Inhaltsverzeichnis

(Ausschnitt)

Messbare Abbildungen I

- Häufig von Interesse:
  Nicht der Ausgang eines Zufallsexperiments selbst, sondern eine **Funktion** dieses Ausgangs, d.h. eine Abbildung $X : \Omega \to \Omega'$ vom Ergebnisraum Ω in eine andere Menge Ω'.
- Beispiele:
  - Augensumme beim gleichzeitigen Werfen von zwei Würfeln
  - Anzahl „Wappen" bei mehrmaligem Münzwurf
  - Resultierender Gewinn zu gegebener Tippreihe beim Lottospiel
  - Anzahl der weißen Kugeln bei wiederholter Ziehung von Kugeln aus Urne mit schwarzen und weißen Kugeln (mit oder ohne Zurücklegen)
- Wahrscheinlichkeitsbegriff des ursprünglichen Zufallsexperiments $(\Omega, \mathcal{F}, P)$ soll in die Wertemenge/Bildmenge Ω' der Abbildung X „transportiert" werden.

Messbare Abbildungen II

- Bestimmung der Wahrscheinlichkeit eines Ereignisses $A' \subseteq \Omega'$ wie folgt:
  - Feststellen, welche Elemente $\omega \in \Omega$ unter der Abbildung X auf Elemente $\omega' \in A'$ abgebildet werden.
  - Wahrscheinlichkeit von A' ergibt sich dann als Wahrscheinlichkeit der erhaltenen Teilmenge $\{\omega \in \Omega \mid X(\omega) \in A'\}$ von Ω, also als $P(\{\omega \in \Omega \mid X(\omega) \in A'\})$.

Definition 8.1 (Urbild)

Es seien $X : \Omega \to \Omega'$ eine Abbildung, $A' \subseteq \Omega'$. Dann heißt

$$X^{-1}(A') := \{\omega \in \Omega \mid X(\omega) \in A'\}$$

das **Urbild** *von A' bzgl. X.*

Messbare Abbildungen III

- Zur Funktionsfähigkeit des „Wahrscheinlichkeitstransports" nötig: Geeignetes Mengensystem ($\rightsquigarrow$ σ-Algebra) $\mathcal{F}'$ über Ω', welches „kompatibel" zur σ-Algebra $\mathcal{F}$ über Ω und zur Abbildung X ist.
- *„Kompatibilitätsanforderung" ergibt sich nach Konstruktion:* Damit $P(\{\omega \in \Omega \,|\, X(\omega) \in A'\}) = P(X^{-1}(A'))$ für $A' \in \mathcal{F}'$ definiert ist, muss $X^{-1}(A') \in \mathcal{F}$ gelten für alle $A' \in \mathcal{F}'$.
- Beschriebene Eigenschaft der Kombination $\mathcal{F}$, $\mathcal{F}'$ und X heißt **Messbarkeit**:

Definition 8.2 (Messbarkeit, messbare Abbildung)

*Es seien $(\Omega, \mathcal{F})$ und $(\Omega', \mathcal{F}')$ zwei Messräume. Eine Abbildung $X : \Omega \to \Omega'$ heißt $\mathcal{F} - \mathcal{F}'-$**messbar**, wenn gilt:*

$$X^{-1}(A') \in \mathcal{F} \text{ für alle } A' \in \mathcal{F}' .$$

Bildwahrscheinlichkeit

Definition 8.3 (Bildwahrscheinlichkeit)

Es seien $(\Omega, \mathcal{F}, P)$ *ein Wahrscheinlichkeitsraum,* $(\Omega', \mathcal{F}')$ *ein Messraum,* X *eine* $\mathcal{F} - \mathcal{F}'-$*messbare Abbildung. Dann heißt das durch*

$$P_X : \mathcal{F}' \to \mathbb{R}; P_X(A') := P(X^{-1}(A'))$$

definierte Wahrscheinlichkeitsmaß P_X **Bildwahrscheinlichkeit** *von* P *bezüglich* X*.* $(\Omega', \mathcal{F}', P_X)$ *ist damit ebenfalls ein Wahrscheinlichkeitsraum.*

- Gilt $\mathcal{F} = \mathcal{P}(\Omega)$, so ist offensichtlich jede Abbildung $X : \Omega \to \Omega'$ $\mathcal{F} - \mathcal{F}'-$messbar für beliebige σ-Algebren $\mathcal{F}'$ über Ω'.

