

Teil 2

Wahrscheinlichkeitsrechnung



Moodle



Lehrbuch

Teil 2: Wahrscheinlichkeitsrechnung

Kapitel 7: Zufallsvorgänge, Ereignisse und Wahrscheinlichkeiten

Kapitel 8: Zufallsvariablen und Verteilungen

Kapitel 9: Verteilungsparameter

Kapitel 10: Gesetz der großen Zahlen und zentraler Grenzwertsatz

Teil 2: Wahrscheinlichkeitsrechnung

Lernziele

- ▶ Kennenlernen zentraler Grundbegriffe der Wahrscheinlichkeitsrechnung.
- ▶ Fähigkeit, mit gebräuchlichen Typen von Zufallsvariablen umzugehen.
- ▶ Erlangung eines fundierten Überblicks über die wichtigsten diskreten und stetigen Verteilungen.
- ▶ Beherrschung wichtiger Verteilungsparameter sowie deren Interpretation.
- ▶ Verständnis des Gesetzes der großen Zahlen sowie des zentralen Grenzwertsatzes.

Kapitel 7: Zufallsvorgänge, Ereignisse und Wahrscheinlichkeiten

Zufallsvorgang

Ein **Zufallsvorgang** ist ein Geschehen, bei welchem sich gegenseitig ausschließende Ausgänge systematisch aber unvorhersehbar eintreten.

Synonyme:

Zufallsabhängiges Geschehen, Zufallsexperiment,
stochastischer Vorgang

Beispiele:

Würfeln, Bestehen einer Klausur durch teilweises Raten,
Platzierung des Schusses beim Elfmeter, ...

Elementarereignisse und Ergebnismenge

Die möglichem sich gegenseitig ausschließenden Ausgänge eines Zufallsvorganges heißen **Elementarereignisse**.

Die Menge aller Elementarereignisse heißt **Ergebnismenge**.

Wir bezeichnen ein Elementarereignis mit ω und die Ergebnismenge mit Ω .

Beispiel Würfel:

$$\Omega = \{\overset{\omega_1}{1}, \overset{\omega_2}{2}, \dots, \overset{\omega_6}{6}\}$$

Beispiel zwei Würfel:

$$\Omega = \{(\overset{\omega_1}{1}, \overset{\omega_2}{1}), (\overset{\omega_2}{1}, \overset{\omega_2}{2}), \dots, (\overset{\omega_{35}}{6}, \overset{\omega_{35}}{5}), (\overset{\omega_{36}}{6}, \overset{\omega_{36}}{6})\}$$

Ereignisse

Alle Elementarereignisse schließen sich gegenseitig aus.
Dies ist bei Ereignissen nicht notwendig der Fall.

Ein **Ereignis** ist eine Menge von Elementarereignissen, also eine Teilmenge von Ω .

Wir schreiben $A \subseteq \Omega$.

Beispiel: Ereignisse beim Würfeln ($\Omega = \{1, 2, \dots, 6\}$)

	Menge
gerade Zahl	$A = \{2, 4, 6\}$
Primzahl	$B = \{2, 3, 5\}$
Zahl > 4	$C = \{5, 6\}$

Ereignissystem

Die Gesamtheit aller bei einem Zufallsvorgang in Betracht kommenden Ereignisse heißt **Ereignissystem**.

Kleinstes Ereignissystem: $\{\emptyset, \Omega\}$

Größtes Ereignissystem:

Potenzmenge $\mathcal{P}(\Omega)$ (Menge aller Teilmengen von Ω)

Ein Ereignissystem muss bestimmte Forderungen erfüllen.

Hierfür müssen wir aber zunächst sinnvolle **Rechenoperationen** für Mengen definieren.

Mengenoperationen

Seien A und B Teilmengen von Ω .

- ▶ $A \cap B$: **Schnittmenge** von A und B
Menge aller $\omega \in \Omega$ die in A und in B enthalten sind.
- ▶ $A \cup B$: **Vereinigungsmenge** von A und B
Menge aller $\omega \in \Omega$ die in A oder B (oder beiden) liegen.
- ▶ \bar{A} : **Komplementärmenge** von A
Menge aller $\omega \in \Omega$ die nicht in A liegen.
- ▶ $A \setminus B$: **Differenzmenge** von A und B
Menge aller $\omega \in \Omega$ die in A aber nicht in B liegen.
- ▶ Zwei Ereignisse A und B heißen **disjunkt**, wenn es keine ω gibt, die in A und B liegen, also wenn $A \cap B = \emptyset$.

Beispiel: Ereignisse beim Würfeln (Fortsetzung)

verbal	Menge
gerade Zahl	$A = \{2, 4, 6\}$
Primzahl	$B = \{2, 3, 5\}$
Zahl > 4	$C = \{5, 6\}$

- ▶ $A \cup B = \{2, 3, 4, 5, 6\}$: gerade Zahl **oder** Primzahl
- ▶ $A \setminus B = \{4, 6\}$: gerade Zahl, **aber keine** Primzahl
- ▶ $B \cap C = \{5\}$: Primzahl **und** größer 4
- ▶ $A \cap B \cap C = \emptyset$: gerade Zahl **und** Primzahl **und** größer 4 - das erfüllt keine Zahl

Eigenschaften eines Ereignissystems

Das Ereignissystem „filtert“ die Informationen, die für die jeweilige Untersuchung relevant sind. Sie spiegelt sozusagen das, was potentiell gewusst werden kann.

Hierfür fordern wir von einem Ereignissystem \mathcal{A} einer Ergebnismenge Ω folgende mathematischen Eigenschaften:

- ▶ Falls $A \in \mathcal{A}$, dann muss auch $\bar{A} \in \mathcal{A}$ sein.
- ▶ Falls $A, B \in \mathcal{A}$, dann muss auch $A \cap B \in \mathcal{A}$ sein.
- ▶ Falls $A, B \in \mathcal{A}$, dann muss auch $A \cup B \in \mathcal{A}$ sein.

Definition Partition oder Zerlegung

Eine Menge von Ereignissen $\{A_1, A_2, \dots, A_k\}$ heißt **Partition** von Ω , falls

- ▶ $A_1 \cup A_2 \cup \dots \cup A_k = \Omega$
Die Ergebnismenge ist die Vereinigung aller Ereignisse
- ▶ $A_i \cap A_j = \emptyset$ für alle $i, j = 1, \dots, k, i \neq j$.
Alle Ereignisse sind paarweise disjunkt.

Für Partitionen von Ω gelten vereinfachte Rechenregeln.

Wahrscheinlichkeit von Ereignissen

Gesucht ist eine Zahl zur Beschreibung der Chance des Eintretens eines bestimmten Ereignisses.

Für unsere Überlegungen benötigen wir zwei Wahrscheinlichkeitsbegriffe (wobei der 2. den 1. umfasst):

- ▶ **Klassischer Wahrscheinlichkeitsbegriff (Laplace)**
- ▶ **Axiomatischer Wahrscheinlichkeitsbegriff**

Definition Klassischer Wahrscheinlichkeitsbegriff

Ein **Laplace-Experiment** ist ein Zufallsvorgang mit endlichem Ω , in dem alle Elementarereignisse mit gleicher Wahrscheinlichkeit auftreten.

- ▶ Jedes Elementarereignis hat die Wahrscheinlichkeit $\frac{1}{|\Omega|}$.
- ▶ Sei $A \subseteq \Omega$ ein Ereignis aus dem endlichen Ereignisraum Ω , dann ist die Wahrscheinlichkeit des Ereignisses A :

$$\mathbb{P}(A) = \frac{|A|}{|\Omega|} = \frac{\text{Anzahl der für } A \text{ günstigen Elementarereignisse}}{\text{Anzahl aller Elementarereignisse}}$$

Rechenregeln für Wahrscheinlichkeiten

Seien A und B Ereignisse eines Laplace-Experiments, dann gilt:

▶ $\mathbb{P}(\Omega) = 1$

▶ $\mathbb{P}(\emptyset) = 0$

▶ $\mathbb{P}(\bar{A}) = 1 - \mathbb{P}(A)$

▶ Falls $A \cap B = \emptyset$, falls also A und B **disjunkt** sind, gilt:

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$$

▶ Allgemein gilt:

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$$

Axiomatischer Wahrscheinlichkeitsbegriff

Definition nach A.N. Kolmogorov (1903–1987)

Sei \mathcal{A} ein Ereignissystem von Ω .

