

Stochastik I

SS 2026

Inhalt

I. Maßtheoretische Grundlagen

- I.1. Mengen, Strukturen, Abbildungen
- I.2. σ -Algebren
- I.3. Maß und Wahrscheinlichkeit
- I.4. Produktmaße
- I.5. Messbare Abbildungen
- I.6. Borel'sche σ -Algebra und Lebesgue'sches Maß
- I.7. Das Lebesgue-Integral

II. Grundprinzipien stochastischer Modellierung

- II.1. Elemente der Kombinatorik
- II.2. Zufallsvariablen und ihre Verteilung, stochastische Unabhängigkeit
- II.3. Verteilungsfunktionen und Momente
- II.4. Erzeugende Funktionen
- II.5. Grundlegende Verteilungen
- II.6. Das Gesetz der Großen Zahlen
- II.7. Der Zentrale Grenzwertsatz
- II.8. Der Satz von Glivenko-Cantelli
- II.9. Bedingte Verteilungen
- II.10. Die mehrdimensionale Normalverteilung

I. Maßtheoretische Grundlagen

I.1. Mengen, Strukturen, Abbildungen

Notationen:

\in, \notin :

als Element (nicht) enthalten

\wedge, \vee, \neg :

logisches "und" / "oder" / "nicht"

\exists, \forall :

es existiert / für alle

$\subset, \subseteq, \supset, \supseteq$:

(echte) Teilmenge / Obermenge von

$\mathfrak{P}(\Omega) = \{A \mid A \subseteq \Omega\}$:

Potenzmenge von Ω

$A^c = \{\omega \in \Omega \mid \omega \notin A\}$:

Komplement der Menge A

$A \setminus B = \{\omega \in \Omega \mid \omega \in A \wedge \omega \notin B\}$:

Differenz der Mengen A und B

$A \cap B = \{\omega \in \Omega \mid \omega \in A \wedge \omega \in B\}$:

Durchschnitt der Mengen A und B

$A \cup B = \{\omega \in \Omega \mid \omega \in A \vee \omega \in B\}$:

Vereinigung der Mengen A und B

$A \oplus B = \{\omega \in \Omega \mid \omega \in A \vee \omega \in B\}$:

Vereinigung der *disjunkten* Mengen A und B

$A \times B = \{(\omega_1, \omega_2) \mid \omega_1 \in A \wedge \omega_2 \in B\}$:

kartesisches Produkt der Mengen A und B

Notationen:

$$\bigcup_{i \in I} A_i = \{\omega \in \Omega \mid \exists i \in I: \omega \in A_i\}:$$

Vereinigung der Mengen $\{A_i\}_{i \in I}$

$$\bigoplus_{i \in I} A_i = \bigcup_{i \in I} A_i:$$

Vereinigung der *paarweise disjunkten* Mengen $\{A_i\}_{i \in I}$

$$\bigcap_{i \in I} A_i = \{\omega \in \Omega \mid \forall i \in I: \omega \in A_i\}:$$

Durchschnitt der Mengen $\{A_i\}_{i \in I}$

$$\prod_{i \in I} A_i = \left\{ f: I \rightarrow \bigcup_{i \in I} A_i \mid \forall i \in I: f(i) \in A_i \right\}:$$

kartesisches Produkt der Mengen $\{A_i\}_{i \in I}$

Auswahlaxiom: Es seien sämtliche Mengen $\{A_i\}_{i \in I}$ sowie die Indexmenge I selbst nicht-leer. Dann ist auch das kartesische Produkt

$$\prod_{i \in I} A_i = \left\{ f: I \rightarrow \bigcup_{i \in I} A_i \mid \forall i \in I: f(i) \in A_i \right\}$$

nicht-leer.

Lemma 1. Es gilt für beliebige, nicht-leere Indexmengen I, J :

$$\left(\bigcup_{i \in I} A_i \right)^c = \bigcap_{i \in I} A_i^c$$

(Regeln von de Morgan)

$$\left(\bigcap_{i \in I} A_i \right)^c = \bigcup_{i \in I} A_i^c$$

$$\bigcap_{j \in J} \times_{i \in I} A_{ij} = \times_{i \in I} \bigcap_{j \in J} A_{ij}$$

$$\left(\bigcup_{i \in I} A_i \right) \cap \left(\bigcup_{j \in J} B_j \right) = \bigcup_{i \in I} \bigcup_{j \in J} (A_i \cap B_j)$$

$$\left(\bigcup_{i \in I} A_i \right) \times \left(\bigcup_{j \in J} B_j \right) = \bigcup_{i \in I} \bigcup_{j \in J} (A_i \times B_j).$$

Definition 1. Eine Teilmenge $R \subseteq \Omega \times \Omega$ heißt *Äquivalenzrelation* über Ω , wenn gilt:

$$\forall x \in \Omega: (x, x) \in R \text{ (Reflexivität)}$$

$$\forall x, y \in \Omega: (x, y) \in R \Rightarrow (y, x) \in R \text{ (Symmetrie)}$$

$$\forall x, y, z \in \Omega: (x, y) \in R \wedge (y, z) \in R \Rightarrow (x, z) \in R \text{ (Transitivität)}$$

Eine Äquivalenzrelation über Ω führt in natürlicher Weise zu einer *Klasseneinteilung*, indem man die zu jedem $x \in \Omega$ gehörige Äquivalenzklasse

$$R_x := \{y \in \Omega \mid y \sim_R x\}$$

betrachtet.

Lemma 2. Ist R eine Äquivalenzrelation über der Menge Ω , so gilt:

$$\forall x, y \in \Omega : R_x \cap R_y \neq \emptyset \Rightarrow R_x = R_y,$$

d.h. alle Äquivalenzklassen R_x, R_y sind entweder disjunkt oder fallen zusammen.

Definition 2. Zwei Mengen A und B heißen *gleichmächtig*, wenn es eine bijektive Abbildung $f: A \rightarrow B$ gibt. Ist eine Menge A gleichmächtig zu der (endlichen) Menge $B = \{1, \dots, n\}$ mit $n \in \mathbb{N}$, so bezeichnen wir mit $\#(A) = n$ die Anzahl der Elemente von A . Ist A zu keiner endlichen Menge gleichmächtig, so wählen wir entsprechend als Anzahl der Elemente $\#(A) = \infty$.

Kontinuumshypothese: Für jede Menge M mit $\mathbb{N} \subseteq M \subseteq \mathbb{R}$ gilt: $M \sim_{\#} \mathbb{N}$ oder $M \sim_{\#} \mathbb{R}$, d.h. es gibt bezüglich der Mächtigkeit keine Menge "zwischen" \mathbb{N} und \mathbb{R} .

I. Maßtheoretische Grundlagen

I.2. σ -Algebren

Definition 3 (σ -Algebra). Ein Mengensystem $\mathcal{A} \subseteq \mathfrak{P}(\Omega)$ heißt σ -Algebra (über Ω), wenn die folgenden drei definierenden Eigenschaften erfüllt sind:

$$\Omega \in \mathcal{A} \quad (\text{Normiertheit})$$

$$A \in \mathcal{A} \Rightarrow A^c \in \mathcal{A} \quad (\text{Komplementstabilität})$$

$$\{A_n\}_{n \in \mathbb{N}} \subseteq \mathcal{A} \Rightarrow \bigcup_{n=1}^{\infty} A_n \in \mathcal{A} \quad (\text{abzählbare Vereinigungsstabilität}).$$

Das Paar (Ω, \mathcal{A}) heißt *Messraum*; die Elemente von \mathcal{A} werden oft auch als (\mathcal{A}) -*messbare* Mengen bezeichnet.

Satz 1. Ein Mengensystem $\mathcal{A} \subseteq \mathfrak{P}(\Omega)$ ist genau dann eine σ -Algebra über Ω , wenn die folgenden vier (definierenden) Eigenschaften erfüllt sind:

$$\Omega \in \mathcal{A}$$

$$A \in \mathcal{A} \Rightarrow A^c \in \mathcal{A}$$

$$A, B \in \mathcal{A} \Rightarrow A \cap B \in \mathcal{A}$$

$$\{A_n\}_{n \in \mathbb{N}} \subseteq \mathcal{A} \text{ paarweise disjunkt} \Rightarrow \bigoplus_{n=1}^{\infty} A_n \in \mathcal{A}.$$

Lemma 3. Es sei Ω eine nicht-leere Menge und \mathfrak{A} ein System von σ -Algebren über Ω . Dann ist auch $\mathcal{B} = \bigcap_{\mathcal{A} \in \mathfrak{A}} \mathcal{A}$ eine σ -Algebra über Ω . Ist $\mathcal{E} \subseteq \mathfrak{P}(\Omega)$ ein beliebiges System von Teilmengen von Ω , so existiert eine bezüglich Inklusion eindeutig bestimmte kleinste σ -Algebra, die \mathcal{E} als Teilmenge enthält, nämlich $\sigma(\mathcal{E}) := \bigcap_{\mathcal{A} \in \mathfrak{A}_{\mathcal{E}}} \mathcal{A}$, wobei $\mathfrak{A}_{\mathcal{E}}$ das System *aller* σ -Algebren über Ω bezeichnet, die \mathcal{E} als Teilmenge enthalten. \mathcal{E} heißt auch *Erzeuger der σ -Algebra $\sigma(\mathcal{E})$* .

Definition 4 (Atom). Es sei (Ω, \mathcal{A}) ein Messraum. Eine nicht-leere Menge $A \in \mathcal{A}$ heißt *Atom* von \mathcal{A} , wenn A minimales Element bezüglich Inklusion ist in $\mathcal{A} \setminus \{\emptyset\}$, d.h. wenn gilt

$$B \in \mathcal{A}, B \neq \emptyset, B \subseteq A \Rightarrow B = A.$$

Die Menge \mathcal{T} aller Atome einer σ -Algebra \mathcal{A} wird als *Atomsystem* bezeichnet.

Ein Atomsystem heißt *erschöpfend*, wenn $\mathcal{A} = \sigma(\mathcal{T})$ gilt; es heißt *vollständig*, wenn es zu jedem $\omega \in \Omega$ ein Atom $T_\omega \in \mathcal{T}$ gibt mit $\omega \in T_\omega$.

Lemma 4. Es sei \mathcal{T} Atomsystem in einem Messraum (Ω, \mathcal{A}) . Dann gilt:

$$\forall S, T \in \mathcal{T} : S \cap T \neq \emptyset \Rightarrow S = T,$$

d.h. je zwei Atome einer σ -Algebra sind entweder disjunkt oder identisch. Ist das Atomsystem vollständig, so wird durch

$$\omega \sim_{\mathcal{T}} \eta \Leftrightarrow \exists T \in \mathcal{T} : \omega \in T \wedge \eta \in T$$

eine Äquivalenzrelation auf Ω definiert.

Satz 2 (Struktursatz für σ -Algebren I). Es sei (Ω, \mathcal{A}) ein Messraum mit *abzählbarer* Grundmenge Ω . Dann gilt:

- a) \mathcal{A} besitzt ein erschöpfendes, vollständiges, *abzählbares* Atomsystem $\mathcal{T} = \{T_i \mid i \in I\}$ mit einer endlichen oder unendlichen Indexmenge $I \subseteq \mathbb{N}$.
- b) Jedes Element A von \mathcal{A} besitzt eine eindeutige Darstellung

$$A = \bigoplus_{i \in I_A} T_i \quad \text{mit einer eindeutigen Indexmenge } I_A \subseteq I;$$

allgemeiner gilt (mit der Konvention $\bigcup_{i \in \emptyset} T_i = \emptyset$):

$$\mathcal{A} = \left\{ \bigoplus_{i \in J} T_i \mid J \subseteq I \right\}.$$

Ist insbesondere Ω *endlich* mit $\#(\mathcal{T}) = \#(I) = n \in \mathbb{N}$, so besitzt \mathcal{A} genau 2^n Elemente.

Satz 3 (Struktursatz für σ -Algebren II). Es sei (Ω, \mathcal{A}) ein *beliebiger* Messraum und $\mathcal{T} = \{T_i \mid i \in I\}$ ein abzählbares Atomsystem mit einer Indexmenge $I \subseteq \mathbb{N}$. Dann gilt:

\mathcal{T} ist vollständig genau dann, wenn \mathcal{T} erschöpfend ist; \mathcal{A} kann dann dargestellt werden als

$$\mathcal{A} = \left\{ \bigoplus_{i \in J} T_i \mid J \subseteq I \right\}.$$

Definition 5 (Produkt- σ -Algebra). Es seien $(\Omega_i, \mathcal{A}_i)$, $i \in I$ Messräume mit einer *endlichen* Indexmenge I . Das Produkt $\mathcal{A} = \bigotimes_{i \in I} \mathcal{A}_i$ der σ -Algebren \mathcal{A}_i , $i \in I$ (Produkt- σ -Algebra) über dem kartesischen Produkt $\Omega = \prod_{i \in I} \Omega_i$ ist definiert durch

$$\bigotimes_{i \in I} \mathcal{A}_i := \sigma(\mathcal{E}) \quad \text{mit} \quad \mathcal{E} := \left\{ \prod_{i \in I} A_i \mid \forall i \in I : A_i \in \mathcal{A}_i \right\}.$$

Das Paar $\left(\prod_{i \in I} \Omega_i, \bigotimes_{i \in I} \mathcal{A}_i \right)$ heißt entsprechend *Produkt-Messraum*.

Satz 4 (Struktursatz für σ -Algebren III). Es seien $(\Omega_i, \mathcal{A}_i)$, $i \in I$ beliebige Messräume mit einer endlichen Indexmenge I , wobei vorausgesetzt sei, dass sämtliche σ -Algebren \mathcal{A}_i vollständige, abzählbare Atomsysteme \mathcal{T}_i , $i \in I$ besitzen. Dann ist

$$\mathcal{T} := \left\{ \prod_{i \in I} T_i \mid \forall i \in I: T_i \in \mathcal{T}_i \right\}$$

das eindeutig bestimmte, erschöpfende, vollständige, abzählbare Atomsystem für die Produkt- σ -Algebra $\mathcal{A} = \bigotimes_{i \in I} \mathcal{A}_i$.

Lemma 5 (Spur- σ -Algebra). Es sei (Ω, \mathcal{A}) ein beliebiger Mesraum und $\Xi \subset \Omega$ eine nicht-leere Teilmenge. Dann wird durch

$$\mathcal{A}_\Xi := \{\Xi \cap A \mid A \in \mathcal{A}\}$$

eine σ -Algebra über Ξ definiert. Diese heißt Spur- σ -Algebra von \mathcal{A} in Ξ .

Falls bereits $\Xi \in \mathcal{A}$, so gilt einfacher

$$\mathcal{A}_\Xi := \{\Xi \cap A \mid A \in \mathcal{A}\} = \{B \in \mathcal{A} \mid B \subseteq \Xi\}.$$

I. Maßtheoretische Grundlagen

I.3. Maß und Wahrscheinlichkeit

Definition 6 (Maß und Maßraum). Es sei \mathcal{A} eine σ -Algebra über einer nicht-leeren Menge Ω . Eine Abbildung $\mu: \mathcal{A} \rightarrow \bar{\mathbb{R}} := \mathbb{R} \cup \{\infty\}$ heißt *Maß*, wenn folgende drei definierenden Eigenschaften gegeben sind:

$$\mu(\emptyset) = 0 \quad (\text{Normiertheit})$$

$$\forall A \in \mathcal{A} : \mu(A) \geq 0 \quad (\text{Nichtnegativität})$$

$$\forall \{A_n\}_{n \in \mathbb{N}} \subseteq \mathcal{A} \text{ paarweise disjunkt} : \mu\left(\bigoplus_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n) \quad (\sigma\text{-Additivität}).$$

Das Tripel $(\Omega, \mathcal{A}, \mu)$ heißt *Maßraum*. Dabei mögen in $\bar{\mathbb{R}}$ folgende erweiterte Rechenregeln gelten:

$$\forall x \in \bar{\mathbb{R}} : x + \infty = \infty + x = \infty.$$

In manchen Fällen betrachtet man auch Abbildungen μ auf einer nicht-leeren Teilmenge $\mathcal{S} \subseteq \mathcal{A}$ mit entsprechenden Eigenschaften, sofern alle beteiligten Mengen in \mathcal{S} liegen (insbesondere die disjunkte Vereinigung). Man spricht dann von einem *Prämaß* auf \mathcal{S} .

Gilt für das Maß $\mu(\Omega) = 1$, so heißt μ *Wahrscheinlichkeitsmaß* (bzw. synonym: *Wahrscheinlichkeitsverteilung*); es wird in der Regel mit dem Symbol P bezeichnet (lat.: *probabilitas*, engl.: *probability*, frz.: *probabilité*).

Jedes nicht-triviale endliche Maß μ (d.h. mit der Eigenschaft $\mu(\Omega) > 0$) kann durch Normierung zu einem Wahrscheinlichkeitsmaß P gemacht werden vermöge

$$P(A) := \frac{\mu(A)}{\mu(\Omega)}, \quad A \in \mathcal{A}.$$

Unter dem trivialen Maß μ auf \mathcal{A} verstehen wir das Maß mit der Eigenschaft $\forall A \in \mathcal{A} : \mu(A) = 0$.

Das triviale Maß wird auch als *Nullmaß* bezeichnet; allgemeiner bezeichnen wir Mengen $A \in \mathcal{A}$ mit der Eigenschaft $\mu(A) = 0$ als (μ) -*Nullmengen*. Beim trivialen Maß sind also alle messbaren Mengen Nullmengen.

Ist $A = \{\omega \in \Omega \mid \omega \text{ besitzt die Eigenschaft } \mathfrak{E}\} \in \mathcal{A}$ ein Ereignis mit $\mu(A^c) = 0$, so sagt man, die Eigenschaft \mathfrak{E} bestehe μ -fast überall. Ist (Ω, \mathcal{A}, P) ein Wahrscheinlichkeitsraum, so sagt man auch, die Eigenschaft \mathfrak{E} bestehe P -fast sicher.

Ist ferner $\Xi \in \mathcal{A}$ eine nicht-leere Teilmenge, so heißt das durch

$$\nu(A) := \mu(\Xi \cap A) \text{ für alle } A \in \mathcal{A}$$

definierte Maß das *Spurmaß* von μ bezüglich Ξ .

Lemma 6. Das Tripel $(\Omega, \mathfrak{P}(\Omega), \#)$ ist für jede nicht-leere Menge Ω stets ein Maßraum.

Definition 7 (Gleichverteilung). Ist Ω endlich, so heißt das durch die obige Normierung erhaltene Wahrscheinlichkeitsmaß P mit

$$P(A) := \frac{\#(A)}{\#(\Omega)}, \quad A \in \mathfrak{P}(\Omega)$$

die *diskrete Gleichverteilung* über Ω (auch: Laplace-Verteilung).

Satz 5. Es sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum. Dann gilt für alle Ereignisse $A, B \in \mathcal{A}$, $\{A_n\}_{n \in \mathbb{N}} \subseteq \mathcal{A}$:

$$\mu(A \oplus B) = \mu(A) + \mu(B) \text{ für } A \cap B = \emptyset \quad (\text{endliche Additivität})$$

$$A \subseteq B \Rightarrow \mu(A) \leq \mu(B) \quad (\text{Monotonie})$$

$$A \subseteq B, \mu(A) < \infty \Rightarrow \mu(B \setminus A) = \mu(B) - \mu(A) \quad (\text{Subtraktivität})$$

$$\mu(A \cup B) + \mu(A \cap B) = \mu(A) + \mu(B)$$

bzw. allgemeiner, unter der einschränkenden Voraussetzung $\mu(A_i) < \infty$, $1 \leq i \leq n$:

$$\mu\left(\bigcup_{i=1}^n A_i\right) = \sum_{i=1}^n (-1)^{i-1} \sum_{1 \leq k_1 < k_2 < \dots < k_i \leq n} \mu\left(\bigcap_{j=1}^i A_{k_j}\right) \quad (\text{Siebformel von Sylvester-Poincaré})$$

$$A_1 \subseteq A_2 \subseteq A_3 \subseteq \dots, \bigcup_{n \in \mathbb{N}} A_n = A \Rightarrow \mu(A) = \lim_{n \rightarrow \infty} \mu(A_n) \quad (\text{Stetigkeit von unten})$$

$$A_1 \supseteq A_2 \supseteq \dots, \mu(A_1) < \infty, \bigcap_{n \in \mathbb{N}} A_n = A \Rightarrow \mu(A) = \lim_{n \rightarrow \infty} \mu(A_n) \quad (\text{Stetigkeit von oben})$$

$$\mu\left(\bigcup_{i=1}^n A_i\right) \leq \sum_{i=1}^n \mu(A_i), \quad n \in \mathbb{N} \cup \{\infty\} \quad ((\sigma\text{-})\text{Sub-Additivität})$$

Satz 6 (kleiner Maßfortsetzungssatz). Es sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum und $\mathcal{T} = \{T_i \mid i \in I\}$ ein vollständiges, abzählbares Atomsystem für \mathcal{A} mit einer abzählbaren Indexmenge I . Dann ist μ bereits eindeutig durch die Werte $\mu(T_i)$, $i \in I$ bestimmt. Ist umgekehrt $\{\mu_i\}_{i \in I}$ eine Folge nicht-negativer erweiterter reeller Zahlen, so wird durch

$$\mu(T_i) := \mu_i, \quad i \in I$$

in eindeutiger Weise ein Maß μ auf \mathcal{A} definiert. Gilt zusätzlich $\sum_{i \in I} \mu_i = 1$, so ist μ ein Wahrscheinlichkeitsmaß auf \mathcal{A} . Falls Ω selbst höchstens abzählbar ist, kann μ sogar (im Allgemeinen allerdings nicht mehr eindeutig) auf die maximale σ -Algebra $\mathfrak{P}(\Omega)$ fortgesetzt werden.

