\phi^{-1}(\alpha).$$

$Y_1 + Y_2$: Für alle $\alpha \geq 0$ gilt

$$\begin{aligned} \mathbb{P}(Y_1 + Y_2 \geq \alpha) &= \mathbb{P}(Y_1 + VY_1 \geq \alpha | V=1)\mathbb{P}(V=1) + \mathbb{P}(Y_1 + VY_1 \geq \alpha | V=-1)\mathbb{P}(V=-1) \\ &= 0.5\mathbb{P}(Y_1 \geq \frac{\alpha}{2}) + 0.5\underbrace{\mathbb{P}(0 \geq \alpha)}_{=0} = \frac{1}{2}\bar{\phi}(\frac{\alpha}{2}). \end{aligned}$$

$$\begin{aligned} \bar{\phi}_{Y_1+Y_2}(\alpha) &= \frac{1}{2}\bar{\phi}(\frac{\alpha}{2}) \\ 1 - \phi_{Y_1+Y_2}(\alpha) &= \frac{1}{2}(1 - \phi(\frac{\alpha}{2})) \\ \phi_{Y_1+Y_2}(\alpha) &= \frac{1}{2} + \frac{1}{2}\phi(\frac{\alpha}{2}) \end{aligned}$$

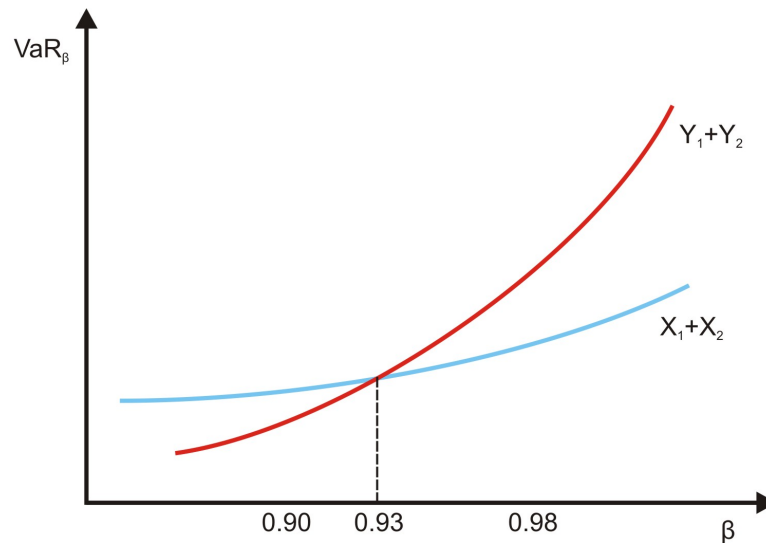
$\phi_{Y_1+Y_2}(\alpha) = \beta \Leftrightarrow \alpha = \phi_{Y_1+Y_2}^{-1}(\beta)$:

$$\frac{1}{2} + \frac{1}{2}\phi(\frac{\alpha}{2}) = \beta \Rightarrow 1 + \phi(\frac{\alpha}{2}) = 2\beta \Rightarrow \phi(\frac{\alpha}{2}) = 2\beta - 1 \Rightarrow \alpha = 2\phi^{-1}(2\beta - 1).$$

Insgesamt gilt also

$$\begin{aligned} \phi_{X_1+X_2}^{-1}(\beta) &= \sqrt{2}\phi^{-1}(\beta) \\ \phi_{Y_1+Y_2}^{-1}(\beta) &= 2\phi^{-1}(2\beta - 1) \end{aligned}$$

und $2\beta - 1 \geq 0.5 \Leftrightarrow \beta \geq 0.75$.



Fazit: Aus den Verlustverteilungen der zwei Teile eines Portfolios und aus der Korrelation der jeweiligen Verluste lassen sich keine Schlüsse über die Verlustverteilung des Gesamtportfolios ziehen.

◇

6.3 Kendall's Tau und Spearman's Rho

Seien $(x, y)^\top$ und $(\tilde{x}, \tilde{y})^\top$ zwei Beobachtungen von einem Zufallsvektor $(X, Y)^\top$. $(x, y)^\top$ und $(\tilde{x}, \tilde{y})^\top$ heißen *übereinstimmend*, falls $(x - \tilde{x})(y - \tilde{y}) > 0$, und *nicht übereinstimmend*, falls $(x - \tilde{x})(y - \tilde{y}) < 0$.

Definition 6.3.1. Sei $(X_1, X_2)^\top$ ein Zufallsvektor. *Kendall's Tau* ist für $(X_1, X_2)^\top$ folgendermaßen definiert:

$$\rho_\tau(X_1, X_2) = \mathbb{P}((X_1 - X'_1)(X_2 - X'_2) > 0) - \mathbb{P}((X_1 - X'_1)(X_2 - X'_2) < 0),$$

wobei $(X'_1, X'_2)^\top$ eine unabhängige Kopie von $(X_1, X_2)^\top$ ist.

Äquivalent: $\rho_\tau(X_1, X_2) = \mathbb{E}(\text{sgn}[(X_1 - X'_1)(X_2 - X'_2)])$.

Im d -dimensionalen Fall $X \in \mathbb{R}^d$: $\rho_\tau(X) = \text{Cov}(\text{sgn}(X - X'))$, wobei $X' \in \mathbb{R}^d$ eine unabhängige Kopie von $X \in \mathbb{R}^d$ ist.

Kendall's Tau der Stichprobe:

Sei $\{(x_1, y_1)^\top, (x_2, y_2)^\top, \dots, (x_n, y_n)^\top\}$ eine Stichprobe von n Beobachtungen des Zufallsvektors $(X, Y)^\top$, dessen Randverteilungen stetig sind. Sei c die Anzahl der übereinstimmenden Paare und d die Anzahl der nicht übereinstimmenden Paare aus der Stichprobe.

$$\tilde{\rho}_\tau(X, Y) = \frac{c - d}{c + d} \stackrel{\text{a.s.}}{=} \frac{c - d}{n(n-1)/2}.$$

Definition 6.3.2. Sei $(X_1, X_2)^\top$ ein Zufallsvektor. *Spearman's Rho* ist für $(X_1, X_2)^\top$ folgendermaßen definiert:

$$\rho_S(X_1, X_2) = 3(\mathbb{P}((X_1 - X'_1)(X_2 - X''_2) > 0) - \mathbb{P}((X_1 - X'_1)(X_2 - X''_2) < 0)),$$

wobei $(X'_1, X'_2)^\top, (X''_1, X''_2)^\top$ unabhängige Kopien von $(X_1, X_2)^\top$ sind.

Äquivalent: Seien F_1 und F_2 die stetigen Randverteilungen von $(X_1, X_2)^\top$. Es gilt $\rho_S(X_1, X_2) = \rho_L(F_1(X_1), F_2(X_2))$, d.h. Spearman's Rho ist die lineare Korrelation der eindeutigen Copula von $(X_1, X_2)^\top$.

Im d -dimensionalen Fall $X \in \mathbb{R}^d$: $\rho_S(X) = \rho_L(F_1(X_1), F_2(X_2), \dots, F_d(X_d))$ ist die Korrelationsmatrix der eindeutigen Copula von X , wobei F_1, F_2, \dots, F_d die stetigen Randverteilungen von X sind.

Satz 6.8. (ohne Beweis)

Sei $(X_1, X_2)^\top$ ein Zufallsvektor mit stetigen Randverteilungen und eindeutiger Copula C . Für die Rangkorrelationen $\rho_\tau(X_1, X_2)$ und $\rho_S(X_1, X_2)$ gilt:

$$\begin{aligned}\rho_\tau(X_1, X_2) &= 4 \int_0^1 \int_0^1 C(u_1, u_2) dC(u_1, u_2) - 1 \\ \rho_S(X_1, X_2) &= 12 \int_0^1 \int_0^1 (C(u_1, u_2) - u_1 u_2) du_1 du_2 = 12 \int_0^1 \int_0^1 C(u_1, u_2) du_1 du_2 - 3\end{aligned}$$

Eigenschaften von ρ_τ und ρ_S

- ρ_τ und ρ_S sind symmetrische Abhängigkeitsmaße mit Wertebereich $[-1, 1]$.
- Falls X_1, X_2 unabhängig, dann $\rho_\tau(X_1, X_2) = \rho_S(X_1, X_2) = 0$.
Die Umkehrung gilt i.a. nicht.
- X_1, X_2 co-monoton dann und nur dann, wenn $\rho_\tau(X_1, X_2) = \rho_S(X_1, X_2) = 1$. X_1, X_2 anti-monoton dann und nur dann, wenn $\rho_\tau(X_1, X_2) = \rho_S(X_1, X_2) = -1$.
- Seien F_1, F_2 die stetigen Randverteilungen von $(X_1, X_2)^\top$ und T_1, T_2 zwei streng monotone Funktionen in $[-\infty, \infty]$. Dann gilt $\rho_\tau(X_1, X_2) = \rho_\tau(T_1(X_1), T_2(X_2))$ und $\rho_S(X_1, X_2) = \rho_S(T_1(X_1), T_2(X_2))$.

(Siehe Embrechts et al., 2002).

6.4 Tail-Abhängigkeit

Definition 6.4.1. Sei $(X_1, X_2)^\top$ ein Zufallsvektor mit Randverteilungen F_1 und F_2 . Der Koeffizient der oberen Tail-Abhängigkeit von $(X_1, X_2)^\top$ wird folgendermaßen definiert:

$$\lambda_U(X_1, X_2) = \lim_{u \rightarrow 1^-} \mathbb{P}(X_2 > F_2^{\leftarrow}(u) | X_1 > F_1^{\leftarrow}(u)),$$

vorausgesetzt der Limes existiert.

Der Koeffizient der unteren Tail-Abhängigkeit von $(X_1, X_2)^\top$ wird folgendermaßen definiert:

$$\lambda_L(X_1, X_2) = \lim_{u \rightarrow 0^+} \mathbb{P}(X_2 \leq F_2^{\leftarrow}(u) | X_1 \leq F_1^{\leftarrow}(u)),$$

vorausgesetzt der Limes existiert.

Wenn $\lambda_U > 0$ ($\lambda_L > 0$) heißt es, $(X_1, X_2)^\top$ hat eine obere (untere) Tail-Abhängigkeit.

Definition 6.4.2. Sei die Copula C die Verteilungsfunktion von (U_1, U_2, \dots, U_d) mit $U_i \sim U[0, 1]$, $i = 1, 2, \dots, d$. Die Verteilungsfunktion von $(1 - U_1, 1 - U_2, \dots, 1 - U_d)$ heißt "survival" Copula von C und wird mit \hat{C} bezeichnet.

Lemma 6.4.3. Sei X ein Zufallsvektor mit multivariater Tail-Funktion \bar{F} ($\bar{F}(x_1, x_2, \dots, x_d) = \mathbb{P}(X_1 > x_1, X_2 > x_2, \dots, X_d > x_d)$) und Randverteilungsfunktionen F_i , $i = 1, 2, \dots, d$. Sei $\bar{F}_i = 1 - F_i$, $i = 1, 2, \dots, d$. Es gilt

$$\bar{F}(x_1, x_2, \dots, x_d) = \hat{C}(\bar{F}_1(x_1), \bar{F}_2(x_2), \dots, \bar{F}_d(x_d)).$$

Lemma 6.4.4. Es gilt $\hat{C}(1 - u_1, 1 - u_2) = 1 - u_1 - u_2 + C(u_1, u_2)$.

Satz 6.9. Sei $(X_1, X_2)^\top$ ein Zufallsvektor mit stetigen Randverteilungsfunktionen und Copula C . Es gilt

$$\begin{aligned} \lambda_U(X_1, X_2) &= \lim_{u \rightarrow 1^-} \frac{1 - 2u + C(u, u)}{1 - u} \text{ und} \\ \lambda_L(X_1, X_2) &= \lim_{u \rightarrow 0^+} \frac{C(u, u)}{u}, \end{aligned}$$

vorausgesetzt die Limita existieren.

Beweis.

$$\begin{aligned} \lambda_L(X_1, X_2) &\stackrel{\text{Def.}}{=} \lim_{u \rightarrow 0^+} \mathbb{P}(X_2 \leq F_2^{\leftarrow}(u) | X_1 \leq F_1^{\leftarrow}(u)) = \lim_{u \rightarrow 0^+} \frac{\mathbb{P}(X_2 \leq F_2^{\leftarrow}(u), X_1 \leq F_1^{\leftarrow}(u))}{\underbrace{\mathbb{P}(X_1 \leq F_1^{\leftarrow}(u))}_{=F_1(F_1^{\leftarrow}(u))=u}} \\ &= \lim_{u \rightarrow 0^+} \frac{F(F_1^{\leftarrow}(u), F_2^{\leftarrow}(u))}{u} \stackrel{\text{Korr. 6.1.4}}{=} \lim_{u \rightarrow 0^+} \frac{C(u, u)}{u}. \end{aligned}$$

$\lambda_U(X_1, X_2)$ folgt analog. □

Beispiel 6.4.5. Die Gumbel Familie von Copulas:

$$C_\theta^{Gu}(u_1, u_2) = \exp\left(-\left[(-\ln u_1)^\theta + (-\ln u_2)^\theta\right]^{1/\theta}\right), \quad \theta \geq 1$$

Es gilt $\lambda_U = 2 - 2^{1/\theta}$, $\lambda_L = 0$ (also keine untere Tailabhängigkeit), denn

$$C_\theta^{Gu}(u, u) = \exp\left(-\left[2(-\ln u)^\theta\right]^{1/\theta}\right) = \exp\left(2^{1/\theta} \ln u\right) = u^{2^{1/\theta}}.$$

$$\begin{aligned} \lambda_L^{\theta, Gu} &= \lim_{u \rightarrow 0^+} \frac{u^{2^{1/\theta}}}{u} \stackrel{\text{de l'H.}}{=} \lim_{u \rightarrow 0^+} \frac{2^{1/\theta} u^{2^{1/\theta}-1}}{1} = 0 \\ \lambda_U^{\theta, Gu} &= \lim_{u \rightarrow 1^-} \frac{1 - 2u + C_\theta^{Gu}(u, u)}{1 - u} = \lim_{u \rightarrow 1^-} \frac{1 - 2u + u^{2^{1/\theta}}}{1 - u} \\ &\stackrel{\text{de l'H.}}{=} \lim_{u \rightarrow 1^-} \frac{-2 + 2^{1/\theta} u^{2^{1/\theta}-1}}{-1} = 2 - 2^{1/\theta} \neq 0. \end{aligned}$$

◇

Beispiel 6.4.6. Die Clayton Familie von Copulas:

$$C_{\theta}^{Cl}(u_1, u_2) = (u_1^{-\theta} + u_2^{-\theta} - 1)^{1/\theta}, \quad \theta > 0$$

Es gilt $\lambda_U = 0$, $\lambda_L = 2^{-1/\theta}$ (also keine obere Tailabhängigkeit). (Analog zu Bsp. 6.4.5.) ◇

In Abb. 6.2 sind zwei Stichproben mit Gamma(3,1)-Randverteilungen und linearer Korrelation 0.5 abgebildet. $(X_1, Y_1)^{\top}$ und $(X_2, Y_2)^{\top}$ haben unterschiedliche Abhängigkeitsstrukturen und so kommt es, dass $(X_1, Y_1)^{\top}$ eine Gumbel-Copula und $(X_2, Y_2)^{\top}$ eine Clayton-Copula hat.

