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 $\mathcal{F} \subseteq \mathcal{P}(\Omega)$ hat folgende 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. $A, B \in \mathcal{F} \implies A \cap B \in \mathcal{F}$

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 $\mathbb{P}: \mathcal{F} \mapsto [0,1]$, $A \mapsto \mathbb{P}[A]$ mit den Eigenschaften:

1. $\mathbb{P}[\Omega] = 1$ und $\mathbb{P}[\varnothing] = 0$
2. $\mathbb{P}[A] = \sum_{i=1}^{\infty} P[A_i]$, falls $A = \bigsqcup_{i=1}^{\infty} A_i$
3. $\mathbb{P}[A^{\complement}] = 1 - \mathbb{P}[A]$
4. $\mathbb{P}[A \cup B] = \mathbb{P}[A] + P[B] - \mathbb{P}[A \cap B]$
5. $A \subseteq B \implies \mathbb{P}[A] \leq \mathbb{P}[B]$ (Monotonie)
6. $\mathbb{P}[\bigcup_{i=1}^{\infty} A_i] \leq \sum_{i=1}^{\infty} \mathbb{P}[A_i]$ (Union Bound)
7. Falls $A_1, \ldots A_n$ paarweise disjunkt, so gilt:

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

Fast sichere Ereignisse

Wir sagen A tritt fast sicher (f.s.) ein, falls $P[A] = 1$.

1.3 Bedingte Wahrscheinlichkeiten

Defintion Bedingte Wahrscheinlichkeit

$$\mathbb{P}[A|B] = \frac{\mathbb{P}[A \cap B]}{\mathbb{P}[B]}, \text{ wobei } \mathbb{P}[B] > 0$$

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

Totale Wahrscheinlichkeit

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

Satz von Bayes

$$\mathbb{P}[B_i|A] = \frac{\mathbb{P}[A|B_i]\ \mathbb{P}[B_i]}{\mathbb{P}[A]} = \frac{\mathbb{P}[A|B_i]\ \mathbb{P}[B_i]}{\sum_{k=1}^{n} \mathbb{P}[A|B_k]\ \mathbb{P}[B_k]}$$

Wobei $\mathbb{P}[A] > 0$ gelten muss.

1.4 Unabhängigkeit von Ereignissen

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

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

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

Unabhängigkeit von mehreren Ereignissen

$A_1, \ldots, A_n \in F$ sind unabhängig, falls:

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

2 Zufallsvariablen

Definition Gewichtsfunktion

$$p_X : \mathbb{R} \mapsto [0,1]: \ p_X(x) = \mathbb{P}[X = x] = \mathbb{P}[\{\omega | X(\omega) = x\}]$$

Es gilt: $\sum_{x \in \mathcal{W}(X)} p_X(x) = 1$ und $\forall x \in \mathcal{W}_X : 0 \leq p_X(x) \leq 1$

Definition Verteilungsfunktion

$$F_X : \ \mathbb{R} \mapsto [0,1] \quad \forall x \in \mathbb{R} \quad F_X(x) = \mathbb{P}[X \leq x]$$

- $F_X(x) = \int_{-\infty}^{x} f_X(x) dx$ bzw. $F_X(x) = \sum_{y \in \mathcal{W}_X, y \leq x} p(y)$
- $a < b \implies \mathbb{P}[a \leq X \leq b] = F_X(b) - F_X(a)$
- F ist monoton wachsend
- F ist rechtsstetig, d.h. $\lim_{t \to 0} F(x+t) = F(x)$
- $\lim_{x \to -\infty} F_X(x) = 0$ und $\lim_{x \to \infty} F_X(x) = 1$
- Median berechnen: x ist Median $\iff F_X(x) = 0.5$

2.1 Unabhängigkeit von ZV

$X_1, \ldots, X_n$ sind unabhängig, ist äquivalent zu:

- Diskreter Fall: $p(x_1, \ldots, x_n) = p_{X_1}(x_1) \cdot \ldots \cdot p_{X_n}(x_n)$
- Stetiger Fall: $f(x_1 \ldots x_n) = f_{X_1}(x_1) \cdot \ldots \cdot f_{X_n}(x_n)$
- $F_{X_1,\ldots,X_n}(x_1 \ldots x_n) = F_{X_1}(x_1) \cdot \ldots \cdot F_{X_n}(x_n)$
- Für alle $\phi_1, \ldots, \phi_n : \mathbb{R} \mapsto \mathbb{R}$ gilt: $\mathbb{E}[\phi_1(x_1) \cdot \ldots \cdot (x_n)] = \mathbb{E}[\phi_1(x_1)] \cdot \ldots \cdot \mathbb{E}[\phi_n(x_n)]$

$X_1, X_2, \ldots$ ist **unabhängig und identisch verteilt (uiv.)**, falls $\forall i, j \quad F_{X_i} = F_{X_j}$ gilt. Seien $X_1, ..., X_n$ (diskret) unabhängig und $g_i : \mathbb{R} \to \mathbb{R}$. Dann sind auch $g_1(X_1), ..., g_n(X_n)$ unabhängig.

2.2 Diskrete Zufallsvariablen

Definition diskrete ZV

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

2.3 Diskrete Verteilungen

(Diskrete) Gleichverteilung (Laplace)

$\mathcal{W} = \{x_1, \ldots, x_N\}$. $\mathbb{P}[X = x_i] = \frac{1}{N}$ und $\mathbb{E}[X] = \frac{\sum_{i=1}^{N} x_i}{N}$

Bernoulli-Verteilung

Hat nur die Ereignisse $\{0, 1\}$. $X \sim \text{Ber}(p)$

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

$$\mathbb{E}[X] = p \text{ und } \operatorname{Var}[X] = p(1-p)$$

Binomialverteilung

Wiederholung von Bernoulli-Exp. $X \sim \text{Bin}(n,p)$

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

$$\mathbb{E}[X] = np \quad \text{und} \quad \text{Var}[X] = np(1-p)$$

Falls $X \sim \text{Bin}(n_x,p)$ und $Y \sim \text{Bin}(n_y,p)$ **unabhängig** sind, dann gilt für $Z := X+Y$, dass $Z \sim \text{Bin}(n_x+n_y,p)$.

Negative Binomialverteilung

Warten auf den r-ten Erfolg von Bernoulli Experimenten (Falls $r=1$, so ist $X \sim \text{Geo}(p)$). $X \sim \text{NB}(r,p)$

$$\mathbb{P}[X=k] = \binom{k-1}{r-1} p^r (1-p)^{k-r}$$

$$\mathbb{E}[X] = \frac{r}{p} \quad \text{und} \quad \text{Var}[X] = \frac{r\cdot(1-p)}{p^2}$$

$(X_i)_{i=1}^r \sim \text{Geo}(p)$ und **unabh.** $\Rightarrow X := X_1 + \dots X_r \sim \text{NB}(r,p)$.

Geometrische Verteilung

Beschreibt das 1. Auftreten eines Erfolges. $X \sim \text{Geom}(p)$

$$\mathbb{P}[X=k] = (1-p)^{k-1} \cdot p \qquad F_X(k) = \mathbb{P}[X \leq k] = 1-(1-p)^k$$

$$\mathbb{E}[X] = \frac{1}{p} \quad \text{und} \quad \text{Var}[X] = \frac{1-p}{p^2}$$

Gedächtnislosigkeit: $\mathbb{P}[X > y+x | X > x] = \mathbb{P}[X > y]$

Hypergeometrische Verteilung

Urne mit n Elementen, davon r vom Typ 1 und $n-r$ vom Typ 2. Ziehe m Elemente daraus, dann beschreibt X die Anzahl Elemente vom Typ 1.

$$\mathbb{P}[X=k] = \frac{\binom{r}{k}\binom{n-r}{m-k}}{\binom{n}{m}} \quad \text{für } k \in \{0,1,...,min(m,r)\}$$

$$\mathbb{E}[X] = m \cdot \frac{r}{n} \quad \text{und} \quad \text{Var}[X] = m \cdot \frac{r}{n}(1-\frac{r}{n})\frac{n-m}{n-1}$$

Cauchy-Verteilung: Sind $Y,Z \sim \mathcal{N}(0,1)$ unabhängig, so ist $X := Y/Z$ Cauchy-verteilt. (EW und Varianz existieren nicht!)
$f_X(x) = \frac{1}{\pi} \cdot \frac{1}{1+x^2}$ und $F_X(x) = \frac{1}{1} + \frac{1}{\pi} \cdot arctan(x)$

Poisson-Verteilung

Annäherung an die Binomialvert. (0-1-Exp.) für grosse n und kleine p (d.h. rare Ereignisse). $X \sim \text{Poisson}(\lambda)$

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

$$\mathbb{E}[X] = \lambda \quad \text{und} \quad \text{Var}[X] = \lambda$$

Sind X und Y **unabhängige** ZV mit $X \sim \text{Po}(\lambda_1)$ und $Y \sim \text{Po}(\lambda_2)$, dann gilt für $Z := X+Y \sim \text{Po}(\lambda_1+\lambda_2)$.

2.4 Stetige Zufallsvariablen

Definition stetige ZV

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

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

Die Verteilungsfunktion F_X sei stetig und stückweise C^1. Dann ist X eine stetige ZV und $f(x) = F_X'(x)$.

Hierbei ist $f: \mathbb{R} \mapsto \mathbb{R}^+$ die **Dichtefunktion** von X. Es gilt:

$$\int_{-\infty}^{+\infty} f(y)dy = 1 \quad \text{und} \quad \forall x \in \mathbb{R}: f_X(x) \geq 0$$

Es gelten folgende Eigenschaften: $\mathbb{P}[a \leq x \leq b] = \mathbb{P}[a < x < b]$ und $\mathbb{P}[X \in [a,b]] = \mathbb{P}[X \in (a,b)]$
Zudem gilt für stetige ZV: $\mathbb{P}[\mathbf{X} = \mathbf{x}] = \mathbf{0}$,$\forall x \in W_x$

2.5 Stetige Verteilungen

(Stetige) Gleichverteilung

Jedes Ereignis hat die gleiche W'keit. $X \sim \mathcal{U}[a,b]$

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

$$F_{a,b}(x) = \int_{-\infty}^x f_{a,b}(t)dt = \begin{cases} 0 & x < a \\ \frac{x-a}{b-a} & a \leq x \leq b \\ 1 & x > b \end{cases}$$

$$\mathbb{E}[X] = \frac{a+b}{2} \quad \text{und} \quad \text{Var}[X] = \frac{(a-b)^2}{12}$$

Exponentialverteilung

Das stetige Gegenstück zur Geometrischen Verteilung (verwendet für Warte- & Überlebenszeiten). $X \sim \text{Exp}(\lambda)$

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

$$F(x) = \int_0^x f(t)dt = \begin{cases} 1-e^{-\lambda \cdot x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

$$\mathbb{E}[X] = \frac{1}{\lambda} \quad \text{und} \quad \text{Var}[X] = \frac{1}{\lambda^2}$$

Skalierung: Für $X \sim \text{Exp}(\lambda)$, $a > 0$ und $Y := aX$: $Y \sim \text{Exp}(\frac{\lambda}{a})$.
Wartezeit: $\forall x \geq 0: \mathbb{P}[X > x] = 1 - F(x) = e^{-\lambda \cdot x}$
Gedächtnislosigkeit: $\mathbb{P}[X > y+x | X > x] = \mathbb{P}[X > y]$

Normalverteilung

Additive Überlagerung einer großen Zahl von unabhängigen Einflüssen (Bsp: Messfehler). $X \sim \mathcal{N}(m,\sigma^2)$

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

$$\mathbb{E}[X] = m \quad \text{und} \quad \text{Var}[X] = \sigma^2$$

$\Rightarrow$ 68.27% der Werte liegen im Intervall $\mu \pm \sigma$, 95,45% in $\mu \pm 2\sigma$ und 99,73% in $\mu \pm 3\sigma$. Zudem gilt: $\mathbb{P}[X \in [\mu - z\sigma, \mu + z\sigma]] = 2\phi(z) - 1$

Sei $X_i \sim \mathcal{N}(m_i, \sigma_i^2)$ **unabhängig**, dann: $Z := (m_0 + \lambda_1 \cdot X_1 + ... + \lambda_n \cdot X_n) \sim \mathcal{N}(m_0 + \lambda_1 \cdot m_1 + ... + \lambda_n \cdot m_n,\ \lambda_1^2 \cdot \sigma_1^2 + \dots + \lambda_n^2 \cdot \sigma_n^2)$.