Jedem Ereignis $A \in \mathcal{A}$ wird eine reelle Zahl $\mathbb{P}(A)$, die Wahrscheinlichkeit von A , zugeordnet, für die gilt:

1. $\mathbb{P}(A) \geq 0$
2. $\mathbb{P}(\Omega) = 1$
3. Ist $A \cap B = \emptyset$, so gilt $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$

Rechenregeln für Zufallsvorgänge

Die 3 Kolmogorov Axiome implizieren:

- ▶ $\mathbb{P}(\emptyset) = 0$
- ▶ $\mathbb{P}(\bar{A}) = 1 - \mathbb{P}(A)$
- ▶ $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$
- ▶ $\mathbb{P}(A \cup B \cup C) = \mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(C)$
- $\mathbb{P}(A \cap B) - \mathbb{P}(A \cap C) - \mathbb{P}(B \cap C)$
+ $\mathbb{P}(A \cap B \cap C)$

- ▶ Falls $\{A_1, A_2, \dots, A_n\}$ Partition von Ω , dann gilt:
 $\mathbb{P}(B) = \mathbb{P}(B \cap A_1) + \dots + \mathbb{P}(B \cap A_n)$

Definition Wahrscheinlichkeitsraum

Das Tripel $\{\Omega, \mathcal{A}, \mathbb{P}\}$ heißt **Wahrscheinlichkeitsraum**, falls

- ▶ Ω eine Ergebnismenge,
- ▶ \mathcal{A} ein Ereignissystem von Ω und
- ▶ $\mathbb{P} : \mathcal{A} \rightarrow \mathbb{R}$ ein Wahrscheinlichkeitsmaß auf \mathcal{A}

ist.

Bedingte Wahrscheinlichkeit

- ▶ Oft hängt die Wahrscheinlichkeit eines Ereignisses A vom Eintreten eines anderen Ereignisses B ab.
- ▶ Uns interessiert dann nur noch die Wahrscheinlichkeit unter dieser „Bedingung“ B .
- ▶ Wir beschränken die „möglichen Elementarereignisse“, die wir bei der Berechnung der Wahrscheinlichkeit betrachten.

Bedingte Wahrscheinlichkeit

Definition

Sei $\mathbb{P}(B) > 0$, dann heißt

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

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

Die bedingte Wahrscheinlichkeit ist nur für bedingende Ereignisse mit $\mathbb{P}(B) > 0$ definiert.

Wahrscheinlichkeiten von Schnittmengen und bedingte Wahrscheinlichkeiten

Seien A, B, \dots, Y, Z Ereignisse eines Zufallsvorgangs. Dann gilt unter der Voraussetzung $\mathbb{P}(A \cap B \cap \dots \cap Y) > 0$:

$$\begin{aligned} \mathbb{P}(A \cap B \cap \dots \cap Z) &= \mathbb{P}(A) \\ &\cdot \mathbb{P}(B|A) \\ &\cdot \mathbb{P}(C|A \cap B) \\ &\dots \\ &\cdot \mathbb{P}(Z|A \cap B \cap \dots \cap Y) \end{aligned}$$

Für $\{A, B\}$, bekommen wir nichts anderes als die (umgeschriebene) Definition der bedingten Wahrscheinlichkeit:

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

Satz von der vollständigen Wahrscheinlichkeit

Seien die Ereignisse A_1, \dots, A_k eine Partition von Ω , d. h.

1. $A_i \cap A_j = \emptyset$ für alle $i \neq j$ (d. h. alle Ereignisse sind disjunkt)
2. $A_1 \cup A_2 \cup \dots \cup A_k = \Omega$,

dann gilt:

$$\mathbb{P}(B) = \sum_{i=1}^k \mathbb{P}(B|A_i) \cdot \mathbb{P}(A_i)$$

Formel von Bayes

Seien die Ereignisse A_1, \dots, A_k eine Partition von Ω und gelte $\mathbb{P}(B) > 0$, dann gilt:

$$\mathbb{P}(A_j|B) = \frac{\mathbb{P}(B|A_j) \cdot \mathbb{P}(A_j)}{\sum_{i=1}^k \mathbb{P}(B|A_i) \cdot \mathbb{P}(A_i)}$$

A-priori- und a-posteriori-Wahrscheinlichkeiten

Sei $\{A_1, \dots, A_m\}$ eine Partition von Ω .

Wir nennen $\mathbb{P}(A_i)$ die **a-priori-Wahrscheinlichkeiten**.

Es trete nun das Ereignis $B \subseteq \Omega$ ein.

Für die möglichen Ereignisse A_i seien die bedingten Wahrscheinlichkeiten $\mathbb{P}(B|A_i)$ für alle A_i bestimmbar.

Gegeben B gilt dann:

$$\mathbb{P}(A_j|B) = \frac{\mathbb{P}(B|A_j) \cdot \mathbb{P}(A_j)}{\sum_{i=1}^m \mathbb{P}(B|A_i) \cdot \mathbb{P}(A_i)}$$

Die $\mathbb{P}(A_j|B)$ heißen **a-posteriori-Wahrscheinlichkeiten**.

Sie beinhalten die Zusatzinformation über die A_i durch die Realisierung von B .

Definition stochastisch unabhängige Ereignisse

Zwei Ereignisse A und B mit positiven Wahrscheinlichkeiten heißen voneinander **stochastisch unabhängig**, wenn gilt:

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

Dies impliziert für die bedingten Wahrscheinlichkeiten:

$$\mathbb{P}(B|A) = \mathbb{P}(B) \text{ und } \mathbb{P}(A|B) = \mathbb{P}(A)$$

Kapitel 8: Zufallsvariablen und Verteilungen

Definition Eindimensionale Zufallsvariable

Für eine gegebene Ergebnismenge Ω eines Zufallsexperimentes heißt eine Funktion

$$\begin{aligned} X &: \Omega \rightarrow \mathbb{R} \\ \omega &\mapsto X(\omega), \end{aligned}$$

eindimensionale Zufallsvariable.

Hierbei heißt (für gegebenes ω) der Funktionswert $X(\omega)$ die **Realisation** dieser Zufallsvariablen.

Wahrscheinlichkeiten von Zufallsvariablen

Für eine Menge $B \subseteq \mathbb{R}$ von Zahlen und eine Zufallsvariable X seien alle Elementarereignisse $\omega \in \Omega$, für welche X einen Wert $X(\omega) \in B$ annimmt, das Ereignis A definiert:

$$A = \{\omega \in \Omega : X(\omega) \in B\}$$

Es gilt also:

$$\omega \in A \Leftrightarrow X(\omega) \in B$$

Wir schreiben $\mathbb{P}(B) = \mathbb{P}(X(\omega) \in B)$:

Wahrscheinlichkeit, dass die Zufallsvariable X einen Wert in B annimmt.

Es gilt hier $\mathbb{P}(B) = \mathbb{P}(A)$.

Definition Mehrdimensionale Zufallsvariable

Für eine gegebene Ergebnismenge Ω eines Zufallsexperimentes und ein $n \in \mathbb{N}$ heißt eine Funktion

$$\begin{aligned} X &: \Omega \rightarrow \mathbb{R}^n \\ \omega &\mapsto X(\omega), \end{aligned}$$

mehrdimensionale Zufallsvariable.

