1. Wahrscheinlichkeitsrechnung

2. Diskrete Zufallsvariable

3. Stetige Zufallsvariable

4. Grenzwertsätze

5. Mehrdimensionale Zufallsvariable

Zufallsvariable

Eine Zufallsvariable X beschreibt eine Größe, die (nur) vom Ergebnis $\omega \in \Omega$ eines Zufallsvorgangs abhängt:

$$X : \Omega \to \mathbb{R}$$

Wenn $\omega \in \Omega$ eintritt, heißt die Zahl $X(\omega)$ *Realisierung* von X.

Durch Zufallsvariable X, Y kann man Ereignisse definieren, z.B.:

$$\{X > 0\} = \{\omega \in \Omega : X(\omega) > 0\}$$

$$\{X \in [-1, 5]\} = \{\omega \in \Omega : X(\omega) \in [-1, 5]\} =$$
$$= \{-1 \leq X \leq 5\} = \{X \geq -1\} \cap \{X \leq 5\}$$

$$\{1 < X \leq 2,\ Y = 3\} = \{\omega \in \Omega : 1 < X(\omega) \leq 2,\ Y(\omega) = 3\} =$$
$$= \{X > 1\} \cap \{X \leq 2\} \cap \{Y = 3\}$$

Unabhängigkeit

Die Zufallsvariablen $X_1, X_2, \ldots, X_k$ heißen (stochastisch) *unabhängig*, wenn

$$P(X_1 \in B_1,\, X_2 \in B_2,\, \ldots,\, X_k \in B_k) = \prod_{i=1}^{k} P(X_i \in B_i)$$

für alle (messbare) $B_1, B_2, \ldots, B_k \subseteq \mathbb{R}$.

Verteilungsfunktion

Die (kumulative) Verteilungsfunktion einer Zufallsvariablen X ist die Funktion $F : \mathbb{R} \to [0, 1]$ mit

$$F(x) = P(X \leq x)$$

für alle $x \in \mathbb{R}$.

Charakterisierende Eigenschaften einer Verteilungsfunktion $F : \mathbb{R} \to [0, 1]$:

1. F ist monoton steigend (d.h. $x < y \Rightarrow F(x) \leq F(y)$)
2. $\lim_{x \downarrow -\infty} F(x) = 0$ und $\lim_{x \uparrow +\infty} F(x) = 1$
3. F ist rechtsstetig (d.h. $\lim_{y \downarrow x} F(y) = F(x)$)

Diskrete Zufallsvariable

Eine Zufallsvariable $X : \Omega \to \mathcal{X}$ heißt *diskret*, wenn $\mathcal{X} \subset \mathbb{R}$ endlich oder abzählbar unendlich ist.

Wahrscheinlichkeitsverteilung: $f : \mathcal{X} \to [0, 1]$ mit $f(x) = P(X = x)$ für alle $x \in \mathcal{X}$.
Daraus folgt: $P(X \in A) = \sum_{x \in A \cap \mathcal{X}} f(x)$ für alle (messbare) $A \subseteq \mathbb{R}$.

Insbesondere: $F(x) = P(X \leq x) = \sum_{x' \in (-\infty, x] \cap \mathcal{X}} f(x') = \sum_{x' \in \mathcal{X} : x' \leq x} f(x')$ für alle $x \in \mathcal{X}$ (d.h. die Verteilungsfunktion F ist eine Treppenfunktion, die in den Stellen $x \in \mathcal{X}$ um den Betrag $f(x)$ nach oben springt, während sie dazwischen konstant verläuft).

Die diskreten Zufallsvariablen $X_1, X_2, \ldots, X_k$ sind genau dann unabhängig, wenn

$$P(X_1 = x_1,\, X_2 = x_2,\, \ldots,\, X_k = x_k) = \prod_{i=1}^{k} P(X_i = x_i)$$

für alle mögliche Werte $x_1, x_2, \ldots, x_k \in \mathbb{R}$.

Erwartungswert

Erwartungswert der diskreten Zufallsvariablen X:

$$E(X) = \sum_{x \in \mathcal{X}} x\, f(x)$$

Der Erwartungswert ist ein Lagemaß ($E(X)$ ist der Schwerpunkt der Verteilung von X).

Eigenschaften des Erwartungswertes für diskrete Zufallsvariable $X, X_1, X_2, \ldots, X_k$:

- $E\left(g(X)\right) = \sum_{x \in \mathcal{X}} g(x)\, f(x)$ für alle Funktionen $g : \mathbb{R} \to \mathbb{R}$ (z.B. $E(X^2) = \sum_{x \in \mathcal{X}} x^2\, f(x)$)
- $X = c$ konstant (d.h. $X(\omega) = c$ für alle $\omega \in \Omega$) $\Rightarrow E(X) = c$
- *Linearität*: $E\left(\sum_{i=1}^k \alpha_i\, X_i\right) = \sum_{i=1}^k \alpha_i\, E(X_i)$ für alle $\alpha_1, \alpha_2, \ldots, \alpha_k \in \mathbb{R}$
- $X_1, X_2, \ldots, X_k$ unabhängig $\Rightarrow E\left(\prod_{i=1}^k X_i\right) = \prod_{i=1}^k E(X_i)$

Varianz

Varianz der diskreten Zufallsvariablen X mit $\mu = E(X)$:

$$Var(X) = E\left((X-\mu)^2\right) = \sum_{x \in \mathcal{X}} (x-\mu)^2 \, f(x)$$

Standardabweichung von X:

$$\sigma = \sqrt{Var(X)}$$

Die Varianz und die Standardabweichung sind Streuungsmaße ($Var(X)$ ist die erwartete quadratische Abweichung der Zufallsvariablen X von ihrem Erwartungswert μ).

