

Wahrscheinlichkeitstheorie und Statistik - Zusammenfassung

Julian Shen

28. Dezember 2021

1 Grundbegriffe

Definition: Ergebnisse und Ereignisse

- **Grundraum** ist eine nicht leere Menge $\Omega \neq \emptyset$ und enthält alle möglichen **Ergebnisse** eines Zufallsexperiments
- **Ereignisse** sind Teilmengen $A \subseteq \Omega$, denen eine Wahrscheinlichkeit zugeordnet werden kann. Falls ein ω Ergebnis ist, dann heißt $\{\omega\}$ **Elementarereignis**

Ereignisse können durch Mengenoperationen logisch verknüpft werden:

- $A \cup B$: Ereignis A oder B tritt ein („inklusives oder“)
- $A \cap B$: Ereignis A und B treffen ein
- $A \setminus B$: Ereignis A tritt ein, aber Ereignis B trifft nicht ein
- B^c : Ereignis B trifft nicht ein
- $A \subseteq B$: Wenn A eintritt, dann tritt auch B ein

Jedem Ereignis kann durch die **relative Häufigkeit** eine Wahrscheinlichkeit zugeordnet werden. Für n Wiederholungen und Ergebnisse $\omega_1, \dots, \omega_n \in \Omega$ gilt:

$$\mathbb{P}_n(A) := \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{\{\omega_i \in A\}}$$

Definition: Diskretes Wahrscheinlichkeitsmaß und Wahrscheinlichkeitsraum

Eine Abbildung $\mathbb{P} : \mathcal{P}(\Omega) \rightarrow [0, 1]$ heißt **diskretes Wahrscheinlichkeitsmaß**, falls

- $\mathbb{P}(\Omega) = 1$
- $\forall A_n \subseteq \Omega, n \in \mathbb{N}, \text{ disjunkt: } \mathbb{P}\left(\bigcup_{n \in \mathbb{N}} A_n\right) = \sum_{i=1}^{\infty} \mathbb{P}(A_n)$ (σ -Additivität)
- es existiert eine abzählbare Menge $\Omega_0 \subseteq \Omega$ mit $\mathbb{P}(\Omega_0) = 1$

Dann heißt (Ω, \mathbb{P}) **diskreter Wahrscheinlichkeitsraum**.

Rechenregeln für diskrete Wahrscheinlichkeitsräume

- $\mathbb{P}(\emptyset) = 0$
- $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$
- $\mathbb{P}(B \setminus A) = \mathbb{P}(B) - \mathbb{P}(A)$
- $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$

Definition: Bernoulliverteilung

Wahrscheinlichkeitsverteilung heißt Bernoulliverteilung Ber_p mit Erfolgswahrscheinlichkeit p , wenn:

- Grundraum $\Omega = \{0, 1\}$
- $\mathbb{P}(1) = p$ für ein $p \in [0, 1]$

Es gilt $\mathbb{P}(\{0\}) = 1 - \mathbb{P}(\{1\}) = 1 - p$

Definition: Gleichverteilung

Das Wahrscheinlichkeitsmaß (Ω, \mathbb{P}) heißt Gleichverteilung oder **Laplace-Verteilung** U_A auf Ω , falls

- $\Omega \neq \emptyset$ endlich
- $\mathbb{P}(A) = \frac{|A|}{|\Omega|}$, für $A \subseteq \Omega$

Urnenmodelle/Fächermodelle

Urnenmodell mit n Kugeln und k Ziehungen	mit Zurücklegen	ohne Zurücklegen	
mit Reihenfolge	n^k	$\frac{n!}{(n-k)!}$	unterscheidbare Murmeln
ohne Reihenfolge	$\binom{n+k-1}{k}$	$\binom{n}{k}$	ununterscheidbare Murmeln
	mit Mehrfachbelegung	ohne Mehrfachbelegung	Verteilung von k Murmeln auf n Fächer

Urnenmodelle ermöglichen es, die Wahrscheinlichkeiten zu bestimmen, falls von einer Gleichverteilung ausgegangen werden kann!

Definition: Zähldichte

Sei (Ω, \mathbb{P}) ein diskreter Wahrscheinlichkeitsraum. Dann wird die Funktion

$$f : \Omega \rightarrow [0, 1], f(\omega) = \mathbb{P}(\{\omega\})$$

Wahrscheinlichkeitsfunktion oder **Zähldichte** von \mathbb{P} genannt.

Diese besitzt folgende Eigenschaften:

- $\Omega_T := \{\omega \in \Omega \mid f(\omega) > 0\}$ ist abzählbar und heißt **Träger** von \mathbb{P} bzw. von f
- $\sum_{\omega \in \Omega} f(\omega) = 1$

Die Zähldichte ist eindeutig!

Definition: Binomialverteilung

Das Wahrscheinlichkeitsmaß $\mathbb{P} = \text{Bin}_{(n,p)}$ auf $\{0, \dots, n\}$ mit der Zähldichte

$$f(k) = \binom{n}{k} p^k \cdot (1-p)^{n-k} \quad \forall k \in \{0, \dots, n\}$$

heißt **Binomialverteilung** mit Parametern $n \in \mathbb{N}$ und $p \in [0, 1]$.

Definition: Geometrische Verteilung

Das Wahrscheinlichkeitsmaß $\mathbb{P} = \text{Geo}_p$ auf \mathbb{N}_0 mit der Zähldichte

$$f(k) = (1-p)^k \cdot p \quad \forall k \in \mathbb{N}_0$$

heißt **geometrische Verteilung** mit Parameter $p \in (0, 1]$.

