Dieses Cheatsheet basiert auf dem WuS Cheatsheet von xyquadrat.

1 Grundbegriffe

Definition Wahrscheinlichkeitraum

Ein Wahrscheinlichkeitraum ist ein Tupel $(\Omega, \mathcal{F}, \mathbb{P})$:

- Die Menge Ω nenen wir **Grundraum**. Ein $\omega \in \Omega$ nennen wir Elementarereignis.
- $\mathcal{F} \subseteq \mathbb{P}(\Omega)$ ist eine **Sigma-Algebra**.
- $\mathbb{P}$ ist ein **Wahrscheinlichkeitsmass** definiert auf $(\Omega, \mathcal{F})$.

Dabei ist $A \subseteq \Omega$ ein Ereignis.

1.1 Sigma-Algebra

Eine Sigma-Algebra ist ein Subset $\mathcal{F} \subseteq \mathcal{P}(\Omega)$ mit den folgenden Eigenschaften:

1. $\Omega \in \mathcal{F}$
2. $A \in \mathcal{F} \implies A^\complement \in \mathcal{F}$
3. $A_1, A_2, \ldots \in \mathcal{F} \implies \bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$
4. $\varnothing \in \mathcal{F}$
5. $A_1, A_2, \ldots \in \mathcal{F} \implies \bigcap_{i=1}^{\infty} A_i \in \mathcal{F}$
6. $A, B \in \mathcal{F} \implies A \cup B \in \mathcal{F}$
7. $A, B \in \mathcal{F} \implies A \cap B \in \mathcal{F}$

Nützlich ist ausserdem die De-Morgan-Regel:

$$(\cup_{i=1}^{\infty} A_i)^\complement = \cap_{i=1}^{\infty} (A_i)^\complement$$

1.2 Wahrscheinlichkeitsmass

Ein Wahrscheinlichkeitsmass $\mathbb{P}$ ist eine Abbildung

$$\begin{aligned} \mathbb{P} : \mathcal{F} &\mapsto [0,1] \\ A &\mapsto \mathbb{P}[A] \end{aligned}$$

mit den Eigenschaften

1. $\mathbb{P}[\Omega] = 1$
2. $\mathbb{P}[A] = \sum_{i=1}^{\infty} P[A_i]$, falls $A = \bigsqcup_{i=1}^{\infty} A_i$
3. $\mathbb{P}[\varnothing] = 0$
4. $\mathbb{P}[A^\complement] = 1 - \mathbb{P}[A]$
5. $\mathbb{P}[A \cup B] = \mathbb{P}[A] + P[B] - \mathbb{P}[A \cap B]$
6. $A \subseteq B \implies \mathbb{P}[A] \leq \mathbb{P}[B]$ (Monotonie)
7. $\mathbb{P}[\bigcup_{i=1}^{\infty} A_i] \leq \sum_{i=1}^{\infty} \mathbb{P}[A_i]$ (Union Bound)

Wenn $A_1, \ldots A_n$ paarweise disjunkt sind, gilt:

$$\mathbb{P}[A_1 \cup \ldots \cup A_n] = \mathbb{P}[A_1] + \ldots + \mathbb{P}[A_n]$$

1.3 Bedingte Wahrscheinlichkeiten

Sei $(\Omega, \mathcal{F}, \mathbb{P})$ ein Wahrscheinlichkeitsraum mit $A, B \in \mathcal{F}$ und $\mathbb{P}[B] > 0$. Die bedingte Wahrscheinlichkeit von A gegeben B ist definiert als:

$$\mathbb{P}[A|B] = \frac{\mathbb{P}[A \cap B]}{\mathbb{P}[B]}$$

Totale Wahrscheinlichkeit

Sei $B_1, \ldots, B_n \in \mathcal{F}$ eine Partition von Ω mit $\mathbb{P}[B_i] > 0$ für alle $1 \leq i \leq n$. Dann gilt:

$$\forall A \in \mathcal{F} \quad \mathbb{P}[A] = \sum_{i=1}^{n} \mathbb{P}[A|B_i]\ \mathbb{P}[B_i]$$

Satz von Bayes

Sei $B_1, \ldots, B_n \in \mathcal{F}$ eine Partition von Ω mit $\mathbb{P}[B_i] > 0$ für jedes i. Für jeden Event A mit $\mathbb{P}[A] > 0$ gilt:

$$\forall i = 1, \ldots, n \quad \mathbb{P}[B_i|A] = \frac{\mathbb{P}[A|B_i]\ \mathbb{P}[B_i]}{\sum_{k=1}^{n} \mathbb{P}[A|B_k]\ \mathbb{P}[B_k]}$$

1.4 Unabhängigkeit

Definition Unabhängigkeit

Zwei Ereignisse $A, B \in \mathcal{F}$ sind unabhängig, falls gilt:

$$\mathbb{P}[A \cap B] = \mathbb{P}[A] \cdot \mathbb{P}[B]$$

Alternative Definitionen sind:

$$\mathbb{P}[A|B] = \mathbb{P}[A] \text{ bzw. } \mathbb{P}[B|A] = \mathbb{P}[B]$$

- Falls $\mathbb{P}[A] \in \{0, 1\}$, dann ist A zu jedem Ereignis unabhängig.
- Wenn ein Ereignis A unabhängig zu sich selbst ist, dann folgt $\mathbb{P}[A] \in \{0, 1\}$.
- Wenn A, B unabhängig sind, dann sind auch $A, B^\complement$ unabhängig.

Unabhängigkeit für mehrere Ereignisse

Seien $A_1, \ldots, A_n \in F$, dann sind die Ereignisse unabhängig, falls gilt:

$$\forall I \subseteq \{1, \ldots, n\} \quad \mathbb{P}[\bigcap_{i \in I} A_i] = \prod_{i \in I} \mathbb{P}[A_i]$$

2 Zufallsvariablen

Eine Zufallsvariable ist eine Abbildung $X : \Omega \mapsto \mathbb{R}$ mit

$$\forall x \in \mathbb{R} \quad \{\omega \in \Omega \mid X(\omega) \leq x\} \in \mathcal{F}$$

Dabei lassen wir das ω oft weg und schreiben nur X.

2.1 Verteilungsfunktion

Definition Verteilungsfunktion

Die Verteilungsfunktion ist die Abbildung $F_X : \mathbb{R} \mapsto [0, 1]$ definiert durch

$$\forall x \in \mathbb{R} \quad F_X(x) = \mathbb{P}[X \leq x]$$

Die Verteilungsfunktion hat folgende Eigenschaften:

1. $a < b \implies \mathbb{P}[a < X \leq b] = F_X(b) - F_X(a)$
2. F ist monoton wachsend
3. F ist rechtsstetig, d.h. $\lim_{t\to 0} F_{x+t} = F(x)$
4. $\lim_{x\to -\infty} F_X(x) = 0$ und $\lim_{x\to\infty} F_X(x) = 1$

Linksstetigkeit

Die Definition der Linksstetigkeit ist

$$F(x-) = \lim_{t\to 0} F(x-t)$$

Es gilt allerdings *nicht* immer $F(x-) = F(x)$, d.h. nicht jede Verteilungsfunktion ist linksstetig. Allerdings gilt:

$$\forall x \in \mathbb{R} \quad F(x) - F(x-) = \mathbb{P}[X = x]$$

Daraus lässt sich für stetige ZV $P[X = x] = 0$ folgern.

