



Inferenzstatistik

BSc Psychologie SoSe 2021

Prof. Dr. Dirk Ostwald

(2) Wahrscheinlichkeitsräume und Zufallsvariablen

Literaturhinweise

Die in diesem Abschnitt vorgestellten Konzepte gehören zum Standard der Wahrscheinlichkeitstheorie. Zu Wahrscheinlichkeitsräumen wird die Lektüre von Georgii (2009, Abschnitte 1.1, 1.2, 3.1 und 3.3) empfohlen. Zu Zufallsvariablen wird die Lektüre von Georgii (2009, Abschnitte 1.3, 2.1, 2.6) empfohlen. In englischer Sprache finden sich entsprechende Darstellungen in Wasserman (2004, Sections 2.1 - 2.8) und DeGroot and Schervish (2012, Chapter 3).

Wahrscheinlichkeitsräume und Zufallsvariablen

- Wahrscheinlichkeitsräume
- Elementare Wahrscheinlichkeiten
- Konstruktion, Definition, und Notation von Zufallsvariablen
- Wahrscheinlichkeitsmassefunktionen
- Wahrscheinlichkeitsdichtefunktionen
- Kumulative Verteilungsfunktionen
- Selbstkontrollfragen

Wahrscheinlichkeitsräume und Zufallsvariablen

- **Wahrscheinlichkeitsräume**
- Elementare Wahrscheinlichkeiten
- Konstruktion, Definition, und Notation von Zufallsvariablen
- Wahrscheinlichkeitsmassefunktionen
- Wahrscheinlichkeitsdichtefunktionen
- Kumulative Verteilungsfunktionen
- Selbstkontrollfragen

Definition (Wahrscheinlichkeitsraum)

Ein *Wahrscheinlichkeitsraum* (*W-Raum*) ist ein Triple $(\Omega, \mathcal{A}, \mathbb{P})$, wobei

- Ω eine Menge *Ergebnismenge* ist, d.h. eine Menge von Elementarereignissen ω ,
- $\mathcal{A}$ eine σ -Algebra ist, d.h., $\mathcal{A}$ ist eine Menge mit den Eigenschaften
 - $\Omega \in \mathcal{A}$,
 - $\mathcal{A}$ ist abgeschlossen unter der Bildung von Komplementärereignissen, d.h., wenn $A \in \mathcal{A}$, dann ist auch $A^c := \Omega \setminus A \in \mathcal{A}$ für alle $A \in \mathcal{A}$,
 - $\mathcal{A}$ ist abgeschlossen unter der abzählbaren Vereinigung von Ereignissen, d.h., wenn $A_1, A_2, \dots \in \mathcal{A}$, dann folgt $\bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$, und
- $\mathbb{P}$ ein *Wahrscheinlichkeitsmaß* ist, d.h. eine Abbildung der Form $\mathbb{P} : \mathcal{A} \rightarrow [0, 1]$ mit
 - $\mathbb{P}(A) \geq 0$ für alle $A \in \mathcal{A}$ (Nicht-Negativität)
 - $\mathbb{P}(\Omega) = 1$ (Normiertheit)
 - $\mathbb{P}(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mathbb{P}(A_i)$ für paarweise disjunkte $A_i \in \mathcal{A}$ (σ -Additivität).

Wahrscheinlichkeitsräume

ERGEBNISSE DER MATHEMATIK UND IHRER GRENZGEBIETE

HERAUSGEGEBEN VON DER SCHRIFTLEITUNG
DES
„ZENTRALBLATT FÜR MATHEMATIK“
ZWEITER BAND

GRUNDBEGRIFFE DER WAHRSCHEINLICHKEITS- RECHNUNG

VON
A. KOLMOGOROFF



BERLIN
VERLAG VON JULIUS SPRINGER
1933

306/ 1. Terminologische Vorbemerkungen

erniedrigt. Ist jedoch die Anzahl der Behauptungen sehr groß, so lassen sich aus der praktischen Sicherheit jeder einzelnen dieser Behauptungen in bezug auf die Richtigkeit der simultanen Behauptung überhaupt keine Schlüsse ziehen. Deshalb folgt aus dem Prinzip A nach folgendem Weg, daß bei einer sehr großen Anzahl von Serien von Versuchen, von denen jede Serie aus n Versuchen besteht, in jeder Serie der Quotient n wie sich von $P(A)$ wenig unterscheiden wird.

Bemerkung II. Dem unmöglichen Ereignis (des leeren Menge) entspricht kraft unserer Axiome die Wahrscheinlichkeit $P(\emptyset) = 0$ *, während umgekehrt aus $P(A) = 0$ die Unmöglichkeit des Ereignisses A durchaus nicht zu folgen braucht; nach dem Prinzip B folgt aus dem Nullwerden der Wahrscheinlichkeit nur, daß bei einer einmaligen Realisation der Bedingungen $\bar{G}$ das Ereignis A praktisch unmöglich ist. Das bedeutet jedoch keineswegs, daß auch bei einer genügend langen Reihe von Versuchen das Ereignis A nicht auftreten wird. Andererseits kann man nach dem Prinzip A nur behaupten, daß bei $P(A) = 0$ und sehr großen n der Quotient n sehr klein wird (er kann z. B. gleich 1 sein).

§ 3. Terminologische Vorbemerkungen.

Wir haben die eigentlichen Objekte unserer weiteren Betrachtungen — die zufälligen Ereignisse — als Mengen definiert. Mehrere mengentheoretische Begriffe bezeichnet man aber in der Wahrscheinlichkeitsrechnung mit anderen Namen. Wir wollen hier ein kurzes Verzeichnis solcher Begriffe geben.

Mengentheoretisch.

1. A und B sind disjunkt, d. h.

$AB = \emptyset$.

2. $AB \dots N = \emptyset$.

3. $AB \dots N = X$.

4. $A + B + \dots + N = X$.

5. Die Komplementärmenge $\bar{A}$.

6. $A = \emptyset$.

7. $A = E$.

8. A ist unmöglich.

9. A muß notwendig vorkommen.

10. A ist unmöglich.

11. A muß notwendig vorkommen.

* Vgl. § 1. Formel (10).

Im Falle der zufälligen Ereignisse.

1. Die Ereignisse A und B sind

unvereinbar.

2. Die Ereignisse $A, B, \dots, N$

sind unvereinbar.

3. Das Ereignis X besteht in der

gleichzeitigen Realisation aller

Ereignisse $A, B, \dots, N$.

4. Das Ereignis X besteht in der Er-

scheinung mindestens eines un-

ter den Ereignissen $A, B, \dots, N$.

5. Das entgegengesetzte Ereignis

$\bar{A}$ besteht in der Nichter-

scheinung des Ereignisses A .

6. A ist unmöglich.

7. A muß notwendig vorkommen.

6 1. Die elementare Wahrscheinlichkeitsrechnung. /306

8. Ein System $\mathcal{H}$ der Mengen $A_1, A_2, \dots, A_n$ bildet eine Zerlegung der Menge E , wenn $A_1 + A_2 + \dots + A_n = E$ ist (das setzt bereits voraus, daß die Mengen A_i paarweise disjunkt sind).

9. B ist eine Untermenge von A : $B \subset A$.

§ 4. Unmittelbare Folgerungen aus den Axiomen, bedingte Wahrscheinlichkeiten, der Satz von BAYES.

Aus $A + A = E$ und den Axiomen IV und V folgt

(1) $P(A) + P(\bar{A}) = 1$.

(2) $P(\bar{A}) = 1 - P(A)$.

Da $E = \emptyset$ ist, erhält man insbesondere

(3) $P(\emptyset) = 0$.

Wenn $A, B, \dots, N$ unvereinbar sind, so folgt aus dem Axiom IV die Formel

(4) $P(A + B + \dots + N) = P(A) + P(B) + \dots + P(N)$

(der Additionssatz).

Wenn $P(A) > 0$ ist, so nennt man den Quotienten

(5) $P_A(B) = \frac{P(AB)}{P(A)}$

die bedingte Wahrscheinlichkeit des Ereignisses B unter der Bedingung A .

Aus (5) folgt unmittelbar

(6) $P(AB) = P(A)P_A(B)$.

Ein Induktionsschluß ergibt sodann die allgemeine Formel

(7) $P(A_1 A_2 \dots A_n) = P(A_1)P_{A_1}(A_2)P_{A_1 A_2}(A_3) \dots P_{A_1 A_2 \dots A_{n-1}}(A_n)$

(der Multiplikationssatz).

Man beweist auch leicht folgende Formeln:

(8) $P_A(B) \geq 0$.

(9) $P_A(B) \leq 1$.

(10) $P_A(B + C) = P_A(B) + P_A(C)$.

