

**Endbericht zum Vorhaben
FKZ: 3609S50002**

**Vergleichende Betrachtung der
probabilistischen/stochastischen
und deterministischen
Modellierung von Expositionen im
Hinblick auf die Belastbarkeit des
Modellergebnisses und die
Anforderungen an die Qualität der
Eingangsdaten
- Anhänge A, B, C -**

BS-Projekt-Nr. 0903-01

erstellt im Auftrag des
Bundesamt für Strahlenschutz
Willy-Brandt-Straße 5
38226 Salzgitter

durch die

Brenk Systemplanung GmbH
Heider-Hof-Weg 23
D-52080 Aachen

Aachen, 07.09.2012

Hinweis:

Der Bericht gibt die Auffassung und Meinung des Auftragnehmers wieder und muss nicht mit der Meinung des Auftraggebers (Bundesminister für Umwelt, Naturschutz und Reaktorsicherheit) übereinstimmen.

INHALTSVERZEICHNIS ZU DEN ANHÄNGEN

ANHANG A: GLOSSAR	456
ANHANG B GRUNDLAGEN DER STATISTIK	462
B.1 GRUNDLEGENDE BEGRIFFE UND DEFINITIONEN	463
B.1.1 Wahrscheinlichkeitsdichte und Verteilungsfunktion	463
B.1.2 Erwartungswert und Varianz	464
B.1.3 Multivariate Zufallsgrößen, Kovarianz, Korrelationskoeffizient	465
B.1.4 Zerlegung der Varianz einer Funktion von zwei Variablen	467
B.2 PUNKTSCHÄTZER	468
B.2.1 Stichprobenmittelwert, Stichprobenvarianz, Stichprobenkorrelationskoeffizient	468
B.2.2 Schiefe und Wölbung	470
B.2.3 Abhängigkeitsmaße	471
B.3 COPULAS	472
B.3.1 Grundlegende Begriffe und Aussagen der Copula-Theorie	473
B.3.2 Familien von Copulas	478
B.3.2.1 Copulas von elliptischen Verteilungen	479
B.3.2.2 Archimedische Copulas	481
B.3.2.3 MARSHALL-OLKIN-Copula	488
B.3.3 Berechnung der empirischen Copula und KENDALL-Verteilung	491
B.4 STOCHASTISCHE PROZESSE UND DIFFERENTIALGLEICHUNGEN	492
B.4.1 Zeitkontinuierliche stochastische Prozesse	492
B.4.1.1 Definition von LÉVY-Prozessen	492
B.4.1.2 Standard-WIENER-Prozess (SWP)	493
B.4.1.3 Elementarer POISSON-Prozess (EPP)	495
B.4.1.4 Zusammengesetzte POISSON-Prozesse (ZPP)	499
B.4.1.5 Darstellung von LÉVY-Prozessen	501
B.4.1.6 Gammaprozess	503
B.4.2 Stochastische Differentialgleichungen (SDE)	506
B.4.2.1 SDE mit WIENER-Prozess als Rauschquelle	507
B.4.2.2 SDE mit POISSON-Prozess als Rauschquelle	510
ANHANG C UNIVARIATE PARAMETRISCHE VERTEILUNGEN	512
C.1 VORBEMERKUNGEN	513
C.2 STETIGE VERTEILUNGEN MIT UNBEGRENZTEM DEFINITIONSBEREICH ($-\infty < X < \infty$)	513
C.2.1 Normalverteilung $X \sim N(\mu, \sigma)$	513
C.2.2 STUDENT'sche t -Verteilung $X \sim t(\nu)$	515
C.3 STETIGE VERTEILUNGEN FÜR POSITIV DEFINIERTE ZUFALLSVARIABLE ($x \geq 0$)	517
C.3.1 Chi-Quadrat-Verteilung $X \sim \chi^2(\nu)$	517
C.3.2 Lognormalverteilung $X \sim \text{Ln}(\mu, \sigma)$	518
C.3.3 Exponentialverteilung $X \sim \text{Ex}(\lambda)$	519
C.3.4 Gammaverteilung $X \sim \text{Gam}(\alpha, \lambda)$ oder $X \sim \text{Ga}(\alpha, \beta)$	520
C.3.5 WEIBULL-Verteilung $X \sim \text{We}(\alpha, \beta)$	524

C.3.6	LOMAX-Verteilung $X \sim Lo(\alpha, \eta)$	526
C.3.7	WALD-Verteilung $X \sim Wa(\mu, \lambda)$	528
C.4	STETIGE VERTEILUNGEN FÜR BEIDSEITIG BEGRENZTE ZUFALLSVARIABLE ($0 \leq x \leq 1$)	531
C.4.1	Gleichverteilung $X \sim U$	531
C.4.2	Betaverteilung $X \sim Bt(\alpha, \beta)$	531
C.4.3	KUMARASWAMY-Verteilung $X \sim Ku(\alpha, \beta)$	534
C.4.4	JOHNSON- S_B -Verteilung $X \sim JS_B(\gamma, \delta)$	535
C.4.5	Log-PEARSON-3-Verteilung $X \sim LP3(\alpha, \lambda)$	536
C.4.6	Dreiecksverteilung $Tr(a, b, m)$	537
C.5	DISKRETE VERTEILUNGEN	539
C.5.1	BERNOULLI-Verteilung $X \sim Ber(p)$	539
C.5.2	Geometrische Verteilung $X \sim Geo(p)$	539
C.5.3	Binomialverteilung $X \sim Bin(n, p)$	539
C.5.4	POISSON-Verteilung $X \sim Poi(\lambda)$	540
C.5.5	Negative Binomialverteilung $X \sim NBin(r, p)$ oder $NegB(\alpha, \beta)$ oder $NB(\lambda, \delta)$	542

ANHANG A: GLOSSAR

- A-D-Test:** (Abkürzung für Anderson-Darling-Test): Anpassungstest zur Prüfung der Hypothese, dass eine gegebene iid Zufallsstichprobe einer bestimmten parametrischen Verteilung zugeordnet werden kann, der auf einem funktionalen Maß für den Abstand zwischen der empirischen und der angenommenen Verteilungsfunktion basiert, wobei die Randbereiche der statistischen Verteilung verstärkt gewichtet werden.
- aleatorische Unsicherheit:** (siehe Variabilität)
- Chi-Quadrat-Test:** Anpassungstest (nach PEARSON) zur Prüfung der Hypothese, dass eine gegebene iid Zufallsstichprobe einer bestimmten Verteilung zugeordnet werden kann. Er basiert für festzulegende Intervalle des Definitionsbereichs der Zufallsgröße auf dem Vergleich erwarteter und beobachteter Häufigkeit von Realisierungen in diesen Intervallen. Die Entscheidung erfolgt mit der asymptotisch (für sehr große Stichprobenumfänge) geltenden χ^2 -Verteilung.
- Dosiskoeffizient:** Der Dosiskoeffizient ist der Quotienten einer Dosis zur Aktivitätszufuhr eines radioaktiven Stoffes. Dieser Begriff (bzw. auch der Begriff Dosis-konversionsfaktor) dienen auch zur Bezeichnung anderer Koeffizienten, die Aktivitäten oder Aktivitätskonzentrationen mit Dosen oder Dosisleistungen verknüpfen (z. B. die externe Dosisleistung in einer bestimmten Entfernung oberhalb einer Fläche mit einer Deposition einer bestimmten Aktivität pro Flächeneinheit eines bestimmten Radionuklids auf dieser Fläche [BFS 09]). Zur Berechnung effektiver Folgedosen infolge Ingestion bzw. Inhalation von Radionukliden sind u. a. in [EU 96] entsprechende Dosiskoeffizienten für potentiell relevante Radionuklide für 6 Altersgruppen der allgemeinen Bevölkerung sowie für Beschäftigte zusammengestellt.
- effektive Dosis:** Die effektive Dosis E ist die Summe der gewichteten Organdosen H_T ,
- $$E = \sum_T w_T \cdot H_T,$$
- wobei mit w_T die Organ- bzw. Gewebewichtungsfaktoren bezeichnet sind (siehe Tabelle 3 und Anhang A in [ICRP 07]). Die Einheit von E ist $J \cdot kg^{-1}$ und trägt den speziellen Namen Sievert (Sv).
- effektive Folgedosis:** Die effektive Folgedosis $E(\tau)$ ist analog zur effektiven Dosis die gewichtete Summe der Organfolgedosen $H_T(\tau)$, wobei τ den Integrationszeitraum nach einer Aktivitätszufuhr bezeichnet.
- epistemische Unsicherheit:** (siehe Unsicherheit)

Exposition:	Konzentration oder Menge eines speziellen Stoffes, der ein Zielorganismus, ein System oder eine (Sub-)Population mit einer bestimmten Häufigkeit und Dauer erreicht, bzw. Einwirkung einer Strahlung
Expositionsroute:	Ingestion, Inhalation und dermale Aufnahme sowie äußere Exposition durch Strahlung, kennzeichnet den Mechanismus der Exposition
Expositionspfad:	Weg, den ein Schadstoff von einer Quelle (z. B. der Ableitung aus einer Anlage oder Einrichtung) über einen Ausbreitungs-/Transportvorgang in den Umweltmedien bis zur Exposition des Menschen nimmt.
Expositionsszenario:	Gesamtheit der Bedingungen und Annahmen zu Quellen, Expositionspfaden und Mengen bzw. Konzentrationen relevanter Schadstoffe, der Art, Anzahl bzw. Charakteristika belasteter Organe, Personen oder Personengruppen, die zur Modellierung und Abschätzung von Expositionen in einer gegebenen oder einer zukünftigen Situation herangezogen werden.
Häufigkeitsverteilung:	Tabellarische oder grafische Zusammenstellung/Darstellung zur Häufigkeit der Realisierungen einer Zufallsgröße.
iid	unabhängig und gleichartig verteilt (englisch: independent and identically distributed). Ist eine grundlegende Annahme zu Werten einer Stichprobe für die statistische Inferenz zu parametrischen Verteilungen.
K-S-Test:	(Abkürzung für Kolmogorov-Smirnov-Test): Anpassungstest zur Prüfung der Hypothese, dass eine gegebene iid Zufallsstichprobe einer bestimmten parametrischen Verteilung zugeordnet werden kann, der auf dem maximalen Abstand zwischen der empirischen und angenommenen Verteilungsfunktion basiert.
Likelihood-Funktion:	Produkt der Wahrscheinlichkeitsdichten einer parametrischen Verteilung mit Werten der Zufallsgröße X für die einzelnen Stichprobendaten, als Funktion der unbekanntem Verteilungsparametern der Population betrachtet wird; stellt ein Maß für die Glaubwürdigkeit dafür dar, dass die Stichprobe aus der betrachteten Verteilungsfamilie stammt (in Abhängigkeit von den möglichen Werten der Verteilungsparameter).
LL-Test:	(Abkürzung für Log-Likelihood-Test): Test zum Vergleich der Plausibilität von unterschiedlichen Verteilungsfamilien, die als generierende Population einer gegebenen iid Zufallsstichprobe in Betracht gezogen werden, der auf der Relation der Maxima der jeweiligen Likelihood-Funktionen bzw. der Differenz ihrer Logarithmen basiert.
Medianwert:	50%-Perzentil (x_{50}) der statistischen Verteilung einer Zufallsgröße X .
Messunsicherheit:	Zu dem aus einer Messung resultierenden Schätzwert einer physikalischen Größe grenzt die Messunsicherheit einen Wertebereich ein, in welchem der wahre Wert der Messgröße mit einer vorgegebenen Wahrscheinlichkeit liegt. Das Ergebnis einer Messung ist somit durch Schätzer und Messunsicherheit definiert; es wird in der Form (Schätzwert \pm Messunsicherheit) angegeben.

- Mittelwert:** Der Mittelwert bzw. Erwartungswert $E[X]$ einer Zufallsgröße X beschreibt den durchschnittlichen Wert der Realisierungen von X und stellt ein Maß für die Lage des Schwerpunktes der statistischen Verteilung dar, er kann als 1. zentrale Moment der Zufallsgröße berechnet werden.
- mittlere Energiedosis:** Die mittlere Energiedosis $D_{T,R}$ ist die durch eine (ionisierende) Strahlung R auf eine Gewebe bzw. Organ T übertragene mittlere Energie, bezogen auf die Masse dieses Gewebes oder Organs. Die Einheit von $D_{T,R}$ ist $J \cdot kg^{-1}$ und trägt den speziellen Namen Gray (Gy).
- ML-Schätzer** (Abkürzung für Maximum-Likelihood-Schätzer); schätzt die Parameter einer parametrischen Verteilungsfamilie durch Bestimmung des Maximums der Likelihood-Funktion unter der Annahme, dass die Stichprobenwerte als iid Zufallsgrößen einer bestimmten Population angesehen werden können.
- Modalwert:** Der Modalwert (auch Modus genannt) ist ein Wert aus der statistischen Verteilung einer Zufallsgröße, dem die größte Häufigkeit beizumessen ist bzw. bei dem die Wahrscheinlichkeitsdichte ein Maximum erreicht. Es können auch mehrere Modalwerte in einer statistischen Verteilung auftreten (insbesondere in gemixten Verteilungen), die dann als multimodale Verteilungen bezeichnet werden.
- Modellunsicherheit:** Auswirkung von fehlendem Wissen, unvollständiger Information oder von vereinfachenden Annahmen bei der Modellierung eines Expositionspfades, einer Extrapolation von Modellen über den Rahmen hinaus, in dem diese anhand von empirischen Untersuchungen/Daten validiert wurden. Hierzu gehören aber auch die unzureichende bzw. nicht adäquate Berücksichtigung von Korrelationen zwischen Modellparametern sowie die Auswirkungen der Wahl idealisierter statistischer Verteilungen für einzelne Modellparameter.
- Monte-Carlo-Simulation:** Numerisches Verfahren, mit dem für alle Modellvariablen, die durch Wahrscheinlichkeitsverteilungen spezifiziert sind, aus ihren Verteilungen zufällige Werte generiert werden, wonach zufällige Realisierungen der Zielgröße resultieren. Nach sehr vielen Wiederholungen dieser Prozedur erhält man für die Zielgröße eine statistische Verteilung der möglichen Werte, die mit zunehmender Anzahl ausgeführter Berechnungen die Wahrscheinlichkeitsdichte bzw. Verteilungsfunktion der Zielgröße verbessert approximiert.
- Organdosis:** Die Organdosis H_T (auch als Äquivalentdosis bezeichnet) ist die Summe der in einem Organ oder Gewebe T durch Strahlungen R verursachten mittleren Energiedosen $D_{T,R}$,
- $$H_T = \sum_R w_R \cdot D_{T,R} ,$$
- gewichtet mit Strahlungswichtungsfaktoren w_R nach Tabelle 2 in [ICRP 07]. Die Einheit von H_T ist $J \cdot kg^{-1}$ mit dem speziellen Namen Sievert (Sv).

- Organfolgedosis:** Die Organfolgedosis $H_T(\tau)$ ist das Zeitintegral der Organdosisleistung in einem bestimmten Gewebe oder Organ T, die eine Referenzperson nach der Zufuhr eines radioaktiven Stoffes in den Körper erhalten wird, wobei τ den Integrationszeitraum nach einer Aktivitätszufuhr bezeichnet. Er beträgt für Erwachsene 50 Jahre und erstreckt sich für Kinder bis zu einem Alter von 70 Jahren. $H_T(\tau)$ wird nur für interne Strahlenexpositionen berechnet und berücksichtigt die biologische und physikalische Halbwertszeit zugeführter Radionuklide sowie ggf. den Aufbau von radioaktiven Tochternukliden.
- Parameterunsicherheit:** Ungewissheit über den wahren Wert eines Modellparameters oder die realen Kenngrößen seiner statistischen Verteilung. Parameterunsicherheiten haben überwiegend ihre Ursache in der realen Variabilität der Parameter, mit denen z. B. primäre Strahlenquellen, die Ausbreitung von Radionukliden in den Umweltmedien und in der Biosphäre sowie schließlich die Exposition von Personen modelliert werden, da immer nur eine begrenzte Stichprobe zur Schätzung von Modell- bzw. Verteilungsparametern zur Verfügung steht. Parameterunsicherheiten werden aber auch durch Messunsicherheiten beeinflusst. Daneben können qualitative Aspekte, wie die Übertragung von Daten aus früheren Untersuchungen auf eine aktuelle Situation (ungewisse Repräsentativität) oder die Bezugnahme auf Expertenschätzungen (partielle Unkenntnis, Unschärfe der Information) zur Unsicherheit der Schätzung von Parameterwerten beitragen.
- Perzentil:** Das Perzentil der Ordnung P (P%-Perzentil) der statistischen Verteilung der Zufallsgröße X ist der kleinste Wert x_p , der von P % der Realisierungen von X nicht überschritten wird (wird P nicht in Prozent sondern einen Anteil p, $0 < p < 1$, angegeben, so spricht man auch vom p-Quantil).
- Punktschätzer:** Wert einer Stichprobenfunktion/Schätzfunktion für eine Kenngröße, die zur Charakterisierung der statistischen Verteilung einer Zufallsgröße verwendet wird (z. B. für den Mittelwert/Erwartungswert, die Varianz oder Perzentile).
- Referenzmensch:** Idealisierter Mann bzw. idealisierte Frau mit charakteristischen Merkmalen, die von der ICRP für Zwecke des Strahlenschutzes definiert wurden, d. h. mit anatomischen und physiologischen Eigenschaften, wie sie in [ICRP 02] definiert sind.
- Referenzperson:** Eine idealisierte Person, für die Organdosen errechnet werden, indem die entsprechenden Dosen des männlichen und weiblichen Referenzmenschen gemittelt werden. Die Organdosen der Referenzperson werden mit den Gewebewichtungsfaktoren multipliziert, um effektive Dosen zu berechnen. Für Referenzpersonen werden Verhaltensparameter definiert.
- Schiefe:** Kenngröße der statistischen Verteilung einer Zufallsgröße X, die als Maß der Asymmetrie der Wahrscheinlichkeitsdichte bzw. Häufigkeitsverteilung durch das 3. zentrale Moment der Zufallsgröße bestimmt ist.

- Stratifizierung:** Unter Stratifizierung (bzw. Abschichtung) versteht man die Aufteilung einer Gesamtheit in verschiedene Kategorien, Gruppen, Teilpopulationen (Strata), wie z. B. die Einteilung von Personen der allgemeinen Bevölkerung in sechs Altersgruppen für Expositionsabschätzungen oder die Zuordnung meteorologischer Gegebenheiten zu sechs Stabilitätsklassen für die Modellierung von Schadstoffausbreitungen über den Luftpfad. Durch eine Stratifizierung wird die Variabilität relevanter Parameter für die einzelnen Strata gegenüber der ursprünglichen Gesamtheit reduziert.
- Szenariounsicherheit:** Auswirkung von fehlendem Wissen oder unvollständiger Information bzgl. einer adäquaten Beschreibung von Expositionsszenarien, wie sie z. B. aus dem Unwissen über zukünftige natürliche oder anthropogene Einwirkungen auf Schadstoffquellen und Migrationswege, über die Lebensbedingungen von aktuell oder potentiell exponierten Personen oder durch unvollständige Analyse möglicher Expositionspfade resultieren können.
- Unsicherheit:** (wird auch als **epistemische Unsicherheit** bezeichnet) Auswirkung von fehlendem Wissen oder unvollständiger Information bzgl. spezieller Aspekte, Faktoren oder Parameter der Beschreibung des aktuellen bzw. der Prognose eines zukünftigen Zustandes eines untersuchten Systems. Die Unsicherheit resultiert aus subjektiven Merkmalen der Systemanalyse und kann grundsätzlich durch zusätzliche Untersuchungen reduziert werden. Bei Expositionsabschätzungen kann zwischen Szenariounsicherheiten, Modellunsicherheiten und Parameterunsicherheiten unterschieden werden.
- Wölbung:** Kenngröße der statistischen Verteilung einer Zufallsgröße X , die als Maß der Krümmung der Wahrscheinlichkeitsdichte bzw. Häufigkeitsverteilung durch das 4. zentrale Moment der Zufallsgröße bestimmt ist.
- Variable:** Beobachtbarer oder messbarer Modellparameter, der in Abhängigkeit von konkreten Rahmenbedingungen unterschiedliche Werte annehmen kann, die räumlich und/oder zeitlich bedingt sind bzw. einer statistischen Verteilung innerhalb einer Grundgesamtheit unterliegen (z. B. Verhaltensparameter in einer Personengruppe; Aktivitätskonzentration in einem Umweltmedium).
- Variabilität:** (wird auch als **aleatorische Unsicherheit** bezeichnet) Heterogenität, Zeitabhängigkeit oder auch Diversität der Werte eines Modellparameters, die auch durch zusätzliche Untersuchungen/Messungen nicht reduziert, aber mit höherer Konfidenz quantifiziert werden kann.
- Varianz:** Die Varianz (σ^2) ist ein Maß der Variabilität einer Zufallsgröße, das durch ihr 2. zentrales Moment (mittlere quadratische Streuung um den Mittelwert) bestimmt ist.
- Verteilungsfunktion:** Die Verteilungsfunktion $F_X(x)$ (genauer: kumulative Verteilungsfunktion) zeigt die Variabilität einer Zufallsgröße X in tabellarischer oder grafischer Form bzw. als mathematische Funktion bestimmter Verteilungsparameter auf. Sie ist eine monoton nicht fallende Funktion die jedem möglichen Wert

x des Definitionsbereiches von X eine bestimmte Wahrscheinlichkeit p ($0 \leq p \leq 1$) zuordnet, mit der dieser Wert durch die Realisierungen von X unterschritten wird. Man unterscheidet parametrische Verteilungsfunktionen (sie werden durch analytische Funktionen beschrieben, die von bestimmten Parametern abhängen) und empirische Verteilungsfunktionen (diese werden durch die einzelnen Werte einer Stichprobe bestimmt, denen allen gleiche Wahrscheinlichkeiten des Auftretens beigemessen werden).

Wahrscheinlichkeitsdichte:

Mathematische Funktion $f_X(x)$, mit der Variabilität einer Zufallsgröße X sowie die Wahrscheinlichkeit/Häufigkeit des Auftretens von Realisierungen aus Intervallen des Definitionsbereiches von X charakterisiert wird. Sie stellt für stetige Verteilungen die 1. Ableitung der Verteilungsfunktion $F_X(x)$ dar. Somit entspricht die Wahrscheinlichkeit der Realisierungen von X im Intervall $[x_a, x_e]$ dem Integral der Funktion $f_X(x)$ über dieses Intervall.

Zufallsvariable:

Als Zufallsvariable (auch Zufallsgröße) bezeichnet man eine Funktion, die den Ergebnissen eines Zufallsexperiments (d. h. einer zufälligen Stichprobe) so genannte Realisationen als Werte zuordnet. Es wird zwischen diskreten Zufallsvariablen (die Funktion kann nur abzählbar endlich oder unendlich viele Werte annehmen) und kontinuierlichen/stetigen Zufallsvariablen (die Funktion kann alle Werte aus einem Kontinuum annehmen und durch eine Wahrscheinlichkeitsdichte charakterisiert werden) unterschieden.

ANHANG B

GRUNDLAGEN DER STATISTIK

B.1 Grundlegende Begriffe und Definitionen

B.1.1 Wahrscheinlichkeitsdichte und Verteilungsfunktion

Ist X eine kontinuierliche Zufallsvariable, deren Verteilung durch eine Wahrscheinlichkeitsdichte $f_X(x|\boldsymbol{\theta})$ spezifiziert ist (engl.: probability density function, pdf),

$$f_X(x|\boldsymbol{\theta}) \geq 0 \quad \text{mit} \quad \int_{-\infty}^{+\infty} f_X(x|\boldsymbol{\theta}) dx = 1, \quad (\text{B-1})$$

wobei $\boldsymbol{\theta} = \{\theta_1, \theta_2, \dots\}$ der Vektor der Verteilungsparameter ist, so ist die Wahrscheinlichkeit dafür, dass X nicht größer als ein bestimmter Wert x ist¹, durch ihre kumulative Verteilungsfunktion (engl.: cumulative distribution function, cdf) $F_X(x|\boldsymbol{\theta})$ gegeben:

$$F_X(x|\boldsymbol{\theta}) = \Pr(X \leq x | \boldsymbol{\theta}) = \int_{-\infty}^x f_X(t|\boldsymbol{\theta}) dt. \quad (\text{B-2})$$

$F_X(x|\boldsymbol{\theta})$ ist eine positive, nicht fallende Funktion von x mit

$$0 \leq F_X(x|\boldsymbol{\theta}) \leq 1 \quad \text{und} \quad \lim_{x \rightarrow \infty} (F_X(x|\boldsymbol{\theta})) = 1. \quad (\text{B-3})$$

Eine kurze Schreibweise zur Kennzeichnung der Art der Verteilung einer Zufallsgröße ist $X \sim \Phi(\boldsymbol{\theta})$, wobei Φ für den Typ der parametrischen Verteilung steht und $\boldsymbol{\theta}$ die Verteilungsparameter angibt (z. B. bezeichnet $X \sim N(\mu, \sigma)$ eine normal verteilte Zufallsgröße mit dem Erwartungswert $\mu = E[X]$ und der Standardabweichung $\sigma = \sqrt{\text{Var}[X]}$; vgl. Anhang C).

Das p -Quantil x_p (zur Wahrscheinlichkeit p , $0 < p < 1$) bzw. das $P\%$ -Perzentil x_P (mit $P = p \cdot 100$) ist jener Wert der Zufallsvariablen X , der mit der Wahrscheinlichkeit p nicht überschritten wird:

$$F_X(x_p|\boldsymbol{\theta}) = p. \quad (\text{B-4})$$

Eine Zufallsvariable X heißt diskret, wenn sie nur endlich oder abzählbar unendlich viele Werte x_i ($i = 1, 2, \dots$) annehmen kann. Sie wird durch eine Wahrscheinlichkeitsmassenfunktion (engl.: probability mass function, pmf) beschrieben,

$$f_X(x_i|\boldsymbol{\theta}) = \Pr(X = x_i|\boldsymbol{\theta}) = p_i \quad \text{mit} \quad 0 \leq p_i \leq 1 \quad \text{und} \quad \sum_i p_i = 1, \quad (\text{B-5})$$

wobei p_i die Wahrscheinlichkeiten der möglichen Realisierungen x_i bei gegebenen Verteilungsparametern $\boldsymbol{\theta}$ sind. Die kumulative Verteilungsfunktion für diskrete Zufallsvariable ist:

$$F_X(x|\boldsymbol{\theta}) = \Pr(X \leq x | \boldsymbol{\theta}) = \sum_{x_i \leq x} \Pr(X = x_i | \boldsymbol{\theta}) = \sum_{i: x_i \leq x} p_i. \quad (\text{B-6})$$

¹ In der Wahrscheinlichkeitstheorie/Statistik ist es üblich, Zufallsgrößen mit Großbuchstaben und ihre Realisierungen mit kleinen Buchstaben zu bezeichnen. Im vorliegenden Bericht schließen wir uns dieser Konvention möglichst an.

Eine Stichprobe vom Umfang n , $\{X\}_n$, besteht aus n Zufallsgrößen X_1, X_2, \dots, X_n (bzw. den nach der Größe aufsteigend geordneten Größen $X_{(1)}, X_{(2)}, \dots, X_{(n)}$). Die empirische Verteilungsfunktion ist mit diesen Werten definiert durch:

$$F_n(x) = \begin{cases} 0 & \text{für } x < X_{(1)} \\ i/n & \text{für } X_{(i)} \leq x < X_{(i+1)} \quad i = 1, 2, \dots, n-1. \\ 1 & \text{für } X_{(n)} \leq x \end{cases} \quad (\text{B-7})$$

Anmerkung: Für graphische Darstellungen und Regressionsberechnungen für Realisierungen $\{x\}_n$ ist es oft ausreichend bzw. auch angebracht, die empirische Verteilungsfunktion nur an den Stellen $x_{(i)}$ zu betrachten und hierfür mittlere Funktionswerte $F_n(x_{(i)}) = (i - 0,5)/n$ anzusetzen ($i = 1, 2, \dots, n$).

B.1.2 Erwartungswert und Varianz

Die beiden wichtigsten Kenngrößen statistischer Verteilungen sind ihr Erwartungswert und ihre Varianz, die für kontinuierliche Zufallsvariable durch folgende Integrale definiert sind:

$$E_X = E[X] = \int_{-\infty}^{+\infty} x \cdot f_X(x | \theta) dx, \quad (\text{B-8})$$

$$\text{Var}[X] = E[(X - E_X)^2] = \int_{-\infty}^{+\infty} (x - E_X)^2 \cdot f_X(x | \theta) dx. \quad (\text{B-9})$$

Neben der Definition der Varianz nach Gl. (B-9) ist folgende alternative Formel nützlich, die sich unmittelbar nach Auflösung des Quadrats $(X - E_X)^2$ ergibt.

$$\text{Var}[X] = E[(X - E_X)^2] = E[X^2] - E_X^2. \quad (\text{B-10})$$

Der Erwartungswert einer durch X mittels einer Funktion $g(X)$ transformierten Zufallsvariable ist:

$$E[g(X)] = \int_{-\infty}^{+\infty} g(x) \cdot f_X(x | \theta) dx. \quad (\text{B-11})$$

Die mit Gl. (B-9) definierte Varianz ist ein spezieller Fall von Gl. (B-11), und zwar für die Funktion $g(X) = (X - E_X)^2$.

Anmerkung: Der Erwartungswert E_X ist der Wert der Zufallsgröße X , für den sich die kleinste mittlere quadratische Abweichung von X ergibt, die dann durch $\text{Var}[X]$ bestimmt ist. Es sei c eine Konstante. Beachtet man, dass gemäß Gl. (B-11) der Erwartungswert einer Summe von X abhängiger Funktionen gleich der Summe der Erwartungswerte dieser Funktionen ist und dass nach Gl. (B-1) der Erwartungswert einer festen Größe/Konstante gleich dieser Größe ist, so folgt:

$$E[(X - c)^2] = E[(X - E_X + E_X - c)^2] = E[(X - E_X)^2] + (E_X - c)^2 + 2 \cdot E[X - E_X] \cdot (E_X - c).$$

Der letzte Term der rechten Seite dieser Gleichung verschwindet wegen $E[X - E_X] = E_X - E_X = 0$. Der erste Term ist nicht beeinflussbar; er stellt gemäß Gl. (B-9) die Varianz der Zufallsgröße X dar. Der mittlere Term ist eine nichtnegative Konstante, die mit $c = E_X$ zu Null wird. Dies beweist die obige Feststellung.

Aus den obigen Definitionen des Erwartungswertes und der Varianz der Zufallsgröße X resultieren unmittelbar auch die beiden folgenden nützlichen Beziehungen für lineare Transformationen von X :

$$E[a \cdot X + b] = a \cdot E_X + b, \quad (\text{B-12})$$

$$\text{Var}[a \cdot X + b] = a^2 \cdot \text{Var}[X]. \quad (\text{B-13})$$

Für den Erwartungswert und die Varianz einer diskreten Zufallsgröße gelten analog zu den Gleichungen (B-8) und (B-9) die Beziehungen

$$E_X = E[X] = \sum_i x_i \cdot p_i, \quad (\text{B-14})$$

$$\text{Var}[X] = E[(X - E_X)^2] = \sum_i (x_i - E_X)^2 \cdot p_i. \quad (\text{B-15})$$

B.1.3 Multivariate Zufallsgrößen, Kovarianz, Korrelationskoeffizient

Sind für probabilistische Expositionsabschätzungen mehrere Modellparameter als Zufallsvariable definiert, so sind diese mit multivariaten Wahrscheinlichkeitsdichten bzw. Verteilungsfunktionen zu beschreiben. Man spricht dann auch von multivariaten Zufallsgrößen. Im bivariaten Fall wird die zweidimensionale Zufallsgröße (X_1, X_2) durch die gemeinsame Wahrscheinlichkeitsdichte

$$f_{1,2}(x_1, x_2) \geq 0 \quad \text{mit} \quad \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f_{1,2}(x_1, x_2) dx_1 dx_2 = 1, \quad (\text{B-16})$$

charakterisiert (zur Verkürzung der folgenden Darstellungen wird hier auf die explizite Angabe der Verteilungsparameter θ verzichtet). Die gemeinsame kumulative Verteilungsfunktion ist:

$$F_{1,2}(x_1, x_2) = \Pr(X_1 \leq x_1; X_2 \leq x_2) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} f_{1,2}(u, v) du dv. \quad (\text{B-17})$$

$F_{1,2}(x_1, x_2)$ ist eine positive, nicht fallende Funktion von x_1 und x_2 mit

$$0 \leq F_{1,2}(x_1, x_2) \leq 1 \quad \text{und} \quad \lim_{x_1 \rightarrow \infty, x_2 \rightarrow \infty} F_{1,2}(x_1, x_2) = 1. \quad (\text{B-18})$$

Aus der gemeinsamen Wahrscheinlichkeitsdichte $f_{1,2}(x_1, x_2)$ können durch Integration über x_1 bzw. x_2 die Randdichten der beiden Zufallsvariablen berechnet werden,

$$f_1(x_1) = \int_{-\infty}^{+\infty} f_{1,2}(x_1, x_2) dx_2 \quad \text{und} \quad f_2(x_2) = \int_{-\infty}^{+\infty} f_{1,2}(x_1, x_2) dx_1. \quad (\text{B-19})$$

Aus diesen beiden (univariaten) Verteilungen lassen sich mit den Gleichungen (B-8) und (B-9) die Erwartungswerte und Varianzen der beiden Zufallsvariablen X_1 und X_2 berechnen.

Eine wichtige Kenngröße bivariater Verteilungen ist die Kovarianz der beiden Zufallsvariablen:

$$\text{Cov}[X_1, X_2] = E[(X_1 - E_{X_1})(X_2 - E_{X_2})] = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x_1 - E_{X_1})(x_2 - E_{X_2}) f_{1,2}(x_1, x_2) dx_1 dx_2 \quad (\text{B-20})$$

$$\text{Cov}[X_1, X_2] = E[X_1 \cdot X_2] - E_{X_1} \cdot E_{X_2}$$

Ein deskriptives Maß für den linearen Zusammenhang der beiden Zufallsvariablen X_1 und X_2 ist der Korrelationskoeffizient,

$$\rho_{1,2} = \frac{\text{Cov}[X_1, X_2]}{\sqrt{\text{Var}[X_1] \cdot \text{Var}[X_2]}}, \quad (\text{B-21})$$

der nur Werte von -1 bis +1 annehmen kann. X_1 und X_2 heißen unkorreliert, wenn $\rho_{1,2} = 0$ ist. Dies gilt insbesondere dann, wenn die beiden Zufallsvariablen voneinander unabhängig sind.²

Die beiden Zufallsgrößen X_1 und X_2 sind voneinander unabhängig, wenn ihre gemeinsame Wahrscheinlichkeitsdichte gleich dem Produkt der beiden Randdichten ist, d.h. wenn

$$f_{1,2}(x_1, x_2) = f_1(x_1) \cdot f_2(x_2) \quad (\text{B-22})$$

ist. Für unabhängige Zufallsgrößen vereinfacht sich die Berechnung der Wahrscheinlichkeitsdichten transformierter Zufallsgrößen $Y = T(X_1, X_2)$ erheblich. Für die einfachen arithmetischen Funktionen gelten (für den allgemeinen Fall) die folgenden Berechnungsvorschriften:

$$Y = X_1 + X_2 : \quad f_Y(y) = \int_{-\infty}^{+\infty} f_{1,2}(x_1, y - x_1) dx_1, \quad (\text{B-23})$$

$$Y = X_1 - X_2 : \quad f_Y(y) = \int_{-\infty}^{+\infty} f_{1,2}(x_1, x_1 - y) dx_1, \quad (\text{B-24})$$

$$Y = X_1 \cdot X_2 : \quad f_Y(y) = \int_{-\infty}^{+\infty} \frac{1}{|x_1|} \cdot f_{1,2}(x_1, y/x_1) dx_1, \quad (\text{B-25})$$

$$Y = X_1/X_2; X_2 \neq 0 \quad f_Y(y) = \int_{-\infty}^{+\infty} |x_1| \cdot f_{1,2}(x_1 \cdot y, x_1) dx_1. \quad (\text{B-26})$$

² Allerdings darf umgekehrt aus $\rho_{1,2} = 0$ nicht auf die Unabhängigkeit der beiden Zufallsgrößen geschlossen werden, was leider oft übersehen wird.

B.1.4 Zerlegung der Varianz einer Funktion von zwei Variablen

Es sei $Z = T(X_1, X_2)$ eine Funktion der beiden (Input-)Variablen X_1 und X_2 , für die eine gemeinsame Wahrscheinlichkeitsdichte $f_{1,2}(x_1, x_2)$ vorliegt. Fixiert man jeweils eine dieser Variablen, ergeben sich daraus die bedingten Wahrscheinlichkeiten: $f_{1|2}(x_1|x_2)$ ist die bedingte Wahrscheinlichkeit von x_1 (für den festen Wert x_2) und $f_{2|1}(x_2|x_1)$ ist die bedingte Wahrscheinlichkeit von x_2 (für den festen Wert x_1). Nach dem Satz über bedingte Wahrscheinlichkeiten (siehe z. B. [RIN 97]) kann die gemeinsame Wahrscheinlichkeitsdichte der beiden Variablen $f_{1,2}(x_1, x_2)$ als Produkt einer bedingten Verteilung mit der Marginalverteilung der jeweils anderen Variablen dargestellt werden:

$$f_{1,2}(x_1, x_2) = f_{1|2}(x_1|x_2) \cdot f_2(x_2) = f_{2|1}(x_2|x_1) \cdot f_1(x_1). \quad (\text{B-27})$$

Nachfolgend wird der linke Teil von Gl. (B-27) betrachtet (die beiden Indizes 1 und 2 können aber ohne Verlust der Gültigkeit der Ergebnisse vertauscht werden).

Der Erwartungswert von Z $E_Z = E[Z]$, $z = T(x_1, x_2)$, kann mit Gl. (B-27) wie folgt dargestellt werden:

$$E_Z = \iint z \cdot f_{1,2}(x_1, x_2) dx_1 dx_2 = \int \left(\int z \cdot f_{1|2}(x_1 | x_2) dx_1 \right) \cdot f_2(x_2) dx_2 \quad (\text{B-28a})$$

$$E_Z = E_{X_2} [E_{Z|X_2}] \text{ mit} \quad (\text{B-28b})$$

$$E_{Z|X_2} = \int z \cdot f_{1|2}(x_1 | x_2) dx_1. \quad (\text{B-29})$$

$E_{Z|X_2}$ ist der bedingte Erwartungswert von Z bzgl. X_2 . Die bedingte Varianz bzgl. X_2 ist

$$\text{Var}_{Z|X_2} = \int (z - E_{Z|X_2})^2 \cdot f_{1|2}(x_1 | x_2) dx_1. \quad (\text{B-30})$$

Zur Berechnung/Zerlegung der Varianz von Z , Var_Z wird die "produktive Null" $\pm E_{Z|X_2}$ verwendet:

$$\begin{aligned} \text{Var}_Z &= \iint (z - E_Z)^2 \cdot f_{1,2}(x_1, x_2) dx_1 dx_2 = \int \left(\int (z - E_Z)^2 \cdot f_{1|2}(x_1 | x_2) dx_1 \right) \cdot f_2(x_2) dx_2 \\ \text{Var}_Z &= \int \left(\int (z - E_{Z|X_2})^2 \cdot f_{1|2}(x_1 | x_2) dx_1 \right) \cdot f_2(x_2) dx_2 \\ &\quad + \int \left(\int (E_{Z|X_2} - E_Z)^2 \cdot f_{1|2}(x_1 | x_2) dx_1 \right) \cdot f_2(x_2) dx_2 + \\ &\quad - 2 \cdot \int (E_{Z|X_2} - E_Z) \cdot \left(\int (z - E_{Z|X_2}) \cdot f_{1|2}(x_1 | x_2) dx_1 \right) \cdot f_2(x_2) dx_2 \end{aligned} \quad (\text{B-31})$$

In Gl. (B-31) ist der 1. Summand der Erwartungswert (äußeres Integral über x_2) der bzgl. X_2 bedingten Varianz von Z , $\text{Var}_{Z|X_2}$ (siehe Gl. (B-30)). Im 2. Summanden hängt $(E_{Z|X_2} - E_Z)^2$ nicht von x_1 ab und kann vor das innere Integral gezogen werden, das dann den Wert 1 ergibt. Das noch verbleibende Integral über x_2 ergibt die Varianz des bedingten Erwartungswertes $E_{Z|X_2}$. Der letzte Term von Gl. (B-31) verschwindet, weil das innere Integral gemäß Gl. (B-29) eine Null ergibt. Damit resultiert für die Varianz von Z die folgende Zerlegung:

$$\text{Var}_Z = E_{X_2} [\text{Var}_{Z|X_2}] + \text{Var}_{X_2} [E_{Z|X_2}]. \quad (\text{B-32a})$$

Ist $\{X\}_m$ eine Auswahl von m aus n Inputvariablen eines Modells $Z = T(X_1, X_2, \dots, X_n)$, dann kann analog zu Gl. (B-32a) die folgende Varianzzerlegung abgeleitet werden:

$$\text{Var}_Z = E_{\{X\}_m} [\text{Var}_{Z|\{X\}_m}] + \text{Var}_{\{X\}_m} [E_{Z|\{X\}_m}]. \quad (\text{B-32b})$$

In der Gl. (B-32b) beschreibt der zweite Summand die durch $\{X\}_m$ erklärte Varianz von Z und der erste Summand die Restvarianz.

B.2 Punktschätzer

Es sei X eine kontinuierliche Zufallsgröße, die durch ihre Wahrscheinlichkeitsdichte $f_X(x|\theta)$ charakterisiert sei und für die eine Stichprobe $\{X\}_n = (X_1, X_2, \dots, X_n)$ vom Umfang n gewonnen werden kann, auf deren Basis bestimmte Kenngrößen von X (Mittelwert, Varianz, etc.) geschätzt werden sollen.

Die n Zufallsvariablen X_1, X_2, \dots, X_n bilden eine zufällige Stichprobe, wenn die X_i gegenseitig unabhängig sind und jeweils durch die gleiche Wahrscheinlichkeitsdichte $f_X(x|\theta)$ charakterisiert sind. Sie sind dann unabhängig und gleichartig verteilt (kurz: iid = independent and identically distributed). Damit eine Stichprobe $\{X\}_n$ n iid Zufallsgrößen liefert, ist eine entsprechende Planung und Durchführung der Stichprobennahme erforderlich (vgl. [GIL 87]).

Analog zu dem in Gl. (B-22) betrachteten Fall von zwei unabhängigen Zufallsgrößen ist die gemeinsame Wahrscheinlichkeitsdichte der n iid Proben dann gegeben durch

$$f_{\{X\}_n} = f_X(x_1 | \theta) \cdot f_X(x_2 | \theta) \cdot \dots \cdot f_X(x_n | \theta) = \prod_{i=1}^n f_X(x_i | \theta). \quad (\text{B-33})$$

Ein Punktschätzer ist, allgemein gesprochen, eine definierte Stichprobenfunktion/Schätzfunktion

$$W = \Psi(X_1, \dots, X_n). \quad (\text{B-34a})$$

Der Wert w , den diese Funktion für eine Realisation (x_1, \dots, x_n) von n Stichprobenwerten der Zufallsvariablen X annimmt, wird als Punktschätzung oder Schätzwert bezeichnet,

$$w = \Psi(x_1, \dots, x_n). \quad (\text{B-34b})$$

B.2.1 Stichprobenmittelwert, Stichprobenvarianz, Stichprobenkorrelationskoeffizient

Da eine zufällige Stichprobe (X_1, X_2, \dots, X_n) aus der Grundgesamtheit der Zufallsgröße X n iid Zufallsgrößen bildet, sind die Wahrscheinlichkeiten der einzelnen Realisierungen x_i gleich, so dass in Analogie zu Gl. (B-14) mit $p_i = 1/n$ das arithmetische Mittel der Stichprobenwerte

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad (\text{B-35b})$$

einen Schätzwert des Erwartungswerts $E[X]$ der Grundgesamtheit, aus der die Stichprobe gewonnen wurde, darstellt. Es sei vermerkt, dass die Schätzfunktion (kurz: der Schätzer)

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \quad (\text{B-35a})$$

selbst wieder eine Zufallsgröße ist. Statistische Eigenschaften dieser Zufallsgröße \bar{X} resultieren aus folgender Feststellung (siehe z. B. [CAS 02]): Ist (X_1, X_2, \dots, X_n) eine iid Stichprobe aus der

Grundgesamtheit der Zufallsgröße X und $g(X)$ eine Funktion, für die der Erwartungswert $E[g(X)]$ und die Varianz $\text{Var}[g(X)]$ existieren, dann ist

$$E\left[\sum_{i=1}^n g(X_i)\right] = n \cdot E[g(X)], \quad (\text{B-36})$$

$$\text{Var}\left[\sum_{i=1}^n g(X_i)\right] = n \cdot \text{Var}[g(X)]. \quad (\text{B-37})$$

Aus Gl. (B-36) und Gl. (B-37) folgen für den Erwartungswert und die Varianz des Stichprobenmittelwertes \bar{X} (für $g(X) = X/n$) unter Beachtung von Gl. (B-12) und Gl. (B-13) die Beziehungen

$$E[\bar{X}] = E[X], \quad (\text{B-38})$$

$$\text{Var}[\bar{X}] = \frac{\text{Var}[X]}{n}. \quad (\text{B-39})$$

Der arithmetische Mittelwert der Stichprobenwerte ist also ein erwartungstreuer Schätzer für den Erwartungswert der Grundgesamtheit, aus der die Stichprobe gewonnen wurde, wobei die Varianz des Stichprobenmittelwertes \bar{X} proportional zur Varianz der Grundgesamtheit ist und umgekehrt proportional mit dem Stichprobenumfang fällt. Dies ist eine zentrale Aussage zu den Möglichkeiten einer Reduzierung von Stichprobenunsicherheiten durch Erhöhung des Stichprobenumfangs.