2.2 Unabhängigkeit

Die Zufallsvariablen $X_1, \dots, X_n$ sind unabhängig, falls:

$$\begin{aligned}\forall x_1, \dots, x_n \in \mathbb{R}: \quad & \mathbb{P}[X_1 \leq x_1, \dots, X_n \leq x_n] = \\ & \mathbb{P}[X_1 \leq x_1] \cdot \ldots \cdot \mathbb{P}[X_n \leq x_n]\end{aligned}$$

Eine Folge von Zufallsvariablen $X_1, X_2, \dots$ ist unabhängig, falls $\forall n$ $X_1, \dots X_n$ unabhängig sind. Sie ist zusätzlich identisch verteilt (uiv.), falls ausserdem $\forall i, j \quad F_{X_i} = F_{X_j}$ gilt.

2.3 Transformationen

Sei $\varphi : \mathbb{R} \mapsto \mathbb{R}$ und X eine Zufallsvariable, so ist

$$\varphi(X) = \varphi \circ X$$

auch eine Zufallsvariable. Seien $X_1, \dots, X_n$ ZVs mit $\phi : \mathbb{R}^n \mapsto \mathbb{R}$, so ist

$$\phi(X_1, \dots, X_n) = \phi \circ (X_1, \dots, X_n)$$

ebenfalls eine Zufallsvariable.

2.4 Konstruktion einer Zufallsvariable

Gegeben eine Verteilungsfunktion F_X wollen wir die entsprechende ZV X konstruieren.

Kolmogorov-Theorem

Es existiert $(\Omega, \mathcal{F}, \mathbb{P})$ und ZVs $X_1, X_2, \dots$, sodass $X_1, X_2, \dots$ uiv. Bernoullivariablen mit $p = 0.5$ sind.

Sei $X_1, X_2, \dots \sim \text{Ber}(1/2)$ eine unendliche Folge, dann ist

$$U = \sum_{n=1}^{\infty} 2^{-n} \cdot X_n$$

gleichverteilt auf $[0, 1]$.

Aufgrund der Eigenschaften der Verteilungsfunktion F wissen wir, dass eine eindeutige Inverse F^{-1} existiert. Wir können die generalisierte Inverse definieren als:

$$\forall \alpha \in [0, 1] \quad F^{-1}(\alpha) = \inf\{x \in \mathbb{R} \mid F(x) \geq \alpha\}$$

Sei nun F eine Verteilungsfunktion und U eine gleichverteilte ZV in $[0, 1]$. Dann besitzt $X = F^{-1}(U)$ genau die Verteilungsfunktion $F_X = F$.

Fast sichere Ereignisse

Ein Ereignis $A \in \mathcal{F}$ tritt fast sicher (f.s) ein, falls $\mathbb{P}[A] = 1$. Seien X, Y ZV, so schreiben wir $X \leq Y$ f.s. $\iff \mathbb{P}[X \leq Y] = 1$.

2.5 Diskrete Zufallsvariablen

Definition diskrete ZV

Eine ZV X heisst diskret, falls $\exists W \subset \mathbb{R}$ endlich oder abzählbar, so dass $X \in W$ f.s. Falls Ω endlich oder abzählbar ist, dann ist X immer diskret.

Die Verteilungsfunktion einer diskreten ZV ist:

$$(p(x))_{x\in W} \text{ wobei } \sum_{x\in W} p(x) = 1$$

Die Gewichtsfunktion einer diskreten ZV ist:

$$\forall x \in W \quad p(x) = \mathbb{P}[X = x]$$

2.6 Diskrete Verteilungen

Bernoulli-Verteilung ($X \sim \text{Ber}(p)$): Hat nur die Ereignisse $\{0, 1\}$. Sie ist definiert als

$$\mathbb{P}[X = 0] = 1 - p \text{ und } \mathbb{P}[X = 1] = p$$

Binomialverteilung ($X \sim \text{Bin}(n, p)$): Die Wiederholung von Bernoulli-Experimenten. Sie ist definiert als

$$\forall k \in \{0, \dots, n\} \quad \mathbb{P}[X = k] = \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k}$$

Geometrische Verteilung ($X \sim \text{Geom}(p)$): Beschreibt das erste Auftreten eines Erfolges. Sie ist definiert als

$$\forall k \in \mathbb{N} - \{0\} \quad \mathbb{P}[X = k] = (1-p)^{k-1} \cdot p$$

Poisson-Verteilung ($X \sim \text{Poisson}(\lambda)$): Annäherung an die Binomialverteilung für grosse n und kleine p (d.h. rare Ereignisse). Sie ist definiert als

$$\forall k \in \mathbb{N}, \lambda > 0 \quad \mathbb{P}[X = k] = \frac{\lambda^k}{k!} \cdot e^{-\lambda}$$

2.7 Stetige Zufallsvariablen

Definition stetige ZV

Eine ZV X heisst stetig, wenn ihre Verteilungsfunktion F_X wie folgt geschrieben werden kann:

$$\forall x \in \mathbb{R} \quad F_X(x) = \int_{-\infty}^{x} f(y)\, dy$$

Hierbei ist $f : \mathbb{R} \mapsto \mathbb{R}_+$ die Dichte von X. Für die Dichte gilt:

$$\int_{-\infty}^{+\infty} f(y)\, dy = 1$$

Es gelten folgende Eigenschaften:

1. $\mathbb{P}[a \leq x \leq b] = \mathbb{P}[a < x < b]$
2. $\mathbb{P}[X = x] = 0$
3. $\mathbb{P}[X \in [a, b]] = \mathbb{P}[X \in (a, b)]$

2.8 Stetige Verteilungen

Gleichverteilung ($X \sim \mathcal{U}[a,b]$): Jedes Ereignis hat die gleiche Wahrscheinlichkeit. Sie ist definiert als

$$f_{a,b}(x) = \begin{cases} 0 & x \notin [a,b] \\ \frac{1}{b-a} & x \in [a,b] \end{cases}$$

Exponentialverteilung ($X \sim \mathrm{Exp}(\lambda)$): Das stetige Gegenstück zur Geometrischen Verteilung. Sie ist definiert als

$$f_{a,b}(x) = \begin{cases} 0 & x < 0 \\ \lambda \cdot e^{-\lambda x} & x \geq 0 \end{cases}$$