2 Bedingte Wahrscheinlichkeit und Unabhängigkeit

Definition: Bedingte Wahrscheinlichkeit

Für (Ω, \mathbb{P}) diskreter Wahrscheinlichkeitsraum und $A, B \in \Omega$ mit $\mathbb{P}(B) > 0$ heißt

$$\mathbb{P}(A \mid B) := \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$$

die **bedingte Wahrscheinlichkeit** von A gegeben B.

Multiplikationsformel

Seien $A_1, \dots, A_n \subseteq \Omega$ Ereignisse mit $\mathbb{P}(A_1 \cap \dots \cap A_{n-1}) > 0$, dann gilt

$$\mathbb{P}(A_1 \cap \dots \cap A_{n-1}) = \mathbb{P}(A_1) \cdot \mathbb{P}(A_2 \mid A_1) \cdot \mathbb{P}(A_3 \mid A_1 \cap A_2) \cdots \mathbb{P}(A_n \mid A_1 \cap \dots \cap A_{n-1})$$

Im Fall von $n=2$ gilt: $\mathbb{P}(A \cap B) = \mathbb{P}(B) \cdot \mathbb{P}(A \mid B)$

Satz von der totalen Wahrscheinlichkeit und Satz von Bayes

Sei (Ω, \mathbb{P}) ein diskreter Wahrscheinlichkeitsraum, I eine abzählbare Indexmenge, $B_i \subseteq \Omega, i \in I$, disjunkt mit $\mathbb{P}(B_i) > 0$ und $\bigcup_{i \in I} B_i = \Omega$ und $A \subseteq \Omega$ beliebig.

- Es gilt der **Satz von der totalen Wahrscheinlichkeit**:

$$\mathbb{P}(A) = \sum_{i \in I} \mathbb{P}(A \mid B_i) \cdot \mathbb{P}(B_i)$$

- Falls $\mathbb{P}(A) > 0$ und $k \in I$, dann gilt der **Satz von Bayes**:

$$\mathbb{P}(B_k \mid A) = \frac{\mathbb{P}(A \mid B_k) \cdot \mathbb{P}(B_k)}{\mathbb{P}(A)} = \frac{\mathbb{P}(A \mid B_k) \cdot \mathbb{P}(B_k)}{\sum_{i \in I} \mathbb{P}(A \mid B_i) \cdot \mathbb{P}(B_i)}$$

Definition: Stochastische Unabhängigkeit

Sei (Ω, \mathbb{P}) ein diskreter Wahrscheinlichkeitsraum. Zwei Ereignisse $A, B \subseteq \Omega$ heißen **stochastisch unabhängig**, falls

$$\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B)$$

Ereignisse $A_1, \dots, A_n \subseteq \Omega$ heißen **stochastisch unabhängig**, wenn für jede Indexmenge $I \subseteq \{1, \dots, n\}, I \neq \emptyset$, gilt

$$\mathbb{P}\left(\bigcap_{i \in I} A_i\right) = \prod_{i \in I} \mathbb{P}(A_i)$$

Achtung: Mehr als zwei Ereignisse A_1, \dots, A_n sind im Allgemeinen nicht stochastisch unabhängig, wenn nur $\mathbb{P}\left(\bigcap_{i=1}^n A_i\right) = \prod_{i=1}^n \mathbb{P}(A_i)$ gilt! Gleiches gilt, wenn jeweils nur zwei der Ereignisse stochastisch unabhängig sind.

3 Zufallsvariablen und ihre Verteilungen

Definition: S -wertige Zufallsvariable

Ist (Ω, \mathbb{P}) ein diskreter Wahrscheinlichkeitsraum und $S \neq \emptyset$ eine beliebige Menge, so wird die Abbildung $X : \Omega \rightarrow S$ **S -wertige Zufallsvariable** genannt.

Definition: Verteilung

Ist $X : \Omega \rightarrow S$ eine Zufallsvariable auf einem diskreten Wahrscheinlichkeitsraum (Ω, \mathbb{P}) , dann wird durch

$$\mathbb{P}^X(B) := \mathbb{P}(X^{-1}(B)) \quad \forall B \subseteq S$$

ein Wahrscheinlichkeitsmaß \mathbb{P}^X auf S definiert, welches Verteilung von X genannt wird. (S, \mathbb{P}^X) ist ein diskreter Wahrscheinlichkeitsraum. Notation für Urbilder:

- $\{X \in B\} := \{\omega \in \Omega \mid X(\omega) \in B\} = X^{-1}(B)$
- $\{X = x\} := \{\omega \in \Omega \mid X(\omega) = x\} = X^{-1}(\{x\})$
- $\{X > x\} := \{\omega \in \Omega \mid X(\omega) > x\} = X^{-1}((x, \infty))$

Zudem schreibt man $\mathbb{P}(X \in B) := \mathbb{P}(\{X \in B\})$

Definition: Stochastische Unabhängigkeit von Zufallsvariablen

Sei (Ω, \mathbb{P}) ein diskreter Wahrscheinlichkeitsraum und $S_i, i \in \{1, \dots, n\}$ nichtleere Zufallsvariablen. Zufallsvariablen $X_i : \Omega \rightarrow S_i, i \in \{1, \dots, n\}$ heißen **stochastisch unabhängig**, wenn für beliebige $B_i \subseteq S_i$ die Ereignisse $\{X_1 \in B_1\}, \dots, \{X_n \in B_n\}$ stochastisch unabhängig sind.

