

Wahrscheinlichkeit und Statistik

by dcamenisch

Grundbegriffe

Wir definieren einen **Wahrscheinlichkeitsraum** als das Tupel $(\Omega, \mathcal{F}, \mathbb{P})$:

Der **Grundraum** Ω ist eine nicht leere Menge, wobei $\omega \in \Omega$ ein Elementarereignis ist.

Eine **Sigma-Algebra** $\mathcal{F} \subseteq \mathbb{P}(\Omega)$ erfüllt die Bedingungen:

- $\Omega \in \mathcal{F}$
- $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$
- $A_1, A_2, \dots \in \mathcal{F} \Rightarrow \bigcup_{i=1}^\infty A_i \in \mathcal{F}$

Ein **Wahrscheinlichkeitsmass** $\mathbb{P}$ auf $(\Omega, \mathcal{F})$ ist eine Abbildung $\mathbb{P} : \mathcal{F} \mapsto [0, 1], A \mapsto \mathbb{P}[A]$, so dass:

- $\mathbb{P}[\Omega] = 1$
- $\mathbb{P}[A] = \sum_{i=1}^\infty \mathbb{P}[A_i]$, falls $A = \bigsqcup_{i=1}^\infty A_i$

Aus diesen Definitionen ergeben sich folgende nützliche Eigenschaften:

- $\emptyset \in \mathcal{F}$
- $A_1, A_2, \dots \in \mathcal{F} \Rightarrow \bigcap_{i=1}^\infty A_i \in \mathcal{F}$
- $A, B \in \mathcal{F} \Rightarrow A \cup B \in \mathcal{F}$
- $A, B \in \mathcal{F} \Rightarrow A \cap B \in \mathcal{F}$

und

- $\mathbb{P}[\emptyset] = 0$
- $\mathbb{P}[A^c] = 1 - \mathbb{P}[A]$
- $\mathbb{P}[A \cup B] = \mathbb{P}[A] + \mathbb{P}[B] - \mathbb{P}[A \cap B]$

Daraus ergibt sich, dass wenn $A_1, \dots, A_n$ paarweise disjunkt sind:

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

Satz. Monotonie: Seien $A, B \in \mathcal{F}$ dann gilt

$$A \subseteq B \Rightarrow \mathbb{P}[A] \leq \mathbb{P}[B]$$

Satz. Union Bound: Seien $A_1, A_2, \dots \in \mathcal{F}$ dann gilt

$$\mathbb{P}[\bigcup_{i=1}^\infty A_i] \leq \sum_{i=1}^\infty \mathbb{P}[A_i]$$

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:

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

Satz. Totale Wahrscheinlichkeit: Sei $A_1, \dots, A_n \in \mathcal{F}$ eine Partition von Ω mit $\mathbb{P}[A_i] > 0$ für alle $1 \leq i \leq n$. Dann gilt:

$$\forall B \in \mathcal{F}. \quad \mathbb{P}[B] = \sum_{i=1}^n \mathbb{P}[B|A_i] \cdot \mathbb{P}[A_i]$$

Satz. Formel von Bayes: Sei $A_1, \dots, A_n \in \mathcal{F}$ eine Partition von Ω mit $\mathbb{P}[A_i] > 0$ für alle $1 \leq i \leq n$. Für jedes $B \in \mathcal{F}$ mit $\mathbb{P}[B] > 0$ gilt:

$$\forall i = 1, \dots, n. \quad \mathbb{P}[A_i|B] = \frac{\mathbb{P}[B|A_i] \cdot \mathbb{P}[A_i]}{\sum_{k=1}^n \mathbb{P}[B|A_k] \cdot \mathbb{P}[A_k]}$$

Unabhängigkeit

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

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

Daraus folgt, dass wenn $A \in \{0, 1\}$ zu jedem Ereignis B unabhängig ist. Weiter gilt, wenn A, B unabhängig sind, so müssen auch A, B^c unabhängig sein.