Die Stichprobenvarianz S^2 wird in Analogie zu Gl. (B-15) durch folgenden Schätzer definiert,

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2. \quad (\text{B-40a})$$

Für eine konkrete Realisierung (x_1, x_2, \dots, x_n) ergibt sich damit zur Varianz der Grundgesamtheit der Zufallsgröße X der Schätzwert

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \quad (\text{B-40b})$$

mit \bar{x} gemäß Gl. (B-35b).

Anmerkung: Abweichend vom arithmetischen Mittel der quadratischen Abweichungen $(X_i - \bar{X})^2$ wird üblich für die Stichprobenvarianz als Faktor vor der Summe nicht mit den Wahrscheinlichkeiten $1/n$, sondern den etwas größeren Werten $1/(n-1)$ gerechnet. Dadurch wird erreicht, dass S^2 ein erwartungstreuer Schätzer der Varianz der Grundgesamtheit der Zufallsgröße X ist³.

³ Aus Gl. (B-40a) folgt unter Beachtung der Gleichungen (B-12), (B-36), (B-38) und (B-39):

$$\begin{aligned} E[S^2] &= \frac{1}{n-1} E\left[\sum_{i=1}^n (X_i^2 - 2 \cdot X_i \cdot \bar{X} + \bar{X}^2)\right] = \frac{1}{n-1} E\left[\sum_{i=1}^n X_i^2 - n \cdot \bar{X}^2\right] = \frac{1}{n-1} (n \cdot E[X^2] - n \cdot E[\bar{X}^2]) = \\ &= \frac{1}{n-1} (n \cdot (\text{Var}[X] + E[X]^2) - n \cdot (\text{Var}[\bar{X}] + E[\bar{X}]^2)) = \frac{1}{n-1} (n \cdot (\text{Var}[X] + E[X]^2) - \text{Var}[X] - n \cdot E[X]^2) = \text{Var}[X] \end{aligned}$$

B.2.2 Schiefe und Wölbung

Zwei wichtige Kenngrößen, mit denen die Form von statistischen Verteilungen charakterisiert wird, sind die Schiefe und die Wölbung. Stichprobenschätzer dieser beiden Kenngrößen können für die Auswahl von Familien parametrischer Verteilungen herangezogen werden, die zur Modellierung bzw. Approximation von empirischen Verteilungen in Betracht gezogen werden. Beide Kenngrößen werden durch Relationen des 3. bzw. 4. zentralen Momentes der Verteilung zu entsprechenden Potenzen der Standardabweichung σ normiert.

Das für die Schiefe der Verteilung einer Zufallsgröße X am häufigsten verwendete Funktional ist definiert durch:

$$\gamma_1 = E \left[\left(\frac{X - \mu}{\sigma} \right)^3 \right], \quad (\text{B-41})$$

wobei μ den Erwartungswert von X und σ dessen Standardabweichung bezeichnen. Verteilungen mit positiver Schiefe bezeichnet man als "linkssteil" bzw. als "rechtsschief". Bei ihnen treten Werte, die kleiner als der Erwartungswert sind, häufiger auf (der Modalwert liegt unter dem Median, und dieser ist kleiner als der Erwartungswert, wie z. B. bei der Lognormalverteilung). Verteilungen mit einer negativen Schiefe bezeichnet man entsprechend als "rechtssteil" oder "linksschief". Die Schiefe von symmetrischen Verteilungen ist gleich Null. Für eine Stichprobe (x_1, x_2, \dots, x_n) aus n Einzelwerten mit dem Mittelwert \bar{x} gemäß Gl. (B-35b) ergibt sich für die mit Gl. (B-41) definierte (auf den Erwartungswert bezogene) Schiefe der folgende Schätzwert g_1 :

$$g_1 = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^3}{\left(\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \right)^{3/2}}, \quad (\text{B-42a})$$

der jedoch nicht erwartungstreu ist. Ein erwartungstreuer Schätzer der Schiefe ist:

$$g'_1 = \frac{\sqrt{n \cdot (n-1)}}{n-2} \cdot g_1 = \frac{n}{(n-1) \cdot (n-2)} \cdot \sum_{i=1}^n \left(\frac{x_i - \bar{x}}{s} \right)^3, \quad (\text{B-42b})$$

wobei s die Stichproben-Standardabweichung bezeichnet, also die Wurzel der Stichprobenvarianz nach Gl. (B-40b).

Es sei vermerkt, dass es weitere Schiefemaße gibt, die wesentliche Anforderungen an ein solches Maß erfüllen und z. T. einfacher zu berechnen sind, wie z. B. das in [ARN 95] definierte Maß γ_M , das auf den Modalwert Mod_X bezogene ist:

$$\gamma_M = 1 - 2 \cdot F_X(\text{Mod}_X). \quad (\text{B-43})$$

Die Kurtosis (Wölbung) der Verteilung einer Zufallsgröße X ist definiert durch:

$$\gamma_2 = E \left[\left(\frac{X - \mu}{\sigma} \right)^4 \right]. \quad (\text{B-44})$$

Zur quantitativen Einschätzung der Wölbung einer Wahrscheinlichkeitsdichte wird sie häufig mit der Wölbung der Normalverteilung verglichen, deren Wert 3 beträgt. Die Größe

$$\gamma_{2,\text{Exzess}} = \gamma_2 - 3 \quad (\text{B-45})$$

wird als "Exzesskurtosis" bzw. auch als "zentriertes Wölbungsmaß" bezeichnet. Verteilungen mit einer Wölbung unter 3 heißen platykurtisch (flach gewölbt), Verteilungen mit einer Wölbung über 3 heißen leptokurtisch (stark gewölbt). Ein erwartungstreuer Schätzer der Exzesskurtosis ist:

$$g'_{2,\text{Exzess}} = \frac{n \cdot (n+1)}{(n-1) \cdot (n-2) \cdot (n-3)} \cdot \sum_{i=1}^n \left(\frac{x_i - \bar{x}}{s} \right)^4 - \frac{3 \cdot (n-1)^2}{(n-2) \cdot (n-3)}, \quad (\text{B-46})$$

wobei s wieder die Stichproben-Standardabweichung bezeichnet.

B.2.3 Abhängigkeitsmaße

Der Schätzer für den Korrelationskoeffizienten $\rho_{X,Y}$ aus Gl. (B-21) bei einem Stichprobenumfang für X und Y von n ist der PEARSON'sche Korrelationskoeffizient:

$$r_{x,y} = \frac{\sum_{i=1}^n (x_i - \bar{x}) \cdot (y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \cdot \sum_{i=1}^n (y_i - \bar{y})^2}} = \frac{\sum_{i=1}^n (x_i - \bar{x}) \cdot (y_i - \bar{y})}{(n-1) \cdot s_x \cdot s_y}, \quad (\text{B-47})$$

wobei s_x und s_y die nach Gl. (B-40b) geschätzten Standardabweichungen der beiden Variablen sind.

Zur Beschreibung der statistischen Abhängigkeit zwischen zwei Zufallsgrößen X und Y , die durch nichtlineare, jedoch monotonen Funktionen miteinander verknüpft sind, kann der SPEARMAN'sche Rang-Korrelationskoeffizient $\rho_{x,y}$ benutzt werden, der die Korrelation der Ränge $rg(x_i)$ und $rg(y_i)$ der Stichprobenwerte bestimmt ($rg(x_i) = k$, wenn x_i der k -größte Wert der Stichprobe ist; k von 1 bis n). Seine Berechnung erfolgt analog zu Gl. (B-47), nur dass statt der Werte x_i und y_i ihre Ränge $rg(x_i)$ und $rg(y_i)$ verwendet werden. Sind alle Stichprobenwerte ungleich (damit auch ihre Ränge), so vereinfacht sich die Berechnung des SPEARMAN'schen Rang-Korrelationskoeffizienten erheblich:

$$\rho_{x,y} = 1 - \frac{6 \cdot \sum_{i=1}^n d_i^2}{n \cdot (n^2 - 1)} \quad \text{mit} \quad d_i = rg(x_i) - rg(y_i). \quad (\text{B-48})$$

Anmerkung: Ist $Y = g(X)$, wobei g eine nichtlineare, monotone Funktion ist, so liegt eine perfekte Abhängigkeit zwischen beiden Variablen vor, der PEARSON'sche Korrelationskoeffizient $r_{x,y}$ nimmt aber im Unterschied zum SPEARMAN'schen Korrelationskoeffizienten $\rho_{x,y}$ nicht die Werte 1 (g monoton steigend) bzw. -1 (g monoton fallend) an.

Ein weiteres Maß zur Quantifizierung der Abhängigkeit zwischen zwei Zufallsgrößen X und Y ist der KENDALL'sche Konkordanzkoeffizient $\tau_{X,Y}$ (wird auch KENDALL-Rang-Korrelationskoeffizient genannt). Zwei Paare (x_i, y_i) und (x_j, y_j) zufälliger Realisierungen von (X, Y) heißen

konkordant, wenn $(x_i - x_j) \cdot (y_i - y_j) > 0$, (B-49)

diskordant, wenn $(x_i - x_j) \cdot (y_i - y_j) < 0$.

Das in Gleichung (B-49) stehende Produkt kann natürlich auch gleich Null sein, wenn $x_i = x_j$ oder $y_i = y_j$ oder beides zutrifft. In diesem Fall spricht man von Bindungen. Wir betrachten zunächst den einfachen Fall, dass in einer Stichprobe (x, y) mit dem Umfang n weder für X noch für Y Bindungen vorliegen. Insgesamt liegen $n \cdot (n-1)/2$ Paare (x_i, y_i) und (x_j, y_j) von Beobachtungen vor. Ist nun n_c die Anzahl der konkordanten Paare und n_d die Anzahl der diskordanten Paare, dann ist

$$\tau_{x,y} = \frac{2 \cdot (n_c - n_d)}{n \cdot (n-1)} = \frac{n_c - n_d}{n_c + n_d} \quad (\text{B-50a})$$

ein Maß für die Abhängigkeit zwischen X und Y , das Werte zwischen -1 (perfekt gegenläufig) und $+1$ (perfekt gleichläufig) annimmt, wobei im Fall von unabhängigen Zufallsgrößen X und Y ein Wert von Null (bzw. ≈ 0) angenommen wird. Wie der Spearman'sche Rangkorrelationskoeffizient ρ ist auch der KENDALL'sche Konkordanzkoeffizient τ unabhängig gegenüber monoton steigenden Transformationen bzgl. X und/oder Y .

Liegen in der Stichprobe für die Variable X n_x und für die Variable Y n_y Bindungen vor, so ist eine Modifizierung von Gl. (50-a) erforderlich, mit der diese Bindungen berücksichtigt werden. Hierzu gibt es verschiedene Möglichkeiten. Im Rahmen des vorliegenden Berichts wird die Beziehung

$$\tau_{x,y} = \frac{n_c - n_d}{\sqrt{(n_c + n_d + n_x) \cdot (n_c + n_d + n_y)}} \quad (\text{B-50b})$$

verwendet, die als Verallgemeinerung von Gl. (B-50a) sehr gebräuchlich ist.

Zwischen den drei betrachteten Abhängigkeitsmaßen gibt es verschiedene Zusammenhänge, wozu insbesondere auf [KRU 58] und [SCHW 81] verwiesen sei.

B.3 Copulas

Mit der auf dem Satz von SKLAR [SKL 59]/[SKL 73] basierenden Copula-Theorie kann zur Analyse und Simulation von multivariaten Verteilungen, die eine gemeinsame Dichte $f_X(x_1, x_2, \dots, x_d)$ bzw. Verteilungsfunktion $F_X(x_1, x_2, \dots, x_d)$ aufweisen, hinsichtlich der Struktur von Abhängigkeiten zwischen den d Zufallsgrößen X_i (i von 1 bis d) von ihren Randverteilungen abstrahiert werden. Zur Vereinfachung der Darstellungen beschränken sich die nachfolgenden Darstellungen überwiegend auf den Fall von zwei Zufallsgrößen ($d = 2$). Zunächst werden in Abschnitt B.3.1 grundlegende Begriffe und Aussagen der Copula-Theorie zusammengefasst. Der Abschnitt B.3.2 enthält Angaben zu ausgewählten Klassen bzw. Familien von Copulas. Auf Methoden zur Simulation voneinander abhängiger Zufallsgrößen mit Hilfe von Copulas wird an entsprechender Stelle eingegangen. Die Berechnung der empirischen Copula-Funktion wird in Abschnitt B.3.3 erläutert.

B.3.1 Grundlegende Begriffe und Aussagen der Copula-Theorie

Definition: Eine d-dimensionale Copula ist eine Funktion $C(\mathbf{u}) = C(u_1, u_2, \dots, u_d)$, die den d-dimensionalen Einheitskubus $I^d = [0, 1]^d$ auf das Einheitsintervall $I = [0, 1]$ abbildet, d. h. $0 \leq u_i \leq 1$ für i von 1 bis d und $0 \leq C(\mathbf{u}) \leq 1$, wenn die folgenden drei Bedingungen erfüllt sind:

$C(\mathbf{u})$ ist gleich Null, wenn eine Variable den Wert Null hat:

$$C(u_1, u_2, \dots, u_{i-1}, 0, u_{i+1}, \dots, u_d) = 0 \text{ für beliebige } u_i = 0; \quad (\text{B-51a})$$

- die Randverteilung jeder der d Variablen ist eine Gleichverteilung:

$$C(1, 1, \dots, 1, u_i, 1, 1, \dots, 1) = u_i \text{ für jede Variable } u_i; \quad (\text{B-51b})$$

- $C(\mathbf{u})$ ist in allen d Variablen wachsend (d-wachsend), d. h. mit $a_i \leq b_i$ für alle i von 1 bis d und den Bezeichnungen $u_{i,1} = a_i$ und $u_{i,2} = b_i$ gilt:

$$\sum_{j_1=1}^2 \dots \sum_{j_d=1}^2 (-1)^{j_1+\dots+j_d} C(u_{1,j_1}, \dots, u_{d,j_d}) \geq 0. \quad (\text{B-51c})$$

Die Funktion $C(\mathbf{u})$ ist dann auch in jeder einzelnen Variable wachsend:

$$C(u_1, u_2, \dots, u_{i-1}, a_i, u_{i+1}, \dots, u_d) \leq C(u_1, u_2, \dots, u_{i-1}, b_i, u_{i+1}, \dots, u_d) \text{ für } a_i \leq b_i. \quad (\text{B-51d})$$

Fall $d = 2$: Insbesondere Gl. (B-51c) ist für beliebige Dimensionen d nicht mehr sehr anschaulich. Deshalb werden auch die im Fall $d = 2$ resultierenden Gleichungen angegeben, wobei es üblich ist, für eine von nur 2 Variablen abhängige Copula die Variablen mit u (für u_1) und v (für u_2) zu bezeichnen. Die Definitionsgleichungen (B-51a) bis (B-51c) vereinfachen sich damit wie folgt:

$$C(u, 0) = 0 \text{ und } C(0, v) = 0, \quad (\text{B-52a})$$

$$C(u, 1) = u \text{ und } C(1, v) = v, \quad (\text{B-52b})$$

$$C(u_2, v_2) - C(u_2, v_1) - C(u_1, v_2) + C(u_1, v_1) \geq 0 \text{ für } u_1 \leq u_2 \text{ und } v_1 \leq v_2. \quad (\text{B-52c})$$

Mit einer Copula $C(u, v)$ wird jedem Rechteck $[0, u] \cdot [0, v]$ aus I^2 ein nicht-negativer Wert aus dem Einheitsintervall I zugeordnet, der mit beiden Variablen wächst. Die Gl. (B-52c) sagt dann aus, dass auch jedem Rechteck $[u_1, u_2] \cdot [v_1, v_2]$ ($u_1 \leq u_2$ und $v_1 \leq v_2$) ein nicht-negativer Wert zugeordnet wird.

Satz von SKLAR: Sind $F(x_1, x_2, \dots, x_d)$ die gemeinsame multivariate Verteilungsfunktion von d Zufallsgrößen X_1 bis X_d und $F_i(x_i)$ ($i = 1, \dots, d$) die jeweiligen Randverteilungen, dann existiert eine Copula C , für die

$$F(x_1, x_2, \dots, x_d) = C(F_1(x_1), F_2(x_2), \dots, F_d(x_d)) \quad (\text{B-53})$$

gilt. Wenn alle d Randverteilungen $F_i(x_i)$ stetig sind, so ist die Copula C eindeutig. Umgekehrt wird durch eine Copula $C(u_1, u_2, \dots, u_d)$ und d Randverteilungen $F_i(x_i)$ ($i = 1, \dots, d$) durch die Gleichung (B-53) eine gemeinsame Verteilungsfunktion für die d Zufallsgrößen definiert, welche diese Randverteilungen aufweist.

Per Definition stellt eine Copula $C(u_1, u_2, \dots, u_d)$ die gemeinsame kumulative Verteilungsfunktion von d auf dem Einheitsintervall $I = [0, 1]$ gleichverteilten Zufallsgrößen dar. Nach dem Satz von SKLAR wird die Struktur der Abhängigkeiten zwischen d stetigen Zufallsgrößen X_i , welche die multivariate Verteilungsfunktion $F(x_1, x_2, \dots, x_d)$ haben, durch eine Copula dargestellt, in der die gleichverteilten Größen u_i den Randverteilungen $F_i(x_i)$ entsprechen.

Im Fall $d = 2$, auf den sich die weiteren Erläuterungen konzentrieren, werden für die Zufallsgrößen die Bezeichnungen X und Y , für ihre Randverteilungen die Bezeichnungen $F(x)$ und $G(y)$ sowie für ihre gemeinsame Verteilungsfunktion die Bezeichnung $H(x, y)$ verwendet. Gl. (B-53) lautet damit:

$$H(x, y) = C(F(x), G(y)). \quad (\text{B-54a})$$

Umgekehrt gilt mit den Quantilen (Inversen) der beiden Randverteilungen $x = F^{-1}(u)$ und $y = G^{-1}(v)$:

$$C(u, v) = H(F^{-1}(u), G^{-1}(v)). \quad (\text{B-54b})$$

Die Gleichung (B-54b) beschreibt auch den Weg zur Berechnung von empirischen Copulas auf der Basis von Stichproben für Wertepaare (x_i, y_i) mit i von 1 bis n .

Copulas für monoton transformierte Zufallsvariable: Es sei $C_{X,Y}(u, v)$ gemäß Gl. (B-54) die Copula für die Zufallsvariablen X und Y . Für stetig transformierte Zufallsvariable $\alpha(X)$ und $\beta(X)$ lassen sich die Copulas $C_{\alpha(X),\beta(Y)}(u, v)$ auf die Copula $C_{X,Y}(u, v)$ zurückführen, wenn die Transformationen α und β strikt monoton (s. m.) steigende oder fallende Funktionen sind:

$$C_{\alpha(X),\beta(Y)}(u, v) = C_{X,Y}(u, v), \quad \text{wenn } \alpha(X) \text{ und } \beta(X) \text{ s. m. steigen;} \quad (\text{B-55a})$$

$$C_{\alpha(X),\beta(Y)}(u, v) = u - C_{X,Y}(u, 1-v), \quad \text{wenn } \alpha(X) \text{ s. m. steigt und } \beta(X) \text{ s. m. fällt;} \quad (\text{B-55b})$$

$$C_{\alpha(X),\beta(Y)}(u, v) = v - C_{X,Y}(1-u, v), \quad \text{wenn } \alpha(X) \text{ s. m. fällt und } \beta(X) \text{ s. m. steigt;} \quad (\text{B-55c})$$

$$C_{\alpha(X),\beta(Y)}(u, v) = u + v - 1 + C_{X,Y}(1-u, 1-v), \quad \text{wenn } \alpha(X) \text{ und } \beta(X) \text{ s. m. fallen.} \quad (\text{B-55d})$$

Copula Π für unabhängige Zufallsgrößen: Die Zufallsgrößen X und Y sind genau dann voneinander unabhängig, wenn $H(x, y) = F(x) \cdot G(y)$ gilt (entspricht Gl. (B-22)). Aus Gl. (B-54b) folgt damit die Copula für unabhängige Zufallsgrößen, die auch Produkt-Copula genannt und mit dem Buchstaben Π bezeichnet wird:

$$\Pi(u, v) = F(F^{-1}(u)) \cdot G(G^{-1}(v)) = u \cdot v. \quad (\text{B-56})$$

FRÉCHET-HOEFFDING-Schranken: Alle Copulas werden durch eine obere und eine untere Schranke begrenzt, die einer perfekten Co- bzw. Contra-Monotonie der Abhängigkeit von Zufallsvariablen entsprechen (vollständige Konkordanz bzw. Diskordanz, vgl. Abschnitt B.2.3). Ist $Y = \gamma(X)$ und γ s. m. steigend (X und Y sind perfekt konkordant), dann folgt mit Gl. (B-55a):

$$C_{X,Y}(u, v) = C_{X,\gamma(X)}(u, v) = C_{X,X}(u, v) = \Pr(F(X) \leq u, F(X) \leq v) = \min(u, v).$$

Diese obere Schranke für den Fall der perfekten Co-Monotonie wird üblich mit $M(u, v)$ bezeichnet.

Ist die Transformation $Y = \gamma(X)$ dagegen s. m. fallend (X und Y sind perfekt diskordant), so folgt mit Gl. (B-55b) für diesen Fall der Contra-Monotonie:

$$C_{X,Y}(u, v) = C_{X,Y(X)}(u, v) = u - C_{X,X}(u, 1-v) = u - \Pr(F(X) \leq u, F(X) \leq 1 - v) = u - \min(u, 1-v) = \max(u+v-1, 0).$$

Diese untere Schranke wird üblich mit $W(u, v)$ bezeichnet. Damit gelten für $d = 2$ die Schranken:

$$W(u, v) \leq C(u, v) \leq M(u, v) \quad \text{mit} \quad (\text{B-57a})$$

$$W(u, v) = \max(u+v-1, 0), \quad (\text{B-57b})$$

$$M(u, v) = \min(u, v). \quad (\text{B-57c})$$

Man kann zeigen, dass die mit Gleichung (B-57b) und (B-57c) definierten unteren und oberen Schranken ebenfalls Copulas sind. Im allgemeinen Fall von d Zufallsvariablen, gelten für die FRÉCHET-HOEFFDING-Schranken die folgenden Beziehungen:

$$W(u_1, u_2, \dots, u_d) \leq C(u_1, u_2, \dots, u_d) \leq M(u_1, u_2, \dots, u_d) \quad \text{mit} \quad (\text{B-58a})$$

$$W(u_1, u_2, \dots, u_d) = \max \left\{ \sum_{i=1}^d u_i + 1 - d, 0 \right\}, \quad (\text{B-58b})$$

$$M(u_1, u_2, \dots, u_d) = \min \{ u_1, u_2, \dots, u_d \}. \quad (\text{B-58c})$$

Während die obere Schranke $M(\mathbf{u})$ für beliebige Dimensionen d eine Copula ist, trifft dies bzgl. der unteren Schranke $W(\mathbf{u})$ nur im Fall $d = 2$ zu.

Die Copulas $W(u, v)$, $\Pi(u, v)$ und $M(u, v)$ sind in Abbildung B-1 dargestellt. Sie sind Grenzfälle von diversen Copula-Familien.

Copula-Dichte: Da Copulas Verteilungsfunktionen von gleichverteilten Zufallsgrößen sind, können auch ihre Dichten betrachtet werden, die (hier für den Fall $d = 2$ betrachtet) durch

$$c(u, v) = \frac{\partial^2}{\partial u \partial v} C(u, v) \quad (\text{B-59})$$

definiert sind. Copulas $C(u, v)$ können einen absolut stetigen Anteil $A_C(u, v)$ und einen singulären Anteil $S_C(u, v)$ enthalten,

$$C(u, v) = A_C(u, v) + S_C(u, v) \quad \text{mit} \quad A_C(u, v) = \int_0^u \int_0^v c(s, t) \, ds \, dt, \quad (\text{B-60})$$

wobei im Einzelfall auch nur einer dieser beiden Anteile vorhanden sein kann. Die Produkt-Copula $\Pi(u, v)$ ist z. B. absolut stetig; ihre Dichte ist gleich 1. Die FRÉCHET-HOEFFDING-Schranken $W(u, v)$ und $M(u, v)$ sind dagegen rein singuläre Copulas, wobei der Träger von $W(u, v)$ die Diagonale $v = 1 - u$ ist (entspricht perfekter Antikorrelation/Diskordanz), wogegen der Träger von $M(u, v)$ die Diagonale $v = u$ ist (entspricht perfekter Korrelation/Konkordanz). Zur anschaulichen Darstellung von Copula-Dichten sind Scatterplots für eine größere Anzahl von zufälligen Realisierungen der Wertepaare (u_i, v_i) (mit i von 1 bis n , n ist Stichprobenumfang) geeignet (siehe Abschnitt B.3.2).

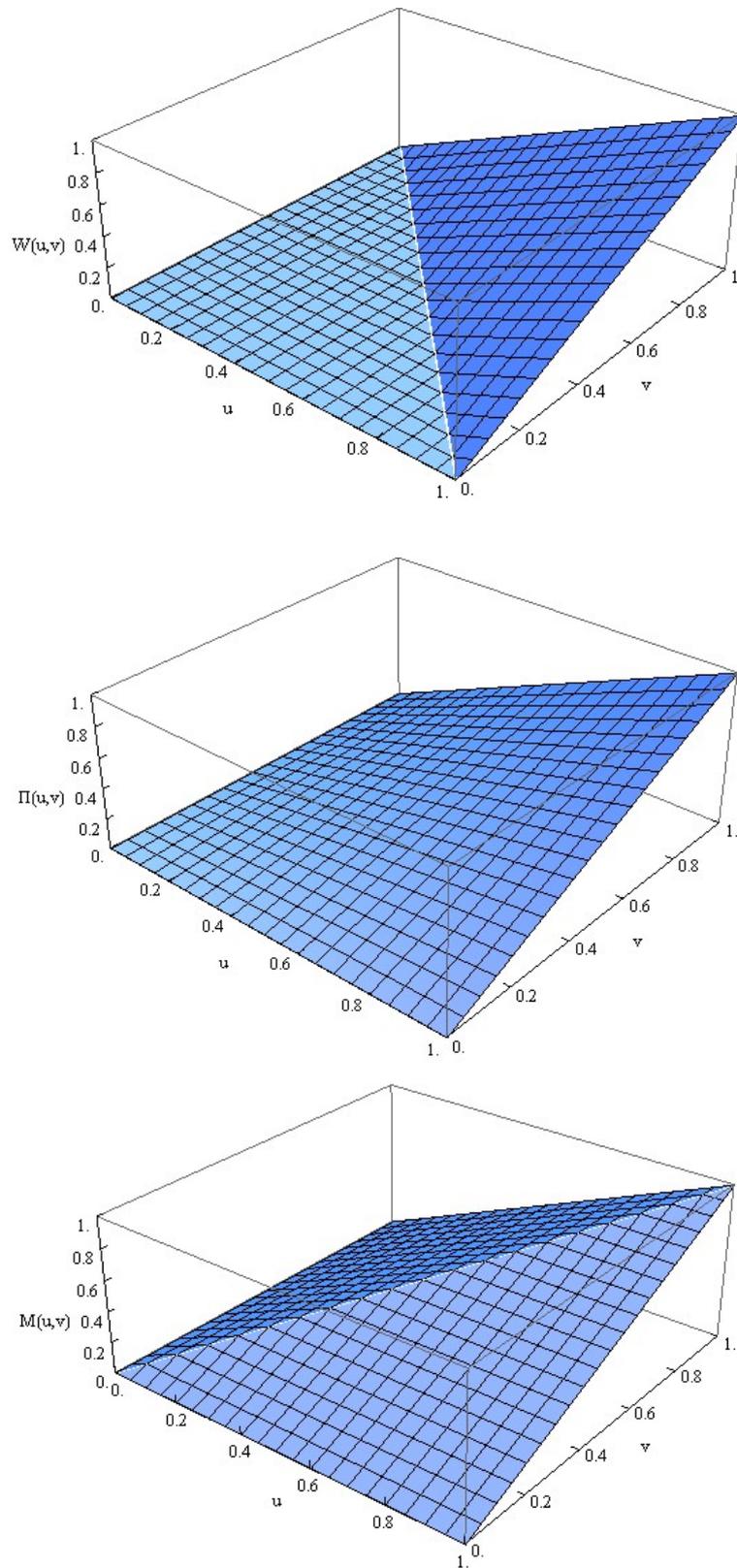


Abbildung B-1: Copulas $W(u, v)$, $\Pi(u, v)$ und $M(u, v)$ gemäß Gl. (B-56) und (B-57b/c)

Simulation aus Copula-Verteilungen: Die Simulation voneinander abhängiger Zufallsgrößen kann mit generell anwendbaren sowie mit Hilfe von speziellen Algorithmen erfolgen (siehe [DEV 86], [NEL 07], [BAL 09]). Für Copulas basiert die allgemeine Methode auf der üblichen Darstellung von Verteilungen als Produkt einer bedingten Wahrscheinlichkeit und einer Randverteilung. Für eine Copula $C(u, v)$ kann die bzgl. einer der beiden Variablen bedingte Wahrscheinlichkeit über eine partielle Ableitung dargestellt werden:

$$\Pr(V \leq v \mid U = u) = \lim_{\Delta u \rightarrow 0} \frac{C(u + \Delta u, v) - C(u, v)}{\Delta u} = \frac{\partial C(u, v)}{\partial u}. \quad (\text{B-61})$$

Die Variable $Z = C_u(v) = \partial C(u, v)/\partial u$ ist wieder eine Verteilungsfunktion und dementsprechend auf dem Intervall $[0, 1]$ gleichverteilt. Damit resultiert zur Erzeugung zufälliger Realisierungen (u, v) aus der Copula $C(u, v)$ ein genereller Algorithmus:

1. Es werden zwei voneinander unabhängige Realisierungen u und z von auf dem Einheitsintervall $[0, 1]$ gleichverteilten Zufallszahlen generiert.
2. Mit der inversen Verteilungsfunktion C_u^{-1} der Größe Z (s. o.) wird $v = C_u^{-1}(z)$ gesetzt.
3. Das Wertepaar (u, v) ist eine zufällige Realisierung der Copula $C(u, v)$.

Diese Methode verlangt eine analytische Darstellung der inversen Verteilungsfunktion C_u^{-1} . Für Zufallsgrößen X und Y mit den Randverteilungen $F(x)$ und $G(y)$ können dann die abhängigen Wertepaare (x, y) gemäß Gl. (B-54) über $x = F^{-1}(u)$ und $y = G^{-1}(v)$ berechnet werden.

Abhängigkeitsmaße: Zur Quantifizierung des durch eine Copula $C(u, v)$ bestimmten Maßes der Abhängigkeit zwischen den beiden Zufallsvariablen U und V werden als einfache globale Maße der SPEARMAN'sche Rang-Korrelationskoeffizient ρ und der KENDALL'sche Konkordanz-Koeffizient τ verwendet. Sie sind invariant gegenüber streng monoton wachsenden Transformationen und können Werte von -1 bis $+1$ annehmen. ρ und τ hängen zwar zusammen, die Beziehung zwischen diesen beiden Kennwerten ist jedoch spezifisch für Copula-Familien (siehe u. a. [SCHW 81], [FRED 07], [NEL 07]). Für gegebene Populationen mit stetig differenzierbaren Copulas $C(u, v)$ können ρ und τ über folgende Integrale berechnet werden:

$$\rho = 12 \cdot \int_0^1 \int_0^1 C(u, v) \, du \, dv - 3 = 3 - 6 \cdot \int_0^1 \int_0^1 \left(u \cdot \frac{\partial C(u, v)}{\partial u} + v \cdot \frac{\partial C(u, v)}{\partial v} \right) \, du \, dv, \quad (\text{B-62})$$

$$\tau = 4 \cdot \int_0^1 \int_0^1 C(u, v) \frac{\partial^2 C(u, v)}{\partial u \partial v} \, du \, dv - 1 = 1 - 4 \cdot \int_0^1 \int_0^1 \frac{\partial C(u, v)}{\partial u} \cdot \frac{\partial C(u, v)}{\partial v} \, du \, dv. \quad (\text{B-63})$$

Der KENDALL'sche Konkordanz-Koeffizient τ spielt in der Copula-Theorie eine besondere Rolle. Er steht in einem engen Zusammenhang mit der Verteilungsfunktion $K_C(t)$,

$$K_C(t) = \Pr[H(X, Y) \leq t] = \Pr[C(U, V) \leq t], \quad (\text{B-64})$$

die im "Copula-Raum" zwischen den beiden FRÉCHET-HOEFFDING-Schranken $W(u, v)$ und $M(u, v)$ eine Ordnungsstruktur definiert. Sie wird als KENDALL'sche Verteilungsfunktion bezeichnet und steht mit τ in folgendem Zusammenhang (siehe [GEN 93], [GEN 01], [NEL 03], [NEL 09]):

$$\tau = 3 - 4 \cdot \int_0^1 K_C(t) dt. \quad (\text{B-65})$$

Für die untere Schranke $W(u, v)$ folgt aus Gl. (B-57b) und Gl. (B-64): $K_W(t) = 1$. Aus Gl. (B-65) folgt $\tau = -1$ (Fall der perfekten Diskordanz). Für die obere Schranke $M(u, v)$ folgt aus Gl. (B-57c) und Gl. (B-64) (siehe hierzu auch Abbildung B-1): $K_M(t) = t$. Aus Gl. (B-65) folgt damit $\tau = 1$ (Fall der perfekten Konkordanz). Für andere Copulas verläuft die KENDALL'sche Verteilungsfunktion $K_C(t)$ zwischen diesen beiden Extremen: $t \leq K_C(t) \leq 1$. Die Verteilungsfunktion $K_C(t)$ ist wesentlich informativer als der Koeffizient τ , der gemäß Gl. (B-65) ein integraler Parameter ist. Mit $K_C(t)$ können auch Besonderheiten zur Abhängigkeit zwischen Zufallsgrößen in den Randbereichen (den Tails) untersucht und dargestellt werden (siehe Abschnitt 4.3).

Für Risikoanalysen ist die mit einer Copula erfasste Tail-Abhängigkeit zwischen Zufallsgrößen von besonderem Interesse. Die Koeffizienten λ_U und λ_L für die obere (upper) bzw. untere (lower) Tail-Abhängigkeit sind durch folgende Grenzwerte für bedingte Wahrscheinlichkeiten definiert:

$$\lambda_U = \lim_{t \rightarrow 1^-} \Pr[Y > G^{-1}(t) \mid X > F^{-1}(t)], \quad (\text{B-66a})$$

$$\lambda_L = \lim_{t \rightarrow 0^+} \Pr[Y \leq G^{-1}(t) \mid X \leq F^{-1}(t)]. \quad (\text{B-66b})$$

Die Koeffizienten λ_U und λ_L können Werte im Bereich von 0 bis 1 annehmen. Ist z. B. λ_U (oder λ_L) gleich Null, dann sind die Zufallsgrößen X und Y für sehr große (kleine) Quantile asymptotisch unabhängig. Je größer die Koeffizienten λ_U bzw. λ_L werden, umso stärker ist die Konkordanz der beiden Zufallsvariablen im jeweiligen Randbereich. Über die Copulas können diese Koeffizienten durch folgende Beziehungen bestimmt werden:

$$\lambda_U = 2 - \lim_{t \rightarrow 1^-} \frac{1 - C(t, t)}{1 - t}, \quad (\text{B-67a})$$

$$\lambda_L = \lim_{t \rightarrow 0^+} \frac{C(t, t)}{t}. \quad (\text{B-67b})$$

Eine Übersicht zu verschiedenen Methoden der nicht-parametrischen Schätzung der Koeffizienten λ_U und λ_L anhand von Stichproben und den dabei zu beachtenden Problemen geben [FRAH 05] und [POU 07]. Zu weiteren Konzepten für die Charakterisierung von Besonderheiten der Abhängigkeit zwischen zwei Zufallsgrößen sei auf [DROU 01], [NEL 07] und [BAL 09] verwiesen.

B.3.2 Familien von Copulas

Aus der großen Vielfalt von Copula-Klassen und ihren Familien werden einige Copula-Familien näher beschrieben. Mit der getroffenen Auswahl sollen wichtige Klassen, wie die der elliptischen und der archimedischen Copulas kurz dargestellt werden, wobei der Fokus auf den Copula-Familien

liegt, die im Rahmen des vorliegenden Berichtes für die Analyse bzw. Simulation von abhängigen Zufallsgrößen genutzt werden. Zudem werden Copulas angesprochen, wie die MARSHALL-OLKIN-Copula, deren Konstruktion auf Modellen zur Genesis von multivariaten Verteilungen basiert. Zur Vereinfachung werden die Darstellungen auf den Fall von zwei Variablen ($d = 2$) beschränkt.

B.3.2.1 Copulas von elliptischen Verteilungen

Wichtige Familien aus der Klasse der so genannten elliptischen Copulas sind die GAUSS-Copula, die auf der multivariaten Normalverteilung basiert, und die t-Copula, die aus der multivariaten STUDENT'schen t-Verteilung abgeleitet ist.

Für $d = 2$ ist die GAUSS-Copula nach Gl. (B-53) definiert durch:

$$C_G(u, v | r) = \frac{1}{2 \cdot \pi \cdot \sqrt{1-r^2}} \cdot \int_{-\infty}^{\Phi^{-1}(u)} \int_{-\infty}^{\Phi^{-1}(v)} \exp\left(-\frac{s^2 - 2 \cdot r \cdot s \cdot t + t^2}{2 \cdot (1-r^2)}\right) ds dt, \quad (\text{B-68})$$

wobei Φ die Verteilungsfunktion der univariaten Standardnormalverteilung $N(0, 1)$ ist und r den linearen Korrelationskoeffizienten bezeichnet. Der KENDALL'sche Konkordanzkoeffizient τ und der Spearman'sche Rangkorrelationskoeffizient ρ sind mit r über folgende Beziehungen verknüpft:

$$\tau = \frac{2}{\pi} \cdot \arcsin(r), \quad \rho = \frac{6}{\pi} \cdot \arcsin(r/2). \quad (\text{B-69})$$

Die GAUSS-Copula hat keine Tail-Abhängigkeit, d. h. $\lambda_U = \lambda_L = 0$.

Für die STUDENT-Copula (t-Copula) gilt gemäß Gl. (B-53) im Fall $d = 2$ folgende Darstellung:

$$C_S(u, v | r) = \frac{1}{2 \cdot \pi \cdot \sqrt{1-r^2}} \cdot \int_{-\infty}^{t_v^{-1}(u)} \int_{-\infty}^{t_v^{-1}(v)} \left(1 + \frac{s^2 - 2 \cdot r \cdot s \cdot t + t^2}{v \cdot (1-r^2)}\right)^{-(v+2)/2} ds dt, \quad (\text{B-70})$$

wobei t_v die Verteilungsfunktion der STUDENT'schen t-Verteilung mit v Freiheitsgraden und r den Korrelationskoeffizienten der bivariaten t-Verteilung bezeichnen. Zur Beziehung zwischen r und τ bzw. ρ gilt für die t-Copula nach [EMB 03] ebenfalls Gl. (B-69). Im Unterschied zur GAUSS-Copula weist die t-Copula eine Tail-Abhängigkeit auf [EMB 03]:

$$\lambda_L = \lambda_U = 2 \cdot \left(1 - t_{v+1}(\sqrt{v+1} \cdot \sqrt{1-r} / \sqrt{1+r})\right). \quad (\text{B-71})$$

Zu weiteren Eigenschaften der t-Copula sei auf [DEMA 05] verwiesen.

Die Erzeugung zufälliger Realisierungen (u, v) aus der GAUSS- und der t-Copula kann entsprechend dem in Abschnitt beschriebenen generellen Algorithmus erfolgen.

Simulation der GAUSS-Copula:

1. Erzeugung von unabhängigen Realisierungen x und z aus der Standardnormalverteilung $N(0, 1)$.
2. Es wird $y = x \cdot r + z \cdot \sqrt{1-r^2}$ berechnet.
3. Mit $u = \Phi(x)$ und $v = \Phi(y)$ ist (u, v) eine zufällige Realisierung aus der Copula $C_G(u, v)$.

Simulation der STUDENT-Copula:

1. Erzeugung von unabhängigen Realisierungen x und z aus der Standardnormalverteilung $N(0, 1)$ sowie einer Zufallszahl w aus der χ^2 -Verteilung mit v Freiheitsgraden.
2. Es wird $y = x \cdot r + z \cdot \sqrt{1 - r^2}$ berechnet.
3. Mit $u = t_v(\sqrt{v/w} \cdot x)$ und $v = t_v(\sqrt{v/w} \cdot y)$, wobei t_v die Verteilungsfunktion der STUDENT'schen t -Verteilung bezeichnet, ist (u, v) eine zufällige Realisierung aus der Copula $C_S(u, v)$.

Zur Veranschaulichung der Unterschiede zwischen der GAUSS- und der STUDENT-Copula sind in Abbildung B-2 Scatterplots zu den Dichten der beiden Copulas mit jeweils $n = 2000$ zufälligen Realisierungen zu den Parameterwerten $r = 0.8$ und $v = 3$ dargestellt. Die Populationswerte zu den Koeffizienten τ und ρ gemäß Gl. (B-69) sind $\tau = 0,590$ und $\rho = 0,786$. Mit den in Abschnitt B.2.3 genannten Beziehungen für die Punktschätzer dieser Koeffizienten resultieren für die Stichproben folgende Werte zur bivariaten GAUSS-Verteilung: $r_{x,y} = 0,802$; $\tau_{x,y} = 0,599$; $\rho_{x,y} = 0,797$; bzw. zur bivariaten STUDENT-Verteilung: $r_{x,y} = 0,778$; $\tau_{x,y} = 0,580$; $\rho_{x,y} = 0,758$.