Einbettung der deskriptiven Statistik in die Wahrscheinlichkeitsrechnung

- Ist Ω die (endliche) Menge von Merkmalsträgern einer deskriptiven statistischen Untersuchung, $\mathcal{F} = \mathcal{P}(\Omega)$ und P die Laplace-Wahrscheinlichkeit

  $$P : \mathcal{P}(\Omega) \to \mathbb{R}; B \mapsto \frac{\#B}{\#\Omega} ,$$

  so kann jedes Merkmal X mit Merkmalsraum $A = \{a_1, \ldots, a_m\}$ als $\mathcal{P}(\Omega) - \mathcal{P}(A)-$messbare Abbildung $X : \Omega \to A$ verstanden werden.
- $(A, \mathcal{P}(A), P_X)$ ist dann ein diskreter Wahrscheinlichkeitsraum mit Wahrscheinlichkeitsfunktion $p(a_j) = r(a_j)$ bzw. — äquivalent — $P_X(\{a_j\}) = r(a_j)$ für $j \in \{1, \ldots, m\}$.
- Durch $(A, \mathcal{P}(A), P_X)$ wird damit die Erhebung des Merkmalswerts eines rein zufällig (gleichwahrscheinlich) ausgewählten Merkmalsträgers modelliert.

Inhaltsverzeichnis

(Ausschnitt)

Borelsche σ-Algebra I

- Häufiger Wertebereich von Abbildungen aus Wahrscheinlichkeitsräumen: $\mathbb{R}$
- $\mathcal{P}(\mathbb{R})$ als σ-Algebra (aus technischen Gründen) aber ungeeignet!
- Alternative σ-Algebra über $\mathbb{R}$: „Borelsche" σ-Algebra $\mathcal{B}$
- $\mathcal{B}$ ist die kleinste σ-Algebra über $\mathbb{R}$, die alle Intervalle

$$\begin{array}{cccc} (a,b) & (a,b] & [a,b) & [a,b] \\ (-\infty,a) & (-\infty,a] & (a,\infty) & [a,\infty) \end{array}$$

für $a, b \in \mathbb{R}$ (mit $a < b$) enthält.

Borelsche σ-Algebra II

- Aufgrund der Eigenschaften von σ-Algebren sind auch alle
  - Einpunktmengen $\{x\}$ für $x \in \mathbb{R}$,
  - endliche Mengen $\{x_1, \ldots, x_m\} \subseteq \mathbb{R}$,
  - abzählbar unendliche Mengen sowie endliche und abzählbare Schnitte und Vereinigungen von Intervallen

  in $\mathcal{B}$ enthalten.
- Abbildungen $X : \Omega \to \mathbb{R}$ aus einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$ heißen (eindimensionale) **Zufallsvariablen**, wenn sie $\mathcal{F} - \mathcal{B} -$messbar sind:

Eindimensionale Zufallsvariablen und deren Verteilung I

Definition 9.1 (Zufallsvariable, Verteilung, Realisation)

Seien $(\Omega, \mathcal{F}, P)$ *ein Wahrscheinlichkeitsraum,* $X : \Omega \to \mathbb{R}$ *eine* $\mathcal{F} - \mathcal{B}-$*messbare Abbildung. Dann heißen* X **(eindimensionale) Zufallsvariable** *über* $(\Omega, \mathcal{F}, P)$ *und die gemäß Definition 8.3 gebildete Bildwahrscheinlichkeit*

$$P_X : \mathcal{B} \to \mathbb{R}; B \mapsto P(X^{-1}(B))$$

Wahrscheinlichkeitsverteilung *oder kürzer* **Verteilung** *von* X. $(\mathbb{R}, \mathcal{B}, P_X)$ *ist damit ebenfalls ein Wahrscheinlichkeitsraum.*
Liegt nach Durchführung des Zufallsexperiments $(\Omega, \mathcal{F}, P)$ *das Ergebnis* $\omega \in \Omega$ *vor, so heißt der zugehörige Wert* $x = X(\omega)$ *die* **Realisierung** *oder* **Realisation** *von* X.

