Hasso Plattner Institut
Digital Engineering · Universität Potsdam

Mathe III

Ralf Herbrich

Eindimensionale Zufallsvariablen

Überblick

1. Zufallsvariable
2. Dichte und Zähldichte
3. Verteilungsfunktion
4. Transformation und Simulation von Zufallsvariablen
5. Erwartungswert
6. Varianz
7. Ungleichungen von Markov und Tschebyscheff

Überblick

1. **Zufallsvariable**
2. Dichte und Zähldichte
3. Verteilungsfunktion
4. Transformation und Simulation von Zufallsvariablen
5. Erwartungswert
6. Varianz
7. Ungleichungen von Markov und Tschebyscheff

Zufallsvariablen

- In der Praxis sind wir oft nicht an allen Details des Experiments interessiert, sondern nur dem **numerischen Wert** einer Messgröße.
- **Beispiel (Zwei Würfel).** Wir werfen zwei faire Würfel und zählen die Summe der beiden Zahlen. Welche Verteilung hat die Summe X?
  - $P_X(X = 2) = P_{\text{naive}}(\{(1,1)\}) = \frac{1}{36}$
  - $P_X(X = 3) = P_{\text{naive}}(\{(1,2), (2,1)\}) = \frac{2}{36}$
  - $P_X(X = 4) = P_{\text{naive}}(\{(1,3), (2,2), (3,1)\}) = \frac{3}{36}$
  - $P_X(X = 5) = P_{\text{naive}}(\{(1,4), (2,3), (3,2), (4,1)\}) = \frac{4}{36}$
  - $P_X(X = 6) = P_{\text{naive}}(\{(1,5), (2,4), (3,3), (4,2), (5,1)\}) = \frac{5}{36}$
  - $P_X(X = 7) = P_{\text{naive}}(\{(1,6), (2,5), (3,4), (4,3), (5,2), (6,1)\}) = \frac{6}{36}$
  - $P_X(X = 8) = P_{\text{naive}}(\{(2,6), (3,5), (4,4), (5,3), (6,2)\}) = \frac{5}{36}$
  - $P_X(X = 9) = P_{\text{naive}}(\{(3,6), (4,5), (5,4), (6,3)\}) = \frac{4}{36}$
  - $P_X(X = 10) = P_{\text{naive}}(\{(4,6), (5,5), (6,4)\}) = \frac{3}{36}$
  - $P_X(X = 11) = P_{\text{naive}}(\{(5,6), (6,5)\}) = \frac{2}{36}$
  - $P_X(X = 12) = P_{\text{naive}}(\{(6,6)\}) = \frac{1}{36}$

Zufallsvariablen

- **Was ist hier passiert?**

Summe

$\Omega_1 = \{1,2,3,4,5,6\}^2$

$\Omega_2 = \{2,3,4,5,6,7,8,9,10,11,12\} \subset \mathbb{N}$

$$\omega = (n_1, n_2) \in \Omega_1 \xrightarrow{X: \Omega_1 \to \Omega_2} X(\omega) = n_1 + n_2$$

$$X^{-1}(i) = \{(n_1, n_2) \mid n_1 + n_2 = i\} \xleftarrow{X^{-1}: \Omega_2 \to \mathcal{P}(\Omega_1)} i \in \Omega_2$$

$$P_{\text{naive}}(\omega) = \frac{1}{36}$$

$$P_X(X = i) = P_{\text{naive}}\left(X^{-1}(i)\right)$$

Zufallsvariablen: Formal

- **Definition (Zufallsvariable).** Gegeben einen Ereignisraum $(\Omega_1, \mathcal{F}_1)$ und einen Ereignisraum $(\Omega_2, \mathcal{F}_2)$. Wir nennen die Funktion $X: \Omega_1 \to \Omega_2$ eine Zufallsvariable, wenn für jedes Ereignis $A \in \mathcal{F}_2$ gilt, dass $X^{-1}(A) \in \mathcal{F}_1$, d.h. wenn X messbar ist.
- **Definition (Diskrete und Reelle Zufallsvariable).** Wenn Ω_2 endlich oder abzählbar unendlich ist, so bezeichnen wir X als diskrete Zufallsvariable. Wenn $\Omega_2 = \mathbb{R}$, so bezeichnen wir X als stetige oder reelle Zufallsvariable.
- **Definition (Verteilung von Zufallsvariablen).** Gegeben ein Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$ und ein Ereignisraum $(\Omega_X, \mathcal{F}_X)$ sowie eine Zufallsvariable $X: \Omega \to \Omega_X$. Dann ist die Wahrscheinlichkeitsverteilung P_X