Auch der Vektor $(X_1, \dots, X_n) : \Omega \rightarrow S_1 \times \dots \times S_n$ ist eine Zufallsvariable mit Verteilung $\mathbb{P}^{(X_1, \dots, X_n)}$ auf $S_1 \times \dots \times S_n$.

Bemerkung zur Schreibweise: $(X_1 \in B_1, X_2 \in B_2) = (X_1 \in B_1) \cap (X_2 \in B_2)$

Satz für stochastisch unabhängige Zufallsvariablen

Es sind äquivalent:

- X_1, \dots, X_n sind stochastisch unabhängig
- $\forall B_i \subseteq S_i : \mathbb{P}(X_i \in B_i \mid \forall 1 \leq i \leq n) = \prod_{i=1}^n \mathbb{P}(X_i \in B_i)$
- Bezeichne mit f_{X_i} die Zähldichte von \mathbb{P}^{X_i} auf S_i . Dann hat die Zähldichte $f_{(X_1, \dots, X_n)}$ von $\mathbb{P}^{(X_1, \dots, X_n)}$ die Form: $f_{(X_1, \dots, X_n)}(t_1, \dots, t_n) = \prod_{i=1}^n f_{X_i}(t_i) \quad \forall t_i \in S_i$

Definition: Hypergeometrischen Verteilung

Das Wahrscheinlichkeitsmaß $\mathbb{P} = \text{Hyp}_{(N,M,n)}$ auf \mathbb{N}_0 gegeben durch die Zähldichte

$$\mathbb{P}^X(\{m\}) = \mathbb{P}(X = m) = \frac{\binom{M}{m} \cdot \binom{N-M}{n-m}}{\binom{N}{n}} \mathbb{1}_S(m) \quad \forall m \in \mathbb{N}_0$$

heißt **hypergeometrische Verteilung**.

Zusammenhang hypergeometrische Verteilung und Binomialverteilung

- Die **hypergeometrische Verteilung** $\text{Hyp}_{(N,M,n)}$ beschreibt die Anzahl der markierten Gegenstände bei n -maligem **Ziehen ohne Zurücklegen** aus N Gegenständen, von denen M markiert sind
- Die **Binomialverteilung** $\text{Bin}_{(n,M/N)}$ beschreibt die Anzahl der markierten Gegenstände bei n -maligem **Ziehen mit Zurücklegen** aus N Gegenständen, von denen M markiert sind

Falls $n \ll N$, dann ist Ziehen mit oder ohne Zurücklegen fast identisch und daher $\text{Hyp}_{(N,M,n)}(\{m\}) \approx \text{Bin}_{(n, \frac{M}{N})}(\{m\}) \quad \forall 0 \leq m \leq n$

Poisson'scher Grenzwertsatz

Für eine große Anzahl an Experimenten n und eine kleine Erfolgswahrscheinlichkeit p kann $\text{Bin}_{(n,p)}$ durch eine strukturell einfachere Verteilung approximiert werden:

$$\lim_{n \rightarrow \infty} \text{Bin}_{(n,p)}(\{k\}) = \lim_{n \rightarrow \infty} \binom{n}{k} p^k (1-p)^{n-k} = e^{-\lambda} \cdot \frac{\lambda^k}{k!} =: \text{Poiss}_\lambda(\{k\}) \quad k \in \mathbb{N}_0$$

Das Wahrscheinlichkeitsmaß Poiss_λ heißt Poissonverteilung mit Parameter λ .

Definition: Faltung

Sind X, Y \mathbb{R} -wertige Zufallsvariablen auf einem diskreten Wahrscheinlichkeitsraum mit Zähldichten f_X von \mathbb{P}^X und f_Y von \mathbb{P}^Y , dann heißt

$$(f_X * f_Y)(z) = \sum_{x \in \mathbb{R}: f_X(x) > 0} f_X(x) \cdot f_Y(z - x) \quad \forall z \in \mathbb{R}$$

die **Faltung** von f_X und f_Y . Hierbei ist $f_X * f_Y$ wieder eine Zähldichte mit Träger $\Omega_T := \{z \in \mathbb{R} \mid \exists x, y \in \mathbb{R} : z = x + y, f_X(x) > 0, f_Y(y) > 0\}$ und die zugehörige diskrete Verteilung $\mathbb{P}^X * \mathbb{P}^Y$ heißt **Faltung** von \mathbb{P}^X und \mathbb{P}^Y .

Satz für die Faltung

Sind X, Y unabhängige \mathbb{R} -wertige Zufallsvariablen auf einem diskreten Wahrscheinlichkeitsraum, so gilt

$$\mathbb{P}^X * \mathbb{P}^Y = \mathbb{P}^{X+Y}$$

4 Wahrscheinlichkeitsmaße auf \mathbb{R}

Kontinuierliche Ergebnisse lassen sich nicht mehr durch eine abzählbare Anzahl an Versuchsausgängen beschreiben. Wahrscheinlichkeiten kann man dann nur noch „gutartigen Mengen“ zuordnen, u.a.:

- Intervalle sind gutartig
- Komplemente gutartiger Mengen sind gutartig
- Abzählbare Vereinigungen gutartiger Mengen sind gutartig

Bezeichne nun mit $\mathcal{A} \subseteq \mathcal{P}(\Omega)$ das System aller „gutartigen“ Mengen.