Definition 8 (Semi-Ring). Es sei Ω eine nicht-leere Menge. Ein Mengensystem $\mathcal{S} \subseteq \mathfrak{P}(\Omega)$ heißt *Semi-Ring* (über Ω), wenn die folgenden drei definierenden Eigenschaften erfüllt sind:

$$\emptyset \in \mathcal{S}$$

$$S, T \in \mathcal{S} \Rightarrow S \cap T \in \mathcal{S}$$

$$S, T \in \mathcal{S} \Rightarrow \exists n \in \mathbb{N}, S_1, \dots, S_n \in \mathcal{S} \text{ paarweise disjunkt mit } S \setminus T = \bigoplus_{i=1}^n S_i.$$

Lemma 7. Es sei $\mathcal{S} \subseteq \mathfrak{P}(\Omega)$ ein Semi-Ring über einer nicht-leeren Menge Ω und $S_1, \dots, S_n \in \mathcal{S}$ für ein $n \in \mathbb{N}$. Dann existieren endlich viele Mengen $T_1, \dots, T_m \in \mathcal{S}$ mit $m \in \mathbb{N}$ derart, dass

$$\bigcup_{i=1}^n S_i = \bigoplus_{j=1}^m T_j$$

gilt. Ist allgemeiner $\{S_n\}_{n \in \mathbb{N}} \subseteq \mathcal{S}$ eine Folge von Mengen, so existiert eine weitere Folge $\{T_n\}_{n \in \mathbb{N}} \subseteq \mathcal{S}$ paarweise disjunkter Mengen mit

$$\bigcup_{n=1}^{\infty} S_n = \bigoplus_{n=1}^{\infty} T_n.$$

Definition 9. Es sei (Ω, \mathcal{A}) ein beliebiger Messraum. Eine Folge $\{A_n\}_{n \in \mathbb{N}} \subseteq \mathcal{A}$ heißt Ω *abzählbar überdeckend*, wenn gilt:

$$\Omega = \bigcup_{n=1}^{\infty} A_n.$$

Eine Mengenfunktion ν auf einem beliebigen Teilsystem $\mathcal{U} \subseteq \mathcal{A}$ heißt (dort) σ -*endlich*, wenn es eine Ω abzählbar überdeckende Folge $\{U_n\}_{n \in \mathbb{N}} \subseteq \mathcal{U}$ gibt mit der Eigenschaft

$$\forall n \in \mathbb{N}: \nu(U_n) < \infty.$$

Satz 7 (großer Maßfortsetzungssatz). Es sei (Ω, \mathcal{A}) ein beliebiger Messraum und $\mathcal{S} \subseteq \mathfrak{P}(\Omega)$ ein \mathcal{A} erzeugender Semi-Ring über Ω , d.h. es gelte $\mathcal{A} = \sigma(\mathcal{S})$. Ferner sei $\mu: \mathcal{S} \rightarrow \overline{\mathbb{R}}$ ein Prämaß, d.h. μ ist normiert, nicht-negativ und σ -additiv auf \mathcal{S} . Dann gibt es ein (nicht notwendig eindeutig bestimmtes) Maß $\tilde{\mu}$ auf $\mathcal{A} = \sigma(\mathcal{S})$ mit $\tilde{\mu}(S) = \mu(S)$ für alle $S \in \mathcal{S}$, d.h. $\tilde{\mu}$ ist Fortsetzung von μ auf \mathcal{A} . Die Fortsetzung $\tilde{\mu}$ von μ auf \mathcal{A} ist eindeutig, wenn μ auf \mathcal{S} σ -endlich ist.

I. Maßtheoretische Grundlagen

I.4. Produktmaße

Satz 8. Es seien $(\Omega_i, \mathcal{A}_i, \mu_i)$, $i \in I$ Maßräume mit einer endlichen Indexmenge I . $\mathcal{A} = \bigotimes_{i \in I} \mathcal{A}_i$ bezeichne wie üblich die Produkt- σ -Algebra der σ -Algebren \mathcal{A}_i , $i \in I$ über dem kartesischen Produkt $\Omega = \prod_{i \in I} \Omega_i$. Die Maße μ_i , $i \in I$ seien sämtlich σ -endlich. Dann gelten folgende Aussagen:

- a) Das System $\mathcal{S} := \left\{ \prod_{i \in I} A_i \mid \forall i \in I: A_i \in \mathcal{A}_i \right\}$ bildet einen Ω abzählbar überdeckenden Semi-Ring.
- b) Es existiert ein eindeutig bestimmtes Maß μ auf \mathcal{A} mit der Eigenschaft:

$$\forall A_i \in \mathcal{A}_i, i \in I: \mu\left(\prod_{i \in I} A_i\right) = \prod_{i \in I} \mu_i(A_i).$$

Definition 10 (Produktmaß). Das unter den Voraussetzungen von Satz 8 eindeutig bestimmte Maß μ heißt *Produktmaß* der Maße μ_i , $i \in I$ und wird mit dem Symbol

$$\mu = \bigotimes_{i \in I} \mu_i$$

bezeichnet.

I. Maßtheoretische Grundlagen

I.5. Messbare Abbildungen

Definition 11 (messbare Abbildung). Es seien $(\Omega_i, \mathcal{A}_i)$, $i = 1, 2$ Messräume. Eine Abbildung $T: \Omega_1 \rightarrow \Omega_2$ heißt $(\mathcal{A}_1, \mathcal{A}_2)$ -messbar (oder, wenn keine Verwechslungen möglich sind, kurz auch nur *messbar*), falls die Urbildabbildung T^{-1} die Eigenschaft

$$\forall A_2 \in \mathcal{A}_2 : T^{-1}(A_2) := \{\omega_1 \in \Omega_1 \mid T(\omega_1) \in A_2\} \in \mathcal{A}_1$$

besitzt.

Lemma 8. Es seien Ω_1 und Ω_2 nicht-leere Mengen und $T: \Omega_1 \rightarrow \Omega_2$ eine beliebige Abbildung.

a) Es sei \mathcal{A}_2 eine σ -Algebra über Ω_2 . Dann ist das System

$$\mathcal{A}_1 := T^{-1}(\mathcal{A}_2) := \{T^{-1}(A_2) \mid A_2 \in \mathcal{A}_2\}$$

aller Urbilder messbarer Mengen eine σ -Algebra über Ω_1 , so dass T (\mathcal{A}_1 - \mathcal{A}_2)-messbar ist.

b) Es sei \mathcal{A}_1 eine σ -Algebra über Ω_1 . Dann ist das System

$$\mathcal{A}_2 := \{A_2 \in \mathfrak{P}(\Omega_2) \mid T^{-1}(A_2) \in \mathcal{A}_1\}$$

eine σ -Algebra über Ω_2 , so dass T (\mathcal{A}_1 - \mathcal{A}_2)-messbar ist.

Lemma 9. Es seien Ω_1 und Ω_2 nicht-leere Mengen, $T: \Omega_1 \rightarrow \Omega_2$ eine beliebige Abbildung und \mathcal{A}_2 eine σ -Algebra über Ω_2 mit einem vollständigen (bzw. erschöpfenden), abzählbaren Atomsystem $\mathcal{T}_2 = \{T_{2i} \mid i \in I_2\}$ mit abzählbarer Indexmenge I_2 . Dann gilt:

$$\mathcal{T}_1 := T^{-1}(\mathcal{T}_2) \setminus \{\emptyset\} = \{T^{-1}(T_{2i}) \mid i \in I_2\} \setminus \{\emptyset\}$$

ist ein vollständiges (bzw. erschöpfendes), abzählbares Atomsystem für die σ -Algebra $\mathcal{A}_1 = T^{-1}(\mathcal{A}_2)$.

Lemma 10. Es seien $(\Omega_i, \mathcal{A}_i)$, $i = 1, 2$ Messräume; \mathcal{E}_2 sei ein (beliebiger) Erzeuger von \mathcal{A}_2 . Eine Abbildung $T: \Omega_1 \rightarrow \Omega_2$ ist genau dann $(\mathcal{A}_1 - \mathcal{A}_2)$ -messbar, wenn gilt:

$$T^{-1}(\mathcal{E}_2) = \{T^{-1}(E_2) \mid E_2 \in \mathcal{E}_2\} \subseteq \mathcal{A}_1.$$

Lemma 11. Es seien $(\Omega_i, \mathcal{A}_i)$, $i = 1, 2, 3$ Messräume sowie T_1 und T_2 messbare Abbildungen mit

$$T_1: (\Omega_1, \mathcal{A}_1) \rightarrow (\Omega_2, \mathcal{A}_2) \quad \text{und} \quad T_2: (\Omega_2, \mathcal{A}_2) \rightarrow (\Omega_3, \mathcal{A}_3).$$

Dann ist die Komposition $T := T_2 \circ T_1: \Omega_1 \rightarrow \Omega_3$ $(\mathcal{A}_1, \mathcal{A}_3)$ -messbar.

Lemma 12. Es seien $T_i: \Omega \rightarrow \Omega_i$ Abbildungen mit endlicher Indexmenge $I = \{1, \dots, n\}$ für $n \in \mathbb{N}$. Ferner sei \mathcal{A} eine σ -Algebra über Ω , und \mathcal{A}_i sei eine σ -Algebra über Ω_i für alle $i \in I$. Genau dann ist der Vektor (T_1, \dots, T_n) messbar bezüglich der Produkt- σ -Algebra $\bigotimes_{i \in I} \mathcal{A}_i$, wenn alle T_i $(\mathcal{A}, \mathcal{A}_i)$ -messbar sind.

Satz 9 (Faktorisierungssatz I). Es seien $T: (\Omega_1, \mathcal{A}_1) \rightarrow (\Omega_2, \mathcal{A}_2)$ und $S: (\Omega_1, \mathcal{A}_1) \rightarrow (\Omega_3, \mathcal{A}_3)$ messbare Abbildungen derart, dass $T(\Omega_1)$ abzählbar ist. Die σ -Algebra \mathcal{A}_3 enthalte alle einelementigen Teilmengen $\{\omega_3\}$ mit $\omega_3 \in \Omega_3$ als Elemente. Genau dann ist S bezüglich der von T erzeugten σ -Algebra $T^{-1}(\mathcal{A}_2)$ messbar, wenn es eine Abbildung $g: (\Omega_2, \mathcal{A}_2) \rightarrow (\Omega_3, \mathcal{A}_3)$ gibt mit der Eigenschaft

$$S = g \circ T.$$

Satz 10. Es sei $T : (\Omega_1, \mathcal{A}_1) \rightarrow (\Omega_2, \mathcal{A}_2)$ eine messbare Abbildung. Dann wird für jedes Maß μ_1 auf \mathcal{A}_1 ein Maß μ_2 auf \mathcal{A}_2 definiert vermöge

$$\forall A_2 \in \mathcal{A}_2 : \mu_2(A_2) := \mu_1(T^{-1}(A_2)).$$

Definition 12 (Bildmaß). Es sei $T : (\Omega_1, \mathcal{A}_1) \rightarrow (\Omega_2, \mathcal{A}_2)$ eine messbare Abbildung und μ_1 ein Maß auf \mathcal{A}_1 . Das nach Satz 10 eindeutig bestimmte Maß μ_2 auf \mathcal{A}_2 heißt *Bildmaß von μ_1 unter T* . Es wird mit $\mu_2 = \mu_1^T$ bezeichnet.

Definition 13 (allgemeine Produkt- σ -Algebra). Es seien $(\Omega_i, \mathcal{A}_i), i \in I$ Messräume mit einer beliebigen, nicht-leeren Indexmenge I . Dann heißt

$$\bigotimes_{i \in I} \mathcal{A}_i := \sigma(\pi_j; j \in I) = \sigma\left(\bigcup_{j \in I} \pi_j^{-1}(\mathcal{A}_j)\right)$$

die Produkt- σ -Algebra der σ -Algebren $\mathcal{A}_i, i \in I$ (über $\prod_{i \in I} \Omega_i$). Dabei bezeichnen die $\pi_j: \prod_{i \in I} \Omega_i \rightarrow \Omega_j$ für $j \in I$ die so genannten *Projektionen*, die in vereinfachter Schreibweise definiert sind über $\pi_j((\omega_i)_{i \in I}) := \omega_j$ für $(\omega_i)_{i \in I} \in \prod_{i \in I} \Omega_i$.

Satz 11 (Struktursatz für Produkt- σ -Algebren). In dem Produkt-Messraum $(\prod_{i \in I} \Omega_i, \bigotimes_{i \in I} \mathcal{A}_i)$ gilt: Zu jedem $A \in \bigotimes_{i \in I} \mathcal{A}_i$ existiert eine abzählbare Menge $J \subseteq I$ und eine Menge $B \in \bigotimes_{j \in J} \mathcal{A}_j$ derart, dass gilt: $A = B \times \prod_{i \in I \setminus J} \Omega_i$, d.h. "A hängt nur von abzählbar vielen Koordinaten ab".

Satz 12 (Produktmaßsatz). In dem Produkt-Messraum $\left(\prod_{i \in I} \Omega_i, \otimes_{i \in I} \mathcal{A}_i\right)$ mit Wahrscheinlichkeitsmaßen P_i auf $\mathcal{A}_i, i \in I$ gilt: Es existiert genau ein Wahrscheinlichkeitsmaß P auf $\otimes_{i \in I} \mathcal{A}_i$ mit der Eigenschaft

$$P(A) = P\left(B \times \prod_{i \in I \setminus J} \Omega_i\right) = \prod_{j \in J} P_j(B_j) \text{ für } A = B \times \prod_{i \in I \setminus J} \Omega_i, B = \prod_{j \in J} B_j$$

mit $B_j \in \mathcal{A}_j, j \in J \subseteq I, J$ abzählbar.

Diese allgemeinen Produktmaße liefern für die Stochastik das "Herzstück" der Modellierung, denn sie geben den formalen Rahmen dafür ab, was wir **stochastische Unabhängigkeit** nennen. Das bereits in der historischen Vorbemerkung mehrfach genannte **Gesetz der Großen Zahlen** sowie der damit verwandte **Zentrale Grenzwertsatz** beruhen ganz wesentlich auf diesem Begriff; erst diese mathematischen Lehrsätze ermöglichen eine sinnvolle Verbindung zwischen "Theorie" und "Praxis", indem sie Möglichkeiten aufzeigen, wie aus "Daten" oder "Stichproben" Rückschlüsse auf die zu Grunde liegenden Wahrscheinlichkeiten gezogen werden können. Ohne den Begriff der stochastischen Unabhängigkeit gäbe es keine schließende Statistik als wissenschaftliche Disziplin und damit letztlich auch keine empirische Überprüfbarkeit naturwissenschaftlicher Erkenntnisse.

I. Maßtheoretische Grundlagen

I.6. Borel'sche σ -Algebra und Lebesgue'sches Maß

Definition 14 (Standard-Intervall / Borel'sche σ -Algebra). Für $\mathbf{a} = (a_1, \dots, a_d)$, $\mathbf{b} = (b_1, \dots, b_d) \in \mathbb{R}^d$ heißt die Menge

$$I_{\mathbf{a}, \mathbf{b}} := \prod_{i=1}^d (a_i, b_i]$$

Standard-Intervall im \mathbb{R}^d . Dabei sei vereinbarungsgemäß $(a, b] = \emptyset$ für $b \leq a$, $a, b \in \mathbb{R}$. Die von $\mathcal{E}^d := \{I_{\mathbf{a}, \mathbf{b}} \mid \mathbf{a}, \mathbf{b} \in \mathbb{R}^d\}$ erzeugte σ -Algebra über \mathbb{R}^d heißt *Borel'sche σ -Algebra* der Dimension d und wird üblicherweise mit dem Symbol \mathcal{B}^d bezeichnet.

Lemma 13. Der Erzeuger \mathcal{E}^d ist in jeder Dimension $d \in \mathbb{N}$ ein \mathbb{R}^d abzählbar überdeckender Semi-Ring.

Satz 13. Die folgenden Mengensysteme sind – neben anderen – ebenfalls Erzeuger der Borel'schen σ -Algebra \mathcal{B}^d :

$$\mathcal{E}_O^d := \{O \subseteq \mathbb{R}^d \mid O \text{ offen}\}$$

$$\mathcal{E}_A^d := \{A \subseteq \mathbb{R}^d \mid A \text{ abgeschlossen}\}$$

$$\mathcal{E}_K^d := \{K \subseteq \mathbb{R}^d \mid K \text{ kompakt}\}.$$

Insbesondere enthält \mathcal{B}^d auch alle "gemischten" Intervalle I der Form $I = \prod_{i=1}^d E_i$ mit $E_i \in \{(a_i, b_i), (a_i, b_i], [a_i, b_i), [a_i, b_i]\}$ und wird auch von den entsprechenden Mengensystemen erzeugt.

Definition 15. Die für Standard-Intervalle vermöge

$$m^d(I_{\mathbf{a},\mathbf{b}}) = m^d\left(\prod_{i=1}^d (a_i, b_i]\right) := \prod_{i=1}^d (b_i - a_i), \quad \mathbf{a}, \mathbf{b} \in \mathbb{R}^d, \quad a_j \leq b_j \text{ für } j = 1, \dots, d$$

erklärte Mengenfunktion m^d heißt *Elementarinhalt* auf dem Semi-Ring \mathcal{E}^d .

Satz 14 (Fortsetzungssatz für den Elementarinhalt). Es existiert genau eine Fortsetzung des Elementarinhalts m^d von \mathcal{E}^d auf die Borel'sche σ -Algebra \mathcal{B}^d zu einem σ -endlichen Maß. Dieses wird unter Beibehaltung der Notation m^d als *Lebesgue-Maß* bezeichnet.

Lemma 15. Das Lebesgue-Maß m^d ist translationsinvariant, d.h. es gilt

$$m^d = (m^d)^{T_{\mathbf{a}}} \text{ mit den Abbildungen } T_{\mathbf{a}} : \mathbb{R}^d \rightarrow \mathbb{R}^d : \mathbf{x} \mapsto \mathbf{x} + \mathbf{a} \text{ für jedes feste } \mathbf{a} \in \mathbb{R}^d.$$

Satz 15 (Vitali 1905). Für jedes $d \in \mathbb{N}$ gilt: $\mathcal{B}^d \neq \mathfrak{P}(\mathbb{R}^d)$.

I. Maßtheoretische Grundlagen

I.7. Das Lebesgue-Integral

Definition 16. Es sei (Ω, \mathcal{A}) ein Messraum und $A \in \mathcal{A}$ ein Ereignis. Die durch

$$\mathbb{1}_A : \omega \mapsto \begin{cases} 1, & \text{falls } \omega \in A \\ 0, & \text{falls } \omega \notin A \end{cases}$$

definierte reellwertige Funktion heißt *Indikatorfunktion* zum Ereignis A .

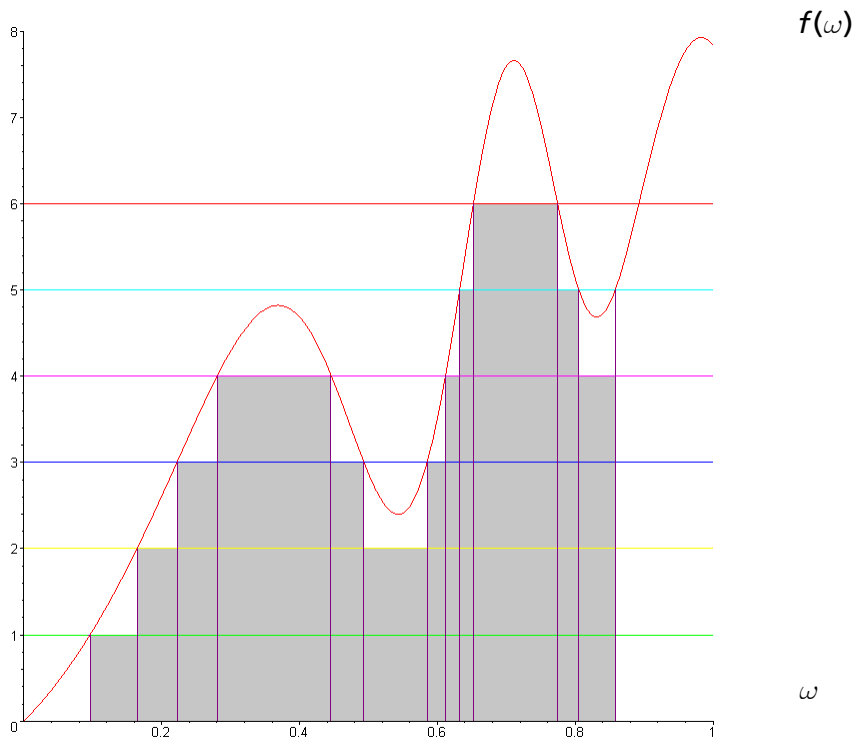
Definition 17. Es sei (Ω, \mathcal{A}) ein Messraum und f eine messbare, reellwertige Funktion auf Ω mit endlichem Wertebereich $f(\Omega)$. Dann heißt f *Elementarfunktion*.

Lemma 16. Es sei (Ω, \mathcal{A}) ein Messraum und f eine $(\mathcal{A}-\mathcal{B}^1)$ -messbare Elementarfunktion auf Ω . Dann besitzt f eine (im Allgemeinen nicht eindeutig bestimmte) Darstellung

$$f(\omega) = \sum_{i=1}^n a_i \mathbb{1}_{A_i}(\omega), \quad \omega \in \Omega$$

mit $n \in \mathbb{N}$, $a_1, \dots, a_n \in \mathbb{R}$, $A_1, \dots, A_n \in \mathcal{A}$. Umgekehrt ist jede Abbildung f mit einer solchen Darstellung eine Elementarfunktion.