- Wenn $Y \sim \mathrm{Exp}(\lambda)$, dann ist $c \cdot Y \sim \mathrm{Exp}(\frac{\lambda}{c})$
- $X_1, X_2 \sim \mathcal{N}(0,1)$ uiv. $\implies \chi_2 = X_1^2 + X_2^2 \sim \mathrm{Exp}(\frac{1}{2})$

Normalverteilung ($X \sim \mathcal{N}(m, \sigma^2)$): Die wohl wichtigste Verteilung. Sie ist definiert als

$$f_{m,\sigma}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{(x-m)^2}{2\sigma^2}}$$

$X \sim \mathcal{N}(0,1)$ wird auch Standardnormalverteilung genannt. Für eine standardnormalverteilte ZV X gilt

$$Z = m + \sigma \cdot X \sim \mathcal{N}(m, \sigma^2)$$

3 Erwartungswert

Definition Erwartungswert

Sei $X : \Omega \mapsto \mathbb{R}_+$ eine ZV mit nicht-negativen Werten. Dann ist

$$\mathbb{E}[X] = \int_0^\infty 1 - F_X(x)\, dx$$

der Erwartungswert von X.

Wenn $\mathbb{E}[|X|] < \infty$, dann ist der Erwartungswert definiert als

$$\mathbb{E}[X] = \mathbb{E}[X_+] - \mathbb{E}[X_-]$$

3.1 Erwartungswert diskreter ZV

Sei X eine diskrete ZV mit $X \in W$ f.s. Sei $\phi : \mathbb{R} \mapsto \mathbb{R}$ eine Abbildung. Falls die Summe wohldefiniert ist, gilt

$$\mathbb{E}[\phi(X)] = \sum_{x \in W} \phi(x) \cdot \mathbb{P}[X = x]$$

Sei $\phi = \mathrm{id}$, gilt

$$\mathbb{E}[X] = \sum_{x \in W} x \cdot \mathbb{P}[X = x].$$

3.2 Erwartungswert stetiger ZV

Sei X eine stetige ZV mit Dichtefunktion f. Sei $\phi : \mathbb{R} \mapsto \mathbb{R}$ eine Abbildung, sodass $\phi(x)$ eine Zufallsvariable ist. Sofern das Integral wohldefiniert ist, gilt

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

Die Definition für $\phi = \mathrm{id}$ ist analog zum diskreten Fall.

3.2.1 Bestimmen der Dichte von $f(X)$

1. Sei $\phi : \mathbb{R} \mapsto \mathbb{R}$ stückweise beschränkt und stetig.
2. $\mathbb{E}[\phi(Y)] = \mathbb{E}[\phi(f(X))] = \mathbb{E}[\tau(X)]$
3. Dichte von X in vorheriger Gleichung einsetzen.
4. Variablenwechsel $u = f(x)$ & Grenzen anpassen

3.3 Rechnen mit Erwartungswerten

Linearität des Erwartungswertes

Seien X, Z ZV mit $\lambda \in \mathbb{R}$. Falls die Erwartungswerte wohldefiniert sind, gilt

$$\mathbb{E}[\lambda \cdot X] = \lambda \cdot \mathbb{E}[X]$$
$$\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y]$$

Falls zwei ZV X, Y unabhängig sind, gilt auch

$$\mathbb{E}[X \cdot Y] = \mathbb{E}[X] \cdot \mathbb{E}[Y]$$

Für Divison hingegen gilt:

$$\mathbb{E}[\frac{X}{Y}] = \mathbb{E}[X] \cdot \mathbb{E}[\frac{1}{Y}]$$

Alternativdefinition unabhängige ZV

Seien $X_1, \ldots, X_n$ diskrete ZV. Dann sind folgende Aussagen äquivalent:

1. $X_1, \ldots, X_n$ sind unabhängig.
2. Für jedes $\phi_1, \ldots, \phi_n : \mathbb{R} \mapsto \mathbb{R}$ beschränkt gilt:

$$\mathbb{E}[\phi_1(X_1) \cdot \ldots \cdot \phi_n(X_n)] = \mathbb{E}[\phi_1(X_1)] \cdot \ldots \cdot \mathbb{E}[\phi_n(X_n)]$$

3.4 Extremwertformel

Sei X eine diskrete ZV mit Werten in $\mathbb{N}$. Dann gilt folgende Identität:

$$\mathbb{E}[X] = \sum_{n=1}^{\infty} \mathbb{P}[X \geq n]$$

Sei X eine stetige ZV mit $X \geq 0$ f.s., dann gilt:

$$\mathbb{E}[X] = \int_0^\infty \mathbb{P}[X > x]\, dx$$

3.5 Ungleichungen

Monotonie

Seien X, Y ZV sodass $X \leq Y$ f.s., dann gilt $\mathbb{E}[X] \leq \mathbb{E}[Y]$.

Markov-Ungleichung

Sei X eine ZV mit $X \geq 0$ f.s., dann gilt für jedes $a > 0$:

$$\mathbb{P}[X \geq a] \leq \frac{\mathbb{E}[X]}{a}$$

Jensen-Ungleichung

Sei X eine ZV und $\phi : \mathbb{R} \mapsto \mathbb{R}$ eine konvexe Funktion, dann gilt:

$$\phi(\mathbb{E}[X]) \leq \mathbb{E}[\phi(X)]$$

3.6 Varianz

Sei X eine ZV sodass $\mathbb{E}[X^2] < \infty$. Die Varianz von X ist definiert durch

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

wobei $m = \mathbb{E}[X]$. Dabei wird σ_X auch die Standardabweichung von X genannt und beschreibt die typische Distanz eines Wertes $x \in X$ zu $\mathbb{E}[X]$.

Chebychev-Ungleichung

Wenn X eine ZV mit $\mathbb{E}[X^2] < \infty$ ist, dann gilt für jedes $a \geq 0$:

$$\mathbb{P}[|X - \mathbb{E}[X]| \geq a] \leq \frac{\sigma_X^2}{a^2}$$

1. Wenn $\mathbb{E}[X^2] < \infty$, dann $\sigma_X^2 = \mathbb{E}[X^2] - \mathbb{E}[X]^2$.
2. Wenn $\mathbb{E}[X^2] < \infty$ und $\lambda \in \mathbb{R}$, dann $\sigma_{\lambda X}^2 = \lambda^2 \sigma_X^2$.
3. Wenn $S = X_1 + \ldots + X_n$, wobei $X_1, \ldots, X_n$ paarweise unabhängig sind, dann gilt $\sigma_S^2 = \sigma_{X_1}^2 + \ldots + \sigma_{X_n}^2$.

3.7 Kovarianz

Wir können mit der Kovarianz die Abhängigkeit von zwei Zufallsvariablen messen.