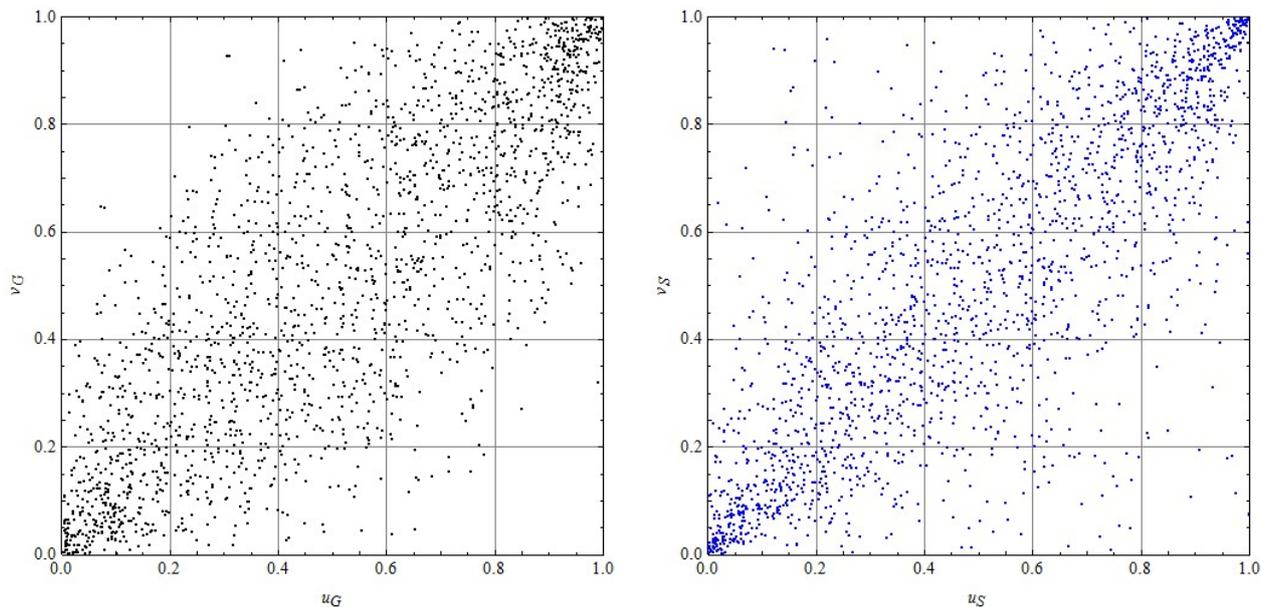


Abbildung B-2: Scatterplots zu den Copulas $C_G(u, v)$ (links) und $C_S(u, v)$ (rechts) zu den Parameterwerten $r = 0,8$ und $v = 3$ mit 2000 zufälligen Realisierungen

Aus Abbildung B-2 ist insbesondere zu erkennen, dass die STUDENT-Copula im Bereich der Tails (für $u < 0,2$ und $v < 0,2$ bzw. für $u > 0,8$ und $v > 0,8$) eine größere Konkordanz aufweist als die GAUSS-Copula. Die Unterschiede zwischen beiden Copulas werden noch deutlicher, wenn man aus ihnen bivariate Verteilungen $H(x, y)$ erzeugt und für die Randverteilungen $G(x)$ und $F(y)$ jeweils die Standardnormalverteilung angesetzt wird, wie in Abbildung B-3 dargestellt. Die Wertepaare $(x_{G(G)}, y_{G(G)})$ bzw. $(x_{G(S)}, y_{G(S)})$ basieren auf der GAUSS- bzw. der STUDENT-Copula.

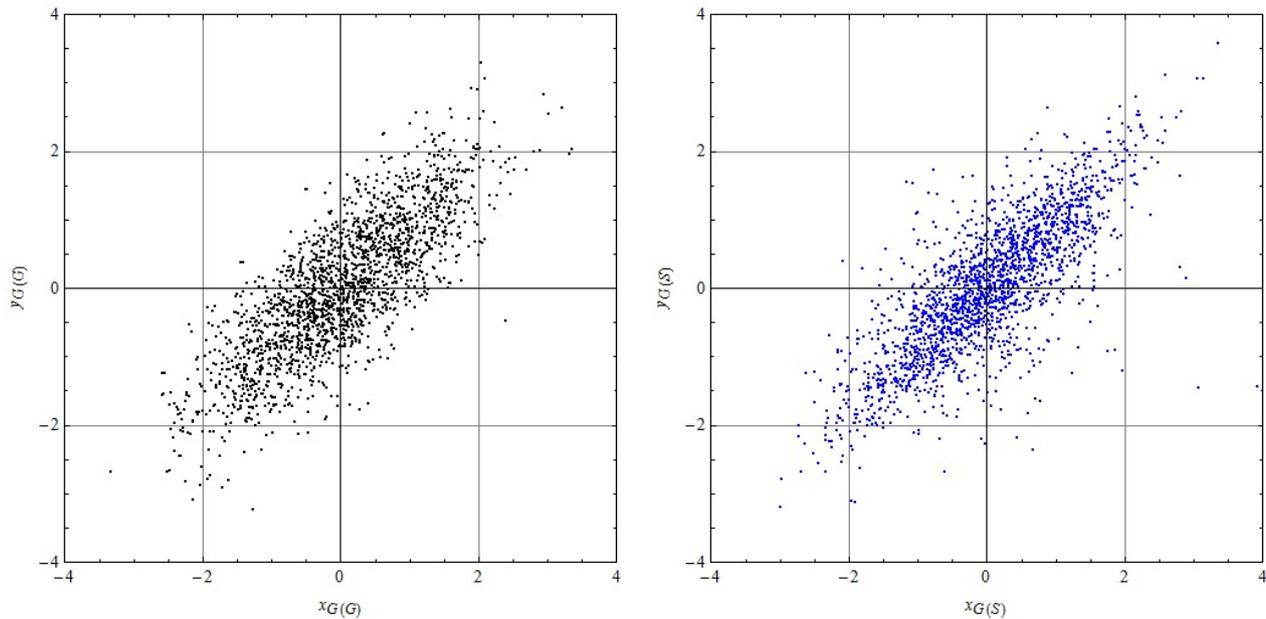


Abbildung B-3: Scatterplots zur gemeinsamen Verteilung $H(x, y)$ mit den gleichen Randverteilungen der Zufallsgrößen $X \sim N(0, 1)$ und $Y \sim N(0, 1)$, deren Abhängigkeit jedoch durch die GAUSS-Copula (links) bzw. die STUDENT-Copula (rechts) gemäß Abbildung B-2 definiert ist.

Aus Abbildung B-3 ist klar ersichtlich, dass bei gleichen Werten der Korrelationskoeffizienten r und ρ bzw. des Konkordanzkoeffizienten τ sehr unterschiedliche Abhängigkeiten zwischen zwei Zufallsgrößen X und Y vorliegen können. Die in kommerziellen Programmen zur MC-Simulation übliche Vereinfachung der Abhängigkeitsstruktur auf GAUSS-Copulas kann deshalb substantielle Einschränkung bzgl. der Erfassung realer Abhängigkeiten zwischen Zufallsvariablen bewirken. Dies wird anhand der im folgenden Abschnitt betrachteten Copulas, die geringere Ähnlichkeiten als die GAUSS- und die STUDENT-Copula aufweisen, noch deutlicher.

B.3.2.2 Archimedische Copulas

Eine umfassende Darstellung der mathematischen Grundlagen und viele Beispiele zur Klasse der so genannten archimedischen Copulas findet man in [NEL 07]. In den folgenden Erläuterungen werden einige mathematische Aspekte vereinfacht behandelt, indem nur solche Beispiele betrachtet werden, für die bestimmte "technische Feinheiten" nicht von Belang sind.

Archimedische Copulas $C(u, v)$ sind relativ einfache Konstrukte, welche die Definitionsgleichungen (B-52a) bis (B-52c) erfüllen. Sie basieren auf generierenden Funktionen $\varphi(t)$, die stetig, strikt fallend und konvex sind (d. h. $\varphi''(t) > 0$) und das Einheitsintervall $[0, 1]$ auf die Menge der positiven reellen Zahlen mit $\varphi(1) = 0$ abbilden. Die inverse Funktion $\varphi^{-1}(w)$ mit $0 \leq w < \infty$ ist dann eindeutig bestimmt und nimmt Werte im Einheitsintervall $[0, 1]$ an. Dann ist die Funktion

$$C(u, v) = \varphi^{-1}(\varphi(u) + \varphi(v)) \quad (\text{B-72})$$

eine archimedische Copula.

Archimedische Copulas sind gemäß Gl. (B-72) symmetrisch. Mit Hilfe der Transformationen nach Gl. (B-55b) bzw. (B-55c) können daraus jedoch auch Copulas gebildet werden, deren Dichten bzgl. der Gerade $v = 1 - u$ symmetrisch sind. Mit der Transformation Gl. (B-55d) resultiert eine Copula, deren Dichte im Vergleich zu $C(u, v)$ bzgl. der Gerade $v = 1 - u$ gespiegelt ist. Es sei angemerkt, dass alle archimedischen Copulas auch assoziativ sind, d. h. $C(C(u, v), w) = C(u, C(v, w))$ und auch für mehrere Variable u_1 bis u_d mit $d > 2$ definiert werden können [NEL 07].

Die generierenden Funktionen $\varphi(t)$ können durch einen oder auch mehrere Verteilungsparameter spezifiziert sein. Hier werden nur arithmetische Copula-Familien mit einem Verteilungsparameter θ betrachtet. In [NEL 07] ist eine Liste von 22 solcher Copula-Familien aufgeführt und charakterisiert, in der neben gebräuchlichen/bekannten Copulas, wie der CLAYTON-Copula [CLAY 78] und der ALI-MIKHAIL-HAQ-Copula [ALI 78], auch diverse namenlose Copulas enthalten sind, auf die in der Fachliteratur über die lfd. Nummer der Liste von [NEL 07] Bezug genommen wird. Für die im Rahmen des vorliegenden Berichts durchgeführten Berechnungen zur Anwendung von Copulas für die Analyse und Simulation von Abhängigkeiten zwischen Zufallsvariablen (siehe Abschnitt 4.3) wurden aus dieser Liste vier archimedische Copulas ausgewählt, die sich hinsichtlich der Tail-Abhängigkeit deutlich unterscheiden. Dies sind die in der Tabelle 4.1 von [NEL 07] unter der Nr. 5 aufgeführte FRANK-Copula [FRA 79], [GEN 87], die unter Nr. 6 aufgeführte JOE-Copula [JOE 93], [JOE 97] sowie zwei weiter unter Nr. 12 bzw. Nr. 19 aufgeführte Copulas. In Tabelle B-1 sind diese vier Copulas und ihre generierenden Funktionen zusammengestellt. Zusätzlich enthalten die beiden letzten Zeilen von Tabelle B-1 die entsprechenden Angaben für die Produkt-Copula Π (Gl. (B-56)) sowie für eine weitere spezielle archimedische Copula, die hier mit Σ bezeichnet wurde.

Tabelle B-1: Bestimmungsgleichungen zu ausgewählten archimedischen Copulas

Copula	$C(u, v \theta)$	$\varphi(t \theta)$	zulässige θ	Grenzfälle
C_F FRANK	$-\frac{1}{\theta} \cdot \ln \left(1 + \frac{(e^{-\theta u} - 1) \cdot (e^{-\theta v} - 1)}{e^{-\theta} - 1} \right)$	$-\ln \left(\frac{e^{-\theta t} - 1}{e^{-\theta} - 1} \right)$	$-\infty < \theta < \infty$ *)	W: $\theta \rightarrow -\infty$ Π: $\theta \rightarrow 0$ M: $\theta \rightarrow \infty$
C_J JOE	$1 - \left((1-u)^\theta + (1-v)^\theta - (1-u)^\theta \cdot (1-v)^\theta \right)^{1/\theta}$	$-\ln(1 - (1-t)^\theta)$	$1 \leq \theta < \infty$	Π: $\theta \rightarrow 1$ M: $\theta \rightarrow \infty$
C_{12}	$1 - \left((1-u)^\theta + (1-v)^\theta - (1-u)^\theta \cdot (1-v)^\theta \right)^{1/\theta}$	$(t^{-1} - 1)^\theta$	$1 \leq \theta < \infty$	Σ: $\theta \rightarrow 1$ M: $\theta \rightarrow \infty$
C_{19}	$\theta / \ln(e^{\theta/u} + e^{\theta/v} - e^\theta)$	$e^{\theta/t} - e^\theta$	$0 < \theta < \infty$	Σ: $\theta \rightarrow 0$ M: $\theta \rightarrow \infty$
Π	$u \cdot v$	$-\ln(t)$	entfällt	entfällt
Σ ^{**)}	$u \cdot v / (u + v - u \cdot v)$	$1/t - 1$	entfällt	entfällt

*) Der Wert $\theta = 0$ ist streng genommen aus dem Definitionsbereich der FRANK-Copula ausgenommen. Im Sinne des Grenzübergangs für $\theta \rightarrow 0$ existieren jedoch sowohl die generierende Funktion als auch die Copula und stimmen in diesem Grenzfalle mit der Produkt-Copula (φ_Π bzw. Π) überein.

***) In [NEL 07] wird für diese spezielle archimedische Copula, die als Grenzfalle von anderen archimedischen Copulas auftritt, die Bezeichnung $\Pi/(\Sigma-\Pi)$ verwendet.

Archimedische Copulas haben nicht zuletzt wegen ihrer relativ einfachen Handhabung eine breite Anwendung gefunden. Für sie können die KENDALL-Verteilungsfunktion $K_C(t)$ nach Gl. (B-64), der Konkordanzkoeffizient τ nach Gl. (B-65) und die Koeffizienten λ_L und λ_U zur Tail-Abhängigkeit nach Gl. (B-67) direkt auf die generierenden Funktionen $\varphi(t|\theta)$ zurückgeführt werden, wodurch sich (z. T. einfache) analytische Funktionen des Verteilungsparameters θ ergeben. Zudem vereinfachen sich Algorithmen für die Simulation.

Für archimedische Copulas resultiert zur KENDALL'schen Verteilungsfunktion $K_C(t)$ die Darstellung

$$K_C(t) = t - \frac{\varphi(t)}{\varphi'(t)}, \quad (\text{B-73})$$

wonach sich gemäß Gl. (B-65) auch für τ eine vereinfachte Berechnungsformel ergibt [GEN 93],

$$\tau = 1 + 4 \cdot \int_0^1 \frac{\varphi(t)}{\varphi'(t)} dt. \quad (\text{B-74})$$

Aus den Gleichungen (B-67a/b) folgen mit Gl. (B-72) die Darstellungen

$$\lambda_U = 2 - \lim_{t \rightarrow 1^-} \frac{1 - \varphi^{-1}(2 \cdot \varphi(t))}{1 - t} = 2 - \lim_{z \rightarrow 0^+} \frac{1 - \varphi^{-1}(2 \cdot z)}{1 - \varphi^{-1}(z)}, \quad (\text{B-75a})$$

$$\lambda_L = \lim_{t \rightarrow 0^+} \frac{\varphi^{-1}(2 \cdot \varphi(t))}{t} = \lim_{z \rightarrow \infty} \frac{\varphi^{-1}(2 \cdot z)}{\varphi^{-1}(z)}. \quad (\text{B-75b})$$

In Tabelle B-2 sind für die 6 Copulas aus Tabelle B-1 die nach Gl. (B-73) bis (B-75) resultierenden Beziehungen zusammengestellt.

Tabelle B-2: KENDALL'sche Verteilungsfunktion $K_C(t|\theta)$, Konkordanzkoeffizient $\tau_C(\theta)$, und Tail-Abhängigkeitskoeffizienten $\lambda_{L;C}(\theta)$ und $\lambda_{U;C}(\theta)$ für einige archimedische Copulas

Copula	$K_C(t \theta)$	$\tau_C(\theta)$	$\lambda_{L;C}(\theta)$	$\lambda_{U;C}(\theta)$
C_F FRANK	$t - \frac{(e^{\theta \cdot t} - 1)}{\theta} \cdot \ln\left(\frac{e^{-\theta \cdot t} - 1}{e^{-\theta} - 1}\right)$	$1 - \frac{4}{\theta} \cdot (1 - D_1(\theta))$	0	0
C_J JOE	$t + \frac{1}{\theta} \left((1-t)^\theta - 1 \right) \cdot (1-t)^{1-\theta} \cdot \ln(1 - (1-t)^\theta)$	$\frac{\theta - 4 + 2 \cdot (\gamma + \psi(1 + 2/\theta))}{\theta - 2}$	0	$2 - 2^{1/\theta}$
C_{12}	$\frac{t \cdot (1-t+\theta)}{\theta}$	$1 - \frac{2}{3 \cdot \theta}$	$2^{-1/\theta}$	$2 - 2^{1/\theta}$
C_{19}	$\frac{t}{\theta} \cdot (t + \theta - t \cdot e^{-\theta/t})$	$\frac{1}{3} (1 + 2 \cdot \theta + 2 \cdot e^\theta \cdot \theta^2 \cdot \text{Ei}(-\theta))$	1	0
Π	$t \cdot (1 - \ln(t))$	0	0	0
Σ	$t \cdot (2 - t)$	1/3	1/2	0

Einige in der 3. Spalte von Tabelle B-2 für $\tau_C(\theta)$ angegebenen Formeln enthalten spezielle Symbole bzw. Funktionen. In der Formel zur FRANK-Copula stellt $D_1(z)$ die DEBYE-Funktion erster Ordnung dar:

$$D_1(z) = \frac{1}{z} \int_0^z \frac{t}{e^t - 1} dt. \quad (\text{B-76})$$

In der Formel zur JOE-Copula ist $\gamma = 0,577216$ die EULER'sche Konstante. $\psi(z)$ bezeichnet die Digammafunktion,

$$\psi(z) = \frac{\Gamma'(z)}{\Gamma(z)} \quad \text{mit der Gammafunktion} \quad \Gamma(z) = \int_0^{\infty} t^{z-1} \cdot e^{-t} dt \quad (\text{B-77})$$

In der Formel zur Copula C_{19} ist $Ei(z)$ die Integralexponentialfunktion

$$Ei(z) = - \int_{-z}^{\infty} \frac{e^{-t}}{t} dt. \quad (\text{B-78})$$

In Abbildung B-4 ist $\tau_C(\theta)$ in Abhängigkeit vom jeweiligen Verteilungsparameter θ für die Copulas C_F , C_J , C_{12} und C_{19} dargestellt (der Verlauf mit $\theta < 0$ für die FRANK-Copula ist nicht dargestellt). Die Abbildung B-5 zeigt mit jeweils 2000 zufälligen Realisierungen Scatterplots zu den Dichten der Copulas C_F , C_J , C_{12} und C_{19} bei Ansatz von $\tau = 0,59$ (analog zur Abbildung B-2).

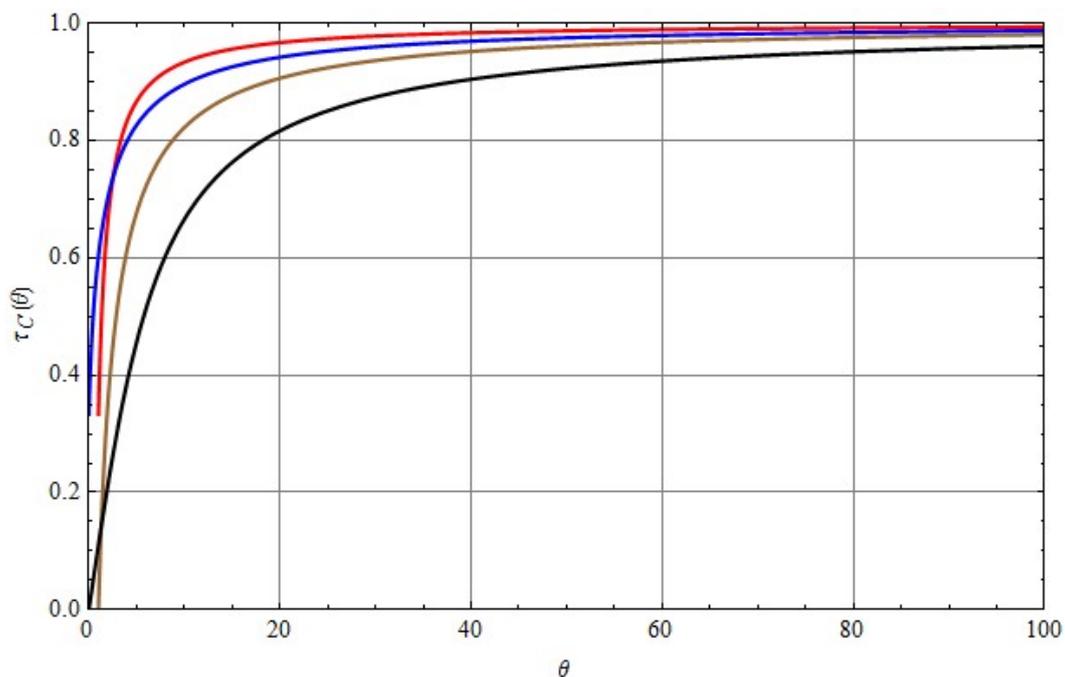


Abbildung B-4: Konkordanzkoeffizienten $\tau_C(\theta)$ für die FRANK-Copula (schwarz), die JOE-Copula (braun), sowie für die Copulas C_{12} (rot) und C_{19} (blau)

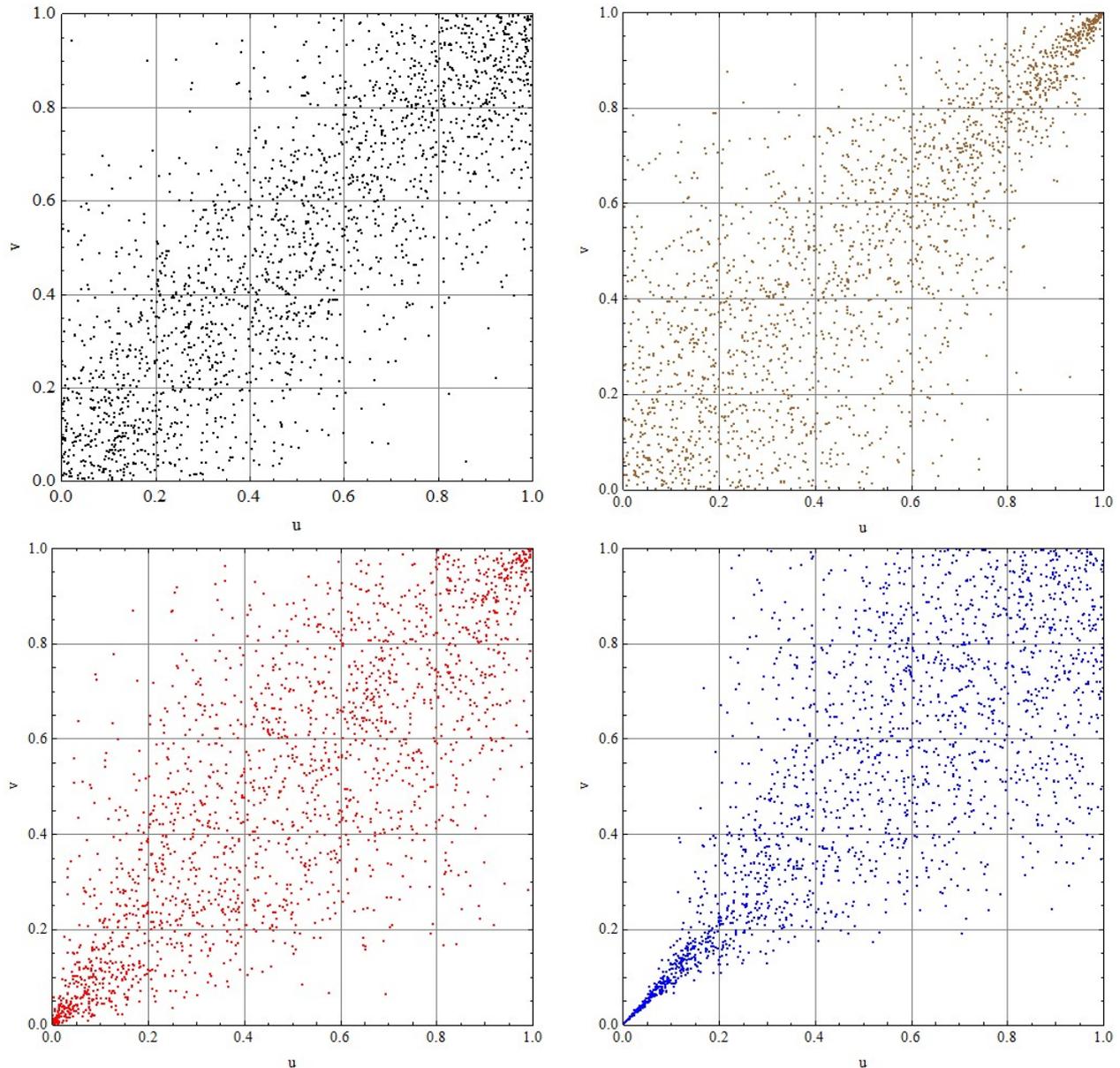


Abbildung B-5: Scatterplots zu den Copulas C_F (links oben, schwarz), C_J (rechts oben, braun), C_{12} (links unten, rot) und C_{19} (rechts unten, blau) für $\tau = 0,59$ mit 2000 zufälligen Realisierungen

Der Konkordanzkoeffizient $\tau = 0,59$ entspricht für diese vier Copulas den folgenden Werten des jeweiligen Verteilungsparameters θ : $\theta_F = 7,669$; $\theta_J = 3,708$; $\theta_{12} = 1,626$ und $\theta_{19} = 1,626$. Die von θ abhängigen Werte von Koeffizienten der Tail-Abhängigkeit (vgl. Tabelle B-2) sind: $\lambda_{U;J} = 0,794$; $\lambda_{L;12} = 0,653$ und $\lambda_{U;12} = 0,468$.

Für den Vergleich mit Abbildung B-3 wurden für die in Abbildung B-5 dargestellten Copulas gemäß Gl. (B-54) die Verteilungen von Zufallsgrößen X und Y erzeugt, deren Randverteilungen $F(x)$ und $G(y)$ Standardnormalverteilungen sind. Die entsprechenden Scatterplots zeigt Abbildung B-6.

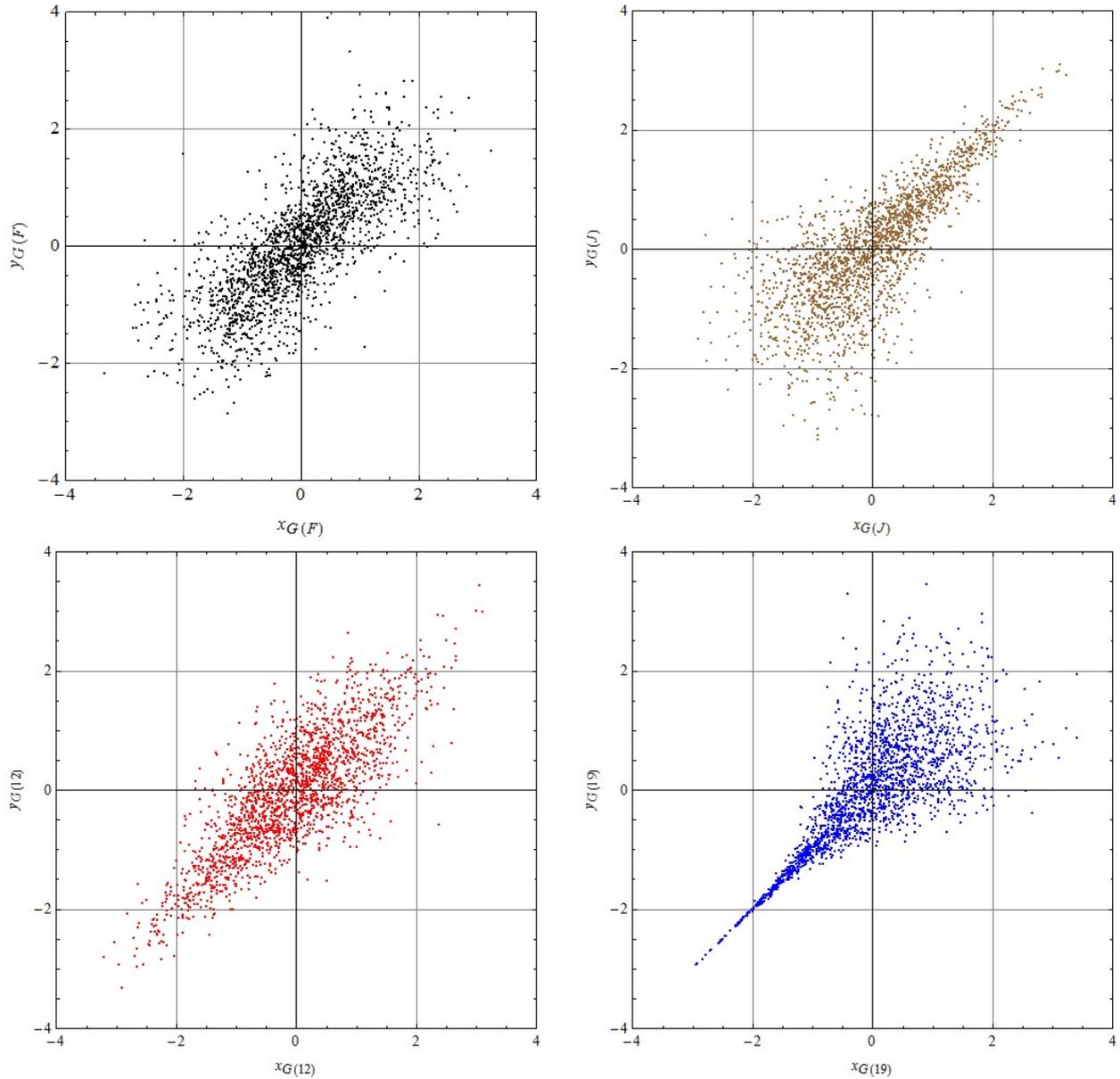


Abbildung B-6: Scatterplots zur gemeinsamen Verteilung $H(x, y)$ mit den gleichen Randverteilungen der Zufallsgrößen $X \sim N(0,1)$ und $Y \sim N(0,1)$, deren Abhängigkeit jedoch durch die Frank-Copula (links oben), die Joe-Copula (rechts oben) bzw. die Copulas C_{12} (links unten) und C_{19} (rechts unten,) gemäß Abbildung B-5 definiert ist

Die (linearen) Korrelationskoeffizienten $r_{x,y}$ der in Abbildung B-6 dargestellten Verteilungen liegen in der Nähe des für die bivariate GAUSS-Verteilung von Abbildung B-3 angesetzten Wertes von 0,8. Die Werte sind 0,752 (basierend auf C_F), 0,795 (basierend auf C_J), 0,807 (basierend auf C_{12}) und 0,759 (basierend auf C_{19}). Dieser Vergleich zeigt, dass die alleinige Angabe von Korrelationskoeffizienten bzgl. der Struktur/Form der Abhängigkeit zwischen Zufallsgrößen, insbesondere für die Randbereiche, nicht informativ ist.

Zum Vergleich sind in Abbildung B-7 für die beiden nicht-parametrischen Copulas Π und Σ aus den Tabellen B-1 und B-2 Scatterplots zur jeweiligen Copula sowie den daraus resultierenden bivariaten Verteilungen von Zufallsgrößen X und Y mit Randverteilungen $N(0, 1)$ dargestellt. Die Produkt-Copula Π beschreibt voneinander unabhängige Zufallsgrößen; Σ beschreibt für diverse archimedische Copulas einen Grenzfall, für den $\tau = 1/3$ und $\lambda_L = 1/2$ gilt.

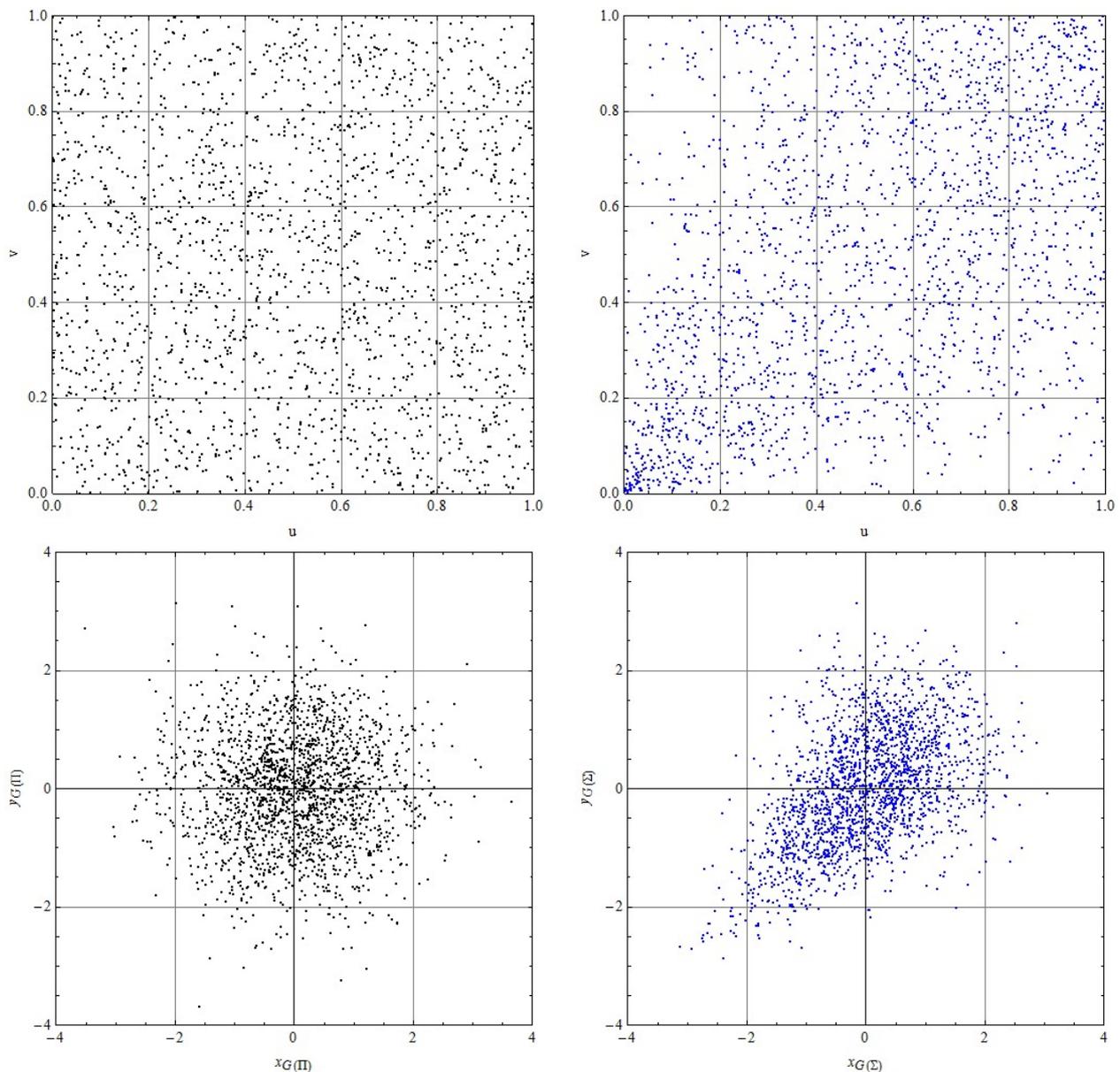


Abbildung B-7: Scatterplots zu den Copulas Π (links oben, schwarz) und Σ (rechts oben, blau) mit 2000 zufälligen Realisierungen sowie daraus resultierende gemeinsame Verteilung $H(x, y)$ mit den gleichen Randverteilungen der Zufallsgrößen $X \sim N(0, 1)$ und $Y \sim N(0, 1)$, deren Unabhängigkeit durch die Produkt-Copula (links unten) bzw. Abhängigkeit durch die Copula Σ , (rechts unten) gemäß Gl. (B-54) definiert ist.

Simulation aus archimedischen Copula-Verteilungen: Die Simulation zufälliger Realisierungen aus archimedischen Copulas vereinfacht sich im Vergleich zu dem in Abschnitt B.3.1 beschriebenen allgemeinen Algorithmus, weil die KENDALL'sche Verteilungsfunktion $K_C(t)$ (siehe Gl. B-73) die Verteilung der Zufallsgröße $C(U, V)$ ist [GEN 93]. Der generelle Algorithmus nimmt damit für archimedische Copulas die folgende Form an [NEL 07]:

1. Es werden zwei voneinander unabhängige Realisierungen s und t von auf dem Einheitsintervall $[0, 1]$ gleichverteilten Zufallszahlen generiert.
2. Mit der inversen K-Verteilungsfunktion K_C^{-1} wird $w = K_C^{-1}(t)$ berechnet.
3. Mit der generierenden Funktion φ der Copula C und ihrer Inversen φ^{-1} werden $u = \varphi^{-1}(s \cdot \varphi(w))$ sowie $v = \varphi^{-1}((1-s) \cdot \varphi(w))$ berechnet. Das Wertepaar (u, v) ist eine zufällige Realisierung der Copula $C(u, v)$.

Diese Methode verlangt eine analytische Darstellung der Inversen der K-Verteilungsfunktion, K_C^{-1} .

Ein spezieller Algorithmus, der nur auf die generierende Funktion φ und ihrer Ableitung φ' sowie die Inversen dieser beiden Funktionen φ^{-1} und $\varphi^{(-1)}$ zurückgreift, wurde in [GEN 86] begründet:

1. Es werden zwei voneinander unabhängige Realisierungen u und t von auf dem Einheitsintervall $[0, 1]$ gleichverteilten Zufallszahlen generiert.
2. Es wird $w = \varphi^{(-1)}(\varphi'(u)/t)$ berechnet.
3. Es wird $v = \varphi^{-1}(\varphi(w) - \varphi(u))$ berechnet. Das Wertepaar (u, v) ist eine zufällige Realisierung der Copula $C(u, v)$.

Welcher dieser beiden Algorithmen für eine bestimmte Copula am zweckmäßigsten eingesetzt wird hängt davon ab, welche der beiden Inversen K_C^{-1} bzw. $\varphi^{(-1)}$ sich einfacher berechnen lässt.

B.3.2.3 MARSHALL-OLKIN-Copula

Die in [MAR 67A]/[MAR 67B] entwickelte MARSHALL-OLKIN-Copula ist ein Beispiel zur Ableitung bi- bzw. multivariater Verteilungen auf der Basis eines fachbezogenen Modells zur Abhängigkeit zwischen Zufallsgrößen. Im einfachsten Fall beschreibt sie die Lebensdauer eines aus zwei parallelen (redundanten) Einheiten bestehenden Systems unter Einfluss von seltenen Ereignissen, die gleichzeitig auf beide Einheiten einwirken können. Beschreibt man das Auftreten von seltenen Ereignissen mit fataler Wirkung durch einen POISSON-Prozess, so ist die (Über-)Lebenszeit eine exponentiell verteilte Zufallsgröße (siehe Anhang C, Abschnitte C.3.3 und C.5.4). Für ein System aus zwei Einheiten wird angenommen, dass die Auswirkungen von Ereignissen (Schocks) durch drei unabhängige POISSON-Prozesse mit den Parametern λ_1 , λ_2 und λ_{12} beschrieben werden können, wobei λ_1 die Ausfallrate der ersten Einheit, λ_2 die Ausfallrate der zweiten Einheit, und λ_{12} die Rate des gleichzeitigen Ausfalls beider Einheiten beim Auftreten seltener Ereignisse beschreiben. Die Lebenszeiten X und Y der beiden Einheiten können dann durch drei unabhängige, exponentiell verteilte Zufallsgrößen Z_1 , Z_2 und Z_{12} (mit den Parametern λ_1 , λ_2 bzw. λ_{12}) beschrieben werden:

$$X = \min\{Z_1, Z_{12}\} \quad \text{und} \quad Y = \min\{Z_2, Z_{12}\}. \quad (\text{B-79})$$

Bezeichnet man mit $H(x, y)$ die gemeinsame Verteilungsfunktion von X und Y , welche die Randverteilungen $F(x)$ und $G(y)$ haben, so kann man die gemeinsame Überlebensfunktion

$$\begin{aligned}\bar{H}(x, y) &= \Pr[X > x, Y > y] = \Pr[Z_1 > X] \cdot \Pr[Z_2 > y] \cdot \Pr[Z_{12} > \max\{x, y\}] \\ &= e^{-\lambda_1 \cdot x - \lambda_2 \cdot y - \lambda_{12} \cdot \max\{x, y\}}\end{aligned}\quad (\text{B-80})$$

durch eine so genannte Überlebens-Copula darstellen. Mit den Rand-Überlebensfunktionen

$$\bar{F}(x) = \Pr[X > x] = 1 - F(x) \quad \text{und} \quad \bar{G}(y) = \Pr[Y > y] = 1 - F(y) \quad (\text{B-81})$$

folgt zunächst

$$\begin{aligned}\bar{H}(x, y) &= \Pr[X > x, Y > y] = 1 - F(x) - G(y) + H(x, y) \\ &= \bar{F}(x) + \bar{G}(y) - 1 + H(x, y) \\ &= \bar{F}(x) + \bar{G}(y) - 1 + C(F(x), G(y)) \\ &= \bar{F}(x) + \bar{G}(y) - 1 + C(1 - \bar{F}(x), 1 - \bar{G}(y))\end{aligned}\quad (\text{B-82})$$

wobei $H(x, y)$ nach dem Satz von SKLAR gemäß Gl. (B-54a) durch eine Copula $C(F(x), G(y))$ dargestellt wurde. Für die auf dem Einheitsintervall $[0, 1]$ gleichverteilten Variablen

$$u = \bar{F}(x) \quad \text{und} \quad v = \bar{G}(y) \quad (\text{B-83})$$

gilt dann:

$$\bar{H}(x, y) = u + v - 1 + C(1 - u, 1 - v). \quad (\text{B-84})$$

Der Vergleich mit Gl. (B-55d) zeigt, dass die rechte Seite von Gl. (B-84) eine Copula ist. Wie die multivariate Verteilung $H(x, y)$ kann also auch die Überlebensfunktion $\bar{H}(x, y)$ durch eine Copula dargestellt werden. Sie wird Überlebens-Copula genannt [NEL 07] und mit $\hat{C}(u, v)$ bezeichnet:

$$\bar{H}(x, y) = \hat{C}(\bar{F}(x), \bar{G}(x)). \quad (\text{B-85})$$

Ersetzt man nun in Gl. (B-80) $\max\{x, y\}$ durch $x + y - \min\{x, y\}$, so folgt mit

$$\bar{F}(x) = e^{-(\lambda_1 + \lambda_{12}) \cdot x} \quad \text{und} \quad \bar{G}(y) = e^{-(\lambda_2 + \lambda_{12}) \cdot y} :$$

$$\bar{H}(x, y) = \bar{F}(x) \cdot \bar{G}(x) \cdot \min\{e^{\lambda_{12} \cdot x}, e^{\lambda_{12} \cdot y}\}.$$

Für u und v gemäß Gl. (B-83) und mit den Bezeichnungen

$$\alpha = \frac{\lambda_{12}}{\lambda_1 + \lambda_{12}} \quad \text{und} \quad \beta = \frac{\lambda_{12}}{\lambda_2 + \lambda_{12}}, \quad 0 \leq \alpha, \beta \leq 1, \quad (\text{B-86})$$

resultiert dann folgende Darstellung der Überlebens-Copula $\hat{C}(u, v)$ (MARSHALL-OLKIN-Copula):

$$\hat{C}(u, v) \equiv C_{\alpha, \beta}(u, v) = \min\{u^{1-\alpha} \cdot v, u \cdot v^{1-\beta}\} = \begin{cases} u^{1-\alpha} \cdot v & \text{für } u^\alpha \geq v^\beta \\ u \cdot v^{1-\beta} & \text{für } u^\alpha \leq v^\beta \end{cases} \quad (\text{B-87})$$

Grenzfälle sind die Produkt-Copula, $C_{\alpha, 0} = C_{0, \beta} = \Pi$, und die obere FRÉCHET-HOEFFDING-Schranke, $C_{1, 1} = M$ (siehe Gl. (B-56) und (B-57c)). Für $\alpha = \beta$ resultiert die CUADRAS-AUGÉ-Copula [CUA 81].

Die MARSHALL-OLKIN-Copula besitzt einen stetigen und einen singulären Anteil (siehe Gl. (B-60)), wobei der singuläre Anteil

$$S_{\alpha,\beta}(u, v) = \frac{\alpha \cdot \beta}{\alpha + \beta - \alpha \cdot \beta} \cdot \left(\min\{u^\alpha, v^\beta\} \right)^{\frac{\alpha+\beta-\alpha\beta}{\alpha\beta}} \quad (\text{B-88})$$

auf der Kurve $u^\alpha = v^\beta$ liegt. Die Größe $S_{\alpha,\beta}(1, 1) = \alpha \cdot \beta / (\alpha + \beta - \alpha \cdot \beta)$ ist das Wahrscheinlichkeitsmaß für die singuläre Komponente $\Pr[U^\alpha = V^\beta]$. Für den Konkordanzkoeffizienten erhält man gemäß Gl. (B-63) den gleichen Wert: $\tau_{\alpha,\beta} = \alpha \cdot \beta / (\alpha + \beta - \alpha \cdot \beta)$. Die Koeffizienten zur Tail-Abhängigkeit gemäß Gl. (B-67) lassen sich direkt berechnen: $\lambda_L = 0$ und $\lambda_U = \min\{\alpha, \beta\}$.