Definition: σ -Algebra

Sei $\Omega \neq \emptyset$ ein beliebiger Grundraum. Eine Menge $\mathcal{A} \subseteq \mathcal{P}(\Omega)$ heißt **σ -Algebra** auf Ω , falls:

- $\emptyset, \Omega \in \mathcal{A}$
- $A \in \mathcal{A} \implies A^c \in \mathcal{A}$
- $A_n \in \mathcal{A} \forall n \in \mathbb{N} \implies \bigcup_{n \in \mathbb{N}} A_n \in \mathcal{A}$

(Ω, \mathcal{A}) heißt dann **messbarer Raum**. Die Mengen $A \in \mathcal{A}$ heißen **Ereignisse**.

Definition: Borel- σ -Algebra

Die **Borel- σ -Algebra** \mathcal{B} auf \mathbb{R} beschreibt die kleinste σ -Algebra, welche alle Intervalle $(a, b]$ für beliebige $a, b \in \mathbb{R}$ enthält.

Definition: Wahrscheinlichkeitsmaß und Wahrscheinlichkeitsraum

Sei (Ω, \mathcal{A}) ein Messraum mit Grundraum $\Omega \neq \emptyset$ und σ -Algebra \mathcal{A} . Eine Abbildung $\mathbb{P} : \mathcal{A} \rightarrow [0, 1]$ heißt **Wahrscheinlichkeitsmaß** auf (Ω, \mathcal{A}) , falls

- $\mathbb{P}(\Omega) = 1$
- $A_n \in \mathcal{A}, n \in \mathbb{N}, \text{ disjunkt} \implies \mathbb{P}\left(\bigcup_{n \in \mathbb{N}} A_n\right) = \sum_{n \in \mathbb{N}} \mathbb{P}(A_n)$ (σ -Additivität)

$(\Omega, \mathcal{A}, \mathbb{P})$ heißt dann **Wahrscheinlichkeitsraum**.

Sätze und Definitionen für allgemeine Wahrscheinlichkeitsräume

Folgende Sätze und Definitionen übertragen sich sinngemäß, wobei als Ereignisse jeweils nur Mengen aus \mathcal{A} betrachtet werden:

- Rechenregeln für diskrete Wahrscheinlichkeitsmaße
- Bedingte Wahrscheinlichkeiten
- Satz von der Totalen Wahrscheinlichkeit und von Bayes
- Stochastische Unabhängigkeit von Ereignissen

Unterschied: Während diskrete Wahrscheinlichkeitsmaße \mathbb{P} vollständig durch die Zähldichte $f(\omega) := \mathbb{P}(\{\omega\}), \omega \in \Omega$ bestimmt sind, ist dies für allgemeine Wahrscheinlichkeitsmaße falsch!

Definition: Verteilungsfunktion

Ist \mathbb{P} ein Wahrscheinlichkeitsmaß auf $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$, so gilt für die durch

$$F : \mathbb{R} \rightarrow [0, 1], F(x) := \mathbb{P}((-\infty, x])$$

definierte **Verteilungsfunktion** von \mathbb{P} :

- F ist monoton steigend
- F ist rechtsseitig stetig
- $F(\infty) := \lim_{x \rightarrow \infty} F(x) = 1, F(-\infty) := \lim_{x \rightarrow -\infty} F(x) = 0$

Ist umgekehrt $F : \mathbb{R} \rightarrow [0, 1]$ eine Funktion, die obige Punkte erfüllt, so existiert genau ein Wahrscheinlichkeitsmaß auf $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$, das F als Verteilungsfunktion besitzt. Für ein diskretes Wahrscheinlichkeitsmaß \mathbb{P} ist die Verteilungsfunktion eine Treppenfunktion.

Definition: Dichten

Sei \mathbb{P} ein Wahrscheinlichkeitsmaß auf $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$. Existiert eine integrierbare Funktion $f : \mathbb{R} \rightarrow [0, \infty)$, sodass

$$F(x) = \mathbb{P}((-\infty, x]) = \int_{-\infty}^x f(t) dt \quad \forall x \in \mathbb{R}$$

so heißt f Dichte von \mathbb{P} bzw. von der zugehörigen Verteilungsfunktion F . Für $A \in \mathcal{B}_{\mathbb{R}}$ gilt dann

$$\mathbb{P}(A) = \int_{\mathbb{R}} f(x) \cdot \mathbf{1}_A(x) dx =: \int_A f(x) dx$$

Umgekehrt ist jede integrierbare Funktion $f : \mathbb{R} \rightarrow [0, \infty)$ mit $\int_{-\infty}^{\infty} f(t)dt = 1$ Dichte eines Wahrscheinlichkeitsmaßes auf $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$, das durch obiges F eindeutig festgelegt ist. Falls eine Dichte existiert, ist F stetig.

Achtung: Dichten dürfen nicht mit Zähldichten verwechselt werden!

Definition: Gleichverteilung

Für $-\infty < a < b < \infty$ heißt das Wahrscheinlichkeitsmaß auf $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ zur Dichte $f := \frac{1}{b-a} \mathbb{1}_{(a,b]}$ **Gleichverteilung** $U_{(a,b]}$ auf $(a, b]$ und für $a \leq c < d \leq b$ gilt:
 $U_{(a,b]}((c, d]) = \frac{d-c}{b-a}$.