Wir können Unabhängigkeit auch für mehr als zwei Ereignisse definieren. Seien $A_1, \dots, A_n \in \mathcal{F}$, so sind die Ereignisse unabhängig falls gilt:

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

Zufallsvariablen

Eine **Zufallsvariable** ist eine messbare Abbildung $X : \Omega \mapsto \mathbb{R}$, wobei Ω die Ereignismenge eines Wahrscheinlichkeitsraums $(\Omega, \mathcal{F}, \mathbb{P})$ ist.

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

Hierbei schreiben wir oftmals nur X .

Verteilungsfunktion

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

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

Satz. Aus $a < b$ folgt:

$$\mathbb{P}[a < X \leq b] = F_X(b) - F_X(a)$$

Die Verteilungsfunktion F hat folgende Eigenschaften:

- F ist monoton wachsend
- F ist rechtsstetig, d.h. $\lim_{t \rightarrow 0} F(x+t) = F(x)$.
- $\lim_{x \rightarrow -\infty} F_X(x) = 0$ und $\lim_{x \rightarrow \infty} F_X(x) = 1$

Unabhängigkeit von ZV

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

$$\forall x_1, \dots, x_n \in \mathbb{R}. \quad \mathbb{P}[X_1 \leq x_1, \dots, X_n \leq x_n] = \mathbb{P}[X_1 \leq x_1] \cdot \dots \cdot \mathbb{P}[X_n \leq x_n]$$

Eine Folge von Zufallsvariablen $X_1, X_2, \dots$ ist:

- unabhängig, falls $\forall n. \quad X_1, \dots, X_n$ unabhängig sind
- unabhängig und identisch verteilt (uiv.), falls sie unabhängig sind und $\forall i, j. \quad F_{X_i} = F_{X_j}$

Transformation von ZV

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

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

auch eine ZV. Seien $X_1, \dots, X_n$ ZV 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 ZV.

Konstruktion einer ZV

Sei eine gültige Verteilungsfunktion F_X gegeben, nun wollen wir eine dazugehörige ZV X konstruieren. Dafür brauchen wir:

Kolmogorov Theorem

$\exists(\Omega, \mathcal{F}, \mathbb{P})$ und $\exists X_1, X_2, \dots$ ZV in $(\Omega, \mathcal{F}, \mathbb{P})$ 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 $F_X = F$.

Diskrete und Stetige ZV

Per Definition ist eine Verteilungsfunktion ist immer rechtsstetig, analog dazu können wir die Linksstetigkeit definieren:

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

Jedoch ist $F(x-) = F(x)$ nicht immer wahr, d.h. nicht jede Verteilungsfunktion ist linksstetig.

Satz. $\forall x \in \mathbb{R}. \quad F(x) - F(x-) = \mathbb{P}[X = x]$. Daraus lässt sich für stetige ZV $\mathbb{P}[X = x] = 0$ folgern.

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. $\Leftrightarrow \mathbb{P}[X \leq Y] = 1$.

Diskrete ZV

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

Die **Verteilungsfunktion** einer diskreten ZV ist definiert als:

$$F(x) = \mathbb{P}[X \leq x] = \sum_{y \in W} p(y) \cdot \mathbb{I}_{y \leq x}$$

Die **Gewichtsfunktion** einer diskreten ZV ist definiert als:

$$\forall x \in W. \quad p(x) = \mathbb{P}[X = x] \text{ wobei } \sum_{x \in W} p(x) = 1$$

Diskrete Verteilungen

Bernoulli-Verteilung: Eine Bernoulli-Verteilte ZV kennt nur die Ereignisse $\{0, 1\}$, wir schreiben auch $X \sim \text{Ber}(p)$.

Binomialverteilung: Dies beschreibt die Wiederholung von Bernoulli-Experimenten. Wir schreiben $X \sim \text{Bin}(n, p)$.

Geometrische Verteilung: Eine Geometrische Verteilung beschreibt das erste Auftreten eines Erfolges. Wir schreiben $X \sim \text{Geom}(p)$.