$$P_X(A) = P\left(X^{-1}(A)\right) \qquad \text{für alle } A \in \mathcal{F}_X$$

- **Bemerkung (Zufallsvariablen).** Die Verteilung P_X ist eindeutig und wird Verteilung von X genannt. Wir schreiben hierfür auch kurz $X \sim P_X$. Zusammen mit $(\Omega_X, \mathcal{F}_X)$ bildet P_X einen Wahrscheinlichkeitsraum $(\Omega_X, \mathcal{F}_X, P_X)$.

Verteilung von Zufallsvariablen

- **Bemerkungen (Notation).** Gegeben zwei Wahrscheinlichkeitsräume $(\Omega, \mathcal{F}, P)$ und $(\Omega_X, \mathcal{F}_X, P_X)$ mit einer Zufallsvariable $X: \Omega \to \Omega_X$.
  - Für ein Ereignis $A \in \mathcal{F}_X$ schreiben wir statt $P_X(A)$ oder $P(X^{-1}(A))$ auch $P(X \in A)$.
  - Für ein Ergebnis $\omega \in \Omega_X$ schreiben wir statt $P_X(\{\omega\})$ oder $P(X^{-1}(\{\omega\}))$ auch $P(X = \omega)$.
  - Ist X eine reelle Zufallsvariable, schreiben wir statt $P_X\big((-\infty, c]\big)$ oder $P\left(X^{-1}\big((-\infty, c]\big)\right)$ auch kurz $P(X \leq c)$.
  - Analoge Schreibweisen lassen sich für $\neq, <, \geq$ und $>$ definieren.
  - Ebenso möglich ist die Anwendung logischer Operatoren.
    - $P(a < X \leq b) = P(X > a \wedge X \leq b) = P\left(X^{-1}\big((a, \infty)\big) \cap X^{-1}\big((-\infty, b]\big)\right)$
    - $P(b \leq X < a) = P(X < a \wedge X \geq b) = P\left(X^{-1}\big((-\infty, a)\big) \cap X^{-1}\big([b, \infty)\big)\right)$

Identisch verteilte Zufallsvariablen

- **Definition (Identisch verteilte Zufallsvariablen).** Sei $(\Omega, \mathcal{F})$ ein Ereignisraum, in welchen zwei Zufallsvariablen X und Y mit der jeweiligen Verteilung P_X und P_Y abbilden. Gilt $P_X = P_Y$, so nennen wir X und Y identisch verteilt. Wir schreiben hierfür auch kurz $X \sim Y$.
- **Bemerkungen (Identisch verteilte Zufallsvariablen)**
  - Identisch verteilt bedeutet nicht dasselbe wie gleichverteilt.
  - Wenn für Zufallsvariablen X und Y gilt $X = Y$, so gilt $X \sim Y$. Die Umkehrung gilt nicht!
  - **Beispiel (Identisch verteilte Zufallsvariablen).** Sei X eine Zufallsvariable, welche bei zweifachem Münzwurf die Anzahl der Münzen zählt, die Zahl zeigen. Sei analog dazu Y eine Zufallsvariable, welche die Anzahl der Münzen zählt, die Kopf zeigen.

$$P(X = 0) = P(\{(\text{Kopf}, \text{Kopf})\}) = \frac{1}{4} \qquad P(Y = 0) = P(\{(\text{Zahl}, \text{Zahl})\}) = \frac{1}{4}$$

$$P(X = 1) = P(\{(\text{Kopf}, \text{Zahl}), (\text{Zahl}, \text{Kopf})\}) = \frac{1}{2} \qquad P(Y = 1) = P(\{(\text{Kopf}, \text{Zahl}), (\text{Zahl}, \text{Kopf})\}) = \frac{1}{2}$$

$$P(X = 2) = P(\{(\text{Zahl}, \text{Zahl})\}) = \frac{1}{4} \qquad P(Y = 2) = P(\{(\text{Kopf}, \text{Kopf})\}) = \frac{1}{4}$$

    - Gleichzeitig gilt $X((\text{Kopf}, \text{Kopf})) = 0$ und $Y((\text{Kopf}, \text{Kopf})) = 2$ und folglich $X \neq Y$.