Definition: Exponentialverteilung

Die **Exponentialverteilung** Exp_{λ} mit Parameter $\lambda > 0$ ist gegeben durch die Dichte

$$f_{\lambda}(x) := \frac{1}{\lambda} e^{-x/\lambda} \cdot \mathbb{1}_{[0,\infty)}(x), \quad x \in \mathbb{R}$$

Die Verteilungsfunktion ist gegeben durch $F_{\lambda}(x) = 1 - e^{-x/\lambda}$, $\forall x \geq 0$
 und $F_{\lambda}(x) = 0$ für alle $x < 0$.

Exponentialverteilungen beschreiben Lebensdauern von Dingen, die nicht altern, d.h. die Wahrscheinlichkeit noch y Jahre zu überleben, gegeben dass bereits x Jahre überlebt wurden, hängt nicht von x ab.

Definition: Normalverteilung

Die **Normalverteilung** $N_{(\mu, \sigma^2)}$ mit Parametern $\mu \in \mathbb{R}, \sigma > 0$ ist gegeben durch die Dichte

$$f(x) := \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), \quad x \in \mathbb{R}$$

Die Verteilung $N_{(0,1)}$ wird **Standardnormalverteilung** genannt. Die Verteilungsfunktion von $N_{(0,1)}$ bezeichnet man mit

$$\Phi(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt$$

Aufgrund der Symmetrie der Standardnormalverteilung zur y-Achse, gilt $\Phi(-x) = 1 - \Phi(x)$ für alle $x \in \mathbb{R}$. Insbesondere ist $\Phi(0) = \frac{1}{2}$.

Für die Verteilungsfunktion der Normalverteilung $N_{(\mu, \sigma^2)}$ gilt

$$F(x) = \Phi\left(\frac{x-\mu}{\sigma}\right)$$

Um Wahrscheinlichkeiten für eine beliebige Normalverteilung zu berechnen, genügt also die Verteilungsfunktion Φ der Standardnormalverteilung:

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
0.00	0.5000	0.76	0.7764	1.52	0.9357	1.84	0.9671	2.28	0.9887
0.02	0.5080	0.78	0.7823	1.54	0.9382	1.86	0.9686	2.30	0.9893
0.04	0.5160	0.80	0.7881	1.56	0.9406	1.88	0.9699	2.32	0.9898
0.06	0.5239	0.82	0.7939	1.58	0.9429	1.90	0.9713	2.34	0.9904
0.08	0.5319	0.84	0.7995	1.60	0.9452	1.92	0.9726	2.36	0.9909
0.10	0.5398	0.86	0.8051	1.62	0.9474	1.94	0.9738	2.38	0.9913
0.12	0.5478	0.88	0.8106	1.64	0.9495	1.96	0.9750	2.40	0.9918
0.14	0.5557	0.90	0.8159	1.66	0.9515	1.98	0.9761	2.42	0.9922
0.16	0.5636	0.92	0.8212	1.68	0.9535	2.00	0.9772	2.44	0.9927
0.18	0.5714	0.94	0.8264	1.70	0.9554	2.02	0.9783	2.46	0.9931
0.20	0.5793	0.96	0.8315	1.72	0.9573	2.04	0.9793	2.48	0.9934
0.22	0.5871	0.98	0.8365	1.74	0.9591	2.06	0.9803	2.50	0.9938
0.24	0.5948	1.00	0.8413	1.76	0.9608	2.08	0.9812	2.52	0.9941
0.26	0.6026	1.02	0.8461	1.78	0.9625	2.10	0.9821	2.54	0.9945
0.28	0.6103	1.04	0.8508	1.80	0.9641	2.12	0.9830	2.56	0.9948
0.30	0.6179	1.06	0.8554	1.82	0.9656	2.14	0.9838	2.58	0.9951

Definition: Zufallsvariable, Messbarkeit und Verteilung

Eine Zufallsvariable X ist nur noch definiert, wenn das Urbild von X eine „gutartige“ Menge ist.

Seien (Ω, \mathcal{A}) , (S, \mathcal{C}) messbare Räume und $X : \Omega \rightarrow S$ eine Abbildung. X heißt S -wertige **Zufallsvariable**, falls $X^{-1}(C) \in \mathcal{A}$ für alle $C \in \mathcal{C}$ ist.

Man schreibt $X : (\Omega, \mathcal{A}) \rightarrow (S, \mathcal{C})$ und sagt, dass X $(\mathcal{A}, \mathcal{C})$ -**messbar** ist. Die sogenannte **Verteilung** von X unter \mathbb{P} wird durch das Wahrscheinlichkeitsmaß

$$\mathbb{P}^X(C) := \mathbb{P}(X^{-1}(C)) = \mathbb{P}(X \in C), \quad C \in \mathcal{C}$$

auf (S, \mathcal{C}) definiert.

Definition: Verteilungsfunktion und Dichte von Zufallsvariablen

Sei $X : (\Omega, \mathcal{A}) \rightarrow (S, \mathcal{C})$ eine \mathbb{R} -wertige Zufallsvariable auf dem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, \mathbb{P})$.

- Die Verteilungsfunktion der Verteilung \mathbb{P}^X von X

$$F_X : \mathbb{R} \rightarrow [0, 1], \quad x \mapsto \mathbb{P}^X((-\infty, x]) = \mathbb{P}(X \leq x)$$

wird auch **Verteilungsfunktion** von X genannt.