Poisson-Verteilung: Diese Verteilung ist eine Annäherung an die Binomialverteilung für grosse n und kleine p . Sie nimmt nur Werte in $\mathbb{N}$ an. Wir schreiben $X \sim \text{Poisson}(\lambda)$.

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. $\mathbb{P}[X \in [a, b]] = \mathbb{P}[X \in (a, b)]$

Stetige Verteilungen

Gleichverteilung: Dies beschreibt die Situation wobei jedes Ereignis die gleiche Wahrscheinlichkeit hat. Wir schreiben $X \sim \mathcal{U}([a, b])$.

Exponentialverteilung: Dies ist das stetige Pendant zur Geometrischen Verteilung. Wir schreiben $X \sim \text{Exp}(\lambda)$.

Normalverteilung: Wir schreiben $X \sim \mathcal{N}(m, \sigma^2)$.

Standard Normalverteilung: $\mathcal{N}(0, 1)$. Weder für die zugehörige Dichte $\varphi(t)$ noch die Verteilungsfunktion $\Phi(t)$ gibt es geschlossene Ausdrücke, aber die Verteilung

$$\Phi(t) = \int_{-\infty}^t \varphi(s) ds = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^t e^{-\frac{s^2}{2}}$$

ist tabelliert, wobei gilt:

$$\Phi(t) = 1 - \Phi(-t) \text{ bzw. } \Phi(-t) = 1 - \Phi(t)$$

Sei $X \sim \mathcal{N}(\mu, \sigma^2)$, so ist $\frac{X-\mu}{\sigma} \sim \mathcal{N}(0, 1)$. Für $Z \sim \mathcal{N}(0, 1)$ gilt:

$$\mathbb{E}\left[Z^2\right] = 1 \quad \text{und} \quad \mathbb{E}\left[Z^4\right] = 3$$

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 .

Erwartungswert diskreter ZV

Sei $X : \Omega \mapsto \mathbb{R}$ 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 $\phi = \text{id}$, gilt

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

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:

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

Auch hier können wir analog den Erwartungswert für $\phi = \text{id}$ definieren.

Rechnen mit Erwartungswerten

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

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

Wir nennen dies auch die **Linearität des Erwartungswertes**.

Falls zwei ZV X, Y **unabhängig** sind, dann gilt auch:

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

Dies gilt nicht für die Division, hier müssen wir wie folgt vorgehen

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

und dabei $\mathbb{E}[1/Y]$ individuell berechnen. Daraus ergibt sich dann die folgende Eigenschaft:

Seien $X_1, \dots, X_n$ diskrete ZV. Dann sind die folgenden Aussagen äquivalent:

- 1. $X_1, \dots, X_n$ sind unabhängig
- 2. Für jedes $\phi_1 : \mathbb{R} \mapsto \mathbb{R}, \dots, \phi_n : \mathbb{R} \mapsto \mathbb{R}$ (messbar) beschränkt gilt:

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

Extremwertformel

Sei X eine diskrete ZV mit Werten in $\mathbb{N}$. Dann gilt folgende Identität, auch **Tailsum-Formel** genannt (**Achtung!** $n = 1$):

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

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

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

Ungleichungen

Satz. 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:

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

Varianz

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

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

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

Sei X eine ZV sodass $\mathbb{E}[X^2] < \infty$. Dann gilt für jedes $a \geq 0$:

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

Satz.

- 1. Sei X ein ZV sodass $\mathbb{E}[X^2] < \infty$ und $\lambda \in \mathbb{R}$:
$$\text{Var}[a \cdot X + b] = a^2 \cdot \text{Var}[X]$$
- 2. $S = X_1 + \dots + X_n$, wobei $X_1, \dots, X_n$ paarweise unabhängig sind, so gilt:
$$\text{Var}[S] = \text{Var}[X_1] + \dots + \text{Var}[X_n]$$

Kovarianz

Die **Kovarianz** kann verwendet werden, um die Abhängigkeit zweier ZV zu messen.