Wir verstehen X als Vektor von n eindimensionalen Zufallsvariablen:

$$X = (X_1, X_2, \dots, X_n)$$

Definition Unabhängigkeit von Zufallsvariablen

Die Zufallsvariablen X_1, X_2, \dots, X_n heißen **unabhängig**, falls für beliebige Teilmengen $B_1, B_2, \dots, B_n \subseteq \mathbb{R}$ gilt:

$$\mathbb{P}(X_1 \in B_1, \dots, X_n \in B_n) = \mathbb{P}(X_1 \in B_1) \cdot \dots \cdot \mathbb{P}(X_n \in B_n)$$

Definition Verteilungsfunktion

Die Funktion $F : \mathbb{R} \rightarrow \mathbb{R}$ mit $F(x) = \mathbb{P}(X \leq x)$ heißt **Verteilungsfunktion** von X .

Sie gibt die Wahrscheinlichkeit an, mit welcher die Zufallsvariable X einen Wert von höchstens x annimmt.

Eigenschaften:

- ▶ $\lim_{x \rightarrow -\infty} F(x) = 0$
- ▶ $\lim_{x \rightarrow \infty} F(x) = 1$
- ▶ $F(x_1) \leq F(x_2)$ für alle x_1, x_2 mit $x_1 < x_2$ (monoton wachsend)
- ▶ $\lim_{x \rightarrow x_0^+} F(x) = F(x_0)$ für alle x_0 (rechtseitig stetig)
- ▶ $\mathbb{P}(a < X \leq b) = F(b) - F(a)$ für alle a, b mit $a < b$

Eindimensionale diskrete Zufallsvariablen

Eine **diskrete** Zufallsvariable kann nur **endlich** oder **abzählbar unendlich** viele Werte annehmen.

Der Wertebereich besitzt also die Gestalt

$$\{x_1, x_2, x_3, \dots\}$$

Zu jedem x_i gibt es dann eine Zahl $p_i > 0$ mit $p_i = \mathbb{P}(X(\omega) = x_i)$ und $p_1 + p_2 + p_3 + \dots = 1$.

Wahrscheinlichkeits- und Verteilungsfunktion von diskreten Zufallsvariablen

Wahrscheinlichkeitsfunktion

$$f(x) = \mathbb{P}(X = x) = \begin{cases} p_i & \text{falls } x = x_i \\ 0 & \text{sonst} \end{cases}$$

Verteilungsfunktion

$$F(x) = \mathbb{P}(X \leq x) = \sum_{x_i \leq x} f(x_i)$$

Die Verteilungsfunktion diskreter Zufallsvariablen ist eine **Treppenfunktion**.

Wir nennen die Verteilungsfunktion manchmal auch **kumulierte Wahrscheinlichkeitsfunktion**.

Wichtige diskrete Verteilungen

- ▶ Bernoulliverteilung
- ▶ Binomialverteilung
- ▶ Poissonverteilung

Die Bernoulliverteilung $B(1, p)$

Für ein gegebenes Ereignis $A \subset \Omega$ nennen wir eine Zufallsvariable X **Bernoulli-verteilt**, falls sie nur zwei Werte annehmen kann:

$$X(\omega) = \begin{cases} 1 & \text{falls } \omega \in A \\ 0 & \text{falls } \omega \notin A \end{cases}$$

Die Wahrscheinlichkeit, dass $X(\omega) = 1$ bezeichnen wir mit

$$\mathbb{P}(X = 1) = \mathbb{P}(A) = p$$

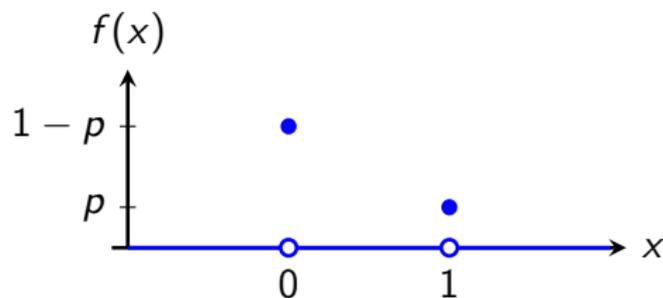
und die entsprechende Gegenwahrscheinlichkeit mit

$$\mathbb{P}(X = 0) = \mathbb{P}(\bar{A}) = 1 - \mathbb{P}(A) = 1 - p$$

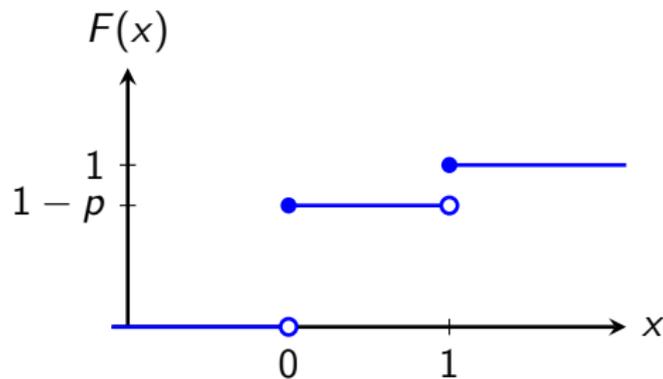
Eine Bernoulli-verteilte Zufallsvariable ist allein durch den Parameter $p \in [0, 1]$ charakterisiert. Wir schreiben $X \sim B(1, p)$.

Die Bernoulliverteilung: f und F

Wahrscheinlichkeitsfunktion f



Verteilungsfunktion F



Die Binomialverteilung $B(n, p)$

Für n unabhängige und identische Bernoulli-verteilte Zufallsvariablen X_1, X_2, \dots, X_n heißt deren Summe

$$X = \sum_{i=1}^n X_i$$

binomialverteilte Zufallsvariable.

$X \sim B(n, p)$ kann die Werte $0, 1, \dots, n$ annehmen.

Die Binomialverteilung $B(n, p)$

Es gibt...

- ▶ eine Möglichkeit, dass $X = 0$.
- ▶ n Möglichkeiten, dass $X = 1$.
- ▶ $\frac{n(n-1)}{2}$ Möglichkeiten, dass $X = 2$.
- ▶ \vdots
- ▶ $\binom{n}{k}$ Möglichkeiten, dass $X = k$ für $k = 0, 1, \dots, n$.

Die Binomialverteilung: Wahrscheinlichkeiten

Wie wahrscheinlich ist es

$$(\overset{x_1}{1}, \dots, \overset{x_k}{1}, \overset{x_{k+1}}{0}, \dots, \overset{x_n}{0})$$

zu beobachten?

$$\mathbb{P}(X_1 = 1, \dots, X_k = 1, X_{k+1} = 0, \dots, X_n = 0) = ?$$

Die Binomialverteilung: Wahrscheinlichkeiten

Wie wahrscheinlich ist es

$$(\overset{x_1}{1}, \dots, \overset{x_k}{1}, \overset{x_{k+1}}{0}, \dots, \overset{x_n}{0})$$

zu beobachten?

Aus unabhängig und identisch verteilt folgt:

$$\mathbb{P}(X_1 = 1, \dots, X_k = 1, X_{k+1} = 0, \dots, X_n = 0)$$

$$= \underbrace{\underbrace{\mathbb{P}(X_1 = 1)}_p \cdot \dots \cdot \underbrace{\mathbb{P}(X_k = 1)}_p}_{k\text{-mal}} \cdot \underbrace{\underbrace{\mathbb{P}(X_{k+1} = 0)}_{1-p} \cdot \dots \cdot \underbrace{\mathbb{P}(X_n = 0)}_{1-p}}_{n-k\text{-mal}}$$

$$= p^k \cdot (1 - p)^{n-k}$$

Die Binomialverteilung: Wahrscheinlichkeiten

Wie viele Möglichkeiten gibt es, genau k Einsen zu beobachten?