- X heißt **stetige Zufallsvariable**, falls F_X eine Dichte f_X besitzt:

$$\mathbb{P}(X \leq x) = \int_{-\infty}^x f_X(t) dt \quad \forall x \in \mathbb{R}$$

f_X heißt dann auch **Dichte** von X .

k σ -Regeln für die Normalverteilung

Für $X \sim N(\mu, \sigma^2)$ und alle $t > 0$ gilt

$$\mathbb{P}(|X - \mu| \leq \sigma t) = 2 \cdot \Phi(t) - 1$$

und insbesondere

$$\mathbb{P}(|X - \mu| \leq k\sigma) = \begin{cases} 0,6827, & k = 1 \\ 0,9545, & k = 2 \\ 0,9973, & k = 3 \end{cases}$$

Definition: Borel- σ -Algebren auf \mathbb{R}^n

Borel- σ -Algebren, Verteilungsfunktion und Dichten kann man analog auch für $\Omega = \mathbb{R}^n$ definieren. Die **Borel- σ -Algebra** $\mathcal{B}_{\mathbb{R}^n}$ auf \mathbb{R}^n wird definiert als die kleinste σ -Algebra, die alle Rechteckmengen $\bigtimes_{i=1}^n (a_i, b_i]$ mit $-\infty < a_i < b_i < \infty$ für alle $i \in \{1, \dots, n\}$ enthält.

Definition: Multivariate Verteilungsfunktion

Ist \mathbb{P} ein Wahrscheinlichkeitsmaß auf $(\mathbb{R}^n, \mathcal{B}_{\mathbb{R}^n})$, so wird die zugehörige **multivariate Verteilungsfunktion** F definiert durch

$$F(x_1, \dots, x_n) := \mathbb{P}\left(\bigtimes_{i=1}^n (-\infty, x_i]\right), \quad (x_1, \dots, x_n) \in \mathbb{R}^n$$

Definition: Dichten auf \mathbb{R}^n

Sei \mathbb{P} ein Wahrscheinlichkeitsmaß auf $(\mathbb{R}^n, \mathcal{B}_{\mathbb{R}^n})$ mit multivariater Verteilungsfunktion F . Existiert eine Abbildung $f: \mathbb{R}^n \rightarrow [0, \infty)$, sodass für alle $(x_1, \dots, x_n) \in \mathbb{R}^n$

$$F(x_1, \dots, x_n) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} \cdots \int_{-\infty}^{x_n} f(t_1, \dots, t_n) dt_n \cdots dt_2 dt_1$$

gilt, so heißt f **Dichte** von \mathbb{P} bzw. von F .

Es gilt dann $\forall B \in \mathcal{B}_{\mathbb{R}^n}$

$$\mathbb{P}(B) = \int_B f(t) dt = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \mathbb{1}_B(t_1, \dots, t_n) \cdot f(t_1, \dots, t_n) dt_n \dots dt_1$$

Insbesondere gilt für jede Dichte f : $\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(t_1, \dots, t_n) dt_n \dots dt_1 = 1$
und für $B = (a_1, b_1] \times \cdots \times (a_n, b_n]$, $a_i < b_i$: $\mathbb{P}(B) = \int_{a_1}^{b_1} \cdots \int_{a_n}^{b_n} f(t_1, \dots, t_n) dt_n \dots dt_1$.

Definition: Zufallsvektoren

Ist (Ω, \mathcal{A}) ein messbarer Raum und $X_i : \Omega \rightarrow \mathbb{R}$, $1 \leq i \leq n$, so gilt $\forall 1 \leq i \leq n$:

$$X = (X_1, \dots, X_n) : \Omega \rightarrow \mathbb{R}^n \text{ } (\mathcal{A}, \mathcal{B}_{\mathbb{R}^n})\text{-messbar} \iff X_i : \Omega \rightarrow \mathbb{R} \text{ } (\mathcal{A}, \mathcal{B}_{\mathbb{R}})\text{-messbar}$$

In dem Fall wird X auch n -dimensionaler **Zufallsvektor** genannt. Die multivariate Verteilungsfunktion von \mathbb{P}^X

$$F_X(x_1, \dots, x_n) := \mathbb{P}(X_1 \leq x_1, X_2 \leq x_2, \dots, X_n \leq x_n)$$

heißt **gemeinsame Verteilungsfunktion** von X_1, \dots, X_n . Besitzt F_X eine Dichte f_X , dann heißt X **stetig verteilt** und f_X heißt **gemeinsame Dichte** von X_1, \dots, X_n . Es gilt dann $\mathbb{P}^X(B) = \mathbb{P}(X \in B) = \int_B f_X(t) dt$, $B \in \mathcal{B}_{\mathbb{R}^n}$.

Definition: Randverteilungen

Ist $X = (X_1, \dots, X_n)$ ein Zufallsvektor auf $(\Omega, \mathcal{A}, \mathbb{P})$, so heißen die Verteilungen \mathbb{P}^{X_i} **Randverteilungen/Marginalverteilungen**. Die Verteilungsfunktion F_i von \mathbb{P}^{X_i} bzw. X_i berechnet sich wie folgt:

$$\begin{aligned} F_i(x) &= \mathbb{P}(X_1 < \infty, \dots, X_{i-1} < \infty, X_i \leq x, X_{i+1} < \infty, \dots, X_n < \infty) \\ &= F(\infty, \dots, \infty, x, \infty, \dots, \infty), \end{aligned}$$

wobei x an der i -ten Stelle steht.