Die Simulation kann z. B. mit folgendem Algorithmus vorgenommen werden:

1. Es werden drei voneinander unabhängige Realisierungen r, s und t von auf dem Einheitsintervall $[0, 1]$ gleichverteilten Zufallszahlen generiert.
2. Es werden $a = \min\{-\ln(r)/(1-\alpha); -\ln(t)/\alpha\}$ und $b = \min\{-\ln(s)/(1-\beta); -\ln(t)/\beta\}$ berechnet.
3. Dann ist mit $u = e^{-a}$ und $v = e^{-b}$ das Wertepaar (u, v) eine zufällige Realisierung der MARSHALL-OLKIN-Copula $C_{\alpha,\beta}(u, v)$.

Die Abbildung B-8 zeigt Scatterplots zur MARSHALL-OLKIN-Copula für zwei Parametersätze (mit $\tau_{\alpha,\beta} = 1/16$ bzw. $1/4$), welche die obigen Erläuterung veranschaulichen. Abschließend sei vermerkt, dass die MARSHALL-OLKIN-Copula nicht nur in Modellen mit exponentiellen Randverteilungen genutzt wird. Eine Übersicht zu diversen Anwendungen enthält [BAL 09].

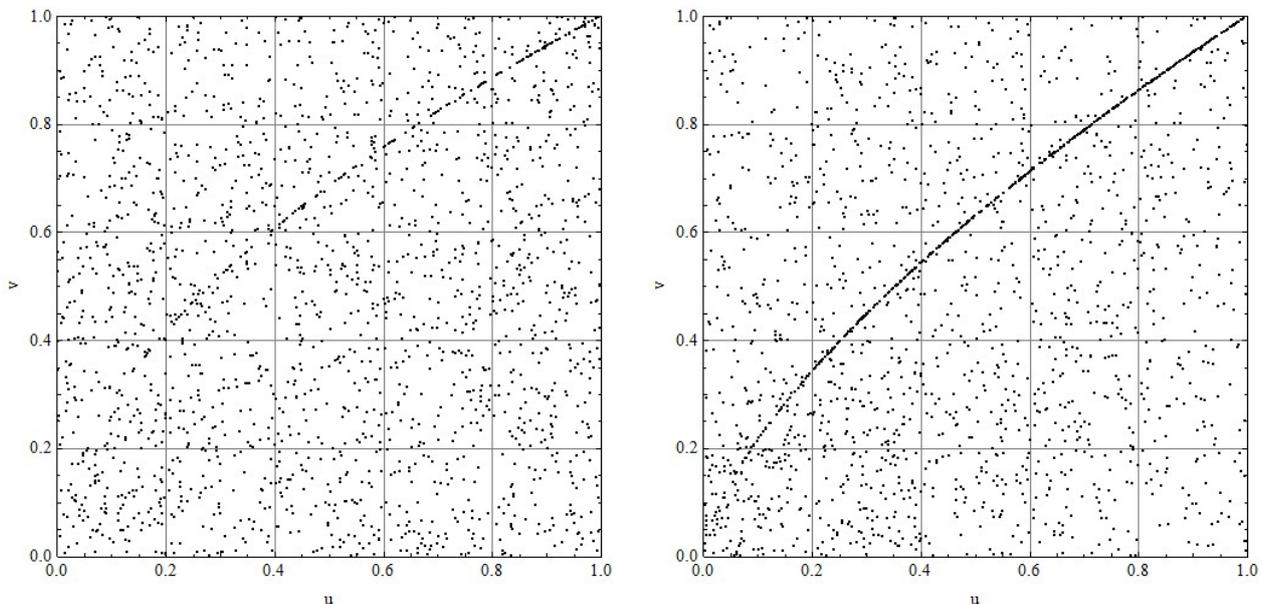


Abbildung B-8: Scatterplots zur MARSHALL-OLKIN-Copula mit 2000 zufälligen Realisierungen für zwei Parametersätze mit $\lambda_1 = 2; \lambda_2 = 1$;
 linkes Bild: $\lambda_{12} = 1/5$; somit $\alpha = 1/11; \beta = 1/6; \tau_{\alpha,\beta} = 1/16$ (für Stichprobe $\tau_{u,v} = 0,0693$);
 rechtes Bild: $\lambda_{12} = 1$; somit $\alpha = 1/3; \beta = 1/2; \tau_{\alpha,\beta} = 1/4$ (für Stichprobe $\tau_{u,v} = 0,248$)

B.3.3 Berechnung der empirischen Copula und KENDALL-Verteilung

Zur Anpassung von parametrischen Copulas an Stichproben, die mit Hilfe der empirischen Randverteilungsfunktionen der jeweiligen Zufallsgrößen auf das Einheitsintervall $[0, 1]$ transformiert wurden, ist die empirische Copula von zentraler Bedeutung. Sie wird analog zur empirischen Verteilungsfunktion univariater Zufallsgrößen definiert (siehe z. B. [NEL 07]).

Es sei $\{(x_k, y_k)\}_n$ eine Stichprobe des Umfangs n von Wertepaaren (x_k, y_k) der Zufallsgrößen X und Y (k von 1 bis n). Mit $x_{(i)}$ und $y_{(j)}$ werden die in aufsteigender Ordnung nummerierten Werte von X bzw. Y bezeichnet: $x_{(1)}$ ist der kleinste und $x_{(n)}$ der größte Wert der Zufallsgröße X in der Stichprobe, wie auch $y_{(1)}$ den kleinsten und $y_{(n)}$ den größten Stichprobenwert der Zufallsgröße Y bezeichnen. Dann ist die empirische Copula $C_n(i/n, j/n)$ wie folgt definiert:

$$C_n\left(\frac{i}{n}, \frac{j}{n}\right) = \frac{\text{Anzahl der Paare } (x_k, y_k) \text{ der Stichprobe mit } x_k \leq x_{(i)} \text{ und } y_k \leq y_{(j)}}{n}. \quad (\text{B-89})$$

Die empirische Copula-Massendichte, $c_n(i/n, j/n)$ ist definiert durch:

$$c_n\left(\frac{i}{n}, \frac{j}{n}\right) = \begin{cases} 1/n & \text{wenn } (x_{(i)}, y_{(j)}) \text{ ein Element der Stichprobe ist} \\ 0 & \text{sonst} \end{cases}. \quad (\text{B-90})$$

Zwischen der empirischen Copula(-verteilungsfunktion) C_n und ihrer Massendichte c_n bestehen folgende Beziehungen:

$$C_n\left(\frac{i}{n}, \frac{j}{n}\right) = \sum_{k=1}^i \sum_{l=1}^j c_n\left(\frac{k}{n}, \frac{l}{n}\right), \quad (\text{B-91})$$

$$c_n\left(\frac{i}{n}, \frac{j}{n}\right) = C_n\left(\frac{i}{n}, \frac{j}{n}\right) - C_n\left(\frac{i-1}{n}, \frac{j}{n}\right) - C_n\left(\frac{i}{n}, \frac{j-1}{n}\right) + C_n\left(\frac{i-1}{n}, \frac{j-1}{n}\right). \quad (\text{B-92})$$

Für den empirischen Konkordanzkoeffizienten τ_n resultiert die Gleichung:

$$\tau_n = \frac{2 \cdot n}{n-1} \sum_{i=2}^n \sum_{j=2}^n \left[C_n\left(\frac{i}{n}, \frac{j}{n}\right) \cdot C_n\left(\frac{i-1}{n}, \frac{j-1}{n}\right) - C_n\left(\frac{i-1}{n}, \frac{j}{n}\right) \cdot C_n\left(\frac{i}{n}, \frac{j-1}{n}\right) \right]. \quad (\text{B-93})$$

Gemäß Gl. (B-64) ergibt sich die folgende empirische KENDALL-Verteilungsfunktion:

$$K_n(t) = \frac{\text{Anzahl der Paare } (x_k, y_k), \text{ für deren Ränge } (i, j) \text{ } C_n(i/n, j/n) \leq t \text{ gilt}}{n}, \quad (\text{B-94})$$

wobei $0 \leq t \leq 1$ ist. Offensichtlich gilt $K_n(0) = 0$, $K_n(1) = 1$, und für $t_1 \leq t_2$ auch $K_n(t_1) \leq K_n(t_2)$.

Zu alternativen Darstellungen der Berechnung von τ_n sei auf [NEL 03] verwiesen.

B.4 Stochastische Prozesse und Differentialgleichungen

In den Abschnitten B.1 bis B.3 wurden Zufallsgrößen betrachtet, die bei "Experimenten" zufällige Wert aus einer Menge Ω der möglichen Ergebnisse annehmen. Für univariate stetige Zufallsgrößen kann Ω z. B. die Menge \mathbb{R} aller reellen Zahlen sein (oder der nicht-negativen reellen Zahlen \mathbb{R}_+ bzw. Zahlen $\mathbb{R}_{[a,b]}$ aus einem Intervall $[a, b]$). Für diskrete Zufallsgrößen kann Ω z. B. die Menge \mathbb{N} aller natürlichen Zahlen sein. Ein Ereignis ist dann eine gewisse Untermenge aus Ω , für dessen Eintreten ein Wahrscheinlichkeitsmaß bestimmt werden kann, $\Pr(X \leq x) = F_X(x|\theta)$, wobei $F_X(x|\theta)$ eine Verteilungsfunktion mit dem Verteilungsparameter θ ist (siehe Beispiele in Anhang C).

Stochastische Prozesse sind zufällige Funktionen $X(t)$, wobei die Variable t als Zeit betrachtet werden kann, $t \in [0, T]$.⁴ Eine Realisierung $x(t)$ wird als Trajektorie oder Stichprobenfunktion bezeichnet (engl.: sample path). Für einen fixierten Zeitpunkt t ist $X_t = X(t)$ eine übliche Zufallsgröße, für die eine Wahrscheinlichkeitsverteilung $\Pr(X_t \leq x) = F_{X_t}(x|\theta, t)$ existiert.

Die Zeit t wird oft nicht als kontinuierlicher Parameter, sondern als diskrete Reihe $t = t_0, t_1, \dots, t_N$ mit (i. A. konstanten) Zeitschritten $\Delta t = t_i - t_{i-1} = T/N$, i von 1 bis N) betrachtet. In diesem Fall wird von zeitdiskreten stochastischen Prozessen bzw. von stochastischen Zeitreihen X_t gesprochen.

Von praktischer Bedeutung sind stochastische Prozessen insbesondere dann, wenn sie an die Stelle üblicher Parameter von deterministischen Modellen treten, die durch eine Differentialgleichung beschrieben werden. Dann wird auch die Lösung der Differentialgleichung, welche den zeitlichen Verlauf der Zielgröße beschreibt, zu einem stochastischen Prozess. Aus gewöhnlichen Differentialgleichungen werden stochastische Differentialgleichungen (bzw. Differenzgleichungen im Fall von zeitdiskreten Prozessen). Dbzgl. Beispiele mit Bezug zu radiologisch relevanten Fragen werden in Kapitel 5 betrachtet. Nachfolgend werden einige grundlegende mathematische Begriffe und Aussagen zu stochastischen Prozessen und Differentialgleichungen erläutert, um das Verständnis dieser Beispiele und der daraus abgeleiteten Schlussfolgerungen zu erleichtern.

Zunächst wird in Abschnitt B.4.1 auf die wichtigsten zeitkontinuierlichen stochastische Prozesse eingegangen. Spezielle Aspekte, die im Vergleich zu gewöhnlichen Differentialgleichungen (ODE, ordinary differential equation) bei der Integration von stochastischen Differentialgleichungen (SDE, stochastic differential equation) zu beachten sind, werden in Abschnitt B.4.2 erläutert.

B.4.1 Zeitkontinuierliche stochastische Prozesse

Die zwei grundlegenden stochastischen Prozesse sind der Standard-WIENER-Prozess (SWP) und der elementare POISSON-Prozess (EPP). Ausführliche Definitionen und Beschreibungen dieser Prozesse findet man in den Fachbüchern zu stochastischen Differentialgleichungen (siehe z. B. [KLOE 95], [ØKS 03], [ALL 07], [JAC 10]). Der SWP und der EPP sind auch die Bausteine für die Klasse der LÉVY-Prozesse, die als Kombinationen von WIENER-Prozessen und zusammengesetzten POISSON-Prozessen dargestellt werden können. Die damit verbundene Flexibilität gibt LÉVY-Prozessen eine besondere Bedeutung für die Modellierung von stochastischen Prozessen. Bevor auf einige spezielle LÉVY-Prozesse näher eingegangen wird, sei deren allgemeine Definition erläutert.

B.4.1.1 Definition von LÉVY-Prozessen

⁴ Analog werden räumliche stochastische Prozesse als zufällige Funktionen der kartesischen Koordinaten betrachtet.

Eine prägnante Darstellung der LÉVY-Prozesse enthält [JAC 10]. Ausführlich sind sie u. a. in [SATO 99], [CONT 04], [KYP 06] und [APP 09] beschrieben.

Ein zeitkontinuierlicher LÉVY-Prozess ist

- (1) ein rechtsstetiger stochastischer Prozess $L(t)$ mit linken Grenzwerten⁵ und $L(0) = 0$, bei dem
- (2) für zwei beliebige Zeitpunkte $0 \leq t_n \leq t_{n+1}$ die Differenz $L(t_{n+1}) - L(t_n)$ die gleiche statistische Verteilung aufweist wie $L(t_{n+1} - t_n)$,
- (3) die Differenz $L(t_{n+1}) - L(t_n)$ für drei beliebige Zeitpunkte $0 \leq t_{n-1} \leq t_n \leq t_{n+1}$ unabhängig von $L(t_{n-1})$ ist, und
- (4) für beliebig kleine Werte $\varepsilon > 0$ die Wahrscheinlichkeit $\Pr(L(t + \Delta t) - L(t) > \varepsilon)$ für den Grenzwert $\Delta t \rightarrow 0$ gleich Null ist.

Die Bedingungen (2) und (3) besagen, dass die Inkremente $L(t + \Delta t) - L(t)$ von LÉVY-Prozessen iid Zufallsgrößen mit einer stationären Verteilung sind und dass sie stationäre MARKOV-Prozesse sind. Die als "stochastische Stetigkeit" bezeichnete Bedingung (4) besagt nicht, dass die Trajektorien stetige Funktionen sein müssen (so enthalten POISSON-Prozesse Sprünge an zufälligen Zeitpunkten), schließt aber Sprünge zu determinierten Zeitpunkten (z. B. "Kalendereffekte") aus. Die Definition der LÉVY-Prozesse als "càdlàg" besagt, dass für einen Zeitpunkt t , für den der Wert der realisierten Trajektorie $l(t)$ bekannt ist, die weitere Entwicklung zufällig verläuft (auch für ein infinitesimales Zeitinkrement dt nicht determiniert ist).

Zur Modellierung von realen stochastischen Prozessen sei vermerkt, dass sich LÉVY-Prozesse durch ihre MARKOV-Eigenschaft ("Gedächtnislosigkeit") oft nicht direkt eignen. Sie finden jedoch breite Anwendung als "Rauschquelle" in stochastischen Differentialgleichungen (siehe Abschnitt B.4.2),

B.4.1.2 Standard-WIENER-Prozess (SWP)

Der (Standard-) WIENER-Prozess $W(t)$ ist für $t \geq 0$ mit $W(0) = 0$ durch die Normalverteilung seines Inkrements

$$dW \sim N(0, \sqrt{dt}) \quad (\text{B-95})$$

für infinitesimale Zeitintervalle dt definiert. Aus Gl. (B-95) folgt:

$$W(t) = \int_0^t dW \quad \text{mit} \quad E[W(t)] = 0, \quad \text{Var}[W(t)] = t, \quad W(t) \sim N(0, \sqrt{t}). \quad (\text{B-96})$$

Im Unterschied zu anderen LÉVY-Prozessen ist der WIENER-Prozess stetig (links- und rechtseitige Grenzwerte stimmen für jeden Zeitpunkt t überein). Er ist jedoch nicht differenzierbar. Dies ist durch die folgende Grenzwertbetrachtung unter Beachtung von Gl. (B-96) einsichtig:

⁵ Eine Funktion $f(t)$, $[0, T] \rightarrow \mathbb{R}$, ist rechtsstetig mit linken Grenzwerten, wenn für beliebige $t \in [0, T]$ linke und rechte Grenzwerte $f_l(t) = \lim_{\delta \rightarrow 0} f(t - \delta)$ und $f_r(t) = \lim_{\delta \rightarrow 0} f(t + \delta)$ existieren und $f(t) = f_l(t)$ gilt. In der Fachliteratur wird diese Eigenschaft von Funktionen als "càdlàg" bezeichnet. Dies ist ein französisches Akronym (continue à droite, limitée à gauche).

$$\frac{dW(t)}{dt} = \lim_{\Delta t \rightarrow 0} \frac{W(t + \Delta t) - W(t)}{\Delta t} = \lim_{\Delta t \rightarrow 0} \frac{W(\Delta t)}{\Delta t} \sim \lim_{\Delta t \rightarrow 0} \frac{\sqrt{\Delta t}}{\Delta t} \rightarrow \infty. \quad (\text{B-97})$$

In Gl. (B-97) steht das Zeichen "~" für "in der Größenordnung von", wobei die letzte Aussage aus $\text{Var}[W(\Delta t)] = E[W^2(\Delta t)] = \Delta t$ folgt. Die Eigenschaft des WIENER-Prozesses, nicht differenzierbar zu sein, verlangt spezielle Methoden für nichtlineare Transformationen und die Integration, worauf in Abschnitt B.4.2 eingegangen wird. In Abbildung B-9 sind zur Veranschaulichung fünf zufällige Realisierungen $w(t)$ des SWP dargestellt. Die numerische Simulation von zufälligen Realisierungen des SWP kann sehr einfach mit folgendem Algorithmus vorgenommen werden:

1. Es wird $w(0) = 0$ gesetzt, und für den Simulationszeitraum T wird für eine (beliebige) große Anzahl I von Simulationsschritten der Zeitschritt $\Delta t = T/I$ festgelegt.
2. Es werden I normalverteilte Zufallszahlen $\Delta w_i \sim N(0, \sqrt{\Delta t})$ erzeugt.
3. Dann ergibt sich mit $w(i \cdot \Delta t) = w((i-1) \cdot \Delta t) + \Delta w_i$ für i von 1 bis I eine Trajektorie des SWP.

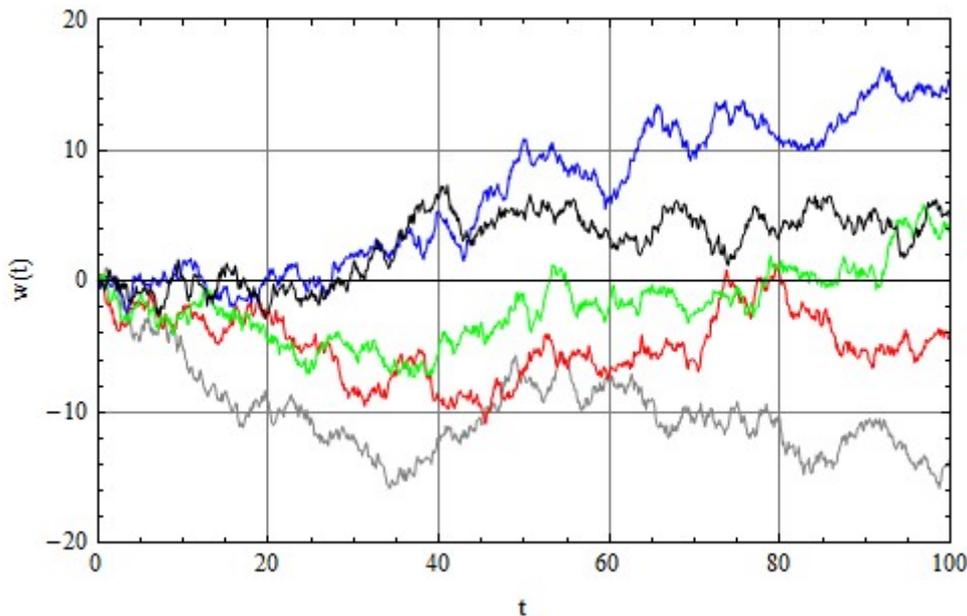


Abbildung B-9: Fünf zufällige Realisierungen des Standard-WIENER-Prozesses $W(t)$; mit Zeitschritt der numerischen Simulation von $\Delta t = 0,1$ ($T = 100$; Anzahl der Zeitschritte: $I = 1000$).

Für einen linear transformierten WIENER-Prozess gilt:

$$B(t) = \mu t + \sigma W(t) \Rightarrow B(t) \sim N(\mu t, \sigma^2 t). \quad (\text{B-98})$$

Der so aus einem SWP definierte Prozess $B(t)$ wird als verallgemeinerter WIENER-Prozess oder als BROWN'sche Bewegung mit Drift μ und Volatilität σ bezeichnet. $B(t)$ hat den Erwartungswert $\mu \cdot t$ und die Varianz $\sigma^2 \cdot t$. Die Summe von zwei SWP ist eine BROWN'sche Bewegung mit $\mu = 0$ und $\sigma^2 = 2$. Diese Additivität ist ein generelles Merkmal aller LÉVY-Prozesse (siehe Abschnitt B.4.1.5).

Der SWP ist zudem ein Martingal. Mit diesem Begriff werden stochastische Prozesse bezeichnet, für die bei einem zum Zeitpunkt u bekannten Zufallswert X_u der bzgl. bedingte Erwartungswert von $X(t)$ für zukünftige Zeitpunkte $t > u$ gleich X_u ist, d. h., $E_X[X_t | X_u] = X_u$ gilt. Für den SWP kann diese Eigenschaft unmittelbar gezeigt werden. Sie folgt aus $W_t - W_u \sim N(0, \sqrt{t-u})$:

$$E[W_t | W_u] = \int_{-\infty}^{\infty} \frac{W_t \exp\left(-\frac{(W_t - W_u)^2}{2(t-u)}\right)}{\sqrt{2\pi(t-u)}} dW_t = W_u. \quad (\text{B-99})$$

In Abbildung B-10 ist dies anhand von fünf für $t > u$ möglichen weiteren Verläufen des SWP $W(t)$ bei Kenntnis der Trajektorie bis zum Zeitpunkt u verdeutlicht, wobei der genaue Verlauf bis zum Zeitpunkt u wegen der MARKOV-Eigenschaft des SWP unerheblich ist. Wichtig ist nur der Wert W_u .

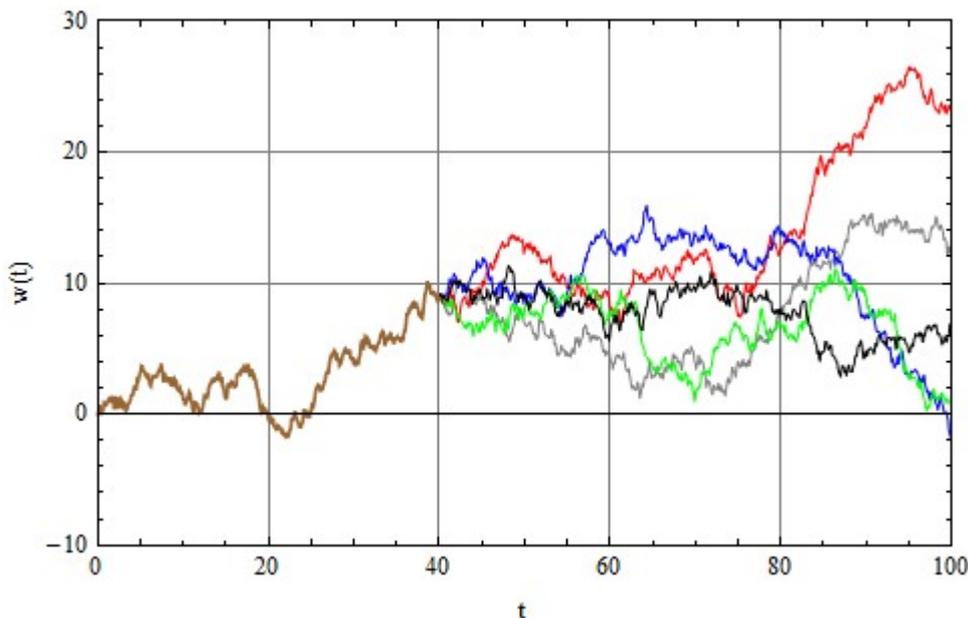


Abbildung B-10: Fünf zufällige Realisierungen des Standard-WIENER-Prozesses $W(t)$ bei bekanntem Verlauf bis zum Zeitpunkt $u = 40$ (braun); mit Zeitschritt der numerischen Simulation von $\Delta t = 0,1$ ($T = 100$; Anzahl der Zeitschritte: $I = 1000$).

B.4.1.3 Elementarer POISSON-Prozess (EPP)

Der elementare POISSON-Prozess (EPP) ist ein Zählprozess mit mittlerer Rate (Intensität) λ , dessen Zustandsgröße zum Zeitpunkt t eine zufällige natürliche Zahl $N(t)$ ist, mit der die Anzahl von Ereignissen im Zeitintervall von 0 bis t gezählt wird. Es gilt $N(0) = 0$. Für einen EPP mit Intensität λ ist die Wahrscheinlichkeit dafür, dass in einem infinitesimalen Zeitintervall dt ein Sprung erfolgt und somit $dN = N(t+dt) - N(t) = 1$ ist, unabhängig von der Zeit t und gleich $\lambda \cdot dt$. Die Wahrscheinlichkeit dafür, dass im Zeitintervall dt kein Sprung erfolgt ist dann gleich $1 - \lambda \cdot dt$, da eines der beiden sich gegenseitig ausschließenden Ereignisse ($dN = 1$ oder $dN = 0$) realisiert wird (das Auftreten von mehreren Sprüngen in einem infinitesimalen Zeitintervall ist ausgeschlossen).

Wird ein Zeitintervall $[0, t]$ in I Intervalle gleicher Länge $\Delta t = t/I$ aufgeteilt, kann die Wahrscheinlichkeit dafür, dass bis t kein Sprung erfolgt, durch folgenden Grenzübergang berechnet werden:

$$\Pr(N(t) = 0 | \lambda) = \lim_{I \rightarrow \infty} \prod_{i=1}^I (1 - \lambda \Delta t) = \lim_{I \rightarrow \infty} (1 - \lambda t/I)^I = e^{-\lambda t}.$$

Analog ergibt sich die Wahrscheinlichkeit dafür, dass n Ereignisse in der Zeit t erfolgen, wozu alle Kombinationen von Zeitintervallen mit bzw. ohne Sprung zu berücksichtigen sind:

$$\Pr(N(t) = n | \lambda) = \lim_{I \rightarrow \infty} \frac{I!}{n!(I-n)!} \cdot (\lambda t/I)^n \cdot (1 - \lambda t/I)^{I-n} = \frac{(\lambda t)^n \cdot e^{-\lambda t}}{n!}. \quad (\text{B-100})$$

Aus dem Vergleich von Gl. (B-100) mit Gl. (C-82) folgt: Die Wahrscheinlichkeit dafür, dass durch einen EPP mit dem Parameter λ im Zeitintervall $[0, t]$ genau n Sprünge erfolgen, ist durch die POISSON-Verteilung mit dem Parameter $\lambda \cdot t$ bestimmt:

$$\Pr(N(t) = n | \lambda) = f_{\text{Poi}}(n | \lambda t). \quad (\text{B-101})$$

Ist $N(t)$ ein EPP mit einer Intensität λ , so gilt dementsprechend (siehe Gl. (C-83)):

$$E[N(t)] = \lambda t, \quad \text{Var}[N(t)] = \lambda t. \quad (\text{B-102})$$

Für die Ableitung einer einfachen Darstellung des EPP

$$N(t) = \int_0^t dN \quad \text{mit} \quad \Pr(dN = 1 | \lambda) = \lambda \cdot dt \quad \text{und} \quad \Pr(dN = 0 | \lambda) = 1 - \lambda \cdot dt, \quad (\text{B-103})$$

aus der sich auch ein effektiver Algorithmus zur Simulation dieses stochastischen Prozesses ergibt, wird die Tatsache genutzt, dass die Zeit τ zwischen zwei Ereignissen eine Zufallsgröße ist, die einer Exponentialverteilung mit dem gleichen Parameter λ folgt: $\tau \sim \text{Ex}(\lambda)$.

Anmerkung: Mit diesem Ansatz kann auch Gl. (B-100) auf einfache Weise abgeleitet werden. Es seien τ_i ($1 \leq i \leq k$) k iid Zufallsgrößen der Exponentialverteilung $\text{Ex}(\lambda)$. Ihre Summe

$$T_k = \sum_{i=1}^k \tau_i, \quad \tau_i \sim \text{Ex}(\lambda) \quad (\text{B-104})$$

ist dann ebenfalls eine Zufallsgröße, welche die Zeit bis zum Eintreten des k -ten POISSON-Ereignisses angibt. Sie genügt der Gammaverteilung⁶ $T_k \sim \text{Gam}(k, \lambda)$ (siehe Abschnitte C.3.3, C.3.4). Damit kann die Wahrscheinlichkeit $\Pr(N(t) = n | \lambda)$ durch eine elementare partielle Integration bestimmt werden:

$$\Pr(N(t) = n | \lambda) = \Pr(T_n < t < T_{n+1} | \lambda) = \int_0^t f_{\text{Gam}}(T | n, \lambda) dT - \int_0^t f_{\text{Gam}}(T | n+1, \lambda) dT, \quad (\text{B-105})$$

wobei $f_{\text{Gam}}(T | \alpha, \lambda)$ die Wahrscheinlichkeitsdichte der Gammaverteilung $\text{Gam}(\alpha, \lambda)$ ist; siehe Gl. (C-26). Eine partielle Integration des zweiten Integrals auf der rechten Seite von Gl. (B-104) ergibt

⁶ Da der Formparameter α der Gammaverteilung $\text{Gam}(\alpha, \lambda)$ in diesem Fall eine natürliche Zahl ist, spricht man auch von einer ERLANG-Verteilung.

$$\begin{aligned} \int_0^t f_{\text{Gam}}(T | n+1, \lambda) dT &= \int_0^t \frac{\lambda^{n+1} \cdot T^n \cdot e^{-\lambda T}}{\Gamma(n+1)} dT = -\frac{\lambda^n \cdot t^n \cdot e^{-\lambda t}}{\Gamma(n+1)} + \int_0^t \frac{\lambda^n \cdot T^{n-1} \cdot e^{-\lambda T}}{\Gamma(n)} dT \\ &= -\frac{(\lambda t)^n \cdot e^{-\lambda t}}{n!} + \int_0^t f_{\text{Gam}}(T | n, \lambda) dT \end{aligned}$$

Setzt man diese Beziehung in Gl. (B-105) ein, so folgt:

$$\Pr(N(t) = n | \lambda) = \frac{(\lambda t)^n \cdot e^{-\lambda t}}{n!}. \quad (\text{B-100}')$$

Mit der in Gl. (B-104) definierten Zufallsgröße T_k (Zeitpunkt des Eintretens des k -ten POISSON-Ereignisses) lässt sich der Prozess $N(t)$, der die Anzahl dieser Ereignisse zählt, wie folgt darstellen:

$$N(t) = \sum_{k \geq 1} 1_{t \geq T_k}(t), \quad (\text{B-106})$$

wobei $1_B(t)$ die Indikatorfunktion bezeichnet, die auf dem im Index angegebene Bereich B den Wert 1 annimmt und außerhalb dieses Bereichs gleich 0 ist,

$$1_B(t) = \begin{cases} 1 & \text{für } t \in B \\ 0 & \text{für } t \notin B \end{cases}. \quad (\text{B-107})$$

Man kann zeigen, dass der mit Gl. (B-104)/(B-106) dargestellte POISSON-Prozess ein stationärer LÉVY-Prozess ist (siehe z. B. [CONT 04]).

Die Summe von zwei unabhängigen POISSON-Prozessen $N_1(t)$ und $N_2(t)$ mit Intensitäten von λ_1 und λ_2 ist ein POISSON-Prozess $N(t)$ mit einer Intensität von $(\lambda_1 + \lambda_2)$. Dies kann durch Faltung der Wahrscheinlichkeitsmassenfunktionen (analog zu Gl. (B-23)) gezeigt werden:

$$\begin{aligned} f_{\text{Poi}}(N_1 | \lambda_1 t) * f_{\text{Poi}}(N_2 | \lambda_2 t) &= \sum_{k=0}^n \frac{(\lambda_1 t)^k \cdot e^{-\lambda_1 t}}{k!} \cdot \frac{(\lambda_2 t)^{n-k} \cdot e^{-\lambda_2 t}}{(n-k)!} \\ &= \frac{(\lambda_1 + \lambda_2)^n t^n \cdot e^{-(\lambda_1 + \lambda_2)t}}{n!} = f_{\text{Poi}}(N | (\lambda_1 + \lambda_2)t) \end{aligned}, \quad (\text{B-108})$$

wobei "*" das übliche mathematische Symbol für die Operation der Faltung ist. Noch einfacher kann diese Eigenschaft der POISSON-Verteilung mit Hilfe ihrer charakteristischen Funktion gezeigt werden (siehe Abschnitt B.4.1.5).

Die numerische Simulation zufälliger Realisierungen $n(t)$ des EPP kann nach Gl. (B-104)/(B-106) mit folgendem Algorithmus vorgenommen werden:

1. Es wird $n(0) = 0$ gesetzt und ein Simulationszeitraum T festgelegt.
2. Es werden nacheinander exponentiell verteilte Zufallszahlen τ_i erzeugt, deren Summe gemäß Gl. (B-104) die Zeitpunkte T_k bestimmt, zu denen (zufällig) POISSON-Ereignisse stattfinden. Entsprechend wird iterativ der der Zähler $n(T_k) = n(T_{k-1}) + 1$ erhöht.
3. Wenn $T_{k'} > T$ wird, endet die Simulation; es wird $n(T) = n(T_{k'-1})$ gesetzt.

Der EPP ist kein Martingal. Durch Subtraktion des Erwartungswertes $\lambda \cdot t$ kann jedoch ein Martingal erzeugt werden, das als kompensierter POISSON-Prozess bezeichnet wird,

$$\tilde{N}(t | \lambda) = N(t | \lambda) - \lambda t. \quad (\text{B-109})$$

Abbildung B-11 zeigt zur Veranschaulichung fünf zufällige Realisierungen $n(t)$ des EPP mit einer Intensität $\lambda = 0,02$ für einem Zeitraum von $T = 1000$ und die zugehörigen kompensierten POISSON-Prozesse.

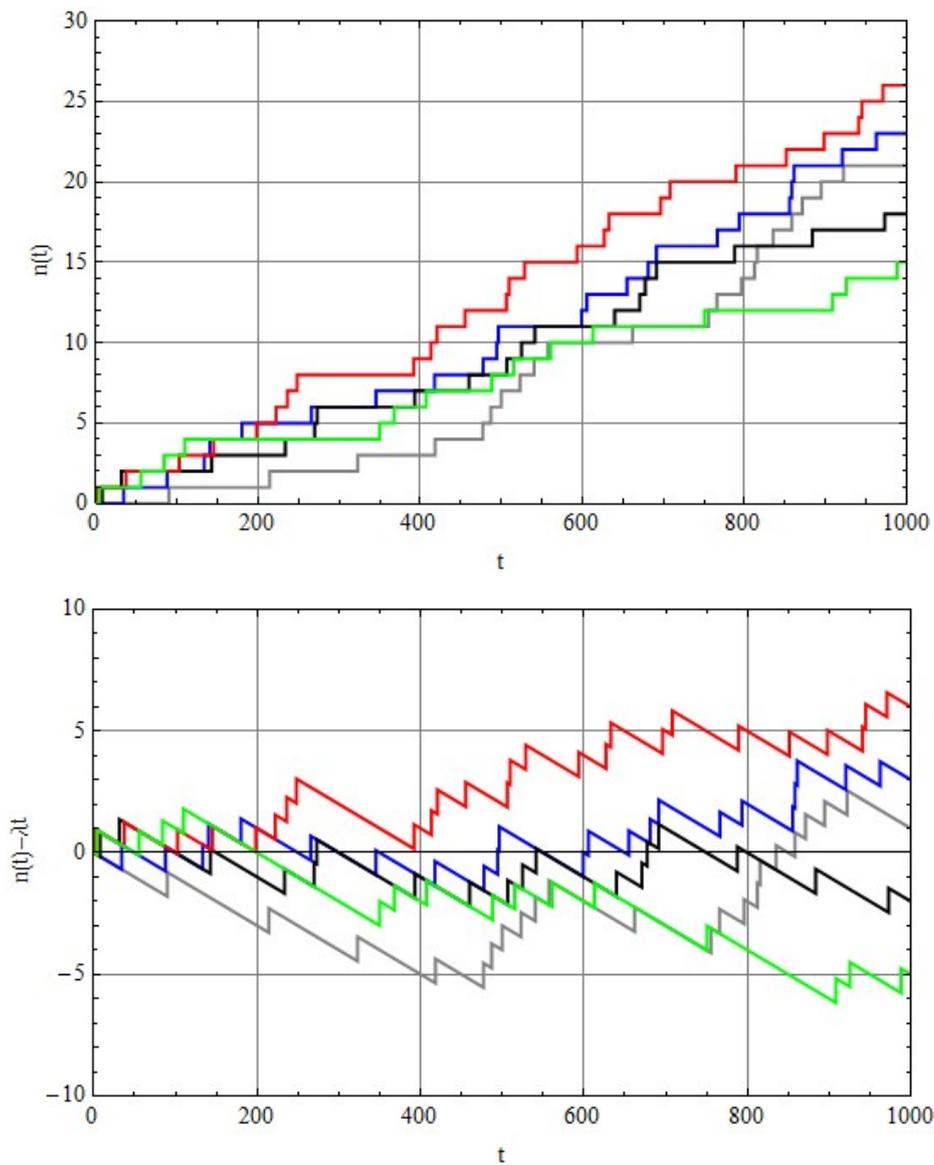


Abbildung B-11: Fünf zufällige Realisierungen des elementaren POISSON-Prozesses $N(t)$ mit Intensität von $\lambda = 0,02$ für einen Zeitraum von $T = 1000$ (oberes Bild) und zugehörige nach Gl. (B-109) kompensierte POISSON-Prozesse (unteres Bild)

B.4.1.4 Zusammengesetzte POISSON-Prozesse (ZPP)

Für Anwendungen haben neben dem EPP so genannte zusammengesetzte POISSON-Prozesse (ZPP, englisch: compound POISSON process) eine große Bedeutung, weil mit ihnen die Modellierung von stochastisch eintretenden Prozessen, deren Auswirkungen durch Zufallsgröße darzustellen sind, flexibel an die jeweiligen Gegebenheiten angepasst werden können. Ein EPP $N(t)$ ist durch zwei Eigenschaften geprägt:

- (1) Die Wahrscheinlichkeit des Eintretens von Ereignissen ist durch die Intensität λ bestimmt.
- (2) Bei Eintreten eines Ereignisses besteht dessen Auswirkung in einer Erhöhung des "Zählers" $N(t)$ um den Wert 1; $dN(T_k) = 1$, wenn T_k der (zufällige) Zeitpunkt eines Ereignisses ist.

Mit Hilfe von ZPP können diese beiden Merkmale von EPP verallgemeinert werden, wobei statt der Sprunghöhe 1 des EPP nun Sprunghöhen definiert werden, die einer probabilistischen Verteilung folgen. Außerdem können Prozesse mit unterschiedlicher Intensität überlagert werden.

Zunächst sei ein ZPP $R(t)$ betrachtet, der auf einem EPP $N(t)$ mit festem λ basiert. Es sei Z eine Zufallsgröße mit Wahrscheinlichkeitsdichte $f_Z(z|\theta)$. Sind Z_k (k von 1 bis $N(t)$) iid Zufallswerte aus dieser Verteilung, dann wird auf Basis von Gl. (B-104)/(B-106) des EPP $N(t) \sim \text{Poi}(\lambda \cdot t)$ mit

$$R(t) = \sum_{k \geq 1} Z_k \cdot 1_{t \geq T_k}(t) = \sum_{k=1}^{N(t)} Z_k, \quad R(0) = 0, \quad (\text{B-110})$$

ein ZPP definiert, für den die Zeitpunkte des Eintretens von Ereignissen durch den EPP $N(t)$ bestimmt werden, die Wirkung eines Ereignisses jedoch zufällig ist und durch die Verteilung der Zufallsgröße Z bestimmt wird. Dabei sind die Höhen Z_k der einzelnen Sprünge unabhängig von $N(t)$, d. h. unabhängig vom Verlauf des zugrunde liegenden EPP.

Für den Erwartungswert und die Varianz der Zufallsgröße $R(t)$ folgt aus Gl. (B-110) mit den in Abschnitt B.1.4 für Funktionen von zwei Zufallsgrößen dargelegten Beziehungen Gl. (B-28b) bzw. Gl. (B-32a) und Beachtung der Gleichheit von Erwartungswert und Varianz des EPP, Gl. (B-102):

$$E_R[R(t)] = E_{\text{Poi}}[E_{R|N(t)}[R(t)]] = E_{\text{Poi}}[N(t) \cdot E_Z[Z]] = E_{\text{Poi}}[N(t)] \cdot E_Z[Z] = \lambda t E_Z[Z], \quad (\text{B-111})$$

$$\begin{aligned} \text{Var}_R[R(t)] &= E_{\text{Poi}}[\text{Var}_{R|N(t)}[R(t)] + \text{Var}_{\text{Poi}}[E_{R|N(t)}[R(t)]]] \\ &= E_{\text{Poi}}[N(t) \cdot \text{Var}_Z[Z] + \text{Var}_{\text{Poi}}[N(t)] \cdot E_Z[Z]^2] \\ &= E_{\text{Poi}}[N(t)] \cdot \text{Var}_Z[Z] + (E_Z[Z])^2 \cdot \text{Var}_{\text{Poi}}[N(t)] \\ &= \lambda t \cdot (\text{Var}_Z[Z] + (E_Z[Z])^2) = \lambda t \cdot E_Z[Z^2] \end{aligned} \quad (\text{B-112})$$

Für einen ZPP, bei dem die Wirkung bei Eintreten eines Ereignisses z. B. durch eine Gamma-Verteilung bestimmt ist, $Z \sim \text{Ga}(\alpha, \beta)$, resultieren aus Gl. (B-111) und (B-112) sowie den bekannten Beziehungen für den Erwartungswert und die Varianz von Zufallsgrößen mit Gammaverteilung (siehe Abschnitt C.3.4) die Beziehungen

$$E[R(t)] = \alpha \beta \lambda t \quad \text{und} \quad \text{Var}[R(t)] = \alpha (1 + \alpha) \beta^2 \lambda t.$$

Für eine Beispielrechnung zu diesem ZPP sei $\alpha = 2$ und $\beta = 0,5$. Für den Parameter des EPP wird mit $\lambda = 0,02$ gerechnet. Die Parameter der Gammaverteilung wurden für den Vergleich der Prozesse $N(t)$ und $R(t)$ so gewählt, dass deren Erwartungswerte wegen $\alpha \cdot \beta = 1$ übereinstimmen. Die Varianz von $R(t)$ ist jedoch um den Faktor $\alpha \cdot (1 + \alpha) \cdot \beta^2 = 1,5$ größer als die Varianz von $N(t)$.

In Abbildung B-12 sind zweimal je 5 zufällige Realisierungen $n(t)$ des EPP $N(t)$ und darauf basierende zufällige Realisierungen $r(t)$ des ZPP $R(t)$ dargestellt. Die "Aufweitung" der Pfade des ZPP durch die Variabilität der einzelnen Sprunghöhen (gemäß Gammaverteilung von Z gegenüber den Sprunghöhen von $dN = 1$ des EPP) ist daraus deutlich erkennbar.