Diese Anzahl gibt der Binomialkoeffizient $\binom{n}{k}$ an:

$$\binom{n}{k} = \frac{n!}{(n-k)!k!}$$

Daher ist die Wahrscheinlichkeit, genau k Einsen zu beobachten:

$$\mathbb{P}(X = k) = \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k}$$

Die Binomialverteilung: f und F

Wahrscheinlichkeitsfunktion:

$$f(x) = \begin{cases} \binom{n}{x} \cdot p^x \cdot (1-p)^{n-x} & \text{falls } x \in \{0, 1, \dots, n\} \\ 0 & \text{sonst} \end{cases}$$

Verteilungsfunktion:

$$\begin{aligned} F(x) &= \sum_{k \in \mathbb{N}: k \leq x} f(k) \\ &= \sum_{k \in \mathbb{N}: k \leq x} \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k} \end{aligned}$$

Offensichtlich: F ist für große n und k nur schwer zu berechnen!

→ [Interaktive Binomialverteilung](#)

Poissonverteilung $P(\lambda)$

Eine Zufallsvariable X mit Wertebereich $\{0, 1, 2, \dots\}$ heißt **poissonverteilt** mit Parameter $\lambda > 0$, $X \sim \mathcal{P}(\lambda)$, falls

$$\mathbb{P}(X = k) = \begin{cases} \frac{\lambda^k}{k!} \cdot \exp(-\lambda) & \text{falls } k = 0, 1, 2, \dots \\ 0 & \text{sonst} \end{cases}$$

Poissonverteilung und Binomialverteilung

Falls die Stichprobengröße n sehr groß wird und dabei die Erfolgswahrscheinlichkeit p sehr klein, und zwar so, dass $n \cdot p = \lambda$, dann kann die Binomialverteilung durch die Poissonverteilung mit Parameter λ approximiert werden:

$$B(n, p) \xrightarrow[n \rightarrow \infty, p \rightarrow 0]{\text{mit } n \cdot p = \lambda} P(\lambda)$$

Als Faustregel kann die Poissonverteilung benutzt werden, falls

- ▶ $n > 50$
- ▶ $p < 0,1$
- ▶ $n \cdot p < 10$

Eindimensionale stetige Zufallsvariablen

Wir nennen eine Zufallsvariable **stetig**, falls sie überabzählbar unendlich viele Werte annehmen kann und jeder einzelne dieser Werte mit Wahrscheinlichkeit null auftritt.

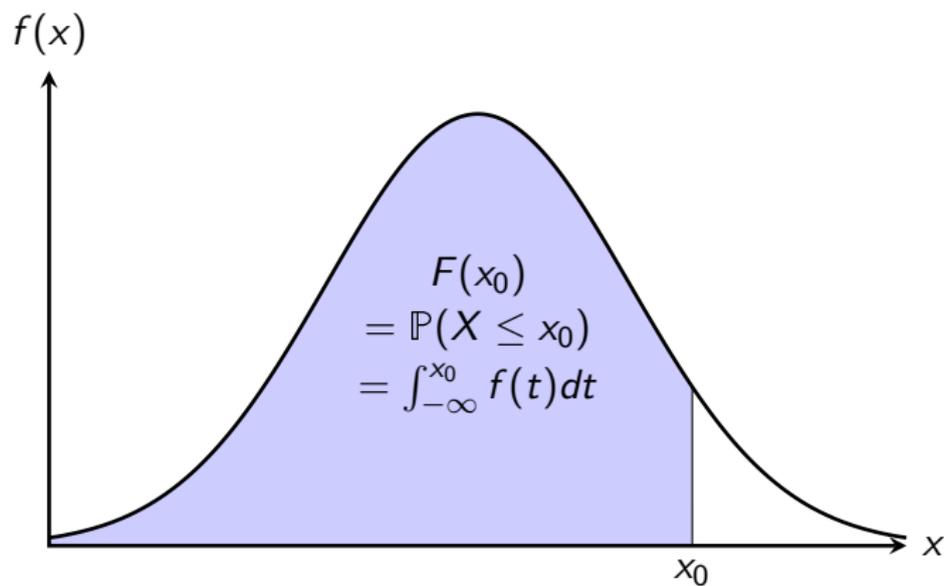
In diesem Fall besitzt sie eine **Dichtefunktion** f , sodass die Verteilungsfunktion F folgende Gestalt hat:

$$F(x) = \int_{-\infty}^x f(t) dt \text{ für jedes } x \in \mathbb{R}$$

Es gilt immer:

$$\int_{-\infty}^{\infty} f(x) dx = 1 \text{ und } f(x) \geq 0 \text{ für alle } x \in \mathbb{R}$$

Dichtefunktion



Eigenschaften der Verteilungsfunktion für stetige Zufallsvariablen

Sei F die Verteilungsfunktion einer stetigen Zufallsvariablen X .

Dann gilt für $a < b$:

▶ $0 \leq F(a) \leq 1$

▶ $F(a) = \mathbb{P}(X \leq a) = \mathbb{P}(X < a)$

▶ $\mathbb{P}(X \geq a) = 1 - \mathbb{P}(X < a) = 1 - \mathbb{P}(X \leq a) = 1 - F(a)$

▶ $\mathbb{P}(a \leq X \leq b) = \mathbb{P}(a < X < b) = F(b) - F(a)$

Wichtige stetige Verteilungen

- ▶ Gleichverteilung
- ▶ Normalverteilung

Gleichverteilung

Sind a, b reelle Zahlen mit $a < b$, so heißt eine Zufallsvariable X mit der Dichtefunktion

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{falls } a \leq x \leq b \\ 0 & \text{sonst} \end{cases}$$

gleichverteilt (oder **uniformverteilt**), $X \sim U(a, b)$.

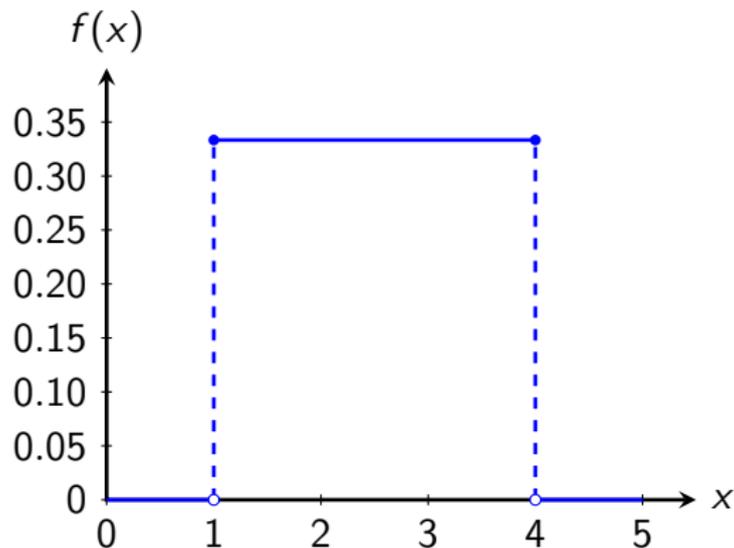
Eigenschaft: Intervalle gleicher Länge sind gleich wahrscheinlich.