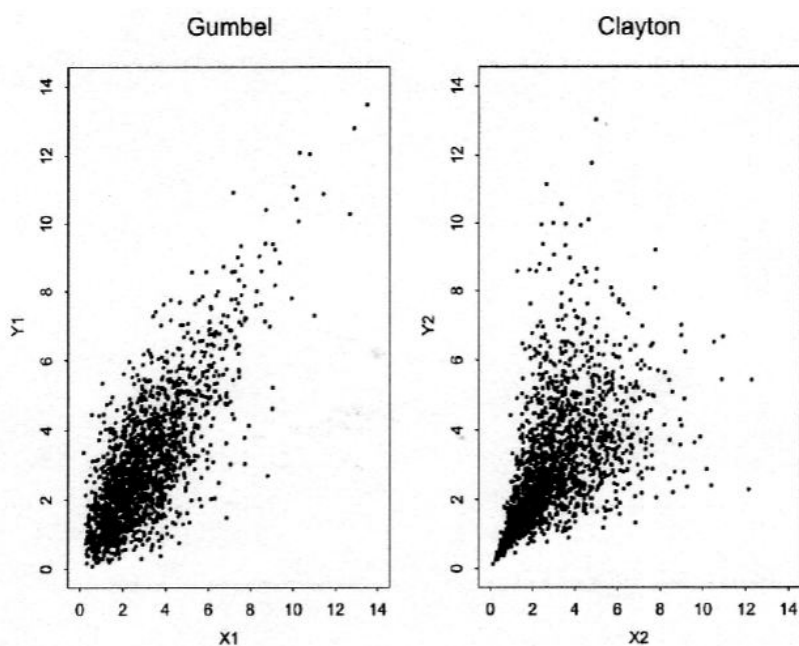


Abbildung 6.2: Zwei Stichproben mit Gamma(3,1)-Randverteilungen und linearer Korrelation 0.5, aber unterschiedlichen Copulas

6.5 Elliptische Copulas

Definition 6.5.1. Sei X ein d -dimensionaler Zufallsvektor, seien $\mu \in \mathbb{R}^d$ und $\Sigma \in \mathbb{R}^{d \times d}$ zwei Konstanten, und sei $\psi: [0, \infty) \rightarrow \mathbb{R}$ eine Funktion. Wenn $\phi_{X-\mu} = \psi(t^{\top} \Sigma t)$ gilt, wobei $\phi_{X-\mu}$ die charakteristische Funktion von $X - \mu$ ist, dann ist X ein *elliptisch* verteilter Zufallsvektor mit Parameter μ, Σ, ψ : $X \sim E_d(\mu, \Sigma, \psi)$.

ψ heißt *erzeugende Funktion* (oder *Generator*) von X .

Für $d = 1$ stimmen die elliptischen Verteilungen mit den symmetrischen Verteilungen überein, denn wenn X elliptisch verteilt ist, besitzt es die stochastische Darstellung $X = \mu + RAS$. Es gilt also

$$X - \mu = |X - \mu| \operatorname{sgn}(X - \mu)$$

und wenn $X - \mu \stackrel{d}{=} \mu - X$ (X also um μ symmetrisch ist), dann gilt

$$\mathbb{P}(\operatorname{sgn}(X - \mu) = 1) = \mathbb{P}(\operatorname{sgn}(X - \mu) = -1) = 0.5$$

und somit mit $RA = |X - \mu|$ und $S = \operatorname{sgn}(X - \mu)$

$$X - \mu = RAS = |X - \mu|\operatorname{sgn}(X - \mu).$$

Satz 6.10. (Stochastische Darstellung)

Ein d -dimensionaler Zufallsvektor X ist elliptisch verteilt, $X \sim E_d(\mu, \Sigma, \psi)$, mit $\operatorname{rang}(\Sigma) = k$, dann und nur dann, wenn es eine Matrix $A \in \mathbb{R}^{d \times k}$, $AA^\top = \Sigma$, sowie eine nichtnegative Zufallsvariable R und einen k -dimensionalen, auf der Einheitskugel $\mathcal{S}^{k-1} = \{z \in \mathbb{R}^k : z^\top z = 1\}$ gleichverteilten, Zufallsvektor S gibt, sodass R und S unabhängig sind und $X \stackrel{d}{=} \mu + RAS$.

Definition 6.5.2. Sei X ein d -dimensionaler Zufallsvektor mit Erwartungsvektor μ .

X heißt *radial symmetrisch* genau dann, wenn $X - \mu \stackrel{d}{=} \mu - X$.

Anmerkung: Eine elliptische Verteilung X ist radial symmetrisch, denn es existieren R , A und S , sodass $X - \mu = RAS$ und somit

$$\mu - X = -RAS = RA(-S) \stackrel{d}{=} RAS = X - \mu.$$

Definition 6.5.3. Sei $X \sim E_d(\mu, \Sigma, \psi)$ mit Verteilungsfunktion F und stetigen Randverteilungen F_1, F_2, \dots, F_d . Die eindeutige Copula C von F , $C(u) = F(F_1^{\leftarrow}(u_1), \dots, F_d^{\leftarrow}(u_d))$, heißt *elliptische Copula*.

Beispiel 6.5.4. Gauss'sche Copulas

Sei C_R^{Ga} die Copula einer d -dimensionalen Standardnormalverteilung mit Korrelationsmatrix R :

$$C_R^{Ga}(u) \stackrel{\text{Sklar}}{=} \phi_R^d(\phi^{-1}(u_1), \dots, \phi^{-1}(u_d)),$$

wobei ϕ_R^d die Gesamtverteilungsfunktion einer d -dimensionalen Normalverteilung mit Erwartungsvektor 0 und Korrelationsmatrix R und ϕ^{-1} die Inverse der Verteilungsfunktion einer univariaten Standardnormalverteilung ist. C_R^{Ga} heißt *Gauss'sche Copula*.

Da die Normalverteilung eine elliptische Verteilung ist, ist die Gauss'sche Copula C_R^{Ga} eine elliptische Copula.

Im bivariaten Fall gilt:

$$C_R^{Ga}(u_1, u_2) = \int_{-\infty}^{\phi^{-1}(u_1)} \int_{-\infty}^{\phi^{-1}(u_2)} \frac{1}{2\pi(1-\rho^2)^{1/2}} \exp\left\{\frac{-(x_1^2 - 2\rho x_1 x_2 + x_2^2)}{2(1-\rho^2)}\right\} dx_1 dx_2,$$

wobei $\rho \in (-1, 1)$. ◇

Definition 6.5.5. Sei $X \stackrel{d}{=} \mu + \frac{\sqrt{\nu}}{\sqrt{S}}Z \sim t_d(\mu, \nu, \Sigma)$, wobei $\mu \in \mathbb{R}^d$, $\nu \in \mathbb{N}$, $\nu > 1$, $S \sim \chi_\nu^2$ und $Z \sim N_d(0, \Sigma)$, und S und Z unabhängig sind. Es heißt, X hat eine d -dimensionale t -Verteilung mit Mittelwert μ (für $\nu > 1$) und Kovarianzmatrix $\operatorname{Cov}(X) = \frac{\nu}{\nu-2}\Sigma$ (für $\nu > 2$). $\operatorname{Cov}(X)$ existiert nicht für $\nu \leq 2$.

Definition 6.5.6. Die Copula $C_{\nu,R}^t$ von X heißt t -Copula. Für die t -Copula gilt:

$$C_{\nu,R}^t(u) = t_{\nu,R}^d(t_{\nu}^{-1}(u_1), \dots, t_{\nu}^{-1}(u_d)).$$

$R_{ij} = \frac{\Sigma_{ij}}{\sqrt{\Sigma_{ii}\Sigma_{jj}}}$, $i, j = 1, 2, \dots, d$, ist die Korrelationsmatrix von Y , $t_{\nu,R}^d$ ist die Verteilungsfunktion von $\frac{\sqrt{\nu}}{\sqrt{S}}Y$, wobei $S \sim \chi_{\nu}^2$ und $Y \sim N_d(0, R)$ unabhängig sind, und t_{ν} sind die Randverteilungen von $t_{\nu,R}^d$.

Bivariater Fall ($d = 2$):

$$C_{\nu,R}^t(u_1, u_2) = \int_{-\infty}^{t_{\nu}^{-1}(u_1)} \int_{-\infty}^{t_{\nu}^{-1}(u_2)} \frac{1}{2\pi(1-\rho^2)^{1/2}} \left\{ 1 + \frac{x_1^2 - 2\rho x_1 x_2 + x_2^2}{\nu(1-\rho^2)} \right\}^{-(\nu+2)/2} dx_1 dx_2,$$

für $\rho \in (-1, 1)$. R_{12} ist der lineare Korrelationskoeffizient der dazugehörigen bivariaten t_{ν} -Verteilung für $\nu > 2$.

6.6 Weitere Eigenschaften von Copulas

Definition 6.6.1. (Radiale Symmetrie oder Kugel-Symmetrie)

Ein Zufallsvektor X (oder eine Verteilungsfunktion) heißt *radial-symmetrisch* (oder *kugel-symmetrisch*) um den Punkt a , wenn $X - a \stackrel{d}{=} a - X$.

Beispiel 6.6.2. Ein elliptisch-verteilter Zufallsvektor $X \sim E_d(\mu, \Sigma, \psi) \in \mathbb{R}^d$ ist radial-symmetrisch um μ . ◇

Definition 6.6.3. (Radiale Symmetrie von Copulas)

Eine Copula C heißt radial-symmetrisch wenn

$$(U_1 - 0.5, \dots, U_d - 0.5) \stackrel{d}{=} (0.5 - U_1, \dots, 0.5 - U_d) \iff U \stackrel{d}{=} 1 - U,$$

wobei (U_1, U_2, \dots, U_d) ein Zufallsvektor mit Verteilungsfunktion C ist.

Für eine radial-symmetrische Copula gilt: $C = \hat{C}$.

Beispiel 6.6.4. Elliptische Copulas sind radial-symmetrisch.

Die Gumbel und Clayton Copulas sind es nicht, da die obere und untere Tail-Abhängigkeit nicht gleich ist. Es gilt nämlich für radial-symmetrische Copulas:

$$\lambda_U = \lim_{u \rightarrow 1-} \frac{\hat{C}(1-u, 1-u)}{1-u} = \lim_{t \rightarrow 0+} \frac{C(t, t)}{t} = \lambda_L.$$

◇

Definition 6.6.5. (Bedingte Copula Verteilungen)

Sei C eine zweidimensionale Copula und (U_1, U_2) ein Zufallsvektor mit Verteilungsfunktion C .

$$\begin{aligned} C_{U_2|U_1}(u_2|u_1) &:= \mathbb{P}(U_2 \leq u_2 | U_1 = u_1) \stackrel{f.s.}{=} \lim_{\delta \rightarrow 0+} \mathbb{P}(U_2 \leq u_2 | u_1 \leq U_1 \leq u_1 + \delta) \\ &= \lim_{\delta \rightarrow 0+} \frac{C(u_1 + \delta, u_2) - C(u_1, u_2)}{\delta} \stackrel{f.s.}{=} \frac{\partial C(u_1, u_2)}{\partial u_1}. \end{aligned}$$

$C_{U_2|U_1}$ ist eine Verteilung in $[0, 1]$; sie ist eine Gleichverteilung d.u.n.d., wenn C die Unabhängigkeitscopula ist (siehe Nelsen 1999).

Die Dichtefunktion einer Copula

Copulas haben nicht immer eine Dichtefunktion. Z.B. die Co-Monotonie-Copula M bzw. die Anti-Monotonie-Copula W haben keine Dichtefunktion.

Wenn die Dichtefunktion c einer Copula C existiert, dann gilt

$$c(u_1, u_2, \dots, u_d) = \frac{\partial C(u_1, u_2, \dots, u_d)}{\partial u_1 \partial u_2 \dots \partial u_d}.$$

Sei C die Copula einer Gesamtverteilung F mit stetigen Randverteilungsfunktionen F_1, \dots, F_d . Dann kann die Gleichung

$$C(u_1, \dots, u_d) = F(F_1^{\leftarrow}(u_1), \dots, F_d^{\leftarrow}(u_d))$$

differenziert werden um die Dichte c von C zu erhalten:

$$c(u_1, \dots, u_d) = \frac{f(F_1^{-1}(u_1), \dots, F_d^{-1}(u_d))}{f_1(F_1^{-1}(u_1)) \dots f_d(F_d^{-1}(u_d))}. \quad (6.3)$$

f ist die Gesamtdichtefunktion, f_i sind die Dichtefunktionen der Randverteilungen, $1 \leq i \leq d$, und F_i^{-1} ist die inverse Funktion von F_i .

Definition 6.6.6. Ein Zufallsvektor X heißt *vertauschbar* ("exchangeable"), wenn $(X_1, \dots, X_d) \stackrel{d}{=} (X_{\pi(1)}, \dots, X_{\pi(d)})$ für jede Permutation $(\pi(1), \pi(2), \dots, \pi(d))$ von $(1, 2, \dots, d)$.

Definition 6.6.7. Eine Copula C heißt *vertauschbar*, wenn sie die Gesamtverteilung eines vertauschbaren Zufallsvektors (mit Gleichverteilungen als Randverteilungen) ist.

Für eine solche Copula gilt:

$$C(u_1, u_2, \dots, u_d) \stackrel{d}{=} C(u_{\pi(1)}, u_{\pi(2)}, \dots, u_{\pi(d)})$$

für jede Permutation $(\pi(1), \pi(2), \dots, \pi(d))$ von $(1, 2, \dots, d)$.

Beispiele von vertauschbaren Copulas sind z.B.: Gumbel, Clayton, Gauss'sche Copula C_R^{Ga} , t-Copula $C_{\nu, R}^{\top}$, für den Fall, dass R eine Äqui-Korrelationsmatrix ist: $R = \rho J_d + (1 - \rho) I_d$. $J_d \in \mathbb{R}^{d \times d}$ ist eine Matrix bestehend aus lauter Einsen und $I_d \in \mathbb{R}^{d \times d}$ ist die d -dimensionale Einheitsmatrix.

Für bivariate vertauschbare Copulas gilt:

$$\mathbb{P}(U_2 \leq u_2 | U_1 = u_1) = \mathbb{P}(U_1 \leq u_2 | U_2 = u_1).$$

Satz 6.11. Sei $(X_1, X_2)^{\top}$ ein normalverteilter Zufallsvektor. Dann gilt:

$$\lambda_U(X_1, X_2) = \lambda_L(X_1, X_2) = 0.$$

Beweis. Sei $R = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$ die Korrelationsmatrix und C_ρ^{Ga} die Copula von $(X_1, X_2)^\top$.

$$\begin{aligned} \lambda_L &= \lim_{u \rightarrow 0^+} \frac{C(u, u)}{u} = \lim_{u \rightarrow 0^+} \frac{dC(u, u)}{du} = \lim_{u \rightarrow 0^+} \frac{1}{du} \left(\frac{\partial C(u, u)}{\partial u_1} du + \frac{\partial C(u, u)}{\partial u_2} du \right) \\ &= \lim_{u \rightarrow 0^+} [\mathbb{P}(U_2 \leq u | U_1 = u) + \mathbb{P}(U_1 \leq u | U_2 = u)] = 2 \lim_{u \rightarrow 0^+} \mathbb{P}(U_2 \leq u | U_1 = u) \\ &\stackrel{\phi^{-1}(U_i)=X_i}{=} 2 \lim_{u \rightarrow 0^+} \mathbb{P}(\underbrace{\phi^{-1}(U_2)}_{=:t} \leq \phi^{-1}(u) | \phi^{-1}(U_1) = \phi^{-1}(u)) = 2 \lim_{t \rightarrow -\infty} \mathbb{P}(X_2 \leq t | X_1 = t) \end{aligned}$$

oBdA: $\mu = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, $\Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \Rightarrow (X_1, X_2) \sim N_2(0, \Sigma) \Rightarrow X_2 | X_1 = x \sim N(\rho x, 1 - \rho^2)$.

$$\lambda_L = 2 \lim_{t \rightarrow -\infty} \phi \left(\frac{t - \rho t}{\sqrt{1 - \rho^2}} \right) = 2 \lim_{t \rightarrow -\infty} \phi \left(\frac{(\sqrt{1 - \rho})^2 t}{\sqrt{1 - \rho} \sqrt{1 + \rho}} \right) = 2 \lim_{t \rightarrow -\infty} \phi \left(t \sqrt{\frac{1 - \rho}{1 + \rho}} \right) = 0.$$

$N_2(\mu, \Sigma)$ ist radial-symmetrisch $\Rightarrow \lambda_U = \lambda_L = 0$. □

Korollar 6.6.8. Sei $(X_1, X_2)^\top$ ein Zufallsvektor mit stetigen Randverteilungen und einer Gauss'schen Copula C_ρ^{Ga} , wobei ρ der Koeffizient der linearen Korrelation zwischen X_1 und X_2 ist. Dann gilt: $\lambda_U(X_1, X_2) = \lambda_L(X_1, X_2) = 0$.