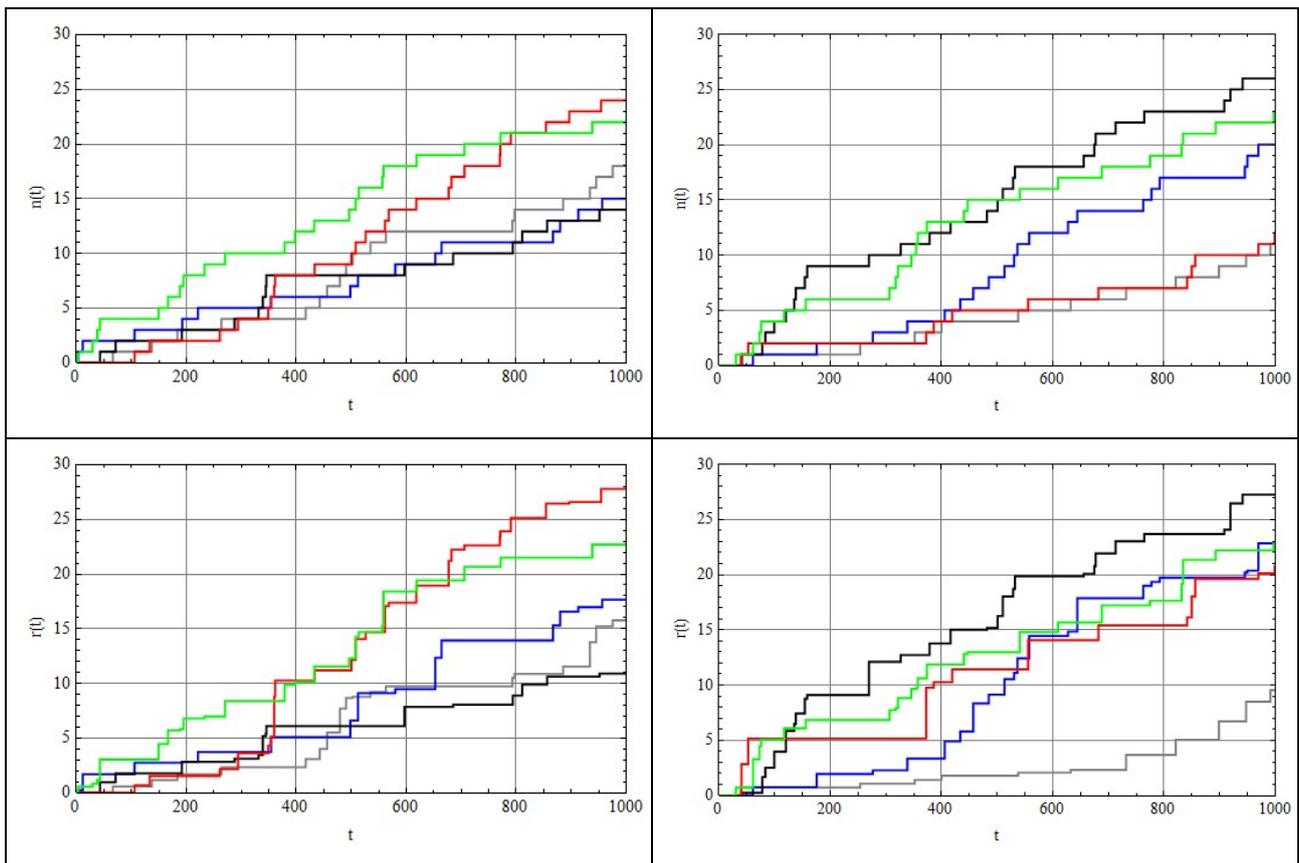


Abbildung B-12: Realisierungen des EPP $N(t)$ mit $\lambda = 0,02$ (obere Bilder) und daraus gemäß Gl. (B-110) für den ZPP $R(t)$ mit $Z \sim \text{Ga}(\alpha, \beta)$, $\alpha = 2$, $\beta = 0,5$ resultierende Verläufe (untere Bilder)

Zur weiteren Spezifizierung von zusammengesetzten POISSON-Prozessen wird die Darstellung von Zufallsgrößen bzw. von stochastischen Prozessen durch ihre charakteristischen Funktionen benötigt. Ist X eine Zufallsgröße mit Wahrscheinlichkeitsdichte $f_X(x|\theta)$, so ist ihre charakteristische Funktion $\phi_X(s|\theta)$ als Erwartungswert von $e^{i \cdot s \cdot X}$ ($i = \sqrt{-1}$) definiert:

$$\phi_X(s | \boldsymbol{\theta}) = E_X[e^{isX}] = \int_{-\infty}^{\infty} e^{isx} \cdot f_X(x | \boldsymbol{\theta}) dx. \quad (\text{B-113})$$

Ist X eine diskrete Zufallsgröße, so ist in Gl. (B-113) das Integral durch eine Summe zu ersetzen, wobei $f_X(x|\boldsymbol{\theta})$ dann die Wahrscheinlichkeitsmassenfunktion bezeichnet.

Für den elementaren POISSON-Prozess $N(t)$ resultiert mit Gl. (B-100) die charakteristische Funktion

$$\begin{aligned} \phi_N(s | \lambda) &= E_{\text{Poi}}[e^{isN(t)}] = \sum_{n=0}^{\infty} e^{isn} \cdot \frac{(\lambda t)^n e^{-\lambda t}}{n!} = e^{-\lambda t} \cdot \sum_{n=0}^{\infty} \frac{(\lambda t e^{is})^n}{n!} \\ &= e^{-\lambda t} \cdot \exp(\lambda t e^{is}) = \exp(\lambda t \cdot (e^{is} - 1)) \end{aligned} \quad (\text{B-114})$$

Zur Berechnung der charakteristischen Funktion des mit Gl. (B-110) definierten ZPP wird von der Tatsache Gebrauch gemacht, dass das Faltungsintegral Gl. (B-23), mit dem die Wahrscheinlichkeitsdichte der Summe von zwei unabhängigen Zufallsgrößen bestimmt werden kann, dem Produkt ihrer charakteristischen Funktionen entspricht. Für die Summe aus n iid Zufallsgrößen Z_k

$$X = \sum_{k=1}^n Z_k, \quad Z_k \text{ sind iid mit der charakteristischen Funktion } \phi_Z(s|\boldsymbol{\theta}), \quad (\text{B-115})$$

ergibt sich dann die charakteristische Funktion

$$\phi_X(s | \boldsymbol{\theta}) = (\phi_Z(s | \boldsymbol{\theta}))^n. \quad (\text{B-116})$$

Zur charakteristischen Funktion des in Gl. (B-110) definierten ZPP $R(t)$ resultiert nach Gl. (B-113) unter Beachtung der Beziehung (B-28b) folgende Darstellung:

$$\begin{aligned} \phi_{R(t)}(s | \boldsymbol{\theta}, \lambda) &= E_R[e^{isR(t)}] = E_{\text{Poi}}[E_{R(t)|N(t)}[e^{isR(t)}]] = E_{\text{Poi}}[(\phi_Z(s | \boldsymbol{\theta}))^{N(t)}] \\ &= \sum_{n=0}^{\infty} \frac{(\phi_Z(s | \boldsymbol{\theta}))^n (\lambda t)^n e^{-\lambda t}}{n!} = \exp\{\lambda t (\phi_Z(s | \boldsymbol{\theta}) - 1)\} \\ &= \exp\left\{\lambda t \int_{-\infty}^{+\infty} (e^{isz} - 1) f_Z(z | \boldsymbol{\theta}) dz\right\} \end{aligned} \quad (\text{B-117})$$

Es sei angemerkt, dass für $f_Z(z|\boldsymbol{\theta}) = \delta(z-1)$ aus Gl. (B-117) durch Integration über die DIRAC'sche δ -Funktion wieder die charakteristische Funktion des EPP Gl. (B-114) resultiert.

B.4.1.5 Darstellung von LÉVY-Prozessen

Überlagert man nun viele voneinander unabhängige zusammengesetzte POISSON-Prozesse, deren charakteristische Funktionen jeweils Darstellungen entsprechend Gl. (B-117) haben, die jedoch unterschiedliche Intensitäten λ_j mit Verteilungsfunktionen $f_j(z)$ für die Sprunghöhen aufweisen (mit j von 1 bis J), dann ergibt sich für diesen komplexen Prozess die charakteristische Funktion aus dem Produkt der nach Gl. (B-117) beschriebenen ZPP, wobei $\lambda \cdot f_Z(z)$ durch die Summe der Produkte $\lambda_j \cdot f_j(z)$ über j von 1 bis J zu ersetzen ist. Eine Verallgemeinerung dieser Konstruktion erfolgt durch

Einführung einer Funktion $\lambda(z)$, die ein Kontinuum der Überlagerung von ZPP beschreibt. Daraus resultiert folgende Darstellung für stochastische Sprungprozesse:

$$\phi_{J(t)}(s) = \exp \left\{ t \int_{-\infty}^{+\infty} (e^{is z} - 1) \lambda(z) dz \right\}, \quad (\text{B-118})$$

(der Index J steht für "Jump Process") Es sei vermerkt, dass $\lambda(z)$ keine Wahrscheinlichkeitsdichte ist. Der EPP entspräche z. B. der Funktion $\lambda(z) = \lambda \cdot \delta(z-1)$, deren Integral gleich λ ist, das beliebige positive Werte haben kann.

Die Sprungratenfunktion $\lambda(z)$ muss bestimmten Anforderungen genügen, damit der Sprung-Prozess $J(t)$ aus Gl. (B-118) einem stochastischen Prozess mit den Eigenschaften von LÉVY-Prozessen entspricht. Zunächst muss das Integral von $\lambda(z)$ über den Bereich großer Sprunghöhen z einen endlichen Wert haben, weil der Prozess $J(t)$ sonst umgehend ins Unendliche divergieren würde. LÉVY-Prozesse können jedoch unbeschränkt große Sprungraten aufweisen, wenn die zugehörigen Sprunghöhen gegen Null gehen. Hierzu wird gefordert, dass das Integral von $z^2 \cdot \lambda(z)$ über kleine z -Werte (i. A. wird hierzu der Bereich $|z| < 1$ festgelegt) ebenfalls endlich ist. Die beiden Forderungen an $\lambda(z)$ sind dann (siehe z. B. [SATO 99], [CONT 04], [KYP 06], [APP 09], [JAC 10]):

$$\int_{|z|>1} \lambda(z) dz < \infty \quad \text{und} \quad \int_{|z|<1} z^2 \lambda(z) dz < \infty. \quad (\text{B-119})$$

Zerlegt man den unter dem Integral in Gl. (B-118) stehenden Ausdruck ($e^{is z} - 1$) in eine TAYLOR-Reihe, ($e^{is z} - 1$) = $i \cdot s \cdot z - s^2 \cdot z^2 / 2 - i \cdot s^3 \cdot z^3 / 3! + \dots$, so muss der Term $i \cdot s \cdot z$ für den Fall $\lambda(z) \rightarrow \infty$ bei $z \rightarrow 0$ speziell betrachtet werden. Wenn

$$\int_{|z|<1} z \lambda(z) dz = a < \infty. \quad (\text{B-120})$$

gilt, so resultiert daraus nach Gl. (B-118) für die charakteristische Funktion $\phi_{J(t)}(s)$ ein Faktor $e^{i \cdot t \cdot s \cdot a}$, der dem Drift-Term der BROWN'schen Bewegung $B(t)$ entspricht. Deren charakteristische Funktion $\phi_{B(t)}(s | \mu, \sigma)$ lässt sich mit Gl. (B-98) nach Gl. (B-113) leicht berechnen:

$$\phi_{B(t)}(s | \mu, \sigma) = \int_{-\infty}^{\infty} e^{is x} \cdot \frac{1}{\sqrt{2 \pi t \sigma}} \exp \left(- \frac{(x - \mu t)^2}{2 \sigma^2 t} \right) dx = e^{t (i \mu s - \frac{\sigma^2 s^2}{2})}. \quad (\text{B-121})$$

Funktionen $\lambda(z)$, die Gl. (B-120) nicht erfüllen, müssen ausgeschlossen werden. Damit resultiert aus Gl. (B-118) und Gl. (B-121) folgende generelle Darstellung für LÉVY-Prozesse $L(t)$, die immer als Superposition einer BROWN'schen Bewegung mit Sprungprozessen, die auf zusammengesetzten POISSON-Prozessen basieren, dargestellt werden können (LÉVY-KHINCHIN-Darstellung):

$$L(t) = B(t) + J(t),$$

$$\phi_{L(t)}(s) = \exp \left\{ t \left(i \mu s - \frac{\sigma^2 s^2}{2} \right) + t \left(\int_{|z| \leq 1} (e^{is z} - 1 - i z s) \lambda(z) dz + \int_{|z| > 1} (e^{is z} - 1) \lambda(z) dz \right) \right\}. \quad (\text{B-122})$$

Die Eignung univariater Verteilungen zur Modellierung von stochastischen Prozessen, welche die Eigenschaften von LÉVY-Prozesse haben, hängt von ihrer unendlichen Teilbarkeit ab [SATO 99]. Diese Eigenschaft lässt sich gut anhand der charakteristischen Funktion $\phi_X(s|\theta)$ einer Zufallsgröße X erkennen. Gemäß Gl. (B-115) und (B-116) ist die charakteristische Funktion der Summe Z von n iid Zufallsgrößen X gleich der n -ten Potenz der charakteristischen Funktion von X . Die Verteilung einer Zufallsgröße X heißt unendlich teilbar, wenn für beliebige natürliche Zahlen $n = 1, 2, \dots$ eine Folge von iid Zufallsgrößen $X_{n,1}, X_{n,2}, \dots, X_{n,k}, \dots, X_{n,n}$ existiert, deren Summe von $k = 1$ bis n wie X verteilt ist. Für die charakteristische Funktion ϕ_X bedeutet dies, dass für jedes $n > 0$ eine charakteristische Funktion ϕ_{X_n} existiert, für die $\phi_X = (\phi_{X_n})^n$ gilt. In Tabelle B-3 sind vier Beispiele unendlich teilbarer Verteilungen aufgeführt, die außerdem reproduktiv sind, was bedeutet, dass die charakteristischen Funktionen ϕ_{X_n} auch zur gleichen Verteilungsfamilie gehören wie ϕ_X .

Tabelle B-3: Vier Beispiele für unendlich teilbare Verteilungen univariater Zufallsgrößen

Verteilung	$f_X(x \theta)$	$\phi_X(s \theta)$	θ_n in $\phi_{X_n}(s \theta_n)$
Normalverteilung $X \sim N(\mu, \sigma)$	$\frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]$	$\exp\left[is\mu + \frac{s^2\sigma^2}{2}\right]$	$\mu_n = \mu/n$ $\sigma_n = \sigma/\sqrt{n}$
Gammaverteilung $X \sim \text{Ga}(\alpha, \beta)$	$\frac{x^{\alpha-1}}{\beta^\alpha \Gamma(\alpha)} \cdot \exp(-x/\beta)$	$\frac{1}{(1-is\beta)^\alpha}$	$\alpha_n = \alpha/n$ $\beta_n = \beta$
WALD-Verteilung $X \sim \text{Wa}(\mu, \lambda)$	$\sqrt{\frac{\lambda}{2\pi x^3}} \cdot \exp\left[\frac{-\lambda(x-\mu)^2}{2\mu^2 x}\right]$	$\exp\left[\frac{\lambda}{\mu} \cdot \left(1 - \sqrt{1 - \frac{2is\mu^2}{\lambda}}\right)\right]$	$\mu_n = \mu/n$ $\lambda_n = \lambda/n^2$
POISSON-Verteilung $X \sim \text{Poi}(\lambda)$	$\frac{e^{-\lambda} \lambda^x}{x!}$	$\exp[(e^{is} - 1) \cdot \lambda]$	$\lambda_n = \lambda/n$

Neben den in Tabelle B-3 genannten Verteilungen gibt es weitere, die unendlich teilbar und zudem reproduktiv sind, z. B. die Negative Binomialverteilung. Unendlich teilbar aber nicht reproduktiv ist z. B. die Exponentialverteilung $X \sim \text{Ex}[\lambda]$, die sich aus der Gammaverteilung mit $\alpha = 1$ ergibt (siehe Abschnitt C.3.3). Ihre charakteristische Funktion ist $1/(1-i\cdot s/\lambda)$. Durch Potenzierung mit $1/n$ resultiert daraus für X_n eine charakteristische Funktion ϕ_{X_n} , die der Gammaverteilung mit $\beta_n = 1/\lambda$ und $\alpha_n = 1/n$ entspricht, also zu einer anderen Verteilungsfamilie gehört.

B.4.1.6 Gammaprozess

Ähnlich wie auf Basis der Normalverteilung und der POISSON-Verteilung kann auch ein auf der Gammaverteilung basierender stochastischer Prozess definiert werden. Aus der letzten Spalte von Tabelle B-3 ist ersichtlich, dass dieser Prozess, der hier mit $G(t)$ bezeichnet sei, wie folgt definiert werden kann:

$$dG \sim \text{Ga}(dt \cdot \alpha, \beta), \quad G(t) = \int_0^t dG, \quad E[G(t)] = \alpha \cdot \beta \cdot t, \quad \text{Var}[G(t)] = \alpha \cdot \beta^2 \cdot t. \quad (\text{B-123})$$

Der Gammaprozess kann entsprechend Gl. (B-118) als ein komplexer stochastischer Sprungprozess mit folgender Sprungratenfunktion dargestellt werden:

$$\lambda_{\text{Ga}}(z) = \frac{\alpha}{z} \cdot e^{-z/\beta} \cdot 1_{z \geq 0}, \quad \alpha > 0, \beta > 0. \quad (\text{B-124})$$

Die beiden Bedingungen aus Gl. (B-119) werden durch die Funktion $\lambda_{\text{Ga}}(z)$ erfüllt:

$$\int_1^{\infty} \lambda_{\text{Ga}}(z) dz = \alpha \cdot \Gamma(0, 1/\beta) < \infty \quad \text{und} \quad \int_0^1 z^2 \lambda_{\text{Ga}}(z) dz = \alpha \beta \cdot (\beta - (1 + \beta) e^{-1/\beta}) < \infty.$$

Da auch die Bedingung Gl. (B-120) erfüllt ist,

$$\int_0^1 z \lambda_{\text{Ga}}(z) dz = \alpha \beta \cdot (1 - e^{-1/\beta}) < \infty, \quad (\text{B-125})$$

ist die in der allgemeinen LÉVY-KHINCHIN-Darstellung Gl. (B-122) enthaltene spezielle Behandlung des Bereiches sehr kleiner Sprunghöhen ($|z| < 1$) mit unbegrenzt hohen Sprungraten ($\lambda(z) \rightarrow \infty$ für $z \rightarrow 0$) nicht erforderlich. Zur Berechnung der charakteristischen Funktion des Gammaprozesses kann die Darstellung nach Gl. (B-118) direkt angewendet werden⁷:

$$\phi_{\text{Ga}(t)}(s) = \exp \left\{ t \int_{-\infty}^{+\infty} (e^{i s z} - 1) \lambda_{\text{Ga}}(z) dz \right\} = (1 - i s \beta)^{-\alpha t}. \quad (\text{B-126})$$

Die Wahrscheinlichkeitsdichte des Gammaprozesses $G(t)$ ergibt sich dann durch Ausführung der inversen FOURIER-Transformation von Gl. (B-126):

$$f_{\text{Ga}}(g | t, \alpha, \beta) = \frac{g^{\alpha t - 1} \cdot e^{-g/\beta}}{\beta^{\alpha t} \cdot \Gamma(\alpha t)}. \quad (\text{B-127})$$

Der mit Gl. (B-123) definierte Gammaprozess ist also eine Überlagerung von zusammengesetzten POISSON-Prozessen mit der in Gl. (B-124) angegebenen Sprungratenfunktion. In Abbildung B-13 sind die Verläufe der Funktion $\lambda_{\text{Ga}}(z)$ für zwei Konstellationen der Verteilungsparameter α und β dargestellt, die jeweils die gleiche Varianz von $\text{Var}[G(t)] = t$ ergeben, d. h., für die $\alpha \cdot \beta^2 = 1$ ist. Die Erwartungswerte dieser beiden Prozesse, $E[G(t)] = \alpha \cdot \beta \cdot t$, müssen dementsprechend unterschiedlich sein. Ähnlich wie beim POISSON-Prozess $N(t)$ können jedoch auch "kompensierte Gammaprozesse" definiert werden:

$$\tilde{G}(t | \alpha, \beta) = G(t | \alpha, \beta) - \alpha \beta t. \quad (\text{B-128})$$

Der Erwartungswert eines kompensierten Gammaprozesses ist gleich Null, die Varianz entspricht für jeden fixierten Zeitpunkt der Varianz des zugrunde liegenden Gammaprozesses $G(t)$. In Abbildung B-14 sind zum Vergleich jeweils fünf zufällige Realisierungen von kompensierten Gammaprozessen für die Verteilungsparameter $(\alpha, \beta) = (1, 1)$ bzw. $(\alpha, \beta) = (4, 0,5)$ dargestellt.

⁷ Diese ist äquivalent zur LÉVY-KHINCHIN-Darstellung Gl. (B-122) mit einem Drift-Term μ nach Gl. (B-125).

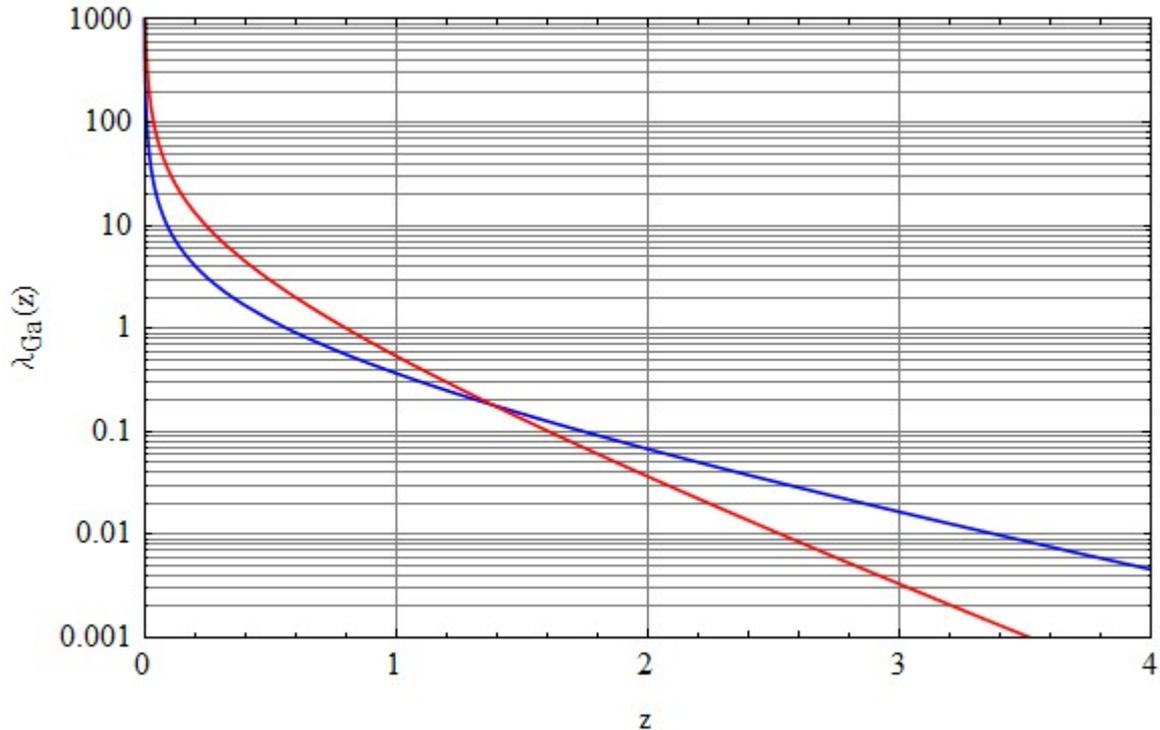


Abbildung B-13: Vergleich von zwei Sprungratenfunktionen $\lambda_{\text{Ga}}(z|\alpha,\beta)$, Gl. (B-124), für Gammaprozesse mit gleicher Varianz; blau: $(\alpha,\beta) = (1, 1)$; rot: $(\alpha,\beta) = (4, 0,5)$

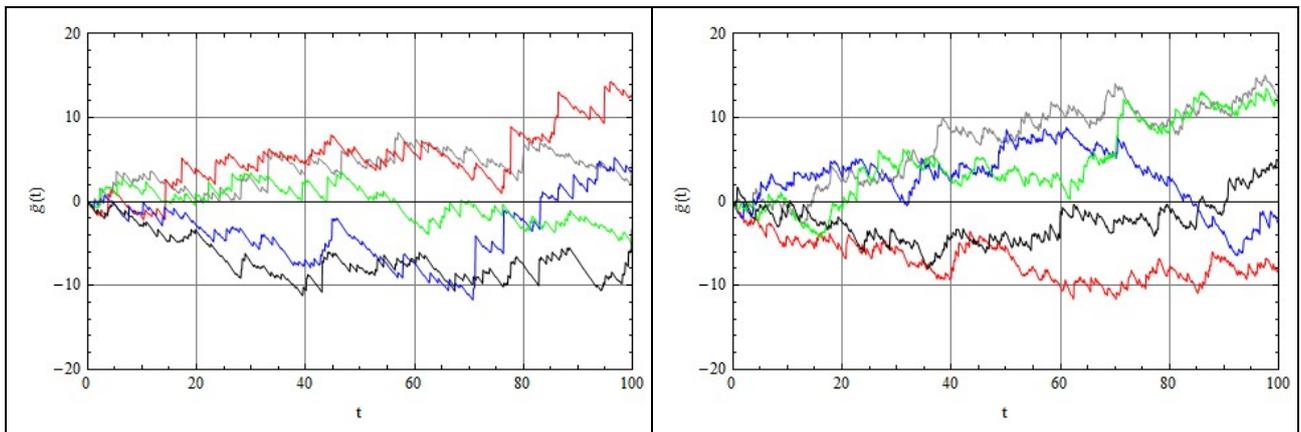


Abbildung B-14: Gegenüberstellung von je fünf zufälligen Realisierungen kompensierter Gammaprozesse mit gleicher Varianz (nach Gl. (B-128)); links: $(\alpha,\beta) = (1, 1)$; rechts: $(\alpha,\beta) = (4, 0,5)$; mit Zeitschritten der numerischen Simulation von $\Delta t = 0,1$ für Zeitraum von $T = 100$;

Aus Abbildung B-14 ist gut erkennbar, dass der kompensierte Gammaprozess mit den Parametern $(\alpha,\beta) = (1, 1)$ einen höheren Anteil größerer Sprünge aufweist als der zweite, mit $(\alpha,\beta) = (4, 0,5)$ definierte Prozess. Dieser hat schon relativ große Ähnlichkeit zum SWP (vgl. mit Abbildung B-9).

B.4.2 Stochastische Differentialgleichungen (SDE)

Genügt eine Variable $x(t)$ einer gewöhnlichen Differentialgleichung (ODE, ordinary differential equation) mit Parametern, die als Konstanten oder als deterministische zeitabhängige Funktionen definiert sind, so kann bei gegebenen Anfangsbedingungen für jeden zukünftigen Zeitpunkt t der Wert der Funktion $x(t)$ berechnet werden. Sind dagegen bestimmte Parameter der Differentialgleichung stochastische Prozesse, so wird die Variable $x(t)$ eine Zufallsfunktion $X(t)$, die für fixierte Zeitpunkte t eine Zufallsgröße ist⁸. Werden zufällige Schwankungen stochastischer Parameter auch für infinitesimale Zeitintervalle dt zugelassen, wie bei den in Abschnitt B.4.1 betrachteten LÉVY-Prozessen, dann genügt $X(t)$ einer stochastischen Differentialgleichung (SDE, stochastic differential equation). Für deren Lösung sind im Vergleich zur Lösung von ODE bestimmte Besonderheiten zu beachten⁹.

Zeitkontinuierliche stochastischen Prozesse sind nicht differenzierbar (siehe dbzgl. Anmerkung in Abschnitt B.4.1.2 zum SWP $W(t)$, Gl. (B-97)). SDE werden deshalb nicht mit Ableitungen der Variablen $X(t)$ nach der Zeit, sondern ausschließlich über Differentiale $dX(t)$ formuliert. Hierbei werden nur Differentiale 1. Ordnung betrachtet. Ebenso wie ODE höherer Ordnung als Systeme von ODE 1. Ordnung dargestellt werden können, ist aber eine analoge Modellierung durch Systeme von SDE möglich. Die folgenden Erläuterungen beschränken sich auf einfache SDE.

Eine stochastische Differentialgleichung beschreibt die zeitliche Änderung des Prozesses $X(t)$ für ein infinitesimales Zeitintervall dt ,

$$dX(t) = X(t+dt) - X(t),$$

wobei üblich statt " $dX(t)$ " verkürzt nur " dX " geschrieben wird. Für einen LÉVY-Prozess $L(t)$ kann die Änderung im Zeitintervall dt ,

$$dL = L(t+dt) - L(t) = L(dt),$$

direkt durch eine mit dt parametrisierte Zufallsgröße dargestellt werden (siehe Abschnitt 4.1.1).

Die allgemeine Form einer SDE mit einem LÉVY-Prozess ist

$$dX = a(X, t) \cdot dt + b(X, t) \cdot dL, \tag{B-129}$$

wobei $a(X, t)$ und $b(X, t)$ bestimmte Funktionen sind. Mit $b(X, t) = 0$ geht Gl. (B-129) von einer stochastischen in eine gewöhnliche Differentialgleichung über. Der zweite Term in Gl. (B-129), mit dem stochastisch variierende Beiträge zum Inkrement dX erfasst werden, wird als "Rauschquelle" bezeichnet. Wenn $b(X, t)$ eine Konstante oder eine nur von t abhängige Funktion ist, so spricht man von Fremderregung der Zustandsgröße X ; andernfalls liegt eine Parametererregung vor.

⁸ Zufallsfunktionen werden hier analog zu Zufallsgrößen mit Großbuchstaben symbolisiert, wogegen ihre (zufälligen) Realisierungen wieder mit kleinen Buchstaben bezeichnet werden.

⁹ $X(t)$ wird auch dann ein stochastischer Prozess, wenn ein stochastisch zeitlich variierender Parameter durch eine zeitdiskrete Folge von Zufallswerten aus einer bestimmten statistischen Verteilung approximiert wird, die z. B. aus Messungen des Parameters für aufeinander folgende feste Zeitintervalle Δt resultieren. In diesem Fall wird der Parameter als stückweise konstante Funktion mit zufälligen Werten dargestellt. Der (stochastische) Verlauf von $X(t)$ kann dann mit der entsprechenden ODE berechnet werden. Diese Art der Modellierung zeitdiskreter stochastischer Prozesse wird in Abschnitt 5.1.1 behandelt.

Es sei vermerkt, dass SDE mehrere Rauschquellen enthalten können, die auf unterschiedliche Weise die Änderung des Prozesses $X(t)$ beeinflussen. So wäre z. B.

$$dX = a(X, t) \cdot dt + b(X, t) \cdot dW + c(X, t) \cdot dN$$

die SDE für einen Prozess $X(t)$, der sowohl durch einen WIENER-Prozess $W(t)$ als auch durch einen POISSON-Prozess $N(t)$ beeinflusst wird. Im Rahmen des vorliegenden Berichts werden jedoch nur SDE mit einem bestimmten stochastischen Prozess als Rauschquelle betrachtet. Nachfolgend wird auf Besonderheiten der Lösung solcher SDE eingegangen, deren Rauschquelle einem der beiden grundlegenden stochastischen Prozesse entspricht, dem Standard-WIENER-Prozess (SWP) bzw. einem elementaren POISSON-Prozess (EPP).

B.4.2.1 SDE mit WIENER-Prozess als Rauschquelle

Für eine analytische Lösung von SDE sind oft - analog zur Lösung von ODE - nichtlineare Transformationen der Variablen erforderlich. Zur Lösung von ODE können dabei höhere Potenzen von dt vernachlässigt werden. Ist $x(t)$ eine stetige Funktion, so ist auch $(x(t) \cdot dt)^2$ vernachlässigbar. Zur Lösung von SDE mit einem Standard-WIENER-Prozess (SWP; $dW \sim N(0, \sqrt{dt})$); siehe Abschnitt B.4.1.2) können höheren Potenzen von dt sowie Produkte von dt mit dW ebenfalls vernachlässigt werden; für das Quadrat von dW gilt jedoch die so genannte ITO-Regel:

$$(dW)^2 = dt. \tag{B-130}$$

Wäre der Prozess $W(t)$ stetig differenzierbar, d. h. $dW/dt = \xi(t)$, so könnte auch $(dW)^2 = (\xi(t))^2 \cdot (dt)^2$ vernachlässigt werden. Dies trifft jedoch auf den SWP nicht zu.

Zum Verständnis von Gl. (B-130) sei zunächst vermerkt, dass nach Definition von dW für den Erwartungswert von $(dW)^2$ die Beziehung

$$E[(dW)^2] = \text{Var}[dW] = dt$$

gilt. Die Varianz von $(dW)^2$ lässt sich elementar durch Integration des Produkts der 4. Potenz von dW mit der Wahrscheinlichkeitsdichte der Normalverteilung $N(0, \sqrt{dt})$ berechnen:

$$\text{Var}[(dW)^2] = E[(dW)^4] - (E[(dW)^2])^2 = 3 \cdot (dt)^2 - (dt)^2 = 2 \cdot (dt)^2.$$

Da höhere Potenzen von dt vernachlässigt werden, ist die Varianz von $(dW)^2$ gleich Null zu setzen, so dass $(dW)^2$ als deterministische Größe mit dem Erwartungswert dt zu behandeln ist (eine ausführlichere Diskussion dieser Problematik enthält z. B. [JAC 10]). Die oben angesprochenen Vernachlässigbarkeit des Produkts $dt \cdot dW$ kann auf ähnliche Weise begründet werden. Da dt eine deterministische Größe ist, gilt

$$E[dt \cdot dW] = dt \cdot E[dW] = 0 \quad \text{und} \quad \text{Var}[dt \cdot dW] = (dt)^2 \cdot \text{Var}[dW] = (dt)^3,$$

wonach $dt \cdot dW$ wegen der Vernachlässigbarkeit höherer Potenzen von dt gleich Null zu setzen ist.

Gl. (B-130) kann bei allen nichtlinearen Transformationen stochastischer Prozesse, die durch einen WIENER-Prozess fremd- bzw. parametererregt sind, angewandt werden. Ist $X(t)$ ein solcher stochastischer Prozess und $Y = f(X, t)$ eine nach t sowie zweifach nach X stetig differenzierbare Funktion, dann ist zur Bestimmung der SDE für den Prozess $Y(t)$ die ITO-Formel

$$dY = \left(\frac{\partial f}{\partial t} \right) \cdot dt + \left(\frac{\partial f}{\partial X} \right) \cdot dX + \frac{1}{2} \cdot \left(\frac{\partial^2 f}{\partial X^2} \right) \cdot (dX)^2 \quad (\text{B-131})$$

anzuwenden. Sie folgt aus der TAYLOR-Entwicklung von $f(X, t)$, wobei höhere Ableitungen als in Gl. (B-131) angegeben wegen $(dt)^2 = 0$ und $dt \cdot dW = 0$ entfallen. Der quadratische Term ist dagegen wegen $(dW)^2 = dt$ gemäß Gl. (B-130) zu berücksichtigen.

Genügt $X(t)$ z. B. der SDE

$$dX = a \cdot dt + b \cdot dW,$$

wobei a und b Konstanten sind, dann gilt gemäß Gl. (B-131) für den Prozess $Y = X^2$ die SDE

$$dY = 2 \cdot X \cdot dX + (dX)^2,$$

in der wegen Gl. (B-130) der letzte Term im Unterschied zu einer ODE nicht mehr verschwindet, sondern den Beitrag $b^2 \cdot (dW)^2 = b^2 \cdot dt$ liefert. $Y(t)$ genügt dementsprechend der SDE

$$dY = (2 \cdot a \cdot X + b^2) \cdot dt + 2 \cdot b \cdot X \cdot dW.$$

Während für einige einfache Fälle für SDE mit WIENER-Rauschterm,

$$dX = a(X, t) \cdot dt + b(X, t) \cdot dW, \quad (\text{B-132})$$

eine analytische Lösung existiert, muss i. A. eine numerische Integration dieser Gleichung erfolgen (siehe z. B. [RÜM 82], [KLOE 95], [ALL 07]). Diese basieren auf der Definition der stochastischen Integration mit dem so genannten ITO-Integral. Das formale Integral der Gl. (B-132),

$$X(t) = X(0) + \int_0^t a(X(u), u) du + \int_0^t b(X(u), u) dW(u), \quad (\text{B-133})$$

wird dazu als Grenzwert einer Summe über $I = t/\Delta$ Zeitintervalle der Länge Δ definiert, wobei für die Funktionen $a(X(u), u)$ und $b(X(u), u)$ jeweils die Werte am Anfang eines Intervalls der Länge Δ gesetzt werden, wogegen für $dW(u)$ die Differenz des WIENER-Prozesses $W(u+\Delta) - W(u)$ zwischen dem Ende und dem Anfang eines Intervalls $[(i-1) \cdot \Delta, i \cdot \Delta]$ mit i von 1 bis I gesetzt wird, d. h.,

$$X(t) = X(0) + \lim_{I \rightarrow \infty} \left(\sum_{i=1}^I a(X(u_{i-1}), u_{(i-1)}) \cdot \Delta + \sum_{i=1}^I b(X(u_{i-1}), u_{i-1}) \cdot (W(u_i) - W(u_{i-1})) \right) \quad (\text{B-134})$$

$$\text{mit } \Delta = t/I \text{ und } u_i = i \cdot \Delta.$$

Die Definition des ITO-Integrals entspricht der obigen Definition von SDE, die deshalb häufig auch als ITO-SDE bezeichnet wird.¹⁰

¹⁰ Eine andere Definition der SDE und stochastischer Integrale stammt von STRATONOVICH [STRA 66]. Dabei wird in Gl. (B-134) das Mittel der Werte der Funktionen a und b an den beiden Enden der Intervalle angesetzt.

Zur Veranschaulichung der Anwendung des ITO-Kalküls zur Lösung von SDE sei die in Abschnitt 2.6.5 angegebene LANGEVIN-Gleichung betrachtet, die in der Fachliteratur zu SDE nach [UHL 30] auch als ORNSTEIN-UHLENBECK-Prozess bezeichnet wird (siehe z. B. [GILL 96], [JAC 10]):

$$dX = -\lambda X dt + b dW, \quad X(0) = X_0, \quad (\text{B-135})$$

wobei λ und b positive Konstanten und $dW \sim N(0, \sqrt{dt})$ den SWP bezeichnen. Ohne den Rauschterm $b \cdot dW$ wäre $X(t) = X_0 \cdot e^{-\lambda t}$ die bekannte Lösung der Zerfallsgleichung. Zur Lösung von Gl. (B-135) wird demgemäß die Transformation

$$Y(t) = X(t) \cdot e^{\lambda t} \quad (\text{B-136})$$

angewandt. Für $Y(t)$ gilt die Anfangsbedingung $Y(0) = X(0) = X_0$. Gemäß Gl. (B-131) resultiert für die transformierte Zufallsgröße die SDE

$$dY = \lambda X e^{\lambda t} dt + e^{\lambda t} dX = \lambda X e^{\lambda t} dt + e^{\lambda t} (-\lambda X dt + b dW) = b e^{\lambda t} dW, \quad (\text{B-137})$$

deren Lösung elementar ist:

$$Y(t) = Y(0) + b \int_0^t e^{\lambda s} dW(s),$$

Nach Rücktransformation entsprechend Gl. (B-136) ergibt sich für $X(t)$ die Lösung

$$X(t) = X_0 e^{-\lambda t} + b \int_0^t e^{-\lambda(t-s)} dW(s). \quad (\text{B-138})$$

Dies ist eine Überlagerung der deterministischen Lösung der Gl. (B-135) ohne Rauschterm mit einem stochastischen Prozess. Da das stochastische Integral nach Gl. (B-134) als Grenzwert einer Summe von normalverteilten Zufallsgrößen zu verstehen ist, stellt $X(t)$ für einen fixen Zeitpunkt t eine normalverteilte Zufallsgröße dar. Wegen $E[dW(s)] = 0$ resultiert für $X(t)$ der Erwartungswert

$$E[X(t)] = X_0 e^{-\lambda t}. \quad (\text{B-139})$$

Zur Berechnung der Varianz von $X(t)$ kann ausgenutzt werden, dass die Varianz des Produkts einer deterministischen Größe mit einer normalverteilten Zufallsgröße gleich dem Produkt des Quadrats der deterministischen Größe mit der Varianz der normalverteilten Zufallsgröße ist. Folglich gilt für beliebige deterministische Funktionen $g(s, t)$, dass das stochastische Integral

$$S(t) = \int_0^t g(s, t) dW(s) \quad (\text{B-140})$$

eine normalverteilte Zufallsgröße mit Erwartungswert $E[S(t)] = 0$ und Varianz

$$\text{Var}[S(t)] = \int_0^t (f(s, t))^2 ds \quad (\text{B-141})$$

ist. Die Anwendung von Gl. (B-141) auf Gl. (B-138) liefert zur Varianz von $X(t)$ die Beziehung

$$\text{Var}[X(t)] = b^2 \int_0^t e^{-2\lambda(t-s)} ds = \frac{b^2}{2\lambda} (1 - e^{-2\lambda t}). \quad (\text{B-142})$$

Hinsichtlich der Lösung von SDE mit Parametererregung durch einen WIENER-Prozess sei auf das Beispiel 5.3 in Abschnitt 5.1.2.1 verwiesen.

B.4.2.2 SDE mit POISSON-Prozess als Rauschquelle

Ähnlich zu den im vorigen Abschnitt betrachteten SDE, die durch einen WIENER-Prozess fremd- oder parametererregt sind, lassen sich auch SDE, denen ein elementarer POISSON-Prozess (EPP) zugrunde liegt, relativ einfach behandeln und z. T. analytisch lösen. Die Form solcher SDE ist analog zu Gl. (B-132):

$$dX = a(X, t) \cdot dt + b(X, t) \cdot dN, \quad (\text{B-145})$$

wobei dN das infinitesimale Inkrement des EPP mit $N(t) \sim \text{Poi}(\lambda \cdot t)$ ist (siehe Abschnitt B.4.1.3). Da zur Lösung der SDE i. A. wieder nicht-lineare Transformationen erforderlich sind, wird eine zur ITO-Formel Gl. (B-131) analoge Beziehungen benötigt. Sie basiert auf der Besonderheit des EPP, wonach dN nur die Werte 1 und 0 annehmen kann ($\text{Pr}(dN = 1) = \lambda \cdot dt$ und $\text{Pr}(dN = 0) = 1 - \lambda \cdot dt$). Folglich gilt unabhängig davon, ob in einem infinitesimalen Zeitintervall dt ein POISSON-Ereignis stattfindet oder nicht, immer die Beziehung

$$dN^2 = dN. \quad (\text{B-146})$$

Sie spielt für den EPP eine ähnliche Rolle wie Gl. (B-130) für den SWP. Während für den WIENER-Prozess Potenzen $(dW)^k$ mit $k > 2$ zu vernachlässigen sind (wegen $(dW)^2 = dt$ und $dW \cdot dt = 0$), gilt für das Inkrement des EPP $(dN)^k = dN$ auch für alle Potenzen $k > 2$. Außerdem gilt $dN \cdot dt = 0$.

Ist $X(t)$ ein durch POISSON-Prozesse bestimmter stochastischer Prozess und $Y = f(X, t)$ eine nach t sowie beliebig oft nach X stetig differenzierbare Funktion, dann ist zur Ableitung der SDE für den Prozess $Y(t)$ die Formel

$$dY = \left(\frac{\partial f}{\partial t} \right) \cdot dt + \sum_{k=1}^{\infty} \frac{1}{k!} \cdot \frac{\partial^k f}{\partial X^k} \cdot (dX)^k = \left(\frac{\partial f}{\partial t} \right) \cdot dt + \frac{\partial f}{\partial X} \cdot dX + \sum_{k=2}^{\infty} \frac{(b(X, t))^k}{k!} \cdot \frac{\partial^k f}{\partial X^k} \cdot dN \quad (\text{B-147})$$

anzuwenden. Im Unterschied zur Gl. (B-131) kann die in Gl. (B-147) aus der TAYLOR-Entwicklung resultierende Reihe wegen $(dN)^k = dN$ (s. o.) also nicht abgebrochen werden.

Zur Veranschaulichung der Anwendung von Gl. (B-147) zur Lösung einer SDE mit EPP sei

$$dX = -a \cdot X \cdot dt - b \cdot X \cdot dN, \quad \text{mit } a > 0, \quad 0 < b < 1 \quad \text{und} \quad X(0) = X_0, \quad (\text{B-148})$$

betrachtet, wobei $dN \sim \text{Poi}(\lambda \cdot t)$ den EPP bezeichnet. In diesem Fall ist $X(t)$ durch eine stochastische Parametererregung gestört.