Verteilungsfunktion:

$$F(x) = \begin{cases} 0 & \text{falls } x < a \\ \frac{x-a}{b-a} & \text{falls } a \leq x \leq b \\ 1 & \text{falls } x > b \end{cases}$$

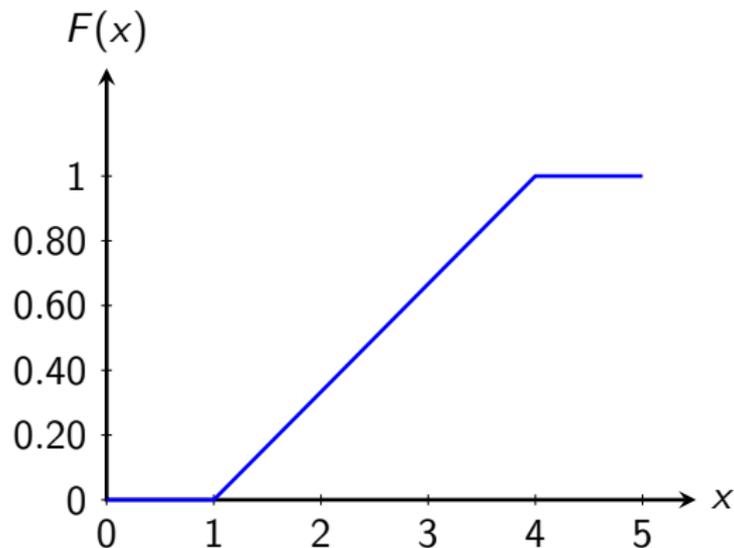
Dichtefunktion der Gleichverteilung

Sei $X \sim U(1, 4)$:



Verteilungsfunktion der Gleichverteilung

Sei $X \sim U(1, 4)$:



Normalverteilung $\mathcal{N}(\mu, \sigma^2)$

Eine Zufallsvariable X mit der Dichtefunktion

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

wobei $\mu \in \mathbb{R}$ und $\sigma > 0$ ist, heißt **normalverteilt**, $X \sim \mathcal{N}(\mu, \sigma^2)$.

Falls $\mu = 0$ und $\sigma = 1$, so heißt X **standard normalverteilt**,
 $X \sim \mathcal{N}(0, 1)$:

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

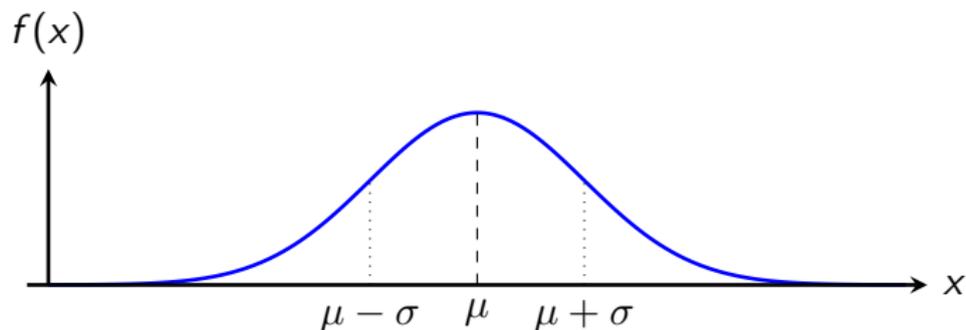
Wir bezeichnen die Dichtefunktion der Standardnormalverteilung mit ϕ .

Eigenschaften der Dichte der Normalverteilung

Die Dichte f der $\mathcal{N}(\mu, \sigma^2)$ -Verteilung ist **symmetrisch** zu μ :

$$f(\mu - x) = f(\mu + x) \text{ für alle } x \in \mathbb{R}$$

f besitzt ein globales Maximum im Punkt $x = \mu$ und zwei Wendepunkte an den Stellen $x = \mu - \sigma$ und $x = \mu + \sigma$.



Standardisierte Zufallsvariable $Z = \frac{X - \mu}{\sigma}$

Ist die Zufallsvariable X gemäß $\mathcal{N}(\mu, \sigma^2)$ verteilt, so ist die **standardisierte Zufallsvariable**

$$Z = \frac{X - \mu}{\sigma}$$

gemäß $\mathcal{N}(0, 1)$ verteilt.

Verteilungsfunktion der Normalverteilung

Wie bei allen stetigen Zufallsvariablen gilt auch hier:

$$F(x) = \int_{-\infty}^x f(t) dt$$

Leider lässt sich dieses Integral für $f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$ nicht explizit als Funktion von x darstellen.

Die Werte der Verteilungsfunktion Φ der Standardnormalverteilung finden sich als Tabelle im Anhang aller Statistik-Bücher.

Tabelle mit Werten der Standardnormalverteilung Φ

Sei $Z \sim \mathcal{N}(0, 1)$ standardnormalverteilt.

	0.00	0.01	0.02	...
0.0	0.5000	0.5040	0.5080	...
0.1	0.5398	0.5438	0.5478	...
0.2	0.5793	0.5832	0.5871	...
0.3	0.6179	0.6217	0.6255	...
⋮	⋮	⋮	⋮	

Die Tabelle zeigt $\mathbb{P}(Z \leq x + y)$, wobei x die Zeile und y die Spalte bezeichnet.

Die Wahrscheinlichkeit, dass Z einen Wert annimmt, der kleiner oder gleich 0,22 ist, lautet **0,5871**, also $\Phi(0,22) = 58,71\%$.

Verteilungsfunktion für Normalverteilung $\mathcal{N}(\mu, \sigma^2)$

$F(x)$ gibt die W'keit an, dass X höchstens den Wert x hat:

$$F(x) = \mathbb{P}(X \leq x)$$

Wir formen nun die Ungleichung $X \leq x$ um:

$$X \leq x \quad \Leftrightarrow \quad X - \mu \leq x - \mu \quad \Leftrightarrow \quad \frac{X - \mu}{\sigma} \leq \frac{x - \mu}{\sigma}$$

Also gilt:

$$F(x) = \mathbb{P}\left(\frac{X - \mu}{\sigma} \leq \frac{x - \mu}{\sigma}\right)$$

Mit $Z = \frac{X - \mu}{\sigma}$ gilt demnach:

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

Funktionen von Zufallsvariablen

Funktionen $g(X)$ oder $g(X_1, X_2, \dots)$ von Zufallsvariablen X oder X_1, X_2, \dots sind selbst wiederum Zufallsvariablen.

Lineare Funktionen von normalverteilten Zufallsvariablen sind ebenfalls normalverteilt:

- ▶ Falls $X \sim \mathcal{N}$, dann gilt: $Y = a + bX \sim \mathcal{N}$.
- ▶ Die Summe normalverteilter Zufallsvariablen ist ebenfalls normalverteilt.
- ▶ Durch die lineare Transformation verändern sich aber μ und σ^2 , dazu später mehr.

Verteilung mehrdimensionaler Zufallsvariablen

Die gemeinsame Verteilungsfunktion

Ist (X_1, \dots, X_n) eine n -dimensionale Zufallsvariable, so heißt die Funktion F , die jedem Vektor (x_1, \dots, x_n) die Wahrscheinlichkeit

$$F(x_1, \dots, x_n) = \mathbb{P}(X_1 \leq x_1, \dots, X_n \leq x_n)$$

zuordnet, die **gemeinsame Verteilungsfunktion** von X_1, \dots, X_n .

Gemeinsame Wahrscheinlichkeitsfunktion von mehrdimensionalen diskreten Zufallsvariablen

$$\mathbb{P}(X_1 = x_1, \dots, X_n = x_n) = f(x_1, \dots, x_n)$$

Beispiel für $n = 2$:

y	1	2	3
x			
1	0,05	0,05	0,00
2	0,05	0,15	0,10
3	0,00	0,15	0,45

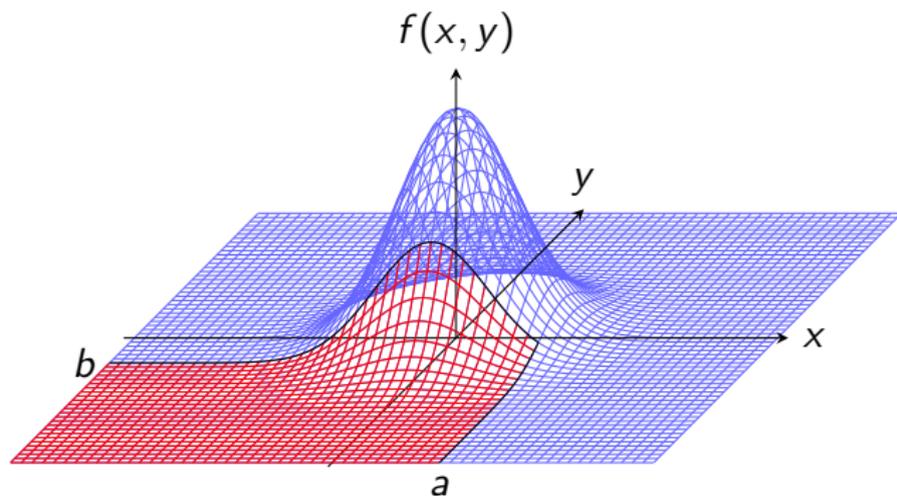
Die Wahrscheinlichkeit, dass $X = 2$ und $Y = 3$, beträgt $\mathbb{P}(X = 2, Y = 3) = 10\%$.