Satz 6.12. Sei $(X_1, X_2)^\top$ ein t -verteilter Zufallsvektor mit ν Freiheitsgraden, Mittelwert 0 und linearer Korrelationsmatrix R : $(X_1, X_2)^\top \sim t_2(0, \nu, R)$. Für $R_{12} > -1$ gilt:

$$\lambda_U(X_1, X_2) = \lambda_L(X_1, X_2) = 2\bar{t}_{\nu+1} \left(\sqrt{\nu+1} \frac{\sqrt{1 - R_{12}}}{\sqrt{1 + R_{12}}} \right).$$

Beweis. Ähnlich wie der Beweis von Satz 6.11. Hinweis:

$$X_2 | X_1 = x \sim \left(\frac{\nu+1}{\nu+x^2} \right)^{1/2} \frac{X_2 - \rho x}{\sqrt{1 - \rho^2}} \sim t_{\nu+1}.$$

$$\begin{aligned} \lambda_L &= \lim_{u \rightarrow 0^+} \frac{C(u, u)}{u} = \lim_{u \rightarrow 0^+} \frac{dC(u, u)}{du} = \lim_{u \rightarrow 0^+} \frac{1}{du} \left(\frac{\partial C(u, u)}{\partial u_1} du + \frac{\partial C(u, u)}{\partial u_2} du \right) \\ &= \lim_{u \rightarrow 0^+} [\mathbb{P}(U_2 \leq u | U_1 = u) + \mathbb{P}(U_1 \leq u | U_2 = u)] = 2 \lim_{u \rightarrow 0^+} \mathbb{P}(U_2 \leq u | U_1 = u) \\ &\stackrel{\phi^{-1}(U_i)=X_i}{=} 2 \lim_{u \rightarrow 0^+} \mathbb{P}(\underbrace{\phi^{-1}(U_2)}_{=:x} \leq \phi^{-1}(u) | \phi^{-1}(U_1) = \phi^{-1}(u)) = 2 \lim_{x \rightarrow -\infty} \mathbb{P}(X_2 \leq x | X_1 = x) \\ &= 2 \lim_{x \rightarrow -\infty} t_{\nu+1} \left(\left(\frac{\nu+1}{\nu+x^2} \right)^{1/2} \frac{x - \rho x}{\sqrt{1 - \rho^2}} \right) = 2 \lim_{x \rightarrow -\infty} t_{\nu+1} \left(\frac{\sqrt{\nu+1}}{-x\sqrt{1+\nu/x^2}} \frac{x(1-\rho)}{\sqrt{1-\rho}\sqrt{1+\rho}} \right) \\ &= 2 \lim_{x \rightarrow -\infty} t_{\nu+1} \left(-\frac{\sqrt{\nu+1}}{\sqrt{1+\nu/x^2}} \sqrt{\frac{1-\rho}{1+\rho}} \right) = 2\bar{t}_{\nu+1} \left(\sqrt{\nu+1} \sqrt{\frac{1-\rho}{1+\rho}} \right). \end{aligned}$$

□

Korollar 6.6.9. Sei $(X_1, X_2)^\top$ ein Zufallsvektor mit stetigen Randverteilungen und einer t -copula $C_{\nu,R}^t$ mit ν Freiheitsgraden und einer Korrelationsmatrix R . Dann gilt:

$$\lambda_U(X_1, X_2) = \lambda_L(X_1, X_2) = 2\bar{t}_{\nu+1} \left(\sqrt{\nu+1} \frac{\sqrt{1-R_{12}}}{\sqrt{1+R_{12}}} \right).$$

Satz 6.13. Sei $(X_1, X_2)^\top$ ein Zufallsvektor mit stetigen Randverteilungen und einer Gauss'schen Copula C_ρ^{Ga} , wobei ρ der Koeffizient der linearen Korrelation zwischen X_1 und X_2 ist. Dann gilt:

$$\rho_\tau(X_1, X_2) = \frac{2}{\pi} \arcsin R_{12} \quad \text{und} \quad \rho_S(X_1, X_2) = \frac{6}{\pi} \arcsin \frac{R_{12}}{2}.$$

Satz 6.14. Sei $(X_1, X_2)^\top$ ein Zufallsvektor mit stetigen Randverteilungen und einer elliptischen Copula $C_{\mu,\Sigma,\psi}^E$. Dann gilt:

$$\rho_\tau(X_1, X_2) = \frac{2}{\pi} \arcsin R_{12}, \quad \text{wobei} \quad R_{12} = \frac{\Sigma_{12}}{\sqrt{\Sigma_{11}\Sigma_{22}}}.$$

Korollar 6.6.10. Sei $X \sim E_d(\mu, \Sigma, \psi)$ ein elliptisch-verteilter Vektor mit stetigen Randverteilungsfunktionen. Dann gilt:

$$\rho_\tau(X_i, X_j) = \frac{2}{\pi} \arcsin R_{ij}, \quad \text{wobei} \quad R_{ij} = \frac{\Sigma_{ij}}{\sqrt{\Sigma_{ii}\Sigma_{jj}}} \quad \text{für } i, j = 1, 2, \dots, d.$$

Beweise von Satz 6.13, Satz 6.14 und Korollar 6.6.10: siehe McNeil et al. (2005).

6.7 Archimedische Copulas

Nachteile elliptischer Copulas:

- I.A. keine Darstellung in geschlossener Form möglich
- kugel-symmetrisch

6.7.1 Bivariate Archimedische Copulas

Definition 6.7.1. Sei $\phi: [0, 1] \rightarrow [0, +\infty]$ stetig und streng monoton fallend, sodass $\phi(1) = 0$. Die pseudo-inverse Funktion $\phi^{[-1]}: [0, \infty] \rightarrow [0, 1]$ von ϕ wird folgendermaßen definiert:

$$\phi^{[-1]}(t) = \begin{cases} \phi^{-1}(t) & 0 \leq t \leq \phi(0) \\ 0 & \phi(0) \leq t \leq \infty \end{cases}$$

$\phi^{[-1]}$ ist stetig und monoton fallend in $[0, \infty]$, streng monoton fallend in $[0, \phi(0)]$ und es gilt:

$$\phi^{[-1]}(\phi(u)) = u \text{ für } u \in [0, 1]$$

$$\phi(\phi^{[-1]}(t)) = \begin{cases} t & 0 \leq t \leq \phi(0) \\ \phi(0) & \phi(0) \leq t \leq +\infty \end{cases}$$

Falls $\phi(0) = +\infty$, dann $\phi^{[-1]} = \phi^{-1}$.

Satz 6.15. Sei $\phi: [0, 1] \rightarrow [0, +\infty]$ stetig und streng monoton fallend in $[0, 1]$, sodass $\phi(1) = 0$, und sei $\phi^{[-1]}$ die pseudo-inverse Funktion von ϕ . Sei $C: [0, 1]^2 \rightarrow [0, 1]$, sodass $C(u_1, u_2) = \phi^{[-1]}(\phi(u_1) + \phi(u_2))$.
 C ist eine Copula dann und nur dann, wenn ϕ konvex ist. Copulas dieser Form heißen Archimedische Copulas. ϕ heißt Generator von C . Falls $\phi(0) = +\infty$, dann $\phi^{[-1]} = \phi^{-1}$ und $C(u_1, u_2) = \phi^{-1}(\phi(u_1) + \phi(u_2))$.

Beweis. Siehe Nelsen 1999. □

Beispiel 6.7.2. Gumbel Copulas

Sei $\phi(t) = (-\ln t)^\theta$, $\theta \geq 1$, $t \in [0, 1]$.

$$C_\theta^{Gu}(u_1, u_2) = \exp\left(-[(-\ln u_1)^\theta + (-\ln u_2)^\theta]^{1/\theta}\right)$$

ist die Gumbel Copula mit Parameter θ .

Für $\theta = 1$: $C_1^{Gu} = u_1 u_2$.

$\lim_{\theta \rightarrow \infty} C_\theta^{Gu} = M(u_1, u_2) = \min\{u_1, u_2\}$.

Die Gumbel Copulas haben eine obere Tail Abhängigkeit. ◇

Beispiel 6.7.3. Clayton Copulas

Sei $\phi(t) = (t^{-\theta} - 1)/\theta$, $\theta > 0$.

$$C_\theta^{Cl}(u_1, u_2) = \left(u_1^{-\theta} + u_2^{-\theta} - 1\right)^{-1/\theta}$$

ist die Clayton Copula mit Parameter θ .

$\lim_{\theta \rightarrow 0} C_\theta^{Cl} = u_1 u_2$ und $\lim_{\theta \rightarrow \infty} C_\theta^{Cl} = M = \min\{u_1, u_2\}$.

Die Clayton Copulas haben eine untere Tail Abhängigkeit. ◇

Beispiel 6.7.4. $\phi(t) = 1 - t$, $t \in [0, 1]$. $\phi^{[-1]}(t) = \max\{1 - t, 0\}$. $\phi(t)$ ist stetig, streng monoton fallend und konvex, erfüllt also die Voraussetzungen von Satz 6.15 und es gilt somit

$$C_\phi(u_1, u_2) = \max\{u_1 + u_2 - 1, 0\} = W(u_1, u_2).$$

D.h. die untere Fréchet Schranke ist eine Archimedische Copula. ◇

Satz 6.16. Sei $(X_1, X_2)^\top$ ein Zufallsvektor mit stetigen Randverteilungen und einer Archimedischen Copula C , generiert von ϕ . Dann gilt $\rho_\tau(X_1, X_2) = 1 + 4 \int_0^1 \frac{\phi(t)}{\phi'(t)} dt$.

Beweis. Siehe Nelsen 1999. □

Beispiel 6.7.5. Kendall's Tau für Gumbel und Clayton Copulas

Gumbel Copulas: $\phi(t) = (\ln t)^\theta, \theta \geq 1$.

$$\rho_\tau(\theta) = 1 + 4 \int_0^1 \frac{\phi(t)}{\phi'(t)} dt = 1 - \frac{1}{\theta}.$$

Clayton Copulas: $\phi(t) = (t^{-\theta} - 1)/\theta, \theta > 0$.

$$\rho_\tau(\theta) = 1 + 4 \int_0^1 \frac{\phi(t)}{\phi'(t)} dt = \frac{\theta}{\theta+2}. \quad \diamond$$

6.7.2 Multivariate Archimedische Copulas

Definition 6.7.6. Eine Funktion $g: [0, \infty) \rightarrow [0, \infty)$ heißt *vollständig monoton*, wenn g stetig ist und folgende Ungleichung für $k = 0, 1, 2, \dots$ gilt:

$$(-1)^k \left(\frac{d^k}{ds^k} g(s) \right) \Big|_{s=t} \geq 0, \forall t \in (0, \infty).$$

Satz 6.17. (Kimberling 1974)

Sei $\phi: [0, 1] \rightarrow [0, \infty]$ eine stetige und streng monoton fallende Funktion, sodass $\phi(0) = \infty$ und $\phi(1) = 0$. Die Funktion $C: [0, 1]^d \rightarrow [0, 1]$, $C(u) = \phi^{-1}(\phi(u_1) + \phi(u_2) + \dots + \phi(u_d))$, ist eine Copula für $d \geq 2$, dann und nur dann, wenn ϕ^{-1} vollständig monoton in $[0, \infty)$ ist.

Lemma 6.7.7. Eine Funktion $\psi: [0, \infty) \rightarrow [0, \infty)$ ist die Laplace-Stieltjes Transformation einer Verteilungsfunktion G in $[0, \infty)$ ($\psi(s) = \int_0^\infty e^{-sx} dG(x), s \geq 0$), dann und nur dann, wenn ψ vollständig monoton und $\psi(0) = 1$.

Satz 6.18. Sei G eine Verteilungsfunktion in $[0, \infty)$, sodass $G(0) = 0$ und sei ψ die Laplace-Stieltjes Transformation von G ,

$$\psi(s) = \int_0^\infty e^{-sx} dG(x), \text{ für } s \geq 0.$$

Sei X eine ZV mit Verteilungsfunktion G und seien U_1, U_2, \dots, U_d bedingt unabhängige Zufallsvariablen in $[0, 1]$ für ein gegebenes $X = x$ mit folgender bedingter Verteilungsfunktion:

$$F_{U_k|X=x}(u) = \exp(-x\psi^{-1}(u)) \text{ für } u \in [0, 1].$$

Es gilt dann

$$\mathbb{P}(U_1 \leq u_1, U_2 \leq u_2, \dots, U_d \leq u_d) = \psi(\psi^{-1}(u_1) + \psi^{-1}(u_2) + \dots + \psi^{-1}(u_d)).$$

und die Verteilungsfunktion von $U = (U_1, U_2, \dots, U_d)$ ist eine Archimedische Copula mit Generator ψ^{-1} .

Beweis. Sei F die Gesamtverteilungsfunktion von (U_1, U_2, \dots, U_d) .

Z.z.: $F(u_1, u_2, \dots, u_d) = \psi(\psi^{-1}(u_1) + \psi^{-1}(u_2) + \dots + \psi^{-1}(u_d)), \forall (u_1, \dots, u_d) \in [0, 1]^d$.

$$\begin{aligned} F(u_1, \dots, u_d) &= \mathbb{P}(U_1 \leq u_1, \dots, U_d \leq u_d) = \int_0^\infty \mathbb{P}(U_1 \leq u_1, \dots, U_d \leq u_d) dG(x) \\ &\stackrel{(*)}{=} \int_0^\infty \prod_{i=1}^d \exp\{-x\psi^{-1}(u_i)\} dG(x) = \int_0^\infty \exp\{-x \sum_{i=1}^d \psi^{-1}(u_i)\} dG(x) \\ &= \psi\left(\sum_{i=1}^d \psi^{-1}(u_i)\right). \end{aligned}$$

Das ist laut Satz 6.17 eine Copula.

(*) U_i ist bedingt unabhängig mit gegebener bed. Verteilung. □

6.7.3 Vorteile und Nachteile Archimedischer Copulas

- Modellierung einer breiteren Klasse von Abhängigkeitsstrukturen
- Darstellung in geschlossener Form möglich
- Wenige freie Parameter vorhanden
- Technische Voraussetzungen für die Generator Funktionen multivariater Archimedische Copulas.

6.8 Simulation elliptischer und archimedischer Copulas

6.8.1 Simulation von Gauss'schen Copulas und t-Copulas

Sei $R \in \mathbb{R}^{d \times d}$ eine symmetrische positiv definite Matrix. Sei $AA^\top = R$ die Cholesky-Zerlegung von R ($A \in \mathbb{R}^{d \times d}$). Falls $Z_1, Z_2, \dots, Z_d \sim N(0, 1)$ unabhängig sind, dann gilt $\mu + AZ \sim N_d(\mu, R)$.

Algorithmus 6.1 zur Erzeugung eines Zufallsvektors $U = (U_1, U_2, \dots, U_d)$, dessen Verteilungsfunktion die Copula C_R^{Ga} ist.

- Berechne die Cholesky-Zerlegung A von R : $R = AA^\top$.
 - Simuliere d unabhängige standard-normalverteilte ZV $Z_1, Z_2, \dots, Z_d \sim N(0, 1)$.
 - Setze $X = AZ \sim N_d(0, R)$.
 - Setze $U_k = \phi(X_k) \sim U(0, 1)$ für $k = 1, 2, \dots, d$, wobei ϕ die Verteilungsfunktion der standard-Normalverteilung ist.
 - $U = (U_1, U_2, \dots, U_d)$ hat Verteilungsfunktion C_R^{Ga} (Satz (6.2)).
-

Algorithmus 6.2 zur Erzeugung eines Zufallsvektors $U = (U_1, U_2, \dots, U_d)$, dessen Verteilungsfunktion die Copula $C_{\nu, R}^t$ ist.

- Berechne die Cholesky-Zerlegung A von R : $R = AA^\top$.
 - Simuliere d unabhängige standard-normalverteilte ZV $Z_1, Z_2, \dots, Z_d \sim N(0, 1)$.
 - Simuliere eine ZV $S \sim \chi_\nu^2$ unabhängig von Z_1, \dots, Z_d .
 - Setze $Y = AZ \sim N_d(0, R)$.
 - Setze $X = \frac{\sqrt{\nu}}{\sqrt{S}}Y$ (Def. einer multivariaten t -Verteilung).
 - Setze $U_k = t_\nu(X_k) \sim U(0, 1)$ für $k = 1, 2, \dots, d$, wobei t_ν die Verteilungsfunktion einer standard- t -Verteilung mit ν Freiheitsgraden ist.
 - $U = (U_1, U_2, \dots, U_d)$ hat Verteilungsfunktion $C_{\nu, R}^t$ (Satz (6.2)).
-

6.8.2 Simulation von Gumbel und Clayton Copulas

Aus Satz 6.18 lässt sich ein Algorithmus zur Erzeugung dieser Copulas konstruieren.