Definition Kovarianz

Wenn X, Y zwei ZV mit $\mathbb{E}[X^2] < \infty, \mathbb{E}[Y^2] < \infty$, dann ist die Kovarianz zwischen X, Y definiert als

$$\mathrm{Cov}(X,Y) = \mathbb{E}[X \cdot Y] - \mathbb{E}[X] \cdot \mathbb{E}[Y]$$

- $\mathrm{Cov}(X,X) = \sigma_X^2$
- X, Y unabhängig $\implies \mathrm{Cov}(X,Y) = 0$

4 Gemeinsame Verteilungen

4.1 Diskrete gemeinsame Verteilungen

Definition gemeinsame Verteilung

Seien $X_1, \ldots, X_n$ diskrete Zufallsvariablen wobei $X_i \in W_i$ f.s. für $W_i \subset \mathbb{R}$. Die gemeinsame Verteilung (GV) von $X_1, \ldots, X_n$ ist die Familie $p = (p(x_1, \ldots, x_n))_{x_1 \in W_1, \ldots, x_n \in W_n}$ definiert durch

$$p(x_1, \ldots, x_n) = \mathbb{P}[X_1 = x_1, \ldots, X_n = x_n]$$

Seien $X_1, \ldots, X_n$ diskrete ZV mit $X_i \in W_i$ f.s. für $W_i \subset \mathbb{R}$ und $\phi : \mathbb{R}^n \mapsto \mathbb{R}$, so ist $Z = \phi(X_1, \ldots, X_n)$ eine diskrete ZV mit Werten in $W = \phi(W_1 \times \ldots \times W_n)$ und folgender Verteilung:

$$\forall z \in W.\ \mathbb{P}[Z = z] = \sum_{\substack{\phi(x_1,\ldots,x_n) \\ = z}} \mathbb{P}[X_1 = x_1, \ldots, X_n = x_n]$$

Seien $X_1, \ldots, X_n$ diskrete ZV mit $X_i \in W_i$ f.s. mit GV p. Dann ist die Randverteilung $\forall z \in W_i$:

$$\mathbb{P}[X_i = z] = \sum_{\substack{x_1, \ldots, x_{i-1}, \\ x_{i+1}, \ldots, x_n}} p(x_1, \ldots, x_{i-1}, z, x_{i+1}, \ldots, x_n)$$

Seien $X_1, \ldots, X_n$ diskrete ZV mit GV p und $\phi : \mathbb{R}^n \mapsto \mathbb{R}$. Dann ist der Erwartungswert definiert als:

$$\mathbb{E}[\phi(X_1, \ldots, X_n)] = \sum_{x_1, \ldots, x_n} \phi(x_1, \ldots, x_n) \cdot p(x_1, \ldots, x_n)$$

Seien $X_1, \ldots, X_n$ diskrete ZV mit GV p, dann sind die folgenden Aussagen äquivalent:

1. $X_1, \ldots, X_n$ sind unabhängig
2. Für alle $x_1 \in W_1, \ldots, x_n \in W_n$ gilt: $p(x_1, \ldots, x_n) = \mathbb{P}[X_1 = x_1] \cdot \ldots \cdot \mathbb{P}[X_n = x_n]$

Die gemeinsame Verteilung von $X_1, \ldots, X_n$ erfüllt

$$\sum_{x_1 \in W_1, \ldots, x_n \in W_n} p(x_1, \ldots, x_n) = 1$$

4.2 Stetige gemeinsame Verteilungen

Definition gemeinsame Verteilung

Seien $X_1, \ldots, X_n$ stetige ZV, so haben sie eine gemeinsame Verteilung, falls eine Funktion $f : \mathbb{R}^n \mapsto \mathbb{R}_+$ existiert, die für jedes $a_1, \ldots, a_n \in \mathbb{R}$ folgende Eigenschaft erfüllt:

$$\mathbb{P}[X_1 \leq a_1, \ldots, X_n \leq a_n]$$
$$= \int_{-\infty}^{a_1} \cdots \int_{-\infty}^{a_n} f(x_1, \ldots, x_n)\, dx_n \ldots dx_1$$

Dann ist f die **gemeinsame Dichte**.

Seien $X_1, \ldots, X_n$ stetige ZV mit einer gemeinsamen Dichte f und $\phi : \mathbb{R}^n \mapsto \mathbb{R}$. Dann ist der Erwartungswert definiert als

$$\mathbb{E}[\phi(X_1, \ldots, X_n)]$$
$$= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \phi(x_1, \ldots, x_n) f(x_1, \ldots, x_n)\, dx_n \ldots dx_1$$

Falls $X_1, \ldots, X_n$ eine gemeinsame Dichte f besitzen, ist die Randverteilung

$$f_i(z) = \int_{x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n \in \mathbb{R}^{n-1}}$$
$$f(x_1, \ldots, x_{i-1}, z, x_{i+1}, \ldots, x_n)\, dx_n \ \ldots dx_1$$

Wenn $X_1, \ldots, X_n$ stetige ZV mit Dichten $f_1, \ldots, f_n$ sind, dann sind die folgenden Aussagen äquivalent:

1. $X_1, \ldots, X_n$ sind unabhängig
2. $X_1, \ldots, X_n$ sind stetig mit gemeinsamer Dichte
$$f(x_1, \ldots, x_n) = f_1(x_1) \cdot \ldots \cdot f_n(x_n)$$
3. Für alle $\phi_1, \ldots, \phi_n : \mathbb{R} \mapsto \mathbb{R}$ gilt:
$$\mathbb{E}[\phi_1(x_1) \cdot \ldots \cdot (x_n)] = \mathbb{E}[\phi_1(x_1)] \cdot \ldots \cdot \mathbb{E}[\phi_n(x_n)]$$

5 Grenzwertsätze

Sei $X_1, X_2, \ldots$ eine unendliche Sequenz an uiv. ZV. Wir betrachten die Teilsumme $S_n = X_1 + \ldots + X_n$.

Gesetz der grossen Zahlen

Sei $\mathbb{E}[|X_1|] < \infty$ und $m = \mathbb{E}[X_1]$, so gilt

$$\lim_{n\to\infty} \frac{X_1 + \ldots + X_n}{n} = m \quad \text{f.s.}$$

Da die ZV uiv. sind, gilt $\mathbb{E}[|X_i|] < \infty$ und $m = \mathbb{E}[X_i]$ auch für alle i.