Zur Lösung von Gl. (B-148) wird zunächst die Transformation

$$Y = \ln(X), \quad (\text{B-149})$$

durchgeführt. Mit

$$\frac{\partial^k \ln(X)}{\partial X^k} = \frac{(-1)^{k+1} \cdot (k-1)!}{X^k}$$

folgt nach Gl. (B-147) mit Gl. (B-148):

$$dY = \sum_{k=1}^{\infty} \frac{(-1)^{k+1} \cdot (k-1)!}{X^k} \cdot \frac{(dX)^k}{k!} = - \sum_{k=1}^{\infty} \frac{(a dt + b dN)^k}{k}$$

In der Summe auf der rechten Seite dieser Gleichung können höhere Potenzen von dt vernachlässigt werden, wogegen für Potenzen von dN nach Gl. (B-146) $(dN)^k = dN$ zu setzen ist. Damit folgt

$$dY = -a dt - dN \sum_{k=1}^{\infty} \frac{b^k}{k} = -a dt + \ln(1-b) \cdot dN.$$

Diese SDE kann direkt integriert werden,

$$Y(t) = Y(0) - a t + \ln(1-b) \cdot N(t),$$

woraus durch Rücktransformation nach Gl (B-149) unter Beachtung der Anfangsbedingung für X(t) die folgende Lösung resultiert:

$$X(t) = X_0 \cdot e^{-a t + \ln(1-b) \cdot N(t)} = X_0 \cdot e^{-a t} \cdot (1-b)^{N(t)}. \quad (\text{B-150})$$

Einer radiologisch relevanten Anwendung der SDE Gl. (B-148) und physikalische Interpretation der Lösung Gl. (B-150) enthält das Beispiel 5.5-1 in Abschnitt 5.1.2.2. Der Erwartungswert und die Varianz von X(t) ergeben sich durch Mittelungen mit der Wahrscheinlichkeitsmassenfunktion der POISSON-Verteilung $N(t) \sim \text{Poi}(\lambda \cdot t)$:

$$E[X(t)] = X_0 \cdot e^{-a t} \sum_{n=0}^{\infty} \frac{(1-b)^n (\lambda t)^n \cdot e^{-\lambda t}}{n!} = X_0 \cdot e^{-a t - b \lambda t}. \quad (\text{B-151})$$

$$\begin{aligned} \text{Var}[X(t)] &= E[X^2(t)] - (E[X(t)])^2 \\ &= X_0^2 e^{-2 a t} \sum_{n=0}^{\infty} \frac{(1-b)^{2n} (\lambda t)^n \cdot e^{-\lambda t}}{n!} - X_0^2 e^{-2 a t - 2 b \lambda t} \\ &= X_0^2 e^{-2 a t} e^{-2 b \lambda t + b^2 \lambda t} - X_0^2 e^{-2 a t - 2 b \lambda t} = X_0^2 e^{-2 a t - 2 b \lambda t} (e^{b^2 \lambda t} - 1) \\ &= (E[X(t)])^2 \cdot (e^{b^2 \lambda t} - 1) \end{aligned} \quad (\text{B-152})$$

ANHANG C

Univariate parametrische Verteilungen

C.1 Vorbemerkungen

In den nachfolgenden Abschnitten dieses Anhangs werden einige wichtige univariate statistische Verteilungen beschrieben. Die Erläuterungen zu stetigen und diskreten Verteilungen beschränken sich auf Verteilungen, die im Hauptteil des vorliegenden Berichts für Beispiele herangezogen werden. Sie streben demgemäß keine vollständige Übersicht aller gängigen Verteilungen an, die auch für Expositionsabschätzungen relevant sind bzw. sein können. Aus der Vielfalt der stetigen Verteilungen wurden neben grundlegenden "Familien" auch einige weniger bekannte Verteilungen ausgesucht, die für spezielle Fragestellungen zunehmend Anwendung finden und/oder wegen ihrer relativ einfachen mathematischen Handhabbarkeit als sinnvolle Alternativen zu den bekannteren Verteilungsarten in Betracht gezogen werden können. Für weiterführende Beschreibungen der hier erläuterten sowie weiterer Verteilungen sei auf die Fachliteratur verwiesen, insbesondere auf die Standardwerke [JOH 94], [JOH 95], [JOH 05] sowie einige statistische Hand- und Fachbücher (siehe z. B. [RIN 97], [BURY 99], [CAS 02], [BAL 03], [KRIS 06], [EVE 10] und [FOR 11]).

Zwischen diversen univariaten Verteilungen bestehen funktionelle Zusammenhänge, die z. B. aus Transformationen, Grenzfallbetrachtungen bzw. für spezielle Parameterkonstellationen oder auch für bestimmte Stichprobenfunktionen resultieren. Auf solche Zusammenhänge wird nachfolgend zumindest teilweise hingewiesen. Eine dbzgl. sehr umfangreiche Darstellung enthält [LEEM 08].

C2 Stetige Verteilungen mit unbegrenztem Definitionsbereich ($-\infty < X < \infty$)

C.2.1 Normalverteilung $X \sim N(\mu, \sigma)$

Wahrscheinlichkeitsdichte und Verteilungsfunktion normal verteilter Variabler $X \sim N(\mu, \sigma)$ sind:

$$f_N(x | \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right], \quad -\infty < x, \mu < \infty; \quad \sigma > 0; \quad (C-1)$$

$$F_N(x | \mu, \sigma) = \Phi\left(\frac{x-\mu}{\sigma}\right) = \frac{1}{2} \left(1 + \operatorname{erf}\left(\frac{x-\mu}{\sigma\sqrt{2}}\right)\right), \quad (C-2)$$

wobei $\Phi(x)$ die Verteilungsfunktion der Standardnormalverteilung $N(0,1)$ bezeichnet und $\operatorname{erf}(x)$ die so genannte Fehlerfunktion ist,

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x \exp[-z^2] dz = \frac{2}{\sqrt{\pi}} \exp[-x^2] \cdot \sum_{n=0}^{\infty} \frac{2^n \cdot x^{2n+1}}{(2n+1)!!}. \quad (C-3)$$

Der Erwartungswert und die Varianz von X sind gegeben durch:

$$E[X] = \mu \quad \text{und} \quad \operatorname{Var}[X] = \sigma^2. \quad (C-4)$$

Für Schiefe γ_1 und Kurtosis γ_2 der Normalverteilung folgt gemäß Gl. (B-41) bzw. (B-43):

$$\gamma_1 = 0 \quad \text{und} \quad \gamma_2 = 3. \quad (C-5)$$

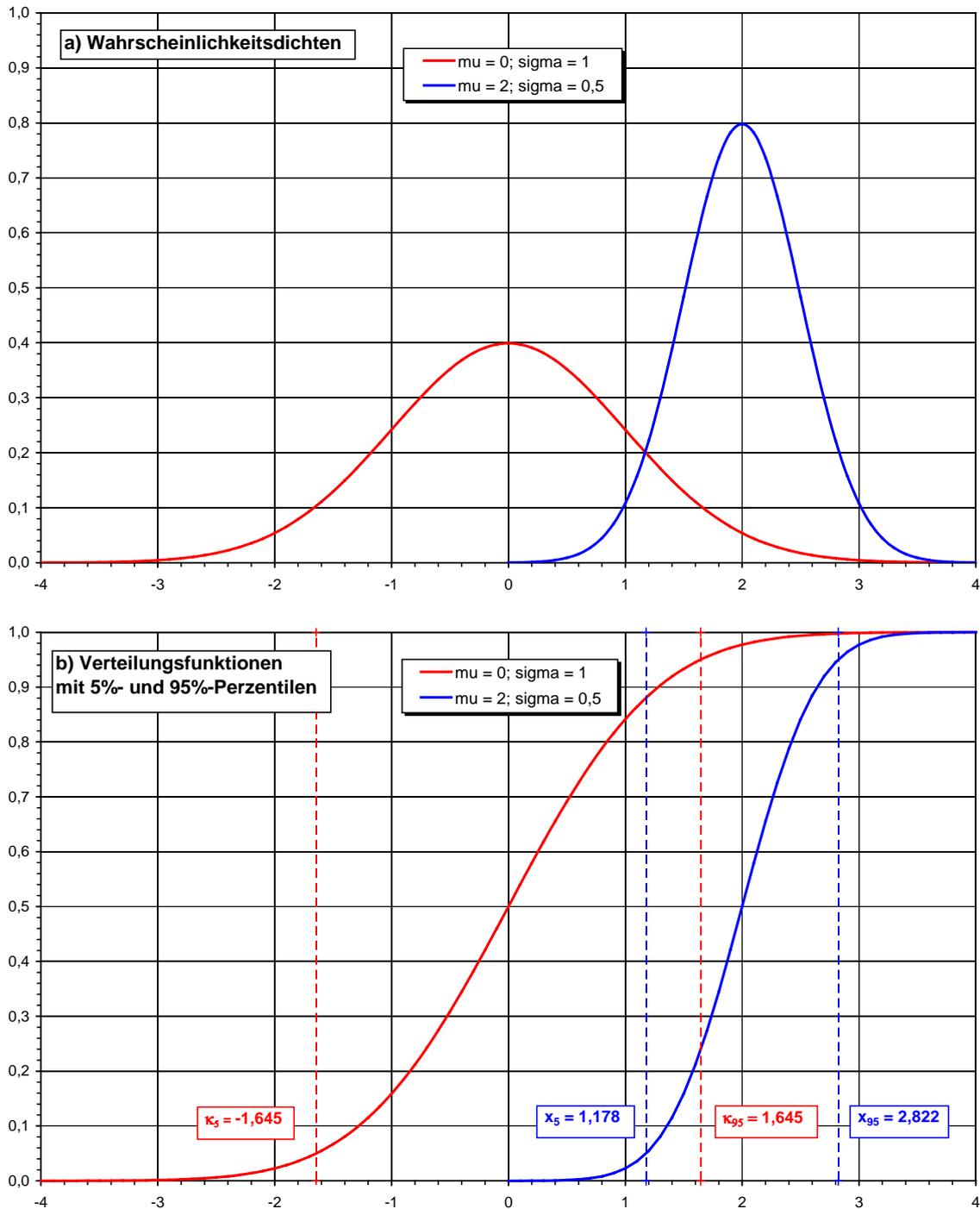


Abbildung C-1: Wahrscheinlichkeitsdichten, Verteilungsfunktionen sowie 5%- und 95%-Perzentile der Standardnormalverteilung $N(0, 1)$ sowie der Normalverteilung mit $\mu = 2; \sigma = 0,5$

Die beiden Verteilungsparameter $\theta_1 = \mu$ und $\theta_2 = \sigma$ einer normal verteilten Zufallsgröße sind somit ihr Erwartungswert und die positive Wurzel ihrer Varianz, die Standardabweichung genannt wird.

In der Abbildung C-1 sind zur Illustration Wahrscheinlichkeitsdichten und Verteilungsfunktionen der Normalverteilung für die Parameter $\mu = 0$ und $\sigma = 1$ bzw. $\mu = 2$ und $\sigma = 0,5$ dargestellt. Im erstgenannten Fall spricht man von der Standardnormalverteilung $N(0,1)$. Die Abbildung C-1b zeigt auch die 5%- und 95%-Perzentile der beiden Verteilungen. Für die p-Quantile der Standardnormalverteilung wird die spezielle Bezeichnung κ_p verwendet. Die Werte von κ_p lassen sich entsprechend Gl. (B-4) mit Gl. (C-2) und (C-3) relativ einfach berechnen; sie sind aber auch in gängigen Fachbüchern tabelliert (siehe z. B. [RIN 97]). Die p-Quantile der allgemeinen Normalverteilungen $N(\mu, \sigma)$ resultieren aus den p-Quantilen der Standardnormalverteilung über folgende Beziehung:

$$x_p = \mu + \sigma \cdot \kappa_p \quad (\text{C-6}).$$

Wegen $\kappa_{0,5} = 0$ ist der Median Med_N (Definition: $F_X(\text{Med}_X) = 0,5$) für eine normal verteilte Größe gleich dem Erwartungswert. Der Modalwert Mod_X einer Zufallsgröße X durch das Maximum ihrer Wahrscheinlichkeitsdichte definiert (wahrscheinlichster Wert). Auch er stimmt für normal verteilte Größen mit dem Erwartungswert überein ($\text{Med}_N = \text{Mod}_N = E_N = \mu$).

Der Ansatz der Normalverteilung für eine Zufallsgröße X ist nach dem zentralen Grenzwertsatz der Wahrscheinlichkeitstheorie nahe liegend, wenn ihre Variabilität aus einer additiven Überlagerung vieler zufälliger Einflüsse resultiert, von denen keines dominiert (Beispiele: Messunsicherheiten, atmosphärische Dispersion von Schadstoffwolken). Die Normalverteilung ist auf der gesamten reellen Achse definiert ($-\infty < x < \infty$), sie kann jedoch näherungsweise auch für Zufallsgrößen, die nur positive Werte annehmen können, verwendet werden, wenn die Standardabweichung wesentlich kleiner als der Mittelwert ist (z. B. für $\mu > 5 \cdot \sigma$). Die Summe $Y = X_1 + X_2$ von zwei unabhängigen, normalverteilten Zufallsgrößen $X_1 \sim N(\mu_1, \sigma_1)$ und $X_2 \sim N(\mu_2, \sigma_2)$ ist ebenfalls eine normalverteilte Größe; $Y \sim N(\mu_1 + \mu_2, \sqrt{\sigma_1^2 + \sigma_2^2})$.

C.2.2 STUDENT'sche t-Verteilung $X \sim t(v)$

Die Wahrscheinlichkeitsdichte der t-Verteilung mit v Freiheitsgraden ist:

$$f_t(t | v) = \frac{\Gamma\left(\frac{v+1}{2}\right) \cdot \left(1 + \frac{t^2}{v}\right)^{-(v+1)/2}}{\sqrt{v \cdot \pi} \cdot \Gamma(v/2)}, \quad -\infty < t < \infty; \quad v \geq 1, \quad (\text{C-7})$$

wobei $\Gamma(\alpha)$ die Gammafunktion bezeichnet:

$$\Gamma(\alpha) = \int_0^{\infty} t^{\alpha-1} \cdot e^{-t} dt \quad \text{mit } \text{Re}(\alpha) > 0. \quad (\text{C-8})$$

Anmerkung: $\Gamma(\alpha+1) = \alpha \cdot \Gamma(\alpha)$; $\Gamma(1) = 1$; $\Gamma(1/2) = \sqrt{\pi}$.

Die t-Verteilung ist insbesondere für die Einschätzung von Unsicherheiten der Schätzung des Erwartungswertes von normalverteilten Größen anhand von Stichproben von Bedeutung.

Bezeichnen \bar{x} den Mittelwert und s die Standardabweichung einer Stichprobe vom Umfang n einer normalverteilten Zufallsgröße, dann genügt die Zufallsvariable

$$t = \frac{\mu - \bar{x}}{s/\sqrt{n}} \quad (\text{C-9})$$

einer STUDENT'schen t -Verteilung mit $\nu = n - 1$ Freiheitsgraden, wobei μ den (unbekannten) Erwartungswert der beprobten Normalverteilung bezeichnet. Da \bar{x} und s sowie n für die betrachtete Stichprobe feste Größen sind, stellt μ eine gemäß Gl. (C-9) transformierte Zufallsgröße im Sinne der BAYES-Theorie dar.

Die Verteilungsfunktion der t -Verteilung ist nicht elementar darstellbar. Die t -Verteilung ist symmetrisch; ihr Mittelwert und ihre Schiefe sind gleich Null. Die t -Verteilung ist leptokurtisch; für ihre Varianz und Kurtosis gelten die Beziehungen:

$$\text{Var}[t] = \nu/(\nu - 2) \quad \text{für } \nu > 2; \quad \gamma_2 = 3 + 6/(\nu - 4) \quad \text{für } \nu > 4. \quad (\text{C-10})$$

Für sehr große Werte des Freiheitsgrades ν (Grenzwert $\nu \rightarrow \infty$) nähert sich die t -Verteilung der Standardnormalverteilung $N(0,1)$ an. Die Abbildung C-2 veranschaulicht diesen Übergang anhand der Wahrscheinlichkeitsdichten der t -Verteilung für die Freiheitsgrade $\nu = \{5, 10, 100\}$. Daraus ist ersichtlich, dass die t -Verteilung bei einem Freiheitsgrad von 100 nur noch sehr geringfügig von der Standardnormalverteilung abweicht. Für kleine Freiheitsgrade (Stichprobenumfänge) sind dagegen die deutlich höheren Wahrscheinlichkeiten in den Randbereichen der t -Verteilung zu beachten.

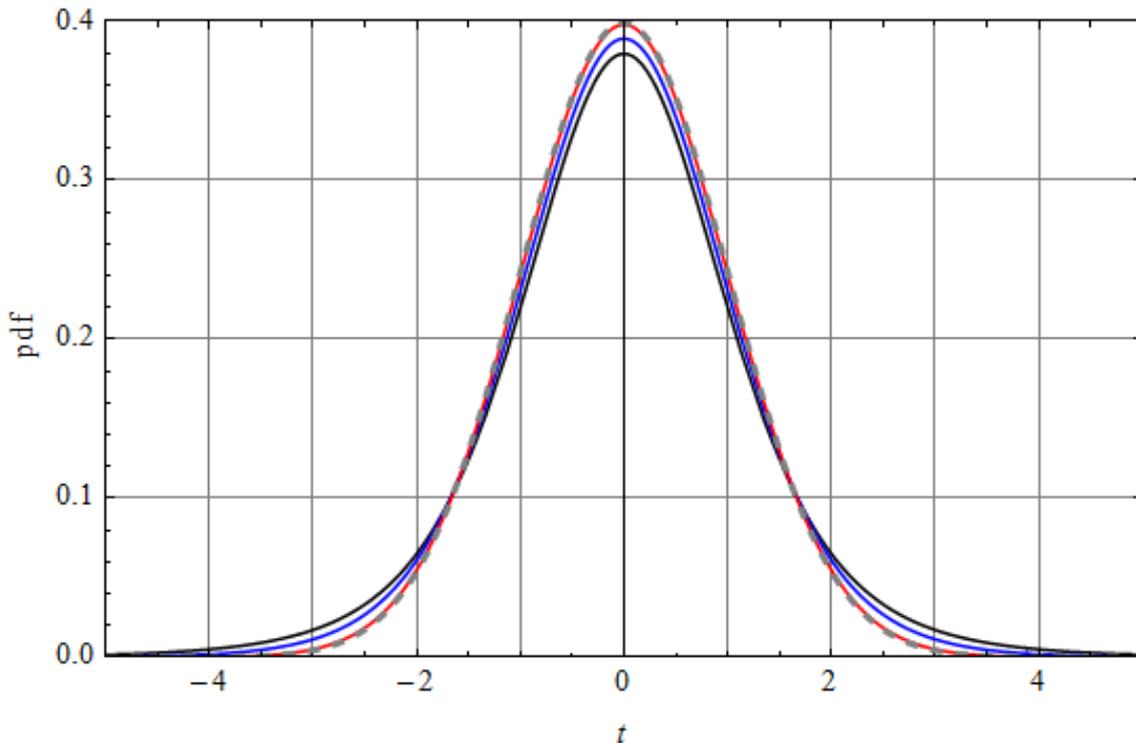


Abbildung C-2: Wahrscheinlichkeitsdichte der t -Verteilung für Freiheitsgrade ν von 5 (schwarz) 10 (blau) und 100 (rot) im Vergleich zur Standardnormalverteilung (grau, gestrichelt)

C.3 Stetige Verteilungen für positiv definierte Zufallsvariable ($x \geq 0$)

C.3.1 Chi-Quadrat-Verteilung $X \sim \chi^2(v)$

Die Wahrscheinlichkeitsdichte der Chi-Quadrat-Verteilung mit v Freiheitsgraden ist:

$$f_{\chi^2}(x | v) = \frac{x^{v/2-1} \cdot e^{-x/2}}{2^{v/2} \cdot \Gamma(v/2)}, \quad 0 \leq x < \infty, v \geq 1 \quad (\text{C-11})$$

wobei $\Gamma(z)$ wieder die Gammafunktion bezeichnet (siehe Gl. (C-8)). Eine wichtige Eigenschaft der Chi-Quadrat-Verteilung ist Ihre Additivität. Sind $X_1 \sim \chi^2(v_1)$ und $X_2 \sim \chi^2(v_2)$ unabhängig, so folgt ihre Summe einer Chi-Quadrat-Verteilung mit $(v_1 + v_2)$ Freiheitsgraden: $X_1 + X_2 \sim \chi^2(v_1 + v_2)$.

Die Chi-Quadrat-Verteilung ist insbesondere zur Einschätzung von Unsicherheiten der Schätzung der Varianz von normalverteilten Größen anhand von Stichproben von Bedeutung. Bezeichnet s die Standardabweichung einer Stichprobe vom Umfang n einer normalverteilten Zufallsgröße, dann genügt die Zufallsvariable

$$x = \frac{(n-1) \cdot s^2}{\sigma^2}$$

einer Chi-Quadrat-Verteilung mit $v = n - 1$ Freiheitsgraden, wobei σ^2 die (unbekannte) Varianz der beprobten Normalverteilung ist. Die Chi-Quadrat-Verteilung ist linkssteil und leptokurtisch. Für sehr große Werte von v nähert sich die Chi-Quadrat-Verteilung einer Normalverteilung $N(v, \sqrt{2 \cdot v})$ an (siehe Abbildung C-3).

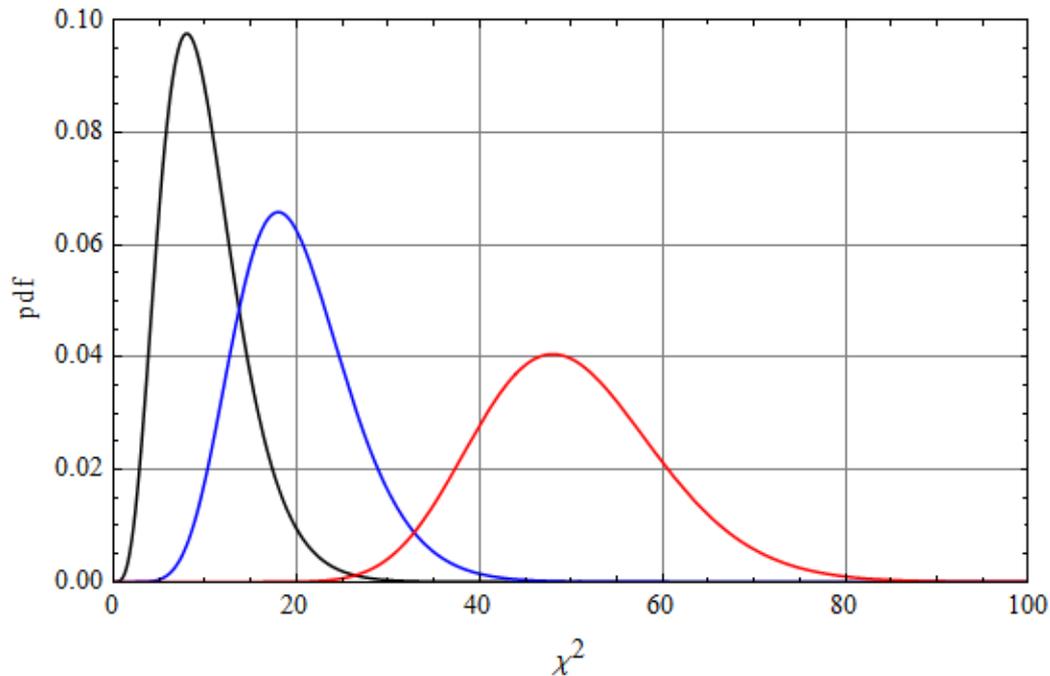


Abbildung C-3: Wahrscheinlichkeitsdichte der Chi-Quadrat-Verteilung für Freiheitsgrade ν von 10 (schwarz), 20 (blau) und 50 (rot)

Ihre wichtigsten statistischen Kenngrößen sind:

$$E_{\chi^2}[X] = \nu; \quad \text{Var}_{\chi^2}[X] = 2 \cdot \nu; \quad \gamma_1 = 2 \cdot \sqrt{2/\nu}; \quad \gamma_2 = 3 + 12/\nu. \quad (\text{C-12})$$

C.3.2 Lognormalverteilung $X \sim \text{Ln}(\mu, \sigma)$

Eine Zufallsgröße X ist lognormal verteilt, wenn ihre logarithmisch transformierte Größe $Y = \ln(X)$ normal verteilt ist. Als Verteilungsparameter werden üblich der Erwartungswert μ und die Standardabweichung σ der logarithmierten Größe $Y \sim N(\mu, \sigma)$ verwendet.

Die Wahrscheinlichkeitsdichte von $X \sim \text{Ln}(\mu, \sigma)$ ist

$$f_{\text{Ln}}(x | \mu, \sigma) = \frac{1}{x \cdot \sigma \sqrt{2\pi}} \exp\left[-\frac{(\ln(x) - \mu)^2}{2\sigma^2}\right] \quad \text{mit } x \geq 0; -\infty < \mu < \infty; \sigma > 0. \quad (\text{C-13})$$

Für die Verteilungsfunktion der Lognormalverteilung ergibt sich die Beziehung

$$F_{\text{Ln}}(x | \mu, \sigma) = \int_0^x f_{\text{Ln}}(t | \mu, \sigma) dt = \frac{1}{2} \left(1 + \text{erf}\left(\frac{\ln(x) - \mu}{\sigma\sqrt{2}}\right) \right). \quad (\text{C-14})$$

Nach Gl. (B-8) bis Gl. (B-10) lassen sich der Erwartungswert und die Varianz analytisch berechnen:

$$E_{\text{Ln}}[X] = \int_0^{\infty} x \cdot f_{\text{Ln}}(x | \mu, \sigma) dx = \exp\left[\mu + \frac{\sigma^2}{2}\right], \quad (\text{C-15})$$

$$\text{Var}_{\text{Ln}}[X] = \int_0^{\infty} x^2 \cdot f_{\text{Ln}}(x | \mu, \sigma) dx - E_{\text{Ln}}^2 = E_{\text{Ln}}^2 \cdot (\exp(\sigma^2) - 1) = \exp(2\mu + \sigma^2) \cdot (\exp(\sigma^2) - 1). \quad (\text{C-16})$$

Die p-Quantile x_p für lognormal verteilte Zufallsgrößen lassen auf einfache Weise aus den p-Quantilen für Normalverteilungen Gl. (C-6) berechnen:

$$x_p = \exp(\mu + \sigma \cdot \kappa_p). \quad (\text{C-17})$$

Wegen $\kappa_{0,5} = 0$ resultiert für den Median ($F_X(\text{Med}_X) = 0,5$) die Formel

$$\text{Med}_{\text{LN}} = \exp[\mu]. \quad (\text{C-18})$$

Der Median einer lognormal verteilten Zufallsgröße ist auch ihr geometrischer Mittelwert und um einen Faktor von $\exp(\sigma^2/2)$ kleiner als der Erwartungswert. Der Modalwert ist noch kleiner:

$$\text{Mod}_{\text{LN}} = \exp(\mu - \sigma^2). \quad (\text{C-19})$$

Die Lognormalverteilung ist linkssteil und leptokurtisch; ihre Schiefe und Kurtosis sind:

$$\gamma_1 = (2 + e^{\sigma^2}) \cdot \sqrt{e^{\sigma^2} - 1} \quad \text{und} \quad \gamma_2 = e^{4\sigma^2} + 2 \cdot e^{3\sigma^2} + 3 \cdot e^{2\sigma^2} - 3. \quad (\text{C-20})$$

Der Ansatz der Lognormalverteilung für eine Zufallsgröße X ist nach dem zentralen Grenzwertsatz der Wahrscheinlichkeitstheorie nahe liegend, wenn ihre Variabilität aus einer multiplikativen Überlagerung vieler zufälliger Einflüsse resultiert, von denen keines dominiert, wobei reziproke Einflussfaktoren, die nur positive Werte annehmen können, eingeschlossen sind (der Logarithmus der Zufallsgröße X ist dann die Summe/Differenz vieler kleiner Einflussfaktoren). Lognormal verteilte Größen stellen eine gute Approximation für viele Parameter der Natur und des Verhaltens von Personen dar (Beispiele: K_d -Werte, Bioakkumulationsfaktoren, natürliche Radonkonzentration in Gebäuden, Trinkwasserverbrauch).

C.3.3 Exponentialverteilung $X \sim \text{Ex}(\lambda)$

Die Wahrscheinlichkeitsdichte und die Verteilungsfunktion der Exponentialverteilung sind:

$$f_{\text{Ex}}(x | \lambda) = \lambda \cdot \exp(-\lambda \cdot x), \quad 0 \leq x < \infty, \quad \lambda > 0, \quad (\text{C-21})$$

$$F_{\text{Ex}}(x | \lambda) = 1 - \exp(-\lambda \cdot x). \quad (\text{C-22})$$

Sie findet in der Zuverlässigkeitsanalyse und anderen statistischen Modellen zur Beschreibung von selten auftretenden Ereignissen Anwendung und beschreibt hier die Dauer zwischen Ereignissen, die mit einer konstanten Rate (λ) auftreten, wie z. B. die Zeit bis zum Zerfall eines Radionuklids bzw. generell bis zum Ausfall ermüdungsfreier Systeme. Die Exponentialverteilung wird auch zur probabilistischen Modellierung von Expositionsfaktoren genutzt (z. B. für die probabilistische Modellierung der Schadstoffaufnahme durch Ingestion und Inhalation; siehe z. B. [LEE 95]).

Die Exponentialverteilung ist die einzige stetige statistische Verteilung, welche die Eigenschaft der "Gedächtnislosigkeit" aufweist. Diese besagt, dass die Überlebenswahrscheinlichkeit eines Systems in Bezug auf einen bestimmten Zeitpunkt nicht von seinem Alter abhängt:

$$\Pr(X \geq x_0 + x \mid X \geq x_0) = \frac{e^{-\lambda \cdot (x_0 + x)}}{e^{-\lambda \cdot x_0}} = e^{-\lambda \cdot x} = \Pr(X \geq x). \quad (\text{C-23})$$

Die Ausfallrate $h(x)$ (Hazardrate), die als Relation der Wahrscheinlichkeitsdichte zum Komplement der Verteilungsfunktion definiert ist, ist demgemäß für die Exponentialverteilung konstant,

$$h_{\text{Ex}}(x \mid \lambda) = \frac{f_{\text{Ex}}(x \mid \lambda)}{1 - F_{\text{Ex}}(x \mid \lambda)} = \lambda. \quad (\text{C-24})$$

Die Exponentialverteilung ist rechtsschief und leptokurtisch. Die wichtigsten Kenngrößen (Erwartungswert, Varianz, Schiefe und Kurtosis) sind:

$$E_{\text{Ex}}[X] = 1/\lambda; \quad \text{Var}_{\text{Ex}}[X] = 1/\lambda^2; \quad \gamma_1 = 2; \quad \gamma_2 = 9. \quad (\text{C-25})$$

Es sei darauf hingewiesen, dass zur Parametrisierung der Exponentialverteilung häufig statt der Rate λ der Skalenparameter $\beta = 1/\lambda$ verwendet wird ($f_{\text{Ex}}(x \mid \beta) = e^{-x/\beta} / \beta$).

Die Summe von n unabhängigen und gleichartig exponentialverteilten Zufallsgrößen ist wieder eine Zufallsgröße, die der so genannten ERLANG-Verteilung genügt. Diese stellt einen speziellen Fall der im folgenden Abschnitt erläuterten Gammaverteilung $\text{Gam}(\alpha, \lambda)$ mit $\alpha = n$ dar, was sich mit Gl. (B-23) durch wiederholte Ausführung des Faltungsintegrals leicht zeigen lässt.

C.3.4 Gammaverteilung $X \sim \text{Gam}(\alpha, \lambda)$ oder $X \sim \text{Ga}(\alpha, \beta)$

Die Wahrscheinlichkeitsdichte und die Verteilungsfunktion der Gammaverteilung sind:

$$f_{\text{Gam}}(x \mid \alpha, \lambda) = \frac{\lambda^\alpha \cdot x^{\alpha-1}}{\Gamma(\alpha)} \cdot \exp(-\lambda \cdot x), \quad 0 \leq x < \infty, \quad \alpha > 0, \lambda > 0, \quad (\text{C-26})$$

$$F_{\text{Gam}}(x \mid \alpha, \lambda) = 1 - \frac{\Gamma(\alpha, \lambda \cdot x)}{\Gamma(\alpha)}, \quad (\text{C-27})$$

wobei in Gl. (C-27) mit $\Gamma(\alpha, z)$ die unvollständige Gammafunktion bezeichnet ist (vgl. Gl. (C-8)):

$$\Gamma(\alpha, z) = \int_z^\infty t^{\alpha-1} \cdot e^{-t} dt \quad \text{mit } \text{Re}(\alpha) > 0 \quad \text{und} \quad z > 0. \quad (\text{C-28})$$

Die Gammaverteilung beschreibt, wenn α eine natürliche Zahl ist, die Summe von unabhängigen, exponentiell verteilten Ereignissen, also z. B. die Zeit bis zum Eintreten einer bestimmten Anzahl von POISSON-verteilten Ereignissen. Für $\alpha > 1$ ist der Verlauf ihrer Wahrscheinlichkeitsdichte ähnlich dem der Lognormalverteilung. Bei gleichem Erwartungswert und gleicher Varianz sind extrem hohe Perzentile der Gammaverteilung jedoch etwas kleiner als die der Lognormalverteilung, wie aus Abbildung C-4b ersichtlich.

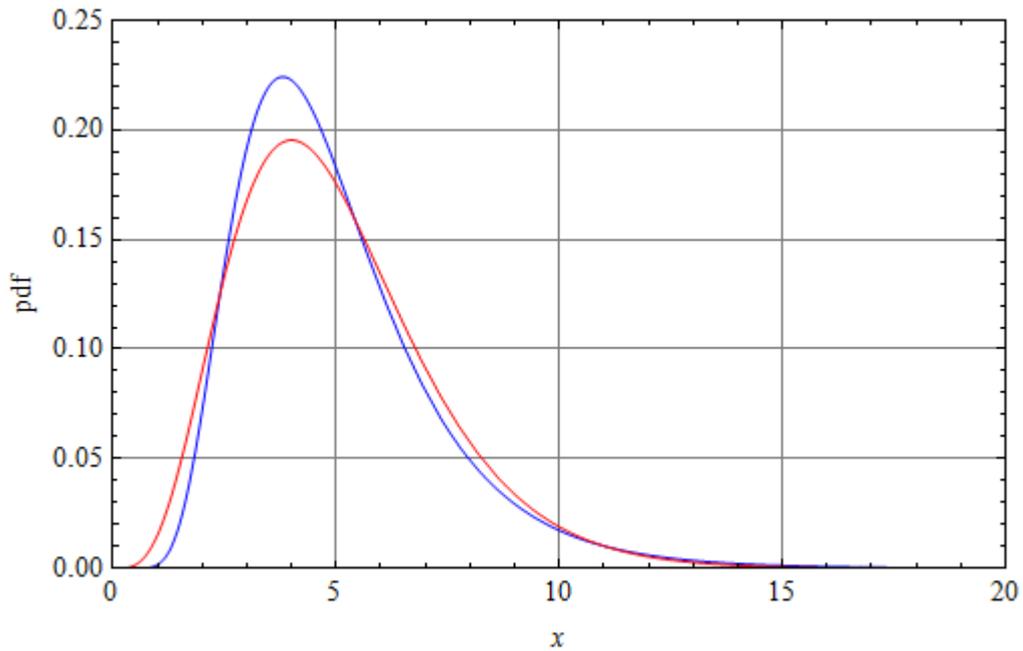
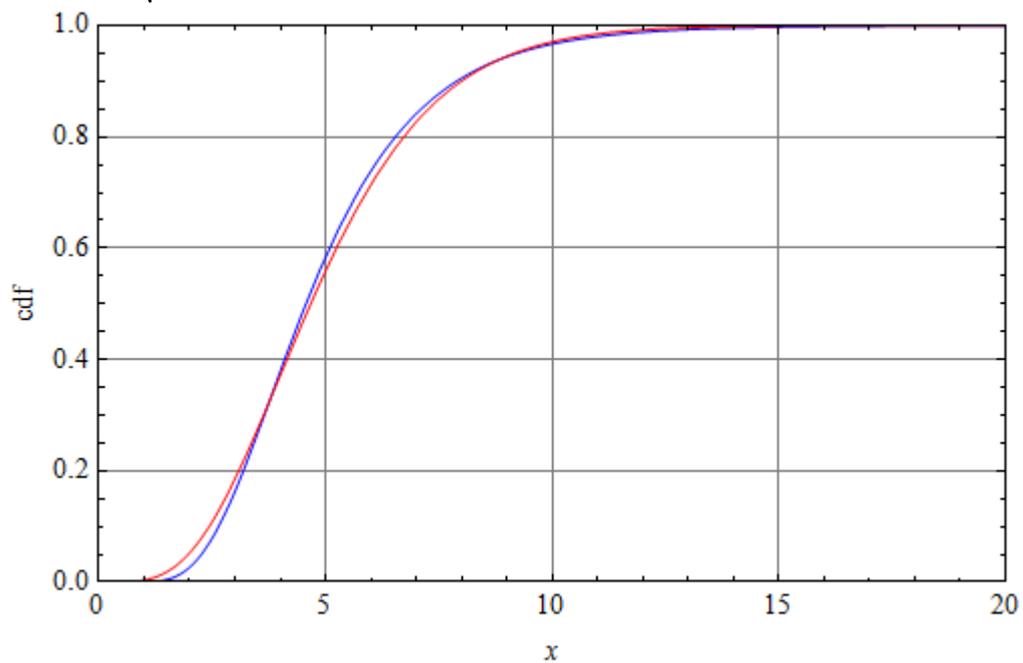


Abbildung C-4a: Vergleich der Wahrscheinlichkeitsdichte der Gammaverteilung $X \sim \text{Gam}(\alpha = 5, \lambda = 1)$ (rot) mit einer bzgl. Erwartungswert (= 5) und Varianz (= 5) angepassten Lognormalverteilung $X \sim \text{Ln}(\mu, \sigma)$; $\mu = \ln(5) - 0,5 \cdot \ln(6/5) \approx 1,52$ und $\sigma = \sqrt{\ln(6/5)} \approx 0,427$ (blau)



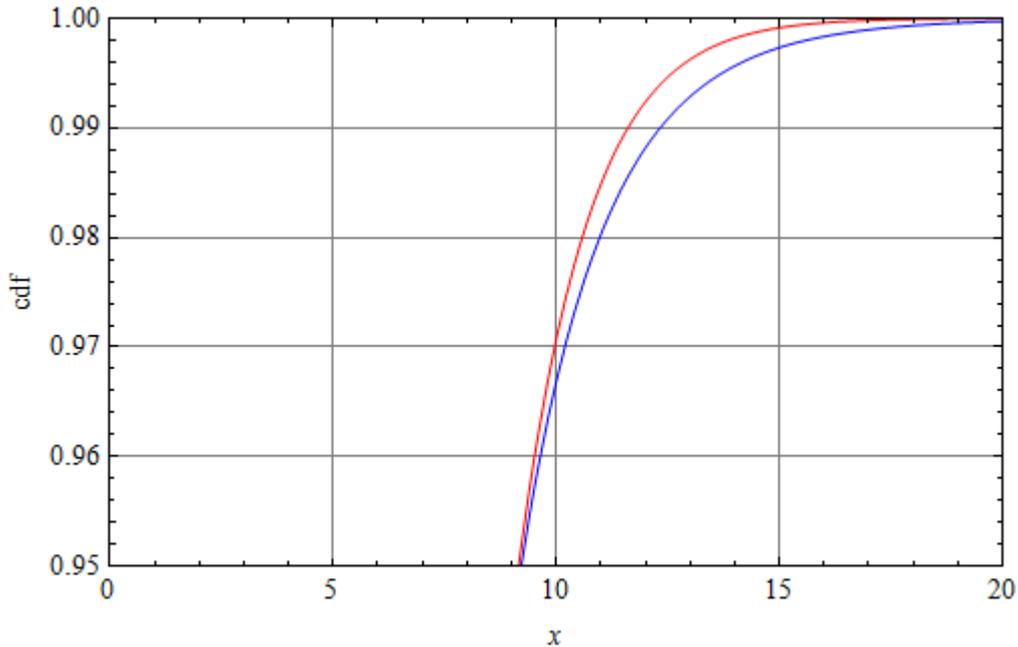


Abbildung C-4b: Vergleich der Verteilungsfunktion der Gammaverteilung $X \sim \text{Gam}(\alpha = 5, \lambda = 1)$ (rot) mit einer bzgl. Erwartungswert (= 5) und Varianz (= 5) angepassten Lognormalverteilung $X \sim \text{Ln}(\mu, \sigma)$; $\mu = \ln(5) - 0,5 \cdot \ln(6/5) \approx 1,5$ und $\sigma = \sqrt{\ln(6/5)} \approx 0,43$ (blau); das untere Bild zeigt zur Verdeutlichung des Unterschieds beider Verteilungen für extrem hohe Perzentile nur den Bereich $0,95 \leq \Pr(X \leq x) \leq 1$.

Die Gammaverteilung ist rechtsschief und leptokurtisch. Für wichtige statistische Kenngrößen, wie Erwartungswert, Varianz, Schiefe und Kurtosis sowie Modalwert gelten folgende Beziehungen:

$$E_{\text{Gam}}[X] = \alpha/\lambda; \quad \text{Var}_{\text{Gam}}[X] = \alpha/\lambda^2; \quad \gamma_1 = 2/\sqrt{\alpha}; \quad \gamma_2 = 3 + 6/\alpha; \quad (\text{C-29})$$

$$\text{Mod}_{\text{Gam}} = (\alpha-1)/\lambda \quad \text{für } \alpha > 1.$$

Spezialfälle der Gammaverteilung sind die Exponentialverteilung ($\alpha = 1$), die ERLANG-Verteilung ($\alpha = n$, natürliche Zahl), sowie die Chi-Quadrat-Verteilung (für $\alpha = \nu/2$ folgt die transformierte Variable $Y = 2 \cdot \lambda \cdot X$ einer χ^2 -Verteilung mit dem Freiheitsgrad ν). Für sehr große Werte von α nähert sich die Gammaverteilung einer Normalverteilung. Die Gammaverteilung ist flexibler als die Lognormalverteilung, was u. a. darin zum Ausdruck kommt, dass mit α -Werten von $0 < \alpha \leq 1$ auch Wahrscheinlichkeitsdichten dargestellt werden können, deren Modalwert bei $x = 0$ liegt bzw. die für $x \rightarrow 0$ singulär werden. Die Abbildung C-5 verdeutlicht die Flexibilität der Gammaverteilung.

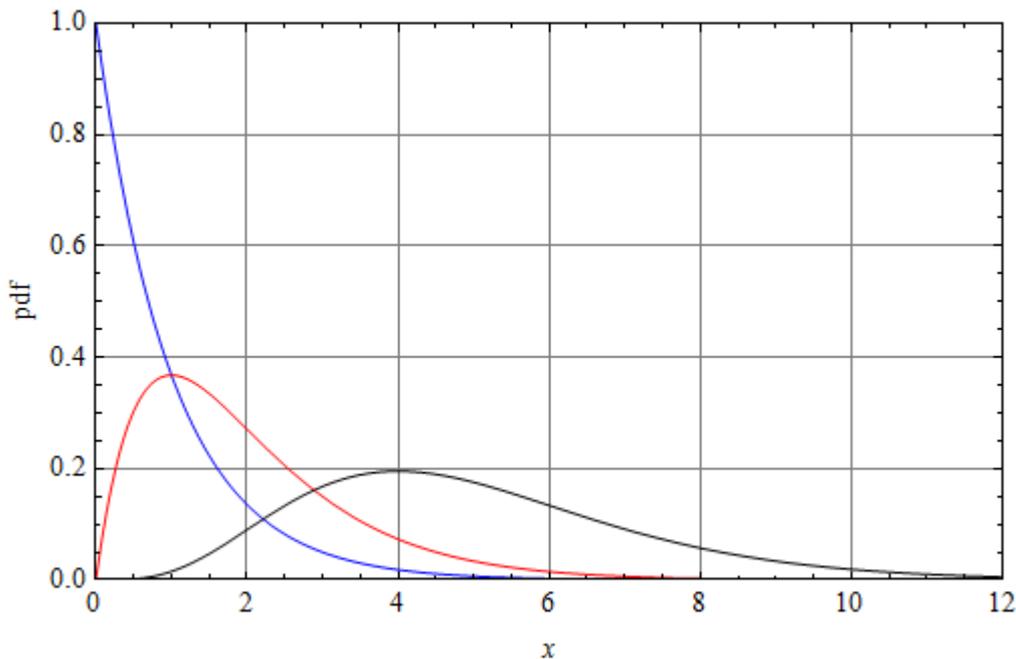


Abbildung C-5: Wahrscheinlichkeitsdichten der Gammaverteilung für $\lambda = 1$ und drei verschiedene Werte des Formparameters α : $\alpha = 1$ (blau), $\alpha = 2$ (rot) und $\alpha = 5$ (schwarz)

Es sei darauf hingewiesen, dass zur **Parametrisierung der Gammaverteilung** häufig statt der Rate λ der Skalenparameter $\beta = 1/\lambda$ verwendet wird (dann wird Bezeichnung $\mathbf{X} \sim \mathbf{Ga}(\alpha, \beta)$ gewählt).