Standardnormalvert.: Sei $Z \sim \mathcal{N}(0,1)$. Für eine normalverteilte ZV X gilt $X = m + \sigma \cdot Z \sim \mathcal{N}(m,\sigma^2)$.
$\Rightarrow \mathbb{P}[X \leq x] = \mathbb{P}[Z \leq \frac{x-\mu}{\sigma}] = \phi(\frac{x-\mu}{\sigma})$
Ist $X \sim \mathcal{N}(\mu,\sigma^2)$, so ist $Y := e^X$ logarithmisch normalvert.

3 Erwartungswert und Varianz

Erwartungswert für diskrete ZV

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

Sei $\phi: \mathbb{R} \mapsto \mathbb{R}$ eine Abbildung. Falls die Summe wohldefiniert ist (abs. konv.), gilt: $\mathbb{E}[\phi(X)] = \sum_{x \in W} \phi(x) \cdot \mathbb{P}[X=x]$
Achtung: EW existiert nicht immer! (Summe muss abs. konv.)

Erwartungswert für stetige ZV

Falls die Verteilung von X absolut stetig ist, gilt:

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

$\Rightarrow$ EW existiert nur, falls das Integral **absolut konvergiert!**
Sei $\phi : \mathbb{R} \mapsto \mathbb{R}$ eine Abbildung, sodass $\phi(x)$ eine Zufallsvariable ist. Es gilt: $\mathbb{E}[\phi(X)] = \int_{-\infty}^{\infty} \phi(x)f(x)dx$
Man kann den EW auch über die Verteilungsfunktion definieren:

$$\mathbb{E}[X] = \int_0^{\infty} (1 - F_X(x))dx - \int_{-\infty}^{0} F_X(x)dx$$

$\Rightarrow$ Für ZV mit nicht-neg. Werten gilt: $\mathbb{E}[X] = \int_0^{\infty}(1-F_X(x))dx$.

3.1 Rechnen mit Erwartungswerten

- **Linearität:** Seien X, Z Zufallsvariabeln mit $a, b \in \mathbb{R}$. Falls die Erwartungswerte wohldefiniert sind, gilt
$$\mathbb{E}[a \cdot X + b] = a \cdot \mathbb{E}[X] + b$$
- **Monotonie:** $X(\omega) \leq Y(\omega), \forall \omega \in \Omega \Rightarrow \mathbb{E}[X] \leq \mathbb{E}[Y]$
- **Mon. Stetigkeit**: $0 \leq X_1 \leq X_2 ... \implies \mathbb{E}[\lim_n X_n] = \lim_n \mathbb{E}[X_n]$
- **Satz von Lebesgue:** Sei $X_1, X_2...$ f.s. konv. Folge mit $|X_i(\omega)| \leq X(\omega)$ für ZV X mit $\mathbb{E}[X] < \infty$, dann: $\mathbb{E}[\lim_n X_n] = \lim_n \mathbb{E}[X_n]$
- Für paarweise disjunkte Ereignisse $(A_i)_{i=1}^n$, $\bigcup_{i=1}^n A_i = \Omega$ und $\forall A_i : \mathbb{P}[A_i] > 0$, gilt:
$$\mathbb{E}[X] = \sum_{i=1}^{n} \mathbb{E}[X|A_i] \cdot \mathbb{P}[A_i]$$
- **Mulitplikativität:** Falls $X_1, ..., X_n$ **unabhängig** $\Rightarrow$ $\mathbb{E}[g_1(X_1) \cdot ... \cdot g_n(X_n)] = \mathbb{E}[g_1(X_1)] \cdot ... \cdot \mathbb{E}[g_n(X_n)]$
- **Extremwertformel:** Sei X eine diskrete ZV mit Werten in $\mathbb{N}$. Dann gilt folgende Identität (gilt auch analog für stetige ZV):
$$\mathbb{E}[X] = \sum_{n=1}^{\infty} \mathbb{P}[X \geq n]$$
- **Waldsche Identität:** Seien $X_1, X_2, \ldots, X_n$ uiv. ZV und N eine ZV mit Werten in $\mathbb{N}$. Dann gilt:
$$\mathbb{E}[\sum_{i=1}^{N} X_i] = \mathbb{E}[N] \cdot \mathbb{E}[X_1]$$
- **Alternativdefinition unabhängige ZV** Seien $X_1, \ldots, X_n$ diskrete ZV. Dann gilt: $X_1, \ldots, X_n$ sind unabhängig $\iff$ Für jedes $\phi_1, \ldots, \phi_n : \mathbb{R} \mapsto \mathbb{R}$ beschränkt, gilt:
$$\mathbb{E}[\phi_1(X_1) \cdot \ldots \cdot \phi_n(X_n)] = \mathbb{E}[\phi_1(X_1)] \cdot \ldots \cdot \mathbb{E}[\phi_n(X_n)]$$

3.2 Bedingter Erwartungswert

Bedingter Erwartungswert

$$\mathbb{E}[X|Y] = \frac{\mathbb{E}[I_A \cdot X]}{\mathbb{P}[Y]} = \sum_{x \in X(\Omega)} x \cdot \mathbb{P}[X = x|Y]$$

Eigenschaften:

- $\mathbb{E}[a_1 \cdot X_1 + a_2 \cdot X_2|A] = a_1 \cdot \mathbb{E}[X_1|A] + a_2 \cdot \mathbb{E}[X_2|A]$
- Ist X unabhängig von Y, so gilt: $\mathbb{E}[X|Y] = \mathbb{E}[X]$

Bedingt auf eine Partition

Sei $\mathcal{B} = (B_i)_{i \in I}$ eine Partition von Ω. Dann definieren wir die **Zufallsvariabel**:

$$\mathbb{E}[X|B](\omega) = \sum_{i \in I} \mathbb{E}[X|B_i] \cdot 1_{B_i}(\omega)$$

Eigenschaften:

- Totaler Erwartungswert: $\mathbb{E}[\mathbb{E}[X|\mathcal{B}]] = \mathbb{E}[X]$
- $\mathbb{E}[1_{B_i} \cdot \mathbb{E}[X|\mathcal{B}]] = \mathbb{E}[1_{B_i} \cdot X]$

Beispiel: Zwei Würfel werden geworfen. Sei $\Omega = \{1, 2, \ldots, 6\}^2$ und $\mathbb{P}[(\omega_1, \omega_2)] = 1/36 \ \forall \omega \in \Omega$. Sei $Y :=$ "Augenzahl des 2. Würfels" und sei $\mathcal{B} = (B_i)_{i=1}^6$ wobei $B_i = \{Y = i\}$. Berechne $\mathbb{E}[Y|\mathcal{B}](2, 4)$.
$\Rightarrow \mathbb{E}[Y|\mathcal{B}](2, 4) = \sum_{i=1}^6 \mathbb{E}[Y|B_i] \cdot 1_{B_i}(2, 4) = \mathbb{E}[Y|B_4] = \frac{\mathbb{E}[1_{Y=4} \cdot Y]}{\mathbb{P}[Y=4]} = \frac{4 \cdot \mathbb{P}[(1,4)] + 4 \cdot \mathbb{P}[(2,4)] + \ldots + 4 \cdot \mathbb{P}[(6,4)]}{\mathbb{P}[Y=4]} = \frac{24/36}{1/6} = 4$

3.3 Varianz

Varianz

Sei X eine ZV sodass $\mathbb{E}[X^2] < \infty$.

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

$$\text{Var}[X] = \begin{cases} \sum_{x \in W_X} (x - \mathbb{E}[X])^2 \cdot \mathbb{P}[X = x] & , X \textbf{ diskret} \\ \int_{-\infty}^{\infty} (x - \mathbb{E}[X])^2 \cdot f_X(x)dx & , X \textbf{ stetig} \end{cases}$$

Eigenschaften:

- $\text{Var}[a \cdot X + b] = a^2 \cdot \text{Var}[X]$
- $X_1, ...X_n$ **unabhängig** $\Rightarrow \text{Var}[\sum_{i=1}^n X_i] = \sum_{i=1}^n \text{Var}[X_i]$
- Allgemein (Cov$(X_i, Y_i) = 0$, wenn X_i und Y_i unabhängig) :
$$\text{Var}[\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n} \text{Var}[X_i] + 2 \cdot \sum_{i<j} \text{Cov}(X_i, Y_i)$$

3.4 Kovarianz

Definition Kovarianz

Sei X, Y zwei ZV mit $\mathbb{E}[X^2], \mathbb{E}[Y^2] < \infty$.

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

Eigenschaften der Kovarianz:

- $\text{Cov}(X, X) = \text{Var}[X]$
- X, Y unabhängig $\implies \text{Cov}(X, Y) = 0$
- $\text{Cov}(X, Y) = \text{Cov}(Y, X)$
- $\text{Cov}(X, a \cdot Y + b) = a \cdot \text{Cov}(X, Y)$ und $\text{Cov}(X, b) = 0$
- $\text{Cov}(X, Y + Z) = \text{Cov}(X, Y) + \text{Cov}(X, Z)$
- $|\text{Cov}(X, Y)| \leq \sigma(X) \cdot \sigma(Y)$

Korrelation

$$\rho(X, Y) := \begin{cases} \frac{\text{Cov}(X,Y)}{\sigma(X) \cdot \sigma(Y)} & , \text{falls } \sigma(X) \cdot \sigma(Y) > 0 \\ 0 & , \text{sonst} \end{cases}$$

Korrelation misst den lin. Zshg. Es gilt $-1 \leq \rho(X, Y) \leq 1$.
Ist $\text{Cov}(X, Y) = 0$, so nennt man X und Y **unkorreliert**.
Es gilt: unabhängig $\Rightarrow$ paarweise unabhängig $\Rightarrow$ unkorreliert

3.5 Momente

Sei X eine Zufallsvariable und $p \in R_+$. Wir definieren:

- *das p-te absolute Moment von X* durch $M_p := \mathbb{E}[|X|^p]$ (kann ∞ sein)
- falls $M_n < \infty$ für ein n, dann ist das *n-te (rohe) Moment von X* durch $m_n := \mathbb{E}[X^n]$ definiert.
- Das *n-te zentralisierte Moment von X* durch $\mu_n := \mathbb{E}[(X - \mathbb{E}[X])^n]$ definiert. (für $n = 2$ erhalten wir die Varianz)

Damit folgt sofort: $M_n < \infty$ für $n \in \mathbb{N} \Rightarrow |m_n| \leq M_n$.
Hat X eine Dichte f_X, dann gilt zudem für das absolute Moment:

$$M_p = \int_{-\infty}^{\infty} |x|^p f_X(x) \ dx$$

Gilt dann $M_n < \infty$ für ein $n \in \mathbb{N}$, dann können wir auch das n-te Moment per Integral bestimmen:

$$m_n = \int_{-\infty}^{\infty} x^n f_X(x) \ dx$$

und für diskrete ZV: $m_n = \sum_{i=1}^{\infty} x_i^n \cdot \mathbb{P}[X = x_i]$

4 Gemeinsame Verteilungen

Diskrete gemeinsame Verteilung

Seien $X_1, \dots, X_n$ diskrete ZV wobei $X_i \in W_i$ für $W_i \subset \mathbb{R}$. Die gemeinsame Verteilung von $X_1, \dots, X_n$ ist:

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

Die **Randverteilung** ist definiert als $F_X(x) := \mathbb{P}[X \leq x] = \mathbb{P}[X \leq x, Y \leq \infty] = \lim_{y \to \infty} F(x,y)$ und die **Randdichte** erhalten wir durch "wegaddieren": $p_X(x) = \sum_{y \in \mathcal{W}(Y)} p(x,y)$
Der **Erwartungswert** ist definiert als ($\phi : \mathbb{R}^n \to \mathbb{R}$):

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

Stetige gemeinsame Verteilung

Falls die gemeinsame Verteilungsfunktion F von $X_1, ..., X_n$ (stetige ZV) sich schreiben lässt als

$$F(x_1, \dots, x_n) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} f(t_1, \dots, t_n) dt_n \dots dt_1$$

für eine Funktion $f : \mathbb{R}^n \to [0, \infty)$, dann ist f die **gemeinsame Dichte**.