Konvergenz in Verteilung

Seien $(X_n)_{n\in\mathbb{N}}$ und X ZV. Wir schreiben

$$X_n \approx X \quad \text{für } n \to \infty$$

falls $\forall x \in \mathbb{R}$ gilt:

$$\lim_{n\to\infty} \mathbb{P}[X_n \le x] = \mathbb{P}[X \le x]$$

Zentraler Grenzwertsatz

Sei $\mathbb{E}[X_1^2] < \infty$ und wohldefiniert. Weiter sei $m = \mathbb{E}[X_1]$ und $\sigma^2 = \mathrm{Var}(X_1)$, so gilt:

$$\mathbb{P}\left[\frac{S_n - nm}{\sqrt{\sigma^2 n}} \le a\right] \xrightarrow[n\to\infty]{} \Phi(a) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{a} e^{\frac{-x^2}{2}}\, dx$$

Der zentrale Grenzwertsatz sagt aus, dass die Verteilung einer ZV

$$Z_n = \frac{S_n - nm}{\sqrt{\sigma^2 n}}$$

wie die Verteilung von $\mathcal{N}(0,1)$ aussieht. Es gilt

$$Z_n \approx Z \quad \text{für } n \to \infty$$

wobei $Z \sim \mathcal{N}(0,1)$. Für normalverteilte ZV $X_1, \ldots, X_n$ ist Z_n immer standardnormalverteilt.

6 Schätzer

Wir treffen folgende Annahmen:

- Parameterraum $\theta \subset \mathbb{R}$
- Familie von Wahrscheinlichkeitsmassen $(\mathbb{P}_\theta)_{\theta\in\Theta}$ auf $(\Omega, \mathcal{F})$; für jedes Element im Parameterraum existiert ein Wahrscheinlichkeitsmodell
- Zufallsvariablen $X_1, \ldots, X_n$ auf $(\Omega, \mathcal{F})$
- Wir nennen die Gesamtheit der beobachteten Daten $x_1, \ldots, x_n$ oder der ZV $X_1, \ldots, X_n$ Stichprobe

Definition Schätzer

Ein Schätzer ist eine Zufallsvariable $T : \Omega \mapsto \mathbb{R}$ von der Form

$$T = t(X_1, \ldots, X_n), \quad t : \mathbb{R}^n \mapsto \mathbb{R}$$

Ein Schätzer T ist **erwartungstreu**, falls für alle $\theta \in \Theta$ gilt:

$$\mathbb{E}_\theta[T] = \theta$$

Sei $\theta \in \Theta$ und T ein Schätzer. Der **Bias** (erwartete Schätzfehler) von T im Modell $\mathbb{P}_\theta$ ist definiert als:

$$\mathbb{E}_\theta[T] - \theta$$

Der mittlere quadratische Schätzfehler (MSE) von T im Modell $\mathbb{P}_\theta$ ist definiert als:

$$\mathrm{MSE}_\theta[T] = \mathbb{E}_\theta[(T-\theta)^2]$$
$$\mathrm{MSE}_\theta[T] = \mathrm{Var}_\theta(T) + (\mathbb{E}_\theta[T] - \theta)^2$$

6.1 Maximum-Likelihood-Methode

6.1.1 Likelihood-Funktion, ML-Schätzer

Die Likelihood-Funktion ist definiert als

$$L(x_1, \ldots, x_n; \theta) = \begin{cases} p(x_1, \ldots, x_n; \theta) & \text{(diskret)} \\ f(x_1, \ldots, x_n; \theta) & \text{(stetig)} \end{cases}$$

Für jedes $x_1, \ldots, x_n \in W$ sei $t_{ML}(x_1, \ldots, x_n)$ der Wert, welcher die Funktion $\Theta \mapsto L(x_1, \ldots, x_n; \theta)$ maximiert. Ein Maximum-Likelihood-Schätzer ist dann definiert als

$$T_{ML} = t_{ML}(X_1, \ldots, X_n)$$

6.1.2 Anwendung der Methode

Die Maximum-Likelihood-Methode ist ein Weg, um systematisch einen Schätzer zu bestimmen.

1. Gemeinsame Dichte/Verteilung der ZV finden
2. Bestimme davon die Log-Likelihood-Funktion $f(\theta) := \ln(L(x_1, \ldots, x_n; \theta))$
3. $f(\theta)$ nach θ ableiten
4. Nullstelle von $f'(\theta)$ finden

Unter dem gefundenen θ ist die Likelihood-Funktion maximal.

7 Konfidenzintervalle

Definition Konfidenzintervall

Sei $\alpha \in [0,1]$. Ein Konfidenzintervall für θ mit Niveau $1-\alpha$ ist ein Zufallsintervall $I = [A, B]$, sodass gilt

$$\forall \theta \in \Theta \quad \mathbb{P}_\theta[A \le \theta \le B] \ge 1 - \alpha$$

wobei A und B Zufallsvariablen der Form $A = a(X_1, \ldots, X_n), B = b(X_1, \ldots, X_n)$ mit $a, b : \mathbb{R}^n \to \mathbb{R}$ sind.

Wenn wir einen Schätzer $T = T_{ML} \sim \mathcal{N}(m, \frac{1}{n})$ haben, suchen wir ein Konfidenzintervall der Form

$$I = [T - c/\sqrt{n}, T + c/\sqrt{n}]$$

Hierbei gilt:

$$\begin{aligned} &\mathbb{P}_\theta[T - c/\sqrt{n} \le m \le T + c/\sqrt{n}] \\ &\quad = \mathbb{P}_\theta[-c \le Z \le c] \end{aligned}$$

wobei $Z = \sqrt{n}(T - m)$ ist.

7.1 Häufige Fälle

Normalverteilt - μ unbekannt, σ^2 bekannt (z-Test)

Erwartungstreuer Schätzer: $\overline{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$
Verteilung unter $\mathbb{P}_\theta$: $\frac{\overline{X}_n - \theta_0}{\sqrt{\sigma^2/n}} \sim \mathcal{N}(0,1)$

1. Modell $X_1, \dots, X_n \sim \mathcal{N}(\theta, \sigma^2)$ uiv. unter $\mathbb{P}_\theta$
2. Hypothesen $H_0 : \theta = \theta_0$, z.B. $H_A : \theta \neq \theta_0$
3. Test $T = \frac{\overline{X}_n - \mu}{\sqrt{\sigma^2/n}} \sim \mathcal{N}(0,1)$
4. Verwerfungsbereich $]-\infty, -c[\, \cup\,]c, \infty$ für $c \geq 0$

Normalverteilt - μ, σ^2 unbekannt (t-Test)

Wir definieren $\vec{\theta} = (\mu, \sigma^2)$ und den Varianz-Schätzer $S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \overline{X}_n)^2$.

1. Modell $X_1, \dots, X_n \sim \mathcal{N}(\theta, \sigma^2)$ uiv. unter $\mathbb{P}_{\vec{\theta}}$
2. Test $T = \frac{\overline{X}_n - \mu_0}{\sqrt{S^2/n}} \sim t_{n-1}$

7.2 Approximatives Konfidenzintervall

Wir können den zentralen Grenzwertsatz benutzen, um eine standardnormalverteilte ZV zu erhalten, und damit die Konfidenzintervalle zu bestimmen.