Vergleicht man diese Formeln (8) bis (10) mit den Axiomen III bis V, so ergibt sich, daß das Mengensystem $\mathcal{G}$ mit der Mengenfunktion $P_A(B)$

“Zweck des vorliegenden Heftes ist eine axiomatische Begründung der Wahrscheinlichkeitsrechnung. Der leitende Gedanke des Verfassers war dabei, die Grundbegriffe der Wahrscheinlichkeitsrechnung, welche noch unlängst für ganz eigenartig galten, natürlicherweise in die Reihe der allgemeinen Begriffsbildungen der modernen Mathematik* einzuordnen.” (*z.B. Mengenlehre, Abbildungen, Lebesgue Integral)

Kolmogorov, A.N. (1933) Grundbegriffe der Wahrscheinlichkeitsrechnung

Ausgewählte Bemerkungen

- W-Räume sind abstrakte Modelle von Zufallsexperimenten.
- W-Räume sind spezielle *Maßräume* $(\Omega, \mathcal{A}, \mu)$.
- Maßräume sind mathematische Modelle zur Beschreibung der Volumina von Mengen.
- Tuple von Ereignismengen und σ -Algebren heißen *Messräume*.
- W-Räume sind Messräume mit Wahrscheinlichkeitsmaßen.
- Elementarereignisse werden anhand von $\mathbb{P}(\{\omega\})$ “gezogen”.
- W-Räume vereinheitlichen endliche, abzählbare, und überabzählbare Ergebnismengen.
- W-Räume vereinheitlichen diskrete und kontinuierliche Wahrscheinlichkeit.
- σ -Algebren sind “vollständige Mengen von Ereignissen”.
- Für endliche Ereignismengen ist die typische σ -Algebra die Potenzmenge $\mathcal{P}(\Omega)$.
- Für die Ereignismenge $\mathbb{R}$ erfüllt die *Borel σ -algebra* die gleiche Funktion.
- Die W-theoretischen Charakteristiken von $(\Omega, \mathcal{A}, \mathbb{P})$ werden durch $\mathbb{P}$ festgelegt.
- W-Räume verschwinden oft hinter Zufallsvariablen und Verteilungen.

Beispiel (Würfel)

- Es ist sinnvoll, die Ereignismenge als $\Omega := \{1, 2, 3, 4, 5, 6\}$ zu definieren.
- Auch die Definition $\Omega := \{\cdot, \cdot, \cdot, \cdot, \cdot, \cdot, \cdot, \cdot, \cdot, \cdot, \cdot, \cdot\}$ wäre möglich.
- Die Potenzmenge $\mathcal{P}(\{1, 2, 3, 4, 5, 6\})$ enthält alle möglichen Ereignisse A_i , z.B.

Es fällt eine beliebige Augenzahl	$A_1 = \Omega$
Es fällt eine Augenzahl größer als 4	$A_2 = \{5, 6\}$
Es fällt eine gerade Augenzahl	$A_3 = \{2, 4, 6\}$
Es fällt eine Sechs	$A_4 = \{6\}$
Eins, drei, oder sechs fällt	$A_5 = \{1, 3, 6\}$
Es fällt keine Augenzahl	$A_6 = \emptyset$

- $\mathbb{P}$ kann durch Festlegung von $\mathbb{P}(\{\omega\})$ für alle $\omega \in \Omega$ definiert $\rightarrow$, die Wahrscheinlichkeiten aller anderen Ereignisse $A \in \mathcal{A}$ können dann mit der σ -Additivität $\mathbb{P}(A) = \sum_{\omega \in A} \mathbb{P}(\{\omega\})$ errechnet werden, weil die $\omega \in \mathcal{A}$ mit $\omega \in \Omega$ paarweise disjunkt sind.
- Im Modell eines unverfälschten Würfels wäre $\mathbb{P}(\{\omega\}) := 1/6$ für alle $\omega \in \Omega$.
- Für das Modell eines verfälschten Würfels könnte zum Beispiel definiert werden, dass $\mathbb{P}(\{1\}) = \mathbb{P}(\{2\}) = \mathbb{P}(\{6\}) := 1/9$, und $\mathbb{P}(\{3\}) = \mathbb{P}(\{4\}) = \mathbb{P}(\{5\}) := 2/9$.

Wahrscheinlichkeitsräume und Zufallsvariablen

- Wahrscheinlichkeitsräume
- **Elementare Wahrscheinlichkeiten**
- Konstruktion, Definition, und Notation von Zufallsvariablen
- Wahrscheinlichkeitsmassefunktionen
- Wahrscheinlichkeitsdichtefunktionen
- Kumulative Verteilungsfunktionen
- Selbstkontrollfragen

Definition (Wahrscheinlichkeitsmaß)

Es seien Ω eine Ereignismenge und $\mathcal{A}$ eine σ -Algebra auf Ω . Eine Abbildung

$$\mathbb{P} : \mathcal{A} \rightarrow [0, 1], A \mapsto \mathbb{P}(A) \quad (1)$$

mit den Eigenschaften

- (1) $\mathbb{P}(A) \geq 0$ für alle $A \in \mathcal{A}$,
- (2) $\mathbb{P}(\Omega) = 1$, und
- (3) wenn $A_1, A_2, \dots \in \mathcal{A}$ paarweise disjunkt sind, dann gilt $\mathbb{P}(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mathbb{P}(A_i)$.

heißt ein *Wahrscheinlichkeitsmaß* oder einfach *Wahrscheinlichkeit*.

Bemerkungen

- Für $A \in \mathcal{A}$ heißt $\mathbb{P}(A)$ die *(totale) Wahrscheinlichkeit von A*.
- Für $A, B \in \mathcal{A}$ heißt $\mathbb{P}(A \cap B)$ die *gemeinsame Wahrscheinlichkeit von A und B*.

Zur Interpretation des Wahrscheinlichkeitsbegriffs

- Die Interpretation des Wahrscheinlichkeitsbegriffs $\mathbb{P}(A)$ ist nicht eindeutig.
- Nach der *frequentistischen Interpretation* ist $\mathbb{P}(A)$ die idealisierte relative Häufigkeit, mit der das Ereignis A unter den gleichen äußeren Bedingungen eintreten pflegt.
- Nach der *Bayesianischen Interpretation* ist $\mathbb{P}(A)$ ist der Grad der Sicherheit, den ein Beobachter aufgrund seiner subjektiven Einschätzung der Lage dem Eintreten des Ereignisses A zumisst.
- Frequentistische und Bayesianische Interpretationen basieren mit der Wahrscheinlichkeitstheorie auf dem gleichen mathematischen Bezugssystem.

Theorem (Eigenschaften von Wahrscheinlichkeiten)

Es sei $(\Omega, \mathcal{A}, \mathbb{P})$ ein W-Raum und $A, B \in \mathcal{A}$. Dann gelten unter anderem

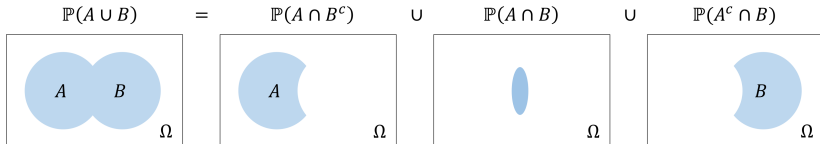
- (1) $\mathbb{P}(\emptyset) = 0$ und $0 \leq \mathbb{P}(A) \leq 1$.
- (2) $A \subset B \Rightarrow \mathbb{P}(A) \leq \mathbb{P}(B)$.
- (3) $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$.
- (4) $A \cap B = \emptyset \Rightarrow \mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$.
- (5) $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$.

Beweis von (5)

Mit der Additivität von $\mathbb{P}$ für disjunkte Ereignisse gilt

$$\begin{aligned}\mathbb{P}(A \cup B) &= \mathbb{P}(A \cap B^c) + \mathbb{P}(A \cap B) + \mathbb{P}(A^c \cap B) \\ &= \mathbb{P}(A \cap B^c) + \mathbb{P}(A \cap B) + \mathbb{P}(A^c \cap B) + \mathbb{P}(A \cap B) - \mathbb{P}(A \cap B) \\ &= \mathbb{P}((A \cap B^c) \cup (A \cap B)) + \mathbb{P}((A^c \cap B) \cup (A \cap B)) - \mathbb{P}(A \cap B) \\ &= \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)\end{aligned}\tag{2}$$

Zur Veranschaulichung der ersten Gleichung



□

Definition (Unabhängige Ereignisse)

Zwei Ereignisse $A \in \mathcal{A}$ and $B \in \mathcal{A}$ heißen (*stochastisch*) *unabhängig*, wenn

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B). \quad (3)$$

Eine Menge von Ereignissen $\{A_i | i \in I\} \subset \mathcal{A}$ mit beliebiger Indexmenge I heißt (*stochastisch*) *unabhängig*, wenn für jede endliche Untermenge $J \subseteq I$ gilt, dass

$$\mathbb{P}(\cap_{j \in J} A_j) = \prod_{j \in J} \mathbb{P}(A_j). \quad (4)$$

Bemerkungen

- Stochastische Unabhängigkeit modelliert die Abwesenheit von gegenseitigen Einflüssen.
- Die Unabhängigkeit bestimmter Ereignissen kann in der Definition eines probabilistischen Modells vorausgesetzt werden, so zum Beispiel die Unabhängigkeit von Fehlervariablen.
- Die Unabhängigkeit von Ereignissen kann aus der Definition eines probabilistischen Modells folgen.
- Disjunkte Ereignisse mit von Null verschiedener Wahrscheinlichkeit sind nie unabhängig:

$$\mathbb{P}(A)\mathbb{P}(B) > 0, \text{ aber } \mathbb{P}(A \cap B) = \mathbb{P}(\emptyset) = 0, \text{ also } \mathbb{P}(A \cap B) \neq \mathbb{P}(A)\mathbb{P}(B).$$

- Die Bedingung der beliebigen Untermengen von I sichert die paarweise unabhängig der $A_i, i \in I$.