Satz 16. Es sei (Ω, \mathcal{A}) ein Messraum und $\{f_n\}_{n \in \mathbb{N}}$ eine punktweise nach oben beschränkte Folge nicht-negativer $(\mathcal{A}-\mathcal{B}^1)$ -messbarer Funktionen auf Ω . Dann ist auch $f := \sup_{n \in \mathbb{N}} f_n$ $(\mathcal{A}-\mathcal{B}^1)$ -messbar. Umgekehrt ist jede nicht-negative reelle $(\mathcal{A}-\mathcal{B}^1)$ -messbare Funktion auf Ω als Supremum einer monoton wachsenden Folge nicht-negativer, punktweise nach oben beschränkter Elementarfunktionen darstellbar.



Lemma 17. Es sei (Ω, \mathcal{A}) ein Messraum und f eine reelle, $(\mathcal{A}-\mathcal{B}^1)$ -messbare Abbildung auf Ω . Dann sind der Positivteil $f^+ := \max\{0, f\}$ und der Negativteil $f^- := -\min\{0, f\} = (-f)^+$ $\mathcal{A}-\mathcal{B}^1$ -messbare Funktionen auf Ω , mit $f = f^+ - f^-$.

Satz 17. Es sei (Ω, \mathcal{A}) ein Messraum, f, g und $\{f_n\}_{n \in \mathbb{N}}$ seien reelle $(\mathcal{A}-\mathcal{B}^1)$ -messbare Abbildungen. Dann sind auch folgende Abbildungen messbar:

- a) $f + g, f - g, f \cdot g, \frac{f}{g}$ (falls $g \neq 0$), $\max(f, g), \min(f, g)$,
- b) $\sup_{n \in \mathbb{N}} f_n, \inf_{n \in \mathbb{N}} f_n$ (falls $\{f_n\}_{n \in \mathbb{N}}$ punktweise nach oben bzw. nach unten beschränkt),
- c) $\limsup_{n \rightarrow \infty} f_n, \liminf_{n \rightarrow \infty} f_n$, und $\lim_{n \rightarrow \infty} f_n$
(falls $\{f_n\}_{n \in \mathbb{N}}$ punktweise beschränkt und konvergent),
- d) $\sum_{n \in \mathbb{N}} f_n$ (falls absolut konvergent).

Definition 18. Es sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum und f eine Elementarfunktion mit einer Darstellung

$$f(\omega) = \sum_{i=1}^n a_i \mathbb{1}_{A_i}(\omega), \quad \omega \in \Omega$$

wie in Lemma 16. Dann ist das μ -Integral (Lebesgue-Integral) von f definiert durch

$$\int f d\mu := \sum_{i=1}^n a_i \mu(A_i).$$

Insbesondere gilt

$$\int \mathbb{1}_A d\mu = \mu(A) \text{ für alle } A \in \mathcal{A}.$$

Lemma 18. Unter den Voraussetzungen von Definition 18 gilt:

- a) Das μ -Integral ist wohldefiniert, d.h. es ist unabhängig von der speziellen Darstellung von f .
- b) Das μ -Integral ist linear und monoton auf der Menge der Elementarfunktionen, d.h. es gilt

$$\int \alpha f + \beta g d\mu = \alpha \int f d\mu + \beta \int g d\mu$$

für alle $\alpha, \beta \in \mathbb{R}$ und Elementarfunktionen f, g sowie

$$\int f d\mu \leq \int g d\mu \text{ für } f \leq g.$$

- c) Das μ -Integral ist auch linear und monoton auf der Menge der Maße auf \mathcal{A} , d.h. ist ν ein weiteres Maß auf \mathcal{A} , so gilt

$$\int f d(\alpha\mu + \beta\nu) = \alpha \int f d\mu + \beta \int f d\nu$$

für alle $\alpha, \beta \geq 0$ und Elementarfunktionen f sowie

$$\int f d\nu \leq \int f d\mu \text{ für } \nu \leq \mu \text{ und } f \geq 0.$$

Lemma 19. Es sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum und f eine nicht-negative reelle $(\mathcal{A}-\mathcal{B}^1)$ -messbare Abbildung. Dann gilt: $\sup_{n \in \mathbb{N}} \int g_n d\mu = \sup_{n \in \mathbb{N}} \int h_n d\mu$ für alle monoton wachsenden Folgen $\{g_n\}_{n \in \mathbb{N}}$, $\{h_n\}_{n \in \mathbb{N}}$ nicht-negativer Elementarfunktionen mit $f = \sup_{n \in \mathbb{N}} g_n = \sup_{n \in \mathbb{N}} h_n$.

Definition 19. Es sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum und f eine nicht-negative reelle, $(\mathcal{A}-\mathcal{B}^1)$ -messbare Abbildung. Dann ist das nach Lemma 19 von der approximierenden Folge unabhängige μ -Integral von f definiert durch

$$\int f d\mu := \sup_{n \in \mathbb{N}} \int g_n d\mu,$$

wobei $\{g_n\}_{n \in \mathbb{N}}$ eine beliebige monoton wachsende Folge nicht-negativer Elementarfunktionen mit $f = \sup_{n \in \mathbb{N}} g_n$ ist.

Definition 20. Es sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum und f eine beliebige reelle, $(\mathcal{A}-\mathcal{B}^1)$ -messbare Abbildung. Dann ist das μ -Integral von f definiert durch

$$\int f d\mu := \int f^+ d\mu - \int f^- d\mu,$$

sofern mindestens eines der beiden rechten Integrale *endlich* ist. In diesem Fall heißt f μ -*integrierbar*. Sind beide rechten Integrale endlich, so heißt f *eigentlich* μ -*integrierbar*. Ist ferner $A \in \mathcal{A}$, so bezeichnet

$$\int_A f d\mu := \int f \cdot \mathbb{1}_A d\mu.$$

Alternativ schreibt man auch: $\int f d\mu = \int f(\omega) \mu(d\omega)$; diese Schreibweise ist insbesondere bei Produktmaßen von Vorteil, weil damit eine eventuelle Integrationsreihenfolge eindeutig bestimmt ist.

Lemma 20. Es sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum, f und g seien reelle, $(\mathcal{A}-\mathcal{B}^1)$ -messbare, μ -integrierbare Abbildungen. Dann gilt:

a) $\int \alpha f + \beta g d\mu = \alpha \int f d\mu + \beta \int g d\mu$ für alle $\alpha, \beta \in \mathbb{R}$ (im Falle der Existenz)

b) $f \leq g \Rightarrow \int f d\mu \leq \int g d\mu$

c) $\left| \int f d\mu \right| \leq \int |f| d\mu$

d) $\forall A \in \mathcal{A}: \mu(A) = 0 \Rightarrow \int_A f d\mu = 0$

e) $f \geq 0$ und $\int f d\mu = 0 \Rightarrow f = 0$ μ -fast überall

f) $f = g$ μ -fast überall $\Rightarrow \int f d\mu = \int g d\mu$

g) $\forall A \in \mathcal{A}: \int_A f d\mu \leq \int_A g d\mu \Rightarrow f \leq g$ μ -fast überall

h) $\forall A \in \mathcal{A}: \int_A f d\mu = \int_A g d\mu \Rightarrow f = g$ μ -fast überall

i) $\forall \varepsilon > 0: \mu(\{\omega \in \Omega \mid |f(\omega)| > \varepsilon\}) \leq \frac{1}{\varepsilon} \int |f| d\mu$ (Markoff-Ungleichung).

Satz 18 (Satz von der monotonen Konvergenz). Es sei $\{f_n\}_{n \in \mathbb{N}}$ eine monoton wachsende Folge nicht-negativer reeller $(\mathcal{A}-\mathcal{B}^1)$ -messbarer Abbildungen auf einem Maßraum $(\Omega, \mathcal{A}, \mu)$. Dann gilt:

$$\int \lim_{n \rightarrow \infty} f_n d\mu = \int \sup_{n \in \mathbb{N}} f_n d\mu = \sup_{n \in \mathbb{N}} \int f_n d\mu = \lim_{n \rightarrow \infty} \int f_n d\mu.$$

Lemma 21. Es sei $\{f_n\}_{n \in \mathbb{N}}$ eine Folge nicht-negativer reeller $(\mathcal{A}-\mathcal{B}^1)$ -messbarer Abbildungen auf einem Maßraum $(\Omega, \mathcal{A}, \mu)$. Dann gilt:

$$\int \sum_{n=1}^{\infty} f_n d\mu = \sum_{n=1}^{\infty} \int f_n d\mu.$$

Lemma 22 (Lemma von Fatou). Es sei $\{f_n\}_{n \in \mathbb{N}}$ eine Folge nicht-negativer reeller $(\mathcal{A}-\mathcal{B}^1)$ -messbarer Abbildungen auf einem Maßraum $(\Omega, \mathcal{A}, \mu)$. Dann gilt:

$$\int \liminf_{n \rightarrow \infty} f_n d\mu \leq \liminf_{n \rightarrow \infty} \int f_n d\mu.$$

Satz 19 (Satz von der majorisierten Konvergenz). Es sei $\{f_n\}_{n \in \mathbb{N}}$ eine Folge reeller $(\mathcal{A}-\mathcal{B}^1)$ -messbarer Abbildungen auf einem Maßraum $(\Omega, \mathcal{A}, \mu)$, die punktweise gegen die (messbare) Abbildung f konvergiere. Ferner existiere eine nicht-negative $(\mathcal{A}-\mathcal{B}^1)$ -messbare und eigentlich μ -integrierbare Abbildung g auf $(\Omega, \mathcal{A}, \mu)$ mit $|f_n| \leq g$ für alle $n \in \mathbb{N}$. Dann gilt:

$$\lim_{n \rightarrow \infty} \int f_n d\mu = \int \lim_{n \rightarrow \infty} f_n d\mu \quad \text{und} \quad \lim_{n \rightarrow \infty} \int |f_n - f| d\mu = 0.$$

Definition 21 (Maße mit Dichten). Es sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum und $f: (\Omega, \mathcal{A}) \rightarrow (\mathbb{R}^1, \mathcal{B}^1)$ eine nicht-negative, messbare Abbildung. Dann wird durch

$$\nu(A) := \int_A f d\mu \quad \text{für alle } A \in \mathcal{A}$$

ein Maß auf \mathcal{A} definiert. Die Abbildung f heißt hier (eine) μ -Dichte von ν .

Satz 20. Es sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum und $g: (\Omega, \mathcal{A}) \rightarrow (\mathbb{R}^1, \mathcal{B}^1)$ eine nicht-negative, messbare Abbildung und ν das Maß mit Dichte g gemäß Definition 21. Dann gilt:

$$\int f d\nu = \int fg d\mu$$

für alle nicht-negativen, messbaren Abbildungen $f: (\Omega, \mathcal{A}) \rightarrow (\mathbb{R}^1, \mathcal{B}^1)$. Die Gleichung bleibt auch dann richtig, wenn f beliebig reellwertig und ν -integrierbar bzw. fg μ -integrierbar ist.

Satz 21 (Satz von Radon und Nikodym, 1930). Es sei (Ω, \mathcal{A}) ein Messraum; ferner seien μ und ν Maße auf \mathcal{A} mit der Eigenschaft: μ ist σ -endlich, und es gilt

$$\forall A \in \mathcal{A}: \mu(A) = 0 \Rightarrow \nu(A) = 0$$

(man sagt auch, μ *dominiert* ν). Dann besitzt ν eine Dichte bezüglich μ .

Satz 22. Es sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum und (Ξ, \mathcal{C}) ein Messraum sowie $T: (\Omega, \mathcal{A}) \rightarrow (\Xi, \mathcal{C})$ eine messbare Abbildung. Dann gilt:

$$\int g \circ T d\mu = \int g d\mu^T$$

für alle nicht-negativen, messbaren Abbildungen $g: (\Xi, \mathcal{C}) \rightarrow (\mathbb{R}^1, \mathcal{B}^1)$. Die Gleichung bleibt auch dann richtig, wenn g beliebig reellwertig und $g \circ T$ μ -integrierbar bzw. g μ^T -integrierbar ist.

Satz 23 (Satz von Fubini, 1907). Es seien $(\Omega_i, \mathcal{A}_i, \mu_i)$ Maßräume für $i = 1, 2$ mit σ -endlichen Maßen μ_i . Ferner sei $f: (\Omega_1 \times \Omega_2, \mathcal{A}_1 \otimes \mathcal{A}_2) \rightarrow (\mathbb{R}^1, \mathcal{B}^1)$ eine nicht-negative, (Produkt-)messbare Abbildung. Dann gilt: die Abbildung $f(\cdot, \omega_2)$ ist für jedes feste $\omega_2 \in \Omega_2$ \mathcal{A}_1 -messbar und die Abbildung $f(\omega_1, \cdot)$ ist für jedes feste $\omega_1 \in \Omega_1$ \mathcal{A}_2 -messbar, und es gilt

$$\int f \, d\mu_1 \otimes \mu_2 = \int \left\{ \int f(\omega_1, \omega_2) \mu_1(d\omega_1) \right\} \mu_2(d\omega_2) = \int \left\{ \int f(\omega_1, \omega_2) \mu_2(d\omega_2) \right\} \mu_1(d\omega_1).$$

Diese Gleichheit bleibt auch gültig, wenn f beliebig reell und eigentlich $\mu_1 \otimes \mu_2$ -integrierbar ist. In diesem Fall existieren die inneren Integrale $\int f(\omega_1, \omega_2) \mu_1(d\omega_1)$ und $\int f(\omega_1, \omega_2) \mu_2(d\omega_2)$ fast überall (als Funktionen von ω_2 bzw. von ω_1).

Satz 24. Es sei $I \in \mathcal{B}^d$ ein nicht-leeres kompaktes Standard-Intervall. Eine beschränkte Funktion $f: I \rightarrow \mathbb{R}$ ist genau dann Riemann-integrierbar, wenn die Menge ihrer Unstetigkeitsstellen eine m^d -Nullmenge ist. In diesem Fall stimmt das Riemann-Integral von f mit dem Lebesgue-Integral überein. Der Satz bleibt richtig, wenn f allgemeiner auf einer beschränkten Jordan-messbaren Menge $M \subseteq \mathbb{R}^d$ definiert ist.

Ist ferner $I \in \mathcal{B}^1$ ein beliebiges Intervall und $f: I \rightarrow \mathbb{R}$ Riemann-integrierbar über jedem kompakten Teilintervall von I , so ist f genau dann (eigentlich) Lebesgue-integrierbar über I , wenn $|f|$ uneigentlich Riemann-integrierbar ist über I . In diesem Fall stimmen wieder beide Integrale überein.

Satz 25 (Transformationsformel). Es sei $d \in \mathbb{N}$ und $U \subseteq \mathbb{R}^d$ eine offene Menge. Ferner sei die Vektorfunktion $g = (g_1, \dots, g_d): U \rightarrow \mathbb{R}^d$ stetig differenzierbar, $V := g(U)$ das Bild von U unter g , $C := \{\mathbf{x} = (x_1, \dots, x_d) \in U \mid \det(\Delta g(\mathbf{x})) = 0\}$ die Menge der kritischen Punkte von g , wobei

$$\Delta g(\mathbf{x}) := \begin{bmatrix} \frac{\partial}{\partial x_1} g_1(\mathbf{x}) & \cdots & \frac{\partial}{\partial x_d} g_1(\mathbf{x}) \\ \vdots & & \vdots \\ \frac{\partial}{\partial x_1} g_d(\mathbf{x}) & \cdots & \frac{\partial}{\partial x_d} g_d(\mathbf{x}) \end{bmatrix}$$

die Funktionalmatrix von g im Punkt $\mathbf{x} \in U$ bezeichne, und g auf der Menge $U \setminus C$ injektiv. Eine Abbildung $f: V \rightarrow \mathbb{R}$ ist genau dann Lebesgue-integrierbar über V , wenn $f \circ g \cdot |\det(\Delta g)|$ Lebesgue-integrierbar über U ist. In diesem Fall gilt:

$$\int_V f \, d\mathbf{m}^d = \int_U f \circ g \cdot |\det(\Delta g)| \, d\mathbf{m}^d.$$

Satz 26 (Transformationsatz für Dichten). Ist unter den Voraussetzungen von Satz 25 die Abbildung f nicht-negativ, die Abbildung g injektiv und die Menge der kritischen Punkte leer, so ist die auf V definierte Abbildung h mit

$$h(\mathbf{y}) := \frac{f(g^{-1}(\mathbf{y}))}{\left| \det(\Delta g(g^{-1}(\mathbf{y}))) \right|}, \quad \mathbf{y} \in V$$

nicht-negativ und Borel-messbar, und es gilt:

$$\int_A h(\mathbf{y}) m^d(d\mathbf{y}) = \int_{g^{-1}(A)} f(\mathbf{x}) m^d(d\mathbf{x}) \quad \text{für alle } A \in V \cap \mathcal{B}^d.$$

II. Grundprinzipien stochastischer Modellierung

II.1. Elemente der Kombinatorik

Definition 22 (Permutationen und Kombinationen). Es sei $\Omega = \{\omega_1, \dots, \omega_n\}$ eine nicht-leere endliche Menge mit $n \in \mathbb{N}$. Für $1 \leq k \leq n$ definieren wir:

- a) Jedes Element $(\eta_1, \dots, \eta_k) \in \Omega^k = \prod_{i=1}^k \Omega$ heißt (k, n) -Permutation aus Ω (mit Wiederholung). Die Menge aller solchen Permutationen wird mit $\text{Perm}_k^n(\Omega; m.W.)$ bezeichnet.
- b) Jede (k, n) -Permutation (η_1, \dots, η_k) mit paarweise verschiedenen Komponenten heißt (k, n) -Permutation aus Ω (ohne Wiederholung). Die Menge dieser Permutationen wird mit $\text{Perm}_k^n(\Omega; o.W.)$ bezeichnet.
- c) Jede (k, n) -Permutation $(\omega_{i_1}, \dots, \omega_{i_k})$ mit $1 \leq i_1 \leq i_2 \leq \dots \leq i_k \leq n$ heißt (k, n) -Kombination aus Ω (mit Wiederholung). Die Menge dieser Kombinationen wird mit $\text{Komb}_k^n(\Omega; m.W.)$ bezeichnet.
- d) Jede (k, n) -Kombination $(\omega_{i_1}, \dots, \omega_{i_k})$ mit paarweise verschiedenen Komponenten heißt (k, n) -Kombination aus Ω (ohne Wiederholung). Die Menge dieser Kombinationen wird mit $\text{Komb}_k^n(\Omega; o.W.)$ bezeichnet.

Lemma 23. Mit den Bezeichnungen aus Definition 22 gilt:

$$\#(\text{Perm}_k^n(\Omega; m.W.)) = n^k$$

$$\#(\text{Perm}_k^n(\Omega; o.W.)) = (n)_k = \binom{n}{k} k! = n(n-1)\cdots(n-k+1)$$

$$\#(\text{Komb}_k^n(\Omega; m.W.)) = \binom{n+k-1}{k} = \binom{n+k-1}{n-1}$$

$$\#(\text{Komb}_k^n(\Omega; o.W.)) = \binom{n}{k}.$$

Problem des französischen Adligen George Brossin Antoine Gombaud, Chevalier de Méré (1607 - 1684):

Warum ist es unvorteilhafter, zum Erreichen einer Doppelsechs mit zwei Würfeln 24 Würfe zu tun, als zum Erreichen einer Sechs mit einem Würfel 4 Würfe zu tun, obwohl das Verhältnis 24 zu 36 (was die Anzahl der Ergebnisse bei 2 Würfeln ist) dasselbe ist wie 4 zu 6 (was die Anzahl der Ergebnisse bei einem Würfel ist)?

Lösung:

Grundmodell: $\Omega = \{1,2,3,4,5,6\}$ für den einfachen Würfelwurf, Ω^m für den m -fachen Würfelwurf. Komplementäres Ereignis:

1. Fall:
$$A^c = \bigtimes_{i=1}^{24} \{(6,6)\}^c \subseteq \Omega^{48} \text{ mit } P(A^c) = \frac{\#(A^c)}{\#(\Omega^{48})} = \frac{35^{24}}{6^{48}} = 0,50859\dots,$$

2. Fall:
$$A^c = \bigtimes_{i=1}^4 \{6\}^c \subseteq \Omega^4 \text{ mit } P(A^c) = \frac{\#(A^c)}{\#(\Omega^4)} = \frac{5^4}{6^4} = 0,48225\dots$$

„Geburtstagsparadox“: In einer Runde von k Personen soll darauf gewettet werden, dass wenigstens zwei Anwesende am gleichen Tag Geburtstag haben. Ab welcher Personenzahl würden Sie diese Wette stellen? Die überraschende Antwort lautet: bereits ab 23 Personen ist die Wahrscheinlichkeit für dieses Ereignis $A_{k,n}$ größer als $1/2$!