Eindimensionale Zufallsvariablen und deren Verteilung II

- $P_X(B)$ gibt nach Definition 8.3 für $B \in \mathcal{B}$ die Wahrscheinlichkeit an, als Ausgang des zugrundeliegenden Zufallsexperiments ein $\omega \in \Omega$ zu erhalten, das zu einer Realisation $X(\omega) \in B$ führt.
- Demzufolge sind folgende Kurzschreibweisen geläufig:
  - $P\{X \in B\} := P(\{X \in B\}) := P_X(B)$ für alle $B \in \mathcal{B}$,
  - $P\{X = x\} := P(\{X = x\}) := P_X(\{x\})$ für alle $x \in \mathbb{R}$,
  - $P\{X < x\} := P(\{X < x\}) := P_X((-\infty, x))$ für alle $x \in \mathbb{R}$, analog für $P\{X \leq x\}$, $P\{X > x\}$ und $P\{X \geq x\}$.
- In vielen Anwendungen interessiert man sich nur noch für die Bildwahrscheinlichkeit P_X der Zufallsvariablen X bzw. den (resultierenden) Wahrscheinlichkeitsraum $(\mathbb{R}, \mathcal{B}, P_X)$. Häufig wird X dann nur noch durch die Angabe von P_X festgelegt und auf die Definition des zugrundeliegenden Wahrscheinlichkeitsraums $(\Omega, \mathcal{F}, P)$ verzichtet.
- Verteilungen P_X von Zufallsvariablen als Abbildungen von $\mathcal{B}$ nach $\mathbb{R}$ allerdings schlecht handhabbar.

Verteilungsfunktionen

- Man kann zeigen, dass Wahrscheinlichkeitsmaße P_X auf $\mathcal{B}$ bereits (z.B.) durch die Angabe aller Wahrscheinlichkeiten der Form $P_X((-\infty, x]) = P(\{X \leq x\})$ für $x \in \mathbb{R}$ eindeutig bestimmt sind!
- Daher überwiegend „Identifikation" der Verteilung P_X mit Hilfe einer Abbildung $\mathbb{R} \to \mathbb{R}$ mit $x \mapsto P_X((-\infty, x])$ für $x \in \mathbb{R}$:

Definition 9.2 (Verteilungsfunktion)

Es seien X eine Zufallsvariable über dem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$ und P_X die Wahrscheinlichkeitsverteilung von X. Dann heißt die Abbildung

$$F_X : \mathbb{R} \to \mathbb{R}; F_X(x) := P_X((-\infty, x]) = P(\{X \leq x\})$$

Verteilungsfunktion *der Zufallsvariablen X.*

Eigenschaften von Verteilungsfunktionen I

- Die Verteilungsfunktion F_X einer (eindimensionalen) Zufallsvariablen X hat folgende Eigenschaften:
  1. F_X ist monoton wachsend, d.h. für $x, y \in \mathbb{R}$ gilt:
  $$x < y \quad \Rightarrow \quad F_X(x) \leq F_X(y)$$
  2. F_X ist rechtsseitig stetig, d.h. für alle $x \in \mathbb{R}$ gilt:
  $$\lim_{\substack{h \to 0 \\ h > 0}} F_X(x+h) = F_X(x)$$
  3. $\lim\limits_{x \to \infty} F_X(x) =: F_X(\infty) = 1$
  4. $\lim\limits_{x \to -\infty} F_X(x) =: F_X(-\infty) = 0$
- Als abkürzende Schreibweise für die linksseitigen Grenzwerte verwendet man

$$F_X(x-0) := \lim_{\substack{h \to 0 \\ h > 0}} F_X(x-h) \qquad \text{für alle } x \in \mathbb{R}.$$