Algorithmus 6.3 zur Erzeugung eines Zufallsvektors $U = (U_1, U_2, \dots, U_d)$, dessen Verteilungsfunktion die Archimedische Copula $C(u) = \phi^{-1}(\phi(u_1) + \phi(u_2) + \dots + \phi(u_d))$ mit Generator ϕ ist.

- Simuliere eine Variable X mit Verteilungsfunktion G , sodass die Laplace-Stieltjes Transformation ψ von G die inverse Funktion des Generators der gesuchten Copula ist, $\psi = \phi^{-1}$.
 - Simuliere unabhängige gleichverteilte Zufallsvariablen V_1, V_2, \dots, V_d in $[0, 1]$.
 - Setze $U = (\psi(-\ln(V_1)/X), \psi(-\ln(V_2)/X), \dots, \psi(-\ln(V_d)/X))$. U hat Verteilungsfunktion $C(u)$.
-

Beweis. von Punkt 3 aus Algorithmus 6.3.

Z.z.: $U = (\psi(-\ln(V_1)/X), \dots, \psi(-\ln(V_d)/X))$ hat die Verteilungsfunktion $C(u) = \phi^{-1}(\sum_{i=1}^d \phi(u_i))$, d.h. die Bedingungen von Satz 6.18 sind erfüllt:

$$F_{U_i|X=x}(u_i) = \exp\{-x\phi^{-1}(u_i)\} \quad \forall i, \forall x.$$

$$\begin{aligned} \mathbb{P}(U_1 \leq u_1, \dots, U_d \leq u_d) &= \mathbb{P}(\psi(-\ln(V_1)/x) \leq u_1, \dots, \psi(-\ln(V_d)/x) \leq u_d) \\ &\stackrel{U_i \searrow}{=} \mathbb{P}(-\ln(V_1)/x \geq \psi^{-1}(u_1), \dots, -\ln(V_d)/x \geq \psi^{-1}(u_d)) \\ &= \mathbb{P}(V_1 \leq \exp\{-x\psi^{-1}(u_1)\}, \dots, V_d \leq \exp\{-x\psi^{-1}(u_d)\}) \\ &\stackrel{V_i \stackrel{u_i}{\sim} U(0,1)}{=} \prod_{i=1}^d \exp\{-x\psi^{-1}(u_i)\}. \quad (**) \end{aligned}$$

Damit gilt:

$$\begin{aligned} F_{U_i|X=x}(u_i) &= \mathbb{P}(U_i \leq u_i | X = x) = \mathbb{P}(U_1 \leq 1, \dots, U_i \leq u_i, \dots, U_d \leq 1 | X = x) \\ &\stackrel{(**)}{=} \exp\{-x\psi^{-1}(u_i)\} \underbrace{\prod_{j=1, j \neq i}^d \exp\{-x\psi^{-1}(1)\}}_{=1} = \exp\{-x\psi^{-1}(u_i)\}. \end{aligned}$$

□

Der Generator $\phi(t) = (t^{-\theta} - 1)/\theta$, $\theta > 0$ erzeugt die Clayton Copula C_θ^{Cl} . Aber auch $\tilde{\phi}(t) = t^{-\theta} - 1$ ist ein Generator der Clayton Copula.

Für $X \sim \text{Gamma}(1/\theta, 1)$, d.h. $f_X(x) = x^{1/\theta-1}e^{-x}/\Gamma(1/\theta)$, gilt:

$$\begin{aligned} \mathbb{E}(e^{-sX}) &= \int_0^\infty e^{-sx} \frac{1}{\Gamma(1/\theta)} x^{1/\theta-1} e^{-x} dx = \int_0^\infty \frac{e^{-(s+1)x} x^{1/\theta-1}}{\Gamma(1/\theta)} dx \\ &\stackrel{(s+1)x=t}{=} \frac{1}{\Gamma(1/\theta)} \int_0^\infty e^{-t} \frac{t^{1/\theta-1}}{(s+1)^{1/\theta-1}} \frac{dt}{s+1} = \frac{(s+1)^{-1/\theta}}{\Gamma(1/\theta)} \underbrace{\int_0^\infty e^{-t} t^{1/\theta-1} dt}_{\Gamma(1/\theta)} \\ &= (s+1)^{-1/\theta} = \tilde{\phi}^{-1}(s). \end{aligned}$$

Algorithmus 6.4 zur Erzeugung eines Zufallsvektors $U = (U_1, U_2, \dots, U_d)$, dessen Verteilungsfunktion die Clayton Copula C_θ^{Cl} ist.

- Simuliere $X \sim \text{Gamma}(1/\theta, 1)$.
- Simuliere unabhängige gleichverteilte Zufallsvariablen V_1, V_2, \dots, V_d in $[0, 1]$.
- Die Verteilungsfunktion des Vektors

$$U = (\psi(-\ln(V_1)/X), \psi(-\ln(V_2)/X), \dots, \psi(-\ln(V_d)/X))$$

ist die Clayton Copula C_θ^{Cl} .

Für die Simulation von Gumbel Copulas C_θ^{Gu} gelte folgendes:

Sei X eine positive stabile ZV, $X \sim St(1/\theta, 1, \gamma, 0)$, mit $\gamma = (\cos(\pi/(2\theta)))^\theta$, $\theta > 1$.

Die Laplace-Stieltjes Transformation von F_X ist $\psi(t) = \exp(-t^{1/\theta})$.

Für eine Simulation von $Z \sim St(\alpha, \beta, 1, 0)$ siehe Nolan 2002.

Für $\gamma \neq 1$ gilt: $X = \delta + \gamma Z \sim St(\alpha, \beta, \gamma, \delta)$.

Eine Alternative zur äußerst komplizierten Simulation von stabilen ZV ist folgende (nur für $d = 2!$):

Sei $\theta \geq 1$ und $\bar{F}(x) = 1 - F(x) = \exp(-x^{1/\theta})$ für $x \geq 0$.

Sei $V \sim U(0, 1)$ und S eine von V unabhängige ZV mit Dichtefunktion

$$h(s) = (1 - 1/\theta + s/\theta) \exp(-s).$$

Sei $(Z_1, Z_2)^\top = (VS^\theta, (1 - V)S^\theta)$.

Die Verteilungsfunktion von $(\bar{F}(Z_1), \bar{F}(Z_2))^\top$ ist C_θ^{Gu} .

Algorithmus 6.5 zur Erzeugung eines Zufallsvektors $U = (U_1, U_2)$, dessen Verteilungsfunktion die Gumbel Copula C_θ^{Gu} ist.

- Simuliere zwei unabhängige ZV $V_1, V_2 \sim U(0, 1)$.
 - Simuliere zwei unabhängige ZV $W_1 \sim \Gamma(1, 1)$, $W_2 \sim \Gamma(2, 1)$
 - Setze $S = I_{V_2 \leq 1/\theta} W_1 + I_{V_2 > 1/\theta} W_2$.
 - Setze $(Z_1, Z_2) = (V_1 S^\theta, (1 - V_1) S^\theta)$.
 - Die Verteilungsfunktion von $U = (\exp(-Z_1^{1/\theta}), \exp(-Z_2^{1/\theta}))^\top$ ist C_θ^{Gu} .
-

6.9 Schätzung von Copulas

Gegeben sei ein Satz multi-dimensionaler Daten. Gesucht ist eine Copula und die Randverteilungen die diesem Datensatz am besten entsprechen.

1. Frage: Welche Familie von (bekannten) Copulas eignet sich am besten?

Antwort: Visueller Vergleich der graphischen Darstellungen von Daten bzw. bekannten Copulas, Berechnung der empirischen Koeffizienten der Tail-Abhängigkeit und Auswahl von dazu passenden Copula Familien.

2. Frage: Schätzung der Parameter einer vorsezifizierten Copula Familie.

Gegeben: Eine Stichprobe $\{X_1, X_2, \dots, X_n\}$ aus einer Gesamtverteilung F mit stetigen Randverteilungen F_1, F_2, \dots, F_d .

Gesucht: Ein Schätzer $\hat{\theta}$ des Parameter-Vektors θ der eindeutigen Copula C_θ , für die $F(x) = C(F_1(x_1), \dots, F_d(x_d))$ gilt.

6.9.1 Die Schätzer $\hat{\theta}$ für C_R^{Ga} , $C_{\nu,R}^t$, C_{θ}^{Cl} und C_{θ}^{Gu}

$$\begin{aligned}
C_R^{Ga} &= \phi_R^d(\phi^{-1}(u_1), \dots, \phi^{-1}(u_d)) & R_{ij} &= \sin(\pi(\rho_{\tau})_{ij}/2) \\
C_{\nu,R}^t &= t_{\nu,R}^d(t_{\nu}^{-1}(u_1), \dots, t_{\nu}^{-1}(u_d)) & R_{ij} &= \sin(\pi(\rho_{\tau})_{ij}/2) \\
C_{\theta}^{Gu}(u) &= \exp(-[(-\ln u_1)^{\theta} + \dots + (-\ln u_d^{\theta})]^{1/\theta}) & \theta &= 1/(1 - (\rho_{\tau})_{ij}) \\
C_{\theta}^{Cl}(u) &= (u_1^{-\theta} + \dots + u_d^{-\theta} - d + 1)^{-1/\theta} & \theta &= 2(\rho_{\tau})_{ij}/(1 - (\rho_{\tau})_{ij}),
\end{aligned}$$

wobei

$$\begin{aligned}
(\rho_{\tau})_{ij} &= \rho_{\tau}(X_{k,i}, X_{k,j}) \\
&= \mathbb{P}((X_{k,i} - X_{l,i})(X_{k,j} - X_{l,j}) > 0) - \mathbb{P}((X_{k,i} - X_{l,i})(X_{k,j} - X_{l,j}) < 0) \\
&= \mathbb{E}(\text{sgn}((X_{k,i} - X_{l,i})(X_{k,j} - X_{l,j}))).
\end{aligned}$$

Standard Schätzer für Kendalls Tau:

$$\hat{\rho}_{\tau ij} = \binom{n}{2}^{-1} \sum_{1 \leq k < l \leq n} \text{sgn}((X_{k,i} - X_{l,i})(X_{k,j} - X_{l,j})).$$

Bei der Schätzung für die Gauss'sche und t-Copula kann es zu Problemen kommen, wenn $\hat{R} = (\hat{R}_{ij})$ nicht symmetrisch und positiv semidefinit ist.

Auch bei der Schätzung für die Gumbel und Clayton Copula können Schwierigkeiten auftreten, da viele verschiedene Werte $\hat{\rho}_{ij}$ für die Berechnung eines einzigen Parameters $\hat{\theta}$ zur Verfügung stehen. Gibt es also eine große Streuung in den Werten von $\hat{\rho}_{ij}$, so ist die Verwendung der Gumbel und Clayton Copula nicht angebracht.

6.9.2 Schätzung von Gauss'schen Copulas und t-Copulas

Wie schon angeführt kann es passieren, dass $\hat{R} = (\hat{R}_{ij})$ nicht positiv definit ist, wobei

$$\hat{R}_{ij} = \sin\left(\frac{\pi}{2}(\hat{\rho}_{\tau})_{ij}\right).$$

\hat{R} wird durch eine Korrelationsmatrix R^* ersetzt, wobei R^* „unweit“ von \hat{R} liegt.

Algorithmus 6.6 Eigenwert-Ansatz, siehe Rousseeuw und Molenberghs 1993

- Berechne die Spektralzerlegung $\hat{R} = \Gamma \Lambda \Gamma^{\top}$ von \hat{R} , wobei Λ eine Diagonalmatrix ist, die die Eigenwerte von \hat{R} enthält, und Γ eine orthogonale Matrix, deren Spalten den Eigenvektoren von \hat{R} entsprechen.
 - Ersetze die negativen Eigenwerte in λ durch eine kleine Zahl $\delta > 0$ um $\tilde{\Lambda}$ zu erhalten.
 - Berechne $\tilde{R} = \Gamma \tilde{\Lambda} \Gamma^{\top}$. \tilde{R} ist symmetrisch und positiv definit, aber nicht unbedingt eine Korrelationsmatrix, weil die Diagonal-Elemente \tilde{R}_{ii} ungleich 1 sein könnten.
 - Setze $\hat{R} = D \tilde{R} D$, wobei D eine diagonale Matrix mit $D_{k,k} = 1/\sqrt{\tilde{R}_{k,k}}$ ist.
-

***t*-Copulas: Schätzung des Parameters ν der Freiheitsgrade**

1. Schätzung der univariaten Randverteilungsfunktionen F_1, F_2, \dots, F_d . Seien $\hat{F}_1, \dots, \hat{F}_d$ die dazugehörigen Schätzer.
2. Bildung einer Pseudo-Stichprobe der Copula:

$$\hat{U}_k = (\hat{U}_{k,1}, \hat{U}_{k,2}, \dots, \hat{U}_{k,d}) := (\hat{F}_1(X_{k,1}), \dots, \hat{F}_d(X_{k,d})),$$

für $k = 1, 2, \dots, n$ (siehe Genest und Rivest 1993).

\hat{F}_k kann folgendermaßen erzeugt werden:

- Parametrische Schätzung: \hat{F}_k ist eine parametrische Verteilungsfunktion, wobei der Parameter zB. mit einem Maximum Likelihood (ML) Ansatz geschätzt wird.
- Nicht Parametrische Schätzung: \hat{F}_i ist die empirische Verteilungsfunktion $\hat{F}_i(x) = \frac{1}{n+1} \sum_{t=1}^n I_{X_{t,i} \leq x}$, $1 \leq i \leq d$.

ML-Schätzung von ν : $\nu = \arg \max_{\xi} \ln L(\xi; \hat{U}_1, \hat{U}_2, \dots, \hat{U}_n)$, wobei

$$L(\xi; \hat{U}_1, \hat{U}_2, \dots, \hat{U}_n) = \prod_{k=1}^n c_{\xi,R}^{\top}(\hat{U}_k)$$

und $c_{\xi,R}^{\top}$ die Dichte der *t*-Copula $C_{\xi,R}^t$ ist.

Daraus folgt:

$$\begin{aligned} \ln L(\xi; \hat{U}_1, \hat{U}_2, \dots, \hat{U}_d) &\stackrel{(6.3)}{=} \sum_{k=1}^n \ln \left(\frac{g_{\xi,R}(t_{\xi}^{-1}(\hat{U}_{k,1}), \dots, t_{\xi}^{-1}(\hat{U}_{k,d}))}{g_{\xi}(t_{\xi}^{-1}(\hat{U}_{k,1})) \cdots g_{\xi}(t_{\xi}^{-1}(\hat{U}_{k,d}))} \right) \\ &= \sum_{k=1}^n \ln \left(g_{\xi,R}(t_{\xi}^{-1}(\hat{U}_{k,1}), \dots, t_{\xi}^{-1}(\hat{U}_{k,d})) \right) - \sum_{k=1}^n \ln \left(\prod_{j=1}^d g_{\xi}(t_{\xi}^{-1}(\hat{U}_{k,j})) \right) \\ &= \sum_{k=1}^n \ln \left(g_{\xi,R}(t_{\xi}^{-1}(\hat{U}_{k,1}), \dots, t_{\xi}^{-1}(\hat{U}_{k,d})) \right) - \sum_{k=1}^n \sum_{j=1}^d \ln \left(g_{\xi}(t_{\xi}^{-1}(\hat{U}_{k,j})) \right), \end{aligned}$$

wobei $g_{\xi,R}$ die Gesamtdichte einer standard d -dimensionalen *t*-Verteilung mit Verteilungsfunktion $t_{\xi,R}^d$, und g_{ξ} die Dichte einer univariaten standard *t*-Verteilung mit ξ Freiheitsgraden ist.

6.10 Die wichtigsten Familien von Copulas im Überblick

Gumbel: $C_{\theta}^{Gu}(u_1, u_2) = \exp\{-[(-\ln u_1)^{\theta} + (-\ln u_2)^{\theta}]^{1/\theta}\}$, $\theta \geq 1$.

$\lambda_U(X_1, X_2) = 2 - 2^{1/\theta}$, $\lambda_L(X_1, X_2) = 0$ und $\rho_{\tau}(\theta) = 1 - \frac{1}{\theta}$.