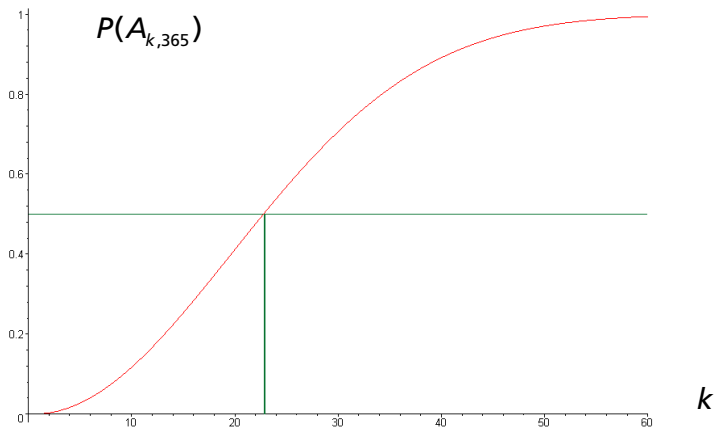
Lösung: $\Omega = \{1, 2, \dots, n\}$, hier mit $n = 365$ für die möglichen Tage im Jahr. Die Menge $Perm_k^n(\Omega; m.W.)$ beschreibt die Menge der Möglichkeiten, wie sich die Geburtstage von k Personen auf das Jahr verteilen können. Wir wählen die Laplace-Verteilung P als Wahrscheinlichkeitsverteilung. Für das komplementäre Ereignis $A_{k,n}^c$, dass alle k Personen an *verschiedenen* Tagen Geburtstag haben, gilt $A_{k,n}^c = Perm_k^n(\Omega; o.W.)$ mit

$$\begin{aligned}
 P(A_{k,n}) &= 1 - P(A_{k,n}^c) = 1 - \frac{\#(Perm_k^n(\Omega; o.W.))}{\#(Perm_k^n(\Omega; m.W.))} = 1 - \frac{(n)_k}{n^k} = 1 - \prod_{i=0}^{k-1} \left(1 - \frac{i}{n}\right) \\
 &= 1 - \exp\left(\sum_{i=0}^{k-1} \ln\left(1 - \frac{i}{n}\right)\right) \geq 1 - \exp\left(-\sum_{i=0}^{k-1} \frac{i}{n}\right) = 1 - \exp\left(-\frac{k(k-1)}{2n}\right) \geq 1 - \exp\left(-\frac{(k-1)^2}{2n}\right).
 \end{aligned}$$

Auflösung dieser der Ungleichung nach k ergibt:

$$k \geq 1 + \sqrt{n \ln 4} \geq 1 + \sqrt{1,386n}.$$

Für $n = 365$ liefert diese approximative Lösung $k \geq 23,49\dots$, die exakte Lösung ist $k \geq 23$ mit $P(A_{23,365}) = 0,50729\dots$. Die folgende Graphik zeigt die Wahrscheinlichkeiten $P(A_{k,365})$ in Abhängigkeit von k im Bereich $1 \leq k \leq 60$.



II. Grundprinzipien stochastischer Modellierung

II.2. Zufallsvariablen und ihre Verteilung, stochastische Unabhängigkeit

Definition 23 (Zufallsvariable / Zufallsvektor). Es sei (Ω, \mathcal{A}, P) ein Wahrscheinlichkeitsraum und $(\mathfrak{X}, \mathcal{B})$ ein Messraum. $X: (\Omega, \mathcal{A}) \rightarrow (\mathfrak{X}, \mathcal{B})$ sei eine messbare Abbildung. Dann heißt X

Zufallsvariable, wenn $\mathfrak{X} = \mathbb{R}$ und \mathcal{B} die zugehörige Borel'sche σ -Algebra ist;

Zufallsvektor, wenn $\mathfrak{X} = \mathbb{R}^d$ für ein $d \in \mathbb{N}$, $d \geq 2$ und \mathcal{B} die zugehörige Borel'sche σ -Algebra ist;

Zufallselement, wenn \mathfrak{X} ein allgemeinerer topologischer Raum und \mathcal{B} die zugehörige Borel'sche (d.h. durch das System der offenen Mengen erzeugte) σ -Algebra ist.

Das durch die Abbildung X induzierte Maß $Q = P^X$ auf \mathcal{B} heißt (*Wahrscheinlichkeits-*)*Verteilung von X* .

Definition 24 (Stochastischer Prozess). Es sei (Ω, \mathcal{A}, P) ein Wahrscheinlichkeitsraum und \mathfrak{X} ein Banach-Raum reeller Funktionen; d.h. \mathfrak{X} ist ein vollständiger, linearer, normierter topologischer Raum. \mathcal{B} bezeichne die von der Topologie des Raumes erzeugte Borel'sche σ -Algebra, $X: (\Omega, \mathcal{A}) \rightarrow (\mathfrak{X}, \mathcal{B})$ sei eine messbare Abbildung. Dann heißt X stochastischer Prozess.

Definition 25 (stochastische Unabhängigkeit). Es sei (Ω, \mathcal{A}, P) ein Wahrscheinlichkeitsraum und $(\mathfrak{X}, \mathcal{B})$ ein Messraum. $X_i: (\Omega, \mathcal{A}) \rightarrow (\mathfrak{X}, \mathcal{B})$ seien messbare Abbildungen mit $i \in I$; die Indexmenge I darf dabei beliebig nicht-leer gewählt sein. Die Familie $\{X_i \mid i \in I\}$ heißt *stochastisch unabhängig*, wenn die auf dem Produktraum $\left(\prod_{i \in I} \mathfrak{X}, \bigotimes_{i \in I} \mathcal{B}\right)$ induzierte Verteilung der Familie das Produktmaß $Q = \bigotimes_{i \in I} P^{X_i}$ ist.

Lemma 24: In der Situation von Definition 25 ist die Familie $\{X_i \mid i \in I\}$ genau dann stochastisch unabhängig, wenn jede endliche Teilfamilie $\{X_j \mid j \in J_m\}$ mit $J_m = \{j_1, \dots, j_m\} \subseteq I$, $m \in \mathbb{N}$ stochastisch unabhängig ist, d.h. wenn gilt

$$P^{(X_{j_1}, \dots, X_{j_m})} = \bigotimes_{k=1}^m P^{X_{j_k}}$$

Lemma 25. In der Situation von Definition 25 ist die Familie $\{X_i \mid i \in I\}$ genau dann stochastisch unabhängig, wenn für jede endliche Teilindexmenge $J_m = \{j_1, \dots, j_m\} \subseteq I$, $m \in \mathbb{N}$ gilt:

$$P(X_{j_1} \in B_1, \dots, X_{j_m} \in B_m) := P\left(\bigcap_{k=1}^m \{X_{j_k} \in B_k\}\right) = \prod_{k=1}^m P(\{X_{j_k} \in B_k\})$$

für alle $B_k \in \mathcal{B}$, $1 \leq k \leq m$.

Hierbei kann die Bedingung $B_k \in \mathcal{B}$ noch ersetzt werden durch $B_k \in \mathcal{S}$, wobei \mathcal{S} einen Semi-Ring bezeichne, der \mathcal{B} erzeugt.

Definition 26 (stochastische Unabhängigkeit von Ereignissen). Es sei $(\Omega, \mathcal{A}, \mathcal{P})$ ein Wahrscheinlichkeitsraum und $\{A_i\}_{i \in I}$ eine beliebige Familie messbarer Mengen. Die Familie heißt *stochastisch unabhängig*, wenn die Familie $\{\mathbb{1}_{A_i}\}_{i \in I}$ der zugehörigen Indikatorfunktionen stochastisch unabhängig ist. Eine Familie $\{\mathcal{F}_i\}_{i \in I}$ mit $\mathcal{F}_i \subseteq \mathcal{A}$ für alle $i \in I$ heißt entsprechend stochastisch unabhängig, wenn jede Familie von Auswahlmengen $\{F_i\}_{i \in I}$ mit $F_i \in \mathcal{F}_i$ für alle $i \in I$ stochastisch unabhängig ist.

Lemma 26. In der Situation von Definition 26 ist die Familie $\{A_i\}_{i \in I}$ genau dann stochastisch unabhängig, wenn für jede endliche Teilindexmenge $J_m = \{j_1, \dots, j_m\} \subseteq I$, $m \in \mathbb{N}$ gilt:

$$P\left(\bigcap_{k=1}^m A_{j_k}\right) = \prod_{k=1}^m P(A_{j_k}) \text{ für alle } 1 \leq k \leq m.$$

Ferner gilt: eine endliche Familie $\{A_i\}_{i \in I}$ von Ereignissen mit $I = \{1, \dots, m\}$, $m \in \mathbb{N}$ ist genau dann stochastisch unabhängig, wenn gilt:

$$P\left(\bigcap_{k=1}^m C_k\right) = \prod_{k=1}^m P(C_k) \text{ für alle Wahlen } C_k \in \{A_k, A_k^c\}, \quad 1 \leq k \leq m.$$

In diesem Fall gilt auch noch

$$P\left(\bigcup_{i=1}^m A_i\right) = 1 - \prod_{i=1}^m (1 - P(A_i)).$$

Satz 27. Es sei (Ω, \mathcal{A}, P) ein Wahrscheinlichkeitsraum mit abzählbarer Grundmenge Ω und $p \in (0,1)$ beliebig, aber fest. Dann existiert keine abzählbare Familie stochastisch unabhängiger Ereignisse $\{A_n \mid n \in \mathbb{N}\} \subseteq \mathcal{A}$ mit $P(A_n) = p$ für alle $n \in \mathbb{N}$.

Definition 27 (elementare bedingte Wahrscheinlichkeit). Es sei (Ω, \mathcal{A}, P) ein beliebiger Wahrscheinlichkeitsraum und $B \in \mathcal{A}$ ein Ereignis mit $P(B) > 0$. Dann heißt das normierte Spurmaß P_B bezüglich B die *durch B induzierte elementare bedingte Verteilung*, in expliziter Form:

$$P_B(A) := \frac{P(A \cap B)}{P(B)} \text{ für alle } A \in \mathcal{A}.$$

Lemma 27. Die stochastische Unabhängigkeit zweier Ereignisse $A, B \in \mathcal{A}$ kann unter den Voraussetzungen von Definition 27 äquivalent charakterisiert werden durch die Gültigkeit wahlweise einer der Beziehungen

$$P(A | B) = P(A) \text{ bzw. } P(B | A) = P(B).$$

Lemma 28 (Unabhängigkeit bei Blockbildung). Es sei (Ω, \mathcal{A}, P) ein Wahrscheinlichkeitsraum und $(\mathcal{X}, \mathcal{B})$ ein Messraum. $X_i : (\Omega, \mathcal{A}) \rightarrow (\mathcal{X}, \mathcal{B})$ seien stochastisch unabhängig (Zufallsvariablen, Zufallsvektoren oder Zufallselemente) mit $i \in I$; die Indexmenge I darf dabei beliebig nicht-leer gewählt sein. Sind dann g_j geeignet messbare Abbildungen auf dem Produktraum $\left(\prod_{i \in I_j} \mathcal{X}, \otimes_{i \in I_j} \mathcal{B} \right)$ mit Werten in einem Messraum $(\mathcal{X}_j, \mathcal{B}_j)$, wobei die I_j , $j \in J$ nichtleere, paarweise disjunkte Teilmengen von I seien, dann ist auch die Familie $\left\{ g_j \left(\{ X_i \}_{i \in I_j} \right) \right\}_{j \in J}$ stochastisch unabhängig.

Satz 28 (von der totalen Wahrscheinlichkeit bzw. von Bayes): Es sei (Ω, \mathcal{A}, P) ein Wahrscheinlichkeitsraum und $\{B_i\}_{i \in I} \subseteq \mathcal{A}$ eine disjunkte Zerlegung von Ω , also $\Omega = \bigoplus_{i \in I} B_i$ mit einer höchstens abzählbaren Indexmenge I . Ferner gelte $P(B_i) > 0$ für alle $i \in I$. Dann gilt:

a) $P(A) = \sum_{i \in I} P(A|B_i) \cdot P(B_i)$ für alle $A \in \mathcal{A}$

b) $P(B_j|A) = \frac{P(A|B_j) \cdot P(B_j)}{\sum_{i \in I} P(A|B_i) \cdot P(B_i)}$ für alle $A \in \mathcal{A}$ mit $P(A) > 0$ und alle $j \in I$.

II. Grundprinzipien stochastischer Modellierung

II.3. Verteilungsfunktionen und Momente

Definition 28 (Verteilungsfunktion). Es sei P ein Wahrscheinlichkeitsmaß auf der Borel'schen σ -Algebra \mathcal{B}^d , $d \in \mathbb{N}$. Die durch

$$F_P(x_1, \dots, x_d) := P\left(\prod_{i=1}^d (-\infty, x_i]\right), \quad (x_1, \dots, x_d) \in \mathbb{R}^d$$

definierte Funktion heißt die zu P gehörige (d -dimensionale) *Verteilungsfunktion*.

Definition 29 (Δ -Monotonie). Eine Funktion $g: \mathbb{R}^d \rightarrow \mathbb{R}$ mit $d \in \mathbb{N}$ heißt *Δ -monoton*, wenn sie die folgende Eigenschaft besitzt:

$$\Delta g_{\mathbf{x}}^{\mathbf{y}} := \sum_{(\varepsilon_1, \dots, \varepsilon_d) \in \{0,1\}^d} (-1)^{\sum_{i=1}^d \varepsilon_i} g(\varepsilon_1 x_1 + (1-\varepsilon_1)y_1, \dots, \varepsilon_d x_d + (1-\varepsilon_d)y_d) \geq 0$$

für alle $\mathbf{x} = (x_1, \dots, x_d)$, $\mathbf{y} = (y_1, \dots, y_d) \in \mathbb{R}^d$ mit $\mathbf{x} \leq \mathbf{y}$, d.h. $x_i \leq y_i$ für $i = 1, \dots, d$.

Lemma 29 (Eigenschaften von Verteilungsfunktionen). Es sei $F = F_P$ die Verteilungsfunktion einer Wahrscheinlichkeitsverteilung P auf der Borel'schen σ -Algebra \mathcal{B}^d , $d \in \mathbb{N}$. Dann gilt:

$\Delta F_{\mathbf{x}}^{\mathbf{y}} \geq 0$ für alle $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$ mit $\mathbf{x} \leq \mathbf{y}$ (Δ -Monotonie von F)

$P\left(\prod_{i=1}^d (x_i, y_i]\right) = \Delta F_{\mathbf{x}}^{\mathbf{y}}$ für alle $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$ mit $\mathbf{x} \leq \mathbf{y}$

$\lim_{y_1 \downarrow x_1, \dots, y_d \downarrow x_d} F(y_1, \dots, y_d) = F(x_1, \dots, x_d)$ für alle $\mathbf{x} \in \mathbb{R}^d$

(rechtsseitige Stetigkeit)

$\lim_{x_1 \uparrow \infty, \dots, x_d \uparrow \infty} F(x_1, \dots, x_d) = 1$, $\lim_{x_i \downarrow -\infty} F(x_1, \dots, x_d) = 0$ für $i = 1, \dots, d$

(Normiertheit / rechtsseitige Stetigkeit).

Satz 29 (Charakterisierungssatz). Es sei F eine Δ -monotone Funktion auf \mathbb{R}^d für $d \in \mathbb{N}$, die die drei Grenzwerteigenschaften von Lemma 29 erfülle. Dann gibt es genau ein Wahrscheinlichkeitsmaß P auf der Borel'schen σ -Algebra \mathcal{B}^d , so dass $F = F_P$ ist.

Ein wichtiger Spezialfall von Verteilungsfunktionen liegt vor, wenn das zugehörige Wahrscheinlichkeitsmaß P eine Dichte f bezüglich des Lebesgue-Maßes m^d besitzt. In diesem Fall gilt nämlich

$$\Delta F_{\mathbf{x}}^{\mathbf{y}} = P\left(\prod_{i=1}^d (x_i, y_i]\right) = \int_{x_d}^{y_d} \cdots \int_{x_1}^{y_1} f(u_1, \dots, u_d) du_1 \cdots du_d.$$

Ist die Dichte f darüber hinaus in dem Punkt $\mathbf{x} = (x_1, \dots, x_d) \in \mathbb{R}^d$ stetig, folgt nach dem (mehrdimensionalen) Hauptsatz der Differential- und Integralrechnung sogar noch

$$f(x_1, \dots, x_d) = \frac{\partial^d}{\partial x_1 \cdots \partial x_d} F(x_1, \dots, x_d).$$

Satz 30 (Verteilungsfunktionen bei Unabhängigkeit). Es seien X_1, \dots, X_d reellwertige Zufallsvariablen auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, \mathcal{P})$ mit den Verteilungsfunktionen F_1, \dots, F_d (d.h. zu den Verteilungen $\mathcal{P}^{X_1}, \dots, \mathcal{P}^{X_d}$). Dann gilt: X_1, \dots, X_d sind stochastisch unabhängig genau dann, wenn für die Verteilungsfunktion F der gemeinsamen Verteilung $\mathcal{P}^{(X_1, \dots, X_d)}$ gilt:

$$F(x_1, \dots, x_d) = \prod_{i=1}^d F_i(x_i), \quad \mathbf{x} = (x_1, \dots, x_d) \in \mathbb{R}^d.$$

Definition 30 (Randverteilungen). Es sei $\left(\prod_{i \in I} \Omega_i, \bigotimes_{i \in I} \mathcal{A}_i\right)$ ein Produkt-Messraum mit einer abzählbaren Indexmenge $I \subseteq \mathbb{N}$ und P ein Wahrscheinlichkeitsmaß auf $\bigotimes_{i \in I} \mathcal{A}_i$. Das für jede Teilindexmenge $J \subseteq I$ auf dem Produkt-Messraum $\left(\prod_{i \in J} \Omega_i, \bigotimes_{i \in J} \mathcal{A}_i\right)$ durch

$$P_J \left(\prod_{j \in J} A_j \right) := P \left(\prod_{i \in I} B_i \right) \text{ mit } B_i := \begin{cases} A_i, & i \in J \\ \Omega_i, & i \in I \setminus J \end{cases}$$

eindeutig bestimmte Wahrscheinlichkeitsmaß P_J heißt Randverteilung der Ordnung J zu P .

Lemma 30. Es sei F die zu einem Wahrscheinlichkeitsmaß P auf \mathcal{B}^d gehörige Verteilungsfunktion auf \mathbb{R}^d für $d \geq 2$ und $J = \{i_1, \dots, i_r\} \subseteq I := \{1, \dots, d\}$ mit einer natürlichen Zahl $r < d$. Ferner sei $K := I \setminus J = \{j_1, \dots, j_{d-r}\}$ die der Größe nach angeordnete Menge der komplementären Indices. Dann ist die zu dem Wahrscheinlichkeitsmaß P_J gehörige Verteilungsfunktion F_J gegeben durch

$$F_J(x_{i_1}, \dots, x_{i_r}) = \lim_{x_{j_1} \rightarrow \infty} \dots \lim_{x_{j_{d-r}} \rightarrow \infty} F(x_1, \dots, x_d), \quad (x_{i_1}, \dots, x_{i_r}) \in \mathbb{R}^r.$$

Besitzt F darüber hinaus eine Dichte f bezüglich des Lebesgue-Maßes m^d , so besitzt F_J ebenfalls eine Dichte (bezüglich des Lebesgue-Maßes m^r), die gegeben ist durch

$$f_J(x_{i_1}, \dots, x_{i_r}) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f(x_1, \dots, x_d) dx_{j_1} \dots dx_{j_{d-r}}, \quad (x_{i_1}, \dots, x_{i_r}) \in \mathbb{R}^r.$$

Satz 31 (Sklar 1958). Es sei F eine beliebige Verteilungsfunktion auf \mathbb{R}^d mit $d \geq 2$. Dann existiert eine Copula C so, dass

$$F(x_1, \dots, x_d) = C(F_1(x_1), \dots, F_d(x_d)) \text{ für alle } \mathbf{x} \in \mathbb{R}^d$$

gilt, wobei F_1, \dots, F_d die zu F gehörigen Randverteilungsfunktionen bezeichnen. Die Copula ist dabei eindeutig bestimmt, wenn die Randverteilungsfunktionen F_1, \dots, F_d *stetig* sind.

Achensee-Paradox: Jemand notiert zwei beliebige *verschiedene* reelle Zahlen auf zwei Zetteln (je eine pro Zettel), die verdeckt auf den Tisch gelegt werden. Eine zweite Person deckt nach Wahl zufällig einen der Zettel auf. Dann gibt es eine Strategie, mit der mit *mehr* als 50%iger Wahrscheinlichkeit entschieden werden kann, ob die noch verdeckte Zahl größer oder kleiner als die aufgedeckte Zahl ist.

Die versprochene Strategie besteht darin, dass man unabhängig von der Größe der aufgedeckten Zahl ein Zufallsexperiment durchführt, in dem man eine Zufallszahl Y erzeugt, die aus einer Wahrscheinlichkeitsverteilung Q stammt, deren zugehörige Verteilungsfunktion F lediglich *stetig* und *streng monoton wachsend* auf \mathbb{R} sein muss.

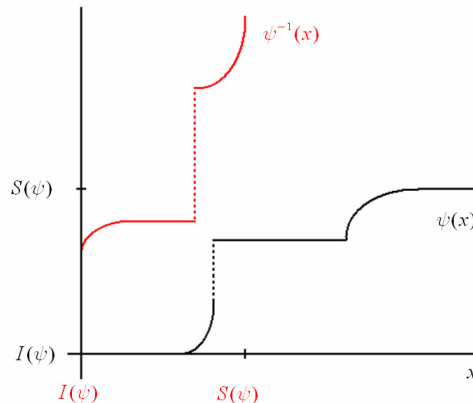
Achensee-Paradox: Jemand notiert zwei beliebige *verschiedene* reelle Zahlen auf zwei Zetteln (je eine pro Zettel), die verdeckt auf den Tisch gelegt werden. Eine zweite Person deckt nach Wahl zufällig einen der Zettel auf. Dann gibt es eine Strategie, mit der mit *mehr* als 50%iger Wahrscheinlichkeit entschieden werden kann, ob die noch verdeckte Zahl größer oder kleiner als die aufgedeckte Zahl ist.

Die Entscheidungsregel zu der Strategie lautet: Ist die Zufallszahl Y *größer* als die aufgedeckte Zahl, so entscheiden wir, dass die noch verdeckte Zahl ebenfalls größer als die aufgedeckte Zahl ist. Ist die Zufallszahl Y *kleiner* als die aufgedeckte Zahl, so entscheiden wir, dass die noch verdeckte Zahl ebenfalls kleiner als die aufgedeckte Zahl ist.