Überblick

1. Zufallsvariable
2. **Dichte und Zähldichte**
3. Verteilungsfunktion
4. Transformation und Simulation von Zufallsvariablen
5. Erwartungswert
6. Varianz
7. Ungleichungen von Markov und Tschebyscheff

Zähldichte (Diskrete Zufallsvariablen)

- **Definition (Zähldichte).** Für eine diskrete Ergebnismenge Ω bezeichnen wir eine Funktion $p: \Omega \to [0,1]$ mit $\sum_{\omega \in \Omega} p(\omega) = 1$ (d.h. p ist normiert) als Zähldichte von Ω. Die Zähldichte wird auch Gewichtung genannt.
  - **Beispiel (Zähldichte).** Für einen fairen Würfel mit $\Omega = \{ 1, 2, 3, 4, 5, 6 \}$ ist die Funktion $p: \Omega \to [0,1], \omega \mapsto \frac{1}{6}$ eine Zähldichte.
  - **Beispiel (Zähldichte).** Für eine Münze mit $\Omega = \{\mathrm{Kopf}, \mathrm{Zahl}\}$ ist die Funktion $p: \Omega \to [0,1], \omega \mapsto \frac{1}{3}$ keine Zähldichte. Warum?
    - Da sie nicht normiert ist.
  - **Beispiel (Zähldichte).** Für eine Münze mit $\Omega = \{\mathrm{Kopf}, \mathrm{Zahl}\}$ ist die Funktion $p: \Omega \to [0,1]$ mit $p(\mathrm{Kopf}) = 2$ und $p(\mathrm{Zahl}) = -1$ keine Zähldichte. Warum?
    - Da sie zwar normiert ist $p(\mathrm{Kopf}) + p(\mathrm{Zahl}) = 1$, allerdings Werte außerhalb des Wertebereiches $[0,1]$ annimmt.

Diskrete und stetige Zufallsversuche

- **Beispiel (Diskrete und stetige Wahrscheinlichkeitsräume)**
  - **Wir betrachten die Augenzahl beim Wurf eines Würfels.**
    - Diskreter Wahrscheinlichkeitsraum ($|\Omega| = 6$)
  - **Wir messen die Körpergröße einer zufällig gewählten Person im Raum.**
    - Stetiger Wahrscheinlichkeitsraum ($\Omega = \mathbb{R}_+$)
  - **Wir zählen die Anzahl von Münzen in einem Brunnen zu einem zufällig gewählten Zeitpunkt.**
    - Diskreter Wahrscheinlichkeitsraum ($\Omega = \mathbb{N}$)
- Angenommen, wir suchen eine Person, welche exakt 180cm groß ist. Aber selbst bei vielfacher Wiederholung werden wir keine Person finden, die **exakt** 180cm groß ist. Warum?
- **Satz (Wahrscheinlichkeit von Einzelergebnissen im stetigen Raum).** In einem stetigen Wahrscheinlichkeitsraum $(\mathbb{R}, \mathcal{F}, P)$ gilt für alle $\omega \in \mathbb{R}$, dass $P(\{\omega\}) = 0$.

Paradox

- Folgt hier aus der Additivität von P, dass

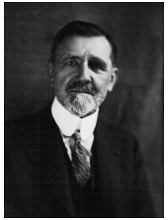
$$0 = \sum_{\omega\in\mathbb{R}} P(\{\omega\}) = P\left(\bigcup_{\omega\in\mathbb{R}}\{\omega\}\right) = P(\Omega) = 1?$$

überabzählbar unendlich!

  - Nein, da die Additivität nur eine Aussage über die **abzählbare** Vereinigung trifft (σ-Additivität)
- Aber welche Ergebnisse sind dann messbar?
  - Da die Wahrscheinlichkeit von Einzelergebnissen stets 0 ist, ist die Betrachtung der vollständigen Potenzmenge als Ereignissystem nicht möglich.
  - Stattdessen wird als Ereignissystem in der Regel die sogenannte **Borelalgebra** $\mathcal{B}(\mathbb{R})$ genutzt.
  - Ereignissystem, welches durch die folgende Ergebnismenge gegeben ist

$$\{(a,b] \mid a,b \in \mathbb{R}, a \neq b\}$$

  - Für uns reicht es zu wissen, dass wir bei stetigen Zufallsversuchen $\mathcal{B}(\mathbb{R})$ verwenden, und dass $\mathcal{B}(\mathbb{R})$ „praktisch alle" Teilmengen von Ω enthält.