Zählt zu den archimedischen und vertauschbaren Copulas.

Clayton: $C_{\theta}^{Cl}(u_1, u_2) = (u_1^{-\theta} + u_2^{-\theta} - 1)^{1/\theta}$, $\theta > 0$.

$$\lambda_U(X_1, X_2) = 0, \lambda_L(X_1, X_2) = 2^{-1/\theta} \text{ und } \rho_\tau(\theta) = \frac{\theta}{\theta+2}.$$

Zählt zu den archimedischen und vertauschbaren Copulas.

Gauss: $C_R^{Ga}(u_1, u_2) = \phi_R^2(\phi^{-1}(u_1), \phi^{-1}(u_2)).$

$$\lambda_U(X_1, X_2) = \lambda_L(X_1, X_2) = 0 \text{ und } \rho_\tau(X_1, X_2) = \frac{2}{\pi} \arcsin R_{12} \text{ mit } R_{12} = \frac{\Sigma_{12}}{\sqrt{\Sigma_{11}\Sigma_{22}}}.$$

Zählt zu den elliptischen und vertauschbaren Copulas.

t: $C_{\nu,R}^t(u_1, u_2) = t_{\nu,R}^2(t_\nu^{-1}(u_1), t_\nu^{-1}(u_2)).$

$$\lambda_U(X_1, X_2) = \lambda_L(X_1, X_2) = 2\bar{t}_{\nu+1} \left(\sqrt{\nu+1} \frac{\sqrt{1-R_{12}}}{\sqrt{1+R_{12}}} \right) \text{ und } \rho_\tau(X_1, X_2) = \frac{2}{\pi} \arcsin R_{12} \text{ mit } R_{12} = \frac{\Sigma_{12}}{\sqrt{\Sigma_{11}\Sigma_{22}}}.$$

Zählt zu den elliptischen Copulas. Ist R eine Äqui-Korrelationsmatrix, so ist die Copula auch vertauschbar.

Kapitel 7

Kreditrisiko

Credit risk is the risk that the value of a portfolio changes due to unexpected changes in the credit quality of issuers or trading partners. This subsumes both losses due to defaults and losses caused by changes in credit quality such as the downgrading of a counterparty in an internal or external rating system

Zitat von McNeil, Frey und Embrechts (2005)

Beispiele für Kreditrisiko-behaftete Finanzinstrumente sind z.B.:

- Portfolios von Unternehmensanleihen
- OTC („over the counter“) Transaktionen
- Handel im Bereich der Kreditderivate

7.1 Modelle

7.1.1 Ein einfaches Modell

Sei P ein Portfolio bestehend aus n risikoreichen Anleihen in der Höhe L_i , $i = 1, 2, \dots, n$ und p_i die Wahrscheinlichkeit, dass Kreditnehmer i zahlungsunfähig wird. Der Anteil des Verlustes aus Anleihe i , falls Kreditnehmer i zahlungsunfähig wird, ist $1 - \lambda_i$. $\lambda_i \in [0, 1]$ heißt „recovery rate“ von Anleihe i .

Der Verlust im Falle von Zahlungsunfähigkeit („loss-given-default“) ist:

$$LGD_i = (1 - \lambda_i)L_i.$$

Sei X_i , der Status des Kreditnehmers i zum Zeitpunkt T , eine Bernoulli-ZV, dh.:

$$X_i = \begin{cases} 1 & \text{Kreditnehmer } i \text{ ist zahlungsunfähig} \\ 0 & \text{Kreditnehmer } i \text{ ist nicht zahlungsunfähig} \end{cases}$$

und $p_i = P(X_i = 1)$. Dann gilt für den Gesamtverlust zum Zeitpunkt T

$$L = \sum_{i=1}^n X_i \cdot LGD_i = \sum_{i=1}^n X_i(1 - \lambda_i)L_i.$$

Die Verteilung von L hängt also von der Gesamtverteilung von $(X_1, \dots, X_n, \lambda_1, \dots, \lambda_n)^\top$ ab.

Das einfachste Modell wäre demnach:

- die L_i sind identisch, d.h. $L_i = L_1, \forall i$,
- die recovery rates sind deterministisch und $\lambda_i = \lambda_1, \forall i$
- die X_i sind i.i.d. mit Wahrscheinlichkeit p .

Dann gilt $L = LGD_1 \cdot N$ mit $N = \sum_{i=1}^n X_i \sim \text{Bin}(n, p)$.

7.1.2 Modelle mit latenten Variablen

Die Kreditnehmer werden in $m + 1$ homogene Kategorien geteilt und alle Kreditnehmer einer Gruppe haben dieselbe Wahrscheinlichkeit zahlungsunfähig zu werden (*default Wahrscheinlichkeit*).

Aus historischen Beobachtungen der Anzahl der Kreditnehmer einer Kategorie, die zahlungsunfähig werden folgt eine Schätzung der default Wahrscheinlichkeit für Kreditnehmer der entsprechenden Kategorie.

Sei $S = (S_1, S_2, \dots, S_n)$, $S_i \in \{0, 1, \dots, m\}$, eine Status Variable für die gilt, dass $S_i = 0$ der Zahlungsunfähigkeit und $S_i = j \in \{1, 2, \dots, m\}$ den unterschiedlichen Einteilungskategorien entspricht (z.B. Rating Klassen). Dann gilt

$$X_i = \begin{cases} 0 & S_i \neq 0 \\ 1 & S_i = 0 \end{cases}$$

Die Status Variable $S = (S_1, S_2, \dots, S_n)^\top$ wird mit Hilfe der latenten Variablen $Y = (Y_1, Y_2, \dots, Y_n)^\top$ modelliert (Y_i könnte z.B. der Wert der Aktien von Kreditnehmer i sein). Seien d_{ij} , $i = 1, 2, \dots, n$, $j = 0, 1, \dots, m + 1$, Schwellwerte, sodass $d_{i,0} = -\infty$ und $d_{i,m+1} = \infty$. Dann gilt:

$$S_i = j \iff Y_i \in (d_{i,j}, d_{i,j+1}].$$

Wenn F_i die Verteilungsfunktion von Y_i ist, dann gilt für die default Wahrscheinlichkeit

$$p_i = F_i(d_{i,1}).$$

Die Wahrscheinlichkeit, dass die ersten k Kreditnehmer zahlungsunfähig werden, ist:

$$\begin{aligned} p_{1,2,\dots,k} &= \mathbb{P}(Y_1 \leq d_{1,1}, Y_2 \leq d_{2,1}, \dots, Y_k \leq d_{k,1}) \\ &= C(F_1(d_{1,1}), F_2(d_{2,1}), \dots, F_k(d_{k,1}), 1, 1, \dots, 1) \\ &= C(p_1, p_2, \dots, p_k, 1, \dots, 1) \end{aligned}$$

D.h. die Gesamt-default-Wahrscheinlichkeit hängt wesentlich von der Copula C ab.

7.1.3 Das univariate KMV Modell (siehe auch www.moodyskmv.com)

Die Status Variablen $S = (S_1, S_2, \dots, S_n)$ können nur zwei Werte 0 und 1 annehmen, d.h. $m = 1$.

Die latenten Variablen $Y = (Y_1, Y_2, \dots, Y_n)^\top$ hängen mit dem Wert der Aktien der jeweiligen Firmen folgendermaßen zusammen:

Das Modell von Merton

Die Bilanz jeder Firma besteht aus 2 Positionen, nämlich Aktiva (Aktien) und Passiva (Liabilities and Equities).

Die Passiva bestehen aus Schulden („Liabilities“) und Stammkapital („Equity“).

$V_{A,i}(T)$: Wert der Aktien der Firma i zum Zeitpunkt T

$K_i(T) =: K_i$: Wert der Schulden der Firma i zum Zeitpunkt T

$V_{E,i}(T)$: Wert des Stammkapitals der Firma i zum Zeitpunkt T

Annahme: Zukünftiger Wert der Aktien wird als geometrische Brown'sche Bewegung modelliert:

$$V_{A,i}(T) = V_{A,i}(t) \exp \left\{ \left(\mu_{A,i} - \frac{\sigma_{A,i}^2}{2} \right) (T - t) + \sigma_{A,i} (W_i(T) - W_i(t)) \right\},$$

$\mu_{A,i}$ ist der Drift, $\sigma_{A,i}$ ist die Volatilität und $(W_i(t): 0 \leq t \leq T)$ ist eine standard Brown'sche Bewegung (Wiener Prozess). Es gilt demnach

$$(W_i(T) - W_i(t)) \sim N(0, T - t).$$

Daraus folgt

$$\begin{aligned} \ln V_{A,i}(T) &= \ln V_{A,i}(t) + \left(\mu_{A,i} - \frac{\sigma_{A,i}^2}{2} \right) (T - t) + \sigma_{A,i} \sqrt{T - t} \underbrace{\frac{W_i(T) - W_i(t)}{\sqrt{T - t}}}_{=: Y_i \sim N(0,1)} \\ &= \mu + \sigma Y_i \end{aligned}$$

und somit $\ln V_{A,i}(T) \sim N(\mu, \sigma^2)$ mit

$$\begin{aligned} \mu &= \ln V_{A,i}(t) + \left(\mu_{A,i} - \frac{\sigma_{A,i}^2}{2} \right) (T - t) \\ \sigma^2 &= \sigma_{A,i}^2 (T - t). \end{aligned}$$

Weiters ist $X_i = I_{(-\infty, K_i)}(V_{A,i}(T))$ und mit $Y_i = \frac{W_i(T) - W_i(t)}{\sqrt{T - t}} \sim N(0, 1)$ gilt schließlich $X_i = I_{(-\infty, K_i)}(V_{A,i}(T)) = I_{(-\infty, -DD_i)}(Y_i)$, wobei

$$DD_i = \frac{\ln V_{A,i}(t) - \ln K_i + \left(\mu_{A,i} - \frac{\sigma_{A,i}^2}{2} \right) (T - t)}{\sigma_{A,i} \sqrt{T - t}}, \quad (7.1)$$

denn

$$X_i = 1 \Leftrightarrow V_{A,i}(T) < K_i \Leftrightarrow \ln V_{A,i}(T) < \ln K_i \Leftrightarrow \mu + \sigma_{A,i} \sqrt{T - t} Y_i < \ln K_i \Leftrightarrow Y_i < \frac{\ln K_i - \mu}{\sigma_{A,i} \sqrt{T - t}}.$$

DD_i heißt *distance to default*.

Die distance to default zu berechnen birgt einige Schwierigkeiten, denn weder $V_{A,i}(t)$ noch

$\sigma_{A,i}(t)$ können beobachtet werden ($\mu_{A,i}(t)$ „kann“ man aus historischen Daten schätzen). Allerdings kann $V_{E,i}(t)$ beobachtet werden! Im KMV Modell besitzen die Geldgeber die Firma solange die Schulden seitens der Stammkapitalbesitzer (equity holders) nicht vollständig bezahlt werden. $V_{E,i}(T)$ ist daher der Preis einer Call Option über die Aktien der Firma mit Strike Price dem Buchwert der Schulden zum Zeitpunkt T :

$$V_{E,i}(T) = \max\{V_{A,i}(T) - K_i, 0\}.$$

Aus der Black-Scholes Formel (Optionspreistheorie) folgt:

$$V_{E,i}(t) = C(V_{A,i}(t), r, \sigma_{A,i}) = V_{A,i}(t)\phi(e_1) - K_i e^{-r(T-t)}\phi(e_2) \quad (7.2)$$

wobei

$$e_1 = \frac{\ln(V_{A,i}(t)) - \ln K_i + (r + \sigma_{A,i}^2/2)(T-t)}{\sigma_{A,i}(T-t)} \quad \text{und} \quad e_2 = e_1 - \sigma_{A,i}(T-t).$$

ϕ ist die Verteilungsfunktion der standard Normalverteilung und r ist der risikolose Zinssatz.

Im KMV Modell gilt weiters:

$$\sigma_{E,i} = g(V_{A,i}(t), \sigma_{A,i}, r). \quad (7.3)$$

Beobachtungen/Schätzungen von $V_{E,i}(t)$ bzw. $\sigma_{E,i}$ aus historischen Daten werden in (7.2) und (7.3) eingesetzt und das Gleichungssystem

$$\begin{aligned} V_{E,i}(t) &= C(V_{A,i}(t), r, \sigma_{A,i}) \\ \sigma_{E,i} &= g(V_{A,i}(t), \sigma_{A,i}, r) \end{aligned} \quad (7.4)$$

wird gelöst um $V_{A,i}(t)$ und $\sigma_{A,i}$ zu ermitteln. Diese Werte werden schließlich zur Berechnung von DD_i aus (7.1) verwendet.

Die erwartete Häufigkeit der Zahlungsunfähigkeit (expected default frequency, EDF)

Das KMV Modell evaluiert nicht direkt die Default Wahrscheinlichkeit $p_i = P(Y_i < -DD_i)$. Stattdessen werden jene Firmen ermittelt, die historisch gesehen je einen „distance to default“ von ca. DD_i hatten. Für diese Firmen wird die Häufigkeit der Zahlungsunfähigkeit als Schätzer für die Default Wahrscheinlichkeit p_i ermittelt. Dieser Schätzer wird *expected default frequency (EDF)* genannt.

Zusammenfassung des univariaten KMV Modells zur Berechnung der Default Wahrscheinlichkeit für eine Firma

- Ermittlung des Aktienwertes $V_{A,i}$ und dessen Volatilität $\sigma_{A,i}$ mit Hilfe der Beobachtungen über Marktwert und Volatilität der Equities ($V_{E,i}$ bzw. $\sigma_{E,i}$) sowie der Schulden K_i als Lösung des Gleichungssystems (7.4).
- Berechnung der „distance to default“ DD_i aus (7.1).
- Berechnung der Default Wahrscheinlichkeiten p_i mit Hilfe einer empirischen Verteilung, die den Zusammenhang zwischen Default Wahrscheinlichkeit und „distance to default“ modelliert (zB. mit Hilfe von EDF).

7.1.4 Das multivariate KMV Modell: Berechnung von multivariaten Default Wahrscheinlichkeiten

Seien $W_j(t)$, $0 \leq t \leq T$, unabhängige standard Brown'sche Bewegungen, $j = 1, 2, \dots, m$. Das grundlegende Modell ist:

$$V_{A,i}(T) = V_{A,i}(t) \exp \left\{ \left(\mu_{A,i} - \frac{\sigma_{A,i}^2}{2} \right) (T-t) + \sum_{j=1}^m \sigma_{A,i,j} (W_j(T) - W_j(t)) \right\},$$

$\mu_{A,i}$ ist der Drift und $\sigma_{A,i}^2 = \sum_{j=1}^m \sigma_{A,i,j}^2$ ist die Volatilität, wobei $\sigma_{A,i,j}$ den Einfluss der Brown'schen Bewegung j auf die Entwicklung des Aktienwertes der Firma i quantifiziert.

Sei $Y_i = \frac{\sum_{j=1}^m \sigma_{A,i,j} (W_j(T) - W_j(t))}{\sigma_{A,i} \sqrt{T-t}}$, dann ist $Y = (Y_1, Y_2, \dots, Y_n) \sim N_n(0, \Sigma)$, wobei $\Sigma_{ij} = \frac{\sum_{k=1}^m \sigma_{A,i,k} \sigma_{A,j,k}}{\sigma_{A,i} \sigma_{A,j}}$.

Es gilt dann $V_{A,i}(T) < K_i \iff Y_i < -DD_i$, wobei

$$DD_i = \frac{\ln V_{A,i}(t) - \ln K_i + \left(\frac{\mu_{A,i} - \sigma_{A,i}^2}{2} \right) (T-t)}{\sigma_{A,i} \sqrt{T-t}}.$$

Die Wahrscheinlichkeit, dass die ersten k Firmen zahlungsunfähig werden, ist

$$\begin{aligned} \mathbb{P}(X_1 = 1, X_2 = 1, \dots, X_k = 1) &= \mathbb{P}(Y_1 < -DD_1, \dots, Y_k < -DD_k) \\ &= C_{\Sigma}^{Ga}(\phi(-DD_1), \dots, \phi(-DD_k), 1, \dots, 1). \end{aligned}$$

C_{Σ}^{Ga} ist die Copula einer multivariaten Normalverteilung mit Kovarianzmatrix Σ . Die Häufigkeit der multivariaten Zahlungsunfähigkeit (joint default frequency) ist

$$JDF_{1,2,\dots,k} = C_{\Sigma}^{Ga}(EDF_1, EDF_2, \dots, EDF_k, 1, \dots, 1),$$

wobei EDF_i die Häufigkeit der Zahlungsunfähigkeit für die Firma i , $i = 1, 2, \dots, k$, ist.