Gemeinsame Dichtefunktion von mehrdimensionalen stetigen Zufallsvariablen

Gemeinsam verteilte Zufallsvariablen X_1, \dots, X_n haben eine gemeinsame Dichtefunktion $f(x_1, \dots, x_n)$, sodass die gemeinsame Verteilungsfunktion F die folgende Gestalt hat:

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

Gemeinsame Dichte- und Verteilungsfunktion von zweidimensionalen stetigen Zufallsvariablen



$$\mathbb{P}(X \leq a, Y \leq b) = \int_{-\infty}^a \int_{-\infty}^b f(t_x, t_y) dt_y dt_x$$

Die Wahrscheinlichkeit, dass $X \leq a$ und $Y \leq b$, entspricht dem **Volumen** unter dem roten Bereich der Dichtefunktion f .

Randverteilungen diskreter Zufallsvariablen

Für eine gegebene gemeinsame Verteilungsfunktion $F(x, y)$ bei welcher X genau k Werte und Y genau m Werte annehmen kann, definieren wir zunächst die **Randwahrscheinlichkeiten**:

$$f_1(x) = \sum_{i=1}^m f(x, y_i) \text{ und } f_2(y) = \sum_{i=1}^k f(x_i, y)$$

und ermitteln die **Randverteilungen** $F_1(x)$ und $F_2(y)$ wie folgt:

$$F_1(x) = \sum_{x_i \leq x} f_1(x_i) \text{ und } F_2(y) = \sum_{y_i \leq y} f_2(y_i)$$

Beispiel diskrete Randverteilungen

Vorgehen: wie bei Kontingenztabelle und Randhäufigkeiten

	y	1	2	3	$f_1(x)$
x					
1		0,05	0,05	0,00	0,1
2		0,05	0,15	0,10	0,3
3		0,00	0,15	0,45	0,6
$f_2(y)$		0,1	0,35	0,55	1

$$F_1(x) = \begin{cases} 0 & \text{falls } x < 1 \\ 0,1 & \text{falls } 1 \leq x < 2 \\ 0,4 & \text{falls } 2 \leq x < 3 \\ 1 & \text{falls } 3 \leq x \end{cases}$$

$$F_2(y) = \begin{cases} 0 & \text{falls } y < 1 \\ 0,1 & \text{falls } 1 \leq y < 2 \\ 0,45 & \text{falls } 2 \leq y < 3 \\ 1 & \text{falls } 3 \leq y \end{cases}$$

Randverteilungen stetiger Zufallsvariablen

Für eine gegebene gemeinsame Verteilungsfunktion $F(x, y)$ definieren wir zuerst die **Randdichten**:

$$f_1(x) = \int_{-\infty}^{\infty} f(x, y) dy$$

$$f_2(y) = \int_{-\infty}^{\infty} f(x, y) dx$$

und ermitteln die **Randverteilungen** $F_1(x)$ und $F_2(y)$ wie folgt:

$$F_1(x) = \int_{-\infty}^x f_1(t) dt =: F(x, \infty)$$

$$F_2(y) = \int_{-\infty}^y f_2(t) dt =: F(\infty, y)$$

Bedingte Wahrscheinlichkeits- bzw. Dichtefunktionen

Für eine gegebene gemeinsame Wahrscheinlichkeits- bzw. Dichtefunktion $f(x, y)$ nennen wir für jedes feste y mit $f_2(y) \neq 0$ den Ausdruck

$$f_1(x|y) = \frac{f(x, y)}{f_2(y)}, \quad x \in \mathbb{R}$$

die **bedingte Wahrscheinlichkeits-** bzw. **Dichtefunktion** von X bei gegebener Realisierung y von Y .

Bedingte Wahrscheinlichkeits- bzw. Dichtefunktionen

Für eine gegebene gemeinsame Wahrscheinlichkeits- bzw. Dichtefunktion $f(x, y)$ nennen wir für jedes feste y mit $f_2(y) \neq 0$ den Ausdruck

$$f_1(x|y) = \frac{f(x, y)}{f_2(y)}, \quad x \in \mathbb{R}$$

die **bedingte Wahrscheinlichkeits-** bzw. **Dichtefunktion** von X bei gegebener Realisierung y von Y .

Ebenso für festes x mit $f_1(x) \neq 0$:

$$f_2(y|x) = \frac{f(x, y)}{f_1(x)}, \quad y \in \mathbb{R}$$

Bedingte Wahrscheinlichkeits- bzw. Dichtefunktion von Y bei gegebenem x von X .

Bedingte Wahrscheinlichkeitsfunktionen im Beispiel

y x	1	$f_1(x y=1)$	2	$f_1(x y=2)$	3	$f_1(x y=3)$
1	0,05	0,50	0,05	0,14	0,00	0,00
2	0,05	0,50	0,15	0,43	0,10	0,18
3	0,00	0,00	0,15	0,43	0,45	0,82
$f_2(y)$	0,10	$\sum = 1$	0,35	$\sum = 1$	0,55	$\sum = 1$

Unabhängigkeit von Zufallsvariablen

Für die Zufallsvariablen X_1, \dots, X_n sei $F(x_1, \dots, x_n)$ die gemeinsame Verteilungsfunktion und $f(x_1, \dots, x_n)$ die gemeinsame Wahrscheinlichkeitsfunktion bzw. Dichte.

Es bezeichnen $F_i(x_i)$ bzw. $f_i(x_i)$ die Randverteilungen bzw. die Randwahrscheinlichkeiten/-Dichten.

Die Zufallsvariablen X_1, \dots, X_n sind **unabhängig**, falls

$$F(x_1, \dots, x_n) = F_1(x_1) \cdot \dots \cdot F_n(x_n) \text{ für alle } x_1, \dots, x_n$$

bzw.

$$f(x_1, \dots, x_n) = f_1(x_1) \cdot \dots \cdot f_n(x_n) \text{ für alle } x_1, \dots, x_n$$

gilt.

Kapitel 9: Verteilungsparameter

Charakterisierung von Zufallsvariablen

Teil 1 „Deskriptive Statistik“:

Kennzahlen, welche die Lage und die Streuung von Stichproben beschreiben

Median, arithmetisches Mittel, mittlere quadratische Abweichung

Teil 2 „Wahrscheinlichkeitsrechnung“:

Kennzahlen, welche die Verteilung von Zufallsvariablen charakterisieren

Teil 3 „Inferenz Statistik“:

Schätzung der theoretischen Kennzahlen durch die korrespondierenden empirischen Kennzahlen

Definition Median und α -Quantil

Sei X eine eindimensionale Zufallsvariable.

Jeder Wert x mit

$$\mathbb{P}(X \geq x) \geq \frac{1}{2} \text{ und } \mathbb{P}(X \leq x) \geq \frac{1}{2}$$

heißt Median, x_{Med} .

Für ein α mit $0 < \alpha < 1$ heißt jeder Wert x_α mit

$$\mathbb{P}(X \geq x_\alpha) \geq 1 - \alpha \text{ und } \mathbb{P}(X \leq x_\alpha) \geq \alpha$$

α -Quantil.

Für jedes α mit $0 < \alpha < 1$ gibt es mindestens ein α -Quantil.