Émile Borel
(1871 – 1956)

Dichte (Reelle Zufallsvariablen)

- **Definition (Dichte)**. Für eine stetige Ergebnismenge $\mathbb{R}$ bezeichnen wir eine Funktion $p: \Omega \to [0, \infty)$ mit $\int_\Omega p(x)\, dx = 1$ (d.h. p ist normiert) als Dichte von Ω. Eine Dichte p eines stetigen Ergebnisraumes $(\Omega, \mathcal{B}(\Omega))$ induziert genau eine Wahrscheinlichkeitsverteilung P mit

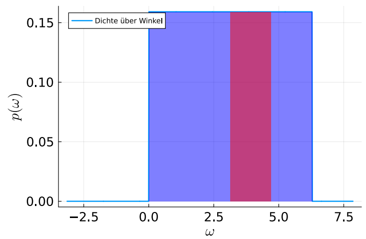
$$\forall A \in \mathcal{B}(\Omega): \quad P(A) = \int_A p(x)\, dx$$

  - **Beispiel (Dichte)**. Betrachte den Ausrichtungswinkel eines Glücksrades mit $\Omega = [0, 2\pi]$ und die Funktion $p: \Omega \to [0, \infty), \omega \mapsto \frac{1}{2\pi}$. Ist dies eine Dichte?

$$\int_0^{2\pi} p(x)\, dx \quad = \frac{x}{2\pi}\bigg|_0^{2\pi} \quad = \frac{2\pi}{2\pi} - \frac{0}{2\pi} \quad = 1$$

  - **Beispiel (Dichte)**. Betrachte nun das Ereignis $A = \left[\pi, \frac{3}{2}\pi\right] \subseteq \Omega$, dass das Glücksrad im dritten Viertel stoppt, gegeben die Dichte p. Wie groß ist die Wahrscheinlichkeit?

$$P(A) \quad = \int_A p(x)\, dx \quad = \int_\pi^{\frac{3}{2}\pi} \frac{1}{2\pi} dx \quad = \frac{x}{2\pi}\bigg|_\pi^{\frac{3}{2}\pi} \quad = \frac{\frac{3}{2}\pi}{2\pi} - \frac{\pi}{2\pi} \quad = \frac{3}{4} - \frac{1}{2} = \frac{1}{4}$$



Überblick

1. Zufallsvariable
2. Dichte und Zähldichte
3. **Verteilungsfunktion**
4. Transformation und Simulation von Zufallsvariablen
5. Erwartungswert
6. Varianz
7. Ungleichungen von Markov und Tschebyscheff

Kumulative Verteilungsfunktion

- **Definition (Kumulative Verteilungsfunktion)**. Für den stetigen Wahrscheinlichkeitsraum $(\mathbb{R}, \mathcal{B}(\mathbb{R}), P)$ bezeichnen wir die Funktion

$$F(c) \coloneqq P\big((-\infty, c]\big)$$

als die kumulative Verteilungsfunktion von P. Besitzt die Verteilung eine Dichte p, so gilt für alle $c \in \mathbb{R}$, dass

$$F(c) = \int_{-\infty}^{c} p(x)\ dx$$

Besitzt die Verteilung eine Zähldichte p, so gilt für alle $c \in \mathbb{R}$, dass

$$F(c) = \sum_{\omega \leq c} p(\omega)$$

- **Bemerkungen (Eigenschaften der kumulativen Verteilungsfunktionen)**
  - Jede Wahrscheinlichkeitsverteilung P hat eine kumulative Verteilungsfunktion.
  - Für jede kumulative Verteilungsfunktion F gilt: $\lim_{c \to -\infty} F(c) = 0$ sowie $\lim_{c \to \infty} F(c) = 1$.
  - Jede kumulative Verteilungsfunktion ist monoton wachsend und rechtsstetig.