Schätzung der Kovarianzen/Korrelationen $\sigma_{A,i,j}$

Schwierigkeiten bei der Schätzung:

- n ist typischerweise sehr groß.
- Es sind wenige historische Daten vorhanden.
- Wenn n groß ist, dann bilden die paarweise geschätzten Korrelationskoeffizienten i.A. keine positiv definite Korrelationsmatrix.

Eine mögliche Lösung ist ein Faktormodell für die latenten Variablen, in dem der Aktienwert durch eine Reihe von gemeinsamen Faktoren (makro-ökonomische globale, regionale, Sektor-, Länder- und Branchen-spezifische Faktoren) und einem firmenspezifischen Faktor bestimmt wird:

$$Y = (Y_1, Y_2, \dots, Y_n)^{\top} = AZ + BU,$$

wobei $Z = (Z_1, \dots, Z_k)^\top \sim N_k(0, \Lambda)$ die k gemeinsamen und $U = (U_1, \dots, U_n)^\top \sim N_d(0, I)$ die firmenspezifischen Faktoren sind. Z und U sind unabhängig und die Konstanten Matrizen $A = (a_{ij}) \in \mathbb{R}^{n \times k}$, $B = \text{diag}(b_1, \dots, b_n) \in \mathbb{R}^{n \times n}$ sind die Modellparameter.

Es gilt dann $\text{Cov}(Y) = A\Lambda A^\top + D$, wobei $D = \text{diag}(b_1^2, \dots, b_n^2) \in \mathbb{R}^{n \times n}$.

7.1.5 Credit Metrics

Dieses Modell wurde bei J.P.Morgan entwickelt und wird in erster Linie für die Evaluierung von Bond Portfolios verwendet (siehe Crouhy et al. (2000), J.P.Morgan Inc. (1997)). Es basiert auf einem Bonitäts-Einstufungssystem (zB. von *Moody* oder *Standard and Poor's*) und berücksichtigt die Veränderungen im PF-Wert aufgrund von Veränderungen in den Bonitäts-Einstufungen.

Sei P ein Portfolio von n Krediten mit einer fixen Laufzeit (zB. 1 Jahr). Sei S_i der Zustand-indikator von Kreditnehmer i . Die möglichen Zustände werden mit $0, 1, \dots, m$ bezeichnet, wobei $S_i = 0$ der Zahlungsunfähigkeit entspricht.

Beispiel 7.1.1. Einstufungssystem von Standard and Poor's

$m = 7$; $S_i = 0$ heißt Zahlungsunfähigkeit; $S_i = 1$ oder *CCC*; $S_i = 2$ oder *B*; $S_i = 3$ oder *BB*; $S_i = 4$ oder *BBB*; $S_i = 5$ oder *A*; $S_i = 6$ oder *AA*; $S_i = 7$ oder *AAA*. \diamond

Für jeden Kreditnehmer wird die Dynamik der Bonitäts-Einstufungen mit Hilfe einer Markovkette mit Zustandsmenge $\{0, 1, \dots, m\}$ und Übergangsmatrix P modelliert. Die Übergangswahrscheinlichkeiten werden mit Hilfe von historischen Daten geschätzt, zB.:

Ursprüngliche Einstufung	Einstufung am Ende des Jahres							Zahlungs- unfähigkeit
	AAA	AA	A	BBB	BB	B	CCC	
AAA	90.81	8.33	0.68	0.06	0.12	0	0	0
AA	0.70	90.65	7.79	0.64	0.06	0.14	0.02	0
A	0.09	2.27	91.05	5.52	0.74	0.26	0.01	0.06
BBB	0.02	0.33	5.95	86.93	5.30	1.17	0.12	0.18
BB	0.03	0.14	0.67	7.73	80.53	8.84	1.00	1.06
B	0	0.11	0.24	0.43	6.48	83.46	4.07	5.20
CCC	0.22	0	0.22	1.30	2.38	11.24	64.86	19.79

Recovery Rates

Im Fall einer Zahlungsunfähigkeit hängt die recovery rate von der Einstufung des Kreditnehmers ab. Der Durchschnittswert und die Standardabweichung der recovery rate werden aufgrund von historischen Daten innerhalb jeder Einstufungsklasse geschätzt.

Evaluierung der Bonds im Falle einer Neueinstufung

Beispiel 7.1.2. Betrachten wir ein BBB Bond mit Laufzeit 5 Jahre. Er zahlt jedes Jahr einen Kupon von 6% aus. Die forward Zinsstrukturkurven (*forward yield curves*) für jede Einstufungsklasse sind wie folgt gegeben (in %):

Einstufung	1. Jahr	2. Jahr	3. Jahr	4. Jahr
AAA	3.60	4.17	4.73	5.12
AA	3.65	4.22	4.78	5.17
A	3.73	4.32	4.93	5.32
BBB	4.10	4.67	5.25	5.63
BB	6.05	7.02	8.03	8.52
CCC	15.05	15.02	14.03	13.52

Für einen Nennwert von 100 zahlt der Bond 6 Währungseinheiten am Ende des 1., 2., 3. und 4. Jahres aus. Am Ende des 5. Jahres zahlt der Bond 106 Währungseinheiten.

Es wird nun die Annahme getroffen, dass am Ende des ersten Jahres der Bond als *A* Bond neu eingestuft wird. Der Wert des Bonds am Ende des ersten Jahres ist somit

$$V = 6 + \frac{6}{1 + 3,73\%} + \frac{6}{(1 + 4,32\%)^2} + \frac{6}{(1 + 4,93\%)^3} + \frac{106}{(1 + 5,32\%)^4} = 108.64$$

Analog wird der Wert des Bonds am Ende des 1. Jahres ermittelt, falls er zu diesem Zeitpunkt zu anderen Klassen eingestuft wird. Es wird eine recovery rate von 51.13% im Falle von Zahlungsunfähigkeit angenommen.

Einstufung am Ende des 1. Jahres	Wert
AAA	109.35
AA	109.17
A	108.64
BBB	107.53
BB	102.01
B	98.09
CCC	83.63
Zahlungsunfähigkeit	51.13

◇

Wert und Risiko eines Bond-Portfolios in Credit Metrics

Die Abhängigkeit der Neueinstufungen unterschiedlicher Bonds und die Wahrscheinlichkeiten von Neueinstufungen von Gruppen von Bonds werden mit Hilfe der dazugehörigen Rendite berechnet. Die Rendite von Bond i wird als Normalverteilung Y_i modelliert.

Seien $d_{Def}, d_{CCC}, \dots, d_{AAA} = +\infty$ Schwellwerte, sodass für einen Kreditnehmer die Wahrscheinlichkeit des Übergangs in eine neue Stufe S_i am Ende einer vordefinierten Periode folgendermaßen gegeben sind:

$$\begin{aligned} \mathbb{P}(S_i = 0) &= \phi(d_{Def}) \\ \mathbb{P}(S_i = CCC) &= \mathbb{P}(d_{Def} < Y_i \leq d_{CCC}) = \phi(d_{CCC}) - \phi(d_{Def}) \\ &\vdots \\ \mathbb{P}(S_i = AAA) &= \mathbb{P}(Y_i > d_{AA}) = 1 - \phi(d_{AA}). \end{aligned}$$

Die Renditen mehrerer Bonds werden mit Hilfe der multivariaten Normalverteilung modelliert. Die Korrelationsmatrix dieser Verteilung wird in **Credit Metrics** mit Hilfe von Faktormodellen berechnet. Es können dann Gesamtwahrscheinlichkeiten wie

$$\mathbb{P}(S_1 = 0, \dots, S_n = 3) = \mathbb{P}(Y_1 \leq d_{Def}, \dots, d_B < Y_n \leq d_{BB})$$

berechnet werden. Als Modell für die Abhängigkeitsstruktur des Vektors (Y_1, Y_2, \dots, Y_n) wird die Gauss'sche Copula(!) verwendet.

Die Risikomaße eines Kreditportfolios werden mit Hilfe von Simulationen berechnet. Es werden viele Szenarien generiert, aufgrund derer der empirische VaR oder CVaR ermittelt wird.

Die Bernoulli Mixture Verteilung

Der 0-1 Zufallsvektor $X = (X_1, \dots, X_n)^\top$ hat eine Bernoulli Mixture Verteilung (BMV), wenn es einen Zufallsvektor $Z = (Z_1, Z_2, \dots, Z_m)^\top$, $m < n$, und Funktionen $f_i: \mathbb{R}^m \rightarrow [0, 1]$, $i = 1, 2, \dots, n$, gibt, sodass X bedingt durch Z ein Vektor von unabhängigen Bernoulli verteilten Zufallsvariablen ist und

$$\mathbb{P}(X_i = 1|Z) = f_i(Z), \quad \mathbb{P}(X_i = 0|Z) = 1 - f_i(Z).$$

Für $x = (x_1, \dots, x_n)^\top \in \{0, 1\}^n$ gilt

$$\mathbb{P}(X = x|Z) = \prod_{i=1}^n f_i(Z)^{x_i} (1 - f_i(Z))^{1-x_i}.$$

Die unbedingte Verteilung ist

$$\mathbb{P}(X = x) = \mathbb{E}(\mathbb{P}(X = x|Z)) = \mathbb{E} \left(\prod_{i=1}^n f_i(Z)^{x_i} (1 - f_i(Z))^{1-x_i} \right).$$

Annahme: alle Funktionen f_i sind identisch, $f_i = f$. Für die Anzahl der Zahlungsunfähigkeitsfällen $N = \sum_{i=1}^n X_i$ gilt dann $N|Z \sim \text{Bin}(n, f(Z))$.

Die Poisson Mixture Verteilung

Der diskrete Zufallsvektor $X = (X_1, \dots, X_n)^\top$ hat eine Poisson Mixture Verteilung (PMV), wenn es einen Zufallsvektor $Z = (Z_1, Z_2, \dots, Z_m)^\top$, $m < n$, und Funktionen $\lambda_i: \mathbb{R}^m \rightarrow (0, \infty)$, $i = 1, 2, \dots, n$, gibt, sodass X bedingt durch Z ein Vektor von unabhängigen Poisson verteilten Zufallsvariablen ist und

$$\mathbb{P}(X_i = x_i|Z) = \frac{\lambda_i(Z)^{x_i}}{x_i!} e^{-\lambda_i(Z)} \quad \text{für } x_i \in \mathbb{N} \cup \{0\}.$$

Für $x = (x_1, \dots, x_n)^\top \in (\mathbb{N} \cup \{0\})^n$ gilt

$$\mathbb{P}(X = x|Z) = \prod_{i=1}^n \frac{\lambda_i(Z)^{x_i}}{x_i!} e^{-\lambda_i(Z)}.$$

Die unbedingte Verteilung ist

$$\mathbb{P}(X = x) = \mathbb{E}(\mathbb{P}(X = x|Z)) = \mathbb{E} \left(\prod_{i=1}^n \frac{\lambda_i(Z)^{x_i}}{x_i!} e^{-\lambda_i(Z)} \right).$$

Annahme: $\tilde{X} = (\tilde{X}_1, \dots, \tilde{X}_n)^\top$ ist PMV mit Faktoren Z , wobei $\tilde{X}_i \in \{0, 1, \dots, m\}$.

Sei $X_i = I_{[1, \infty)}(\tilde{X}_i)$. Dann ist $X = (X_1, \dots, X_n)$ BMV mit $f_i(Z) = 1 - e^{-\lambda_i(Z)}$.

Für $\lambda_i(Z)$ klein gilt $\tilde{N} = \sum_{i=1}^n \tilde{X}_i \approx \sum_{i=1}^n X_i$. $\tilde{N}|Z \sim \text{Poi}(\bar{\lambda}(Z))$, wobei $\bar{\lambda}(Z) = \sum_{i=1}^n \lambda_i(Z)$.

Beispiele von Bernoulli Mixture Modellen

Annahmen:

- Z ist univariat (d.h. es gibt einen Risikofaktor)
- $f_i = f$ für alle i

Es gilt: $\mathbb{P}(X_i = 1|Z) = f(Z), \forall i; N|Z = \sum_{i=1}^n X_i \sim \text{Bin}(n, f(Z))$.

Die unbedingte Wahrscheinlichkeit, dass die ersten k Kreditnehmer zahlungsunfähig werden, ist

$$\begin{aligned} & \mathbb{P}(X_1 = 1, \dots, X_k = 1, X_{k+1} = 0, \dots, X_n = 0) = \\ & \mathbb{E}(\mathbb{P}(X_1 = 1, \dots, X_k = 1, X_{k+1} = 0, \dots, X_n = 0|Z)) = \mathbb{E}(f(Z)^k(1 - f(Z))^{n-k}). \end{aligned}$$

Sei G die Verteilungsfunktion von Z . Dann gilt:

$$\mathbb{P}(X_1 = 1, \dots, X_k = 1, X_{k+1} = 0, \dots, X_n = 0) = \int_{-\infty}^{\infty} f(z)^k(1 - f(z))^{n-k} d(G(z)).$$

Die Verteilung der Anzahl N der Zahlungsunfähigen Kreditnehmer ist

$$\mathbb{P}(N = k) = \binom{n}{k} \int_{-\infty}^{\infty} f(z)^k(1 - f(z))^{n-k} d(G(z)).$$

- Die Beta mixing Verteilung

Es gilt $Z \sim \text{Beta}(a, b)$ und $f(z) = z$. Die Dichte g von Z ist also

$$g(z) = \frac{1}{\beta(a, b)} z^{a-1}(1 - z)^{b-1},$$

für $a, b > 0, z \in (0, 1)$, wobei $\beta(a, b) = \int_0^1 z^{a-1}(1 - z)^{b-1} dz$ die Euler'sche Betafunktion ist.

Die Verteilung der Anzahl der zahlungsunfähigen Kreditnehmer ist

$$\begin{aligned} \mathbb{P}(N = k) &= \binom{n}{k} \int_0^1 z^k(1 - z)^{n-k} g(z) dz \\ &= \binom{n}{k} \frac{1}{\beta(a, b)} \int_0^1 z^{a+k-1}(1 - z)^{n-k+b-1} dz \\ &= \binom{n}{k} \frac{\beta(a + k, b + n - k)}{\beta(a, b)} \end{aligned}$$

und diese entspricht der Beta-Binomial Verteilung.

- Probit-normal Mixture

$Z \sim N(0, 1)$, $f(z) = \Phi(\mu + \sigma z)$, $\mu \in \mathbb{R}, \sigma > 0$ und Φ ist die standard Normalverteilungsfunktion.

- Logit-normal Mixture

$Z \sim N(0, 1)$, $f(z) = (1 + \exp\{\mu + \sigma z\})^{-1}$, $\mu \in \mathbb{R}, \sigma > 0$.

7.1.6 CreditRisk⁺ - Ein Poisson Mixture Modell

CreditRisk⁺ wurde von CSFB (1997) entwickelt, siehe Crouhy et al. (2000), und http://www.credit_suisse.com/investment_banking/research/en/credit_risk.jsp.

Seien m unabhängige Risikofaktoren Z_1, Z_2, \dots, Z_m , $Z_j \sim \Gamma(\alpha_j, \beta_j)$, $j = 1, 2, \dots, m$, gegeben sodass $\mathbb{E}(Z_j) = 1$ und sei weiters

$$\lambda_i(Z) = \bar{\lambda}_i \sum_{j=1}^m a_{ij} Z_j, \quad \sum_{j=1}^m a_{ij} = 1 \quad \text{für } i = 1, 2, \dots, n.$$

$\bar{\lambda}_i > 0$, α_j, β_j sind Konstanten, wobei α_j und β_j meistens so gewählt werden, dass $\mathbb{E}(\lambda_i(Z)) = \bar{\lambda}_i > 0$ gilt. Die Dichte von Z_j ist

$$f_j(z) = \frac{z^{\alpha_j-1} \exp\{-z/\beta_j\}}{\beta_j^{\alpha_j} \Gamma(\alpha_j)}.$$

Der Verlust bei Kredit i durch die Zahlungsunfähigkeit des Kreditnehmers i ist

$$LGD_i = (1 - \lambda_i)L_i, \quad 1 \leq i \leq n,$$

wobei λ_i die erwartete deterministische „Recovery rate“ ist und L_i die Höhe von Kredit i .