Seien X, Y zwei ZV mit $\mathbb{E}[X^2] < \infty, \mathbb{E}[Y^2] < \infty$, dann ist die **Ko-varianz** zwischen X, Y definiert als:

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

X, Y unabhängig impliziert $\text{Cov}[X, Y] = 0$, aber nicht andersherum! Eine Ausnahme ist die 2-dim Normalverteilung, hier gilt die Implikation in beide Richtungen.

Gemeinsame Verteilungen

Diskrete gemeinsame Verteilungen

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

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

Daraus ergibt sich folgende Eigenschaft. Sei $\phi : \mathbb{R}^n \mapsto \mathbb{R}$, so ist $Z = \phi(X_1, \dots, X_n)$ eine diskrete ZV mit Werten in $W = \phi(W_1 \times \dots \times W_n)$ und folgender Verteilung:

$$\forall z \in W. \quad \mathbb{P}[Z = z] = \sum_{\substack{x_1 \in W_1, \dots, x_n \in W_n \\ \phi(x_1, \dots, x_n) = z}} \mathbb{P}[X_1 = x_1, \dots, X_n = x_n]$$

Satz. Die **Randverteilung** gegeben durch:

$$\forall z \in W_i. \mathbb{P}[X_i = z] = \sum_{x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n} p(x_1, \dots, x_{i-1}, z, x_{i+1}, \dots, x_n)$$

Satz. Der Erwartungswert definiert als:

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

Satz. Folgenden Aussagen sind äquivalent:

- $X_1, \dots, X_n$ sind unabhängig
- Für alle $x_1 \in W_1, \dots, x_n \in W_n$ gilt:

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

Stetige gemeinsame Verteilungen

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

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

Wir nennen f die **gemeinsame Dichte**.

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

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

Satz. Falls $X_1, \dots, X_n$ eine gemeinsame Dichte f besitzt. Dann ist die Randverteilung:

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

Satz. Seien $X_1, \dots, X_n$ stetige ZV mit Dichten $f_1, \dots, f_n$, dann sind die folgenden Aussagen äquivalent:

- $X_1, \dots, X_n$ sind unabhängig
- $X_1, \dots, X_n$ sind stetig mit gemeinsamer Dichte:

$$f(x_1, \dots, x_n) = f_1(x_1) \cdot \dots \cdot f_n(x_n)$$

- Für alle $\phi_1, \dots, \phi_n : \mathbb{R} \mapsto \mathbb{R}$ gilt:

$$\mathbb{E}[\phi_1(x_1) \cdot \dots \cdot \phi_n(x_n)] = \mathbb{E}[\phi_1(x_1)] \cdot \dots \cdot \mathbb{E}[\phi_n(x_n)]$$

Grenzwertsätze

Gegeben eine unendliche Sequenz an uiv. ZV $X_1, X_2, \dots$, für jedes n betrachten wir die Teilsumme $S_n = X_1 + \dots + X_n$.

Gesetz der grossen Zahlen

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

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

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

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

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

falls für jedes $x \in \mathbb{R}$ gilt:

$$\lim_{n \rightarrow \infty} \mathbb{P}[X_n \leq x] = \mathbb{P}[X \leq x]$$

Zentraler Grenzwertsatz

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

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

Achtung: m und σ^2 kommen von X_1 und nicht von S_n !

Der Zentrale Grenzwertsatz besagt, dass die Verteilung einer ZV

$$Z_n = \frac{S_n - n \cdot m}{\sqrt{\sigma^2 \cdot n}} = \frac{\frac{1}{n} S_n - m}{\frac{\sigma}{\sqrt{n}}}$$

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

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

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

Statistik

Schätzer

Wir nehmen an, dass wir wie einen Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \mathbb{P})$ und Zufallsvariablen X_i haben. Zudem haben wir einen Parameterraum $\Theta \subset \mathbb{R}$, wobei $(\mathbb{P}_\theta)_{\theta \in \Theta}$ eine Familie von Wahrscheinlichkeitsmassen ist. $\mathbb{P}_\theta$ wird auch **Modell** genannt. Die Gesamtheit der beobachteten Daten nennen wir Stichprobe und die Anzahl n den Stichprobenumfang.