Definition (Bedingte Wahrscheinlichkeit)

Es sei $(\Omega, \mathcal{A}, \mathbb{P})$ ein W-Raum und $A, B \in \mathcal{A}$ mit $\mathbb{P}(B) > 0$. Die *bedingte Wahrscheinlichkeit des Ereignisses $A \in \mathcal{A}$ gegeben Ereignis $B \in \mathcal{A}$* ist definiert als

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}. \quad (5)$$

Für jedes feste B , ist $\mathbb{P}(\cdot|B)$ ein Wahrscheinlichkeitsmaß, d.h. $\mathbb{P}(\cdot|B) \geq 0$, $\mathbb{P}(\Omega|B) = 1$, und für paarweise disjunkte $A_1, A_2, \dots \in \mathcal{A}$ gilt $\mathbb{P}(\cup_{i=1}^{\infty} A_i|B) = \sum_{i=1}^{\infty} \mathbb{P}(A_i|B)$.

Anmerkungen

- $\mathbb{P}(A|B)$ ist die normalisierte Version der gemeinsamen Wahrscheinlichkeit $\mathbb{P}(A \cap B)$.
- Die Festlegung der gemeinsamen Wahrscheinlichkeit $\mathbb{P}(A \cap B)$ legt $\mathbb{P}(A|B)$ fest.
- Die Rechengeln der Wahrscheinlichkeit gelten für die Ereignisse links des Strichs.
- Üblicherweise gilt $\mathbb{P}(A|B) \neq \mathbb{P}(B|A)$, z.B. $\mathbb{P}(\text{Tod}|\text{Erhängen}) \neq \mathbb{P}(\text{Erhängen}|\text{Tod})$.
- Eine Verallgemeinerung für $\mathbb{P}(B) = 0$ ist möglich, aber technisch aufwändig.

Beispiel (Bedingte Wahrscheinlichkeit)

Wir betrachten das Wahrscheinlichkeitsraum Modell des fairen Würfels und wollen die bedingte Wahrscheinlichkeit für das Ereignis “Es fällt eine gerade Augenzahl” gegeben das Ereignis “Es fällt eine Zahl größer als 3” berechnen. Wir halten zunächst fest, dass das Ereignis “Es fällt eine gerade Augenzahl” der Teilmenge $A := \{2, 4, 6\}$ von Ω entspricht, und dass das Ereignis “Es fällt eine Zahl größer als 3” der Teilmenge $B := \{4, 5, 6\}$ von Ω entspricht. Wir halten dann fest, dass unter der Annahme des Modell des fairen Würfels gilt, dass

$$\mathbb{P}(\{2, 4, 6\}) = \mathbb{P}(\{2\}) + \mathbb{P}(\{4\}) + \mathbb{P}(\{6\}) = 1/6 + 1/6 + 1/6 = 3/6 \quad (6)$$

$$\mathbb{P}(\{4, 5, 6\}) = \mathbb{P}(\{4\}) + \mathbb{P}(\{5\}) + \mathbb{P}(\{6\}) = 1/6 + 1/6 + 1/6 = 3/6. \quad (7)$$

Schließlich halten wir fest, dass das Ereignis $A \cap B$, also das Ereignis “Es fällt eine gerade Augenzahl, die größer als 3 ist”, die Wahrscheinlichkeit

$$\mathbb{P}(A \cap B) = \mathbb{P}(\{2, 4, 6\} \cap \{4, 5, 6\}) = \mathbb{P}(\{4, 6\}) = \mathbb{P}(\{4\}) + \mathbb{P}(\{6\}) = 1/6 + 1/6 = 2/6 \quad (8)$$

hat. Nach Definition der bedingten Wahrscheinlichkeit ergibt sich dann

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(\{4, 6\})}{\mathbb{P}(\{4, 5, 6\})} = \frac{2/6}{3/6} = 2/6 \cdot 6/3 = 2/3. \quad (9)$$

Wenn man also weiß, dass eine Augenzahl größer als 3 gefallen ist, ist die Wahrscheinlichkeit, dass es sich dabei um eine gerade Augenzahl handelt $2/3$. Wenn man ersteres nicht weiß, ist die Wahrscheinlichkeit für das Fallen einer geraden Augenzahl (nur) $1/2$.

Theorem (Bedingte Wahrscheinlichkeit unter Unabhängigkeit)

Es seien $(\Omega, \mathcal{A}, \mathbb{P})$ ein W-Raum und $A, B \in \mathcal{A}$ unabhängige Ereignisse mit $\mathbb{P}(\cdot|B) \geq 0$. Dann gilt

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(A)\mathbb{P}(B)}{\mathbb{P}(B)} = \mathbb{P}(A). \quad (10)$$

Bemerkungen

- Bei Unabhängigkeit ändert das Wissen um das Eintreten von B die Wahrscheinlichkeit des Eintretens von A nicht.
- Unabhängigkeit wird als $\mathbb{P}(A)\mathbb{P}(B)$ modelliert, damit $\mathbb{P}(A|B) = \mathbb{P}(A)$ gilt.

Wahrscheinlichkeitsräume und Zufallsvariablen

- Wahrscheinlichkeitsräume
- Elementare Wahrscheinlichkeiten
- **Konstruktion, Definition, und Notation von Zufallsvariablen**
- Wahrscheinlichkeitsmassefunktionen
- Wahrscheinlichkeitsdichtefunktionen
- Kumulative Verteilungsfunktionen
- Selbstkontrollfragen

Konstruktion von Zufallsvariablen und Verteilungen

- Es seien $(\Omega, \mathcal{A}, \mathbb{P})$ ein W-Raum und $X : \Omega \rightarrow \mathcal{X}$ eine Abbildung.
- Es sei $\mathcal{S}$ eine σ -Algebra auf $\mathcal{X}$.
- Für jedes $S \in \mathcal{S}$ sei das *Urbild von S* definiert als

$$X^{-1}(S) := \{\omega \in \Omega \mid X(\omega) \in S\}. \quad (11)$$

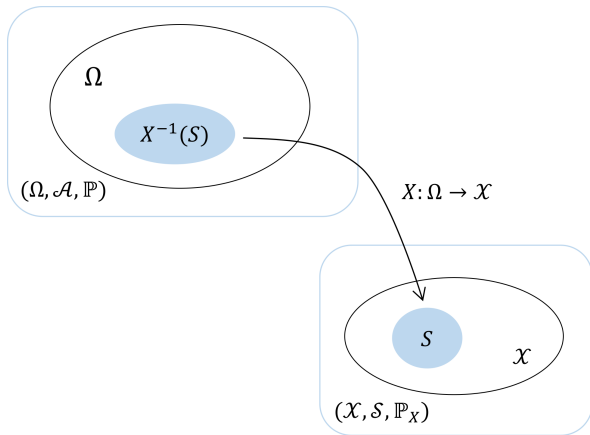
- Wenn $X^{-1}(S) \in \mathcal{A}$ für alle $S \in \mathcal{S}$ gilt, dann heißt X *messbar*.
- $X : \Omega \rightarrow \mathcal{X}$ sei messbar. Allen $S \in \mathcal{S}$ kann die Wahrscheinlichkeit

$$\mathbb{P}_X : \mathcal{S} \rightarrow [0, 1], S \mapsto \mathbb{P}_X(S) := \mathbb{P}(X^{-1}(S)) = \mathbb{P}(\{\omega \in \Omega \mid X(\omega) \in S\}) \quad (12)$$

zugeordnet werden.

- X heißt nun *Zufallsvariable* und $\mathbb{P}_X$ heißt *Bildmaß* oder *Verteilung von X* .
- $(\mathcal{X}, \mathcal{S}, \mathbb{P}_X)$ ist ein W-Raum.
- Mit $\mathcal{X} := \mathbb{R}$ und $\mathcal{S} := \mathcal{B}$ rückt der W-Raum $(\mathbb{R}, \mathcal{B}, \mathbb{P}_X)$ ins Zentrum der Betrachtung.