Ziel ist es nun, die Gesamtverlustverteilung durch eine diskrete Verteilung zu approximieren und für diese die Erzeugendenfunktion zu ermitteln.

Sei Y eine diskrete ZV mit Wertebereich $\{y_1, \dots, y_m\}$ oder eine kontinuierliche ZV mit Dichtefunktion $f(y)$ in \mathbb{R} .

Definition 7.1.3. Die Erzeugendenfunktion von Y ist definiert als

$$g_Y(t) := \mathbb{E}(t^Y) = \sum_{i=1}^m t^{y_i} \mathbb{P}(Y = y_i)$$

bzw.

$$g_Y(t) := \int_{-\infty}^{\infty} t^y f(y) dy$$

für $t \in [0, 1]$.

Einige Eigenschaften von Erzeugendenfunktionen:

- (i) Wenn $Y \sim \text{Bernoulli}(p)$, dann $g_Y(t) = 1 + p(t - 1)$.
- (ii) Wenn $Y \sim \text{Poisson}(\lambda)$, dann $g_Y(t) = \exp\{\lambda(t - 1)\}$.
- (iii) Für unabhängige Zufallsvariablen X_1, \dots, X_n gilt

$$g_{X_1+\dots+X_n}(t) = \prod_{i=1}^n g_{X_i}(t).$$

(iv) Sei Y eine Zufallsvariable mit Dichtefunktion f und sei $g_{X|Y=y}(t)$ die Erzeugendenfunktion von $X|Y = y$. Dann gilt

$$g_X(t) = \int_{-\infty}^{\infty} g_{X|Y=y}(t) f(y) dy.$$

(v) Sei $g_X(t)$ die Erzeugendenfunktion von X . Dann gilt

$$\mathbb{P}(X = k) = \frac{1}{k!} g^{(k)}(0), \text{ wobei } g^{(k)}(t) = \frac{d^k g(t)}{dt^k}.$$

Die Erzeugendenfunktion der Verlustverteilung

Jeder Verlust wird als ganzzahliges Vielfaches einer vordefinierten Verlusteinheit L_0 (zB. $L_0 = 10^6$ Euro) modelliert:

$$LGD_i = (1 - \lambda_i)L_i \approx \left[\frac{(1 - \lambda_i)L_i}{L_0} \right] L_0 = v_i L_0 \quad \text{mit} \quad v_i := \left[\frac{(1 - \lambda_i)L_i}{L_0} \right],$$

wobei $[x] = \arg \min_t \{t - x : t \in \mathbb{Z}, t - x \in (-1/2, 1/2]\}$.

Die Verlustfunktion ist somit $L = \sum_{i=1}^n X_i v_i L_0$.

(a) Ermittlung der Erzeugendenfunktion für $N = X_1 + \dots + X_n$

$X_i|Z \sim \text{Poisson}(\lambda_i(Z)) \forall i \implies g_{X_i|Z}(t) = \exp\{\lambda_i(Z)(t - 1)\} \forall i \implies$

$$g_{N|Z}(t) = \prod_{i=1}^n g_{X_i|Z}(t) = \prod_{i=1}^n \exp\{\lambda_i(Z)(t - 1)\} = \exp\{\mu(t - 1)\} \quad (7.5)$$

mit $\mu := \sum_{i=1}^n \lambda_i(Z) = \sum_{i=1}^n \left(\bar{\lambda}_i \sum_{j=1}^m a_{ij} Z_j \right)$.

$$\begin{aligned} g_N(t) &= \int_0^\infty \dots \int_0^\infty g_{N|Z=(z_1, z_2, \dots, z_m)} f_1(z_1) \dots f_m(z_m) dz_1 \dots dz_m \\ &= \int_0^\infty \dots \int_0^\infty \exp \left\{ \sum_{i=1}^n \left(\bar{\lambda}_i \sum_{j=1}^m a_{ij} z_j \right) (t - 1) \right\} f_1(z_1) \dots f_m(z_m) dz_1 \dots dz_m \\ &= \int_0^\infty \dots \int_0^\infty \exp \left\{ (t - 1) \sum_{j=1}^m \underbrace{\left(\sum_{i=1}^n \bar{\lambda}_i a_{ij} \right)}_{\mu_j} z_j \right\} f_1(z_1) \dots f_m(z_m) dz_1 \dots dz_m \\ &= \int_0^\infty \dots \int_0^\infty \exp\{(t - 1)\mu_1 z_1\} f_1(z_1) dz_1 \dots \exp\{(t - 1)\mu_m z_m\} f_m(z_m) dz_m \\ &= \prod_{j=1}^m \int_0^\infty \exp\{z_j \mu_j (t - 1)\} \frac{1}{\beta_j^{\alpha_j} \Gamma(\alpha_j)} z_j^{\alpha_j - 1} \exp\{-z_j / \beta_j\} dz_j \end{aligned} \quad (7.6)$$

Die Berechnung der einzelnen Integrale in (7.6) ergibt:

$$\int_0^\infty \frac{1}{\Gamma(\alpha_j) \beta_j^{\alpha_j}} \exp\{z_j \mu_j (t - 1)\} z_j^{\alpha_j - 1} \exp\{-z_j / \beta_j\} dz_j$$

$$\begin{aligned}
&= \frac{1}{\beta_j^{\alpha_j} \Gamma(\alpha_j)} \int_0^\infty z_j^{\alpha_j-1} \exp\left\{-z_j \left(\beta_j^{-1} - \mu_j(t-1)\right)\right\} dz_j \\
&= \frac{1}{\beta_j^{\alpha_j} \Gamma(\alpha_j)} \int_0^\infty \exp\{-u_j\} u_j^{\alpha_j-1} \frac{1}{\left(\beta_j^{-1} - \mu_j(t-1)\right)^{\alpha_j}} du_j \quad \text{mit } u_j = z_j \left(\beta_j^{-1} - \mu_j(t-1)\right) \\
&= \frac{1}{\beta_j^{\alpha_j} \Gamma(\alpha_j)} \frac{1}{\left(\beta_j^{-1} - \mu_j(t-1)\right)^{\alpha_j}} \Gamma(\alpha_j) = \left(\frac{1}{1 - \beta_j \mu_j t + \beta_j \mu_j}\right)^{\alpha_j} \\
&= \left(\frac{1 - \delta_j}{1 - \delta_j t}\right)^{\alpha_j},
\end{aligned}$$

wobei

$$\delta_j = \beta_j \mu_j / (1 + \beta_j \mu_j). \quad (7.7)$$

Es gilt also $g_N(t) = \prod_{j=1}^m \left(\frac{1 - \delta_j}{1 - \delta_j t}\right)^{\alpha_j}$.

(b) Ermittlung der Erzeugendenfunktion für $L = \sum_{i=1}^n X_i v_i L_0$.

Der bedingte Verlust aufgrund von Zahlungsunfähigkeit von Kreditnehmer i ist $L_i|Z = v_i(X_i|Z)$; $L_i|Z$ ist unabhängig für $i = 1, 2, \dots, n$.

$$g_{L_i|Z}(t) = \mathbb{E}(t^{L_i}|Z) = \mathbb{E}(t^{v_i X_i}|Z) = g_{X_i|Z}(t^{v_i}).$$

Die Erzeugendenfunktion des gesamten Verlusts bedingt durch Z ist also

$$\begin{aligned}
g_{L|Z}(t) &= g_{L_1+L_2+\dots+L_n|Z}(t) = \prod_{i=1}^n g_{L_i|Z}(t) = \prod_{i=1}^n g_{X_i|Z}(t^{v_i}) = \\
&\exp\left\{\sum_{j=1}^m Z_j \left(\sum_{i=1}^n \bar{\lambda}_i a_{ij} (t^{v_i} - 1)\right)\right\}.
\end{aligned}$$

Ähnlich wie bei der Berechnung von $g_N(t)$ erhalten wir:

$$g_L(t) = \prod_{j=1}^m \left(\frac{1 - \delta_j}{1 - \delta_j \Lambda_j(t)}\right)^{\alpha_j}, \quad \text{wobei } \Lambda_j(t) = \frac{1}{\mu_j} \sum_{i=1}^n \bar{\lambda}_i a_{ij} t^{v_i}.$$

δ_j und μ_j sind wie in (7.7) bzw. (7.5) gegeben.

Beispiel 7.1.4. Kreditportfolio mit $n = 100$ Krediten, Anzahl der Risikofaktoren $m = 1$ oder $m = 5$, $\bar{\lambda}_i = \bar{\lambda} = 0.15$, für $i = 1, 2, \dots, n$, $\alpha_j = \alpha = 1$, $\beta_j = \beta = 1$, $a_{i,j} = 1/m$, $i = 1, 2, \dots, n$, $j = 1, 2, \dots, m$,

$$\mathbb{P}(N = k) = \frac{1}{k!} g_N^{(k)}(0) = \frac{1}{k!} \frac{d^k g_N}{dt^k}.$$

Für die Berechnung von $\mathbb{P}(N = k)$, $k = 0, 1, \dots, 100$, kann folgende rekursive Formel verwendet werden:

$$g_N^{(k)}(0) = \sum_{l=0}^{k-1} \binom{k-1}{l} g_N^{(k-1-l)}(0) \sum_{j=1}^m l! \alpha_j \delta_j^{l+1}, \quad k > 1.$$

$\underline{m} = \underline{1}$: $g_N(t) = \left(\frac{1-\delta_1}{1-\delta_1 t}\right)^{\alpha_1}$, wobei $\alpha_1 = \beta_1 = 1$, $\delta_1 = \frac{\mu_1}{1+\mu_1} = \frac{15}{16}$ und $\mu_1 = \sum_{i=1}^{100} \bar{\lambda}_i a_{ij} = 15$. Es gilt also

$$\begin{aligned} g_N(t) &= \frac{1}{16-15t} \\ g_N(0) &= \mathbb{P}(N=0) = \frac{1}{16} = 0.0625 \\ g'_N(0) &= \mathbb{P}(N=1) = \frac{15}{16^2} = 0.0586 \\ g''_N(0) &= \mathbb{P}(N=2) = \frac{15^2}{16^2} = 0.0549 \\ &\vdots \end{aligned}$$

$\underline{m} = \underline{5}$: $g_N(t) = \prod_{j=1}^5 \left(\frac{1-\delta_j}{1-\delta_j t}\right)^{\alpha_j}$, wobei $\alpha_j = \beta_j = 1$, $\delta_j = \frac{\mu_j}{1+\mu_j} = \frac{3}{4} \forall j$ und $\mu_j = \sum_{i=1}^{100} \bar{\lambda}_i a_{ij} = \frac{1}{5} \sum_{i=1}^{100} 0.015 = 3$.

$$\begin{aligned} g_N(t) &= \left(\frac{1-3/4}{1-3/4t}\right)^5 \\ g_N(0) &= \mathbb{P}(N=0) = 9.7 \cdot 10^{-4} \\ g'_N(0) &= \mathbb{P}(N=1) = 0.0037 \\ g''_N(0) &= \mathbb{P}(N=2) = 0.165 \\ &\vdots \end{aligned}$$

In Abb. 7.1 sind die Wahrscheinlichkeiten $\mathbb{P}(N=k)$ jeweils für $m=1$ und $m=5$ geplottet. Jene für $m=1$ fallen kontinuierlich, wohingegen jene für $m=5$ zuerst ansteigen bevor sie fallen. \diamond

7.2 Monte Carlo Methoden im Kreditrisiko-Management

Gegeben ist ein Kreditportfolio P bestehend aus n Krediten mit der Verlustfunktion $L = \sum_{i=1}^n L_i$, wobei die Verluste L_i unabhängig bedingt einem Vektor ψ von ökonomischen Einflussfaktoren sind.

Gesucht sind

$$\begin{aligned} VaR_\alpha(L) &= q_\alpha(L), \quad CVaR_\alpha = \mathbb{E}(L|L > q_\alpha(L)), \\ CVaR_{i,\alpha} &= \mathbb{E}(L_i|L > q_\alpha(L)). \end{aligned}$$

Bei der Anwendung von Monte Carlo (MC) Simulationen tritt das Problem der Simulation von seltenen Ereignissen auf („rare event simulation“), wie z.B. der VaR für $\alpha = 0,99$. Nur etwa 1% der standard MC Simulationen führt zu einem Verlust L , sodass $L > q_\alpha(L)$.

Der standard MC Schätzer ist

$$\widehat{CVaR}_\alpha^{(MC)}(L) = \frac{1}{\sum_{i=1}^n I_{(q_\alpha, +\infty)}(L_i)} \sum_{i=1}^n L_i I_{(q_\alpha, +\infty)}(L_i),$$

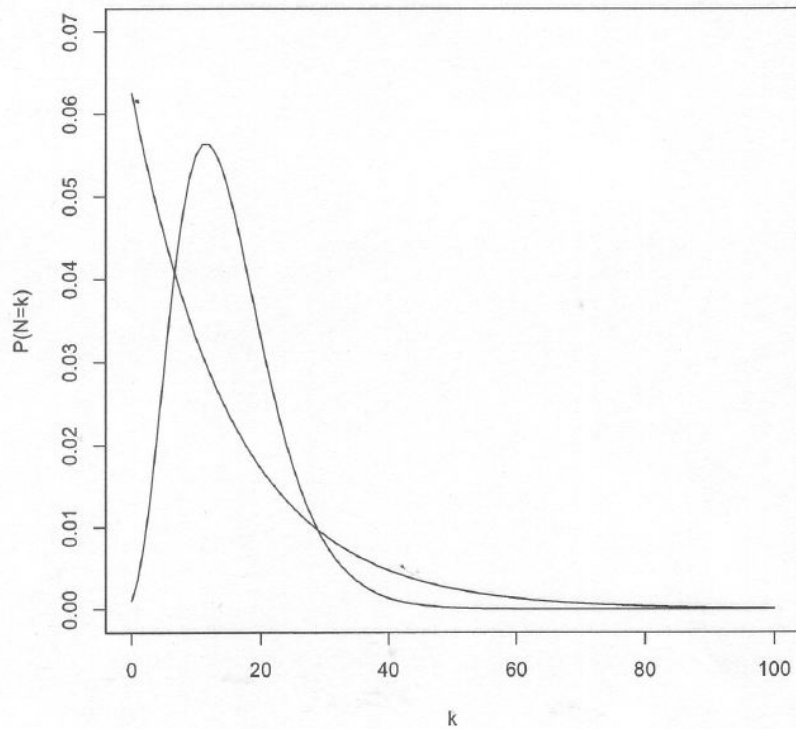


Abbildung 7.1: $\mathbb{P}(N = k)$ für $k = 0, \dots, 100$ für $m = 1$ und $m = 5$

wobei L_i der Verlustwert im i -ten Simulationslauf ist.

$\widehat{CVaR}_\alpha^{(MC)}(L)$ ist sehr instabil, d.h. er hat eine sehr hohe Varianz wenn die Anzahl der Simulationen n nicht sehr groß ist.

7.2.1 Grundlagen von „Importance Sampling“

Sei X eine ZV in einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \mathbb{P})$ mit absolut stetiger Verteilungsfunktion F und Dichtefunktion f .

Gesucht ist

$$\theta = \mathbb{E}(h(X)) = \int_{-\infty}^{\infty} h(x)f(x)dx$$

für eine bekannte Funktion h .

Die Berechnung der Wahrscheinlichkeit eines Ereignisses A erfolgt via $h(x) = I_A(x)$. Die Berechnung von CVaR ist dann

$$h(x) = xI_{x>c}(x) \quad \text{mit} \quad c = VaR(X).$$

Algorithmus 7.1 (Monte Carlo Integration)

- (1) Generiere X_1, X_2, \dots, X_n unabhängige ZV aus der Dichte f .
- (2) Berechne den standard MC Schätzer

$$\hat{\theta}_n^{(MC)} = \frac{1}{n} \sum_{i=1}^n h(X_i).$$

Aus dem starken Gesetz der großen Zahlen folgt

$$\lim_{n \rightarrow \infty} \hat{\theta}_n^{(MC)} = \theta.$$

Im Falle von *seltenen Ereignissen* ($h(x) = I_A(x)$, $\mathbb{P}(A) \ll 1$) ist die Konvergenz sehr langsam!