Die Gammaverteilung wird im Rahmen von probabilistischen Expositionsmodellen sowohl für die Modellierung von meteorologischen Ereignissen (z. B. Niederschläge, Stürme etc.) als auch für die Verteilung von Schadstoffkonzentrationen verwendet. Darüber hinaus ist hervorzuheben, dass sie für verschiedene statistische Modelle von zentraler Bedeutung ist, z. B. als konjugierter Prior für BAYES-Modelle oder als Hyperverteilung für einen Verteilungsparameter elementarer statistischer Verteilungen zur Ableitung von komplexeren Verteilungsfamilien (siehe z. B. Abschnitte C.3.6 und C.5.5). In diesem Zusammenhang sei vermerkt, dass die Wahrscheinlichkeitsdichte $f_{\text{Gam}}(x|\alpha, \lambda)$ gegen eine DIRAC'sche δ -Funktion konvergiert, wenn ihre beiden Verteilungsparameter α und λ gleichzeitig gegen ∞ streben, ihre Relation jedoch konstant bleibt:

$$\lim_{\substack{\alpha, \lambda \rightarrow \infty \\ \alpha/\lambda = \mu = \text{const}}} f_{\text{Gam}}(x|\alpha, \lambda) = \delta(x - \mu). \quad (\text{C-30})$$

Dies ist intuitiv aus Gl. (C-29) einsichtig, da die Wahrscheinlichkeitsdichte der Normalverteilung, $f_{\text{N}}(x|\mu, \sigma)$, für $\sigma \rightarrow 0$ gegen $\delta(x - \mu)$ konvergiert. Abbildung C-6 veranschaulicht die asymptotische Beziehung Gl. (C-30).

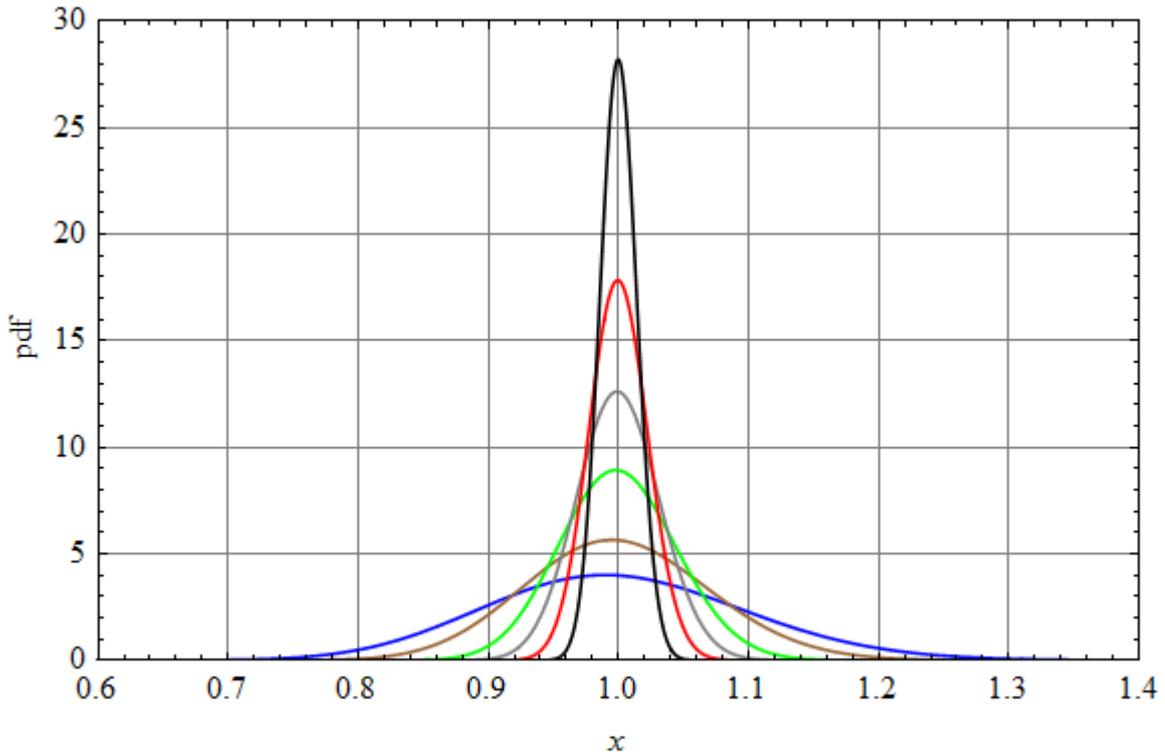


Abbildung C-6: Wahrscheinlichkeitsdichten der Gammaverteilung für Parameterwerte von $\alpha = \lambda$ von 100 (blau), 200 (braun), 500 (grün), 1000 (grau), 2000 (rot) und 5000 (schwarz) zur Illustration der Annäherung an die DIRAC'sche $\delta(x-1)$ gemäß Gl. (C-30)

Eine wichtige Eigenschaft der Gammaverteilung ist ihre Additivität bei gleichem Ratenparameter. Für unabhängige $X_1 \sim \text{Gam}(\alpha_1, \lambda)$ und $X_2 \sim \text{Gam}(\alpha_2, \lambda)$ gilt: $X_1 + X_2 \sim \text{Gam}(\alpha_1 + \alpha_2, \lambda)$.

C.3.5 WEIBULL-Verteilung $X \sim \text{We}(\alpha, \beta)$

Die WEIBULL-Verteilung stellt eine Verallgemeinerung der Exponentialverteilung dar, die aus einer Transformation einer exponentiell verteilten Zufallsgröße $Y \sim \text{Ex}(1/\beta)$ durch eine Potenzfunktion $Y = (X/\beta)^\alpha$ resultiert (für die WEIBULL-Verteilung ist es üblich, die Exponentialverteilung nicht mit Hilfe der Hazardrate λ , sondern mit dem Skalenparameter β zu parametrisieren). Abhängig davon, ob β mitpotenziert wird oder ob nur die Zufallsgröße X potenziert wird, resultieren zur WEIBULL-Verteilung voneinander abweichende Darstellungen. Mit der obigen Transformation ist sie durch die folgenden Gleichungen für die Wahrscheinlichkeitsdichte bzw. Verteilungsfunktion definiert:

$$f_{\text{We}}(x | \alpha, \beta) = \frac{\alpha \cdot (x/\beta)^{\alpha-1}}{\beta} \cdot \exp\left(- (x/\beta)^\alpha\right), \quad 0 \leq x < \infty, \quad \alpha, \beta > 0, \quad (\text{C-31})$$

$$F_{\text{We}}(x | \alpha, \beta) = 1 - \exp\left(- (x/\beta)^\alpha\right). \quad (\text{C-32})$$

Für den Erwartungswert und die Varianz der Weibull-Verteilung ergeben sich die Beziehungen

$$E_{\text{We}}[X] = \beta \cdot \Gamma(1 + 1/\alpha); \quad \text{Var}_{\text{We}}[X] = \beta^2 \cdot \left(\Gamma(1 + 2/\alpha) - \Gamma^2(1 + 1/\alpha) \right). \quad (\text{C-33})$$

Ihr Modalwert ist

$$\text{Mod}_{\text{we}} = \beta \cdot \left(\frac{\alpha - 1}{\alpha} \right)^{1/\alpha} \quad \text{für } \alpha > 1. \quad (\text{C-34})$$

Zur Schiefe und Kurtosis der WEIBULL-Verteilung lassen sich ebenfalls analytische Ausdrücke angeben. Sie ist für $\alpha < 3,60235$ linkssteil und für $\alpha > 3,60235$ rechtssteil (wie z. B. in Abbildung C-6 dargestellt. Für $2,252 < \alpha < 5,773$ ist die WEIBULL-Verteilung platykurtisch, sonst leptokurtisch.

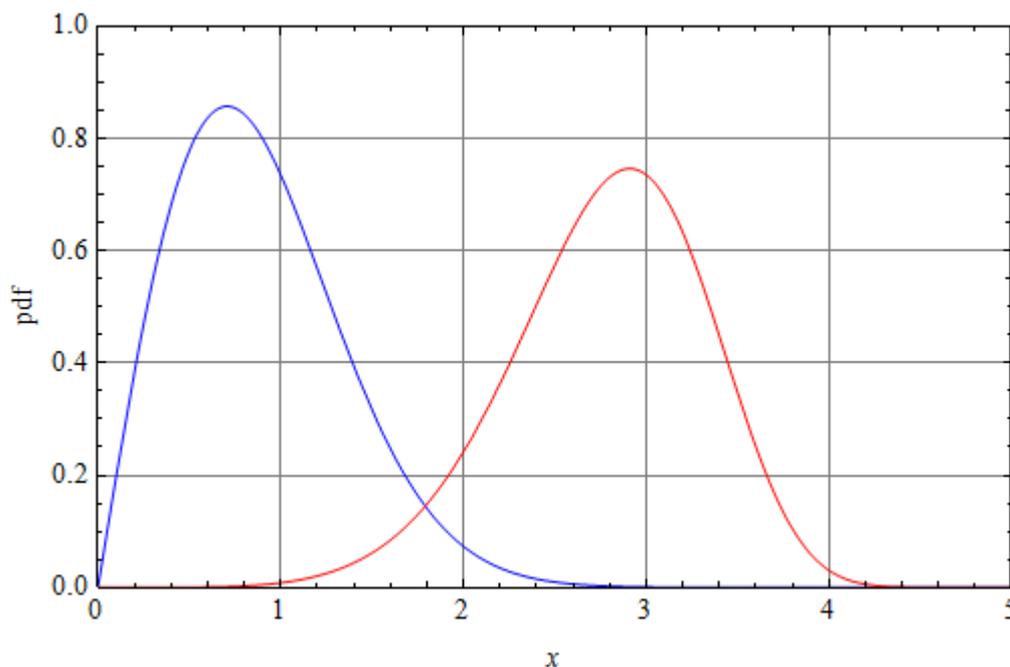


Abbildung C-6: Wahrscheinlichkeitsdichten einer linkssteilen ($X \sim \text{We}(2,1)$, blau gezeichnet) sowie einer rechtssteilen ($X \sim \text{We}(6,3)$, rot gezeichnet) WEIBULL-Verteilung

Die WEIBULL-Verteilung wird wegen ihrer ausgeprägten Flexibilität für unterschiedlichste Zwecke der probabilistischen Modellierung genutzt [WEIB 51], [JOH 94], u. a. für Korngrößenverteilungen, Festigkeitsanalysen, Lebensdauer- bzw. Zuverlässigkeitsanalysen und für diverse meteorologische Parameter. In Abhängigkeit vom Wert des Verteilungsparameters α können stochastische Prozesse mit fallender ($\alpha < 1$) bzw. mit steigender ($\alpha > 1$) Hazardrate

$$h_{\text{we}}(x) = \frac{\alpha}{\beta} \cdot (x/\beta)^{\alpha-1} \quad (\text{C-35})$$

modelliert werden (für $\alpha = 1$ resultiert die Exponentialverteilung mit konstanter Hazardrate). Die Weibull-Verteilung ist zudem eine von drei Arten von Extremwert-Verteilungen [KOTZ 00A].

C.3.6 LOMAX-Verteilung $X \sim \text{Lo}(\alpha, \eta)$

Die LOMAX-Verteilung (auch als PARETO-Verteilung Typ II bekannt; siehe [JOH 94]), wurde in [LOM 54] empirisch aus der statistischen Analyse der Überlebenszeit von Firmen unterschiedlicher Branchen abgeleitet. Ihre Wahrscheinlichkeitsdichte und Verteilungsfunktion sind definiert durch:

$$f_{\text{Lo}}(x | \alpha, \eta) = \frac{\alpha / \eta}{(1 + x / \eta)^{\alpha+1}}, \quad 0 \leq x < \infty, \quad \alpha > 0, \eta > 0, \quad (\text{C-36})$$

$$F_{\text{Lo}}(x | \alpha, \eta) = 1 - \frac{1}{(1 + x / \eta)^\alpha}, \quad (\text{C-37})$$

wobei α einen Form- und η einen Skalenparameter in der üblichen Parametrisierung darstellen. Für den Erwartungswert und die Varianz erhält man die Beziehungen

$$E_{\text{Lo}}[X] = \frac{\eta}{\alpha - 1} \quad (\text{für } \alpha > 1); \quad \text{Var}_{\text{Lo}}[X] = \frac{\eta^2}{(\alpha - 1)^2} \cdot \frac{\alpha}{\alpha - 2} \quad (\text{für } \alpha > 2). \quad (\text{C-38})$$

Zur Schiefe und Kurtosis der LOMAX-Verteilung lassen sich ebenfalls analytische Ausdrücke angeben. Sie ist rechtsschief und leptokurtisch. Das 3. bzw. das 4. zentrale Moment existieren jedoch nur für $\alpha > 3$ bzw. $\alpha > 4$:

$$\gamma_1 = \frac{2 \cdot (\alpha + 1)}{\alpha - 3} \cdot \sqrt{\frac{\alpha - 2}{\alpha}} \quad (\text{für } \alpha > 3); \quad \gamma_2 = 9 + \frac{81}{\alpha - 4} - \frac{32}{\alpha - 3} - \frac{1}{\alpha} \quad (\text{für } \alpha > 4) \quad (\text{C-39})$$

Die LOMAX-Verteilung trägt der Tatsache Rechnung, dass für bestimmte Systeme die Hazardrate mit der Zeit fällt. Dies trifft für Firmen z. B. deshalb zu, weil ihre Überlebenschance am Markt durch im Laufe der Zeit gewonnene Erfahrungen und Verbindungen wächst, die Wahrscheinlichkeit ihres Versagens (die Hazardrate) somit sinkt. Aus Gl. (C-36) und (C-37) resultiert die Hazardrate

$$h_{\text{Lo}}(x | \alpha, \eta) = \frac{f_{\text{Lo}}(x | \alpha, \eta)}{1 - F_{\text{Lo}}(x | \alpha, \eta)} = \frac{\alpha / \eta}{1 + x / \eta}. \quad (\text{C-40})$$

Ein ähnliches Verhalten kann sicherlich auch für den Verbleib von Schadstoffkonzentrationen in einem Aquifer oder von Schadstoffablagerungen auf Pflanzen unterstellt werden. Zuerst werden mit einer höheren Rate jene Spezies abgereichert, die relativ schnell an Bodenpartikel sorbieren bzw. relativ leicht durch Witterungseinflüsse von den Pflanzen entfernt werden. Im Laufe der Zeit verbleiben dann hauptsächlich nur noch solche Spezies im Grundwasser bzw. auf den Pflanzen, deren Reduzierung aufgrund von besonderen chemischen bzw. physikalischen Eigenschaften mit einer geringen Rate erfolgt, die letztlich gegen Null tendiert.

Nachfolgend wird auf Basis dieser Überlegungen eine einfache mathematische Begründung der LOMAX-Verteilung gegeben. Es wird angenommen, dass den einzelnen Individuen einer Population (der statistischen Grundgesamtheit) zwar jeweils eine bestimmte Versagensrate λ zugeschrieben und ihre Überlebenswahrscheinlichkeit mit Hilfe einer Exponentialverteilung dargestellt werden kann, dass diese jedoch innerhalb der Population variieren kann. Geht man davon aus, dass die statistische Verteilung des Parameters λ in der Population durch eine Gammaverteilung hinreichend gut approximiert werden kann, dann ergibt sich die Wahrscheinlichkeitsdichte für die auf diese

Grundgesamtheit bezogene Zufallsgröße X durch die mit einer Gammaverteilung für den Parameter λ gewichtete Mittelung von (individuellen) Exponentialverteilungen:

$$f_{\text{Lo}}(x | \alpha, \eta) = \int_0^{\infty} f_{\text{Ex}}(x | \lambda) \cdot f_{\text{Gam}}(\lambda | \alpha, \eta) d\lambda = \int_0^{\infty} \lambda \cdot e^{-\lambda \cdot x} \cdot \frac{\lambda^{\alpha-1} \cdot \eta^{\alpha}}{\Gamma(\alpha)} \cdot e^{-\eta \cdot \lambda} d\lambda = \frac{\alpha / \eta}{(1 + x / \eta)^{\alpha+1}} \quad (\text{C-41})$$

Zu Gl. (C-41) sei angemerkt, dass der Parameter η in der Gammaverteilung zwar die Bedeutung einer Rate hat, durch die Integration in der resultierenden LOMAX-Verteilung jedoch die Funktion eines Skalenparameter erhalten hat.

Eine Anpassung der Parametrisierung der LOMAX-Verteilung an eine Exponentialverteilung kann für weitere Betrachtungen hilfreich sein. Da die Wahrscheinlichkeitsdichte der Gammaverteilung $f_{\text{Gam}}(\lambda | \alpha, \eta)$ für $\alpha / \eta = \text{const.}$ und $\alpha \rightarrow \infty$ gegen die Dirac'sche δ -Funktion $\delta(\lambda - \alpha / \eta)$ konvergiert (siehe Gl. (C-30)), hätten in diesem Grenzfall alle Individuen der mit Gl. (C-41) beschriebenen Grundgesamtheit die gleiche Versagensrate $\lambda = \alpha / \eta$. Eine entsprechende Substitution $\eta = \alpha / \lambda$ ergibt eine Parametrisierung der LOMAX-Verteilung mit den Parametern (α, λ) , wobei für sehr große Werte von α (näherungsweise, asymptotisch) eine Exponentialverteilung resultiert.

Für die Schätzung der Parameter der LOMAX-Verteilung mit der Maximum-Likelihood-Methode und Vergleiche mit der Exponentialverteilung ist zweckmäßig, die Parametrisierung der LOMAX-Verteilung durch Gleichsetzung der Erwartungswerte beider Verteilungen anzupassen. Hierfür ist gemäß Gl. (C-38) die Substitution $\eta = (\alpha - 1) / \lambda$ vorzunehmen. Zur Unterscheidung der aus dieser Substitution resultierenden Parametrisierung der LOMAX-Verteilung von der üblichen Form wird der Index "Lox" verwendet. Damit ist:

$$f_{\text{Lox}}(x | \alpha, \lambda) = \frac{\alpha}{\alpha - 1} \cdot \frac{\lambda}{(1 + \lambda \cdot x / (\alpha - 1))^{\alpha+1}}, \quad 0 \leq x < \infty, \quad \alpha > 1, \lambda > 0, \quad (\text{C-36}')$$

$$F_{\text{Lox}}(x | \alpha, \lambda) = 1 - \frac{1}{(1 + \lambda \cdot x / (\alpha - 1))^{\alpha}}, \quad (\text{C-37}')$$

$$E_{\text{Lox}}[X] = \frac{1}{\lambda} \quad (\text{für } \alpha > 1); \quad \text{Var}_{\text{Lox}}[X] = \frac{1}{\lambda^2} \cdot \frac{\alpha}{\alpha - 2} \quad (\text{für } \alpha > 2). \quad (\text{C-38}')$$

Für große Werte von α nähert sich die LOMAX-Verteilung der Exponentialverteilung an:

$$\lim_{\alpha \rightarrow \infty} f_{\text{Lox}}(x | \alpha, \lambda) = \lim_{\alpha \rightarrow \infty} \frac{\alpha}{\alpha - 1} \cdot \frac{\lambda}{(1 + \lambda \cdot x / (\alpha - 1))^{\alpha+1}} = \lambda \cdot e^{-\lambda \cdot x} = f_{\text{Ex}}(x | \lambda) \quad (\text{C-42})$$

Die Abbildung C-7 verdeutlicht die Abweichungen der LOMAX- von der Exponentialverteilung.

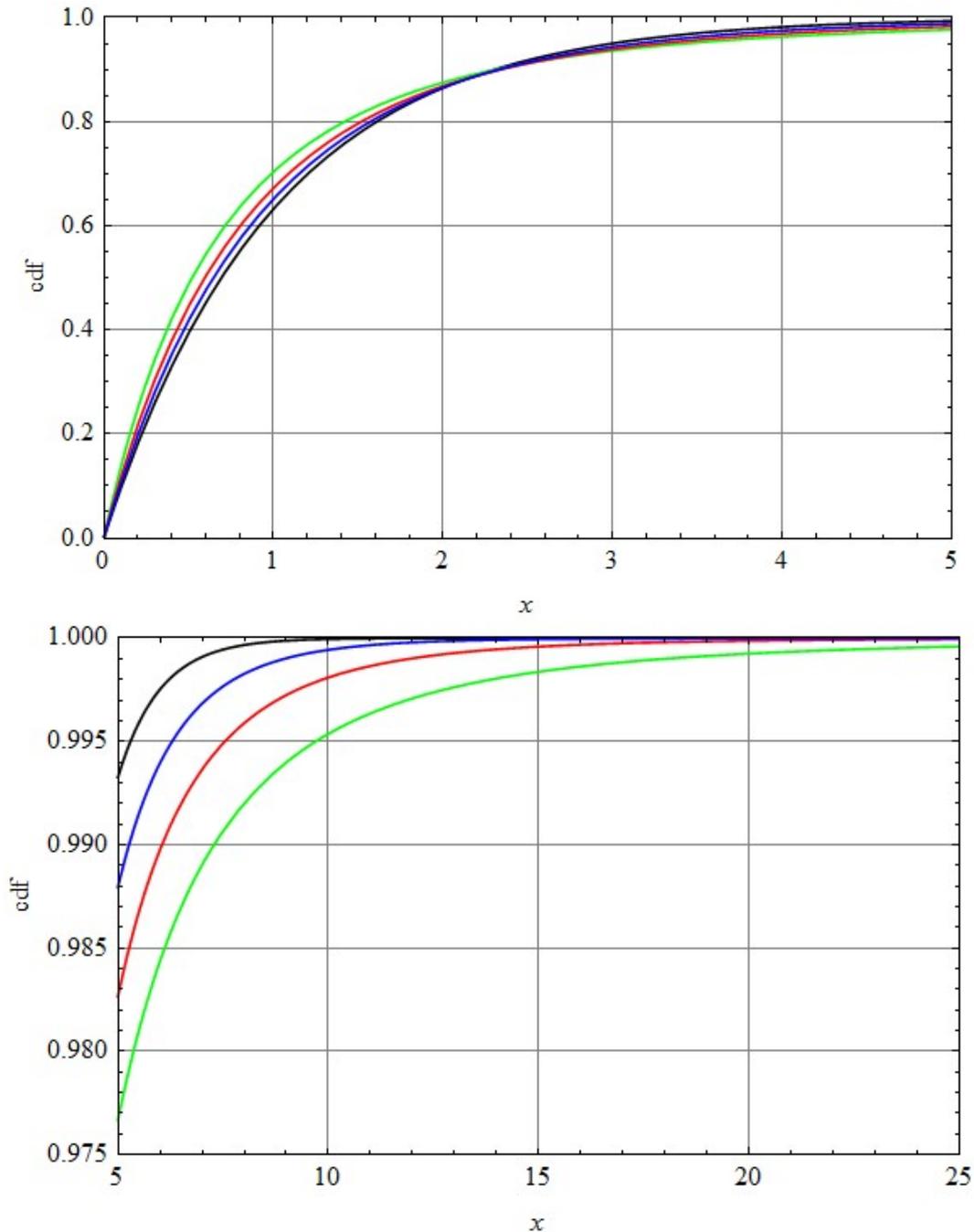


Abbildung C-7: Vergleich von Verteilungsfunktionen $F_{\text{Lomax}}(x|\alpha,\lambda)$ der LOMAX-Verteilung für $\lambda = 1$ und α -Werten von 3 (grün), 5 (rot) und 10 (blau) mit der Exponentialverteilung $F_{\text{Ex}}(x|\lambda=1)$ (schwarz)

C.3.7 WALD-Verteilung $X \sim \text{Wa}(\mu,\lambda)$

Die WALD-Verteilung (üblich auch inverse Normal- bzw. inverse GAUSS-Verteilung genannt) wurde in [SCHR 15] für die statistische Beschreibung der Zeit bis zum ersten Durchgang eines Teilchens bei BROWN'scher Bewegung mit Drift durch einen bestimmten Punkt abgeleitet (sie wurde deshalb auch als "first passage time distribution" bezeichnet). In [WALD 44] wurde diese Verteilung

als asymptotische Verteilung des Stichprobenumfangs für den sequentiellen Test von statistischen Hypothesen abgeleitet (siehe auch [WALD 47]). Wahrscheinlichkeitsdichte und Verteilungsfunktion der WALD-Verteilung sind definiert durch:

$$f_{\text{Wa}}(x | \mu, \lambda) = \sqrt{\frac{\lambda}{2\pi \cdot x^3}} \cdot \exp\left(\frac{-\lambda \cdot (x - \mu)^2}{2\mu^2 \cdot x}\right), \quad 0 \leq x < \infty, \quad \mu, \lambda > 0 \quad (\text{C-43})$$

$$F_{\text{Wa}}(x | \mu, \lambda) = \Phi\left(\sqrt{\frac{\lambda}{x}} \cdot \left(\frac{x}{\mu} - 1\right)\right) + e^{2\lambda/\mu} \cdot \Phi\left(-\sqrt{\frac{\lambda}{x}} \cdot \left(\frac{x}{\mu} + 1\right)\right), \quad (\text{C-44})$$

wobei $\Phi(x)$ die Verteilungsfunktion der Standardnormalverteilung ist, siehe Gl. (C-2). Zum Beweis von Gl. (C-44) sei z. B. auf [CHHI 74] verwiesen.

Für den Erwartungswert und die Varianz sowie für Schiefe und Kurtosis der WALD-Verteilung gelten die folgenden Beziehungen:

$$E_{\text{Wa}}[X] = \mu; \quad \text{Var}_{\text{Wa}}[X] = \frac{\mu^3}{\lambda}; \quad \gamma_1 = 3 \cdot \sqrt{\frac{\mu}{\lambda}}; \quad \gamma_2 = 3 + \frac{15 \cdot \mu}{\lambda}. \quad (\text{C-45})$$

Die WALD-Verteilung ist somit linkssteil und leptokurtisch; sie ist damit ähnlich zu anderen für positive Zufallsgrößen bekannten Verteilungen, wie z. B. Lognormalverteilung, Gammaverteilung und WEIBULL-Verteilung. Ihr Modalwert ist:

$$\text{Mod}_{\text{Wa}} = \mu \cdot \left(\sqrt{1 + \frac{9 \cdot \mu^2}{4 \cdot \lambda^2}} - \frac{3 \cdot \mu}{2 \cdot \lambda} \right). \quad (\text{C-46})$$

Aus Gl. (C-46) ist ersichtlich, dass sich der Modalwert bei fixiertem Erwartungswert (μ) für größere Werte von λ dem Erwartungswert annähert, wobei sich gemäß Gl. (C-45) die Schiefe dem Wert 0 und die Kurtosis dem Wert 3 nähern. Dies zeigt, dass die Wald-Verteilung mit zunehmendem λ gegen eine Normalverteilung mit Erwartungswert μ und Standardabweichung $\sqrt{\mu^3/\lambda}$ konvergiert. Die Abbildung C-8 veranschaulicht diesen Übergang und den Einfluss des Parameters λ auf die Form der WALD-Verteilung. Zu wichtigen statistischen Eigenschaften der WALD-Verteilung und zu Methoden der Schätzung ihrer Parameter anhand von Stichproben sei auf [TWEE 57A], [TWEE 57B] sowie [FOL 78] verwiesen. Die letztgenannte Arbeit enthält auch diverse Anwendungen der WALD-Verteilung, z. B. zur Lebensdaueranalyse und zur Analyse von Niederschlagsdaten (Rate, Abfluss). Neben der Lebensdaueranalyse [CHHI 77] und der statistischen Modellierung des Zeitaufwands für Arbeiten [HAW 97] wird die WALD-Verteilung auf weiteren Gebieten für die Analyse und Modellierung von positiven Zufallsgrößen mit linkssteiler Verteilung genutzt.

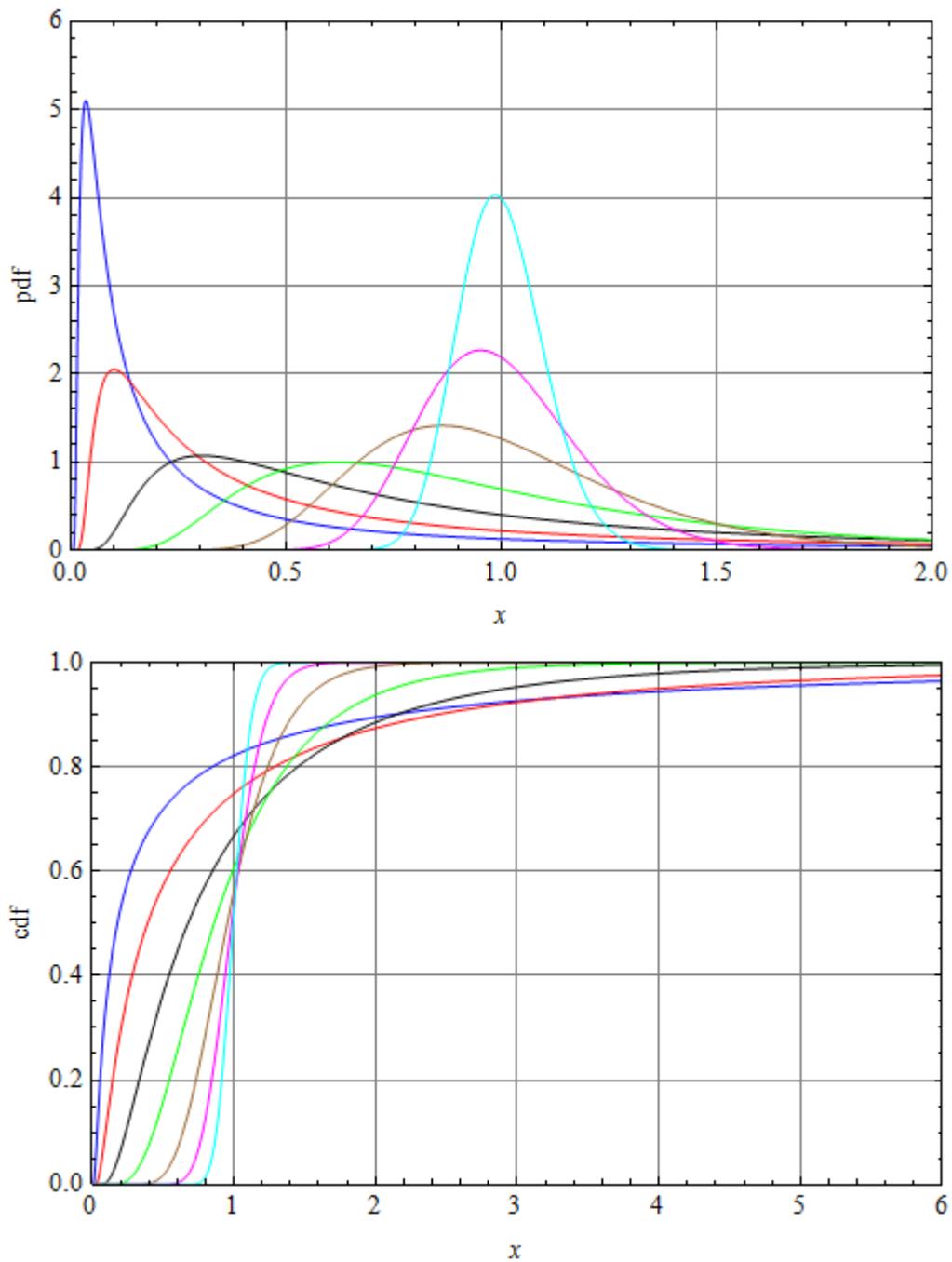


Abbildung C-8: Formen der WALD-Verteilung $X \sim \text{Wa}(\mu, \lambda)$ bei fixiertem Erwartungswert $\mu = 1$ für verschiedene Werte des Verteilungsparameters λ : 0,1 (blau), 0,3 (rot), 1 (schwarz), 3 (grün), 10 (braun), 30 (magenta), 100 (cyan); oberes Bild: Wahrscheinlichkeitsdichte für $0 < x < 2$; unteres Bild: Verteilungsfunktion für $0 < x < 6$

C.4 Stetige Verteilungen für beidseitig begrenzte Zufallsvariable ($0 \leq x \leq 1$)

Eine stetige Zufallsvariable Y , die durch eine Bedingung $a \leq Y \leq b$ beidseitig begrenzt ist, wobei a und b konstante Werte bezeichnen, kann durch die einfache lineare Transformation

$$X = \frac{Y - a}{b - a} \quad (\text{C-47})$$

in eine Zufallsgröße X überführt werden, die auf dem Intervall $[0, 1]$ definiert ist. Deshalb ist es hier ausreichend, ausschließlich Verteilungen für Zufallsvariable X zu beschreiben, für die $0 \leq X \leq 1$ gilt (sie stellen standardisierte Verteilungen dar). Es sei aber darauf hingewiesen, dass im Allgemeinen die konkreten Werte a und b , welche die Zufallsgröße Y begrenzen, nicht bekannt sein müssen, sondern, wie andere Verteilungsparameter auch, aus den vorhandenen Stichprobendaten geschätzt werden müssen. In diesem Fall ist natürlich die Schätzung der Verteilungsparameter aufwendiger als für die nachfolgend betrachteten standardisierten Verteilungen, da zwei weitere Größen (a und b) mitbestimmt werden müssen. Diese Problematik entfällt natürlich, wenn von a priori bekannt ist, dass die zu analysierende Zufallsgröße auf das Intervall $[0, 1]$ begrenzt ist, was z. B. für Variable zutrifft, die Anteile oder ähnliche Relationen darstellen.

C.4.1 Gleichverteilung $X \sim U$

Die Gleichverteilung (engl.: Uniform Distribution) ist die einfachste und in gewissem Sinne auch wichtigste statistische Verteilung, weil sie der Erzeugung anderer Zufallsgrößen zugrunde liegt. Für ihre Wahrscheinlichkeitsdichte und Verteilungsfunktion sowie für die Kenngrößen Erwartungswert, Varianz, Schiefe und Kurtosis gelten folgende Beziehungen:

$$f_U(x) = 1 \quad \text{und} \quad F_U(x) = x \quad (0 \leq x \leq 1); \quad (\text{C-48})$$

$$E_U[X] = 1/2, \quad \text{Var}_U[X] = 1/12, \quad \gamma_1 = 0, \quad \gamma_2 = 9/5. \quad (\text{C-49})$$

Die Gleichverteilung bzw. eine durch die zu Gl. (C-47) inverse Transformation dargestellte Rechteckverteilung $Re(a,b)$ wird auch für die Beschreibung von partieller Unkenntnis bzgl. einer Größe Y verwendet, für die nur der minimal und der maximal mögliche Wert spezifiziert werden können.

C.4.2 Betaverteilung $X \sim Bt(\alpha, \beta)$

Die Betaverteilung wird häufig zur Modellierung der statistischen Verteilungen zufälliger Anteile bzw. Relationen genutzt, weil sie sehr flexibel verschiedene Formen der Wahrscheinlichkeitsdichte ermöglicht (links- oder rechtsteil, lept- oder platykurtisch). Ihre Wahrscheinlichkeitsdichte ist

$$f_{Bt}(x | \alpha, \beta) = \frac{1}{B(\alpha, \beta)} \cdot x^{\alpha-1} \cdot (1-x)^{\beta-1}, \quad 0 \leq x \leq 1, \quad \alpha, \beta > 0, \quad (\text{C-50})$$

wobei $B(\alpha, \beta)$ die so genannte Betafunktion bezeichnet,

$$B(\alpha, \beta) = \int_0^1 x^{\alpha-1} \cdot (1-x)^{\beta-1} dx = \frac{\Gamma(\alpha) \cdot \Gamma(\beta)}{\Gamma(\alpha + \beta)}. \quad (\text{C-51})$$

Für die Verteilungsfunktion der Beta-Verteilung existiert keine einfache analytische Form; sie ist die regularisierte unvollständige Betafunktion:

$$F_{\text{Bt}}(x | \alpha, \beta) = \frac{B_x(\alpha, \beta)}{B(\alpha, \beta)} = \frac{1}{B(\alpha, \beta)} \cdot \int_0^x t^{\alpha-1} \cdot (1-t)^{\beta-1} dt. \quad (\text{C-52})$$

für den Erwartungswert sowie Varianz, Schiefe und Kurtosis gelten folgende Beziehungen:

$$E_{\text{Bt}}[X] = \frac{\alpha}{\alpha + \beta}, \quad \text{Var}_{\text{Bt}}[X] = \frac{\alpha \cdot \beta}{(\alpha + \beta)^2 (\alpha + \beta + 1)}, \quad (\text{C-53})$$

$$\gamma_1 = 2 \frac{(\beta - \alpha) \cdot \sqrt{\alpha + \beta + 1}}{(\alpha + \beta + 2) \cdot \sqrt{\alpha \cdot \beta}}, \quad \gamma_2 = 3 \frac{(\alpha + \beta + 1) \cdot (\alpha \cdot \beta \cdot (\alpha + \beta - 6) + 2 \cdot (\alpha + \beta)^2)}{\alpha \cdot \beta \cdot (\alpha + \beta + 2) \cdot (\alpha + \beta + 3)}. \quad (\text{C-54})$$

Ihr Modalwert ist:

$$\text{Mod}_{\text{Bt}} = \frac{\alpha - 1}{\alpha + \beta - 2} \quad \text{für } \alpha, \beta > 1; \quad (\text{Anti-Modalwert für } \alpha, \beta < 1). \quad (\text{C-55})$$

Die Abbildung C-9 verdeutlicht die Veränderung der Wahrscheinlichkeitsdichte der Betaverteilung bei Variation eines der beiden Formparameter. Für $\alpha = \beta = 1$ entspricht die Betaverteilung der Gleichverteilung. Die für $\alpha = \beta = 1/2$ resultierende Verteilung ist als "ArcSin-Verteilung" bekannt.

Die verbreitete Anwendung der Betaverteilung für die Modellierung von zufälligen Anteilen ist jedoch nicht zwangsläufig, da einige andere Verteilungen ähnliche Eigenschaften aufweisen (siehe folgende Abschnitte C.4.3 bis C.4.5). Genügen zwei positive Zufallsgrößen X und Y allerdings Gammaverteilungen mit gleichem Ratenparameter (bzw. Skalenparameter), d. h. $X \sim \text{Ga}(\alpha, \lambda)$ und $Y \sim \text{Ga}(\beta, \lambda)$, dann folgt die Relation $X/(X+Y)$ exakt einer Betaverteilung, $X/(X+Y) \sim \text{Bt}(\alpha, \beta)$.

Besondere Bedeutung hat die Betaverteilung in der BAYES-Theorie für die statistische Inferenz von Bernoulli-Versuchen als Prior- und Posteriorverteilung für die Erfolgswahrscheinlichkeit p , aber auch in der Ordnungsstatistik, wo sie die statistische Verteilung der nach der Größe geordneten Zufallswerte von Stichproben aus einer Gleichverteilung beschreibt (siehe z. B. [BAL 03]). Für eine Stichprobe $\{x_1, x_2, \dots, x_n\}$ der Zufallsvariablen $X \sim U$ werden die nach aufsteigender Größe geordneten Werte mit $x_{(k)}$ bezeichnet (k von 1 bis n). Der k -größte Wert ist dann eine Zufallsgröße, die einer Betaverteilung folgt: $X_{(k)} \sim \text{Bt}(k, n+1)$.

Schließlich sei vermerkt, dass die Betaverteilung nach entsprechender Transformation auch zur Modellierung von andersartig verteilten Zufallsgrößen genutzt wird. Ist $X \sim \text{Bt}(\alpha, \beta)$, so resultieren z. B. durch die Transformationen $Y = -\ln(X)$ oder $Y = X/(1-X)$ Zufallsgrößen mit positiven Werten ($0 < y < \infty$). Die letztgenannte Transformation führt zur so genannten PEARSON VI Verteilung. Zu dbzgl. Erläuterungen und Beispielen zu Anwendungen sei auf [JOH 95] verwiesen.