Konstruktion von Zufallsvariablen und Verteilungen



$$\mathbb{P}(X^{-1}(S)) = \mathbb{P}(\{\omega \in \Omega | X(\omega) \in S\}) =: \mathbb{P}_X(S)$$

Definition (Zufallsvariable)

Es sei $(\Omega, \mathcal{A}, \mathbb{P})$ ein W-Raum und $(\mathcal{X}, \mathcal{S})$ ein *Messraum*. Eine *Zufallsvariable* (ZV) ist dann definiert als eine Abbildung $X : \Omega \rightarrow \mathcal{X}$ mit der *Messbarkeitseigenschaft*

$$\{\omega \in \Omega | X(\omega) \in S\} \in \mathcal{A} \text{ für alle } S \in \mathcal{S}. \quad (13)$$

Bemerkungen

- ZVen sind weder “zufällig” noch “Variablen”.
- Intuitiv wird $\omega \in \Omega$ “zufällig” anhand von $\mathbb{P}$ gezogen und $X(\omega)$ realisiert.
- Wir nennen $\mathcal{X}$ den *Ergebnisraum* der ZV X .
- Die Verteilungen (Bildmaße) von ZVen sind in der Statistik zentral.
- Der Begriff der Verteilung wird oft auch für W-Maße und Dichten verwendet.

Beispiel (Summe zweier Würfel)

- Für das Werfen zweier Würfel ist ein sinnvolles Wahrscheinlichkeitsraum-Modell
 - $\Omega := \{(d_1, d_2) | d_1 \in \mathbb{N}_6, d_2 \in \mathbb{N}_6\}$
 - $\mathcal{A} := \mathcal{P}(\Omega)$.
 - $\mathbb{P} : \mathcal{A} \rightarrow [0, 1]$ mit $\mathbb{P}(\{(d_1, d_2)\}) = 1/36$ für alle $(d_1, d_2) \in \Omega$.
- Die Augenzahl-Summenbildung wird dann sinnvoller Weise durch die Zufallsvariable

$$X : \Omega \mapsto \mathcal{X}, (d_1, d_2) \mapsto X((d_1, d_2)) := d_1 + d_2. \quad (14)$$

beschrieben, wobei $\mathcal{X} := \{2, 3, \dots, 12\}$.

- $\mathcal{S} := \mathcal{P}(\mathcal{X})$ ist eine sinnvolle σ -Algebra auf $\mathcal{X}$.
- Mithilfe der σ -Additivität von $\mathbb{P}$ können wir die Verteilung $\mathbb{P}_X$ von X für alle Elementarereignisse $\{x\} \in \mathcal{S}$ berechnen, wie auf der nächsten Folie gezeigt.
- Wir haben damit ein weiteres Wahrscheinlichkeitsraum-Modell $(\mathcal{X}, \mathcal{S}, \mathbb{P}_X)$ konstruiert.

Beispiel (Summe zweier Würfel)

Bestimmung der Verteilung $\mathbb{P}_X$ von X

$$\mathbb{P}_X(\{2\}) = \mathbb{P}(X^{-1}(\{2\})) = \mathbb{P}(\{(1, 1)\}) = \frac{1}{36}$$

$$\mathbb{P}_X(\{3\}) = \mathbb{P}(X^{-1}(\{3\})) = \mathbb{P}(\{(1, 2), (2, 1)\}) = \frac{2}{36}$$

$$\mathbb{P}_X(\{4\}) = \mathbb{P}(X^{-1}(\{4\})) = \mathbb{P}(\{(1, 3), (3, 1), (2, 2)\}) = \frac{3}{36}$$

$$\mathbb{P}_X(\{5\}) = \mathbb{P}(X^{-1}(\{5\})) = \mathbb{P}(\{(1, 4), (4, 1), (2, 3), (3, 2)\}) = \frac{4}{36}$$

$$\mathbb{P}_X(\{6\}) = \mathbb{P}(X^{-1}(\{6\})) = \mathbb{P}(\{(1, 5), (5, 1), (2, 4), (4, 2), (3, 3)\}) = \frac{5}{36}$$

$$\mathbb{P}_X(\{7\}) = \mathbb{P}(X^{-1}(\{7\})) = \mathbb{P}(\{(1, 6), (6, 1), (2, 5), (5, 2), (3, 4), (4, 3)\}) = \frac{6}{36}$$

$$\mathbb{P}_X(\{8\}) = \mathbb{P}(X^{-1}(\{8\})) = \mathbb{P}(\{(2, 6), (6, 2), (3, 5), (5, 3), (4, 4)\}) = \frac{5}{36}$$

$$\mathbb{P}_X(\{9\}) = \mathbb{P}(X^{-1}(\{9\})) = \mathbb{P}(\{(3, 6), (6, 3), (4, 5), (5, 4)\}) = \frac{4}{36}$$

$$\mathbb{P}_X(\{10\}) = \mathbb{P}(X^{-1}(\{10\})) = \mathbb{P}(\{(4, 6), (6, 4), (5, 5)\}) = \frac{3}{36}$$

$$\mathbb{P}_X(\{11\}) = \mathbb{P}(X^{-1}(\{11\})) = \mathbb{P}(\{(5, 6), (6, 5)\}) = \frac{2}{36}$$

$$\mathbb{P}_X(\{12\}) = \mathbb{P}(X^{-1}(\{12\})) = \mathbb{P}(\{(6, 6)\}) = \frac{1}{36}$$

Definition (Notation für ZVen)

Es seien $(\Omega, \mathcal{A}, \mathbb{P})$ und $(\mathcal{X}, \mathcal{S}, \mathbb{P}_X)$ W-Räume und $X : \Omega \rightarrow \mathcal{X}$ sei eine ZV. Dann gelten folgende Konventionen:

$$\{X \in S\} := \{\omega \in \Omega | X(\omega) \in S\}, S \subset \mathcal{X},$$

$$\{X = x\} := \{\omega \in \Omega | X(\omega) = x\}, x \in \mathcal{X},$$

$$\{X \leq x\} := \{\omega \in \Omega | X(\omega) \leq x\}, x \in \mathcal{X},$$

$$\{X < x\} := \{\omega \in \Omega | X(\omega) < x\}, x \in \mathcal{X}.$$

Aus diesen Konventionen folgen exemplarisch die folgenden Konventionen für Verteilungen:

$$\mathbb{P}_X(X \in S) = \mathbb{P}(\{X \in S\}) = \mathbb{P}(\{\omega \in \Omega | X(\omega) \in S\}), S \subset \mathcal{X}$$

$$\mathbb{P}_X(X \leq x) = \mathbb{P}(\{X \leq x\}) = \mathbb{P}(\{\omega \in \Omega | X(\omega) \leq x\}), x \in \mathcal{X}.$$

Oft wird zudem auf das ZV Subskript bei Verteilungssymbolen verzichtet, z.B.

$$\mathbb{P}(X \in S) = \mathbb{P}_X(X \in S), S \subset \mathcal{X},$$

$$\mathbb{P}(X \leq x) = \mathbb{P}_X(X \leq x), x \in \mathcal{X}.$$

Zum Zusammenhang von W-Maßen, ZVen, und Dichten

- Verteilungen von Zufallsvariablen sind in der Statistik zentral.
- Verteilungen von Zufallsvariablen werden üblicherweise mithilfe von WMFen und WDFen definiert.
- Welche Rolle spielen dann Wahrscheinlichkeitsräume und Wahrscheinlichkeitsmaße?
- Wahrscheinlichkeitsmaße $\Leftrightarrow$ Verteilungen von Zufallsvariable

$\mathbb{P}$ sei ein Wahrscheinlichkeitsmaß auf der Borel σ -Algebra.

$\Leftrightarrow (\Omega, \mathcal{F}, \mathbb{P}) := (\mathbb{R}, \mathcal{B}, \mathbb{P})$ ist ein Wahrscheinlichkeitsraum

$\Leftrightarrow X : \Omega \rightarrow \mathbb{R}, \omega \mapsto X(\omega) := \omega$ ist eine Zufallsvariable

$\Leftrightarrow \mathbb{P}_X = \mathbb{P}$.

$\Leftrightarrow \mathbb{P}$ ist die Verteilung einer Zufallsvariable.