Für $\alpha = \frac{1}{2}$ ist jedes $x_{\frac{1}{2}}$ -Quantil ein Median.

Mehrere α -Quantile: Welches wählen wir aus?

In manchen Fällen gibt es mehrere α -Quantile.

Es gibt verschiedene Vorschläge, welches α -Quantil dann auszuwählen ist:

- ▶ Wähle das kleinste x , welches ein α -Quantil ist.
- ▶ Wähle das Mittel vom kleinsten und größten α -Quantil.

Falls in einer Klausuraufgabe nach einem α -Quantil gefragt wird und es mehrere α -Quantile gibt, so ist jedes α -Quantil eine richtige Antwort.

α -Quantile von stetigen Zufallsvariablen

Für jede stetige Zufallsvariable X gilt $\mathbb{P}(X = x) = 0$ für alle x und damit:

$$\mathbb{P}(X \geq x) = 1 - \mathbb{P}(X \leq x) = 1 - F(x)$$

Also folgt für jedes α mit $0 < \alpha < 1$ und dessen α -Quantil x_α :

$$\underbrace{1 - F(x_\alpha)}_{\mathbb{P}(X \geq x_\alpha)} \geq 1 - \alpha \quad \text{und} \quad \underbrace{F(x_\alpha)}_{\mathbb{P}(X \leq x_\alpha)} \geq \alpha$$

Dies ist äquivalent zu der einfacheren Bedingung:

$$F(x_\alpha) = \alpha$$

Falls die Dichtefunktion f positiv ist, so ist F streng monoton steigend und jedes α -Quantil ist eindeutig bestimmt.

Definition Erwartungswert

Sei X eine diskrete Zufallsvariable mit Wertebereich $\{x_1, x_2, \dots\}$ und Wahrscheinlichkeitsfunktion f .

Dann heißt

$$\mathbb{E}[X] = \sum_i x_i f(x_i)$$

der **Erwartungswert** von X .

Sei X eine stetige Zufallsvariable mit Dichtefunktion f .

Dann heißt

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

der **Erwartungswert** von X .

Rechenregeln für Erwartungswerte

Im Folgenden untersuchen wir die Auswirkungen auf den Erwartungswert durch:

- ▶ Symmetrie
- ▶ Lineare Transformationen
- ▶ Addition mehrerer Zufallsvariablen
- ▶ Unabhängigkeit von Zufallsvariablen

Erwartungswert bei Symmetrie

Ist die Wahrscheinlichkeits- bzw. Dichtefunktion f symmetrisch bezüglich eines Punktes $x = a$, so gilt:

$$\mathbb{E}[X] = a$$

Bei Symmetrie gilt $f(a - x) = f(a + x)$ für alle x .

Versuche das Resultat anhand dieser Eigenschaft selbst zu begründen.

Erwartungswerte von Funktionen von Zufallsvariablen

Funktionen von Zufallsvariablen sind wiederum Zufallsvariablen.

Wir können Erwartungswerte dieser transformierten Zufallsvariablen ausrechnen:

Sei X eine Zufallsvariable und $g : \mathbb{R} \rightarrow \mathbb{R}$. Dann gilt:

a) Wenn X diskret mit Wahrscheinlichkeitsfunktion f ist:

$$\mathbb{E}[g(X)] = \sum_{i \in \mathcal{I}} g(x_i) \cdot f(x_i)$$

b) Wenn X stetig mit Dichtefunktion f ist:

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

Erwartungswert bei linearer Transformation

Sei X eine Zufallsvariable und $g(X) = a + bX$ eine lineare Transformation von X .

Dann gilt:

$$\mathbb{E}[a + bX] = a + b\mathbb{E}[X]$$

Erwartungswert der Summe mehrerer Zufallsvariablen

Seien X_1, X_2, \dots, X_n Zufallsvariablen mit gemeinsamer Wahrscheinlichkeits- bzw. Dichtefunktion f .

Dann gilt:

$$\mathbb{E} \left[\sum_{i=1}^n X_i \right] = \sum_{i=1}^n \mathbb{E}[X_i]$$

Erwartungswert bei unabhängigen Zufallsvariablen

Seien X_1, X_2, \dots, X_n unabhängige Zufallsvariablen mit gemeinsamer Wahrscheinlichkeits- bzw. Dichtefunktion f .

Dann gilt:

$$\mathbb{E} \left[\prod_{i=1}^n X_i \right] = \prod_{i=1}^n \mathbb{E}[X_i]$$

Wiederholungen, arithmetisches Mittel und Erwartungswert

Mit „Erwartung“ ist Folgendes gemeint:

Bei unabhängiger Wiederholung des gleichen Zufallsexperimentes sollte zu erwarten sein, dass das arithmetische Mittel aller Ausgänge ungefähr dem Erwartungswert der Zufallsvariablen entspricht.

Je häufiger wiederholt wird, desto zuverlässiger kann erwartet werden.

Definition: Varianz und Standardabweichung

Sei X eine Zufallsvariable mit Wahrscheinlichkeits- bzw. Dichtefunktion f .

Dann heißt

$$\text{Var}(X) = \mathbb{E} [(X - \mathbb{E}[X])^2]$$

die **Varianz** von X .

Anstelle von $\text{Var}(X)$ schreiben wir kürzer σ^2 , wenn klar ist, welche Zufallsvariable gemeint ist.

Die positive Wurzel $\sqrt{\text{Var}(X)}$ heißt **Standardabweichung** von X (kürzer: σ).

Der Verschiebungssatz

$$\mathbb{E} [(X - \mathbb{E}[X])^2] = \mathbb{E} [X^2] - \mathbb{E}[X]^2$$

Die Varianz von linearen Transformationen

Sei X eine Zufallsvariable und $a + bX$ eine lineare Transformation.
Dann gilt:

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

Die Varianz der Summe von unabhängigen Zufallsvariablen

Seien X_1, X_2, \dots, X_n unabhängige Zufallsvariablen mit gemeinsamer Wahrscheinlichkeits- bzw. Dichtefunktion f .

Dann gilt:

$$\text{Var} \left(\sum_{i=1}^n X_i \right) = \sum_{i=1}^n \text{Var}(X_i)$$

Erwartungswerte und Varianzen wichtiger Verteilungen

Verteilung von X		$\mathbb{E}[X]$	$\text{Var}(X)$
Bernoulliverteilung	$B(1, p)$	p	$p \cdot (1 - p)$
Binomialverteilung	$B(n, p)$	$n \cdot p$	$n \cdot p \cdot (1 - p)$
Poissonverteilung	$P(\lambda)$	λ	λ
Gleichverteilung	$U(a, b)$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
Normalverteilung	$\mathcal{N}(\mu, \sigma^2)$	μ	σ^2

Weitere Aussagen über Erwartungswert und Varianz

Standardisierung

- ▶ Ist X eine Zufallsvariable mit $\mathbb{E}[X] = \mu$ und $\text{Var}(X) = \sigma^2$, so gilt für die standardisierte Zufallsvariable $Z = \frac{X - \mu}{\sigma}$:

$$\mathbb{E}[Z] = 0 \text{ und } \text{Var}(Z) = 1$$

Stichprobenmittel

- ▶ Sind X_1, \dots, X_n unabhängige und identische Zufallsvariablen mit $\mathbb{E}[X_i] = \mu$ und $\text{Var}(X_i) = \sigma^2$ für $i = 1, \dots, n$, so gilt für $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$:

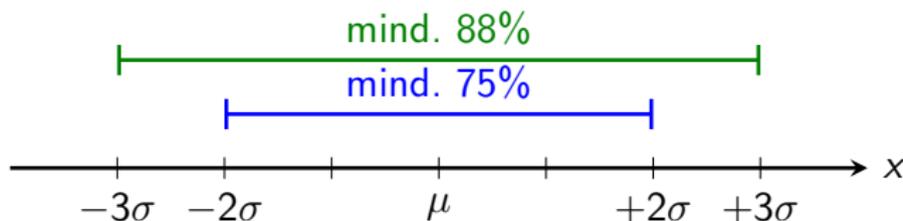
$$\mathbb{E}[\bar{X}] = \mu \text{ und } \text{Var}(\bar{X}) = \frac{1}{n} \sigma^2$$