Es gilt: $\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(t_1, \dots, t_n) dt_n \dots dt_1 = 1$
Seien $X_1, \dots, X_n$ stetige ZV mit einer gemeinsamen Dichte f und $\phi : \mathbb{R}^n \mapsto \mathbb{R}$. Dann ist der **Erwartungswert** definiert als:

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

Die **Randverteilung** ist definiert als $F_X(x) := \mathbb{P}[X \leq x] = \mathbb{P}[X \leq x, Y \leq \infty] = \lim_{y \to \infty} F(x,y)$ und die **Randdichte** erhalten wir durch "wegintegrieren": $f_X(x) = \int_{-\infty}^{\infty} f(x,y) dy$.

Bedingte Dichte, Verteilung und EW

Die bedingte Verteilung für zwei stetige ZV X, Y ist (gilt analog auch für diskrete ZV):

$$f_{X|Y}(x \mid y) := \frac{f(x,y)}{f_Y(y)} \text{ , für } f_Y(y) > 0 \text{ und } 0 \text{ sonst.}$$

$$\mathbb{P}[Y > t \mid Y < a] = \frac{P[t < Y < a]}{P[Y < a]}$$

$$E[X_1 \mid X_2](x_2) = \int x_1 f_{X_1|X_2}(x_1 \mid x_2)\ dx_1$$

4.1 Konvolution (Faltung)

Konvolution (Faltung)

Seien X und Y **unabhängige** kontinuierliche Zufallsvariabeln. Für die Dichte von $Z := X + Y$ gilt:

$$f_Z(z) = (f_X * f_Y)(z) = \int_{-\infty}^{\infty} f(x, z-x) dx \overset{uiv.}{=} \int_{-\infty}^{\infty} f_X(x) \cdot f_Y(z-x) dx$$

und für die Verteilung erhalten wir:

$$F_Z(z) = \int_{-\infty}^{\infty} \int_{-\infty}^{z} f(x, v-x) dv dx = \int_{-\infty}^{\infty} F_X(z-x) f_Y(x) dx$$

Für diskrete Zufallsvariabeln gilt analog:

$$p_Z(z) = p_X(z) * p_Y(z) = \sum_{x \in \mathcal{W}(X)} p_X(x) \cdot p_Y(z-x)$$

Es gilt: $\text{Ber}(p) * \text{Ber}(p) = \text{Bin}(2,p)$, $\text{Bin}(n,p) * \text{Bin}(m,p) = \text{Bin}(n+m,p)$, $\text{Poi}(\lambda_1) * \text{Poi}(\lambda_2) = \text{Poi}(\lambda_1 + \lambda_2)$, $\text{Geo}(p) * \text{Geo}(p) = \text{NB}(2,p)$, $\text{NB}(r,p) * \text{NB}(s,p) = \text{NB}(r+s,p)$, $\mathcal{N}(0,1) * \mathcal{N}(0,1) = \mathcal{N}(0,2)$ und $\mathcal{N}(\mu_1, \sigma_1^2) * \mathcal{N}(\mu_2, \sigma_2^2) = \mathcal{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$
Beispiel: Seien X, Y unabh. und exp.vert. mit $\lambda > 0$, berechne f_Z von $Z := X + Y$. $\Rightarrow f_Z(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z-x) dx = \int_0^z \lambda e^{-\lambda x} \lambda e^{-\lambda(z-x)} dx = \int_0^z \lambda^2 e^{-\lambda z} dx = \lambda^2 z e^{-\lambda z}$, für $z \geq 0$. Wir sehen also $Z \sim Ga(2, \lambda)$.

5 Ungleichungen

Markov-Ungleichung

Sei X eine ZV die nur **nichtnegative** Werte annimmt, dann gilt für jedes $a > 0$:

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

Sei $g : \mathcal{W}(X) \to [0, \infty)$ eine wachsende Funktion. Dann gilt sogar:

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

Chebychev-Ungleichung

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

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

Oder äquivalent: $\mathbb{P}[|X - \mathbb{E}[X]| \geq a \cdot \sqrt{\text{Var}[X]}] \leq \frac{1}{a^2}$

Jensen-Ungleichung

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

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

$\Rightarrow$ Somit gilt: $|\mathbb{E}[X]| \leq \mathbb{E}[|X|]$ und $\mathbb{E}[|X|] \leq \sqrt{\mathbb{E}[X^2]}$

6 Grenzwertsätze

Seien im folgenden $X1, \dots, X_n$ **u.i.v.** ZV (und somit $\mathbb{E}[X_i] = \mathbb{E}[X_j]$ und $\text{Var}[X_i] = \text{Var}[X_j]$ für $i, j \in \{1, \dots, n\}$).

Gesetz der grossen Zahlen

Sei $\mathbb{E}[|X_i|] < \infty$, so gilt (fast sicher):

$$\lim_{n \to \infty} \frac{X_1 + \ldots + X_n}{n} = \mathbb{E}[X_i]$$

Zentraler Grenzwertsatz

Sei $S_n := X_1 + \ldots + X_n$, $\mu = \mathbb{E}[X_i]$ und $\sigma^2 = \text{Var}(X_i)$, so gilt:

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

für alle $a \in \mathbb{R}$, wobei Φ die Verteilungsfunktion der $\mathcal{N}(0,1)$-Verteilung ist.

Der zentrale Grenzwertsatz sagt aus, dass die Verteilung einer ZV $Z_n = \frac{S_n - nm}{\sigma \sqrt{n}}$ ungefähr wie die Verteilung von $\mathcal{N}(0,1)$ aussieht, also $\mathbb{P}[Z_n \leq x] \approx \Phi(x)$ (für grosse n). Und somit auch $S_n \overset{\text{approx.}}{\sim} \mathcal{N}(n\mu, n\sigma^2)$. Zudem: $\lim_{n \to \infty} \mathbb{P}[a \leq Z_n \leq b] = \int_a^b \varphi(x) dx$

7 Random Walk

Definition Simple Random Walk: $0 < p < 1$ und $X_i = \{-1, 1\}$ und $\mathbb{P}[X_i = +1] = p$ bzw. $\mathbb{P}[X_i = -1] = 1 - p$.
Gambler's Ruin: Nehmen wir einen Simple Rand. Walk mit Anfangsposition $a \geq 0$. Das Spiel ist beendet, falls der Zocker ein Kontostand von N oder 0 erreicht. Sei $p_{win}(a) := \mathbb{P}$["Kontostand N erreicht von Startkapital a aus"] (wobei $0 < a < N$). Es gilt folgende Rekurrenz: $p_{win}(a) = p \cdot p_{win}(a+1) + (1-p) \cdot p_{win}(a-1)$
Offensichtlich gilt: $p_{win}(0) = 0$ (bereits verloren) und $p_{win}(N) = 1$ (bereits gewonnen). Wir erhalten für $p_{win}(a)$ (und 1-$p_{win}(a)$ für Wahrscheinlichkeitkeit von Ruin):

$$p_{win}(a) = \begin{cases} \frac{1 - (\frac{1-p}{p})^a}{1 - (\frac{1-p}{p})^N} & \text{wenn } p \neq (1-p) \\ \frac{a}{N} & \text{wenn } p = (1-p) = 1/2 \end{cases}$$

8 Induktive Statistik: Basics

Idee: Man fasst die Daten $x_1, ..., x_n$ auf als Realisierungen $X_1(\omega), ..., X_n(\omega)$ von Zufallsvariablen $X_1, ..., X_n$, und sucht dann (unter geeigneten Zusatzannahmen) Aussagen über die Verteilung von $X_1, ..., X_n$. Die $x_1, ..., x_n$ werden *Stichprobe* genannt, das n ist der *Stichprobenumfang*.

8.1 Schätzer

Sei $\vec{X} = (X_1, \ldots, X_n)$ eine Stichprobe und sei $\Theta \subseteq \mathbb{R}$ der Parameterraum. Wir suchen für den Parameter $\vartheta \in \Theta$ einen Schätzer T aufgrund unserer Stichprobe.

Schätzer

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

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

mit Schätzwert $T(\omega) = t(X_1(\omega), ..., X_n(\omega)) = t(x_1, ..., x_n)$

Ein Schätzer T ist **erwartungstreu** (unbiased) für ϑ, falls für alle $\vartheta \in \Theta$ gilt:

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

$\Rightarrow$ Ist T *erwartungstreu*, so schätzt T *im Mittel* richtig.

Der **Bias** (erwartete Schätzfehler) von T im Modell $\mathbb{P}_\vartheta$ ist:

$$\mathbb{E}_\vartheta[T - \vartheta] = \mathbb{E}_\vartheta[T] - \vartheta$$

$\Rightarrow$ Für *erwartungstreue* Schätzer ist der Bias immer gleich Null.

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

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

$\Rightarrow$ Für erwartungstreue Schätzer sind Varianz und MSE dasselbe.

Eine Schätzvariable T heisst **konsistent**, wenn $\text{MSE}_\vartheta[T] \to 0$ für $n \to 0$ gilt. Für genügend grosse n liegen also die Werte von T beliebig nahe am gesuchten Wert ϑ.

Beispiel: Für $\vartheta = \mathbb{E}[X]$ bzw. $\vartheta = \text{Var}[X]$ eignen sich folgende Schätzer:

$$\overline{X}_n := \frac{1}{n}\sum_{i=1}^{n} X_i \quad \text{und} \quad S_n^2 := \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$$

$\overline{X}_n$ und S_n^2 heissen (empirisches) *Stichprobenmittel* bzw. *Stichprobenvarianz* der Stichprobe $X_1, \ldots, X_n$ u.i.v. und sind *erwartungstreue* Schätzer für den Erwartungswert bzw. Varianz.

8.2 Maximum-Likelihood-Methode

Die Likelihood-Funktion L entspricht der Wahrscheinlichkeit, dass wir die Stichprobe $x_1, ..., x_n$ erhalten, wenn wir den Parameter mit dem Wert ϑ belegen.