Mit anderen Worten: wir besorgen uns ersatzweise die zweite Zahl aus einem eigenen unabhängigen Zufallsexperiment und entscheiden dann so, als wenn die „gezogene“ Zahl tatsächlich die noch verdeckte wäre.

Definition 31 (Pseudo-Inverse). Es sei $\psi: \mathbb{R} \rightarrow \mathbb{R}$ eine schwach monoton wachsende, rechtsseitig stetige Funktion. Ferner sei $I(\psi) := \inf\{\psi(x) \mid x \in \mathbb{R}\}$, $S(\psi) := \sup\{\psi(x) \mid x \in \mathbb{R}\}$. Dann ist auf dem offenen Intervall $(I(\psi), S(\psi))$ die *Pseudo-Inverse* ψ^{-1} von ψ definiert durch

$$\psi^{-1}(y) := \inf\{x \in \mathbb{R} \mid \psi(x) \geq y\}, \quad y \in (I(\psi), S(\psi)).$$



Lemma 31 (Eigenschaften der Pseudo-Inversen). Es seien ψ und ψ^{-1} wie in Definition 31. Dann gilt:

- a) ψ^{-1} ist auf $(I(\psi), S(\psi))$ schwach monoton wachsend und linksseitig stetig.
- b) $\psi(\psi^{-1}(y)) \geq y$ für alle $y \in (I(\psi), S(\psi))$.
- c) Ist ψ in $\psi^{-1}(y)$ stetig für ein $y \in (I(\psi), S(\psi))$, so gilt $\psi(\psi^{-1}(y)) = y$.
- d) $\psi^{-1}(\psi(x)) \leq x$ für alle $x \in \mathbb{R}$ mit $\psi(x) \in (I(\psi), S(\psi))$.
- e) Ist ψ^{-1} in $\psi(x)$ stetig für ein $x \in \mathbb{R}$ mit $\psi(x) \in (I(\psi), S(\psi))$, so gilt $\psi^{-1}(\psi(x)) = x$.

Lemma 32. Es sei X eine Zufallsvariable auf einem Wahrscheinlichkeitsraum (Ω, \mathcal{A}, P) mit stetiger Verteilungsfunktion F . Dann ist die Zufallsvariable $Y := F(X)$ über dem Einheitsintervall $[0, 1]$ stetig gleichverteilt. Ist umgekehrt F eine beliebige Verteilungsfunktion und Y eine über dem Einheitsintervall $[0, 1]$ stetig gleichverteilte Zufallsvariable, so besitzt die Zufallsvariable $X := F^{-1}(Y)$ die Verteilungsfunktion F .

Definition 32 (Erwartungswert). Es sei X eine reellwertige Zufallsvariable auf einem Wahrscheinlichkeitsraum (Ω, \mathcal{A}, P) . Im Falle der Existenz heißt

$$E(X) := \int X dP$$

der *Erwartungswert* von X .

Lemma 33 (Eigenschaften des Erwartungswerts). Es sei (Ω, \mathcal{A}, P) ein Wahrscheinlichkeitsraum, X, Y und $\{X_n\}_{n \in \mathbb{N}}$ seien reellwertige Zufallsvariablen. Dann gilt:

- a) Ist $X = \mathbb{1}_A$ für ein $A \in \mathcal{A}$, so ist $E(X) = P(A)$.
- b) $E(\alpha X + \beta Y) = \alpha E(X) + \beta E(Y)$ für alle $\alpha, \beta \in \mathbb{R}$ (im Falle der Existenz)
- c) $X \leq Y \Rightarrow E(X) \leq E(Y)$
- d) $|E(X)| \leq E(|X|)$
- e) $\forall A \in \mathcal{A}: P(A) = 0 \Rightarrow \int_A X dP = E(\mathbb{1}_A \cdot X) = 0$
- f) $X \geq 0$ und $E(X) = 0 \Rightarrow X = 0$ P -fast sicher
- g) $X = Y$ P -fast sicher $\Rightarrow E(X) = E(Y)$
- h) $\forall A \in \mathcal{A}: E(\mathbb{1}_A \cdot X) \leq E(\mathbb{1}_A \cdot Y) \Rightarrow X \leq Y$ P -fast sicher
- i) $\forall A \in \mathcal{A}: E(\mathbb{1}_A \cdot X) = E(\mathbb{1}_A \cdot Y) \Rightarrow X = Y$ P -fast sicher

Lemma 33 (Eigenschaften des Erwartungswerts). Es sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum, X, Y und $\{X_n\}_{n \in \mathbb{N}}$ seien reellwertige Zufallsvariablen. Dann gilt:

j) $\forall \varepsilon > 0: P(|X| > \varepsilon) \leq \frac{E(|X|)}{\varepsilon}$ (Markoff-Ungleichung).

k) Ist die Folge $\{X_n\}_{n \in \mathbb{N}}$ nicht-negativ und schwach monoton wachsend, so gilt $\lim_{n \rightarrow \infty} E(X_n) = E\left(\lim_{n \rightarrow \infty} X_n\right)$.

l) Ist die Folge $\{X_n\}_{n \in \mathbb{N}}$ punktweise konvergent gegen X und gilt $|X_n| \leq Y$ für alle $n \in \mathbb{N}$ mit $E(Y) < \infty$, so gilt ebenfalls $\lim_{n \rightarrow \infty} E(X_n) = E\left(\lim_{n \rightarrow \infty} X_n\right) = E(X)$ sowie $\lim_{n \rightarrow \infty} E(|X_n - X|) = 0$.

Satz 32 (Berechnung von Erwartungswerten). Es seien X und Y reellwertige Zufallsvariablen auf einem Wahrscheinlichkeitsraum (Ω, \mathcal{A}, P) .

- a) Ist X *diskret* verteilt, d.h. gibt es eine höchstens abzählbare Menge $T \subset \mathbb{R}$ mit $P^X(T) = 1$ und $P(X = x) > 0$ für alle $x \in T$, so ist X genau dann eigentlich P -integrierbar, wenn

$$\sum_{x \in T} |x| P(X = x) < \infty$$

gilt. In diesem Fall lässt sich der Erwartungswert darstellen als

$$E(X) = \sum_{x \in T} x \cdot P(X = x).$$

Ist des Weiteren P selbst eine *diskrete Verteilung*, d.h. gilt $\{\omega\} \in \mathcal{A}$ für alle $\omega \in \Omega$ und $P(\{\omega\}) > 0$ für höchstens abzählbar viele $\omega \in \Omega$, so gilt analog

$$E(X) = \sum_{\omega \in \Omega} X(\omega) P(\{\omega\}).$$

Satz 32 (Berechnung von Erwartungswerten). Es seien X und Y reellwertige Zufallsvariablen auf einem Wahrscheinlichkeitsraum (Ω, \mathcal{A}, P) .

b) Besitzt P^X eine Dichte bezüglich des Lebesgue-Maßes, so ist X genau dann eigentlich P -integrierbar, wenn

$$\int_{-\infty}^{\infty} |x| f(x) dx < \infty$$

gilt. In diesem Fall lässt sich der Erwartungswert darstellen als

$$E(X) = \int_{-\infty}^{\infty} x \cdot f(x) dx.$$

Satz 32 (Berechnung von Erwartungswerten). Es seien X und Y reellwertige Zufallsvariablen auf einem Wahrscheinlichkeitsraum (Ω, \mathcal{A}, P) .

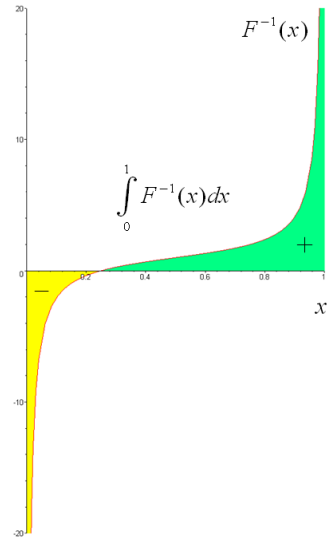
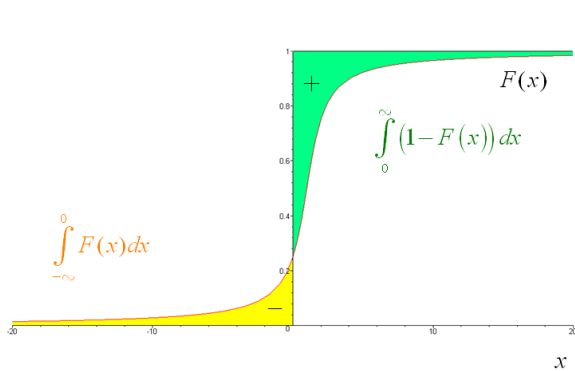
c) Bezeichnet F die Verteilungsfunktion von X und ist X eigentlich P -integrierbar, so gilt

$$E(X) = - \int_{-\infty}^0 F(x) dx + \int_0^{\infty} (1 - F(x)) dx.$$

d) Ist X eigentlich P -integrierbar, so lässt sich der Erwartungswert auch darstellen als

$$E(X) = \int_0^1 F^{-1}(x) dx,$$

wobei F^{-1} wieder die Pseudo-Inverse von F bezeichne.



Veranschaulichung des Erwartungswertes

Lemma 34. Es sei \mathbf{X} ein Zufallsvektor auf einem Wahrscheinlichkeitsraum (Ω, \mathcal{A}, P) mit Werten in \mathbb{R}^d , $d \in \mathbb{N}$. Ferner sei $G: \mathbb{R}^d \rightarrow \mathbb{R}$ eine Borel-messbare Abbildung. Dann gilt: ist $G(\mathbf{X})$ P -integrierbar, so gilt

$$E(G(\mathbf{X})) = \int G(\mathbf{X}) dP = \int G dP^{\mathbf{X}}.$$

Besitzt darüber hinaus die Verteilung $P^{\mathbf{X}}$ eine Dichte f , so gilt auch

$$E(G(\mathbf{X})) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} G(x_1, \dots, x_d) \cdot f(x_1, \dots, x_d) dx_1 \cdots dx_d.$$

Lemma 35. Es seien X_1, \dots, X_d stochastisch unabhängige, reellwertige Zufallsvariablen auf einem Wahrscheinlichkeitsraum (Ω, \mathcal{A}, P) . Sind dann alle X_i P -integrierbar, so auch $Y := \prod_{i=1}^d X_i$, und es gilt

$$E(Y) = E\left(\prod_{i=1}^d X_i\right) = \prod_{i=1}^d E(X_i).$$

Lemma 36 (Jensen'sche Ungleichung). Es sei $\mathbf{X} = (X_1, \dots, X_d)$ ein d -dimensionaler Zufallsvektor auf einem Wahrscheinlichkeitsraum (Ω, \mathcal{A}, P) mit Werten in einer konvexen Menge $\mathcal{M} \in \mathcal{B}^d$ mit $d \in \mathbb{N}$ und $G: \mathcal{M} \rightarrow \mathbb{R}$ eine messbare, konvexe (konkave) Abbildung. $G(\mathbf{X})$ sowie die Komponenten X_1, \dots, X_d von \mathbf{X} seien sämtlich P -integrierbar. Dann ist $(E(X_1), \dots, E(X_d)) \in \mathcal{M}$, und es gilt:

$$E(G(\mathbf{X})) \begin{cases} \geq G(E(X_1), \dots, E(X_m)), & \text{falls } G \text{ konvex} \\ \leq G(E(X_1), \dots, E(X_m)), & \text{falls } G \text{ konkav} \end{cases}$$

Ist G nicht-negativ und konkav, braucht dabei lediglich die Integrierbarkeit aller Komponenten von \mathbf{X} gefordert zu werden.

Definition 33 (Varianz, Kovarianz und höhere Momente). Es seien X und Y Zufallsvariablen auf einem Wahrscheinlichkeitsraum (Ω, \mathcal{A}, P) sowie $k \in \mathbb{N}$.

a) Ist X^2 P -integrierbar, so auch X ; in diesem Fall heißt

$$\text{Var}(X) := E\left(\left(X - E(X)\right)^2\right)$$

die *Varianz* von X bzw. von P^X .

b) Sind X, Y und $X \cdot Y$ jeweils P -integrierbar, so heißt

$$\text{Kov}(X, Y) := E\left[\left(X - E(X)\right)\left(Y - E(Y)\right)\right]$$

die *Kovarianz* von X und Y . Ist $\text{Kov}(X, Y) = 0$, so heißen X und Y *unkorreliert*.

Definition 33 (Varianz, Kovarianz und höhere Momente). Es seien X und Y Zufallsvariablen auf einem Wahrscheinlichkeitsraum (Ω, \mathcal{A}, P) sowie $k \in \mathbb{N}$.

- c) Ist $|X|^k$ P -integrierbar, so auch X ; in diesem Fall heißt
- $E|X|^k$ das k -te absolute Moment von X
 - $E(X^k)$ das k -te Moment von X
 - $E(|X - E(X)|^k)$ das k -te absolute zentrale Moment von X und
 - $E((X - E(X))^k)$ das k -te zentrale Moment von X (bzw. P^X).

Lemma 37 (Eigenschaften von Momenten). Es seien X und Y Zufallsvariablen auf einem Wahrscheinlichkeitsraum (Ω, \mathcal{A}, P) .

a) Sind X^2, Y und $X \cdot Y$ P -integrierbar, so ist auch

$$\text{Var}(X) = \text{Kov}(X, X) = E(X^2) - (E(X))^2$$

und

$$\text{Kov}(X, Y) = E(XY) - E(X)E(Y).$$

Für beliebige $c \in \mathbb{R}$ gilt darüber hinaus:

$$E((X - c)^2) = \text{Var}(X) + (E(X) - c)^2,$$

d.h. $E((X - c)^2)$ ist minimal für $c = E(X)$. Ferner gilt:

$$\text{Var}(X) = 0 \Leftrightarrow X = \text{const } P\text{-fast sicher.}$$

Lemma 37 (Eigenschaften von Momenten). Es seien X und Y Zufallsvariablen auf einem Wahrscheinlichkeitsraum (Ω, \mathcal{A}, P) .

b) Sind X^2, Y und $X \cdot Y$ P -integrierbar, so gilt:

$$\text{Var}(aX + b) = a^2 \cdot \text{Var}(X)$$

$$\text{Kov}(aX + b, cY + d) = ac \cdot \text{Kov}(X, Y)$$

für alle $a, b, c, d \in \mathbb{R}$. Ferner ist

$$\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) + 2 \text{Kov}(X, Y).$$

Sind speziell X und Y unkorreliert, so gilt

$$\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y).$$

Lemma 37 (Eigenschaften von Momenten). Es seien X und Y Zufallsvariablen auf einem Wahrscheinlichkeitsraum (Ω, \mathcal{A}, P) .

c) Ist X^2 P -integrierbar, so gilt

$$P(|X - E(X)| > \varepsilon) \leq \frac{\text{Var}(X)}{\varepsilon^2}$$

für alle $\varepsilon > 0$ (Tschebyscheff-Ungleichung).

d) Existiert für ein $k \in \mathbb{N}$ das k -te absolute Moment von X , so existieren auch alle m -ten Momente von X für $m \leq k$ (d.h. absolute und nicht-absolute, zentrale und nicht-zentrale), und es gilt

$$E((X - c)^m) \leq E(|X - c|^m) \leq \left(E(|X - c|^k)\right)^{m/k} \leq 2^m \left(E(|X|^k) + |c|^k\right)^{m/k}$$

für alle $c \in \mathbb{R}$.

Lemma 37 (Eigenschaften von Momenten). Es seien X und Y Zufallsvariablen auf einem Wahrscheinlichkeitsraum (Ω, \mathcal{A}, P) .

Bezeichnet F die Verteilungsfunktion von X bzw. P^X , so gilt auch die Darstellung

$$E(X^k) = (-1)^k \int_{-\infty}^0 kx^{k-1}F(x)dx + \int_0^{\infty} kx^{k-1}(1-F(x))dx = \int_0^1 (F^{-1}(x))^k dx$$

$$E(|X|^k) = \int_{-\infty}^0 kx^{k-1}F(x)dx + \int_0^{\infty} kx^{k-1}(1-F(x))dx = \int_0^1 |F^{-1}(x)|^k dx.$$

e) Sind $|X|^p$ und $|Y|^q$ P -integrierbar für $p, q > 1$ mit $\frac{1}{p} + \frac{1}{q} = 1$, so ist auch $|XY|$ P -integrierbar, und es gilt

$$E(|X \cdot Y|) \leq \left(E(|X|^p)\right)^{1/p} \cdot \left(E(|Y|^q)\right)^{1/q} \quad (\text{Hölder-Ungleichung}).$$

Lemma 37 (Eigenschaften von Momenten). Es seien X und Y Zufallsvariablen auf einem Wahrscheinlichkeitsraum (Ω, \mathcal{A}, P) .

Ist speziell $p = q = 2$, so gilt auch

$$|\text{Kov}(X, Y)| \leq \sqrt{\text{Var}(X) \cdot \text{Var}(Y)}$$

mit Gleichheit genau dann, wenn X und Y P -f.s. linear voneinander abhängen, d.h. wenn Zahlen $a, b, c \in \mathbb{R}$, a, b nicht beide Null, existieren mit $aX + bY = c$ P -fast sicher.

Definition 34 (Korrelation). Es seien X und Y Zufallsvariablen auf einem Wahrscheinlichkeitsraum (Ω, \mathcal{A}, P) mit $0 < \text{Var}(X), \text{Var}(Y) < \infty$. Dann heißt

$$\text{Korr}(X, Y) = \frac{\text{Kov}(X, Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}}$$

die Korrelation zwischen X und Y .

Definition 35 (Faltung). Es seien X und Y reellwertige, stochastisch unabhängige Zufallsvariablen. Dann heißt die Verteilung der Summe $X + Y$ die Faltung der Verteilungen von X und Y , in Zeichen:

$$P^X * P^Y := P^{X+Y}.$$

Lemma 38. Es seien X und Y stochastisch unabhängige, reellwertige Zufallsvariablen. Ihre Verteilungsfunktionen seien mit F_X und F_Y bezeichnet, bei stetigen Verteilungen ferner ihre Dichten entsprechend mit f_X und f_Y . Die Verteilungsfunktion der Faltung $P^X * P^Y$ sei mit $F_X * F_Y$ bezeichnet, entsprechend im Fall der Existenz die Dichte mit $f_X * f_Y$. Dann gilt:

$$F_X * F_Y(z) = F_Y * F_X(z) = \int_{-\infty}^{\infty} F_X(z - y) P^Y(dy) = \int_{-\infty}^{\infty} F_Y(z - x) P^X(dx)$$

$$F_X * F_Y(z) = F_Y * F_X(z) = \int_{-\infty}^{\infty} F_X(z - y) f_Y(y) dy = \int_{-\infty}^{\infty} F_Y(z - x) f_X(x) dx$$

$$f_X * f_Y(z) = f_Y * f_X(z) = \int_{-\infty}^{\infty} f_X(z - y) f_Y(y) dy = \int_{-\infty}^{\infty} f_Y(z - x) f_X(x) dx, \quad z \in \mathbb{R}.$$

Lemma 39. Es seien X und Y stochastisch unabhängige, reellwertige Zufallsvariablen mit Werten in \mathbb{Z} . Dann gilt:

$$F_X * F_Y(n) = P(X + Y \leq n) = \sum_{k \in \mathbb{Z}} F_Y(n - k) \cdot P(X = k) = \sum_{k \in \mathbb{Z}} F_X(n - k) \cdot P(Y = k)$$

$$P(X + Y = n) = \sum_{k \in \mathbb{Z}} P(X = k) \cdot P(Y = n - k) = \sum_{k \in \mathbb{Z}} P(Y = k) \cdot P(X = n - k)$$

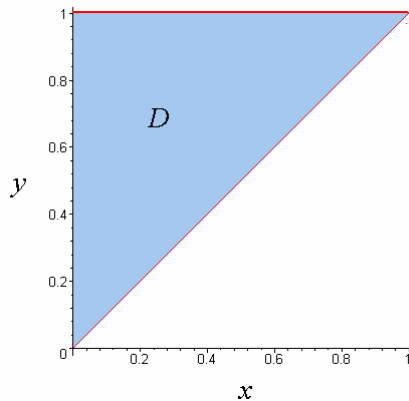
für alle $n \in \mathbb{Z}$.

Beispiel für abhängige Zufallsvariablen:

Der Zufallsvektor $Z = (X, Y)$ besitze eine auf das Dreieck $D := \{(x, y) \mid 0 \leq x \leq y \leq 1\}$ konzentrierte Lebesgue-Dichte der Form

$$f_z(x, y) = c \cdot (x + y) \text{ für } (x, y) \in D$$

mit einer passenden positiven Konstanten c .
Diese erhalten wir aus der Bedingung



$$\begin{aligned} 1 = P(Z \in D) &= \iint_D f_z(x, y) dx dy = c \int_0^1 \int_0^y x + y dx dy \\ &= c \int_0^1 \frac{x^2}{2} + xy \Big|_{x=0}^{x=y} dy = c \int_0^1 \frac{3}{2} y^2 dy = c \frac{y^3}{2} \Big|_{y=0}^{y=1} = \frac{c}{2}, \text{ also } c = 2. \end{aligned}$$

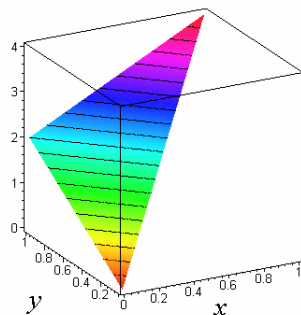
Achtung: Man muss bei der Festlegung der Integrationsgrenzen in der iterativen Auflösung auf die Reihenfolge der Integration achten. Alternativ gilt auch

$$\begin{aligned} 1 = P(Z \in D) &= \iint_D f_z(x, y) dy dx = c \int_0^1 \int_x^1 x + y dy dx = c \int_0^1 xy + \frac{y^2}{2} \Big|_{y=x}^{y=1} dx \\ &= c \int_0^1 \frac{1}{2} + x - \frac{3}{2}x^2 dx = c \left(\frac{x}{2} + \frac{x^2}{2} - \frac{x^3}{2} \Big|_{x=0}^{x=1} \right) = \frac{c}{2}, \text{ also wieder } c = 2. \end{aligned}$$

Die jeweiligen Integrationsgrenzen ergeben sich hierbei aus der das Dreieck charakterisierenden Ungleichung $0 \leq x \leq y \leq 1$.