Satz über die Dichte einer Komponente eines Zufallsvektors

Besitzt die Zufallsvariable $X = (X_1, \dots, X_n)$ eine gemeinsame Dichte $f : \mathbb{R}^n \rightarrow [0, \infty)$, so hat X_i die Dichte $f_i : \mathbb{R} \rightarrow [0, \infty)$ mit

$$f_i(x) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(t_1, \dots, t_{i-1}, x, t_{i+1}, \dots, t_n) dt_n \cdots dt_{i+1} dt_{i-1} \cdots dt_1$$

für alle $x \in \mathbb{R}$.

Definition: Unabhängigkeit von Zufallsvariablen

Zufallsvariablen $X_1, \dots, X_n : (\Omega, \mathcal{A}) \rightarrow (\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ heißen **stochastisch unabhängig**, wenn die Ereignisse $\{X_1 \in B_1\}, \dots, \{X_n \in B_n\}$ für alle $B_1, \dots, B_n \in \mathcal{B}_{\mathbb{R}}$ stochastisch unabhängig sind.

Satz für stochastisch unabhängige Zufallsvariablen

Für Zufallsvariablen $X_i : (\Omega, \mathcal{A}) \rightarrow (\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ mit Verteilungsfunktionen F_{X_i} sind äquivalent:

- X_1, \dots, X_n sind stochastisch unabhängig
- $\mathbb{P}(X_1 \in B_1, \dots, X_n \in B_n) = \prod_{i=1}^n \mathbb{P}(X_i \in B_i) \quad \forall B_1, \dots, B_n \in \mathcal{B}_{\mathbb{R}}$
- $\mathbb{P}(X_1 \leq x_1, \dots, X_n \leq x_n) = \prod_{i=1}^n \mathbb{P}(X_i \leq x_i) = \prod_{i=1}^n F_{X_i}(x_i) \quad \forall x_1, \dots, x_n \in \mathbb{R}$

Zufallsvariablen sind also genau dann unabhängig, wenn die Verteilungsfunktion ihrer gemeinsamen Verteilung das Produkt ihrer Verteilungsfunktionen ist.

Dichten unabhängiger Zufallsvariablen

Seien $X_i : \Omega \rightarrow \mathbb{R}$ für $i \in \{1, \dots, n\}$ Zufallsvariablen auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, \mathbb{P})$. Falls alle Randverteilungen \mathbb{P}^{X_i} jeweils eine Dichte f_i besitzen, dann sind äquivalent:

- X_1, \dots, X_n sind stochastisch unabhängig:
- $\mathbb{P}^{(X_1, \dots, X_n)}$ besitzt eine Dichte f gegeben durch $f(x_1, \dots, x_n) := \prod_{i=1}^n f_i(x_i)$

Blockungslemma

Seien $X_{11}, X_{12}, \dots, X_{1n_1}, X_{21}, \dots, X_{2n_2}, \dots, X_{k1}, \dots, X_{kn_k}$ stochastisch unabhängige Zufallsvariablen und $g_1 : \mathbb{R}^{n_1} \rightarrow \mathbb{R}, g_2 : \mathbb{R}^{n_2} \rightarrow \mathbb{R}, \dots, g_k : \mathbb{R}^{n_k} \rightarrow \mathbb{R}$ Funktionen. Dann sind auch die Zufallsvariablen

$$Y_1 := g_1(X_{11} \dots, X_{1n_1}), Y_2 := g_2(X_{21} \dots, X_{2n_2}), \dots, Y_k := g_k(X_{k1} \dots, X_{kn_k})$$

stochastisch unabhängig. Funktionen von disjunkten Blöcken unabhängiger Zufallsvariablen sind also wieder unabhängig.

Definition: Faltung für Verteilungen mit Dichten

- Sind $X, Y : (\Omega, \mathcal{A}) \rightarrow (\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ unabhängige Zufallsvariablen auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, \mathbb{P})$ so wird die Verteilung \mathbb{P}^{X+Y} die **Faltung** von \mathbb{P}^X und \mathbb{P}^Y genannt. Schreibe dafür: $\mathbb{P}^X * \mathbb{P}^Y$
- Besitzt X eine Dichte f_X und Y eine Dichte f_Y , so ist

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

eine Dichte von $\mathbb{P}^X * \mathbb{P}^Y$, d.h. $f_{X+Y} = f_X * f_Y$

Definition: Gamma-Verteilung

Die Zufallsvariable X hat eine **Gamma-Verteilung** mit Parametern $\alpha > 0$ und $\beta > 0$, falls X die Dichte

$$f(t) = \frac{\beta^\alpha}{\Gamma(\alpha)} t^{\alpha-1} e^{-\beta t} \mathbb{1}_{(0,\infty)}(t), \quad t \in \mathbb{R}$$

mit Gamma-Funktion $\Gamma(t) := \int_0^\infty x^{t-1} e^{-x} dx$, $t > 0$ besitzt.

Additionsgesetze

$X \sim$	$Y \sim$	$X + Y \sim$
$\text{Bin}_{m,p}$	$\text{Bin}_{n,p}$	$\text{Bin}_{m+n,p}$
Pois_λ	Pois_μ	$\text{Pois}_{\lambda+\mu}$
N_{μ,σ^2}	N_{ν,τ^2}	$N_{\mu+\nu,\sigma^2+\tau^2}$
$\Gamma_{\mu,\beta}$	$\Gamma_{\nu,\beta}$	$\Gamma_{\mu+\nu,\beta}$
Exp_λ	Exp_λ	$\Gamma_{2,\lambda}$

Beachte: In jedem Fall sind X und Y als stochastisch unabhängig vorausgesetzt.