Likelihood Funktion L

$$L(x_1, \ldots, x_n; \vartheta) := \begin{cases} p_{\vec{X}}(x_1, ..., x_n; \vartheta) & \text{(diskreter Fall)} \\ f_{\vec{X}}(x_1, \ldots, x_n; \vartheta) & \text{(stetiger Fall)} \end{cases}$$

Wobei $p_{\vec{X}}(x_1, ..., x_n; \vartheta) = \mathbb{P}_\vartheta[X_1 = x_1, ..., X_n = x_n]$ und falls X_i unter $\mathbb{P}_\vartheta$ u.i.v, dann sogar: $p_{\vec{X}}(x_1, ..., x_n; \vartheta) = \prod_{i=1}^n p_X(x_i; \vartheta)$.

Maximum-Likelihood-Schätzer T_{ML}

Der ML-Schätzer T_{ML} maximiert $\vartheta \mapsto L(x_1, \ldots, x_n; \vartheta)$:

$$T_{ML} = t_{ML}(X_1, \ldots, X_n) \in \arg\max_{\vartheta \in \Theta} L(x_1, \ldots, x_n; \vartheta)$$

Oder anders definiert: $\hat{\vartheta}$ ist der ML-Schätzwert, wenn gilt: $L(x_1, ..., x_n; \vartheta) \leq L(x_1, ..., x_n; \hat{\vartheta})$ für alle ϑ.

Anwendung der Methode:

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

$\Rightarrow$ Unter gefundenem ϑ ist die Likelihood-Funktion maximal.

Beispiel: (Bernoulli-Verteilung) Sei $X_1, \ldots, X_n$ u.i.v. $\sim \text{Be}(p)$, also $\vartheta = p$. Ziel: Wir wollen Parameter p schätzen.

1. $\Rightarrow p_X(x; \vartheta) = \mathbb{P}_\vartheta[X = x] = \vartheta^x \cdot (1 - \vartheta)^{1-x}$ für $x \in \{0, 1\}$
2. $L(x_1, ..., x_n; \vartheta) \overset{u.i.v.}{=} \prod_{i=1}^n p_X(x_i; \vartheta) = \vartheta^{\sum_{i=1}^n x_i} \cdot (1-\vartheta)^{n - \sum_{i=1}^n x_i}$
   $\Rightarrow log\ L(x_1, ..., x_n; \vartheta) = \sum_{i=1}^n x_i log(\vartheta) + (n - \sum_{i=1}^n x_i) log(1 - \vartheta)$
3. $\frac{\partial}{\partial \vartheta} log\ L(x_1, ..., x_n; \vartheta) = \frac{1}{\vartheta}\sum_{i=1}^n x_i - \frac{1}{1-\vartheta}(n - \sum_{i=1}^n x_i)$
4. Nullstelle bei: $\vartheta = \frac{1}{n}\sum_{i=1}^n x_i$, da $(1-\vartheta)\sum_{i=1}^n x_i = \vartheta(n - \sum_{i=1}^n x_i)$

$\Rightarrow$ Der ML-Schätzer für ϑ bzw. p ist also: $T = \frac{1}{n}\sum_{i=1}^n X_i =: \overline{X}$

9 Testen von Hypothesen

Grundproblem: Entscheidung zwischen zwei konkurrierenden Modellklassen zu treffen.
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 (z.B. $\Theta_0 = \{\vartheta_0\}$); sonst heissen sie *zusammengesetzt*. Die Hypothesen lauten also:

H_0: "der wahre Parameter ϑ liegt in Θ_0", also $\vartheta \in \Theta_0$
H_A: "der wahre Parameter ϑ liegt in Θ_A", also $\vartheta \in \Theta_A$

Definition Teststatistik

Man hat eine Abb. $t : \mathbb{R}^n \to \mathbb{R}$, $(x_1, ..., x_n) \mapsto t(x_1, ..., x_n)$ und einen *Verwerfungsbereich* $K \subseteq \mathbb{R}$. Die ZV $T := t(X_1, ..., X_n)$ heisst dann *Teststatistik*.

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

Wobei $T(\omega) = t(X_1(\omega), \ldots, X_n(\omega))$ gilt.

Es handelt sich um einen **Fehler 1. Art**, wenn H_0 fälschlicherweise verworfen wird, obwohl sie richtig ist. Wahrscheinlichkeit für einen Fehler 1. Art:

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

Es handelt sich um einen **Fehler 2. Art**, wenn H_0 fälschlicherweise akzeptiert wird, obwohl sie falsch ist:

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

Ziel: Fehler 1. und 2. Art möglichst klein, also $\vartheta \mapsto \mathbb{P}_\vartheta[T \in K]$ auf Θ_0 minimieren und auf Θ_A maximieren. Unser primäres Ziel, ist den Fehler 1. Art zu minimieren (die *Macht* zu maximieren).
$\Rightarrow$ Schwieriger die Hypothese zu verwerfen, anstatt zu behalten. Ein seriöser Test wird deshalb als Hypothese immer die Negation der eigentlich gewünschten Aussage benutzen.

Signifikanzniveau und Macht

Ein Test hat *Signifikanzniveau* $\alpha \in [0, 1]$ falls:

$$\forall \vartheta \in \Theta_0 : \ \mathbb{P}_\vartheta[T \in K] \leq \alpha \iff \underset{\vartheta \in \Theta_0}{sup}\ \mathbb{P}_\vartheta[T \in K] \leq \alpha$$

Die *Macht* wird definiert als:

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

Ziel: $\vartheta \mapsto \beta(\vartheta)$ möglichst gross, bzw. $1 - \beta(\vartheta)$ möglichst klein.

9.1 Konstruktion von Tests

Ziel: Systematischer Ansatz zum finden von K und T.
Sei $\vartheta_0 \in \Theta_0$ und $\vartheta_A \in \Theta_A$. Sei $X_1, \ldots, X_n$ diskret oder gemeinsam stetig unter $\mathbb{P}_{\vartheta_0}$ und $\mathbb{P}_{\vartheta_A}$, wobei $\vartheta_0 \neq \vartheta_A$.

Likelihood-Quotient

$$R(x_1,\dots,x_n;\vartheta_0,\vartheta_A) := \frac{L(x_1,\dots,x_n;\vartheta_A)}{L(x_1,\dots,x_n;\vartheta_0)}$$

(Falls $L(x_1,\dots,x_n;\vartheta_0)=0$ setzen wir $R(x_1,\dots,x_n)=+\infty$.)
$\Rightarrow$ Wenn $R \gg 1$, so gilt $H_A > H_0$ (H_A ist wahrscheinlicher) und analog wenn $R \ll 1$, so gilt $H_A < H_0$.

Likelihood-Quotient-Test

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

$$T = R(x_1,\dots,x_n;\vartheta_0,\vartheta_A) \quad \text{und} \quad K = (c,\infty]$$

$\Rightarrow$ Der LQ-Test ist optimal, wenn Hypothese und Alternative beide *einfach* sind, da jeder andere Test mit kleinerem Signifikanzniveau auch eine kleinere Macht hat (*Neyman-Pearson-Lemma*).

9.2 Häufige Fälle

Einen allgemein approximativen Zugang liefert der ZGS. Oft ist ein Schätzer T eine Funktion einer Summe $\sum_{i=1}^n Y_i$, wobei die Y_i u.i.v. sind. Nach dem ZGS ist dann für grosse n $\sum_{i=1}^n Y_i$ approximativ normalverteilt mit $\mu = n\cdot\mathbb{E}_\vartheta[Y_i]$ und $\sigma^2 = n\cdot\mathrm{Var}_\vartheta[Y_i]$. Für normalverteilte Stichproben hat man exakte Aussagen (siehe z-Test und t-Test).

Approximativer Binomialtest

Annahmen: $X_1,\dots,X_n$ u.i.v. mit $X_i \sim Ber(p)$ mit $\vartheta = p$ unbekannt. Sei n hinreichend gross.
Hypothesen:

1. $H_0: p = p_0$ gegen $H_A: p \neq p_0$
2. $H_0: p \geq p_0$ gegen $H_A: p < p_0$
3. $H_0: p \leq p_0$ gegen $H_A: p > p_0$

Testgrösse:

$$T := \frac{\sum_{i=1}^n X_i - n\cdot p_0}{\sqrt{n\cdot p_0\cdot(1-p_0)}} \sim \mathcal{N}(0,1) \quad \text{nach ZGS}$$

Ablehnungskriterium: für H_0 bei Signifikanzniveau α:

1. $|T| > z_{1-\alpha/2}$
2. $T < z_\alpha$
3. $T > z_{1-\alpha}$

Herleitung: Wir wissen: $T \sim \mathcal{N}(0,1)$ unter $\mathbb{P}_\vartheta$. Seien $K_>, K_<$ und $K_{\neq}$ die Verwerfungsbereiche zu den Fällen 1-3.

1. $\alpha = \mathbb{P}_{\vartheta_0}[T \in K_{\neq}] = \mathbb{P}_{\vartheta_0}[T < -c_{\neq}] + \mathbb{P}_{\vartheta_0}[T > c_{\neq}] = \mathbb{P}_{\vartheta_0}[T < -c_{\neq}] + (1-\mathbb{P}_{\vartheta_0}[T \leq c_{\neq}]) = 2\cdot(1-\Phi(c_{\neq})) \quad \Rightarrow c_{\neq} = z_{1-\alpha/2}$
2. $\alpha = \mathbb{P}_{\vartheta_0}[T \in K_<] = \mathbb{P}_{\vartheta_0}[T < c_<] = \mathbb{P}_{\vartheta_0}[T > -c_<] = 1-\mathbb{P}_{\vartheta_0}[T \leq -c_<] \quad \Rightarrow c_< = -z_{1-\alpha} = z_\alpha$
3. $\alpha = \mathbb{P}_{\vartheta_0}[T \in K_>] = \mathbb{P}_{\vartheta_0}[T > c_>] = 1-\mathbb{P}_{\vartheta_0}[T \leq c_>] = 1-\Phi(c_>) \quad \Rightarrow c_> = \Phi^{-1}(1-\alpha) =: z_{1-\alpha}$

Wichtige Eigenschaften: $\Phi(c) := \mathbb{P}[Z \leq c]$ für $Z \sim \mathcal{N}(0,1)$, $\Phi(-c) = 1-\Phi(c)$, $\Phi(x)^{-1} =: z_x = -z_{1-x}$ und $t_{m,x} = -t_{m,1-x}$

z-Test (Gausstest)

Annahmen: $X_1,\dots,X_n$ u.i.v. mit $X_i \sim \mathcal{N}(\mu,\sigma^2)$, wobei σ^2 bekannt und μ unbekannt ist. $\Rightarrow \overline{X}_n \sim \mathcal{N}(\mu,\sigma^2/n)$
Hypothesen:

1. $H_0: \mu = \mu_0$ gegen $H_A: \mu \neq \mu_0$
2. $H_0: \mu \geq \mu_0$ gegen $H_A: \mu < \mu_0$
3. $H_0: \mu \leq \mu_0$ gegen $H_A: \mu > \mu_0$

Testgrösse:

$$T := \frac{\overline{X}-\mu_0}{\sigma}\cdot\sqrt{n} \sim \mathcal{N}(0,1)$$

Ablehnungskriterium: für H_0 bei Signifikanzniveau α:

1. $|T| > z_{1-\alpha/2}$ (für $\alpha = 5\%$ ist $z_{1-\alpha/2} = 1.96$)
2. $T < z_\alpha$
3. $T > z_{1-\alpha}$

t-Test

Annahmen: $X_1,\dots,X_n$ u.i.v. mit $X_i \sim \mathcal{N}(\mu,\sigma^2)$, σ^2 und μ unbekannt.
Hypothesen:

1. $H_0: \mu = \mu_0$ gegen $H_A: \mu \neq \mu_0$
2. $H_0: \mu \geq \mu_0$ gegen $H_A: \mu < \mu_0$
3. $H_0: \mu \leq \mu_0$ gegen $H_A: \mu > \mu_0$

Testgrösse:

$$T := \frac{\overline{X}-\mu_0}{S}\cdot\sqrt{n} \sim t_{n-1}$$

Ablehnungskriterium: für H_0 bei Signifikanzniveau α:

1. $|T| > t_{n-1,1-\alpha/2}$
2. $T < t_{n-1,\alpha}$
3. $T > t_{n-1,1-\alpha}$

Zwei-Strichproben-t-Test

Annahmen: $X_1,\dots,X_n$ u.i.v. mit $X_i \sim \mathcal{N}(\mu_X,\sigma^2)$ und $Y_1,\dots,Y_m$ u.i.v. mit $Y_i \sim \mathcal{N}(\mu_Y,\sigma^2)$.
Hypothesen:

1. $H_0: \mu_X = \mu_Y$ gegen $H_A: \mu_X \neq \mu_Y$
2. $H_0: \mu_X \geq \mu_Y$ gegen $H_A: \mu_X < \mu_Y$
3. $H_0: \mu_X \leq \mu_Y$ gegen $H_A: \mu_X > \mu_Y$

Testgrösse: Falls $m = n$ (*gepaarter Zweistichproben-Test*), dann sei $Z_i := X_i - Y_i$ und somit $Z_i \sim \mathcal{N}(\mu_X - \mu_Y, 2\sigma^2)$ u.i.v.. Dann führe normal z-Test oder t-Test durch (abhängig ob σ^2 bekannt oder unbekannt).
Falls $m \neq n$ (*ungepaarter Zweistichproben-Test*), so unterscheidet man weiter ob σ^2 bekannt oder unbekannt ist.
Falls σ^2 *bekannt*, so wähle:

$$T_1 := \frac{(\overline{X}_n - \overline{Y}_m) - (\mu_X - \mu_Y)}{\sigma\sqrt{\frac{1}{n}+\frac{1}{m}}} \sim \mathcal{N}(0,1)$$

Falls σ^2 *unbekannt*, so wähle:

$$T_2 := \frac{(\overline{X}_n - \overline{Y}_m) - (\mu_X - \mu_Y)}{S\sqrt{\frac{1}{n}+\frac{1}{m}}} \sim t_{n+m-2}$$

, mit $S^2 = \frac{1}{m+n-2}((n-1)S_X^2 + (m-1)S_Y^2)$

Ablehnungskriterium: für H_0 bei Signifikanzniveau α:

1. $|T_1| > z_{1-\alpha/2}$ bzw. $|T_2| > t_{m+n-2,1-\alpha/2}$
2. $T_1 < z_\alpha$ bzw. $T_2 < t_{m+n-2,\alpha}$
3. $T_1 > z_{1-\alpha}$ bzw. $T_2 > t_{m+n-2,1-\alpha}$

9.3 Vorgehen Tests

1. Wahl des Modells.
2. Formulierung von Hypothese und Alternative.
3. Bestimmung der Teststatistik T und der Form des kritischen Bereichs K; das kann aus einer Herleitung via LQ-Test stammen.
4. Festlegung des Niveaus α liefert (die Grenze für) den kritischen Bereich K; dazu braucht man die Verteilung von T unter $\mathbb{P}_\vartheta$ für alle $\vartheta \in \Theta_0$ (exakt oder approximativ).
5. Berechnen der Teststatistik $T(\omega)$ aus den Daten; ist $T(\omega) \in K$, so wird die Hypothese H_0 abgelehnt, andernfalls wird die Hypothese H_0 nicht verworfen.

10 Konfidenzintervalle

Das Konfidenzintervall gibt den Bereich an, der mit einer gewissen Wahrscheinlichkeit den Parameter ϑ einer Verteilung einer Zufallsvariablen einschließt.
Oftmals verwendet man nur eine Schätzvariavel T und konstruiert ein *symmetrisches Konfidenzintervall* $[T-\delta, T+\delta]$.

Definition Konfidenzintervall

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

$$\forall \vartheta \in \Theta \quad \mathbb{P}_\vartheta[A \leq \vartheta \leq B] \geq 1-\alpha$$

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

Beispiel Normalverteilung: (beidseitig) Sei $X_i \sim \mathcal{N}(\mu,\sigma^2)$ und somit ist $\overline{X}_n = \frac{1}{n}\sum_{i=1}^n X_i \sim \mathcal{N}(\mu, \frac{\sigma^2}{n})$, wobei $\vartheta = \mu$ der unbekannte Parameter und σ^2 bekannt. Dann gilt:

$$\mathbb{P}_\vartheta[-z_{1-\frac{\alpha}{2}} \leq \frac{\overline{X}_n - \mu}{\sigma/\sqrt{n}} \leq z_{1-\frac{\alpha}{2}}] = 1-\alpha$$

Wobei $z_{1-\frac{\alpha}{2}} = \Phi^{-1}(1-\alpha/2)$. Wir formen um und erhalten:

$$\mathbb{P}_\vartheta[\overline{X}_n - z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}} \leq \mu \leq \overline{X}_n + z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}] = 1-\alpha$$

Also erhalten wir für $\vartheta = \mu$ und Niveau $1-\alpha$ das Konfidenzintervall:

$$I = [\ \overline{X}_n - z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}\ ,\ \overline{X}_n + z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}\ ]$$

Beispiel Geometrische Verteilung: Sei $X_i \sim Geo(\vartheta)$, $T_{ML} = \frac{n}{\sum_{i=1}^n X_i}$ und $\alpha = 0.05$. $\Rightarrow E[X_i] = 1/\vartheta$ und $\mathrm{Var}[X_i] = (1-\vartheta)/\vartheta^2$, erhalten wir: $\frac{\sum_{i=1}^n X_i - n/\vartheta}{\sqrt{n(1-\vartheta)/\vartheta^2}} \sim \mathcal{N}(0,1)$ für $n \to \infty$ (folgt nach ZGS). Somit:

$$\mathbb{P}_\vartheta[-z_{1-\frac{\alpha}{2}} \leq \frac{\sum_{i=1}^n X_i - n/\vartheta}{\sqrt{n(1-\vartheta)/\vartheta^2}} \leq z_{1-\frac{\alpha}{2}}] \geq 1-\alpha$$

11 Diverses

Die **momenterzeugende Funktion** einer ZV X ist definiert als: $M_X(t) := \mathbb{E}[e^{tX}]$, für $t \in \mathbb{R}$.

Chernoff Schranken

Seien $X_1,\dots,X_n$ u.i.v. ZV, wobei $M_X(t)$ endlich ist für alle $t \in \mathbb{R}$. Sei $S_n := \sum_{i=1}^n X_i$. Dann gilt für jedes $b \in \mathbb{R}$:

$$\mathbb{P}[S_n \geq b] \leq \exp\ (\inf_{t\in\mathbb{R}} (n \log M_X(t) - tb))$$

Schranke exponentiell in b und in $n \Rightarrow$ sehr gute Abschätzung

Chernoff für Bernoulli-Vert. ZV

Sei nun $X_i \sim Ber(p)$ u.i.v. und somit $S_n \sim Bin(n,p)$. Dann gilt:

$$\mathbb{P}[S_n \geq (1+\delta)\mu_n] \leq \left(\frac{e^\delta}{(1+\delta)^{1+\delta}}\right)^{\mu_n}$$

Beispiel: Wir werfen 100-Mal eine Münze und wollen $\mathbb{P}[S_100 \geq 60]$ berechnen, mit $p_i = 1/2$ für alle i und damit $\mu_n = 50$ und $\delta = 0.2$. Wir erhalten als Abschätzung:

$$\mathbb{P}[S_{100} \geq 60] \leq \left(\frac{e^{0.2}}{1.2^{1.2}}\right)^5 0 = 0.3909$$

11.1 Anwendung: Simulationsalgorithmus

Satz

Sei F eine stetige und streng monoton wachsende Verteilungsfunktion, mit Umkehrfunktion F^{-1}. Ist $X \sim \mathcal{U}(0,1)$ und $Y = F^{-1}(X)$, so hat Y gerade die Verteilungsfunktion F.

Beweis: $F_Y(t) = \mathbb{P}[Y \leq t] = \mathbb{P}[F^{-1}(X) \leq t] = \mathbb{P}[X \leq F(t)] = F(t)$
Algorithmus: Man hat einen "Zufallszahlengenerator", d.h. einen deterministischen Algorithmus, der eine Folge $(x1, x2, ...)$ von Zahlen in $[0,1]$ produziert, die sich in einem gewissen Sinn verhält wie die Realisierung einer Folge von unabhängigen $\mathcal{U}(0,1)$-verteilten Zufallsvariablen. In diesem Sinne simuliert also nach obigem Satz $F^{-1}(Zufallsgenerator)$ die Verteilung F.
Beispiel: Um eine Exponentialverteilung mit Parameter λ zu simulieren, nehmen wir die zugehörige Verteilungsfunktion $F(t) = 1 - e^{-\lambda \cdot t}$ für $t \geq 0$. Die Inverse können wir berechnen und erhalten: $F^{-1}(F(t)) = t = -\frac{log(1-F(t)}{\lambda}$. Mit $X \sim \mathcal{U}(0,1)$ erhalten wir: $Y := F^{-1}(X) = -\frac{log(1-X)}{\lambda} \sim \mathrm{Exp}(\lambda)$

11.2 p-Wert

Der *p-Wert* ist Evidenzmaß für die Glaubwürdigkeit der Nullhypothese. Kleiner p-Wert $\Rightarrow H_0$ unwahrscheinlich.
Sei $X_1,\dots,X_n$ eine Stichprobe vom Umfang n. Wir wollen eine Hypothese $H_0 : \vartheta = \vartheta_0$ gegen eine Alternative $H_A : \vartheta \in \Theta_A$ testen. Sei $T = t(X_1,\dots,X_n)$ eine Teststatistik und $(T, K_t)_{t\geq 0}$ eine Familie von Tests.
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 $H_0 : \vartheta = \vartheta_0$ eine einfache Nullhypothese. Sei $(T, K_t)_{t\geq 0}$ eine geordnete Familie von Tests. Der p-Wert ist eine ZV, mit:

$$G : \mathbb{R}_+ \mapsto [0,1], \quad p-Wert := G(t) = \mathbb{P}_{\vartheta_0}[T \in K_t]$$

Der p-Wert hat folgende Eigenschaften:

1. Sei T stetig und $K_t = (t,\infty)$. Dann ist der p-Wert unter $\mathbb{P}_{\vartheta_0}$ auf $[0,1]$ gleichverteilt.
2. Für einen p-Wert γ gilt, dass alle Tests mit Signifikanzniveau $\alpha > \gamma$ die Nullhypothese verwerfen.

p-Wert ist klein $\implies H_0$ wird wahrscheinlich verworfen

Beispiel: Gegeben wir haben einen t-Test mit 8 Freiheitsgraden ausgeführt, wir wollen den p-Wert berechnen. Gegeben sei $T(w) = -3.4$ dann erhalten wir:
p-Wert$(\omega) = P_{H_0}[T < t_0]|_{t_0 = T(W)} = P_{H_0}[T < -3.4] = 1 - P_{H_0}[T \leq 3.4]$. Ablesen gibt uns die Schätzung $0.995 \leq P \leq 1$.