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

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

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

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

Der **Bias** (erwarteter 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:

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

Maximum-Likelihood-Methode

Nehmen wir an $X_1, \dots, X_n \in W$ sind ZV unter $\mathbb{P}_\theta$. Die **Likelihood Funktion** ist definiert durch:

$$L(x_1, \dots, x_n; \theta) = \begin{cases} p(x_1, \dots, x_n; \theta) & \text{im diskreten Fall} \\ f(x_1, \dots, x_n; \theta) & \text{im stetigen Fall} \end{cases}$$

wobei p respektive f die gemeinsame Gewichtsfunction / Dichtefunktion ist. Falls die X_i uiv. sind unter $\mathbb{P}_\theta$, so gilt:

$$p_X(x_1, \dots, x_n; \theta) = \prod_{i=1}^n p_{X_i}(x_i; \theta), \quad f_X(x_1, \dots, x_n; \theta) = \prod_{i=1}^n f_{X_i}(x_i; \theta)$$

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

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

Die **Log-Likelihood** Funktion hat den Vorteil, dass sie durch eine Summe anstelle eines Produkts gegeben ist und daher oftmals einfach zu berechnen ist.

Konfidenzintervalle

Wir haben nun Methoden für Schätzer von unbekannten Parameter kennengelernt. Nun wollen wir wissen wie weit diese Schätzer vom effektiven Wert p weg liegen.

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

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

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

Wenn wir nun einen $T = T_{\text{ML}} \sim \mathcal{N}(m, 1/n)$ haben, (z.Bsp. T_{ML} mit $X_1, \dots, X_n$ uiv. $\mathcal{N}(m, 1)$) suchen wir einen Konfidenzinterfall der Form:

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

Hierbei gilt:

$$\mathbb{P}_\theta[T - c/\sqrt{n} \leq m \leq T + c/\sqrt{n}] = \mathbb{P}_\theta[-c \leq Z \leq c]$$

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

μ_0	σ^2	Erwartungstreuer Schätzer	Verteilung unter $\mathbb{P}_\theta$
×	✓	$\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$	$\sqrt{n}(\frac{\bar{X}_n - \mu}{\sqrt{\sigma^2}}) \sim \mathcal{N}(0, 1)$
✓	×	$T = \frac{1}{n} \sum_{i=1}^n (X_i - \mu)^2$	$n \frac{T}{\sigma^2} \sim \chi_n^2$
×	×	$\mu : \bar{X}_n,$ $\sigma^2 : S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2$	$\sqrt{n} \frac{\bar{X}_n - \mu}{\sqrt{S^2}} \sim t_{n-1},$ $(n-1) \frac{S^2}{\sigma^2} \sim \chi_{n-1}^2$

Verteilungsaussagen

Eine stetige ZV heisst **χ^2 -verteilt** mit m Freiheitsgrade, falls ihre Dichte gegeben ist durch

$$f_X(x) = \frac{1}{2^{\frac{m}{2}} \Gamma(\frac{m}{2})} x^{\frac{m}{2}-1} e^{-\frac{1}{2}x} \mathbb{1}_{x>0},$$

wobei

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

Für natürliche Zahlen gilt $\Gamma(n) = (n-1)!$. Ein Spezialfall ist $m = 2$, hierbei erhalten wir $X \sim \text{Exp}(\frac{1}{2})$

Satz. Für ZV $X_1, \dots, X_m$ u.i.v. mit $X_i \sim \mathcal{N}(0, 1)$ ist die Summe

$$Y = \sum_{i=1}^m X_i^2 \sim \chi_m^2$$

Eine stetige ZV X heisst **t -verteilt** mit m Freiheitsgrade, falls ihre Dichte gegeben ist durch

$$f_X(x) = \frac{\Gamma(\frac{m+1}{2})}{\sqrt{m\pi} \Gamma(\frac{m}{2})} \left(1 + \frac{x^2}{m}\right)^{-\frac{m+1}{2}},$$