Sei g eine Wahrscheinlichkeitsdichte, sodass $f(x) > 0 \Rightarrow g(x) > 0$.

Definition 7.2.1. Der *Likelihood ratio* wird definiert als

$$r(x) := \begin{cases} \frac{f(x)}{g(x)} & g(x) > 0 \\ 0 & g(x) = 0 \end{cases}.$$

Es gilt damit:

$$\theta = \int_{-\infty}^{\infty} h(x)r(x)g(x)dx = \mathbb{E}_g(h(x)r(x)) \quad (7.8)$$

Algorithmus 7.2 (Importance Sampling)

- (1) Generiere X_1, X_2, \dots, X_n unabhängige ZV aus der Dichte g .
- (2) Berechne den IS Schätzer

$$\hat{\theta}_n^{(IS)} = \frac{1}{n} \sum_{i=1}^n h(X_i)r(X_i).$$

g heißt „Importance Sampling“-Dichte.

Ziel ist es nun eine geeignete „Importance Sampling“-Dichte auszuwählen, sodass die Varianz des IS-Schätzers wesentlich kleiner als die Varianz des standard MC Schätzers ist.

$$\text{Var}_g \left(\hat{\theta}_n^{(IS)} \right) = (1/n) (\mathbb{E}_g(h^2(X)r^2(X)) - \theta^2)$$

$$\text{Var} \left(\hat{\theta}_n^{(MC)} \right) = (1/n) (\mathbb{E}(h^2(X)) - \theta^2)$$

Theoretisch kann die Varianz des IS-Schätzers auf 0 reduziert werden! Sei dazu $h(x) \geq 0, \forall x$. Für $g^*(x) = f(x)h(x)/E(h(x))$ gilt:

$$\hat{\theta}_1^{(IS)} = h(X_1)r(X_1) = \mathbb{E}(h(X)).$$

Der IS-Schätzer gibt den richtigen Wert nach einer einzigen Simulation wieder!

Sei $h(x) = I_{\{X \geq c\}}(x)$, wobei $c \gg \mathbb{E}(X)$ (seltenes Ereignis).

Es gilt $\mathbb{E}(h^2(X)) = \mathbb{P}(X \geq c)$ und aus (7.8) folgt:

$$\mathbb{E}_g(h^2(X)r^2(X)) = \int_{-\infty}^{\infty} h^2(x)r^2(x)g(x)dx = \mathbb{E}_g(r^2(X)|X \geq c) = \quad (7.9)$$

$$= \int_{-\infty}^{\infty} h^2(x)r(x)f(x)dx = \int_{-\infty}^{\infty} h(x)r(x)f(x)dx = \mathbb{E}(r(X)|X \geq c). \quad (7.10)$$

Das Ziel ist es nun g so auszuwählen, dass $\mathbb{E}_g(h^2(X)r^2(X))$ klein wird, oder sodass $r(x)$ für $x \geq c$ klein und das Ereignis $X \geq c$ unter der Dichte g wahrscheinlicher als unter der Dichte f ist.

7.2.2 Exponential tilting: Bestimmung der IS-Dichte für „light tailed“ Variablen

Sei $M_x(t): \mathbb{R} \rightarrow \mathbb{R}$ die momentengenerierende Funktion von X :

$$M_X(t) = \mathbb{E}(e^{tX}) = \int_{-\infty}^{\infty} e^{tx} f(x)dx.$$

Die IS-Dichte ist $g_t(x) = \frac{e^{tx}f(x)}{M_X(t)}$ und der Likelihood Ratio ist $r_t(x) = \frac{f(x)}{g_t(x)} = M_X(t)e^{-tx}$.

Sei $\mu_t = \mathbb{E}_{g_t}(X) = \mathbb{E}(Xe^{tX})/M_X(t)$. Wie wählt man ein geeignetes t für ein bestimmtes IS Problem?

Z.B. für die Schätzung der Tail-Wahrscheinlichkeit: Das Ziel ist es t so zu wählen, dass $\mathbb{E}(r(X)|X \geq c) = \mathbb{E}(I_{X \geq c}M_X(t)e^{-tX})$ klein wird.

$$e^{-tx} \leq e^{-tc}, \quad \text{für } x \geq c, t \geq 0 \quad \Rightarrow \quad \mathbb{E}(I_{X \geq c}M_X(t)e^{-tX}) \leq M_X(t)e^{-tc}.$$

Wir setzen $t = \arg \min\{M_X(t)e^{-tc} : t \geq 0\}$, denn

$$\min_{t \geq 0} M_X(t)e^{-tc} \iff \min_{t \geq 0} \ln(M_X(t)e^{-tc}) \iff \min_{t \geq 0} \underbrace{(\ln M_X(t) - tc)}_{F(t)}$$

$$F'(t) = \underbrace{\frac{\mathbb{E}[xe^{tx}]}{M_X(t)}}_{\mu_t} - c \stackrel{!}{=} 0$$

Daraus folgt $t = t(c)$, wobei $t(c)$ die Lösung der Gleichung $\mu_t = c$ ist. (Eine eindeutige Lösung existiert für alle relevanten Werte von c - ohne Beweis.)

Exponential Tilting für die Normalverteilung

Sei $X \sim N(0, 1)$ mit Dichtefunktion $\phi(x)$.

$$g_t(x) = \frac{e^{tx}\phi(x)}{M_X(t)} = \frac{e^{tx}\phi(x)}{e^{t^2/2}} = \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}(x-t)^2\right\}$$

$$\mu_t = \frac{\mathbb{E}(X \exp\{tX\})}{M_X(t)} = t$$

D.h. unter der Verteilung g_t gilt $X \sim N(t, 1)$. Die Gleichung $\mu_t = c$ lautet $t = c$.

7.2.3 IS im Falle von Wahrscheinlichkeitsmaßen

Seien f und g Wahrscheinlichkeitsdichten. Definiere zwei Wahrscheinlichkeitsmaße \mathbb{P} und \mathbb{Q} :

$$\mathbb{P}(A) = \int_{x \in A} f(x) dx \quad \text{und} \quad \mathbb{Q}(A) = \int_{x \in A} g(x) dx.$$

Die grundlegende Gleichung der IS (7.8) lautet dann:

$$\theta = \mathbb{E}^{\mathbb{P}}(h(X)) = \mathbb{E}^{\mathbb{Q}}(h(X)r(X)).$$

Exponential tilting im Fall von Wahrscheinlichkeitsdichten funktioniert nun analog zu vorhin: Sei X eine ZV in $(\Omega, \mathcal{F}, \mathbb{P})$, sodass $M_X(t) = \mathbb{E}^{\mathbb{P}}(\exp\{tX\}) < \infty, \forall t$. Sei weiters $\mathbb{Q}_t(A) := \mathbb{E}^{\mathbb{P}}\left(\frac{\exp\{tX\}}{M_X(t)} \mid A\right)$ ein Wahrscheinlichkeitsmaß in (Ω, \mathcal{F}) .

Der IS-Algorithmus bleibt gleich: Simuliere unabhängige Realisierungen von X_i in $(\Omega, \mathcal{F}, \mathbb{Q}_t)$ und setze

$$\hat{\theta}_n^{(IS)} = \frac{1}{n} \sum_{i=1}^n h(X_i) r_t(X_i),$$

wobei $r_t(X) = M_X(t) \exp\{-tX\}$.

Anwendung von IS auf Bernoulli Mixture Modelle (siehe Glasserman und Li (2003))

Sei $L = \sum_{i=1}^m e_i Y_i$ die Verlustfunktion eines Kreditportfolios. Y_i sind die Verlustindikatoren mit Default-Wahrscheinlichkeit \bar{p}_i und $e_i = (1 - \lambda_i)L_i$ sind die positiven deterministischen Exposures (λ_i sind recovery rates und L_i sind die Kredithöhen), $i = 1, 2, \dots, m$.

Sei Ψ ein Vektor von ökonomischen Einflussfaktoren, sodass $Y_i | \Psi$ unabhängig sind und $Y_i | \Psi = \psi \sim \text{Bernoulli}(p_i(\psi))$. Ziel ist es nun $\theta = \mathbb{P}(L \geq c)$ mit Hilfe des IS-Ansatzes für ein gegebenes $c, c \gg \mathbb{E}(L)$, zu schätzen.

Vereinfachter Fall

Seien die $Y_i, i = 1, 2, \dots, m$, unabhängig und sei $\Omega = \{0, 1\}^m$ der Raum der Zustände von Y . Das Wahrscheinlichkeitsmaß \mathbb{P} in Ω ist

$$\mathbb{P}(\{y\}) = \prod_{i=1}^m \bar{p}_i^{y_i} (1 - \bar{p}_i)^{1-y_i}, \quad y \in \{0, 1\}^m.$$

Die momentenerzeugende Funktion von L ist $M_L(t) = \prod_{i=1}^m (e^{te_i \bar{p}_i} + 1 - \bar{p}_i)$.

Das Wahrscheinlichkeitsmaß \mathbb{Q}_t ist

$$\mathbb{Q}_t(\{y\}) = \prod_{i=1}^m \frac{\exp\{te_i y_i\}}{\exp\{te_i \bar{p}_i\} + 1 - \bar{p}_i} \bar{p}_i^{y_i} (1 - \bar{p}_i)^{1-y_i}.$$

Seien $\bar{q}_{t,i}$ neue Default-Wahrscheinlichkeiten:

$$\bar{q}_{t,i} := \frac{\exp\{te_i\}\bar{p}_i}{\exp\{te_i\}\bar{p}_i + 1 - \bar{p}_i}.$$

Dann gilt:

$$\mathbb{Q}_t(\{y\}) = \prod_{i=1}^m \bar{q}_i^{y_i} (1 - \bar{q}_i)^{1-y_i}, \quad y \in \{0, 1\}^m.$$

D.h. nach der exponential tilting sind die Default-Indikatoren unabhängig mit neuen Default-Wahrscheinlichkeiten $\bar{q}_{t,i}$. Des weiteren gilt $\lim_{t \rightarrow \infty} \bar{q}_{t,i} = 1$ und $\lim_{t \rightarrow -\infty} \bar{q}_{t,i} = 0$ und somit nimmt $\mathbb{E}^{\mathbb{Q}_t}(L)$ für $t \in \mathbb{R}$ alle Werte in $(0, \sum_{i=1}^m e_i)$ an. Für IS-Anwendungen wählt man t so, dass $\sum_{i=1}^m e_i \bar{q}_{t,i} = c$.

Allgemeiner Fall:

Y_i sind unabhängig bedingt durch Ψ .

1. Schritt: Schätzung der bedingten Überschuss-Wahrscheinlichkeit

$$\theta(\psi) := \mathbb{P}(L \geq c | \Psi = \psi)$$

für eine gegebene Realisierung ψ der ökonomischen Faktoren Ψ , mit Hilfe des im vereinfachten Fall beschriebenen IS-Ansatzes.

Algorithmus 7.3 (IS für die bedingte Verlustverteilung)

- (1) Für ein gegebenes ψ berechne die spezifischen Default-Wahrscheinlichkeiten $p_i(\psi)$ basierend auf dem spezifischen Modell und löse folgende Gleichung:

$$\sum_{i=1}^m \frac{e_i \exp\{te_i\} p_i(\psi)}{\exp\{te_i\} p_i(\psi) + 1 - p_i(\psi)} = c.$$

Die Lösung $t = t(c, \psi)$ gibt den richtigen *tilting-Grad*.

- (2) Erzeuge n_1 bedingte Realisierungen des Vektors der Default-Indikatoren (Y_1, \dots, Y_m) . Die einzelnen Indikatoren Y_i werden unabhängig aus Bernoulli(q_i), $i = 1, 2, \dots, m$, simuliert, wobei

$$q_i = \frac{\exp\{t(c, \psi)e_i\} p_i(\psi)}{\exp\{t(c, \psi)e_i\} p_i(\psi) + 1 - p_i(\psi)}.$$

- (3) Sei $M_L(t, \psi) := \prod [\exp\{te_i\} p_i(\psi) + 1 - p_i(\psi)]$ die bedingte momentenerzeugende Funktion von L . Seien $L^{(1)}, L^{(2)}, \dots, L^{(n_1)}$ die n_1 bedingten Realisierungen von L für die n_1 simulierten Realisierungen von Y_1, Y_2, \dots, Y_m . Berechne den *IS-Schätzer* für die Tail-Wahrscheinlichkeit der bedingten Verlustverteilung:

$$\hat{\theta}_{n_1}^{(IS)}(\psi) = M_L(t(c, \psi), \psi) \frac{1}{n_1} \sum_{j=1}^{n_1} I_{L^{(j)} \geq c} \exp\{-t(c, \psi)L^{(j)}\}.$$

2. Schritt: Schätzung der unbedingten Überschuss-Wahrscheinlichkeit $\theta = \mathbb{P}(L \geq c)$.

Eine naive Vorgangsweise wäre, mehrere Realisierungen ψ der Einflussfaktoren Ψ zu erzeugen und $\hat{\theta}_{n_1}^{(IS)}(\psi)$ für jede dieser Realisierungen zu berechnen. Der gesuchte Schätzer ist der Durchschnittswert der Schätzer $\hat{\theta}_{n_1}^{(IS)}(\psi)$ über alle Realisierungen ψ . Diese Lösung ist nicht die beste, siehe Glasserman und Li (2003).

Eine bessere Herangehensweise ist folgende: IS für die Einflussfaktoren.

Annahme: $\Psi \sim N_p(0, \Sigma)$ (zB. Probit-normal Bernoulli mixture)

Die IS-Dichte g ist die Dichte von $N_p(\mu, \Sigma)$ für einen „neuen“ Erwartungswertvektor $\mu \in \mathbb{R}^p$. Eine gute Wahl von μ sollte zu häufigen Realisierungen ψ führen die zu höheren bedingten Default-Wahrscheinlichkeiten $p_i(\psi)$ führen.

Der Likelihood Ratio ist dann:

$$r_\mu(\Psi) = \frac{\exp\{-\frac{1}{2}\Psi'\Sigma^{-1}\Psi\}}{\exp\{-\frac{1}{2}(\Psi - \mu)'\Sigma^{-1}(\Psi - \mu)\}} = \exp\{-\mu'\Sigma^{-1}\Psi + \frac{1}{2}\mu'\Sigma^{-1}\mu\}.$$

Algorithmus 7.4 (vollständige IS für Bernoulli mixture Modelle mit Gauss'schen Faktoren)

- (1) Erzeuge $\psi_1, \psi_2, \dots, \psi_n \sim N_p(\mu, \Sigma)$ (n ist die Anzahl der Simulationsrunden)
- (2) Für jedes ψ_i berechne $\hat{\theta}_{n_1}^{(IS)}(\psi_i)$ wie in Algorithmus 7.3.
- (3) Berechne den IS-Schätzer für die unbedingte Überschuss-Wahrscheinlichkeit:

$$\hat{\theta}_n^{(IS)} = \frac{1}{n} \sum_{i=1}^n r_\mu(\psi_i) \hat{\theta}_{n_1}^{(IS)}(\psi_i)$$

Die Auswahl von μ :

μ soll so gewählt werden, dass die Varianz des Schätzers klein ist. Die Idee von Glasserman und Li (2003) ist folgende (Skizze):

$\hat{\theta}_{n_1}^{(IS)}(\psi) \approx \mathbb{P}(L \geq c | \Psi = \psi) \implies$ suche eine gute IS-Dichte für die Funktion $\psi \rightarrow \mathbb{P}(L \geq c | \Psi = \psi)$.

Die optimale IS-Dichte g^* ist proportional zu

$$\mathbb{P}(L \geq c | \Psi = \psi) \exp\left\{-\frac{1}{2}\psi'\Sigma^{-1}\psi\right\}.$$

Man sollte also eine IS-Dichte mit demselben Modus wie die optimale Dichte g^* wählen. Das

führt zu folgendem Optimierungsproblem:

$$\mu = \arg \max_{\psi} \left\{ \mathbb{P}(L \geq c | \Psi = \psi) \exp \left\{ -\frac{1}{2} \psi' \Sigma^{-1} \psi \right\} \right\}.$$

Eine exakte Lösung ist hier schwierig, weil $\mathbb{P}(L \geq c | \Psi = \psi)$ i.a. nicht in analytischer Form verfügbar ist (siehe Glasserman und Li (2003) für Lösungsansätze).