11.3 Monte-Carlo Verfahren

Ziel: Integrale approximieren durch die Erzeugung von Zufallszahlen.
Nehmen wir an, wir wollen für eine gegebene Funktion $h : [0,1] \to \mathbb{R}$ das Integral

$$I := \int_0^1 h(x)dx$$

berechnen. Allgemein auch: $I := \int_{[0,1]^d} h(\underline{x})d\underline{x}$.
Idee: Fasse I als EW auf; ist nämlich $U \sim \mathcal{U}([0,1])$, so ist:

$$\mathbb{E}[h(U)] = \int_{-\infty}^{\infty} h(x) f_U(x) dx = \int_0^1 h(x)dx = I$$

Haben wir eine Folge von ZV $U_1, U_2, \dots$ u.i.v. mit $U_i \sim \mathcal{U}([0,1])$ (mit Zufallsgenerator), so liefert das Gesetz der grossen Zahlen:

$$\overline{h(U_n)} = \frac{1}{n}\sum_{i=1}^n h(U_i) \xrightarrow[n\to\infty]{} \mathbb{E}[h(U_1)] = I \quad \text{(f.s)}$$

$\Rightarrow$ Wir erhalten also eine Approximation von I. Der zu erwartende Fehler sinkt mit $\approx \frac{1}{\sqrt{n}}$

11.4 Weitere Verteilungen: Chi- & t-Vert.

Seien $X_i \sim \mathcal{N}(0,1)$, dann ist die Summe $Y = \sum_{i=1}^n X_i^2 \sim \chi_m^2$ (mit n Freiheitsgraden). Falls $n = 2$, haben wir eine Exponentialverteilung mit Parameter $\frac{1}{2}$.

$\mathcal{X}^2$-Verteilung (Chiquadrat-Verteilung)

$X_1, \ldots, X_n$ u.i.v. $\sim \mathcal{N}(0,1)$, so ist $Y := \sum_{i=1}^n X_i^2 \sim \mathcal{X}_n^2$.

$$f_Y(t) = \frac{1}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})} t^{\frac{n}{2}-1} e^{-\frac{t}{2}} \quad t > 0$$

$$\mathbb{E}[\chi_n^2] = n \quad \text{und} \quad \mathrm{Var}[\chi_n^2] = 2n$$

Wobei die Gamma-Funktion für $v \geq 0$ definiert ist:

$$\Gamma(v) := \int_0^\infty t^{v-1} \cdot e^{-t} dt$$

Es gilt: $\Gamma(n) = (n-1)!$ für $n \in \mathbb{N}$ und $\forall r \in \mathbb{R} : r \cdot \Gamma(r) = \Gamma(r+1)$.

t-Verteilung

$X \sim \mathcal{N}(0,1)$ und $Y \sim \chi_n^2$ unabh.: $Z := \frac{X}{\sqrt{\frac{1}{m}Y}}$ t_m-verteilt.

$$f_Z(t) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi} \cdot \Gamma(\frac{n}{2})}\left(1 + \frac{t^2}{n}\right)^{-\frac{n+1}{2}}$$

$$\mathbb{E}[Z] = 0 \text{ für } n > 1 \quad \text{und} \quad \mathrm{Var}[Z] = \frac{n}{n-2} \text{ für } n > 2$$

Es gilt: $t_{n,x} \xrightarrow[n \to \infty]{} z_x$, mit $z_x := \Phi(x)^{-1}$

11.5 MLE Schätzer

- Bernoulli: $\hat{\lambda} = \hat{\mu} = \frac{\sum_{i=1}^n x_i}{n} = \bar{X}$
- Exponential: $\hat{\mu} = \frac{\sum_{i=1}^n x_i}{n} = \bar{X}$ und $\hat{\lambda} = \frac{n}{\sum_{i=1}^n x_i}$
- Geometrisch: $\hat{p} = \frac{n}{\sum_{i=1}^n x_i} = \frac{1}{\bar{X}}$
- Binomial: $\hat{\mu} = \frac{1}{N}\frac{\sum_{i=1}^n x_i}{n}$ und $\hat{p} = \frac{\sum_{i=1}^n x_i}{n} = \overline{X}$
- Normalverteilung: $\hat{\mu} = \frac{\sum_{i=1}^n x_i}{n} = \bar{X}$ und $\hat{\sigma}^2 = S^2$
- Poisson: $\hat{\lambda} = \hat{\mu} = \frac{\sum_{i=1}^n x_i}{n} = \bar{X}$
- Uniform: $\hat{b} = max(x_i), \quad \hat{a} = min(x_i)$

12 Math Stuff

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} \qquad (x+y)^n = \sum_{k=0}^n \binom{n}{k} x^{n-k} y^k$$

$$ax^2 + bx + c = 0 \quad \Rightarrow x_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

(1) $\cos(z) = \cos(-z)$ und $\sin(-z) = -\sin(z)$

(2) $\cos(\pi - x) = -\cos(x)$ und $\sin(\pi - x) = \sin(x)$

(3) $\sin(z+w) = \sin(z)\cos(w) + \cos(z)\sin(w)$

(4) $\cos(z+w) = \cos(z)\cos(w) - \sin(z)\sin(w)$

(5) $\cos(z)^2 + \sin(z)^2 = 1$

(6) $\sin(2z) = 2\sin(z)\cos(z)$

(7) $\cos(2z) = \cos(z)^2 - \sin(z)^2$

$$\sin^2(x) = \frac{1}{2}(1 - \cos(2x)) \;\; \text{und} \;\; \cos^2(x) = \frac{1}{2}(1 + \cos(2x))$$

α	0	30°	45°	60°	90°	120°	150°	180°	270°
	0	$\frac{\pi}{6}$	$\frac{\pi}{4}$	$\frac{\pi}{3}$	$\frac{\pi}{2}$	$\frac{2\pi}{3}$	$\frac{5\pi}{6}$	π	$\frac{3\pi}{2}$
$\sin\alpha$	0	$\frac{1}{2}$	$\frac{\sqrt{2}}{2}$	$\frac{\sqrt{3}}{2}$	1	$\frac{\sqrt{3}}{2}$	$\frac{1}{2}$	0	-1
$\cos\alpha$	1	$\frac{\sqrt{3}}{2}$	$\frac{\sqrt{2}}{2}$	$\frac{1}{2}$	0	$-\frac{1}{2}$	$-\frac{\sqrt{3}}{2}$	-1	0
$\tan\alpha$	0	$\frac{\sqrt{3}}{3}$	1	$\sqrt{3}$	N/A	$-\sqrt{3}$	$-\frac{\sqrt{3}}{3}$	0	N/A

12.1 Typische-Reihen

$\sum_{i=1}^n i = \frac{n(n+1)}{2}$	$\sum_{i=1}^n i^2 = \frac{n(n+1)(2n+1)}{6}$
$\sum_{i=1}^n i^3 = \frac{n^2(n+1)^2}{4}$	$\sum_{i=1}^\infty \frac{1}{n^2} = \frac{\pi^2}{6}$

Die Geometrische Reihe: $\sum_{i=0}^n q^i = \frac{1-q^{n+1}}{1-q}$ konvergiert wenn $|q| < 1$. Dies gilt auch bei $n \to \infty$:

$$\sum_{i=1}^\infty q^i = 1 + q + q^2 + \ldots = \frac{1}{1-q}$$

Die Harmonische Reihe: Die Harmonische Reihe $\sum_{n=1}^\infty \frac{1}{n}$ ist divergent. Die alternierende harmonische Reihe ist jedoch konvergent.

Die Zeta Funktion: Die Riemann-Zeta Funktion $\zeta(s) = \sum_{n=1}^\infty \frac{1}{n^s}$ konvergiert für $s > 1$.

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

Partielle Integration

$$\int_a^b f'(x) \cdot g(x)dx = [f(x) \cdot g(x)]_a^b - \int_a^b f(x) \cdot g'(x)dx$$

Substitution

$$\int_a^b f(g(x)) \cdot g'(x)dx = \int_{g(a)}^{g(b)} f(u)du$$

$$\int_{\varphi(U)} f(\vec{v})d\vec{v} = \int_U f(\varphi(\vec{u})) \cdot \mid \det(D\varphi)(\vec{u}) \mid d\vec{u}$$

13 Aufgaben

(HS17, Aufg. 3) Seien X und X ZV mit:

$$f(x,y) = \begin{cases} \frac{1}{x^2y^2} & \text{für } x \geq 1, y \geq 1 \\ 0 & \text{sonst} \end{cases}$$

Sei $U := \frac{X}{Y}$ und $V := X \cdot Y$.
a) Berechne F_U. b) Berechne F_v.
c) Bestimme f_U. d) Bestimme f_V.

a) Für $u \leq 0$ haben wir $F_U(u) = 0$. Für $u > 1$ erhalten wir: $F_U(u) = \mathbb{P}[\frac{X}{Y} \leq u] = \mathbb{P}[X \leq Y \cdot u] = \int_1^\infty \int_1^{y \cdot u} \frac{1}{x^2y^2} dxdy = \int_1^\infty \frac{1}{y^2}(1 - \frac{1}{uy})dy = \int_1^\infty \frac{1}{y^2} - \frac{1}{uy^3} dy = 1 - \frac{1}{2u}$.
Für $0 < u \leq 1$ erhalten wir: $F_U(u) = \int_{1/u}^\infty \int_1^{yu} \frac{1}{x^2y^2} dxdy = \int_{1/u}^\infty \frac{1}{y^2}(1 - \frac{1}{uy})dy = \int_{1/u}^\infty \frac{1}{y^2} - \frac{1}{uy^3} dy = u - \frac{u}{2} = \frac{u}{2}$.

b) Für $v < 1$ ist $F_V(v) = 0$. Für $v \geq 1$: $F_V(v) = \mathbb{P}[X \cdot Y \leq v] = \mathbb{P}[X \leq v/Y] = \int_1^v \int_1^{v/y} \frac{1}{x^2y^2} dxdy = \int_1^v \frac{1}{y^2} \int_1^{v/y} \frac{1}{x^2} dxdy = \int_1^v \frac{1}{y^2}(1 - \frac{y}{v})dy = \int_1^v \frac{1}{y^2} - \frac{1}{yv} dy = 1 - \frac{1}{v}(1 + log(v))$

c) Für $u > 1$ erhalten wir: $f_U(u) = \frac{dF_U(u)}{du} = \frac{1}{2u^2}$, für $0 < u \leq 1$: $f_U(u) = \frac{1}{2}$ und für $u \leq 0$ haben wir: $f_U(u) = 0$.

d) Für $v \geq 1$ erhalten wir: $f_V(v) = \frac{dF_V(v)}{dv} = \frac{log(v)}{v^2}$ und $f_V(v) = 0$ für $v < 1$.