Wahrscheinlichkeitsräume und Zufallsvariablen

- Wahrscheinlichkeitsräume
- Elementare Wahrscheinlichkeiten
- Konstruktion, Definition, und Notation von Zufallsvariablen
- **Wahrscheinlichkeitsmassefunktionen**
- Wahrscheinlichkeitsdichtefunktionen
- Kumulative Verteilungsfunktionen
- Selbstkontrollfragen

Definition (Diskrete ZV, Wahrscheinlichkeitsmassefunktion)

Eine ZV X heißt *diskret*, wenn ihr Ergebnisraum $\mathcal{X}$ endlich oder abzählbar ist. Die *Wahrscheinlichkeitsmassefunktion (WMF)* einer diskreten ZV ist definiert als

$$p : \mathcal{X} \rightarrow [0, 1], x \mapsto p(x) := \mathbb{P}(X = x). \quad (15)$$

Bemerkungen

- Eine Menge heißt abzählbar, wenn sie bijektiv auf $\mathbb{N}$ abgebildet werden kann.
- WMFen sind nicht-negativ, d.h. $p(x) \geq 0$ für alle $x \in \mathcal{X}$.
- WMFen sind normiert, d.h. $\sum_{x \in \mathcal{X}} p(x) = 1$.
- WMFen heißen im Deutschen üblicherweise *W-Funktionen* oder *Zähldichten*.
- WMFen heißen auf Englisch *probability mass functions (PMFs)*.
- Zur Parallelität mit PMFs und WDFs bevorzugt wird den WMF Begriff.

Definition (Bernoulli-Zufallsvariable)

Es sei X eine Zufallsvariable mit Ergebnisraum $\mathcal{X} = \{0, 1\}$ und WMF

$$p : \mathcal{X} \rightarrow [0, 1], x \mapsto p(x) := \mu^x(1 - \mu)^{1-x} \text{ for } \mu \in [0, 1]. \quad (16)$$

Dann sagen wir, dass X einer *Bernoulli-Verteilung mit Parameter* $\mu \in [0, 1]$ unterliegt und nennen X eine *Bernoulli-Zufallsvariable*. Wir kürzen dies mit $X \sim \text{Bern}(\mu)$ ab. Die WMF einer Bernoulli-Zufallsvariable bezeichnen wir mit

$$\text{Bern}(x; \mu) := \mu^x(1 - \mu)^{1-x}. \quad (17)$$

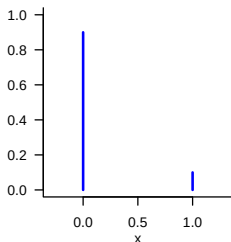
Bemerkungen

- Eine Bernoulli-Zufallsvariable kann als Modell eines Münzwurfs dienen.
- μ ist die Wahrscheinlichkeit dafür, dass X den Wert 1 annimmt,

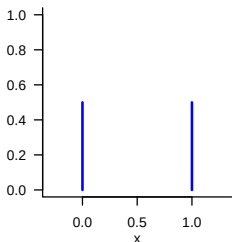
$$\mathbb{P}(X = 1) = \mu^1(1 - \mu)^{1-1} = \mu. \quad (18)$$

Beispiel (Bernoulli-Verteilungen)

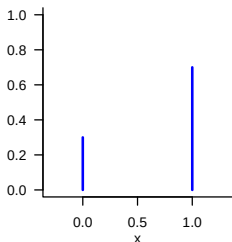
Bern($x;0.1$)



Bern($x;0.5$)



Bern($x;0.7$)



Definition (Binomial-Zufallsvariable)

Es sei X eine Zufallsvariable mit Ergebnisraum $\mathcal{X} := \mathbb{N}_n^0$ und WMF

$$p : \mathcal{X} \rightarrow [0, 1], x \mapsto p(x) := \binom{n}{x} \mu^x (1 - \mu)^{n-x} \text{ for } \mu \in [0, 1]. \quad (19)$$

Dann sagen wir, dass X einer *Binomialverteilung mit Parametern* $\mu \in [0, 1]$ *und* $n \in \mathbb{N}$ *unterliegt* und nennen X eine Binomial-Zufallsvariable. Wir kürzen dies mit $X \sim \text{Bin}(\mu, n)$ ab. Die WMF einer Binomial-Zufallsvariable bezeichnen wir mit

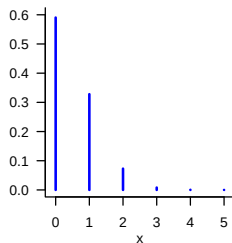
$$\text{Bin}(x; \mu, n) := \binom{n}{x} \mu^x (1 - \mu)^{n-x}. \quad (20)$$

Bemerkung

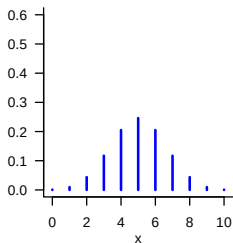
- Es gilt $\text{Bin}(x; \mu, 1) = \text{Bern}(x; \mu)$.

Beispiel (Binomial-Verteilungen)

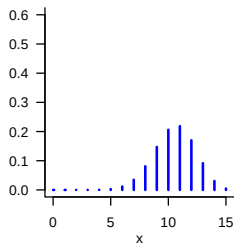
$\text{Bin}(x;5,0.1)$



$\text{Bin}(x;10,0.5)$



$\text{Bin}(x;15,0.7)$



Definition (Diskret-gleichverteilte Zufallsvariable)

Es sei X eine diskrete Zufallsvariable mit Ergebnisraum $\mathcal{X}$ und WMF

$$p : \mathcal{X} \rightarrow \mathbb{R}_{\geq 0}, x \mapsto p(x) := \frac{1}{|\mathcal{X}|}. \quad (21)$$

Dann sagen wir, dass X einer *diskreten Gleichverteilung* unterliegt und nennen X eine *diskret-gleichverteilte Zufallsvariable*. Wir kürzen dies mit $X \sim U(|\mathcal{X}|)$ ab. Die WMF einer diskret-gleichverteilten Zufallsvariable bezeichnen wir mit

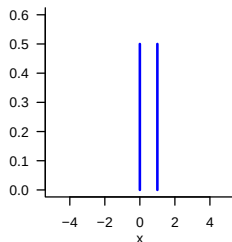
$$U(x; |\mathcal{X}|) := \frac{1}{|\mathcal{X}|}. \quad (22)$$

Bemerkungen

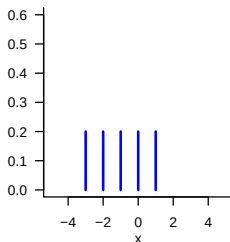
- $\text{Bern}(x; 0.5) = U(x; |\mathcal{X}|)$ für $\mathcal{X} = \{0, 1\}$.
- Es gilt zum Beispiel $\text{Bin}(x; 0.5) = U(x; |\mathcal{X}|)$ für $\mathcal{X} = \{0, 1\}$.

Beispiel (Diskrete Gleichverteilungen)

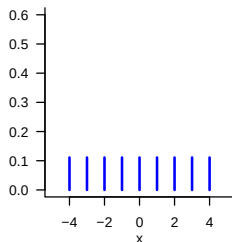
$U(x; X = \{0,1\})$



$U(x; X = \{-3,-2,-1,0,1\})$



$U(x; X = \{-4,-3,-2,-1,0,1,2,3,4\})$



Wahrscheinlichkeitsräume und Zufallsvariablen

- Wahrscheinlichkeitsräume
- Elementare Wahrscheinlichkeiten
- Konstruktion, Definition, und Notation von Zufallsvariablen
- Wahrscheinlichkeitsmassefunktionen
- **Wahrscheinlichkeitsdichtefunktionen**
- Kumulative Verteilungsfunktionen
- Selbstkontrollfragen

Definition (Kontinuierliche ZV, Wahrscheinlichkeitsdichtefunktion)

Eine Zufallsvariable X heißt *kontinuierlich*, wenn eine Funktion

$$p : \mathbb{R} \rightarrow \mathbb{R}_{\geq 0}, x \mapsto p(x) \quad (23)$$

existiert, so dass gilt

- $p(x) \geq 0$ für alle $x \in \mathbb{R}$,
- $\int_{-\infty}^{\infty} p(x) dx = 1$,
- $\mathbb{P}(a \leq X \leq b) = \int_a^b p(x) dx$ für alle $a, b \in \mathbb{R}, a \leq b$.

Eine entsprechende Funktion p heißt *Wahrscheinlichkeitsdichtefunktion (WDF)* von X .

Bemerkungen

- WDFen können Werte größer als 1 annehmen.
- Es gilt $\mathbb{P}(X = a) = \int_a^a p(x) dx = 0$.
- Wahrscheinlichkeiten werden aus WDFen durch Integration berechnet.
- (Wahrscheinlichkeits)Masse = (Wahrscheinlichkeits)Dichte $\times$ (Mengen)Volumen.