8 Tests

Null- und Alternativhypothese

Die Nullhypothese H_0 und die Alternativhypothese H_A sind zwei Teilmengen $\Theta_0 \subseteq \Theta, \Theta_A \subseteq \Theta$ wobei $\Theta_0 \cap \Theta_A = \varnothing$. Eine Hypothese heisst *einfach*, falls die Teilmenge aus einem einzelnen Wert besteht; sonst *zusammengesetzt*.

Definition Test

Ein Test ist ein Tupel (T, K), wobei T eine *ZV* der Form $T = t(X_1, \dots, X_n)$ und $K \subseteq \mathbb{R}$ eine deterministische Teilmenge von $\mathbb{R}$ ist. Wir nennen T die *Teststatistik* und K den *Verwerfungsbereich* oder kritischen Bereich.

Wir wollen nun anhand der Daten $(X_1(\omega), \dots, X_n(\omega))$ entscheiden, ob die Nullhypothese akzeptiert oder verworfen wird. Zuerst berechnen wir die Teststatistik $T(\omega) = t(X_1(\omega), \dots, X_n(\omega))$ und gehen dann wie folgt vor:

- Die Hypothese H_0 wird *verworfen*, falls $T(\omega) \in K$.
- Die Hypothese H_0 wird *akzeptiert*, falls $T(\omega) \notin K$.

Fehler 1. und 2. Art

Ein Fehler 1. Art ist, wenn H_0 fälschlicherweise verworfen wird, obwohl sie richtig ist.

$$\mathbb{P}_\theta[T \in K], \quad \theta \in \Theta_0$$

Ein Fehler 2. Art ist, wenn H_0 fälschlicherweise akzeptiert wird, obwohl sie falsch ist.

$$\mathbb{P}_\theta[T \notin K] = 1 - \mathbb{P}_\theta[T \in K], \quad \theta \in \Theta_A$$

8.1 Signifikanzniveau und Macht

Ein Test hat Signifikanzniveau $a \in [0,1]$ falls

$$\forall \theta \in \Theta_0 \quad \mathbb{P}_\theta[T \in K] \leq a$$

Es ist meist unser primäres Ziel, die Fehler 1. Art zu minimieren.

Das sekundäre Ziel ist, Fehler 2. Art zu vermeiden. Hierfür definieren wir die Macht eines Tests als Funktion:

$$\beta : \Theta_A \mapsto [0,1], \quad \theta \mapsto \mathbb{P}_\theta[T \in K]$$

Zu beachten ist, dass eine kleine Wahrscheinlichkeit für einen Fehler 2. Art einem *grossen* β entspricht.

8.2 Konstruktion von Tests

Wir nehmen an, dass $X_1, \dots, X_n$ diskret oder gemeinsam stetig unter $\mathbb{P}_{\theta_0}$ und $\mathbb{P}_{\theta_A}$ sind, wobei $\theta_0 \neq \theta_A$ einfach sind. Der Likelihood-Quotient ist somit wohldefiniert:

$$R(x_1, \dots, x_n) = \frac{L(x_1, \dots, x_n; \theta_A)}{L(x_1, \dots, x_n; \theta_0)}$$

(Falls $L(x_1, \dots, x_n; \theta_0) = 0$ setzen wir $R(x_1, \dots, x_n) = +\infty$.) Wenn $R \gg 1$, so gilt $H_A > H_0$ und analog $R \ll 1 \implies H_A < H_0$.

Likelihood-Quotient-Test

Der Likelihood-Quotient-Test (LQ-Test) mit Parameter $c \geq 0$ ist definiert durch:

$$T = R(x_1, \dots, x_n) \quad \text{und} \quad K = (c, \infty]$$

Der LQ-Test ist optimal, da jeder andere Test mit kleinerem Signifikanzniveau auch eine kleinere Macht hat (Neyman-Pearson-Lemma).

8.3 p-Wert

Sei $T = t(X_1, \dots, X_n)$ eine Teststatistik und $(T, K_t)_{t \geq 0}$ eine Familie von Tests.

Geordnete Teststatistik

Eine Familie von Tests heisst geordnet bzgl. T falls $K_t \subset \mathbb{R}$ und $s \leq t \implies K_t \subset K_S$. Beispiele:

- $K_t = (t, \infty)$ (rechtsseitiger Test)
- $K_t = (-\infty, -t)$ (linksseitiger Test)
- $K_t = (-\infty, -t) \cup (t, \infty)$ (beidseitiger Test)

Definition p-Wert

Sei $H_0 : \theta = \theta_0$ eine einfache Nullhypothese. Sei $(T, K_t)_{t \geq 0}$ eine geordnete Familie von Tests. Der p-Wert ist definiert als ZV $G(t)$, wobei

$$G : \mathbb{R}_+ \mapsto [0,1], \quad G(t) = \mathbb{P}_{\theta_0}[T \in K_t]$$

Der p-Wert hat folgende Eigenschaften:

1. Sei T stetig und $K_t = (t, \infty)$. Dann ist der p-Wert unter $\mathbb{P}_{\theta_0}$ auf $[0,1]$ gleichverteilt.
2. Für einen p-Wert γ gilt, dass alle Tests mit Signifikanzniveau $\alpha > \gamma$ die Nullhypothese verwerfen.

Insgesamt gilt also:

$$\text{kleiner } p\text{-Wert} \implies H_0 \text{ wird wahrscheinlich verworfen}$$

9 Aufgaben

9.1 Multiple Choice

Seien X, Y zwei ZV mit gemeinsamer Dichte $f_{X,Y}$. Welche Aussage ist korrekt?

- ✓ X, Y sind immer stetig
- ☐ Die ZV sind nicht notwendigerweise stetig.

Seien $(X_i)_{i=1}^n$ uiv. mit Verteilungsfunktion $F_{X_i} = F$. Was ist die Verteilungsfunktion von $M = \max(X_1, ..., X_n)$?

- ✓ $F_M(a) = F(a)^n$
- ☐ $F_M(a) = 1 - F(a)^n$
- ☐ $F_M(a) = (1 - F(a))^n$

Seien X, Y unabhängig und lognormalverteilt ($\ln X, \ln Y$ sind normalverteilt). Welche Aussage ist korrekt?