Verteilungsfunktion von Maximum und Minimum

Seien X_1, \dots, X_n stochastisch unabhängige Zufallsvariablen mit den Verteilungsfunktionen F_{X_1}, \dots, F_{X_n} . Dann gilt:

- $U := \max(X_1, \dots, X_n)$ besitzt die Verteilungsfunktion

$$F_U(t) = \prod_{j=1}^n F_{X_j}(t), \quad t \in \mathbb{R}$$

- $V := \min(X_1, \dots, X_n)$ besitzt die Verteilungsfunktion

$$F_V(t) = 1 - \prod_{j=1}^n (1 - F_{X_j}(t)), \quad t \in \mathbb{R}$$

5 Erwartungswerte und Momente von Zufallsvariablen

Definition: Erwartungswert auf einem diskreten Wahrscheinlichkeitsraum

Der **Erwartungswert** einer \mathbb{R} -wertigen Zufallsvariable $X : \Omega \rightarrow \mathbb{R}$ auf einem diskreten Wahrscheinlichkeitsraum (Ω, \mathbb{P}) ist definiert als:

$$\mathbb{E}_{\mathbb{P}}[X] := \mathbb{E}[X] := \sum_{x \in X(\Omega)} x \cdot \mathbb{P}(X = x) = \sum_{x \in X(\Omega)} x \cdot \mathbb{P}^X(\{x\})$$

falls $\sum_{x \in X(\Omega)} |x| \cdot \mathbb{P}(X = x) < \infty$. $\mathbb{E}_{\mathbb{P}}[X]$ heißt auch **Mittelwert** von \mathbb{P}^X .

Definition: Erwartungswert für stetige Zufallsvariablen

Der **Erwartungswert** einer stetigen Zufallsvariable X mit Dichte f_X wird definiert als:

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} x \cdot f_X(x) dx$$

falls $\int_{-\infty}^{\infty} |x| \cdot f_X(x) dx < \infty$. $\mathbb{E}[X]$ heißt auch **Mittelwert** von \mathbb{P}^X .

Transformationssatz

Sei $(\Omega, \mathcal{A}, \mathbb{P})$ ein Wahrscheinlichkeitsraum, $X : \Omega \rightarrow S$ eine diskrete/stetige Zufallsvariable mit (Zähl-)Dichte f_X und $g : S \rightarrow \mathbb{R}$ eine Funktion. Die Zufallsvariable $g(X) = g \circ X$ besitzt genau dann einen endlichen Erwartungswert bzgl. \mathbb{P} , wenn g einen endlichen Erwartungswert bzgl. \mathbb{P}^X besitzt. In diesem Fall gilt:

$$\mathbb{E}_{\mathbb{P}}[g(X)] = \mathbb{E}_{\mathbb{P}^X}[g] = \begin{cases} \sum_{x \in S} g(x) \cdot f_X(x) & \text{falls } X \text{ diskret} \\ \int_{-\infty}^{\infty} g(x) \cdot f_X(x) dx & \text{falls } X \text{ stetig} \end{cases}$$

Der Transformationssatz gilt auch für Zufallsvektoren (s. Folie 150)

Rechenregeln für Erwartungswerte

Seien X, Y diskrete/stetige Zufallsvariablen auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, \mathbb{P})$, die Erwartungswerte besitzen. Dann gilt:

- $\mathbb{E}[aX + Y] = a \cdot \mathbb{E}[X] + \mathbb{E}[Y]$ für alle $a \in \mathbb{R}$ (**Linearität**)
- Gilt $X \leq Y$, dann folgt $\mathbb{E}[X] \leq \mathbb{E}[Y]$ (**Monotonie**)
- Wenn f_X symmetrisch zu $x = a$ ist, dann gilt $\mathbb{E}[X] = a$

Siebformel von Sylvester-Poincaré

Seien $A_1, \dots, A_n \in \mathcal{A}$ Ereignisse im einem diskreten Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, \mathbb{P})$. Dann gilt:

$$\mathbb{P}\left(\bigcup_{i=1}^n A_i\right) = \sum_{I \subseteq \{1, \dots, n\}, I \neq \emptyset} (-1)^{|I|+1} \cdot \mathbb{P}\left(\bigcap_{i \in I} A_i\right)$$

Darstellungsformel für nicht-negative Zufallsvariablen

Ist X eine \mathbb{N}_0 -wertige oder \mathbb{R}_+ -wertige Zufallsvariable, so gilt

$$\mathbb{E}[X] = \begin{cases} \sum_{n=1}^{\infty} \mathbb{P}(X \geq n) & \text{falls } X \text{ diskret} \\ \int_0^{\infty} \mathbb{P}(X \geq x) & \text{falls } X \text{ stetig} \end{cases}$$

Multiplikationsformel für Erwartungswerte

Sind X und Y stochastisch unabhängige reellwertige Zufallsvariablen mit Erwartungswerten $\mathbb{E}[X]$ und $\mathbb{E}[Y]$, so gilt

$$\mathbb{E}[X \cdot Y] = \mathbb{E}[X] \cdot \mathbb{E}[Y]$$

Die Umkehrung gilt im Allgemeinen nicht!