Definition (Normalverteilte und standardnormalverteilte Zufallsvariablen)

Es sei X eine Zufallsvariable mit Ergebnisraum $\mathbb{R}$ und WDF

$$p : \mathbb{R} \rightarrow \mathbb{R}_{>0}, x \mapsto p(x) := \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}(x - \mu)^2\right). \quad (24)$$

Dann sagen wir, dass X einer *Normalverteilung (oder Gauß-Verteilung)* mit Parametern $\mu \in \mathbb{R}$ und $\sigma^2 > 0$ unterliegt und nennen X eine *normalverteilte Zufallsvariable*. Wir kürzen dies mit $X \sim N(\mu, \sigma^2)$ ab. Die WDF einer normalverteilten Zufallsvariable bezeichnen wir mit

$$N(x; \mu, \sigma^2) := \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}(x - \mu)^2\right). \quad (25)$$

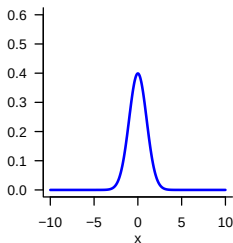
Eine normalverteilte Zufallsvariable mit $\mu = 0$ und $\sigma^2 = 1$ heißt *standardnormalverteilte Zufallsvariable* und wird oft als *Z-Zufallsvariable* bezeichnet.

Bemerkungen

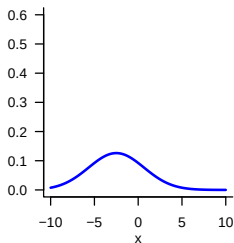
- Der Parameter μ entspricht dem Wert höchster Wahrscheinlichkeitsdichte.
- Der Parameter σ^2 spezifiziert die Breite der WDF.

Beispiel (Normalverteilungen)

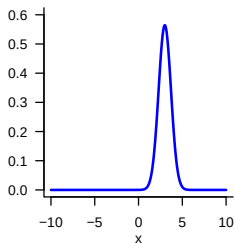
$N(x; 0,1)$



$N(x; -2.5,10)$



$N(x; 3,0.5)$



Definition (Gamma Zufallsvariable)

Es sei X eine Zufallsvariable mit Ergebnisraum $\mathcal{X} := \mathbb{R}_{>0}$ und WDF

$$p : \mathbb{R}_{>0} \rightarrow \mathbb{R}_{>0}, x \mapsto p(x) := \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} \exp\left(-\frac{x}{\beta}\right), \quad (26)$$

wobei Γ die Gammafunktion bezeichne. Dann sagen wir, dass X einer *Gammaverteilung mit Formparameter $\alpha > 0$ und Skalenparameter $\beta > 0$* unterliegt und nennen X eine *gammaverteilte Zufallsvariable*. Wir kürzen dies mit $X \sim G(\alpha, \beta)$ ab. Die WDF einer gammaverteilten Zufallsvariable bezeichnen wir mit

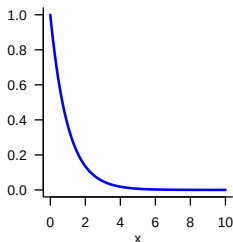
$$G(x; \alpha, \beta) := \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} \exp\left(-\frac{x}{\beta}\right). \quad (27)$$

Bemerkung

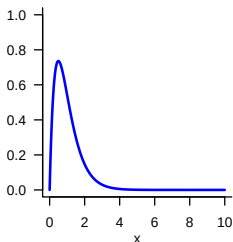
- $G\left(\frac{n}{2}, 2\right)$ heißt auch *Chi-Quadrat (χ^2) Verteilung mit n Freiheitsgraden*.

Beispiel (Gamma Verteilungen)

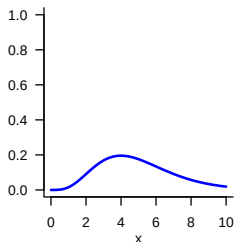
$G(x; 1,1)$



$G(x; 2,2)$



$G(x; 5,1)$



Definition (Beta-Zufallsvariable)

Es sei X eine Zufallsvariable mit Ergebnisraum $\mathcal{X} := [0, 1]$ und WDF

$$p : \mathcal{X} \rightarrow [0, 1], x \mapsto p(x) := \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1} (1-x)^{\beta-1} \text{ for } \alpha, \beta \in \mathbb{R}_{>0}, \quad (28)$$

wobei Γ die Gammafunktion bezeichne. Dann sagen wir, dass X einer *Beta-Verteilung* mit Parametern $\alpha > 0$ und $\beta > 0$ unterliegt, und nennen X eine beta-verteilte Zufallsvariable. Wir kürzen dies mit $X \sim \text{Beta}(\alpha, \beta)$ ab. Die WDF einer beta-verteilten Zufallsvariable bezeichnen wir mit

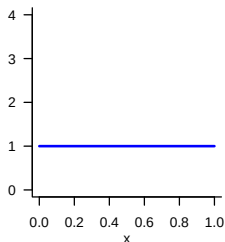
$$\text{Beta}(x; \alpha, \beta) := \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1} (1-x)^{\beta-1}. \quad (29)$$

Bemerkung

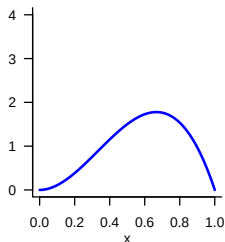
- Für $\alpha < 1, \beta < 1$ ist der Ergebnisraum $\mathcal{X} :=]0, 1[$.

Beispiel (Beta-Verteilungen)

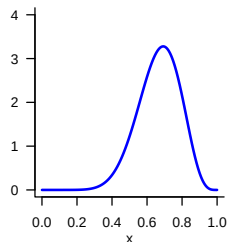
Beta(x; 1,1)



Beta(x; 3,2)



Beta(x; 10,5)



Definition (Gleichverteilte Zufallsvariable)

Es sei X eine kontinuierliche Zufallsvariable mit Ergebnisraum $\mathbb{R}$ und WDF

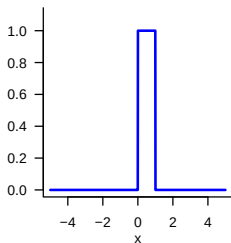
$$p : \mathbb{R} \rightarrow \mathbb{R}_{\geq 0}, x \mapsto p(x) := \begin{cases} \frac{1}{b-a} & x \in [a, b] \\ 0 & x \notin [a, b] \end{cases}. \quad (30)$$

Dann sagen wir, dass X einer *Gleichverteilung mit Parametern a und b* unterliegt und nennen X eine *gleichverteilte Zufallsvariable*. Wir kürzen dies mit $X \sim U(a, b)$ ab. Die WDF einer gleichverteilten Zufallsvariable bezeichnen wir mit

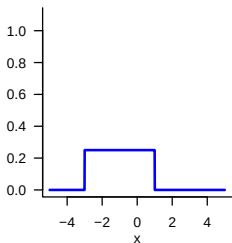
$$U(x; a, b) := \frac{1}{b-a}. \quad (31)$$

Beispiel (Gleichverteilung)

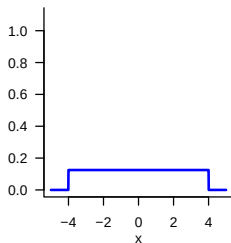
$U(x; 0,1)$



$U(x; -3,1)$



$U(x; -4,4)$



Wahrscheinlichkeitsräume und Zufallsvariablen

- Wahrscheinlichkeitsräume
- Elementare Wahrscheinlichkeiten
- Konstruktion, Definition, und Notation von Zufallsvariablen
- Wahrscheinlichkeitsmassefunktionen
- Wahrscheinlichkeitsdichtefunktionen
- **Kumulative Verteilungsfunktionen**
- Selbstkontrollfragen

Definition (Kumulative Verteilungsfunktion)

Die *kumulative Verteilungsfunktion (KVF)* einer Zufallsvariable X ist definiert als

$$P : \mathbb{R} \rightarrow [0, 1], x \mapsto P(x) := \mathbb{P}(X \leq x). \quad (32)$$

Bemerkungen

- KVFe sind sowohl für diskrete als auch kontinuierliche ZVen definiert.
- $P(x)$ ist für jedes $x \in \mathbb{R}$ definiert, auch wenn $x \notin \mathcal{X}$.
- Mithilfe von KVFe können Intervallwahrscheinlichkeiten angegeben werden

Theorem (Überschreitungswahrscheinlichkeit)

Es sei X eine Zufallsvariable mit Ereignisraum $\mathcal{X}$ und P ihre kumulative Verteilungsfunktion. Dann gilt für die *Überschreitungswahrscheinlichkeit* $\mathbb{P}(X > x)$, dass

$$\mathbb{P}(X > x) = 1 - P(x) \text{ für alle } x \in \mathcal{X}. \quad (33)$$

Beweis

Die Ereignisse $\{X > x\}$ und $\{X \leq x\}$ sind disjunkt und

$$\Omega = \{\omega \in \Omega | X(\omega) > x\} \cup \{\omega \in \Omega | X(\omega) \leq x\} = \{X > x\} \cup \{X \leq x\}. \quad (34)$$