Die Dichte von Z lautet somit:

$$f_z(x, y) = 2(x + y), (x, y) \in D \Leftrightarrow 0 \leq x \leq y \leq 1.$$

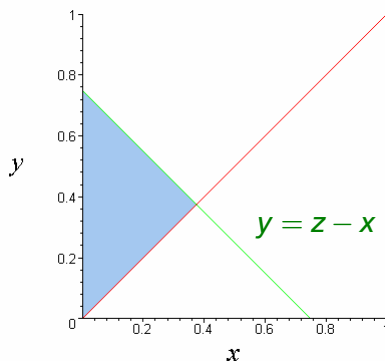


Graph der Dichte von Z

Berechnung der Summendichte für $S := X + Y$:

Fall1: $0 \leq z \leq 1$: Es sind die Ungleichungen $0 \leq x \leq y \leq 1$ und $x + y \leq z$ zu beachten. Diese beschreiben das rechts blau markierte Teildreieck, das bestimmt ist durch die Ungleichungen

$$0 \leq x \leq \frac{z}{2}, \quad x \leq y \leq z - x.$$



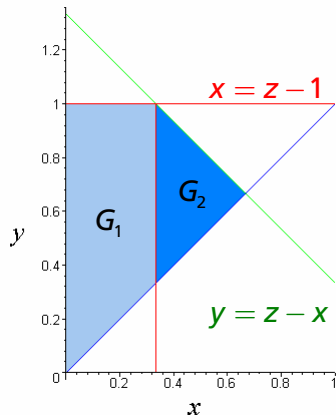
Mit der inneren Integration nach y folgt

$$\begin{aligned} F_S(z) &= P(X + Y \leq z) = \iint_{\substack{x+y \leq z \\ 0 \leq x \leq y \leq 1}} f_z(x, y) dy dx = 2 \int_0^{z/2} \int_x^{z-x} x + y dy dx \\ &= 2 \int_0^{z/2} xy + \frac{y^2}{2} \Big|_{y=x}^{y=z-x} dx = \int_0^{z/2} z^2 - 4x^2 dx = z^2 x - \frac{4}{3} x^3 \Big|_{x=0}^{x=z/2} = \left(\frac{1}{2} - \frac{1}{6} \right) z^3 = \frac{z^3}{3}. \end{aligned}$$

Fall 2: $1 \leq z \leq 2$: Es sind wieder die Ungleichungen $0 \leq x \leq y \leq 1$ und $x + y \leq z$ zu beachten. Diese beschreiben diesmal die Vereinigung der beiden rechts blau markierten Gebiete G_1 und G_2 , die bestimmt sind durch die Ungleichungen

$$0 \leq x \leq z-1, \quad x \leq y \leq 1 \text{ für } G_1$$

$$z-1 \leq x \leq \frac{z}{2}, \quad x \leq y \leq z-x \text{ für } G_2.$$



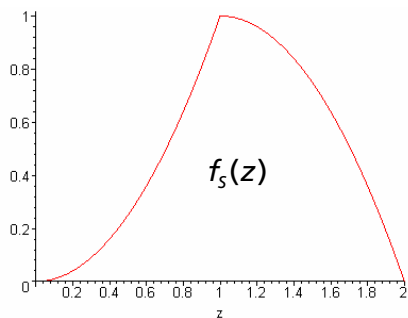
Mit der inneren Integration nach y folgt

$$F_S(z) = P(X + Y \leq z) = 2 \int_0^{z-1} \int_x^1 x + y \, dy \, dx + 2 \int_{z-1}^{z/2} \int_x^{z-x} x + y \, dy \, dx = 2 \int_0^{z-1} \frac{1}{2} + x - \frac{3}{2} x^2 \, dx + \dots$$

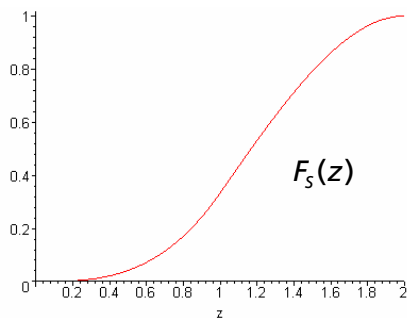
$$\dots + 2 \int_{z-1}^{z/2} z^2 - 4x^2 \, dx = x + x^2 - x^3 \Big|_0^{z-1} + \left(2z^2 x - \frac{8}{3} x^3 \Big|_{z-1}^{z/2} \right) = -\frac{1}{3} + z^2 - \frac{z^3}{3}.$$

Durch Differenzieren erhalten wir damit auch die Summendichte:

$$f_S(z) = \begin{cases} z^2, & 0 \leq z \leq 1 \\ 2z - z^2, & 1 \leq z \leq 2 \end{cases}$$



Dichte von S



Verteilungsfunktion von S

Beispielhafte Berechnung von Momenten:

$$\begin{aligned} E(X) &= \int_D \int x \cdot f_z(x, y) dy dx = 2 \int_0^1 \int_0^y x^2 + xy dx dy \\ &= 2 \int_0^1 \left. \frac{x^3}{3} + \frac{x^2}{2} y \right|_{x=0}^{x=y} dy = \frac{5}{3} \int_0^1 y^3 dy = \frac{5}{12} \end{aligned}$$

$$\begin{aligned} E(Y) &= \int_D \int y \cdot f_z(x, y) dy dx = 2 \int_0^1 \int_0^y xy + y^2 dx dy \\ &= \int_0^1 \left. x^2 y + 2xy^2 \right|_{x=0}^{x=y} dy = 3 \int_0^1 y^3 dy = \frac{3}{4} \end{aligned}$$

Beispielhafte Berechnung von Momenten:

$$\begin{aligned} E(X \cdot Y) &= \iint_D x \cdot y \cdot f_z(x, y) dy dx = 2 \int_0^1 \int_0^y x^2 y + xy^2 dx dy \\ &= 2 \int_0^1 \left. \frac{x^3}{3} y + \frac{x^2}{2} y^2 \right|_{x=0}^{x=y} dy = \frac{5}{3} \int_0^1 y^4 dy = \frac{1}{3} \end{aligned}$$

$$\text{Kov}(X, Y) = E(X \cdot Y) - E(X) \cdot E(Y) = \frac{1}{3} - \frac{5}{12} \cdot \frac{3}{4} = \frac{1}{48}$$

II. Grundprinzipien stochastischer Modellierung

II.4. Erzeugende Funktionen

Definition 36 (erzeugende Funktionen). Es sei X eine reellwertige Zufallsvariable derart, dass für eine Teilmenge $I \subseteq \mathbb{R}$ der Ausdruck

$$\psi_X(t) := E(e^{tx}), \quad t \in I$$

für alle $t \in I$ endlich ist. Dann heißt die auf I definierte Abbildung ψ_X die *momenterzeugende* Funktion zu X bzw. zu der Verteilung P^X .

Die durch

$$\varphi_X(s) := \psi_X(\ln s) = E(s^X), \quad s \in e^I := \{e^t \mid t \in I\}$$

definierte Funktion heißt die *wahrscheinlichkeitserzeugende* Funktion zu X bzw. zu der Verteilung P^X .

Satz 33. Es sei X eine reellwertige Zufallsvariable derart, dass für eine Teilmenge $I \subseteq \mathbb{R}$ die momenterzeugende Funktion ψ_X existiert. Dann gilt:

- a) Es ist stets $\psi_X(0) = \varphi_X(1) = 1$. Existiert ferner $\psi_X(t)$ für ein $t = t^* > 0$ bzw. $t = t_* < 0$, so auch für alle $t \in [0, t^*]$ bzw. $t \in [t_*, 0]$.
Bezeichnet speziell

$$t^+ := \sup\{t \in \mathbb{R} \mid \psi_X(t) < \infty\}, \quad t^- := \inf\{t \in \mathbb{R} \mid \psi_X(t) < \infty\},$$

so existiert $\psi_X(t)$ für alle $t \in (t^-, t^+)$, und $\varphi_X(s)$ existiert für alle $s \in (e^{t^-}, e^{t^+})$ (mit der Konvention $e^{-\infty} = 0$, $e^{\infty} = \infty$). Gilt insbesondere $X \geq 0$ bzw. $X \leq 0$ (mit Wahrscheinlichkeit 1), so ist $t^- = -\infty$ bzw. $t^+ = +\infty$.

Satz 33. Es sei X eine reellwertige Zufallsvariable derart, dass für eine Teilmenge $I \subseteq \mathbb{R}$ die momenterzeugende Funktion ψ_X existiert. Dann gilt:

b) Es sei $0 < \delta < \min\{t^+, -t^-\}$. Dann existieren sämtliche Momente $E(|X|^k)$, $k \in \mathbb{N}$, ψ_X ist im Nullpunkt beliebig oft differenzierbar, und es gilt

$$\psi_X^{(k)}(\mathbf{0}) = E(X^k), \quad k \in \mathbb{N} \text{ und}$$

$$\psi_X(t) = \sum_{k=0}^{\infty} \frac{E(X^k)}{k!} t^k, \quad |t| \leq \delta.$$

Insbesondere ist

$$E(X) = \psi_X'(\mathbf{0}), \quad \text{Var}(X) = \psi_X''(\mathbf{0}) - \{\psi_X'(\mathbf{0})\}^2.$$

Satz 33. Es sei X eine reellwertige Zufallsvariable derart, dass für eine Teilmenge $I \subseteq \mathbb{R}$ die momenterzeugende Funktion ψ_X existiert. Dann gilt:

Ferner ist $\varphi_X(s)$ für $s = 1$ differenzierbar, und es gilt

$$\varphi_X^{(k)}(1) = E\left(\prod_{i=0}^{k-1} (X - i)\right), \quad k \in \mathbb{N} \text{ und}$$
$$\varphi_X(s) = \sum_{k=0}^{\infty} \frac{E\left(\prod_{i=0}^{k-1} (X - i)\right)}{k!} (s-1)^k, \quad |s-1| \leq 1 - e^{-\delta}.$$

Insbesondere ist

$$E(X) = \varphi_X'(1), \quad \text{Var}(X) = \varphi_X''(1) + \varphi_X'(1)\{1 - \varphi_X'(1)\}.$$

Satz 33. Es sei X eine reellwertige Zufallsvariable derart, dass für eine Teilmenge $I \subseteq \mathbb{R}$ die momenterzeugende Funktion ψ_X existiert. Dann gilt:

- c) Gilt $P(X \in \mathbb{Z}^+) = 1$ mit $\mathbb{Z}^+ := \{0, 1, 2, \dots\}$, so lässt sich φ_X fortsetzen, d.h. es existiert $\varphi_X(s) = E(s^X)$ auch für alle $|s| \leq 1$, und es gilt

$$\frac{\varphi_X^{(k)}(0)}{k!} = P(X = k), \quad k \in \mathbb{N} \text{ und}$$

$$\varphi_X(s) = \sum_{k=0}^{\infty} P(X = k) s^k, \quad |s| \leq 1.$$

- d) Sind X und Y stochastisch unabhängige, reellwertige Zufallsvariablen mit momenterzeugenden Funktionen ψ_X und ψ_Y , die beide in derselben Menge $I \subseteq \mathbb{R}$ existieren, so besitzt dort auch die Zufallsvariable $Z = X + Y$ eine momenterzeugende Funktion, und es gilt

$$\psi_{X+Y}(t) = \psi_X(t) \cdot \psi_Y(t), \quad t \in I$$

bzw. auch

$$\varphi_{X+Y}(s) = \varphi_X(s) \cdot \varphi_Y(s), \quad s \in e^I$$

II. Grundprinzipien stochastischer Modellierung

II.5. Grundlegende Verteilungen

p^x	Zähldichte $f(k) = P(X = k)$	$\varphi_x(s)$	$E(X)$	$Var(X)$
U_n	$\frac{1}{n}, k = 1, \dots, n \in \mathbb{N}$	$\frac{s \cdot s^n - 1}{n \cdot s - 1}$	$\frac{n+1}{2}$	$\frac{n^2 - 1}{12}$
$B(n, p)$	$\binom{n}{k} p^k (1-p)^{n-k}, k = 0, \dots, n; p \in [0, 1]$	$(1-p + ps)^n$	np	$np(1-p)$
$NB(\beta, p)$	$\binom{\beta+k-1}{k} p^\beta (1-p)^k, k \in \mathbb{Z}^+; \beta > 0, p \in (0, 1]$	$\left(\frac{p}{1-(1-p)s}\right)^\beta$	$\beta \frac{1-p}{p}$	$\beta \frac{1-p}{p^2}$
$\mathcal{P}(\lambda)$	$e^{-\lambda} \frac{\lambda^k}{k!}, k \in \mathbb{Z}^+; \lambda > 0$	$e^{\lambda(s-1)}$	λ	λ

U_n	Gleichverteilung
$B(n, p)$	Binomialverteilung
$NB(\beta, p)$	negative Binomialverteilung
$\mathcal{P}(\lambda)$	Poisson-Verteilung

Lemma 40. Ist X eine $\mathcal{G}(p)$ -verteilte Zufallsvariable mit $0 < p < 1$, so gilt:

$$P(X = m \mid X \geq n) = P(X = m - n) \text{ für alle } 0 \leq n \leq m \in \mathbb{Z}^+.$$

Satz 34 (Poisson 1837). Es seien $\lambda, \nu > 0$ und $\{p_n\}_{n \in \mathbb{N}}, \{q_n\}_{n \in \mathbb{N}} \subseteq (0, 1)$ Folgen reeller Zahlen mit $\lim_{n \rightarrow \infty} n \cdot p_n = \lambda$, $\lim_{n \rightarrow \infty} n \cdot (1 - q_n) = \nu$. Ferner seien X, Y, X_n, Y_n Zufallsvariablen mit $P^X = \mathcal{P}(\lambda)$, $P^Y = \mathcal{P}(\nu)$, $P^{X_n} = B(n, p_n)$, $P^{Y_n} = NB(n, q_n)$ für alle $n \in \mathbb{N}$. Dann gilt:

$$P(X = k) = \lim_{n \rightarrow \infty} P(X_n = k), \quad P(Y = k) = \lim_{n \rightarrow \infty} P(Y_n = k) \quad \text{für alle } k \in \mathbb{Z}^+.$$

P^x	Dichte $f(x)$	$\psi_x(t)$	$E(X)$	$Var(X)$
$\mathcal{U}[a, b]$	$\frac{1}{b-a}, a \leq x \leq b$	$\frac{e^{bt} - e^{at}}{t(b-a)}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
$\mathcal{E}(\lambda)$	$\lambda e^{-\lambda x}, x \geq 0; \lambda > 0$	$\frac{\lambda}{\lambda - t}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
$\Gamma(\alpha, \lambda)$	$\lambda^\alpha \frac{x^{\alpha-1}}{\Gamma(\alpha)} e^{-\lambda x}, x > 0; \alpha, \lambda > 0$	$\left(\frac{\lambda}{\lambda - t}\right)^\alpha$	$\frac{\alpha}{\lambda}$	$\frac{\alpha}{\lambda^2}$
$\mathcal{N}(\mu, \sigma^2)$	$\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\left(\frac{(x-\mu)^2}{2\sigma^2}\right)}, x, \mu \in \mathbb{R}, \sigma > 0$	$\exp\left(\frac{\sigma^2 t^2}{2} + \mu t\right)$	μ	σ^2

$\mathcal{U}[a, b]$	(stetige) Gleichverteilung
$\mathcal{E}(\lambda)$	Exponentialverteilung
$\Gamma(\alpha, \lambda)$	Gamma-Verteilung
$\mathcal{N}(\mu, \sigma^2)$	Normalverteilung

Lemma 41. Ist X eine $\mathcal{E}(\lambda)$ -verteilte Zufallsvariable mit $\lambda > 0$, so gilt:

$$P(X > z \mid X > x) = P(X > z - x) \text{ für alle } 0 \leq x \leq z \in \mathbb{R}^+.$$

Satz 35. Ist X eine Zufallsvariable mit Werten in \mathbb{R}^+ und einer *stetigen* Verteilungsfunktion F , und gilt

$$P(X > z \mid X > x) = P(X > z - x) \text{ für alle } 0 \leq x \leq z \in \mathbb{R}^+,$$

so ist notwendig X exponentialverteilt mit einem Parameter $\lambda > 0$.

Lemma 42. Ist X eine $\mathcal{E}(\lambda)$ -verteilte Zufallsvariable mit $\lambda > 0$, so ist $Y := \lfloor X \rfloor$ geometrisch verteilt mit Parameter $p = 1 - e^{-\lambda}$. Dabei bezeichne $\lfloor z \rfloor := \max\{m \in \mathbb{Z} \mid m \leq z\}$ für alle $z \in \mathbb{R}$ (größte ganze Zahl unterhalb von z , "Gauß-Klammer"). Ist umgekehrt $\{X_n\}_{n \in \mathbb{N}}$ eine Folge geometrisch verteilter Zufallsvariablen mit Parametern $p_n = 1 - e^{-\lambda/n}$, so ist $Y_n := \frac{X_n}{n}$ asymptotisch $\mathcal{E}(\lambda)$ -verteilt, d.h. hier:

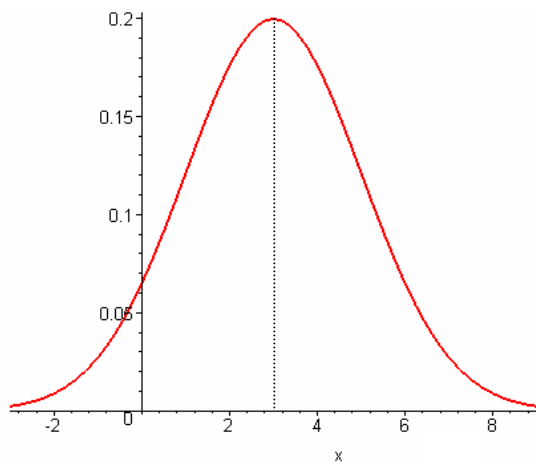
$$\lim_{n \rightarrow \infty} P(Y_n \leq x) = 1 - e^{-\lambda x} \text{ für alle } x \geq 0.$$

Lemma 43. Ist X eine $\mathcal{N}(0,1)$ -verteilte Zufallsvariable, so ist $Y := \mu + \sigma X$ mit $\mu \in \mathbb{R}$, $\sigma > 0$ $\mathcal{N}(\mu, \sigma^2)$ -verteilt.

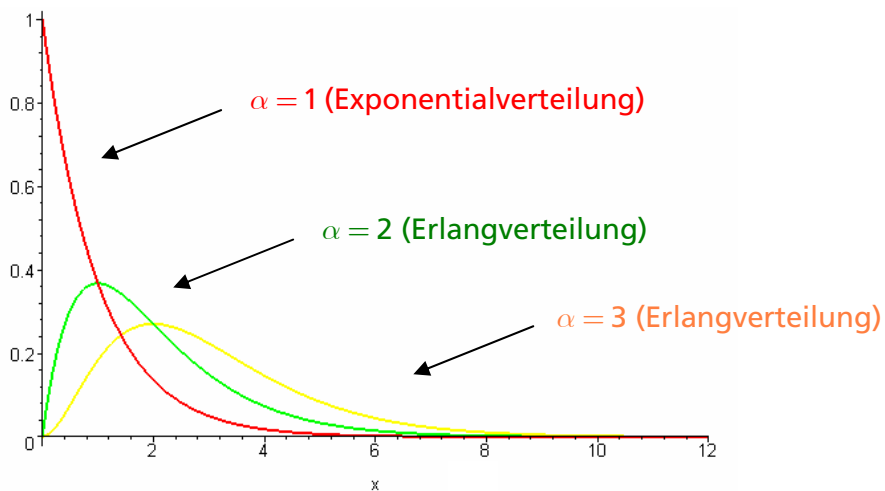
Ist umgekehrt Y $\mathcal{N}(\mu, \sigma^2)$ -verteilt, so ist $X := \frac{Y - \mu}{\sigma}$ $\mathcal{N}(0,1)$ -verteilt.

Lemma 44. Für eine $\mathcal{N}(\mu, \sigma^2)$ -verteilte Zufallsvariable X gilt:

$$\psi_X(t) = \exp\left(\frac{\sigma^2 t^2}{2} + \mu t\right), \quad t \in \mathbb{R}; \quad y \in \mathbb{R}, \quad \sigma > 0.$$



Dichte der Normalverteilung $\mathcal{N}(3,2)$



Dichten verschiedener $\Gamma(\alpha, 1)$ -Verteilungen

p^x	Dichte $f(x)$, $x \in \mathbb{R}$	$E(X)$	$Var(X)$
$\chi_n^2 = \Gamma\left(\frac{n}{2}, \frac{1}{2}\right)$	$\frac{x^{(n-2)/2}}{\sqrt{2^n} \Gamma(n/2)} e^{-x/2}, x > 0$	n	$2n$
$\mathcal{B}(\alpha, \beta)$	$\frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha, \beta)}, 0 \leq x \leq 1$	$\frac{\alpha}{\alpha + \beta}$	$\frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)}$
$\mathcal{LN}(\mu, \sigma^2)$	$\frac{1}{x\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(\ln x - \mu)^2}{2\sigma^2}\right), x > 0$	$e^{\mu + \sigma^2/2}$	$e^{2\mu + \sigma^2} (e^{\sigma^2} - 1)$

$\chi_n^2 = \Gamma\left(\frac{n}{2}, \frac{1}{2}\right)$	χ^2 -Verteilung ($n \in \mathbb{N}$)
$\mathcal{B}(\alpha, \beta)$	Beta-Verteilung ($\alpha, \beta > 0$)
$\mathcal{LN}(\mu, \sigma^2)$	Log-Normalverteilung ($\mu \in \mathbb{R}, \sigma > 0$)

Die χ^2 -Verteilung ergibt sich als Verteilung einer Summe aus Quadraten von n stochastisch unabhängigen standard-normalverteilten Zufallsvariablen.