(HS16, Aufg. 3) Seien X und Y zwei unabhängige ZV, beide exponentialverteilt mit Parameter $\lambda > 0$. Sei $U := \frac{X}{X+Y}$ und $V := X + Y$.
a) Berechne $f_U(u)$ & $F_U(u)$. b) $f_V(v)$ & $F_V(v)$.
c) Bestimme $f_{U,V}$. Sind U und V unabhängig?

a) U hat nur Werte in $(0,1)$, also $u \in (0,1)$. $\mathbb{P}[U \leq u] = \lambda^2 \int_0^\infty e^{-\lambda x}(\int_0^\infty 1_{\{\frac{x}{x+y} \leq u\}} e^{-\lambda x} dy)dx$
$= \lambda^2 \int_0^\infty e^{-\lambda x}(\int_0^\infty 1_{\{\frac{x}{u} - x \leq y\}} e^{-\lambda x} dy)dx$
$= \lambda \int_0^\infty e^{-\lambda x}(\int_{\frac{x}{u}-x}^\infty \lambda e^{-\lambda x} dy)dx$
$= \lambda \int_0^\infty e^{-\lambda x} e^{-\lambda x(\frac{1}{u}-1)} dx = \lambda \int_0^\infty e^{-\frac{\lambda x}{u}} dx = u.$
Somit ist $U \sim \mathcal{U}(0,1)$ und darum: $f_U(u) = 1_{\{0 \leq u \leq 1\}}$.

b) $\mathbb{P}[V \leq v] = \lambda^2 \int_0^\infty e^{-\lambda x}(\int_0^\infty 1_{\{x+y \leq v\}} e^{-\lambda y} dy)dx$
$= \lambda^2 \int_0^\infty e^{-\lambda x} \cdot 1_{\{x \leq v\}}(\int_0^\infty 1_{\{y \leq v-x\}} e^{-\lambda y} dy)dx$
$= \lambda \int_0^v e^{-\lambda x}(\int_0^{v-x} \lambda e^{-\lambda y} dy)dx = \lambda \int_0^v e^{-\lambda x}(1 - e^{-\lambda(v-x)})dx = 1 - e^{-\lambda v} - \lambda v e^{-\lambda v}$ für $v > 0$ und $F_V(v) = 0$ für $v \leq 0$.
Und die Dichte: $f_V(v) = \frac{dF_V(v)}{dv} = \lambda^2 v e^{-\lambda v} \cdot 1_{\{v > 0\}}$.

c) $\mathbb{P}[U \leq u, V \leq v] = \lambda^2 \int_0^\infty e^{-\lambda x}(\int_0^\infty 1_{\{\frac{x}{x+y} \leq u, x+y \leq v\}} e^{-\lambda y} dy)dx$
$= \lambda^2 \int_0^\infty e^{-\lambda x}(\int_0^\infty 1_{\{x(\frac{1}{u}-1) \leq y \leq v-x, x \leq uv\}} e^{-\lambda y} dy)dx$
$= \lambda \int_0^{uv} e^{-\lambda x}(\int_{x(\frac{1}{u}-1)}^{v-x} \lambda e^{-\lambda y} dy)dx$
$= \lambda \int_0^{uv} e^{-\lambda x}(e^{-\lambda x(\frac{1}{u}-1)} - e^{-\lambda(v-x)})dx$
$= \lambda \int_0^{uv}(e^{-\lambda \frac{x}{u}} - e^{-\lambda v})dx = u(1 - e^{-\lambda v} - \lambda v e^{-\lambda v})$.
Somit: $F_{U,V}(u,v) = F_U(u) \cdot F_V(v) \Rightarrow U$ und V sind unabhängig. Und für die Dichte gilt somit:
$f_{U,V} = f_U(u) \cdot f_V(v) = \lambda^2 v e^{-\lambda v} \cdot 1_{\{0 \leq u \leq 1, v > 0\}}$

(HS20) Seien U_1, U_2, U_3 u.i.v mit $U_i \sim \mathcal{U}([0,1])$. Sei $L := min(U_1, U_2, U_3)$ und $M := max(U_1, U_2, U_3)$.
a) Berechne $f_M(m)$. b) Bestimme $f_{M,L}(m,l)$.
c) Berechne $f_{L|M}(l,m)$.

a) $F_M(m) = \mathbb{P}[U_1 \leq m, U_2 \leq m, U_3 \leq m]$

$$\overset{unab.}{=} \Pi_{i=1}^3 \mathbb{P}[U_i \leq m] \overset{i.d.}{=} \begin{cases} 1 & m \geq 1 \\ m^3 & 0 \leq m \leq 1 \\ 0 & \text{sonst} \end{cases}$$

$\Rightarrow f_M(m) = 3m^2 \cdot 1_{\{m \in [0,1]\}}$

b) $\mathbb{P}[M \leq m, L \leq l]$
$= \mathbb{P}[M \leq m] - \mathbb{P}[M < m, L > l]$
$= m^3 - \mathbb{P}[l < U_1 < m, l < U_2 < m, l < U_3 < m]$
$= m^3 - (\mathbb{P}[l < U_1 < m])^3$
$= m^3 - (m-l)^3$, für $0 \leq l \leq m \leq 1$.
$\Rightarrow f_{M,L}(m,l) = 6(m-l) \cdot 1_{\{0 \leq l \leq m \leq 1\}}$

c) $f_{L|M}(l,m) = \frac{f_{M,L}(m,l)}{f_M(m)} = \frac{6(m-l)}{3m^2} \cdot 1_{\{0 \leq l \leq m \leq 1\}}$
$= 2\frac{m-l}{m^2} \cdot 1_{\{0 \leq l \leq m \leq 1\}}$

(HS20) Würfel mit sechs Seiten, vermutlich gezinkt (landet eher auf der 6). Experiment: Zehn Würfe (unabhängig). Sei $X_i = 1$, wenn der i-te Wurf eine sechs ist, und sonst gleich null. Wir wissen: $\sum_{i=1}^{1} 0X_i = 4$ (aus Tabelle).
a) Führe einen Test mit $\alpha = 1\%$ und prüfe ob gezinkt
b) Finde den P-Wert.

a) **i. Modell:** $X_i \overset{i.i.d}{\sim} Bernoulli(p), p$ unbekannt.
ii. Nullhypothese: $H_0 : p = p_0 = 1/6$
iii. Alternativhypothese: $H_A : p > 1/6$
iv. Teststatistik: $T := \sum_{i=1}^{10} X_i$
v. Vert. Testst. unter H_0: $T \sim Bin(10, 1/6)$
vi. Verwerfungsbereich: Wir möchten alle $k \in \{0, ..., 9\}$, sodass: $\mathbb{P}_{p_0}[T > k] \leq 0,01 \iff \mathbb{P}_{p_0}[T \leq k] > 0.99$.
$\Rightarrow K = \{6, 7, 8, 9\}$ (Werte aus Tabelle)
vii. Beob. Werte der Teststat.: $T(x_1, ..., x_{10}) = 4$
viii. Testentscheid: $T(\omega) = 4 \notin \{6, 7, 8, 9\} \Rightarrow$ Wir werfen H_0 nicht.

b) $\mathbb{P}_{p_0}[T \geq 4] = 1 - \mathbb{P}_{p_0}[T < 4] = 1 - 0.93 = 0.07$

((FS21, Aufg. 4, t-Test) Gegeben sind 9 unabhängige Temperaturmessung die $\mathcal{N}(\mu, \sigma^2)$-verteilte ZV sind. Wir möchten auf einem 5%-Niveau testen, ob die erwartetet tägliche Höchsttemperatur im Vergleich zum Wert zum Vorjahr (22 Grad) gesunken ist. Gegeben: $\overline{X}_9 = 18,6$ und $S_9 = 3$.
a) Führe einen Test durch b) bestimme P-Wert

a) **i. Modell:** $X_1, \ldots, X_9$ i.i.d. $\sim \mathcal{N}(\mu, \sigma^2)$ unter $\mathbb{P}_\vartheta$, wobei $\vartheta = (\mu, \sigma^2)$ unbekannt.
ii. Hypothesen: $H_0 : \mu = \mu_0 = 22$ und $H_A : \mu < \mu_0$
iii. Teststatistik & Verteilung unter H_0**:** Da μ und σ^2 unbekannt, führen wir einen t-Test durch:

$$T = \frac{\overline{X}_9 - \mu_0}{S_9/\sqrt{9}} \sim t_8$$

iv. Verwerfungsbereich: Nach der Alternative hat der kritische Bereich die Form $K_< = (-\infty, c_<)$ für ein zu bestimmendes $c_<$. Für $\alpha = 0.05$ wählen wir $c_<$ so, dass: $\alpha = \mathbb{P}[T < c_<] \quad \Rightarrow c_< = t_{n-1,\alpha} = -t_{n-1,1-\alpha} = -t_{8,0.95} = -1.86 \quad \Rightarrow K_< = (-\infty, -1.86)$.
v. Wert der Teststatistik: Durch einsetzen:

$$T(\omega) = t(x_1, ..., x_9) = \frac{\overline{X}_9 - \mu_0}{S_9/\sqrt{9}} = \frac{18.6 - 22}{3/3} = -3.4$$

vi. Testentscheid: Wegen $T \in (-\infty, -1.86) = K_<$ verwerfen wir die Hypothese und nehmen die Alternative an. $\Rightarrow$ Daten sprechen dafür, dass die Temperatur tatsächlich gesunken ist.

b) p-Wert$(\omega) = \mathbb{P}_{H_0}[T < -3.4] = 1 - \mathbb{P}_{H_0}[T < 3.4]$. Wegen $t_{8,0.995} = 0.355 < 3.4$, folgt: $0 \leq$ p-Wert$(\omega) \leq 0.005$.

(FS21, Aufg. 2) Seien X und Y zwei unabh. ZV mit $X \sim Poi(\lambda)$ und $Y \sim Poi(\mu)$, wobei $\lambda, \mu > 0$.
a) Zeige, dass $X + Y \sim Poi(\lambda + \mu)$
b) Zeige, dass $X|X + Y = n \sim Bin(n, \frac{\lambda}{\lambda+\mu})$
c) Bestimme $\mathbb{P}[X + Y = n|Y = l]$ für alle n, l

a) $p_{X+Y}(k) = p_X(k) * p_Y(k) \overset{unab.}{=} \sum_{x=0}^{k} p_X(x) \cdot p_Y(k-x) = \sum_{x=0}^{k} \frac{\lambda^x}{x!} e^{-\lambda} \cdot \frac{\mu^{k-x}}{(k-x)!} e^{-\mu} = \sum_{x=0}^{k} \frac{\lambda^x \mu^{k-x}}{x!(k-x)!} e^{-(\lambda+\mu)} = e^{-(\lambda+\mu)} \frac{1}{k!} \sum_{x=0}^{k} \frac{k!}{x!(k-x)!} \lambda^x \mu^{k-x} \overset{Bin.Satz}{=} e^{-(\lambda+\mu)} \frac{(\lambda+\mu)^k}{k!}$

b) $\mathbb{P}[X = k|X+Y = n] = \frac{\mathbb{P}[X=k, X+Y=n]}{\mathbb{P}[X+Y=n]} = \frac{\mathbb{P}[X=k, Y=n-k]}{\mathbb{P}[X+Y=n]} = \frac{\mathbb{P}[X=k] \cdot [X+Y=n]}{\mathbb{P}[X+Y=n]} = \frac{\frac{\lambda^k}{k!} e^{-\lambda} \cdot \frac{\mu^{n-k}}{(n-k)!} e^{-\mu}}{\frac{(\lambda+\mu)^n}{n!} e^{-(\lambda+\mu)}} = \frac{n!}{k!(n-k)!} \cdot \frac{\lambda^k \mu^{n-k}}{(\lambda+\mu)^n} = \binom{n}{k} \frac{\lambda^k \mu^{n-k}}{(\lambda+\mu)^k (\lambda+\mu)^n - k} = \binom{n}{k} \left(\frac{\lambda}{\lambda+\mu}\right)^k \left(\frac{\mu}{\lambda+\mu}\right)^{n-k}$

c) Für $n \notin \mathbb{N}$ oder $l \notin \mathbb{N}$ oder $l, n \in \mathbb{N}, l > n$ ist $P[X+Y = n|Y = l] = 0$ (trivial).
Für den Fall $l, n \in \mathbb{N}, l \leq n$: $P[X + Y = n|Y = l] = \frac{\mathbb{P}[X+Y=n, Y=l]}{\mathbb{P}[Y=l]} = \frac{\mathbb{P}[X=n-l, Y=l]}{\mathbb{P}[Y=l]} = \frac{\mathbb{P}[X=n-l] \cdot \mathbb{P}[Y=l]}{\mathbb{P}[Y=l]} = \mathbb{P}[X = n - l] = \frac{\lambda^{n-l}}{(n-l)!} e^{-\lambda}$