Mit der σ -Additivität von $\mathbb{P}$ folgt dann

$$\begin{aligned} \mathbb{P}(\Omega) &= 1 \\ \Leftrightarrow \mathbb{P}(\{X > x\} \cup \{X \leq x\}) &= 1 \\ \Leftrightarrow \mathbb{P}(\{X > x\}) + \mathbb{P}(\{X \leq x\}) &= 1 \\ \Leftrightarrow \mathbb{P}(\{X > x\}) &= 1 - \mathbb{P}(\{X \leq x\}) \\ \Leftrightarrow \mathbb{P}(\{X > x\}) &= 1 - P(x). \end{aligned} \quad (35)$$

□

Theorem (Intervallwahrscheinlichkeiten)

Es sei X eine Zufallsvariable mit Ereignisraum $\mathcal{X}$ und P ihre kumulative Verteilungsfunktion. Dann gilt für die *Intervallwahrscheinlichkeit* $\mathbb{P}(X \in]x_1, x_2])$, dass

$$\mathbb{P}(X \in]x_1, x_2]) = P(x_2) - P(x_1) \text{ für alle } x_1, x_2 \in \mathcal{X} \text{ mit } x_1 < x_2. \quad (36)$$

Beweis

Wir betrachten die Ereignisse $\{X \leq x_1\}$, $\{x_1 < X \leq x_2\}$ und $\{X \leq x_2\}$, wobei

$$\{X \leq x_1\} \cap \{x_1 < X \leq x_2\} = \emptyset \text{ und } \{X \leq x_1\} \cup \{x_1 < X \leq x_2\} = \{X \leq x_2\}. \quad (37)$$

gelten. Mit der σ -Additivität von $\mathbb{P}$ gilt dann

$$\begin{aligned} \mathbb{P}(\{X \leq x_1\} \cup \{x_1 < X \leq x_2\}) &= \mathbb{P}(\{X \leq x_2\}) \\ \Leftrightarrow \mathbb{P}(\{X \leq x_1\}) + \mathbb{P}(\{x_1 < X \leq x_2\}) &= \mathbb{P}(\{X \leq x_2\}) \\ \Leftrightarrow \mathbb{P}(\{x_1 < X \leq x_2\}) &= \mathbb{P}(\{X \leq x_2\}) - \mathbb{P}(\{X \leq x_1\}) \\ \Leftrightarrow \mathbb{P}(\{x_1 < X \leq x_2\}) &= P(x_2) - P(x_1) \\ \Leftrightarrow \mathbb{P}(X \in]x_1, x_2]) &= P(x_2) - P(x_1). \end{aligned} \quad (38)$$

□

Theorem (Eigenschaften von kumulative Verteilungsfunktionen)

Es sei X eine Zufallsvariable und P ihre kumulative Verteilungsfunktion. Dann hat P die folgenden Eigenschaften

- (1) P ist *monoton steigend*, i.e., wenn $x_1 < x_2$, dann gilt $P(x_1) \leq P(x_2)$.
- (2) $\lim_{x \rightarrow -\infty} P(x) = 0$ und $\lim_{x \rightarrow \infty} P(x) = 1$.
- (3) P ist *rechtsseitig stetig*, d.h., $P(x) = P(x^+) = \lim_{y \rightarrow x, y > x} P(y)$ für alle $x \in \mathbb{R}$

Bemerkungen

- Die genannten Eigenschaften können auch zur Definition einer KVF genutzt werden.
- (3) $\Leftrightarrow$ Eine KVF hat keine Sprünge, wenn man sich Grenzpunkten von rechts nähert.

Beweis

(1) Wir halten zunächst fest, dass für Ereignisse $A \subset B$ gilt, dass $\mathbb{P}(A) \leq \mathbb{P}(B)$. Wie halten dann fest, dass für $x_1 < x_2$,

$$\{X \leq x_1\} = \{\omega \in \Omega | X(\omega) \leq x_1\} \subset \{\omega \in \Omega | X(\omega) \leq x_2\} = \{X \leq x_2\} \quad (39)$$

Also gilt

$$\mathbb{P}(\{X \leq x_1\}) \leq \mathbb{P}\{X \leq x_2\} \Rightarrow P(x_1) \leq P(x_2). \quad (40)$$

(2) Wir verzichten auf einen Beweis

(3) Wir definieren

$$P(x^+) = \lim_{y \rightarrow x, y > x} P(y). \quad (41)$$

Seien nun $y_1 > y_2 > \dots$ so, dass $\lim_{n \rightarrow \infty} y_n = x$. Dann gilt

$$\{X \leq x\} = \cap_{n=1}^{\infty} \{X \leq y_n\}. \quad (42)$$

Es gilt also

$$P(x) = \mathbb{P}(\{X \leq x\}) = \mathbb{P}(\cap_{n=1}^{\infty} \{X \leq y_n\}) = \lim_{n \rightarrow \infty} \mathbb{P}(\{X \leq y_n\}) = P(x^+), \quad (43)$$

wobei wir die dritte Gleichung unbegründet stehen lassen.

Kumulative Verteilungsfunktionen von diskreten Zufallsvariablen

- Wenn $a < b$ und $\mathbb{P}(a < X < b) = 0$, dann ist P konstant horizontal auf $]a, b[$.
- An jedem Punkt x mit $\mathbb{P}(X = x) > 0$ springt die KVF um den Betrag $\mathbb{P}(X = x)$.
- $\Leftrightarrow$ An jedem Punkt x mit $p(x) > 0$ springt die KVF um den Betrag $p(x) > 0$.
- Generell ist die KVF einer diskreten Zufallsvariable mit Ergebnisraum $\mathbb{N}_0$ durch

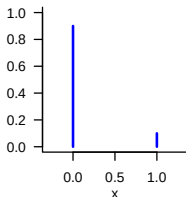
$$P : \mathbb{R} \rightarrow [0, 1], x \mapsto P(x) := \sum_{k=0}^{\lfloor x \rfloor} \mathbb{P}(X = k) \quad (44)$$

gegeben, wobei $\lfloor x \rfloor$ die Abrundungsfunktion bezeichnet.

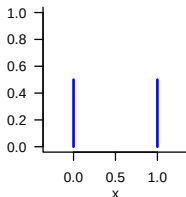
Kumulative Verteilungsfunktionen

Beispiel (Bernoulliverteilungen)

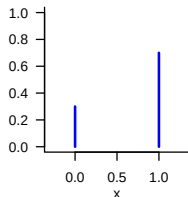
WMF von $\text{Bern}(x;0.1)$



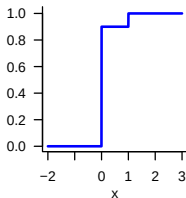
WMF von $\text{Bern}(x;0.5)$



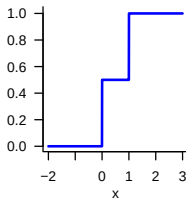
WMF von $\text{Bern}(x;0.7)$



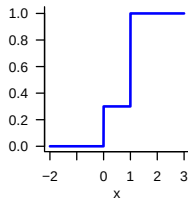
KVF von $\text{Bern}(x;0.1)$



KVF von $\text{Bern}(x;0.5)$



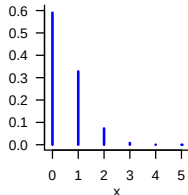
KVF von $\text{Bern}(x;0.7)$



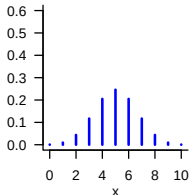
Kumulative Verteilungsfunktionen

Beispiel (Binomialverteilungen)

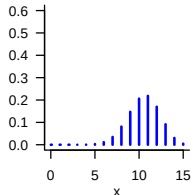
WMF von $\text{Bin}(x;5,0.1)$



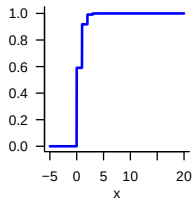
WMF von $\text{Bin}(x;10,0.5)$



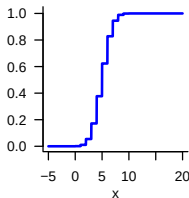
WMF von $\text{Bin}(x;15,0.7)$



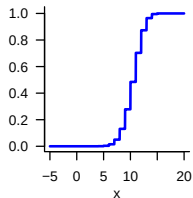
KVF von $\text{Bin}(x;5,0.1)$



KVF von $\text{Bin}(x;10,0.5)$



KVF von $\text{Bin}(x;15,0.7)$



Theorem (Kumulative Verteilungsfunktionen von kontinuierlichen ZVen)

X sei eine kontinuierliche Zufallsvariable mit WDF p und KVF P . Dann gilt

$$P(x) = \int_{-\infty}^x p(t) dt \text{ und } p(x) = \frac{d}{dx} P(x). \quad (45)$$

Beweis

Wir halten zunächst fest, dass weil $\mathbb{P}(X = x) = 0$ für alle $x \in \mathbb{R}$ gilt, die KVF von X keine Sprünge hat, d.h. P ist stetig. Mit der Definitionen von WDF und KVF, folgt, dass P die Form einer Stammfunktion von p hat. Dass p die Ableitung von P ist, folgt dann unmittelbar aus dem Fundamentalsatz der Analysis.