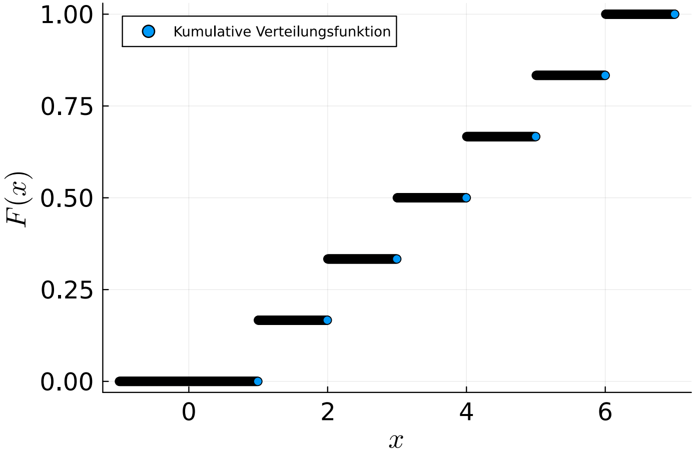
Kumulative Verteilungsfunktion

- **Beispiel (Kumulative Verteilungsfunktion für diskrete Zufallsräume).** Betrachte den Wurf eines fairen Würfels mit Zähldichte

$$p(\omega) = \begin{cases} \frac{1}{6} & \omega \in \{1,2,3,4,5,6\} \\ 0 & \text{sonst} \end{cases}$$

Was ist $F(1)$, $F(3)$, $F(3.48)$, $F(400)$ und $F(-12.7)$?

$$F(1) = \frac{1}{6}$$
$$F(3) = \frac{3}{6}$$
$$F(3.48) = \frac{3}{6}$$
$$F(400) = 1$$
$$F(-12.7) = 0$$



Kumulative Verteilungsfunktion

- **Beispiel (Kumulative Verteilungsfunktion für stetige Zufallsräume).** Betrachte den Ausrichtungswinkel eines Glücksrades mit Dichte

$$p(\omega) = \begin{cases} \frac{1}{2\pi} & \omega \in [0,2\pi] \\ 0 & \text{sonst} \end{cases}$$

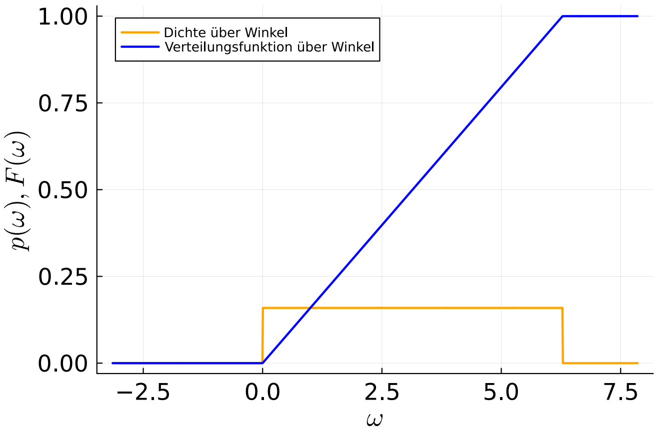
Was ist $F(\pi)$, $F\left(\frac{\pi}{4}\right)$, $F(400)$ und $F(-12.7)$?

$$F(\pi) = \int_0^{\pi} \frac{1}{2\pi} dx = \frac{1}{2}$$

$$F\left(\frac{\pi}{4}\right) = \int_0^{\frac{\pi}{4}} \frac{1}{2\pi} dx = \frac{1}{8}$$

$$F(400) = \int_0^{2\pi} \frac{1}{2\pi} dx = 1$$

$$F(-12.7) = \int_{-\infty}^{-12.7} 0\, dx = 0$$



Kumulative Verteilungsfunktionen und Dichten

- **Beobachtung**: Für eine Dichte p auf einem Ereignisraum $\left(\mathbb{R}, \mathcal{B}(\mathbb{R})\right)$ können wir eine eindeutige Wahrscheinlichkeitsverteilung P und somit auch eine eindeutige kumulative Verteilungsfunktion F ermitteln:

$$F(c) = P\big((-\infty, c]\big) = \int_{-\infty}^{c} p(x)\, dx$$

- **Frage:** Wie können wir umgekehrt eine Dichte p aus einer kumulativen Verteilungsfunktion F bestimmen?
  - Durch Ableiten, denn per Definition

$$dF(x) = p(x)\, dx \quad \Leftrightarrow \quad p(x) = \frac{d}{dx} F(x)$$

- **Bemerkungen (Dichtefunktion aus kumulativer Verteilungsfunktion)**
  - Die Ableitung von F liefert uns *eine mögliche* Dichtefunktion p. Es gibt allerdings mehr als eine Dichtefunktion, welche F beschreibt.
  - Die Verteilung besitzt eine Dichtefunktion genau dann, wenn F stetig ist und an höchstens abzählbar vielen Stellen nicht ableitbar ist.

Viel Spaß bis zur nächsten Vorlesung!