Eigenschaften der Varianz für diskrete Zufallsvariable $X, X_1, X_2, \ldots, X_k$:

- $Var(X) = E(X^2) - (E(X))^2$
- $X = c$ konstant (d.h. $X(\omega) = c$ für alle $\omega \in \Omega$) $\Rightarrow Var(X) = 0$
- $X_1, X_2, \ldots, X_k$ unabhängig $\Rightarrow Var\left(\sum_{i=1}^k \alpha_i \, X_i\right) = \sum_{i=1}^k \alpha_i^2 \, Var(X_i)$ für alle $\alpha_1, \alpha_2, \ldots, \alpha_k \in \mathbb{R}$

Gleichverteilung

Eine (diskrete) Zufallsvariable X heißt *gleichverteilt* auf $\mathcal{X} = \{x_1, x_2, \ldots, x_k\}$, wenn $P(X = x_i) = \frac{1}{k}$ für alle $i \in \{1, 2, \ldots, k\}$.

Eigenschaften:

- $E(X) = \frac{1}{k} \sum_{i=1}^{k} x_i$ (d.h. $E(X)$ ist das arithmetische Mittel von $x_1, x_2, \ldots, x_k$)
- $Var(X) = \left(\frac{1}{k} \sum_{i=1}^{k} x_i^2\right) - \left(\frac{1}{k} \sum_{i=1}^{k} x_i\right)^2$

Bernoulli-Verteilung

Eine (diskrete) Zufallsvariable X heißt *Bernoulli-verteilt* mit Parameter $p \in [0, 1]$ (kurz: $X \sim Ber(p)$), wenn $P(X = 1) = p$ und $P(X = 0) = 1 - p$.

Eigenschaften:

- $E(X) = p$
- $Var(X) = p\,(1 - p)$
- $(1 - X) \sim Ber(1 - p)$

Binomialverteilung

Eine (diskrete) Zufallsvariable X heißt *binomialverteilt* mit Parametern $n \in \mathbb{N}$ und $p \in [0,1]$ (kurz: $X \sim Bin(n,p)$), wenn $P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$ für alle $k \in \{0, 1, \dots, n\}$.

Eigenschaften:

- $X_1, X_2, \dots, X_n \overset{u.i.v.}{\sim} Ber(p) \Rightarrow \left(\sum_{i=1}^n X_i\right) \sim Bin(n,p)$ (d.h. $Bin(n,p)$ ist die Verteilung der Anzahl des Eintretens eines bestimmten Ereignisses A mit $P(A) = p$ in n unabhängigen Wiederholungen des Zufallsvorgangs)
- $E(X) = n\,p$
- $Var(X) = n\,p\,(1-p)$
- $(n - X) \sim Bin(n, 1-p)$
- $Y \sim Bin(m,p)$ unabhängig von $X \Rightarrow (X + Y) \sim Bin(n+m, p)$

Geometrische Verteilung

Eine (diskrete) Zufallsvariable X heißt *geometrisch verteilt* mit Parameter $p \in (0, 1]$ (kurz: $X \sim Geom(p)$), wenn $P(X = k) = p\,(1 - p)^{k-1}$ für alle $k \in \mathbb{N}$.

Eigenschaften:

- $X_1, X_2, \ldots \overset{u.i.v.}{\sim} Ber(p) \Rightarrow$ $P(X = k) = P(X_1 = 0,\, X_2 = 0,\, \ldots,\, X_{k-1} = 0,\, X_k = 1)$ für alle $k \in \mathbb{N}$ (d.h. $Geom(p)$ ist die Verteilung der Anzahl der unabhängigen Wiederholungen des Zufallsvorgangs bis zum Eintreten eines bestimmten Ereignisses A mit $P(A) = p$)
- $E(X) = \frac{1}{p}$
- $Var(X) = \frac{1-p}{p^2}$

Hypergeometrische Verteilung

Eine (diskrete) Zufallsvariable X heißt *hypergeometrisch verteilt* mit Parametern $n \in \{1, 2, \dots, N\}$, $M \in \{0, 1, \dots, N\}$, und $N \in \mathbb{N}$ (kurz: $X \sim Hyper(n, M, N)$), wenn $P(X = k) = \frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}$ für alle $k \in \{\max(0, n + M - N), \min(n, M)\}$.

Eigenschaften:

- $Hyper(n, M, N)$ ist die Verteilung der Anzahl der Einheiten mit einer bestimmten Eigenschaft E in einer Stichprobe von Umfang n gezogen ohne Zurücklegen aus einer Grundgesamtheit von N Einheiten, von denen M die Eigenschaft E besitzen
- $E(X) = n\,\frac{M}{N}$
- $Var(X) = n\,\frac{M}{N}\,\left(1 - \frac{M}{N}\right)\,\frac{N-n}{N-1}$, wenn $N \geq 2$

Poisson-Verteilung

Eine (diskrete) Zufallsvariable X heißt *Poisson-verteilt* mit Parameter $\lambda \in (0, +\infty)$ (kurz: $X \sim Pois(\lambda)$), wenn $P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}$ für alle $k \in \mathbb{N}_0$.

Eigenschaften:

- $X_n \sim Bin\left(n, \frac{\lambda}{n}\right) \Rightarrow \lim_{n \to \infty} P(X_n = k) = P(X = k)$ für alle $k \in \mathbb{N}_0$ (d.h. X_n konvergiert *in Verteilung* gegen X, kurz: $X_n \xrightarrow{d} X$)
- $E(X) = \lambda$
- $Var(X) = \lambda$
- $Y \sim Pois(\mu)$ unabhängig von $X \Rightarrow (X + Y) \sim Pois(\lambda + \mu)$