□

Bemerkungen

- Die KVF ist **eine** Stammfunktion der WDF, die WDF ist die Ableitung der KVF.
- Das *Theorem von Radon-Nikodym* ist eine generalisierte Variante dieser Einsicht.
- KVFe von kontinuierlichen ZV heißen auch kumulative Dichtefunktionen (KDFen).

Beispiel (Normalverteilung)

Es sei $X \sim N(\mu, \sigma^2)$.

- Die WDF von X ist

$$p : \mathbb{R} \rightarrow \mathbb{R}_{>0}, x \mapsto p(x) := \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}(x - \mu)^2\right).$$

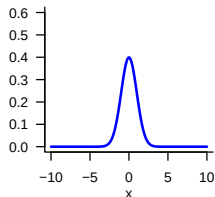
- Die KVF von X ist

$$P : \mathbb{R} \rightarrow]0, 1[, x \mapsto P(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^x \exp\left(-\frac{1}{2\sigma^2}(\xi - \mu)^2\right) d\xi.$$

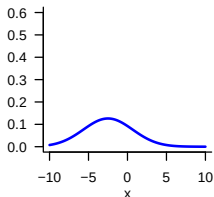
- Die KVF von $X \sim N(\mu, \sigma^2)$ kann nur numerisch, nicht analytisch, berechnet werden.
- Für $\mu = 1, \sigma^2 = 1$, gilt zum Beispiel $p(2) = 0.24$ und $P(2) = 0.84$.
- Die WDF und KVF von $Z \sim N(0, 1)$ werden oft mit ϕ und Φ , respektive, bezeichnet.

Beispiel (Normalverteilung)

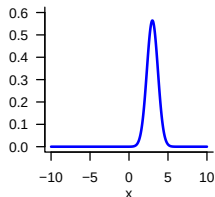
WDF von $N(0,1)$



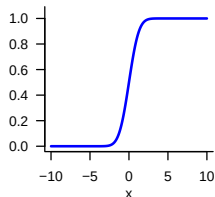
WDF von $N(-2.5,10)$



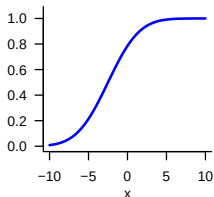
WDF von $N(3,0.5)$



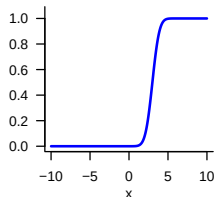
KDF von $N(0,1)$



KDF von $N(-2.5,10)$



KDF von $N(3,0.5)$



Definition (Inverse Kumulative Verteilungsfunktion)

X sei eine kontinuierliche Zufallsvariable mit KVF P . Dann heißt die Funktion

$$P^{-1} :]0, 1[\rightarrow \mathbb{R}, q \mapsto P^{-1}(q) := \{x \in \mathbb{R} | P(x) = q\} \quad (46)$$

die *inverse kumulative Verteilungsfunktion* von X .

Bemerkungen

- P^{-1} ist die Inverse von P , d.h. $P^{-1}(P(x)) = x$.
- Offenbar gilt $P(x) = q \Leftrightarrow \mathbb{P}(X \leq x) = q$.
- Für $q \in]0, 1[$ ist also $P^{-1}(q)$ der Wert x von X , so dass $\mathbb{P}(X \leq x) = q$ gilt.
- Wenn $Z \sim N(0, 1)$ mit KVF Φ ist, dann gilt zum Beispiel $\Phi^{-1}(0.975) = 1.960$.

Beispiel (Normalverteilung)

Es sei $X \sim N(\mu, \sigma^2)$.

- Die KVF von X ist

$$P : \mathbb{R} \rightarrow]0, 1[, x \mapsto \mathbb{P}(X \leq x) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^x \exp\left(-\frac{1}{2\sigma^2}(\xi - \mu)^2\right) d\xi \quad (47)$$

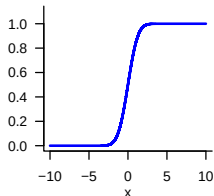
- Die inverse KVF von X ist

$$P^{-1} :]0, 1[\rightarrow \mathbb{R}, q \mapsto P^{-1}(q) = \{x \in \mathbb{R} | P(x) = q\}. \quad (48)$$

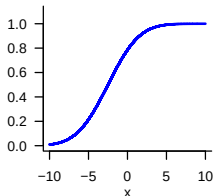
- Für $\mu = 1, \sigma^2 = 1$ gilt z.B., dass $P(2) = 0.84$ und $P^{-1}(0.84) = 2$.
- Die inverse KVF von $X \sim N(0, 1)$ wird oft mit Φ^{-1} bezeichnet.
- Typische Beispielwerte für die KVF und inverse KVF von $N(0, 1)$ sind
 - $\Phi(1.645) = 0.950$, $\Phi^{-1}(0.950) = \Phi^{-1}(1 - 0.050) = 1.640$.
 - $\Phi(1.960) = 0.975$, $\Phi^{-1}(0.975) = \Phi^{-1}(1 - \frac{0.050}{2})$.

Beispiel (Normalverteilung)

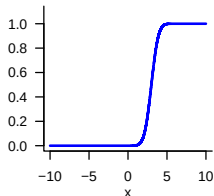
KDF von $N(0,1)$



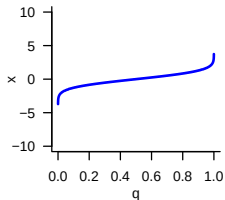
KDF von $N(-2.5,10)$



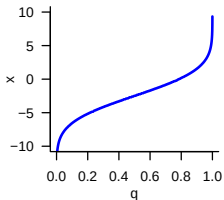
KDF von $N(3,0.5)$



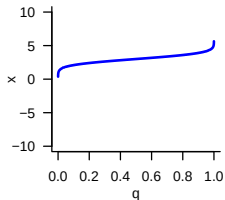
Inverse KDF von $N(0,1)$



Inverse KDF von $N(-2.5,10)$



Inverse KDF von $N(3,0.5)$



Wahrscheinlichkeitsräume und Zufallsvariablen

- Wahrscheinlichkeitsräume
- Elementare Wahrscheinlichkeiten
- Konstruktion, Definition, und Notation von Zufallsvariablen
- Wahrscheinlichkeitsmassefunktionen
- Wahrscheinlichkeitsdichtefunktionen
- Kumulative Verteilungsfunktionen
- **Selbstkontrollfragen**

1. Definieren Sie den Begriff der σ -Algebra
2. Definieren Sie den Begriff des Wahrscheinlichkeitsmaßes.
3. Definieren Sie den Begriff des Wahrscheinlichkeitsraumes
4. Skizzieren Sie das Wahrscheinlichkeitsraummodell eines fairen Würfels.
5. Geben sie zwei Interpretationen der Wahrscheinlichkeit eines Ereignisses.
6. Definieren Sie den Begriff der Unabhängigkeit zweier Ereignisse.
7. Definieren Sie die bedingte Wahrscheinlichkeit eines Ereignisses.
8. Geben Sie das Gesetz von der totalen Wahrscheinlichkeit wieder.
9. Geben Sie das Theorem zur bedingten Wahrscheinlichkeit unter Unabhängigkeit wieder.
10. Definieren Sie den Begriff der Zufallsvariable.

11. Definieren Sie den Begriff der Wahrscheinlichkeitsmassefunktion.
12. Definieren Sie die Begriffe der Wahrscheinlichkeitsdichtefunktion.
13. Definieren Sie den Begriff der kumulativen Verteilungsfunktion.
14. Schreiben sie die Intervallwahrscheinlichkeit einer Zufallsvariable mithilfe ihrer KVF.
15. Definieren Sie die WDF und KVF einer normalverteilten Zufallsvariable.
16. Schreiben Sie den Wert $P(x)$ der KVF einer Zufallsvariable mithilfe ihrer **WDF**.
17. Schreiben Sie den Wert $p(x)$ der WDF einer Zufallsvariable mithilfe ihrer KVF.
18. Nennen sie drei Eigenschaften von kumulativen Verteilungsfunktionen.
19. Definieren Sie den Begriff der inversen Verteilungsfunktion.

Literatur

- DeGroot, M. H. and Schervish, M. J. (2012). *Probability and Statistics*. Addison-Wesley, Boston, 4th ed edition.
- Georgii, H.-O. (2009). *Stochastik: Einführung in die Wahrscheinlichkeitstheorie und Statistik*. DeGruyter-Lehrbuch. de Gruyter, Berlin, 4., überarb. und erw. aufl edition.
- Wasserman, L. (2004). *All of Statistics*.