Eigenschaften von Verteilungsfunktionen II

- Analog zur empirischen Verteilungsfunktion (deskriptive Statistik, Folie 46) können Intervallwahrscheinlichkeiten leicht mit der Verteilungsfunktion F_X einer Zufallsvariablen X berechnet werden.
- Für $a, b \in \mathbb{R}$ mit $a < b$ gilt:
  - $P_X((-\infty, b]) = P(\{X \leq b\}) = F_X(b)$
  - $P_X((-\infty, b)) = P(\{X < b\}) = F_X(b-0)$
  - $P_X([a, \infty)) = P(\{X \geq a\}) = 1 - F_X(a-0)$
  - $P_X((a, \infty)) = P(\{X > a\}) = 1 - F_X(a)$
  - $P_X([a, b]) = P(\{a \leq X \leq b\}) = F_X(b) - F_X(a-0)$
  - $P_X((a, b]) = P(\{a < X \leq b\}) = F_X(b) - F_X(a)$
  - $P_X([a, b)) = P(\{a \leq X < b\}) = F_X(b-0) - F_X(a-0)$
  - $P_X((a, b)) = P(\{a < X < b\}) = F_X(b-0) - F_X(a)$
- Insbesondere gilt für $x \in \mathbb{R}$ auch:

$$P_X(\{x\}) = P(\{X = x\}) = F_X(x) - F_X(x-0)$$

Diskrete Zufallsvariablen

- Einfacher, aber geläufiger Spezialfall für Zufallsvariable X über Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$: **Wertebereich**

$$X(\Omega) := \{x \in \mathbb{R} \mid x = X(\omega) \text{ für (mindestens) ein } \omega \in \Omega\}$$

  ist endlich oder abzählbar unendlich (bzw. etwas allgemeiner: es gibt eine endliche oder abzählbar unendliche Menge B mit $P(\{X \in B\}) = 1$).
- Analog zu Definition 6.2 heißen solche Zufallsvariablen „diskret".

Definition 9.3 (Diskrete ZV, Wahrscheinlichkeitsfunktion, Träger)

Seien $(\Omega, \mathcal{F}, P)$ ein Wahrscheinlichkeitsraum, X eine Zufallsvariable über $(\Omega, \mathcal{F}, P)$ und $B \subseteq \mathbb{R}$ endlich oder abzählbar unendlich mit $P(\{X \in B\}) = 1$. Dann nennt man

- *X eine* **diskrete** *Zufallsvariable,*
- *$p_X : \mathbb{R} \to [0, 1]; p_X(x) := P_X(\{x\})$ die* **Wahrscheinlichkeitsfunktion** *von X,*
- *$T(X) := \{x \in \mathbb{R} \mid p_X(x) > 0\}$ den* **Träger** *von X sowie alle Elemente $x \in T(X)$* **Trägerpunkte** *von X und deren zugehörige Wahrscheinlichkeitsfunktionswerte $p_X(x)$* **Punktwahrscheinlichkeiten**.

Beispiel

Anzahl „Wappen" bei dreimaligem Münzwurf

- Zufallsexperiment: Dreimaliger Münzwurf mit fairer Münze („**W**appen" oder „**Z**ahl")
  - $\Omega = \{WWW, WWZ, WZW, WZZ, ZWW, ZWZ, ZZW, ZZZ\}$
  - $\mathcal{F} = \mathcal{P}(\Omega)$
  - $P : \mathcal{F} \to \mathbb{R}; P(A) = \frac{|A|}{|\Omega|}$
- Zufallsvariable $X : \Omega \to \mathbb{R}$: Anzahl der auftretenden Wappen, also

$$\begin{array}{cccc} X(WWW) = 3, & X(WWZ) = 2, & X(WZW) = 2, & X(WZZ) = 1, \\ X(ZWW) = 2, & X(ZWZ) = 1, & X(ZZW) = 1, & X(ZZZ) = 0. \end{array}$$