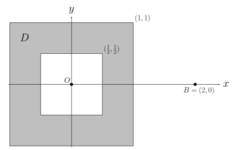
(FS20, Aufg. 2) Zeige, dass für eine diskrete ZV Z mit Werten in $\mathbb{N}$ gilt: $\exists q \in (0, 1) : Z \sim Geo(q) \iff \mathbb{P}[Z > n] = \mathbb{P}[Z > n + k|Z > k] \quad , n, k \geq 1$

- **($\Rightarrow$):** Sei $Z \sim Geo(q), q \in [0, 1]$. Dann gilt: $\mathbb{P}[Z > k] = (1 - q)^k$ und somit für alle $k, n \in \mathbb{N}$, dass: $\mathbb{P}[Z > n + k \mid Z > k] = \frac{\mathbb{P}[Z>n+k, Z>k]}{\mathbb{P}[Z>k]} = \frac{\mathbb{P}[Z>n+k]}{\mathbb{P}[Z>k]} = \frac{(1-q)^{n+k}}{(1-q)^k} = (1 - q)^n = \mathbb{P}[Z > n]$
- **($\Leftarrow$):** Sei nun $\mathbb{P}[Z > n + k \mid Z > k] = \mathbb{P}[Z > n]$ für $n, k \in \mathbb{N}$. Dann gilt: $\mathbb{P}[Z > n] = \mathbb{P}[Z > n + k \mid Z > k] = \frac{\mathbb{P}[z>n+k]}{Z>k}$. Nun sei $f(n) := \mathbb{P}[Z > n]$. Es gilt also: $\forall n, k \in \mathbb{N} : f(n) \cdot f(k) = f(n+k)$. Wegen $f(n+1) = f(n)f(1)$ folgt sofort durch Iteration, dass $f(n) = a^n$ mit $a = f(1)$ und damit: $\mathbb{P}[Z = n] = \mathbb{P}[Z > n-1] - \mathbb{P}[Z > n] = f(n-1) - f(n) = (1 - a)a^{n-1}$. Schliesslich ist $a = f(1) = \mathbb{P}[Z > 1] \in [0, 1]$, also auch $q = 1 - a \in [0, 1]$ und damit $Z \sim Geo(q)$.

(FS20, Aufg. 3) Man wählt zufällig uniform verteilt einen Punkt $A = (X, Y)$ in dem Gebiet $D = \{(x, y) \in \mathbb{R} : 1/2 \leq max(\mid x \mid, \mid y \mid) \leq 1\}$. Somit ist:

$$f_{X,Y}(x, y) = \begin{cases} c & \text{falls } 1/2 \leq\mid x \mid\leq 1 \text{ und } \mid y \mid\leq 1 \\ c & \text{falls } 1/2 \leq\mid y \mid\leq 1 \text{ und } \mid x \mid\leq 1/2 \\ 0 & \text{sonst} \end{cases}$$

Sei $V = (2max(|X|, |Y|))^2$ die Fläche des achsenparallelen Quadrates, welches seinen Mittelpunkt im Ursprung $O = (0, 0)$ hat und bei welchem der Punkt A auf einer der Seitenkanten liegt. Sei weiter σ der Abstand vom Punkt A zum Punkt $B = (2, 0)$. (a) Bestimme $c, f_X(x)$ und $f_y(y)$ (b) Finde $\mathbb{E}[X^2]$ & $\mathbb{E}[\sigma^2]$ (c) Berechne $f_V(v)$ & $\mathbb{E}[V]$



a) Zuerst berechnen wir c: $\int_{-\infty}^{\infty} \int_{\infty}^{\infty} f_{X,Y}(x, y)\, dx\, dy$
$= \int_{-1}^{1} \int_{-1}^{1} c\, dx\, dy - \int_{-\frac{1}{2}}^{\frac{1}{2}} \int_{-\frac{1}{2}}^{\frac{1}{2}} c\, dx\, dy = 4c - c$
$= 3c \overset{!}{=} 1 \quad \Rightarrow c = \frac{1}{3}$ Für die Dichte erhalten wir:

$$f_X(x) = \begin{cases} \int_{-1}^{-\frac{1}{2}} c\, dy + \int_{\frac{1}{2}}^{1} c\, dy = c = \frac{1}{3} & \text{falls } |x| \leq \frac{1}{2} \\ \int_{-1}^{1} c\, dy = 2c = \frac{2}{3} & \text{falls } \frac{1}{2} \leq |x| \leq 1 \\ 0 & \text{sonst} \end{cases}$$

Zudem gilt: $f_X(k) = f_Y(k)$ aus Symmetrie gründen.

b) $\mathbb{E}\left[X^2\right] = \int_{-1}^{1} x^2 \cdot f_X(x)\, dx = 2 \cdot \left(\int_0^{\frac{1}{2}} x^2 \cdot \frac{1}{3}\, dx + \int_{\frac{1}{2}}^{1} x^2 \cdot \frac{2}{3}\, dx\right)$
$= 2 \cdot \left(\left|\frac{1}{9} \cdot x^3\right|_0^{\frac{1}{2}} + \left|\frac{2}{9} \cdot x^3\right|_{\frac{1}{2}}^{1}\right) = 2 \cdot \left(\frac{1}{72} - \frac{2}{72} + \frac{2}{9}\right) = \frac{5}{12}$
Zudem gilt: $\mathbb{E}[X^2] = \mathbb{E}[Y^2]$ aus Symmetrie gründen.

$\mathbb{E}\left[\varrho^2\right] = \mathbb{E}\left[(X-2)^2 + Y^2\right] = \mathbb{E}\left[X^2 - 4X + 4 + Y^2\right]$
$= \mathbb{E}\left[X^2\right] - 4 \cdot E\left[X\right] + 4 + E\left[Y^2\right] = \frac{5}{12} - 4 \cdot 0 + 4 + \frac{5}{12} = \frac{29}{6}$

c) Wir sehen, dass $1 \leq V \leq 4$. Sei $S := 2max(|X|, |Y|)$ (Seitenlänge von V). Nun gilt für $v \in [1, 4]$:
$\mathbb{P}[V \leq v] = \mathbb{P}[S \leq \sqrt{v}] = \mathbb{P}[max(|X|, |Y|) \leq \sqrt{v}/2] = \mathbb{P}[|X| \leq \sqrt{v}/2, |Y| \leq \sqrt{v}/2]$
$= \frac{1}{3} \int_{|y|\leq 1} \int_{1/2\leq|x|\leq 1} 1_{\{|x|\leq\sqrt{v}/2,\ |y|\leq\sqrt{v}/2\}} dxdy$
$+ \frac{1}{3} \int_{|x|\leq 1/2} \int_{1/2\leq|y|\leq 1} 1_{\{|x|\leq\sqrt{v}/2,\ |y|\leq\sqrt{v}/2\}} dydx$
$= \frac{1}{3}\sqrt{v}(\sqrt{v} - 1) + \frac{1}{3}(\sqrt{v} - 1) = \frac{1}{3}(v - 1)$.
Somit ist: $f_V(v) = 1/3$ für $v \in [0, 4]$ und 0 sonst.
Somit erhalten wir: $\mathbb{E}[V] = \frac{1}{3} \int_1^4 v\, dv = \frac{5}{2}$.

(HS19, A3) Sei $P \sim \mathcal{U}(0, 1)$ und $H \mid P \sim \mathcal{U}(0, p)$.
(a) Bestimme $f_{H,P}$ (b) Bestimme f_H
(c) Bestimme $\mathbb{E}[H]$, $\mathbb{E}[P]$. (d) Bestimme $cov(H, P)$.

a) Wir wissen: $f_{H|P}(h \mid p) = \frac{1}{p} \cdot 1_{\{h\in(0,p)\}}$
$\Rightarrow f_{H,P}(h, p) = f_{H|P}(h \mid p) \cdot f_P(p) = \frac{1}{p} \cdot 1_{\{h\in(0,p)\wedge p\in[0,1]\}}$

b) $f_H(h) = \int_{-\infty}^{\infty} f_{H,P}(h, p)dp = \int_0^1 \frac{1}{p} \cdot 1_{\{h\in(0,p)\}}\, dp$
$= \int_h^1 \frac{1}{p}\, dp = -ln(h)$, für $0 \leq h \leq 1$.

c) $\mathbb{E}[H] = -\int_0^1 hlogh\, dh = \left[-\frac{h^2}{2}logh\right]_0^1 + \int_0^1 \frac{h^2}{2}\frac{1}{h}\, dh$
$= 1/4$. Zudem ist $\mathbb{E}[P] = \frac{1}{2}$ und somit $\mathbb{E}[H] = \frac{\mathbb{E}[P]}{2}$.

d) $\mathbb{E}[P \cdot H] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p \cdot h \cdot f_{P,H}(p, h)\, dp\, dh = \int_0^1 \int_0^p h\, dhdp$
$= \int_0^1 \frac{p^2}{2}\, dp = \frac{1}{6}. \quad \Rightarrow Cov(P, H) = \mathbb{E}[PH] - \mathbb{E}[P]\mathbb{E}[H]$
$= \frac{1}{6} - \frac{1}{4} \cdot \frac{1}{2} = \frac{1}{24}$. Somit sind P und H abhängig!

(FS19, A2) Zwei Halbleiter parallel geschaltet, Kontrolllicht leuchtet auf, falls einer der beiden ausfällt. Sei $X_1, X_2 \sim Exp(1/60)$ die Lebensdauer der beiden Halbleiter. Sei Z die Zeit bis die Kontrollleuchte aufleuchtet.
(a) Finde Verteilung von Z
(b) Finde W'keit, dass Bauteil ≥ 35 Mal in 3 Jahren ersetzt wird (wird ersetzt falls Kontroll. leuchtet)

a) Sei $T :=$ Zeit bis Kontrolllicht leuchtet.
$\Rightarrow F_T(t) = \mathbb{P}[T \leq t] = \mathbb{P}[min(X_1, X_2) \leq t]$
$= 1 - \mathbb{P}[min(X_1, X_2) > t] = 1 - \mathbb{P}[X_1 > t] \cdot \mathbb{P}[X_2 > t]$
$= 1 - e^{-2/60} \quad \Rightarrow T \sim Exp(1/30)$

b) Nach einer Zeit T_i (in Tagen) wir das erste Bauteil ersetzt. Es gilt $T_i i.i.d. \sim Exp(1/30)$. Sei $S := \sum_{i=1}^{36} T_i$. Wir wollen $\mathbb{P}[S < 3 \cdot 365]$ finden. Es gilt: $\mathbb{E}[S] = 36 \cdot \mathbb{E}[T_i] = 36 \cdot 30 = 1080$ und $\text{Var}[S] = 36 \cdot \text{Var}[T_i] = 36 \cdot 30^2 = 32400$. Nach dem ZGS folgt: $\mathbb{P}[S < 3 \cdot 365] = \mathbb{P}\left[\frac{S-\mathbb{E}[S]}{\sqrt{\text{Var}[S]}} \leq \frac{3\cdot 365-\mathbb{E}[S]}{\sqrt{\text{Var}[S]}}\right] = \mathbb{P}\left[\frac{S-1080}{180} \leq \frac{15}{180}\right] = \Phi(1/12) = \Phi(0,08) = 53.19\%$