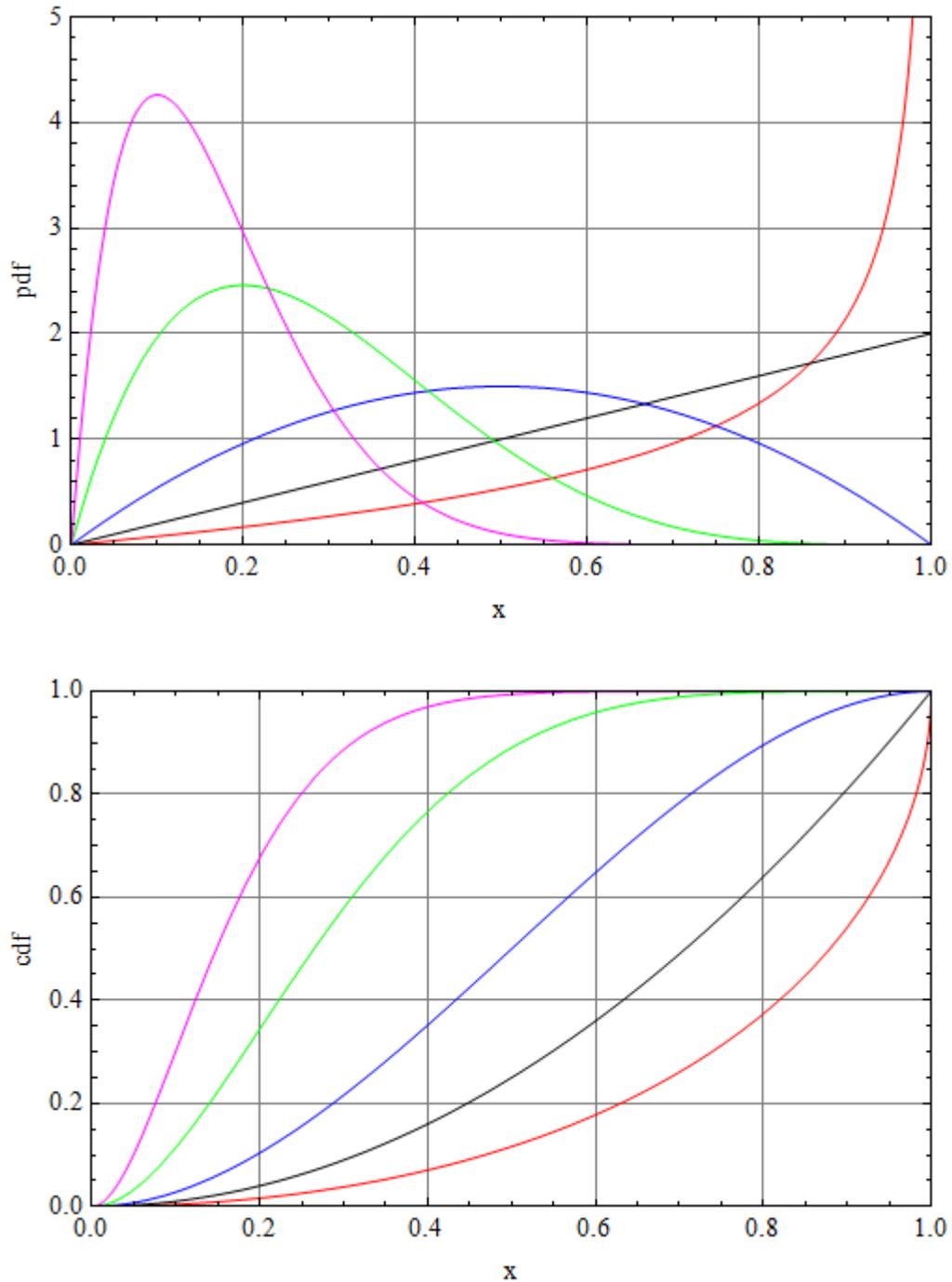


Abbildung C-9: Formen der Betaverteilung $X \sim \text{Bt}(\alpha, \beta)$ bei fixiertem Wert des 1. Formparameters von $\alpha = 2$ für verschiedene Werte des 2. Formparameters β : 0,5 (rot), 1 (schwarz), 2 (blau), 5 (grün) und 10 (magenta); oberes Bild: Wahrscheinlichkeitsdichte; unteres Bild: Verteilungsfunktion

C.4.3 KUMARASWAMY-Verteilung $X \sim \text{Ku}(\alpha, \beta)$

Die KUMARASWAMY-Verteilung wurde in [KUM 80] empirisch für die statistische Modellierung von täglichen Niederschlagsmengen entwickelt (in [LEEM 08] als Minimax-Verteilung bezeichnet). Ihre Wahrscheinlichkeitsdichte und Verteilungsfunktion sind definiert durch:

$$f_{\text{Ku}}(x | \alpha, \beta) = \alpha \cdot \beta \cdot x^{\alpha-1} \cdot (1-x)^{\beta-1}, \quad 0 \leq x \leq 1, \quad \alpha, \beta > 0, \quad (\text{C-56})$$

$$F_{\text{Ku}}(x | \alpha, \beta) = 1 - (1-x)^{\beta}. \quad (\text{C-57})$$

Für den Erwartungswert und die Varianz erhält man die Beziehungen

$$E_{\text{Ku}}[X] = \frac{1}{\alpha} B(1/\alpha, \beta + 1); \quad \text{Var}_{\text{Ku}}[X] = \frac{2}{\alpha} \cdot B(2/\alpha, \beta + 1) - \frac{1}{\alpha^2} \cdot B^2(1/\alpha, \beta + 1), \quad (\text{C-58})$$

wobei $B(*, *)$ die Betafunktion gemäß Gl. (C-51) bezeichnet. Zur Schiefe und Kurtosis lassen sich für die KUMARASWAMY-Verteilung analytische Ausdrücke berechnen, die aber kompliziert werden. Hier sei deshalb nur die Beziehung für die Schiefe angegeben:

$$\gamma_1 = \frac{2 \cdot B^3(1/\alpha, \beta + 1) - 6 \cdot \alpha \cdot B(1/\alpha, \beta + 1) \cdot B(2/\alpha, \beta + 1) + 3 \cdot \alpha^2 \cdot B(3/\alpha, \beta + 1)}{(2 \cdot \alpha \cdot B(2/\alpha, \beta + 1) - B^2(1/\alpha, \beta + 1))^{3/2}}. \quad (\text{C-59})$$

Für den Modalwert der KUMARASWAMY-Verteilung folgt aus Gl. (C-56):

$$\text{Mod}_{\text{Ku}} = \left(\frac{\alpha - 1}{\alpha \cdot \beta - 1} \right)^{1/\alpha} \quad \text{für } \alpha, \beta > 1; \quad (\text{Anti-Modalwert für } \alpha, \beta < 1). \quad (\text{C-60})$$

Die KUMARASWAMY-Verteilung weist einige Gemeinsamkeiten mit der Beta-Verteilung auf. Beide Verteilungen sind für $\{\alpha > 1, \beta > 1\}$ unimodal, für $\{\alpha < 1, \beta < 1\}$ uniantimodal und ergeben für $\alpha = \beta = 1$ die Gleichverteilung. Beide Verteilungen sind für $\{\alpha > 1, \beta \leq 1\}$ mit x wachsend und für $\{\alpha \leq 1, \beta > 1\}$ fallend. Auch verhalten sich beide Wahrscheinlichkeitsdichten an den Grenzen gleichartig, wie $x^{\alpha-1}$ für $x \rightarrow 0$ bzw. $(1-x)^{\beta-1}$ für $x \rightarrow 1$. Ein Vergleich der Wahrscheinlichkeitsdichten der beiden Verteilungen, Gl. (C-50) und Gl. (C-56), macht den direkten Zusammenhang deutlich: aus $X \sim \text{Ku}(\alpha, \beta)$ folgt $X^\alpha \sim \text{Bt}(1, \beta)$.

Ein Vergleich von Vor- und Nachteilen der KUMARASWAMY- und der Beta-Verteilung wurde in [JON 09] vorgenommen. Der wesentliche Vorteil der KUMARASWAMY-Verteilung ist die einfache Form ihrer Verteilungsfunktion, Gl. (C-57), für die auch die inverse Funktion (die Quantilfunktion) analytisch angegeben werden kann, was statistische Datenanalysen und Simulationsrechnungen stark vereinfacht. Da sie – wie die Beta-Verteilung – zur Modellierung sehr unterschiedliche Formen beidseitig begrenzter Zufallsgrößen geeignet ist (siehe Abschnitt 4.1.4.3) kann sie in vielen Fällen zweckmäßig angewandt werden. Ein Nachteil der KUMARASWAMY-Verteilung ist, dass sie nur im trivialen Fall $\alpha = \beta = 1$ symmetrisch ist.

Wie in Abschnitt C.4.2 für die Beta-Verteilung vermerkt, kann natürlich auch die KUMARASWAMY-Verteilung durch entsprechende Transformationen zur Modellierung von andersartig verteilten

Zufallsgrößen genutzt wird. Durch die Transformation $Y = -\ln(X)$ resultiert z.B. die in [GUP 99] als Alternative zu bekannten stetigen Verteilungen für positiv definierte Zufallsvariable (wie Gamma-Lognormal-, und WEIBULL-Verteilung) eingeführte "Verallgemeinerte Exponentialverteilung". Die hierfür entwickelten Methoden der statistischen Inferenz (siehe [GUP 01], [GUP 07], [KUN 08]) können durch Transformation auf die KUMARASWAMY-Verteilung übertragen werden.

C.4.4 JOHNSON-S_B-Verteilung $X \sim JS_B(\gamma, \delta)$

Die statistischen Eigenschaften stetiger Zufallsvariablen sind durch ihre Wahrscheinlichkeitsdichte definiert. Bezeichnet X eine bestimmte Zufallsgröße mit der Wahrscheinlichkeitsdichte $f_X(x|\theta)$, wobei θ für den Vektor der Verteilungsparameter steht, so definiert eine Transformation $Z = g(X)$ eine andere Zufallsgröße Z . In [JOH 49] wurden Transformationen untersucht, mit denen X auf eine Standard-Normalverteilung $Z \sim N(0,1)$ zurückgeführt wird. Ähnlich zum PEARSON'schen System (siehe Kapitel 4 in [JOH 94]) wurde θ mit vier skalare Verteilungsparametern definiert. Eine relativ allgemeine Form der Transformation mit klarer Bedeutung der Verteilungsparameter ergibt sich mit

$$Z = \gamma + \delta \cdot g\left(\frac{X - \xi}{\beta}\right), \quad \text{bzw.} \quad Y = \frac{Z - \gamma}{\delta} = g\left(\frac{X - \xi}{\beta}\right), \quad (\text{C-61a})$$

wobei Y einer Normalverteilung genügt, $Y \sim N(\gamma, \delta)$. Für die Zufallsvariable X stellt ξ einen Lageparameter und β ein Skalenparameter dar; folglich ist $\beta > 0$. Für Y ist γ ein Lage- und δ ein Skalenparameter; folglich ist $\delta > 0$. Betrachtet man die standardisierte Form für die Zufallsgröße X , so kann mit $\xi = 0$ und $\beta = 1$ von der folgenden, vereinfachten Transformation ausgegangen werden:

$$Z = \gamma + \delta \cdot g(X), \quad (\text{C-61b})$$

die nur noch von den zwei Parametern γ und δ abhängen, die für X als Formparametern fungieren (mit $g(X) = \ln(X)$ stellt X z. B. eine Lognormalverteilung $X \sim Ln(\gamma, \delta)$ dar, die als JS_L-Verteilung bezeichnet wird; mit $g(X) = \text{arsinh}(x)$ folgt die JS_U-Verteilung für reelle Zufallsgrößen, $-\infty < x < \infty$).

Eine beidseitig begrenzte Zufallsgröße (Verteilung vom Typ JS_B) entsteht mit $g(X) = \ln(X/(1-X))$. Mit Gl. (C-61b) resultieren für X die folgende Wahrscheinlichkeitsdichte und Verteilungsfunktion:

$$f_{JS_B}(x | \gamma, \delta) = \frac{\delta}{\sqrt{2 \cdot \pi \cdot x \cdot (1-x)}} \cdot \exp\left(-\frac{1}{2} \left(\gamma + \delta \cdot \ln\left(\frac{x}{1-x}\right)\right)^2\right), \quad 0 \leq x \leq 1, \quad \delta > 0, \quad (\text{C-62})$$

$$F_{JS_B}(x | \gamma, \delta) = \Phi\left(\gamma + \delta \cdot \ln\left(\frac{x}{1-x}\right)\right). \quad (\text{C-63})$$

In Gl. (C-63) bezeichnet $\Phi(x)$ die Verteilungsfunktion der Standardnormalverteilung $N(0,1)$. Für den Erwartungswert, die Varianz und die 3. bzw. 4. Momente, die zur Berechnung der Schiefe und der Kurtosis nötig sind, existieren für die JOHNSON-S_B-Verteilung keine analytischen Ausdrücke. Dies trifft auch für den Modalwert zu. Diese erheblichen Nachteile der Johnson-S_B-Verteilung beeinträchtigen jedoch nicht die Schätzung der Verteilungsparameter anhand einer Stichprobe $\{x_i\}_n$, da sich über die Transformation $y_i = \ln(y_i/(1-y_i))$ mit Gl. (C-61b) eine Stichprobe $\{y_i\}_n$ ergibt, die

normalverteilt ist. Anwendungen der JOHNSON- S_B -Verteilung für Expositionsparametern verlangen jedoch oft eine Schätzung aller vier Parameter gemäß Gl. (C-61a). Hierzu sind Quantilmethoden bzw. BAYES-Methoden geeignet (siehe z. B. [MAGE 80], [FLY 06], [TSI 01]).

Die JS_B -Verteilung ist – wie die Betaverteilung – zur Modellierung sehr unterschiedlicher Formen beidseitig begrenzter Zufallsgrößen geeignet (siehe Abschnitt 4.1.4.3). Sie kann deshalb in vielen Fällen als Alternative zur Betaverteilung angewandt werden. Wie die Betaverteilung ist sie auch zur Modellierung symmetrischer Verteilungen mit unterschiedlicher Varianz/Kurtosis geeignet.

C.4.5 Log-PEARSON-3-Verteilung $X \sim LP3(\alpha, \lambda)$

Die in [CON 71] als Log-Gamma-Verteilung untersuchte Transformation der Gammaverteilung wird nachfolgend als Log-PEARSON-3-Verteilung bezeichnet, weil die Bezeichnung "Log-Gamma" inzwischen überwiegend für eine andere Transformation der Gamma-Verteilung genutzt wird (siehe z. B. [LEEM 08] und Programm EasyFit von MathWave Technologies). Die PEARSON-Verteilung Typ 3 ist eine Gammaverteilung (siehe [JOH 94]). Die hier verwendete Bezeichnung LP3 ist somit bzgl. [CON 71] als Synonym zur Vermeidung von Verwechslungen zu verstehen.

Ist Y eine der Gammaverteilung folgende positive Zufallsgröße, $Y \sim Ga(\alpha, \lambda)$, $0 < y < \infty$ ($\alpha, \lambda > 0$), so definiert die Transformation $X = e^{-Y}$ eine auf das Intervall $[0, 1]$ begrenzte Zufallsgröße X . Die Wahrscheinlichkeitsdichte und die Verteilungsfunktion der so definierten LP3-Verteilung sind:

$$f_{LP3}(x | \alpha, \lambda) = \frac{\lambda^\alpha \cdot x^{\lambda-1} \cdot (-\ln(x))^{\alpha-1}}{\Gamma(\alpha)}, \quad 0 \leq x \leq 1, \quad \alpha > 0, \lambda > 0, \quad (C-64)$$

$$F_{LP3}(x | \alpha, \lambda) = \frac{\Gamma(\alpha, -\lambda \cdot \ln(x))}{\Gamma(\alpha)}, \quad (C-65)$$

wobei $\Gamma(\alpha, z)$ die unvollständige Gammafunktion ist (siehe Gl. (C-28)). Für den Erwartungswert und die Varianz erhält man die Beziehungen

$$E_{LP3}[X] = \left(\frac{\lambda}{\lambda + 1} \right)^\alpha; \quad \text{Var}_{LP3}[X] = \left(\frac{\lambda}{\lambda + 2} \right)^\alpha - \left(\frac{\lambda}{\lambda + 1} \right)^{2\alpha}. \quad (C-66)$$

Zur Schiefe und Kurtosis lassen sich für die LP3-Verteilung analytische Ausdrücke berechnen, die aber kompliziert werden. Für den Modalwert ergibt sich jedoch eine einfache Beziehung:

$$\text{Mod}_{LP3} = \text{Exp} \left(-\frac{\alpha - 1}{\lambda - 1} \right) \quad \text{für } \alpha, \lambda > 1; \quad (\text{Anti-Modalwert für } \alpha, \lambda < 1). \quad (C-67)$$

Die LP3-Verteilung ist – wie die Betaverteilung – zur Modellierung sehr unterschiedliche Formen beidseitig begrenzter Zufallsgrößen geeignet (siehe Abschnitt 4.1.4.3). Sie kann deshalb in vielen Fällen als Alternative zur Betaverteilung angewandt werden.

C.4.6 Dreiecksverteilung $\text{Tr}(a,b,m)$

Die Dreiecksverteilung wird oft zur Beschreibung von partieller Unkenntnis für Modellparameter verwendet, für die nur minimale (a), maximale (b) sowie wahrscheinlichste (m) Werte benannt werden können (z. B. durch Expertenschätzungen). Sie wird dementsprechend überwiegend zur Charakterisierung von epistemisch unsicheren Parametern genutzt, für die keine repräsentative Datenbasis zur Bestimmung "üblicher" statistischer Verteilungen (z. B. der Betaverteilung) vorliegt.

Eine auf dem Intervall $[a, b]$ definierte Zufallsgröße X ist dreiecksverteilt, $X \sim \text{Tr}(a,b,m)$, wenn ihre Wahrscheinlichkeitsdichte von a bis zum Modalwert $\text{Mod}_{\text{Tr}} = m$ linear ansteigt und danach von m bis b linear abfällt ($a \leq m \leq b$). Ihre Wahrscheinlichkeitsdichte und Verteilungsfunktion sind:

$$f_{\text{Tr}}(x | a, b, m) = \begin{cases} \frac{2 \cdot (x - a)}{(b - a) \cdot (m - a)} & \text{für } a \leq x \leq m \\ \frac{2 \cdot (b - x)}{(b - a) \cdot (b - m)} & \text{für } m \leq x \leq b \end{cases}, \quad (\text{C-68})$$

$$F_{\text{Tr}}(x | a, b, m) = \begin{cases} \frac{(x - a)^2}{(b - a) \cdot (m - a)} & \text{für } a \leq x \leq m \\ 1 - \frac{(b - x)^2}{(b - a) \cdot (b - m)} & \text{für } m \leq x \leq b \end{cases}. \quad (\text{C-69})$$

Für den Erwartungswert sowie für Varianz, Schiefe und Kurtosis resultieren die Beziehungen

$$E_{\text{Tr}}[X] = (a + b + m)/3; \quad \text{Var}_{\text{Tr}}[X] = (a^2 - a \cdot b + b^2 - am - bm + m^2)/18; \quad (\text{C-70})$$

$$\gamma_1 = \frac{\sqrt{2} \cdot (2a^3 - 3a^2b - 3ab^2 + 2b^3 - 3a^2m + 12abm - 3b^2m - 3am^2 - 3bm^2 + 2 \cdot m^3)}{5 \cdot (a^2 - a \cdot b + b^2 - am - bm + m^2)^{3/2}}; \quad (\text{C-71})$$

$$\gamma_2 = 12/5. \quad (\text{C-72})$$

Die Anwendbarkeit von Dreiecksverteilungen zur Approximation von Betaverteilungen für die probabilistische Modellierung von Parametern eines Expertensystems (PERT, Program Evaluation and Review Technique) wurde in [JOHN 97] untersucht. Es wurde gezeigt, dass für dbzgl. relevante Parameterbereiche der Betaverteilung eine relativ gute Approximation durch Dreiecksverteilungen vorgenommen werden kann. Dies kann z. B. durch Gleichsetzung des Modalwertes der Dreiecksverteilung mit dem Modalwert der Betaverteilung sowie Anpassung des 5%- und 95%-Perzentils der Dreiecksverteilung an die entsprechenden Perzentile der Betaverteilung erfolgen. Dies ist in der Abbildung C-10 anhand einer Anpassung der Dreiecksverteilung an $X \sim \text{Bt}(10,3)$ veranschaulicht.

Umgekehrt können mit analogen Ansätzen natürlich auch Beta- oder ähnliche Verteilungen (s. o) so an Dreiecksverteilungen angepasst werden, dass wesentliche Kenngrößen der Dreiecksverteilung, die auf Expertenschätzungen basieren, auf stetig differenzierbare Wahrscheinlichkeitsverteilungen, die "natürlicher" sind, übertragen werden. Hierzu sind dann allerdings i d. R. die 4-parametrischen Formen z. B. der Beta-, KUMARASWAMY-, JS_B oder LP3-Verteilung erforderlich.

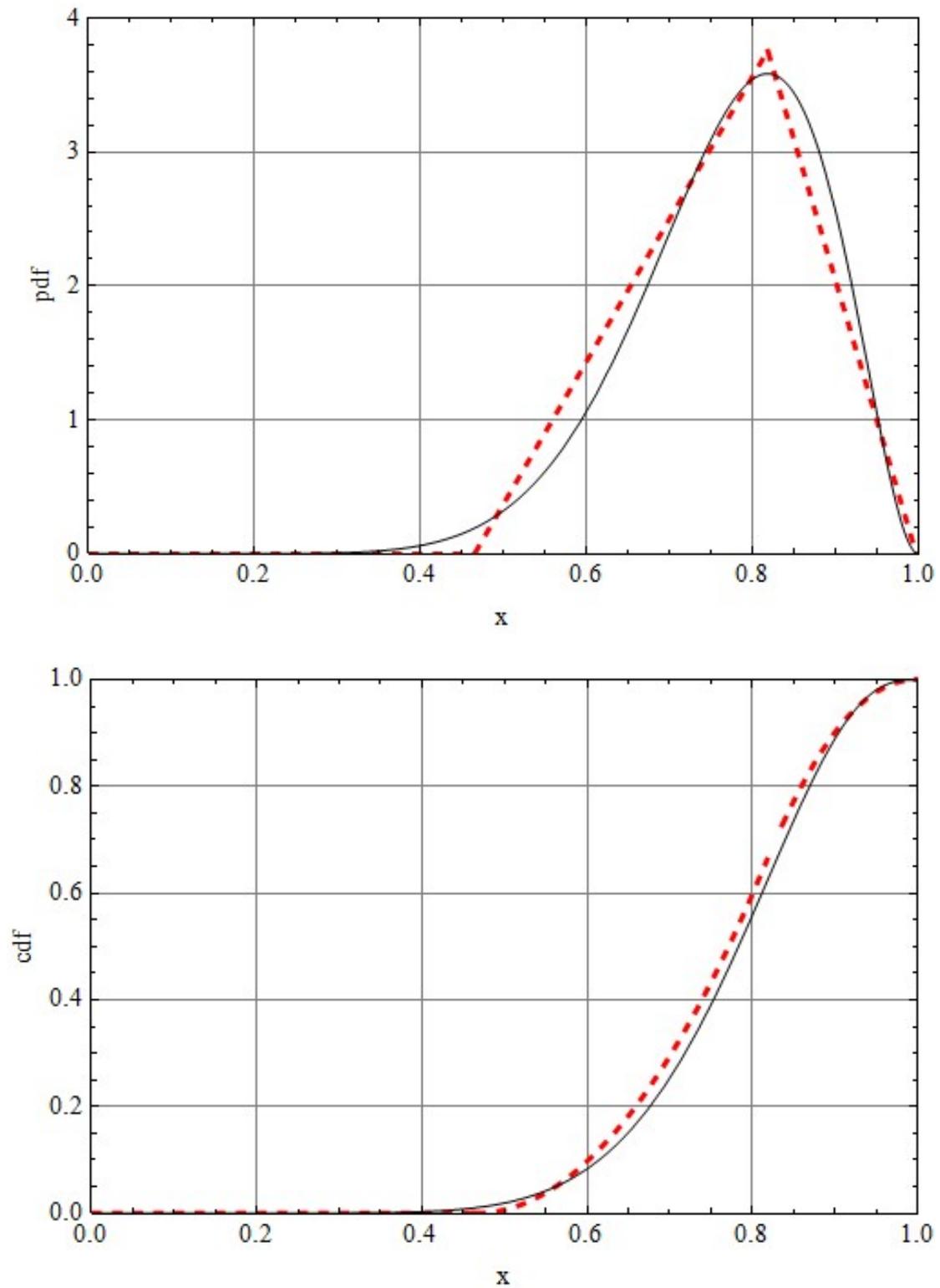


Abbildung C-10: Wahrscheinlichkeitsdichte (oben) und Verteilungsfunktion (unten) einer an die Beta-
verteilung $X \sim \text{Bt}(10,3)$ (schwarz) angepassten Dreiecksverteilung (rot, gestrichelt)
 $X \sim \text{Tr}(0.465, 0.997, 0.818)$.

C.5 Diskrete Verteilungen

C.5.1 BERNOULLI-Verteilung $X \sim \text{Ber}(p)$

Die Bernoulli-Verteilung wird zur Beschreibung von Versuchen genutzt, bei denen es nur zwei mögliche Ergebnisse gibt, die mit $X = 1$ (Erfolg) und $X = 0$ (Misserfolg) bezeichnet werden. Ist p die Erfolgswahrscheinlichkeit des Versuches, so ist

$$f_{\text{Ber}}(x|p) = \Pr(X=x|p) = p^x \cdot (1-p)^{1-x}; \quad x = 0 \text{ oder } 1; \quad 0 < p < 1; \quad (\text{C-73})$$

die Wahrscheinlichkeitsmassenfunktion (pmf). Der Erwartungswert und die Varianz sind:

$$E_{\text{Ber}}[X|p] = p, \quad \text{Var}_{\text{Ber}}[X|p] = p \cdot (1-p). \quad (\text{C-74})$$

Die BERNOULLI-Verteilung stellt die Basis für diverse andere diskrete Verteilungen dar.

C.5.2 Geometrische Verteilung $X \sim \text{Geo}(p)$

Die geometrische Verteilung beschreibt die Anzahl X von erfolglosen Bernoulli-Versuchen, die eine Erfolgswahrscheinlichkeit p haben, bis zum Eintreten des ersten Erfolges im $(X+1)$ -ten Versuch. Die pmf und die Verteilungsfunktion sind:

$$f_{\text{Geo}}(x | p) = (1-p)^x \cdot p, \quad x \geq 0 \text{ (0, 1, 2, ...)}, \quad 0 < p < 1, \quad (\text{C-75})$$

$$F_{\text{Geo}}(x | p) = 1 - (1-p)^{x+1}. \quad (\text{C-76})$$

Erwartungswert und Varianz sind:

$$E_{\text{Geo}}[X|p] = (1-p)/p, \quad \text{Var}_{\text{Geo}}[X|p] = (1-p)/p^2. \quad (\text{C-77})$$

Für die Schiefe und die Kurtosis ergeben sich folgende Beziehungen:

$$\gamma_1 = \frac{2-p}{\sqrt{1-p}}; \quad \gamma_2 = 8-p + \frac{1}{1-p}. \quad (\text{C-78})$$

C.5.3 Binomialverteilung $X \sim \text{Bin}(n, p)$

Die Binomialverteilung beschreibt die Anzahl X von erfolgreichen Versuchen bei n identischen und voneinander unabhängigen BERNOULLI-Versuchen, die eine Erfolgswahrscheinlichkeit von p haben. Die pmf der Binomialverteilung ist:

$$f_{\text{Bin}}(x | n, p) = \frac{n!}{x!(n-x)!} \cdot p^x \cdot (1-p)^{n-x}, \quad 0 \leq x \leq n, \quad 0 < p < 1. \quad (\text{C-79})$$

Erwartungswert und Varianz sind:

$$E_{\text{Bin}}[X|n,p] = n \cdot p, \quad \text{Var}_{\text{Bin}}[X|n,p] = n \cdot p \cdot (1-p). \quad (\text{C-80})$$

Für die Schiefe und die Kurtosis ergeben sich folgende Beziehungen:

$$\gamma_1 = \frac{1-2 \cdot p}{\sqrt{n \cdot p \cdot (1-p)}}; \quad \gamma_2 = 3 - \frac{6}{n} + \frac{1}{n \cdot p \cdot (1-p)}. \quad (\text{C-81})$$

Für sehr große n nähert sich die Binomialverteilung einer POISSON-Verteilung $\text{Poi}(\lambda)$ mit $\lambda = n \cdot p$ an. Dies ist in Abbildung C-11 für Parameterwerte (n, p) der Binomialverteilung von $(10; 0,3)$ sowie $(50; 0,06)$ illustriert.

C.5.4 POISSON-Verteilung $X \sim \text{Poi}(\lambda)$

Die POISSON-Verteilung dient der Beschreibung der Anzahl seltener Ereignisse, die mit konstanter Rate λ (z. B. mittlere Anzahl von Ereignissen pro Zeiteinheit) stattfinden. Sie wurde durch POISSON als Grenzverteilung der Binomialverteilung für $n \rightarrow \infty$ und $p \rightarrow 0$ mit $n \cdot p = \text{const.} = \lambda$ gefunden:

$$\lim_{n \rightarrow \infty} \frac{n!}{x!(n-x)!} \cdot \left(\frac{\lambda}{n}\right)^x \cdot \left(1 - \frac{\lambda}{n}\right)^{n-x} = \frac{e^{-\lambda} \cdot \lambda^x}{x!} = f_{\text{Poi}}(x|\lambda), \quad x = 0, 1, 2, \dots; \lambda > 0. \quad (\text{C-82})$$

Erwartungswert und Varianz der POISSON-Verteilung sind:

$$E_{\text{Poi}}[X|\lambda] = \lambda, \quad \text{Var}_{\text{Poi}}[X|\lambda] = \lambda. \quad (\text{C-83})$$

Ihre Schiefe und Kurtosis sind:

$$\gamma_1 = \frac{1}{\sqrt{\lambda}}; \quad \gamma_2 = 3 + \frac{1}{\lambda}. \quad (\text{C-84})$$

Generell wird ein POISSON-Prozess durch folgende Bedingungen definiert:

- Die Anzahl von Ereignissen in disjunkten Zeitintervallen ist voneinander unabhängig.
- Die Wahrscheinlichkeit des Auftretens von genau einem Ereignis in einem infinitesimal kleinen Zeitabschnitt dt ist $\lambda \cdot dt$ (mit $\lambda > 0$).
- Die Wahrscheinlichkeit des Auftretens von zwei oder mehreren Ereignissen in infinitesimal kleinen Zeitabschnitten ist vernachlässigbar gering.
- Die Wahrscheinlichkeit des Auftretens einer bestimmten Anzahl X von Ereignissen in einem Zeitintervall der Länge Δt hängt nur von der Länge dieses Intervalls, jedoch nicht von ihrem Anfangszeitpunkt ab (ein POISSON-Prozess hat die Eigenschaft der "Gedächtnislosigkeit").

Für die Modellierung von diskreten Zufallsgrößen ist die Reproduktivität der POISSON-Verteilung von besonderer Bedeutung. Sind $X_1 \sim \text{Poi}(\lambda_1)$ und $X_2 \sim \text{Poi}(\lambda_2)$ zwei unabhängige Zufallsgrößen, so folgt ihre Summe ebenfalls einer POISSON-Verteilung: $X_1 + X_2 = Y \sim \text{Poi}(\lambda_1 + \lambda_2)$.

Für große Werte des Parameters λ nähert sich die POISSON-Verteilung einer Normalverteilung mit dem Erwartungswert λ und der Standardabweichung $\sqrt{\lambda}$ an, wie in Abbildung C-12 dargestellt.

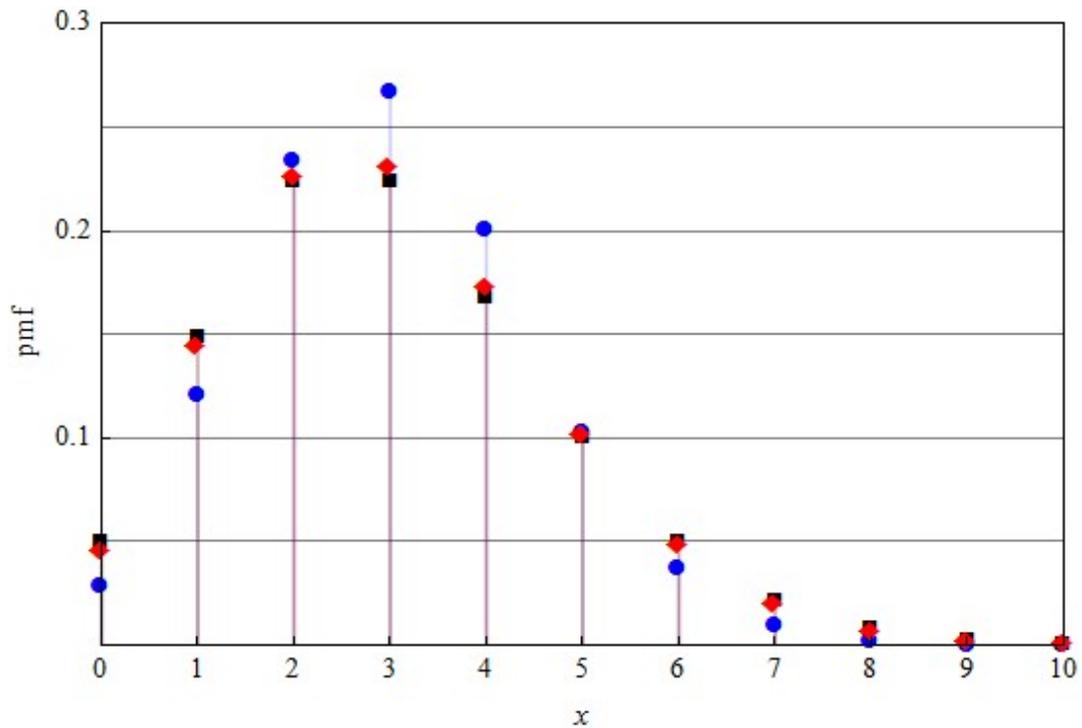


Abbildung C-11: Annäherung der Binomialverteilung an eine POISSON-Verteilung; schwarz: $X \sim \text{Poi}(3)$; blau: $X \sim \text{Bin}(10; 0,3)$; rot: $X \sim \text{Bin}(50; 0,06)$

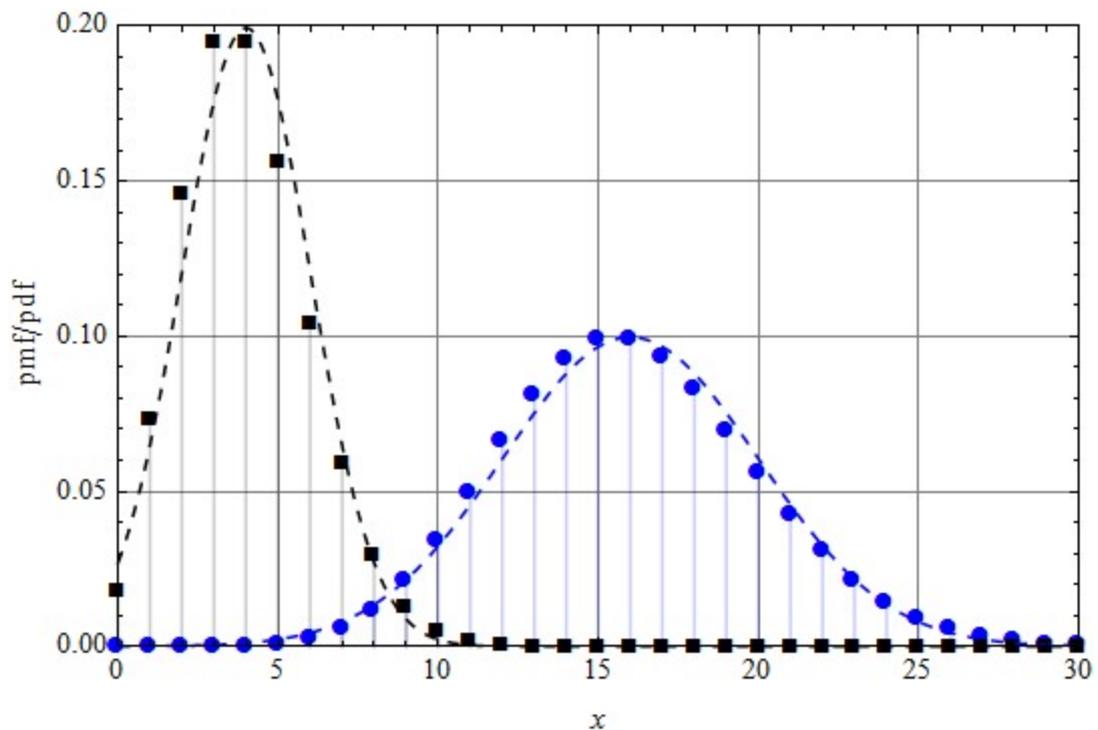


Abbildung C-12: Annäherung der POISSON-Verteilung an eine Normalverteilung; schwarz: $X \sim \text{Poi}(4)$ und $X \sim N(4,2)$ (gestrichelt); blau: $X \sim \text{Poi}(16)$ und $X \sim N(16,4)$ (gestrichelt)

C.5.5 Negative Binomialverteilung $X \sim \text{NBin}(r,p)$ oder $\text{NegB}(\alpha,\beta)$ oder $\text{NB}(\lambda,\delta)$

Nach der klassischen Definition der Negativen Binomialverteilung beschreibt diese die Anzahl X von erfolglosen iid Bernoulli-Versuchen, die eine Erfolgswahrscheinlichkeit p haben, bis zum Eintreten des r -ten Erfolges im $(X+1)$ -ten Versuch. Sie ist damit gleich der Summe von r Zufallsgrößen, die einer geometrischen Verteilung $\text{Geo}(p)$ folgen (siehe Abschnitt C.5.2). Die pmf der Negativen Binomialverteilung ist:

$$f_{\text{NBin}}(x | r, p) = \frac{(r+x-1)!}{x!(r-1)!} \cdot p^r \cdot (1-p)^x, \quad x = 0, 1, 2, \dots; 0 < p < 1. \quad (\text{C-85})$$

Erwartungswert und Varianz sind:

$$E_{\text{NBin}}(X | r, p) = \frac{r \cdot (1-p)}{p}; \quad \text{Var}_{\text{NBin}}(X | r, p) = \frac{r \cdot (1-p)}{p^2}. \quad (\text{C-86})$$

Für die Schiefe und die Kurtosis ergeben sich folgende Beziehungen:

$$\gamma_1 = \frac{2-p}{\sqrt{r \cdot (1-p)}}; \quad \gamma_2 = 3 + \frac{6}{r} + \frac{p^2}{r \cdot (1-p)}. \quad (\text{C-87})$$

In den Gl. (C-85) bis (C-87) ist r eine natürliche Zahl. Die so definierte Negative Binomialverteilung wird auch als PASCAL-Verteilung bezeichnet (siehe z. B. [JOH 05], [LEEM 08]).

Für die probabilistische Modellierung der Anzahl selten auftretender Ereignisse ist eine zweite Bedeutung der Negativen Binomialverteilung von Interesse, die ihren engen Zusammenhang mit der POISSON-Verteilung verdeutlicht. In vielen Fällen kann man nicht davon ausgehen, dass in einer Population für alle Individuen die gleiche Wahrscheinlichkeitsverteilung für das Auftreten seltener Ereignisse vorliegt, wie sie durch eine POISSON-Verteilung mit festem Parameter λ (der mittleren Anzahl von Ereignissen pro Zeiteinheit) beschrieben wird. In [GRE 20] wurde angenommen, dass Risiken für Individuen zwar mit bestimmten λ -Werten beschrieben werden können, dass diese in einer Population jedoch selbst einer statistischen Verteilung unterliegen. Geht man hierzu von einer Gammaverteilung aus, d. h. $\lambda \sim \text{Ga}(\alpha, \beta)$, so resultiert für die gesamte Population zur Anzahl X selten auftretender Ereignisse die Wahrscheinlichkeitsmassenfunktion (pmf):

$$f_{\text{NegB}}(x | \alpha, \beta) = \int_0^{\infty} f_{\text{Poi}}(x | \lambda) \cdot f_{\text{Ga}}(\lambda | \alpha, \beta) d\lambda = \left(\frac{1}{1+\beta} \right)^{\alpha} \cdot \left(\frac{\beta}{1+\beta} \right)^x \cdot \frac{\Gamma(x+\alpha)}{x! \Gamma(\alpha)}, \quad \alpha, \beta > 0, \quad (\text{C-88})$$

wobei $\Gamma(\alpha)$ die Gammafunktion gemäß Gl. (C-8) bezeichnet. Setzt man in Gl. (C-88) $p = 1/(1+\beta)$, so entspricht diese Verteilung eine Negativen Binomialverteilung $\text{NBin}(\alpha, p)$, in der α nun jedoch eine beliebige positive reelle Zahl sein kann. Ist α eine natürliche Zahl, dann ist $\text{NegB}(\alpha, \beta) = \text{NBin}(\alpha, 1/(1+\beta))$. Für den Mittelwert und die Varianz der diskreten Zufallsgröße $X \sim \text{NegB}(\alpha, \beta)$ folgt:

$$E_{\text{NegB}}(X | \alpha, \beta) = \alpha \cdot \beta; \quad \text{Var}_{\text{NegB}}(X | \alpha, \beta) = \alpha \cdot \beta \cdot (1 + \beta). \quad (\text{C-89})$$

Ihre Schiefe und Kurtosis folgen aus Gl. (C-87) mit den Substitutionen $r \rightarrow \alpha$ und $p \rightarrow 1/(1+\beta)$.

Berücksichtigt man gemäß Gl. (C-30) die Konvergenz der Gammaverteilung $Ga(\alpha, \beta)$ zu einer DIRAC'schen δ -Funktion für $\alpha \rightarrow \infty$, $\beta \rightarrow 0$ mit $\alpha \cdot \beta = \lambda = \text{const.}$, so ist klar, dass die Negative Binomialverteilung $\text{NegB}(\alpha, \beta)$ in diesem Grenzfall zu einer POISSON-Verteilung $\text{Poi}(\lambda)$ wird:

$$\lim_{\substack{\alpha \rightarrow \infty, \beta \rightarrow 0 \\ \alpha \cdot \beta = \lambda = \text{const}}} f_{\text{NegB}}(x | \alpha, \beta) = f_{\text{Poi}}(x | \lambda). \quad (\text{C-90})$$

Dieser Übergang ist in Abbildung (C-13) veranschaulicht.

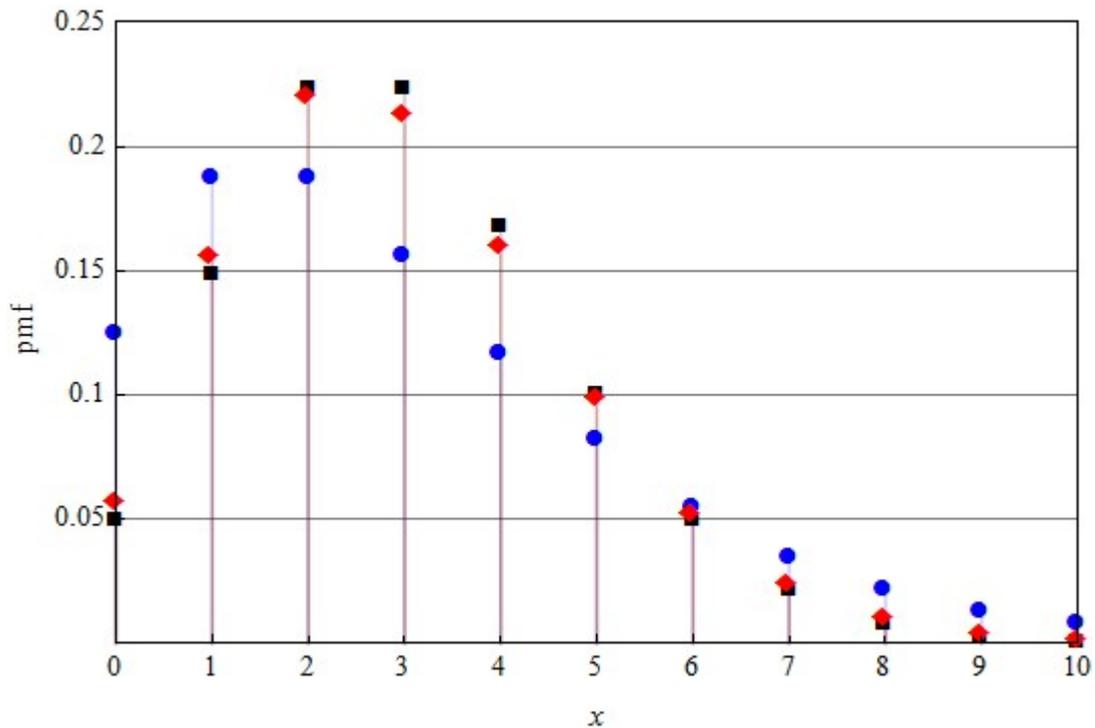


Abbildung C-13: Annäherung der Negativen Binomialverteilung an eine POISSON-Verteilung; schwarz: $X \sim \text{Poi}(3)$; blau: $X \sim \text{NegB}(3; 1)$; rot: $X \sim \text{NegB}(30; 0,1)$

Die Negative Binomialverteilung stellt somit eine Verallgemeinerung der POISSON-Verteilung für diskrete Zufallsgrößen dar, deren Varianz größer als ihr Erwartungswert ist. Die Interpretation als Mischung von POISSON-Verteilungen mit Gammaverteilung des POISSON-Parameters ist ähnlich zu der in Abschnitt C.3.6 angesprochenen Verallgemeinerung der Exponentialverteilung zur LOMAX-Verteilung. Zu weiteren Modellen der Genesis der Negativen Binomialverteilung siehe [JOH 05].

Ausgehend von Gl. (C-90) bietet sich eine Parametrisierung der Negativen Binomialverteilung an, die auf der Substitutionen $\alpha \cdot \beta = \lambda$ basiert, wobei außerdem der Parameter α durch $\delta = 1/\alpha$ ersetzt wird. Für die mit (λ, δ) parametrisierte Negative Binomialverteilung wird die Bezeichnung $X \sim \text{NB}(\lambda, \delta)$ verwendet. Aus Gl. (C-88) resultiert folgende pmf:

$$f_{\text{NB}}(x|\lambda, \delta) = \left(\frac{1}{1 + \delta \cdot \lambda}\right)^{1/\delta} \cdot \left(\frac{\delta \cdot \lambda}{1 + \delta \cdot \lambda}\right)^x \cdot \frac{\Gamma(x + 1/\delta)}{x! \Gamma(1/\delta)}, \quad \lambda, \delta > 0, \quad (\text{C-90})$$

Für den Erwartungswert und die Varianz von X resultieren aus Gl. (C-90) die Beziehungen

$$E_{\text{NB}}(X|\lambda, \delta) = \lambda; \quad \text{Var}_{\text{NB}}(X|\lambda, \delta) = \lambda \cdot (1 + \delta \cdot \lambda). \quad (\text{C-91})$$

Vergleicht man Gl. (C-91) mit den analogen Beziehungen für die POISSON-Verteilung in Gl. (C-83), so ist klar, dass der Parameter δ die Vergrößerung der Varianz der Negativen Binomialverteilung gegenüber der POISSON-Verteilung charakterisiert. Deshalb wird $\delta = 1/\alpha$ als Dispersionsparameter der Negativen Binomialverteilung bezeichnet [CLA 89], [PIE 90]. Er erfasst die erhöhte Dispersion für Realisierungen aus der NB-Verteilung im Vergleich zu Stichproben, die durch einen POISSON-Prozess erzeugt wurden. Analog zu Gl. (C-90) ergibt sich die POISSON-Verteilung als Grenzfall:

$$\lim_{\delta \rightarrow 0} f_{\text{NB}}(x|\lambda, \delta) = \frac{e^{-\lambda} \cdot \lambda^x}{x!} = f_{\text{Poi}}(x|\lambda). \quad (\text{C-90})$$