- Für $X(\Omega) = \{0, 1, 2, 3\}$ gilt offensichtlich $P(X(\Omega)) = 1$, also X diskret.
- Konkreter ist $T(X) = \{0, 1, 2, 3\}$ und die Punktwahrscheinlichkeiten sind

$$p_X(0) = \frac{1}{8}, \qquad p_X(1) = \frac{3}{8}, \qquad p_X(2) = \frac{3}{8}, \qquad p_X(3) = \frac{1}{8}.$$

Diskrete Zufallsvariablen (Forts.) I

- $T(X)$ ist endlich oder abzählbar unendlich, die Elemente von $T(X)$ werden daher im Folgenden häufig mit x_i bezeichnet (für $i \in \{1, \ldots, n\}$ bzw. $i \in \mathbb{N}$), Summationen über Trägerpunkte mit dem Symbol $\sum_{x_i}$.
- Ist $p_X : \mathbb{R} \to \mathbb{R}$ die Wahrscheinlichkeitsfunktion einer diskreten Zufallsvariablen X, so gilt $P_X(A) = \sum_{x_i \in A \cap T(X)} p_X(x_i)$ für alle $A \in \mathcal{B}$.
- Spezieller gilt für die Verteilungsfunktion F_X einer diskreten Zufallsvariablen

$$F_X(x) = \sum_{\substack{x_i \in T(X) \\ x_i \leq x}} p_X(x_i) \qquad \text{für alle } x \in \mathbb{R}.$$

Verteilungsfunktionen diskreter Zufallsvariablen sind damit (vergleichbar mit empirischen Verteilungsfunktionen) Treppenfunktionen mit **Sprunghöhen** $p_X(x_i)$ an den **Sprungstellen** (=Trägerpunkten) $x_i \in T(X)$.

Diskrete Zufallsvariablen (Forts.) II

- *Im Münzwurf-Beispiel:*

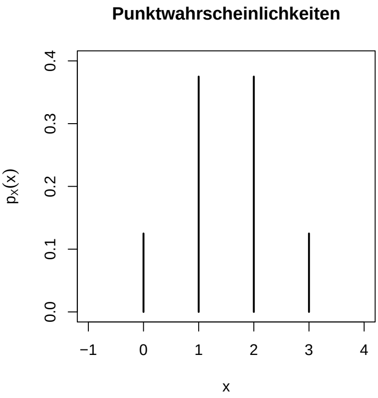
$$F_X(x) = \begin{cases} 0 & \text{für} \quad x < 0 \\ \frac{1}{8} & \text{für} \quad 0 \leq x < 1 \\ \frac{1}{2} & \text{für} \quad 1 \leq x < 2 \\ \frac{7}{8} & \text{für} \quad 2 \leq x < 3 \\ 1 & \text{für} \quad x \geq 3 \end{cases}$$

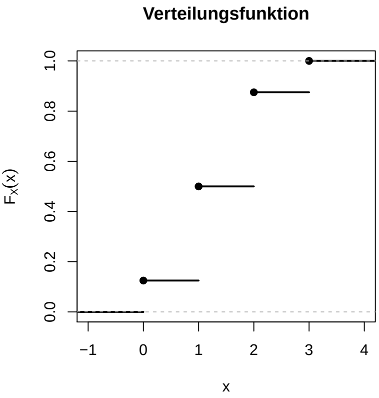
- Ist der Träger $T(X)$ endlich und die Anzahl der Elemente in $T(X)$ klein, so werden die Punktwahrscheinlichkeiten häufig in Tabellenform angegeben.
- *Im Münzwurf-Beispiel:*

x_i	0	1	2	3
$p_X(x_i)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

Diskrete Zufallsvariablen (Forts.) III

- *Grafische Darstellung im Münzwurf-Beispiel:*





Stetige Zufallsvariablen I

- Weiterer wichtiger Spezialfall: **stetige** Zufallsvariablen
- Wertebereich stetiger Zufallsvariablen ist immer ein Kontinuum: es gibt **keine** endliche oder abzählbar unendliche Menge $B \subseteq \mathbb{R}$ mit $P_X(B) = 1$, stattdessen gilt sogar $P_X(B) = 0$ für alle endlichen oder abzählbar unendlichen Teilmengen $B \subseteq \mathbb{R}$.
- Verteilungsfunktionen stetiger Zufallsvariablen sind nicht (wie bei diskreten Zufallsvariablen) als Summe von Wahrscheinlichkeitsfunktionswerten darstellbar, sondern als Integral über eine sogenannte *Dichtefunktion*:

Stetige Zufallsvariablen II

Definition 9.4 (Stetige Zufallsvariable, Dichtefunktion)

Seien $(\Omega, \mathcal{F}, P)$ *ein Wahrscheinlichkeitsraum und* X *eine Zufallsvariable über* $(\Omega, \mathcal{F}, P)$. *Gibt es eine nichtnegative Abbildung* $f_X : \mathbb{R} \to \mathbb{R}$ *mit*

$$F_X(x) = \int_{-\infty}^{x} f_X(t)dt \qquad \text{für alle } x \in \mathbb{R}, \tag{2}$$

so heißt die Zufallsvariable X **stetig**. *Jede nichtnegative Abbildung* f_X *mit der Eigenschaft (2) heißt* **Dichtefunktion** *von* X.

Stetige Zufallsvariablen III

- Aus Definition 9.4 lassen sich weitere Eigenschaften von stetigen Zufallsvariablen bzw. Dichtefunktionen ableiten, zum Beispiel:
  - $P_X(\{x\}) = 0$ für alle $x \in \mathbb{R}$.
  - $F_X(x - 0) = F_X(x)$ für alle $x \in \mathbb{R}$, F_X ist also stetig auf $\mathbb{R}$.
  - $P(\{a \leq X \leq b\}) = P(\{a < X \leq b\}) = P(\{a \leq X < b\}) = P(\{a < X < b\})$ $= F_X(b) - F_X(a) = \int_a^b f_X(t)dt$ für alle $a, b \in \mathbb{R}$ mit $a \leq b$.
  - (Mindestens) in allen Stetigkeitsstellen $x \in \mathbb{R}$ von f_X ist F_X differenzierbar und es gilt $F_X'(x) = f_X(x)$.
- Wahrscheinlichkeit von Intervallen stimmt mit Fläche zwischen Intervall und Dichtefunktion (analog zu Histogrammen bei deskriptiver Statistik mit klassierten Daten) überein.

Stetige Zufallsvariablen IV

- Dichtefunktion f_X zu einer Verteilungsfunktion F_X ist nicht eindeutig bestimmt; Abänderungen von f_X an endlich oder abzählbar unendlich vielen Stellen sind beliebig möglich!
- **Aber:** Ist $f_X : \mathbb{R} \to \mathbb{R}$ nichtnegative uneigentlich integrierbare Abbildung mit $\int_{-\infty}^{+\infty} f_X(x)dx = 1$, so gibt es genau eine Verteilungsfunktion $F_X : \mathbb{R} \to \mathbb{R}$, zu der f_X Dichtefunktion ist.
- Neben diskreten und stetigen Zufallsvariablen sind auch Mischformen (mit einem diskreten und stetigen Anteil) möglich, auf die hier aber nicht näher eingegangen wird!

- **Beispiel:** Die Abbildung
$$f_X : \mathbb{R} \to \mathbb{R}; f_X(x) := \begin{cases} 6(x - x^2) & \text{für } 0 \le x \le 1 \\ 0 & \text{sonst} \end{cases}$$
ist nichtnegativ und uneigentlich integrierbar mit $\int_{-\infty}^{+\infty} f_X(x)dx = 1$, also eine Dichtefunktion.
- Die Verteilungsfunktion F_X zu f_X erhält man als
$$F_X : \mathbb{R} \to \mathbb{R}; F_X(x) = \int_{-\infty}^{x} f_X(t)dt = \begin{cases} 0 & \text{für } x < 0 \\ 3x^2 - 2x^3 & \text{für } 0 \le x \le 1 \\ 1 & \text{für } x > 1 \end{cases} .$$