- ✓ XY ist lognormalverteilt
- ☐ XY ist normalverteilt
- ☐ e^{X+Y} ist normalverteilt

9.2 Aufgaben Wahrscheinlichkeit

Dichte von $\max(X_1, X_2)$

Seien $X_1, X_2 \sim \mathcal{U}[0,1]$ unabhängige ZV und sei $X = \max(X_1, X_2)$. Berechne die Dichtefunktion von X und $\mathbb{P}[X_1 \leq x \mid X \geq y]$.

$$\begin{aligned} F_X(t) &= \mathbb{P}[\max(X_1, X_2) \leq t] \\ &= \mathbb{P}[X_1 \leq t] \cdot \mathbb{P}[X_2 \leq t] = F_{X_1}(t) \cdot F_{X_2}(t) \\ f_X(t) &= \frac{d}{dt} F_{X_1}(t) \cdot F_{X_2}(t) = \frac{d}{dt} t^2 \cdot \mathbb{I}_{0 \leq t \leq 1} = 2t \cdot \mathbb{I}_{0 \leq t \leq 1} \end{aligned}$$

Für die Wahrscheinlichkeit brauchen wir eine Fallunterscheidung:

1. $x < 0$ oder $1 < x$:

$$\mathbb{P}[X_1 \leq x \mid X \geq y] = 0$$

2. $0 \leq x \leq y \leq 1$:

$$\frac{\mathbb{P}[X_1 \leq x \cap X \geq y]}{\mathbb{P}[X \geq y]} = \frac{x(1-y)}{1-y^2}$$

3. $0 \leq y \leq x \leq 1$:

$$\frac{\mathbb{P}[X_1 \leq x \cap X \geq y]}{\mathbb{P}[X \geq y]} = \frac{x-y^2}{1-y^2}$$

Gemeinsame Dichte

Bestimme die gemeinsame Dichte von $P \sim \mathcal{U}[0,1]$ und $H \sim \mathcal{U}[0,P]$. Wir wissen:

$$f_P(p) = \mathbb{I}_{p \in [0,1]} \quad f_{H|P}(h \mid p) = \frac{1}{p} \cdot \mathbb{I}_{h \in [0,p]}$$

Somit ist:

$$f_{P,H}(p,h) = f_P(p) \cdot f_{H|P}(h \mid p) = \frac{1}{p} \cdot \mathbb{I}_{0 \leq h \leq p \leq 1}$$

Zentraler Grenzwertsatz

Sei $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$ mit $X_1, ..., X_n$ uiv. und $\sigma_{X_i} = 6, \mu$. Zeige, dass für n gross genug gilt:

$$\mathbb{P}[|\bar{X} - \mu| \leq 1] \approx 2\Phi\left[\frac{\sqrt{n}}{6}\right] - 1$$

Hierfür verwenden wir den zentralen Grenzwertsatz:

$$\begin{aligned} \mathbb{P}[|\bar{X} - \mu| \leq 1] &= \mathbb{P}[|n \cdot \bar{X} - n \cdot \mu| \leq n] \\ &= \mathbb{P}\left[\left|\frac{n \cdot \bar{X} - n \cdot \mu}{\sqrt{6^2 n}}\right| \leq \frac{n}{\sqrt{6^2 n}}\right] \\ &= \mathbb{P}\left[\left|\frac{n \cdot \bar{X} - n \cdot \mu}{6\sqrt{n}}\right| \leq \frac{\sqrt{n}}{6}\right] \\ &\approx \Phi\left[\frac{\sqrt{n}}{6}\right] - \Phi\left[-\frac{\sqrt{n}}{6}\right] = 2\Phi\left[\frac{\sqrt{n}}{6}\right] - 1 \end{aligned}$$

Dichte via Erwartungswert

Sei $U \sim \mathcal{U}[0,1]$. Berechne die Dichte von

$$U' = a + (b-a)U$$

Wir definieren $\tau(x) = \phi(a + (b-a)U)$. Dann verwenden wir

$$\begin{aligned} \mathbb{E}[\tau(U)] &= \int_{-\infty}^{\infty} \tau(x) \mathbb{I}_{x \in [0,1]} \, dx \\ &= \int_0^1 \phi(a + (b-a)x) \, dx \\ &= \int_a^b \phi(y) \frac{1}{b-a} \, dy \\ &= \int_{-\infty}^{\infty} \phi(y) \frac{1}{b-a} \mathbb{I}_{y \in [a,b]} \, dy \end{aligned}$$

Die Dichte von U' ist also $\frac{1}{b-a} \mathbb{I}_{y \in [a,b]}$.

10 Tabellen

10.1 Grenzwerte

$\lim_{x\to\infty} \frac{e^x}{x^m} = \infty$	$\lim_{x\to-\infty} xe^x = 0$
$\lim_{x\to\infty}(1+x)^{\frac{1}{x}} = 1$	$\lim_{x\to 0}(1+x)^{\frac{1}{x}} = e$
$\lim_{x\to\infty}(1+\frac{1}{x})^b = 1$	$\lim_{x\to\infty} n^{\frac{1}{n}} = 1$
$\lim_{x\to 0} \frac{e^x-1}{x} = 1$	$\lim_{x\to\infty}(1-\frac{1}{x})^x = \frac{1}{e}$
$\lim_{x\to\pm\infty}(1+\frac{k}{x})^{mx} = e^{km}$	$\lim_{x\to\infty}(\frac{x}{x+k})^x = e^{-k}$
$\lim_{x\to 0} \frac{\log 1-x}{x} = -1$	$\lim_{x\to 0} x\log x = 0$
$\lim_{x\to 0} \frac{e^{ax}-1}{x} = a$	$\lim_{x\to 0} \frac{\ln(x+1)}{x} = 1$
$\lim_{x\to 1} \frac{\ln(x)}{x-1} = 1$	$\lim_{x\to\infty} \frac{\log(x)}{x^a} = 0$

Partielle Integration

$$\int f'(x)g(x)\,dx = f(x)g(x) - \int f(x)g'(x)\,dx$$

- Meist gilt: Polynome ableiten ($g(x)$), wo das Integral periodisch ist ($\sin, \cos, e^x, \ldots$) integrieren ($f'(x)$)
- Teils: mit 1 multiplizieren, um partielle Integration anwenden zu können (z.B. im Fall von $\int \log(x)\,dx$)

Substitution

Um $\int_a^b f(g(x))\,dx$ zu berechnen: Ersetze $g(x)$ durch u und integriere $\int_{g(a)}^{g(b)} f(u)\frac{du}{g'(x)}$.

- $g'(x)$ muss sich herauskürzen, sonst nutzlos.
- Grenzen substituieren nicht vergessen.
- Alternativ: unbestimmtes Integral berechnet werden und dann u wieder durch x substituieren.