P^X	Dichte $f(x)$, $x \in \mathbb{R}$	$E(X)$	$Var(X)$
\mathcal{L}	$\frac{e^{-x}}{(1+e^{-x})^2}$	0	$\frac{\pi^2}{3}$
\mathcal{LL}	$\frac{\alpha x^{\alpha-1}}{(1+x^\alpha)^2}, x > 0$	$\frac{\pi}{\alpha \sin \frac{\pi}{\alpha}}$	$\frac{\pi \left(\sin \frac{\pi}{\alpha} - \frac{\pi}{\alpha} \cos \frac{\pi}{\alpha} \right)}{\alpha \sin^3 \frac{\pi}{\alpha} \cos \frac{\pi}{\alpha}}$
\mathcal{C}	$\frac{1}{\pi} \frac{1}{1+x^2}$	—	—
$\mathcal{IN}(\mu, \sigma^2)$	$\frac{1}{\sqrt{2\pi\sigma^2}x^3} \exp\left(-\frac{(x-\mu)^2}{2\mu^2\sigma^2x}\right), x > 0$	μ	$\mu^3\sigma^2$

\mathcal{L}	Logistische Verteilung
\mathcal{LL}	Log-Logistische Verteilung ($\alpha > 0$)
\mathcal{C}	Cauchy-Verteilung
$\mathcal{IN}(\mu, \sigma^2)$	Inverse Gauß-Verteilung ($\mu, \sigma > 0$)

p^x	Dichte $f(x)$, $x \in \mathbb{R}$	$E(X)$	$Var(X)$
$IG(\alpha)$	$\frac{e^{-1/x}}{x^{\alpha+1}\Gamma(\alpha)}$, $x > 0$	$\frac{1}{\alpha-1}$	$\frac{1}{(\alpha-1)^2(\alpha-2)}$
$Pa(\alpha)$	$\frac{\alpha}{(1+x)^{\alpha+1}}$, $x > 0$	$\frac{1}{\alpha-1}$	$\frac{\alpha}{(\alpha-1)(\alpha-2)}$
\mathcal{G}	$e^{-x} \cdot \exp(-e^{-x})$	γ	$\frac{\pi^2}{6}$
$\mathcal{F}(\alpha)$	$\frac{\alpha}{x^{\alpha+1}} \exp(-x^{-\alpha})$, $x > 0$	$\Gamma\left(1-\frac{1}{\alpha}\right)$	$\Gamma\left(1-\frac{2}{\alpha}\right) - \Gamma^2\left(1-\frac{1}{\alpha}\right)$
$\mathcal{W}(\alpha)$	$\alpha x^{\alpha-1} \exp(-x^\alpha)$, $x > 0$	$\Gamma\left(1+\frac{1}{\alpha}\right)$	$\Gamma\left(1+\frac{2}{\alpha}\right) - \Gamma^2\left(1+\frac{1}{\alpha}\right)$

$IG(\alpha)$	Inverse Gamma-Verteilung ($\alpha > 0$)
$Pa(\alpha)$	Pareto-Verteilung ($\alpha > 0$)
\mathcal{G}	Gumbel-Verteilung
$\mathcal{F}(\alpha)$	Fréchet-Verteilung ($\alpha > 0$)
$\mathcal{W}(\alpha)$	Weibull-Verteilung ($\alpha > 0$)

\mathcal{P}^x	Name	Verteilungsfunktion $F(x)$, $x \in \mathbb{R}$
\mathcal{L}	Logistische Verteilung	$\frac{1}{1 + e^{-x}}$
\mathcal{LL}	Log-Logistische Verteilung	$\frac{1}{1 + x^{-\alpha}}$, $x > 0$ $(\alpha > 0)$
\mathcal{C}	Cauchy-Verteilung	$\frac{1}{\pi} \arctan(x) + \frac{1}{2}$
$\mathcal{Pa}(\alpha)$	Pareto-Verteilung	$1 - \frac{1}{(1+x)^\alpha}$, $x > 0$ $(\alpha > 0)$
\mathcal{G}	Gumbel-Verteilung	$\exp(-e^{-x})$
$\mathcal{F}(\alpha)$	Fréchet-Verteilung	$\exp(-x^{-\alpha})$, $x > 0$ $(\alpha > 0)$
$\mathcal{W}(\alpha)$	Weibull-Verteilung	$1 - \exp(-x^\alpha)$, $x > 0$ $(\alpha > 0)$

P^X	Transformation	P^Y
$\mathcal{U}[0,1]$	$Y = a + (b - a)X$	$\mathcal{U}[a,b]$
$\mathcal{U}[0,1]$	$Y = -\frac{1}{\lambda} \ln(X)$	$\mathcal{E}(\lambda)$
$\mathcal{N}(\mu, \sigma^2)$	$Y = e^X$	$\mathcal{LN}(\mu, \sigma^2)$
$\mathcal{U}[0,1]$	$Y = \ln\left(\frac{X}{1-X}\right)$	\mathcal{L}
\mathcal{L}	$Y = e^{X/\alpha}$	\mathcal{LL}
$\mathcal{U}[0,1]$	$Y = -\cot(\pi X)$	\mathcal{C}
$\Gamma(\alpha, 1)$	$Y = \frac{1}{X}$	$\mathcal{IG}(\alpha)$

P^X	Transformation	P^Y
$\mathcal{U}[0,1]$	$Y = X^{-1/\alpha} - 1$	$\mathcal{Pa}(\alpha)$
$\mathcal{U}[0,1]$	$Y = -\ln(-\ln X)$	\mathcal{G}
$\mathcal{E}(1)$	$Y = -\ln(X)$	\mathcal{G}
\mathcal{G}	$Y = e^{X/\alpha}$	$\mathcal{F}(\alpha)$
$\mathcal{F}(\alpha)$	$Y = \frac{1}{X}$	$\mathcal{W}(\alpha)$
$\mathcal{E}(1)$	$Y = X^{1/\alpha}$	$\mathcal{W}(\alpha)$
$\mathcal{W}(\alpha)$	$Y = \frac{1}{X}$	$\mathcal{F}(\alpha)$

Dichte von P^X	Bereich	Transformation	Dichte von P^Y	Bereich
$f(x)$	$x > 0$	$Y = \frac{1}{X}$	$\frac{1}{x^2} f\left(\frac{1}{x}\right)$	$x > 0$
$f(x)$	$x > 0$	$Y = X^\alpha, \alpha > 0$	$\frac{1}{\alpha} x^{1/\alpha - 1} f(x^{1/\alpha})$	$x > 0$
$f(x)$	$x > 0$	$Y = \ln(X)$	$e^x \cdot f(e^x)$	$x \in \mathbb{R}$
$f(x)$	$x \in \mathbb{R}$	$Y = e^x$	$\frac{1}{x} f(\ln x)$	$x > 0$

II. Grundprinzipien stochastischer Modellierung

II.6. Das Gesetz der Großen Zahlen

Definition 37 (Stochastische Konvergenz): Eine Folge $\{X_n\}_{n \in \mathbb{N}}$ reeller Zufallsvariablen auf einem Wahrscheinlichkeitsraum (Ω, \mathcal{A}, P) heißt *stochastisch konvergent* gegen eine reelle Zufallsvariable X , wenn gilt

$$\lim_{n \rightarrow \infty} P(|X_n - X| > \varepsilon) = 0 \quad \text{für alle } \varepsilon > 0,$$

in Zeichen: $P\text{-}\lim_{n \rightarrow \infty} X_n = X$ oder $X_n \xrightarrow{P} X$ ($n \rightarrow \infty$). Die Folge heißt *gleichmäßig stochastisch konvergent* gegen eine reelle Zufallsvariable X , wenn gilt

$$\lim_{n \rightarrow \infty} P\left(\sup_{m \geq n} |X_m - X| > \varepsilon\right) = 0 \quad \text{für alle } \varepsilon > 0.$$

Definition 38 (fast sichere Konvergenz): Eine Folge $\{X_n\}_{n \in \mathbb{N}}$ reeller Zufallsvariablen auf einem Wahrscheinlichkeitsraum (Ω, \mathcal{A}, P) heißt *fast sicher konvergent* gegen eine reelle Zufallsvariable X , wenn gilt

$$P\left(\lim_{n \rightarrow \infty} X_n = X\right) = P\left(\bigcap_{k=1}^{\infty} \bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} \left\{|X_m - X| \leq \frac{1}{k}\right\}\right) = 1,$$

in Zeichen: $\lim_{n \rightarrow \infty} X_n = X$ f.s. oder $X_n \xrightarrow{\text{f.s.}} X$ ($n \rightarrow \infty$).

Definition 39: Es sei (Ω, \mathcal{A}) ein Messraum und $\{A_n\}_{n \in \mathbb{N}} \subseteq \mathcal{A}$ eine beliebige Folge von Ereignissen. Dann heißt

$\limsup_{n \rightarrow \infty} A_n := \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} A_m$ der Limes Superior der Mengenfolge,

$\liminf_{n \rightarrow \infty} A_n := \bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} A_m$ der Limes Inferior der Mengenfolge.

Satz 36: Es sei $\{X, X_n\}_{n \in \mathbb{N}}$ eine Folge reeller Zufallsvariablen auf einem Wahrscheinlichkeitsraum (Ω, \mathcal{A}, P) . Dann gelten folgende Aussagen:

a) $\lim_{n \rightarrow \infty} X_n = X$ f.s. $\Rightarrow P\text{-}\lim_{n \rightarrow \infty} X_n = X$,

d.h. *fast sichere* impliziert *stochastische* Konvergenz.

b) $P\text{-}\limsup_{n \rightarrow \infty} \sup_{m \geq n} |X_m - X| = 0 \Rightarrow \lim_{n \rightarrow \infty} X_n = X$ f.s.,

d.h. *gleichmäßige stochastische* Konvergenz impliziert *fast sichere* Konvergenz.

Den letzten Sachverhalt kann man auch so ausdrücken: *fast sichere* und *gleichmäßige stochastische* Konvergenz sind äquivalent.

Satz 37 (schwaches Gesetz der Großen Zahlen): Es sei $\{X_n\}_{n \in \mathbb{N}}$ eine Folge paarweise unkorrelierter quadratisch integrierbarer Zufallsvariablen auf einem Wahrscheinlichkeitsraum (Ω, \mathcal{A}, P) mit

$$E(X_n) = \mu \in \mathbb{R} \text{ für alle } n \in \mathbb{N} \quad \text{und} \quad \lim_{n \rightarrow \infty} \frac{1}{n^2} \sum_{k=1}^n \text{Var}(X_k) = 0.$$

Dann gilt für die arithmetischen Mittel $\bar{X}_n := \frac{1}{n} \sum_{k=1}^n X_k$ dieser Folge:

$$P\text{-}\lim_{n \rightarrow \infty} \bar{X}_n = \mu.$$

Lemma 45: Es sei (Ω, \mathcal{A}, P) ein Wahrscheinlichkeitsraum und $A \in \mathcal{A}$ ein Ereignis mit Eintrittswahrscheinlichkeit $p := P(A) \in (0, 1)$. Betrachtet wird ein Experiment mit unabhängiger Versuchswiederholung, d.h. der Produkt-Wahrscheinlichkeitsraum $(\Omega^*, \mathcal{A}^*, P^*) = \left(\prod_{n \in \mathbb{N}} \Omega, \otimes_{n \in \mathbb{N}} \mathcal{A}, \otimes_{n \in \mathbb{N}} P \right)$ mit der stochastisch unabhängigen Ereignisfolge $\{A_n\}_{n \in \mathbb{N}} \subseteq \mathcal{A}^*$, definiert durch

$$A_n := \prod_{i=1}^{n-1} \Omega \times A \times \prod_{i=n+1}^{\infty} \Omega \text{ für alle } n \in \mathbb{N},$$

d.h. A_n beschreibt das Eintreten von A im n -ten Experiment. Ferner sei $X_n := \mathbb{1}_{A_n}$ für alle $n \in \mathbb{N}$. Dann beschreibt $H_n := \frac{1}{n} \sum_{k=1}^n X_k$ die *relative Häufigkeit* des Eintretens des Ereignisses A in den ersten n Experimenten, und es gilt

$$P\text{-}\lim_{n \rightarrow \infty} H_n = p.$$

Lemma 46 (Borel-Cantelli): Es sei $\{A_n\}_{n \in \mathbb{N}}$ eine Folge von Ereignissen in einem Wahrscheinlichkeitsraum (Ω, \mathcal{A}, P) . Dann gilt:

a) $\sum_{n=1}^{\infty} P(A_n) < \infty \Rightarrow P\left(\limsup_{n \rightarrow \infty} A_n\right) = 0.$

b) Ist die Folge $\{A_n\}_{n \in \mathbb{N}}$ zusätzlich stochastisch unabhängig, so gilt

$$\sum_{n=1}^{\infty} P(A_n) = \infty \Rightarrow P\left(\limsup_{n \rightarrow \infty} A_n\right) = 1.$$

Satz 38 (starkes Gesetz der Großen Zahlen; Etemadi 1981): Es sei $\{X_n\}_{n \in \mathbb{N}}$ eine Folge paarweise unkorrelierter quadratisch integrierbarer Zufallsvariablen auf einem Wahrscheinlichkeitsraum (Ω, \mathcal{A}, P) mit

$$E(X_n) = \mu \in \mathbb{R} \text{ für alle } n \in \mathbb{N} \text{ und } V := \sup_{n \in \mathbb{N}} \{\text{Var}(X_n)\} < \infty.$$

Dann gilt für die arithmetischen Mittel $\bar{X}_n := \frac{1}{n} \sum_{k=1}^n X_k$ dieser Folge:

$$\lim_{n \rightarrow \infty} \bar{X}_n = \mu \text{ fast sicher.}$$

Lemma 47: Unter den Bedingungen von Lemma 45 gilt auch:

$$\lim_{n \rightarrow \infty} H_n = p \text{ fast sicher.}$$

Satz 39 (Gesetz vom iterierten Logarithmus): Es sei $\{X_n\}_{n \in \mathbb{N}}$ eine Folge unabhängiger, identisch verteilter, quadratisch integrierbarer Zufallsvariablen auf einem Wahrscheinlichkeitsraum (Ω, \mathcal{A}, P) mit

$$E(X_n) = \mu \in \mathbb{R} \text{ für alle } n \in \mathbb{N} \text{ und } \text{Var}(X_n) = \sigma^2 > 0.$$

Dann gilt, mit $\sigma = \sqrt{\sigma^2} > 0$:

$$\limsup_{n \rightarrow \infty} \frac{\sum_{k=1}^n (X_k - \mu)}{\sqrt{2n \ln \ln n}} = \sigma, \quad \liminf_{n \rightarrow \infty} \frac{\sum_{k=1}^n (X_k - \mu)}{\sqrt{2n \ln \ln n}} = -\sigma \text{ fast sicher.}$$

II. Grundprinzipien stochastischer Modellierung

II.7. Der Zentrale Grenzwertsatz

Definition 40 (schwache Konvergenz): Es seien $\{Q_n\}_{n \in \mathbb{N}}$ und Q Verteilungen auf der Borel'schen σ -Algebra \mathcal{B}^1 . Die Folge $\{Q_n\}_{n \in \mathbb{N}}$ heißt *schwach* (oder auch *verteilungs-)konvergent* gegen Q , wenn gilt:

$$\lim_{n \rightarrow \infty} \int g dQ_n = \int g dQ \text{ für alle } g \in CB(\mathbb{R}),$$

wobei $CB(\mathbb{R})$ die Menge der reellen, auf \mathbb{R} stetigen und beschränkten Funktionen bezeichne; in Zeichen: $Q = w\text{-}\lim Q_n$ oder $Q_n \xrightarrow{w} Q$ ($n \rightarrow \infty$).

Lemma 48: Es seien $\{Q_n\}_{n \in \mathbb{N}}$ und Q Verteilungen auf der Borel'schen σ -Algebra \mathcal{B}^1 . Die Folge $\{Q_n\}_{n \in \mathbb{N}}$ ist genau dann schwach konvergent gegen Q , wenn gilt:

$$\lim_{n \rightarrow \infty} \int g dQ_n = \int g dQ \text{ für alle } g \in UCB(\mathbb{R}),$$

wobei $UCB(\mathbb{R})$ die Menge der reellen, auf \mathbb{R} *gleichmäßig* stetigen und beschränkten Funktionen bezeichne (aus dem Englischen: **u**niformly **b**ounded and **c**ontinuous).

Satz 40: Es sei $\{X, X_n\}_{n \in \mathbb{N}}$ eine Folge reeller Zufallsvariablen auf einem Wahrscheinlichkeitsraum (Ω, \mathcal{A}, P) . Dann gelten folgende Aussagen:

a) $P\text{-}\lim_{n \rightarrow \infty} X_n = X \Rightarrow w\text{-}\lim_{n \rightarrow \infty} X_n = X,$

d.h. *stochastische* impliziert *schwache* Konvergenz.

b) Ist X fast sicher konstant, gilt hiervon auch die Umkehrung:

$$w\text{-}\lim_{n \rightarrow \infty} X_n = X \Rightarrow P\text{-}\lim_{n \rightarrow \infty} X_n = X.$$

Satz 41. Es seien $\{Q_n\}_{n \in \mathbb{N}}$ und Q Verteilungen auf der Borel'schen σ -Algebra \mathcal{B}^1 . Die zugehörigen Verteilungsfunktionen seien mit F_n , $n \in \mathbb{N}$ bzw. F bezeichnet; $C(F)$ bzw. $C(F^{-1})$ mögen die Mengen der reellen Stetigkeitspunkte von F bzw. der im Intervall $(0,1)$ gelegenen Stetigkeitspunkte der Pseudoinversen F^{-1} bezeichnen. Die folgenden drei Aussagen sind dann äquivalent:

- a) $Q = w\text{-}\lim Q_n$
- b) $F(x) = \lim_{n \rightarrow \infty} F_n(x)$ für alle $x \in C(F)$
- c) $F^{-1}(u) = \lim_{n \rightarrow \infty} F_n^{-1}(u)$ für alle $u \in C(F^{-1})$

Satz 42: Es sei $\{X, X_n\}_{n \in \mathbb{N}}$ eine Folge reeller Zufallsvariablen auf einem Wahrscheinlichkeitsraum (Ω, \mathcal{A}, P) mit zugehörigen momenterzeugenden Funktionen $\{\psi_X, \psi_{X_n}\}_{n \in \mathbb{N}}$, die sämtlich in einem Intervall $[-\delta, \delta]$ für ein $\delta > 0$ existieren mögen. Dann gilt:

$$\lim_{n \rightarrow \infty} \psi_{X_n}(t) = \psi_X(t) \text{ für alle } |t| \leq \delta \quad \Rightarrow \quad P^X = w\text{-}\lim P^{X_n},$$

d.h. die Konvergenz der erzeugenden Funktionen impliziert die schwache Konvergenz der Verteilungen, und es gilt ferner

$$\lim_{n \rightarrow \infty} E(X_n^k) = E(X^k) \text{ für alle } k \in \mathbb{N},$$

d.h. die Konvergenz sämtlicher Momente.

Satz 43 (Zentraler Grenzwertsatz): Es sei $\{X_n\}_{n \in \mathbb{N}}$ eine Folge stochastisch unabhängiger, identisch verteilter, quadratisch integrierbarer Zufallsvariablen auf einem Wahrscheinlichkeitsraum (Ω, \mathcal{A}, P) mit

$$E(X_n) = \mu \in \mathbb{R} \quad \text{und} \quad \text{Var}(X_n) = \sigma^2 > 0 \quad \text{für alle } n \in \mathbb{N}.$$

Die (normalisierten) Zufallsvariablen $\{T_n\}_{n \in \mathbb{N}}$ seien definiert durch

$$T_n := \frac{1}{\sigma\sqrt{n}} \sum_{k=1}^n (X_k - \mu) \quad \text{für alle } n \in \mathbb{N}.$$

Dann gilt:

$$\mathcal{N}(0, 1) = w\text{-}\lim P^{T_n}, \quad \text{d.h. insbesondere} \quad \lim_{n \rightarrow \infty} P \left(\frac{1}{\sigma\sqrt{n}} \sum_{k=1}^n (X_k - \mu) \leq x \right) = \Phi(x) \quad \text{für alle } x \in \mathbb{R},$$

unabhängig von der den $\{X_n\}_{n \in \mathbb{N}}$ zu Grunde liegenden Verteilung.