Die Ungleichung von Tschebyscheff

$$\mathbb{P}(|X - \mathbb{E}[X]| \geq c) \leq \frac{\text{Var}(X)}{c^2} \text{ für alle } c > 0$$

Aus dieser Ungleichung folgen untere Schranken für die Wahrscheinlichkeiten der **$k\sigma$ -Bereiche**:

$$\mathbb{P}(\mu - k\sigma \leq X \leq \mu + k\sigma) \geq 1 - \frac{1}{k^2} = \begin{cases} \frac{3}{4} & \text{für } k = 2 \\ \frac{8}{9} & \text{für } k = 3 \end{cases}$$



Kovarianz und Korrelation zweier Zufallsvariablen

Definition:

Für zwei Zufallsvariablen X und Y heißt

$$\begin{aligned}\text{Cov}(X, Y) &= \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])] \\ &= \mathbb{E}[X \cdot Y] - \mathbb{E}[X] \cdot \mathbb{E}[Y]\end{aligned}$$

die **Kovarianz** von X und Y .

Falls $\text{Var}(X), \text{Var}(Y) \neq 0$, dann heißt

$$\rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)} \cdot \sqrt{\text{Var}(Y)}}$$

der **Korrelationskoeffizient** von X und Y .

Cov und ρ bei linearen Transformationen

Die Kovarianz und der Korrelationskoeffizient von X und Y messen die lineare Abhängigkeit von X und Y .

Ist Y eine lineare Transformation von X , also $Y = a + b \cdot X$, so gilt:

$$\text{Cov}(X, Y) = b\text{Var}(X)$$

und

$$\rho(X, Y) = \frac{b\text{Var}(X)}{\sqrt{\text{Var}(X)} \cdot \sqrt{b^2\text{Var}(X)}} = \frac{b}{|b|} = \begin{cases} 1 & \text{falls } b > 0 \\ -1 & \text{falls } b < 0 \end{cases}$$

Cov und ρ bei unabhängigen Zufallsvariablen

Sind die Zufallsvariablen X und Y unabhängig, so gilt:

$$\text{Cov}(X, Y) = 0$$

und

$$\rho(X, Y) = 0$$

Aus Unabhängigkeit folgt also Unkorreliertheit.

Der Umkehrschluss gilt im Allgemeinen aber nicht!

Für zwei normalverteilte Zufallsvariablen X und Y sind Unabhängigkeit und Unkorreliertheit äquivalent.

Die Varianz der Summe von korrelierten Zufallsvariablen

Seien X_1, \dots, X_n Zufallsvariablen und sei $a + \sum_{i=1}^n b_i X_i$ eine lineare Transformation dieser Zufallsvariablen. Dann gilt

$$\begin{aligned}\text{Var}\left(a + \sum_{i=1}^n b_i X_i\right) &= \sum_{i=1}^n \sum_{j=1}^n b_i b_j \text{Cov}(X_i, X_j) \\ &= \sum_{i=1}^n b_i^2 \text{Var}(X_i) + 2 \sum_{i=1}^n \sum_{j<i} b_i b_j \text{Cov}(X_i, X_j)\end{aligned}$$

Spezialfall:

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

Bedingte Erwartungen

Wir hatten zuvor die bedingten Wahrscheinlichkeits- bzw. Dichtefunktionen

$$f_1(x|y) = \frac{f(x, y)}{f_2(y)} \quad \text{und} \quad f_2(y|x) = \frac{f(x, y)}{f_1(x)}$$

definiert.

Der bedingte Erwartungswert einer Zufallsvariablen verwendet nun anstelle der normalen Wahrscheinlichkeits- bzw. Dichtefunktion die bedingte Wahrscheinlichkeits- bzw. Dichtefunktion.

Definition bedingter Erwartungswert

Seien X und Y zwei diskrete Zufallsvariablen mit der gemeinsamen Wahrscheinlichkeitsfunktion $f(x, y)$.

Dann ist der **bedingte Erwartungswert** von Y gegeben $X = x$ mit $f_1(x) > 0$ definiert als:

$$\mathbb{E}[Y|X = x] = \sum_i y_i \cdot f_2(y_i|x)$$

Seien X und Y zwei stetige Zufallsvariablen mit der gemeinsamen Dichtefunktion $f(x, y)$.

Dann ist der **bedingte Erwartungswert** von Y gegeben $X = x$ mit $f_1(x) > 0$ definiert als:

$$\mathbb{E}[Y|X = x] = \int_{-\infty}^{\infty} y \cdot f_2(y|x) dy$$

Kapitel 10:

Gesetz der großen Zahlen und zentraler Grenzwertsatz

Voraussetzungen

Gegeben seien identisch und unabhängig verteilte Zufallsvariablen X_1, \dots, X_n mit $\mathbb{E}[X_i] = \mu$ und $\text{Var}(X_i) = \sigma^2$ für $i = 1, \dots, n$.

In diesem Kapitel interessieren wir uns für das Stichprobenmittel

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$$

von sehr großen Stichproben, genauer für $n \rightarrow \infty$.

Das schwache Gesetz der großen Zahlen

$$\lim_{n \rightarrow \infty} \mathbb{P} (|\bar{X}_n - \mu| \leq \epsilon) = 1 \text{ für alle } \epsilon > 0$$

Zentraler Grenzwertsatz

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \leq x \right) = \Phi(x)$$

Interpretation des schwachen Gesetzes der großen Zahlen

$$\lim_{n \rightarrow \infty} \mathbb{P} (|\bar{X}_n - \mu| \leq \epsilon) = 1 \text{ für alle } \epsilon > 0$$

Bei wiederholten Zufallsvorgängen konvergiert die Wahrscheinlichkeit, dass das arithmetische Mittel aller Ergebnisse beliebig nahe am Erwartungswert liegt mit steigender Anzahl von Wiederholungen gegen 1.

Bemerkungen

- ▶ Folgt aus der Ungleichung von Tschebyscheff.
- ▶ Setzt keine Informationen über die Verteilung der X_i voraus.
- ▶ Dieses Gesetz gilt auch ohne Unabhängigkeit und ohne identische Verteilungen.

Bemerkungen zum zentralen Grenzwertsatz

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \leq x \right) = \Phi(x)$$

Beachte: $\mathbb{E}[\bar{X}_n] = \mu$ und $\text{Var}(\bar{X}_n) = \sigma^2/n$.

Die standardisierte Zufallsvariable $\frac{\bar{X}_n - \mathbb{E}[\bar{X}_n]}{\sqrt{\text{var}(\bar{X}_n)}}$ ist also standardnormalverteilt, wenn die Stichprobengröße n gegen unendlich strebt.

Der ZGS setzt keine Informationen über die Verteilung der X_i voraus.

Zusammenfassung Teil 2 Wahrscheinlichkeitsrechnung

- ▶ Zufallsvorgänge und Ereignisse
 - ▶ Elementarereignisse ω , Grundmenge Ω
 - ▶ Rechenoperationen \cap , \cup , \setminus
 - ▶ Partition, Ereignissystem
- ▶ Wahrscheinlichkeiten
 - ▶ klassischer & axiomatischer Wahrscheinlichkeitsbegriff
- ▶ Zufallsvariablen und Verteilungen
 - ▶ diskret und stetig
 - ▶ Unabhängigkeit
 - ▶ Wahrscheinlichkeits- bzw. Dichtefunktion f , Verteilungsfunktion F
- ▶ Verteilungsparameter
 - ▶ Erwartungswert
 - ▶ Varianz
 - ▶ Rechenregeln
- ▶ Gesetz der großen Zahlen und zentraler Grenzwertsatz