10.2 Ableitungen

$\mathbf{F(x)}$	$\mathbf{f(x)}$	$\mathbf{f'(x)}$
$\frac{x^{-a+1}}{-a+1}$	$\frac{1}{x^a}$	$\frac{a}{x^{a+1}}$
$\frac{x^{a+1}}{a+1}$	$x^a\ (a\neq 1)$	$a\cdot x^{a-1}$
$\frac{1}{k\ln(a)}a^{kx}$	a^{kx}	$ka^{kx}\ln(a)$
$\ln\lvert x\rvert$	$\frac{1}{x}$	$-\frac{1}{x^2}$
$\frac{2}{3}x^{3/2}$	$\sqrt{x}$	$\frac{1}{2\sqrt{x}}$
$-\cos(x)$	$\sin(x)$	$\cos(x)$
$\sin(x)$	$\cos(x)$	$-\sin(x)$
$\frac{1}{2}(x-\frac{1}{2}\sin(2x))$	$\sin^2(x)$	$2\sin(x)\cos(x)$
$\frac{1}{2}(x+\frac{1}{2}\sin(2x))$	$\cos^2(x)$	$-2\sin(x)\cos(x)$
$-\ln\lvert\cos(x)\rvert$	$\tan(x)$	$\frac{1}{\cos^2(x)}$ $1+\tan^2(x)$
$\cosh(x)$	$\sinh(x)$	$\cosh(x)$
$\log(\cosh(x))$	$\tanh(x)$	$\frac{1}{\cosh^2(x)}$
$\ln\lvert\sin(x)\rvert$	$\cot(x)$	$-\frac{1}{\sin^2(x)}$
$\frac{1}{c}\cdot e^{cx}$	e^{cx}	$c\cdot e^{cx}$
$x(\ln\lvert x\rvert-1)$	$\ln\lvert x\rvert$	$\frac{1}{x}$
$\frac{1}{2}(\ln(x))^2$	$\frac{\ln(x)}{x}$	$\frac{1-\ln(x)}{x^2}$
$\frac{x}{\ln(a)}(\ln\lvert x\rvert-1)$	$\log_a\lvert x\rvert$	$\frac{1}{\ln(a)x}$

10.3 Weitere Ableitungen

$\mathbf{F(x)}$	$\mathbf{f(x)}$
$\arcsin(x)$	$\frac{1}{\sqrt{1-x^2}}$
$\arccos(x)$	$\frac{-1}{\sqrt{1-x^2}}$
$\arctan(x)$	$\frac{1}{1+x^2}$
$x^x\ (x>0)$	$x^x\cdot(1+\ln x)$

10.4 Integrale

$\mathbf{f(x)}$	$\mathbf{F(x)}$
$\int f'(x)f(x)\,dx$	$\frac{1}{2}(f(x))^2$
$\int \frac{f'(x)}{f(x)}\,dx$	$\ln\lvert f(x)\rvert$
$\int_{-\infty}^{\infty} e^{-x^2}\,dx$	$\sqrt{\pi}$
$\int (ax+b)^n\,dx$	$\frac{1}{a(n+1)}(ax+b)^{n+1}$
$\int x(ax+b)^n\,dx$	$\frac{(ax+b)^{n+2}}{(n+2)a^2} - \frac{b(ax+b)^{n+1}}{(n+1)a^2}$
$\int (ax^p+b)^n x^{p-1}\,dx$	$\frac{(ax^p+b)^{n+1}}{ap(n+1)}$
$\int (ax^p+b)^{-1}x^{p-1}\,dx$	$\frac{1}{ap}\ln\lvert ax^p+b\rvert$
$\int \frac{ax+b}{cx+d}\,dx$	$\frac{ax}{c} - \frac{ad-bc}{c^2}\ln\lvert cx+d\rvert$
$\int \frac{1}{x^2+a^2}\,dx$	$\frac{1}{a}\arctan\frac{x}{a}$
$\int \frac{1}{x^2-a^2}\,dx$	$\frac{1}{2a}\ln\left\lvert\frac{x-a}{x+a}\right\rvert$
$\int \sqrt{a^2+x^2}\,dx$	$\frac{x}{2}f(x) + \frac{a^2}{2}\ln(x+f(x))$

10.5 Diskrete Verteilungen

Verteilung	Parameter	$\mathbb{E}[X]$	$\mathrm{Var}(X)$	$p_X(t)$	$F_X(t)$
Gleichverteilung	n: Anzahl Ereignisse x_i: Ereignisse	$\frac{1}{n}\sum_{i=1}^n x_i$	$\frac{1}{n}\sum_{i=1}^n x_i^2 - \frac{1}{n^2}\left(\sum_{i=1}^n x_i\right)^2$	$\frac{1}{n}$	$\frac{\lvert\{k:x_k\leq t\}\rvert}{n}$
Bernoulli	p : ErfolgsWK	p	$p\cdot(1-p)$	$p^t(1-p)^{1-t}$	$1-p$ für $0\leq t<1$
Binomial	n: Anzahl Versuche p : ErfolgsWK	np	$np(1-p)$	$\binom{n}{t}p^t(1-p)^{n-t}$	$\sum_{k=0}^t \binom{n}{k}p^k(1-p)^{n-k}$
Geometrisch	p: ErfolgsWK t : Anzahl Versuche	$\frac{1}{p}$	$\frac{1-p}{p^2}$	$p(1-p)^{t-1}$	$1-(1-p)^t$
Poisson	λ: Erwartungswert und Varianz	λ	λ	$\frac{\lambda^t}{t!}e^{-\lambda}$	$e^{-\lambda}\sum_{k=0}^t \frac{\lambda^k}{k!}$

10.6 Stetige Verteilungen

Verteilung	Parameter	$\mathbb{E}[X]$	$\mathrm{Var}(X)$	$f_X(t)$	$F_X(t)$
Gleichverteilung	$[a,b]$: Intervall	$\frac{a+b}{2}$	$\frac{1}{12}(b-a)^2$	$\begin{cases}\frac{1}{b-a} & a\leq x\leq b\\ 0 & \text{sonst}\end{cases}$	$\begin{cases}0 & x\leq a\\ \frac{t-a}{b-a} & a<x<b\\ 1 & x\geq b\end{cases}$
Exponentialverteilung	$\lambda : \frac{1}{\mathbb{E}[X]}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$	$\begin{cases}\lambda e^{-\lambda t} & t\geq 0\\ 0 & t<0\end{cases}$	$\begin{cases}1-e^{-\lambda t} & t>0\\ 0 & t\leq 0\end{cases}$
Normalverteilung	σ^2: Varianz $\mu : \mathbb{E}[X]$	μ	σ^2	$\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(t-\mu)^2}{2\sigma^2}}$	$\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^t e^{-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2}\mathrm{d}y$
χ^2-Verteilung	n: Freiheitsgrad	n	$2n$	$\frac{1}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})}t^{\frac{n}{2}-1}e^{-\frac{t}{2}}$ für $t>0$	$P\left(\frac{n}{2},\frac{t}{2}\right)$
t-Verteilung	n: Freiheitsgrad	$\begin{cases}0 & n>1\\ \text{undef.} & \text{sonst}\end{cases}$	$\begin{cases}\frac{n}{n-2} & n>2\\ \infty & 1<n\leq 2\\ \text{undef.} & \text{sonst}\end{cases}$	$\frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\cdot\Gamma\left(\frac{n}{2}\right)}\left(1+\frac{t^2}{n}\right)^{-\frac{n+1}{2}}$	I'd rather not