Satz 44 (Zentraler Grenzwertsatz von Lindeberg): Es sei $\{X_n\}_{n \in \mathbb{N}}$ eine Folge stochastisch unabhängiger quadratisch integrierbarer Zufallsvariablen auf einem Wahrscheinlichkeitsraum (Ω, \mathcal{A}, P) mit

$$E(X_n) = \mu \in \mathbb{R} \quad \text{und} \quad \text{Var}(X_n) = \sigma_n^2 > 0 \quad \text{für alle } n \in \mathbb{N}.$$

Die (normalisierten) Zufallsvariablen $\{T_n\}_{n \in \mathbb{N}}$ seien definiert durch

$$T_n := \frac{1}{\sigma^{[n]}} \sum_{k=1}^n (X_k - \mu) \quad \text{für alle } n \in \mathbb{N}$$

mit

$$\sigma^{[n]} = \sqrt{\text{Var}\left(\sum_{k=1}^n X_k\right)} = \sqrt{\sum_{k=1}^n \sigma_k^2}.$$

Dann gilt $\mathcal{N}(0,1) = w\text{-}\lim P^{T_n}$, wenn die so genannte *Lindeberg-Bedingung* erfüllt ist:

$$\lim_{n \rightarrow \infty} \frac{1}{(\sigma^{[n]})^2} \sum_{k=1}^n \int_{\{|x-\mu| \geq \varepsilon \sigma^{[n]}\}} (x-\mu)^2 P^{X_k}(dx) = 0 \quad \text{für ein } \varepsilon > 0.$$

Bemerkung: Die Lindeberg-Bedingung ist in folgenden Situationen erfüllt (und damit der Zentrale Grenzwertsatz gültig):

a) Die Folge $\{X_n\}_{n \in \mathbb{N}}$ ist *identisch verteilt*. (Situation des Satzes 43).

b) Die Folge $\{X_n\}_{n \in \mathbb{N}}$ erfüllt die so genannte *Ljapunoff-Bedingung*:

$$\lim_{n \rightarrow \infty} \frac{1}{(\sigma^{[n]})^{2+\delta}} \sum_{k=1}^n E(|X_k - \mu|^{2+\delta}) = 0.$$

c) Die Folge $\{X_n\}_{n \in \mathbb{N}}$ ist gleichmäßig beschränkt und es gilt $\lim_{n \rightarrow \infty} \sigma^{[n]} = \infty$.

II. Grundprinzipien stochastischer Modellierung

II.8. Der Satz von Glivenko-Cantelli

Empirische Verteilungsfunktion:

Sind X_1, \dots, X_n ($n \in \mathbb{N}$) unabhängige und identisch verteilte Zufallsvariablen mit der (theoretischen) Verteilungsfunktion F , so ist die empirische Verteilungsfunktion \hat{F}_n definiert durch

$$\hat{F}_n(x) := \frac{1}{n} \sum_{k=1}^n \mathbb{1}_{(-\infty, x]}(X_k) \text{ für alle } x \in \mathbb{R}.$$

Interpretation:

Die empirische Verteilungsfunktion \hat{F}_n ist die Verteilungsfunktion einer (zufälligen) diskreten Gleichverteilung auf den „Punkten“ X_1, \dots, X_n .

Satz von *Glivenko-Cantelli* (Zentralsatz der Statistik):

Die empirische Verteilungsfunktion \hat{F}_n konvergiert gleichmäßig gegen die „wahre“ Verteilungsfunktion F für $n \rightarrow \infty$, d.h. es gilt

$$\limsup_{n \rightarrow \infty} \sup_{x \in \mathbb{R}} |\hat{F}_n(x) - F(x)| = 0.$$

Bemerkung:

Die schwächere (punktweise) Aussage $\lim_{n \rightarrow \infty} |\hat{F}_n(x) - F(x)| = 0$ für alle $x \in \mathbb{R}$ ist eine direkte Konsequenz des Gesetzes der großen Zahlen, angewendet auf die Folge $B(1, F(x))$ -binomialverteilter Zufallsvariablen $Y_n(x) := \mathbb{1}_{(-\infty, x]}(X_n)$ für $n \in \mathbb{N}$, $x \in \mathbb{R}$.

II. Grundprinzipien stochastischer Modellierung

II.9. Bedingte Verteilungen

Definition 41 (allgemeine bedingte Wahrscheinlichkeit): Es sei (Ω, \mathcal{A}, P) ein Wahrscheinlichkeitsraum und $\mathcal{C} \subseteq \mathcal{A}$ eine beliebige Teil- σ -Algebra. Jede \mathcal{C} -messbare Zufallsvariable $P(A|\mathcal{C})$ mit $A \in \mathcal{A}$ und den folgenden Eigenschaften

- $P(A|\mathcal{C})(\omega) \in [0,1]$ für alle $\omega \in \Omega$
- $P(\emptyset|\mathcal{C}) = 0, P(\Omega|\mathcal{C}) = 1$
- $P(A \cap B) = \int_B P(A|\mathcal{C}) dP$ für alle $B \in \mathcal{C}$

heißt *bedingte Wahrscheinlichkeit von A unter (der σ -Algebra) \mathcal{C}* .

Definition 42 (reguläre bedingte Verteilungen): Gilt unter den Bedingungen von Definition 42 zusätzlich für alle paarweise disjunkten Folgen $\{A_n\}_{n \in \mathbb{N}} \subseteq \mathcal{A}$:

- $$P\left(\bigcup_{n=1}^{\infty} A_n \mid \mathcal{C}\right) = \sum_{n=1}^{\infty} P(A_n \mid \mathcal{C}) \quad P_{\mathcal{C}}\text{-fast sicher,}$$

so heißt das durch $P(\cdot \mid \mathcal{C})$ $P_{\mathcal{C}}$ -fast sicher bestimmte (bedingte) Wahrscheinlichkeitsmaß auf \mathcal{A} *regulär*.

Satz 45 (Faktorisierungssatz II). Es sei (Ω, \mathcal{A}, P) ein Wahrscheinlichkeitsraum und Y ein Zufallselement mit Werten in einem Messraum $(\mathcal{X}, \mathcal{B})$. \mathcal{C} bezeichne die von Y erzeugte Teil- σ -Algebra von \mathcal{A} . Eine reellwertige Zufallsvariable X auf (Ω, \mathcal{A}, P) ist genau dann \mathcal{C} -messbar, wenn eine messbare Abbildung $h: (\mathcal{X}, \mathcal{B}) \rightarrow (\mathbb{R}^1, \mathcal{B}^1)$ existiert mit $X = h \circ Y$.

Satz 46 (bedingte Verteilungen und Dichten). Es sei \mathbf{X} ein n -dimensionaler und \mathbf{Y} ein m -dimensionaler Zufallsvektor ($n, m \in \mathbb{N}$) auf einem Wahrscheinlichkeitsraum (Ω, \mathcal{A}, P) derart, dass $P^{(\mathbf{X}, \mathbf{Y})}$ eine $(n+m)$ -dimensionale Dichte $f_{(\mathbf{X}, \mathbf{Y})}$ bezüglich eines geeigneten Produkt-Maßes $\mu = \mu_n \otimes \nu_m$ auf der Borel'schen σ -Algebra $\mathcal{B}^{n+m} = \mathcal{B}^n \otimes \mathcal{B}^m$ besitze. In diesem Fall existiert eine reguläre Version der bedingten Verteilung $P^{\mathbf{X}}(\cdot | \mathbf{Y} = \mathbf{y})$; eine solche ist etwa gegeben durch die zugehörige bedingte μ_n -Dichte

$$f_{\mathbf{X}}(\mathbf{x} | \mathbf{Y} = \mathbf{y}) := \begin{cases} \frac{f_{(\mathbf{X}, \mathbf{Y})}(\mathbf{x}, \mathbf{y})}{f_{\mathbf{Y}}(\mathbf{y})}, & \text{falls } f_{\mathbf{Y}}(\mathbf{y}) > 0 \\ f_{\mathbf{X}}(\mathbf{x}), & \text{falls } f_{\mathbf{Y}}(\mathbf{y}) = 0, \end{cases} \quad \mathbf{y} \in \mathbb{R}^m.$$

Hierbei bezeichnet $f_{\mathbf{X}}(\mathbf{x})$ die Randdichte von \mathbf{X} (entsprechend für $f_{\mathbf{Y}}(\mathbf{y})$), die gegeben ist durch

$$f_{\mathbf{X}}(\mathbf{x}) = \int_{\mathbb{R}^m} f_{(\mathbf{X}, \mathbf{Y})}(\mathbf{x}, \mathbf{y}) \nu_m(d\mathbf{y}), \quad \mathbf{x} \in \mathbb{R}^n.$$

Satz 47. In der Situation des Satzes 46 sei μ das $(n+m)$ -dimensionale Lebesgue-Maß und zusätzlich $f_{(\mathbf{x}, \mathbf{y})}(\mathbf{x}, \cdot)$ für alle $\mathbf{x} \in \mathbb{R}^n$ und $f_{\mathbf{y}}$ in einem Punkt $\mathbf{y} \in \mathbb{R}^m$ stetig mit $f_{\mathbf{y}}(\mathbf{y}) > 0$. $U_h(\mathbf{y}) := \prod_{i=1}^m [y_i - h, y_i + h]$ bezeichne für $h > 0$ die Würfel-Umgebung von \mathbf{y} mit Kantenlänge $2h$. Dann gilt für jede Menge $A \in \mathcal{B}^n$:

$$P(\mathbf{X} \in A | \mathbf{Y} = \mathbf{y}) = \lim_{h \downarrow 0} P(\mathbf{X} \in A | \mathbf{Y} \in U_h(\mathbf{y})).$$

Satz 48. Es sei (Ω, \mathcal{A}, P) ein Wahrscheinlichkeitsraum; X, Y, Z seien Zufallselemente mit Werten in Messräumen $(\mathfrak{X}, \mathcal{B}), (\mathfrak{Y}, \mathcal{D}), (\mathfrak{Z}, \mathcal{F})$. Dann gilt:

a) Ist U eine reellwertige Zufallsvariable auf $(\mathfrak{Y} \times \mathfrak{Z}, \mathcal{D} \otimes \mathcal{F})$ und $P(\cdot | Z = z), z \in \mathfrak{Z}$ regulär, so gilt im Falle der Integrierbarkeit:

$$\int U(y, z) P^{(Y, Z)}(dy, dz) = \int \int U(y, z) P^Y(dy | Z = z) P^Z(dz)$$

b) Ist $P(\cdot | Z = z), z \in \mathfrak{Z}$ regulär, so gilt

$$P(X \in A, Y \in B | Z = z) = \int_B P(X \in A | Y = y, Z = z) P^Y(dy | Z = z)$$

P^Z -fast sicher für alle $A \in \mathcal{B}, B \in \mathcal{D}, z \in \mathfrak{Z}$.

c) X und Y sind stochastisch unabhängig genau dann, wenn es eine reguläre Version von $P^X(\cdot | Y = y), y \in \mathfrak{Y}$ bzw. $P^Y(\cdot | X = x), x \in \mathfrak{X}$ gibt, die nicht von y bzw. von x abhängt.

Satz 49 (Ersetzungsformel). Es sei (Ω, \mathcal{A}, P) ein Wahrscheinlichkeitsraum; \mathbf{X} und \mathbf{Y} seien Zufallselemente mit Werten in Messräumen $(\mathcal{X}, \mathcal{B})$ bzw. $(\mathcal{Y}, \mathcal{D})$. Die bedingte Verteilung $P^{\mathbf{X}}(\cdot | \mathbf{Y} = \mathbf{y})$, $\mathbf{y} \in \mathcal{Y}$ sei regulär, $G: (\mathcal{X} \times \mathcal{Y}, \mathcal{B} \otimes \mathcal{D}) \rightarrow (\mathcal{Z}, \mathcal{F})$ eine Abbildung mit Werten in dem Messraum $(\mathcal{Z}, \mathcal{F})$. Dann gilt für alle $A \in \mathcal{F}$:

$$P(G(\mathbf{X}, \mathbf{Y}) \in A | \mathbf{Y} = \mathbf{y}) = P(G(\mathbf{X}, \mathbf{y}) \in A | \mathbf{Y} = \mathbf{y}) \\ := P(G(\mathbf{X}, \mathbf{z}) \in A | \mathbf{Y} = \mathbf{y}) \Big|_{\mathbf{z}=\mathbf{y}} \quad P^{\mathbf{Y}}\text{-f.s.}$$

Sind \mathbf{X} und \mathbf{Y} stochastisch unabhängig, so gilt speziell

$$P(G(\mathbf{X}, \mathbf{Y}) \in A | \mathbf{Y} = \mathbf{y}) = P(G(\mathbf{X}, \mathbf{y}) \in A) \quad P^{\mathbf{Y}}\text{-f.s.}$$

Definition 43 (regulärer bedingter Erwartungswert): Es sei (Ω, \mathcal{A}, P) ein Wahrscheinlichkeitsraum und $\mathcal{C} \subseteq \mathcal{A}$ eine Teil- σ -Algebra von \mathcal{A} ; $P(\cdot | \mathcal{C})$ sei regulär. Ist X eine reellwertige Zufallsvariable auf (Ω, \mathcal{A}, P) , so heißt im Falle der Existenz

$$E(X | \mathcal{C})(\omega) := \int X(\eta) P(d\eta | \mathcal{C})(\omega) = \int X dP(\cdot | \mathcal{C})(\omega), \quad \omega \in \Omega,$$

oder kürzer

$$E(X | \mathcal{C}) = \int X dP(\cdot | \mathcal{C}).$$

Lemma 49 (Zerlegungslemma): Es sei (Ω, \mathcal{A}, P) ein Wahrscheinlichkeitsraum und $\mathcal{C} \subseteq \mathcal{A}$ eine Teil- σ -Algebra von \mathcal{A} ; $P(\cdot | \mathcal{C})$ sei regulär. Ist die reellwertige Zufallsvariable X auf (Ω, \mathcal{A}, P) integrierbar, so ist sie auch bezüglich $P(\cdot | \mathcal{C})$ integrierbar, und es gilt

$$E(X) = E(E(X | \mathcal{C})) = \int E(X | \mathcal{C}) dP.$$

Ist ferner X sogar quadratisch integrierbar, so auch bezüglich $P(\cdot | \mathcal{C})$, und es gilt

$$\text{Var}(X) = E(\text{Var}(X | \mathcal{C})) + \text{Var}(E(X | \mathcal{C})).$$

Satz 50 (Eigenschaften [regulärer] bedingter Erwartungen): Es sei (Ω, \mathcal{A}, P) ein Wahrscheinlichkeitsraum und $\mathcal{C} \subseteq \mathcal{A}$ eine Teil- σ -Algebra von \mathcal{A} ; die reellwertige Zufallsvariable X auf (Ω, \mathcal{A}, P) sei integrierbar. Dann gilt:

- a) Ist Y eine weitere reellwertige Zufallsvariable auf (Ω, \mathcal{A}, P) , die \mathcal{C} -messbar ist und für die das Produkt $X \cdot Y$ integrierbar ist, dann gilt:

$$E(X \cdot Y | \mathcal{C}) = Y \cdot E(X | \mathcal{C}) \quad (\text{Faktor-Regel})$$

Insbesondere gilt im Fall $X \equiv 1$:

$$E(Y | \mathcal{C}) = Y$$

- b) Ist $\mathcal{D} \subseteq \mathcal{C}$ eine weitere Teil- σ -Algebra von \mathcal{A} , so gilt:

$$E[E(X | \mathcal{C}) | \mathcal{D}] = E[E(X | \mathcal{D}) | \mathcal{C}] = E(X | \mathcal{D}) \quad (\text{Turm-Eigenschaft})$$

- c) Sind die σ -Algebren $\sigma(X)$ und \mathcal{C} stochastisch unabhängig (im Sinne von Definition 26), so gilt

$$E(X | \mathcal{C}) = E(X) \quad (\text{Unabhängigkeits-Eigenschaft})$$

Insbesondere gilt im Fall $\mathcal{C} = \{\emptyset, \Omega\}$: $E(X | \mathcal{C}) = E(X)$.

II. Grundprinzipien stochastischer Modellierung

II.10. Die mehrdimensionale Normalverteilung

Definition 44 (mehrdimensionale Normalverteilung): Ein d -dimensionaler Zufallsvektor $X = (X_1, \dots, X_d)^{tr}$ mit Werten in $(\mathbb{R}^d, \mathcal{B}^d)$ besitzt eine d -dimensionale (nicht-entartete) Normalverteilung $\mathcal{N}(\boldsymbol{\mu}, \Sigma)$, wenn er bezüglich des Lebesgue-Maßes m^d die Dichte

$$f(x_1, \dots, x_d) = \frac{1}{\sqrt{(2\pi)^d \det \Sigma}} \exp\left(-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^{tr} \Sigma^{-1} (\mathbf{x} - \boldsymbol{\mu})\right)$$

mit $\mathbf{x} = (x_1, \dots, x_d)^{tr} \in \mathbb{R}^d$, $\boldsymbol{\mu} = (\mu_1, \dots, \mu_d)^{tr} \in \mathbb{R}^d$ besitzt, wobei Σ eine symmetrische positiv-definite $d \times d$ -Matrix ist.

Wegen der positiven Definitheit von Σ existiert (mindestens) eine quadratische invertierbare Matrix A mit $\Sigma = A \cdot A^{tr}$.

Spektralzerlegung der Matrix Σ : voraussetzungsgemäß besitzt Σ d (ggf. mit Vielfachheiten vorkommende) positive *Eigenwerte* $\lambda_1, \dots, \lambda_d$; ferner existiert eine $d \times d$ -Matrix T aus orthonormalen *Eigenvektoren* mit der Eigenschaft

$$\Sigma = T \Delta T^{tr} = T \Delta T^{-1} \text{ mit } \Delta = \begin{bmatrix} \lambda_1 & 0 & 0 & \dots & 0 \\ 0 & \lambda_2 & 0 & \dots & 0 \\ 0 & 0 & \ddots & \dots & \vdots \\ \vdots & \vdots & & \lambda_{d-1} & 0 \\ 0 & 0 & \dots & 0 & \lambda_d \end{bmatrix}.$$

Hier ist $A = T \Delta^{1/2}$ eine mögliche Wahl mit $\Delta^{1/2} = \begin{bmatrix} \sqrt{\lambda_1} & 0 & 0 & \dots & 0 \\ 0 & \sqrt{\lambda_2} & 0 & \dots & 0 \\ 0 & 0 & \ddots & \dots & \vdots \\ \vdots & \vdots & & \sqrt{\lambda_{d-1}} & 0 \\ 0 & 0 & \dots & 0 & \sqrt{\lambda_d} \end{bmatrix}.$

Cholesky-Zerlegung von Σ ; Die gesuchte Matrix A wird dabei als *untere Dreiecksmatrix* angenommen:

$$A = \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{d1} & a_{d2} & \cdots & a_{dd} \end{bmatrix}.$$

Hiermit ergibt sich

$$\Sigma = [\sigma_{ij}] = AA^{tr} = \begin{bmatrix} a_{11}^2 & a_{11}a_{21} & \cdots & a_{11}a_{d1} \\ a_{21}a_{11} & a_{21}^2 + a_{22}^2 & \cdots & a_{21}a_{d1} + a_{22}a_{d2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{d1}a_{11} & a_{d1}a_{21} + a_{d2}a_{22} & \cdots & \sum_{k=1}^d a_{dk}^2 \end{bmatrix}.$$

Diese Gleichung kann rekursiv aufgelöst werden zu

$$a_{11} = \sqrt{\sigma_{11}}, \quad a_{kk} = \sqrt{\sigma_{kk} - \sum_{i=1}^{k-1} a_{ki}^2}, \quad a_{k1} = \frac{\sigma_{k1}}{a_{11}}, \quad a_{kj} = \frac{\sigma_{kj} - \sum_{i=1}^{j-1} a_{ki}a_{ji}}{a_{jj}}, \quad 1 \leq k, j \leq d.$$

Satz 51 (Eigenschaften der multivariaten Normalverteilung I): Es sei \mathbf{X} ein $\mathcal{N}(\boldsymbol{\mu}, \Sigma)$ -verteilter Zufallsvektor. Dann ist der Zufallsvektor $\mathbf{Y} = \mathbf{X} - \boldsymbol{\mu} \sim \mathcal{N}(\mathbf{0}, \Sigma)$ -verteilt. Ist umgekehrt $\mathbf{X} \sim \mathcal{N}(\mathbf{0}, \Sigma)$ -verteilt, dann ist $\mathbf{Y} = \mathbf{X} + \boldsymbol{\mu} \sim \mathcal{N}(\boldsymbol{\mu}, \Sigma)$ -verteilt.

Satz 52 (Eigenschaften der multivariaten Normalverteilung II): Es sei \mathbf{X} ein d -dimensionaler $\mathcal{N}(\boldsymbol{\mu}, \Sigma)$ -verteilter Zufallsvektor und B eine beliebige $m \times d$ -Matrix mit $m < d$ und vollem Rang m . Dann ist der m -dimensionale Zufallsvektor $\mathbf{Y} = B\mathbf{X}$ ebenfalls multivariat normalverteilt mit $P^{\mathbf{Y}} = \mathcal{N}(B\boldsymbol{\mu}, B\Sigma B^{\text{tr}})$. Im speziellen Fall $m=1$ bedeutet dies, dass jede Linearkombination multivariat normalverteilter Zufallsvariablen wieder (univariat) normalverteilt ist. Insbesondere ist jede Komponente X_k von \mathbf{X} selbst univariat normalverteilt mit Erwartungswert μ_k und Varianz σ_k^2 für $k = 1, \dots, d$.

Lemma 50: Für stochastisch unabhängige multivariat normalverteilte Zufallsvektoren \mathbf{X} und \mathbf{Y} derselben Dimension d gilt folgende Rechenregel:

$$P^{\mathbf{X}} = \mathcal{N}(\boldsymbol{\mu}_{\mathbf{X}}, \Sigma_{\mathbf{X}}) \text{ und } P^{\mathbf{Y}} = \mathcal{N}(\boldsymbol{\mu}_{\mathbf{Y}}, \Sigma_{\mathbf{Y}}) \Rightarrow P^{\mathbf{X}+\mathbf{Y}} = \mathcal{N}(\boldsymbol{\mu}_{\mathbf{X}} + \boldsymbol{\mu}_{\mathbf{Y}}, \Sigma_{\mathbf{X}} + \Sigma_{\mathbf{Y}}).$$