Satz. Für X, Y unabhängig mit $X \sim \mathcal{N}(0, 1)$ und $Y \sim \chi_m^2$, ist der Quotient

$$Z = \frac{X}{\sqrt{\frac{Y}{m}}} \sim t_m.$$

Normalverteilung mit μ, σ^2 unbekannt

Satz. Seien $X_1, \dots, X_n$ i.i.v. $\sim \mathcal{N}(\mu, \sigma^2)$. Dann sind

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

unabhängig.

Nun haben wir $X_1, \dots, X_n$ ZV, die alle unter $\mathbb{P}_\theta$ i.i.v. $\sim \mathcal{N}(\mu, \sigma^2)$ sind. Die offensichtlichen Schätzer für μ, σ^2 sind das Stichprobenmittel $\bar{X}_n$ und die Stichprobenvarianz S^2 . Für jedes $\theta \in \Theta$ gilt:

$$\frac{\bar{X}_n - \mu}{S/\sqrt{n}} \sim t_{n-1} \text{ unter } \mathbb{P}_\theta$$

Also wollen wir:

$$1 - \alpha \leq \mathbb{P}_\theta \left[\left| \frac{\bar{X}_n - \mu}{S/\sqrt{n}} \right| \leq \frac{\dots}{S/\sqrt{n}} \right]$$

Um eine kleines Intervall zu erhalten, wollen wir die Bedingung mit Gleichheit erfüllen und nehmen $\frac{\dots}{S/\sqrt{n}} = t_{n-1, 1-\alpha/2}$, somit erhalten wir für μ folgendes Konfidenzintervall:

$$\left[\bar{X}_n - t_{n-1, 1-\alpha/2} \frac{S}{\sqrt{n}}, \bar{X}_n + t_{n-1, 1-\alpha/2} \frac{S}{\sqrt{n}} \right]$$

Um ein Konfidenzintervall für σ^2 zu konstruieren brauchen wir:

$$\frac{1}{\sigma^2} (n-1) S^2 = \frac{1}{\sigma^2} \sum_{i=1}^n (X_i - \bar{X}_n)^2 \sim \chi_{n-1}^2 \text{ unter } \mathbb{P}_\theta$$

Mit der Notation $\chi_{m, \gamma}^2$ für das γ -Quantil einer χ_m^2 Verteilung wollen wir:

$$1 - \alpha = \mathbb{P}_\theta \left[\chi_{n-1, \alpha/2}^2 \leq \frac{1}{\sigma^2} (n-1) S^2 \leq \chi_{n-1, 1-\alpha/2}^2 \right]$$

Somit erhalten wir das Konfidenzintervall:

$$\left[\frac{(n-1) S^2}{\chi_{n-1, 1-\alpha/2}^2}, \frac{(n-1) S^2}{\chi_{n-1, \alpha/2}^2} \right]$$

Fazit: Das wichtigste Tool zur Bestimmung von Konfidenzintervallen sind Verteilungsaussagen über Schätzer. Die ist im Allgemeinen aber schwierig / nicht möglich.

Approximative Konfidenzintervalle

Zur Erinnerung der zentrale Grenzwertsatz besagt, dass wenn X_i in $\mathbb{P}_\theta$ i.i.v. sind, dann ist $\sum_{i=1}^n X_i$ approximativ normalverteilt mit $\mu = n\mathbb{E}[X_i]$ und $\sigma^2 = n\text{Var}_\theta[X_i]$. Insbesondere können wir daraus auch eine standard normalverteilte ZV erhalten (siehe Grenzwertsatz).

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 $\Theta_0 \cap \Theta_A = \emptyset$. Eine Hypothese heisst **einfach**, falls die Teilmenge aus einem einzelnen Wert besteht; sonst zusammengesetzt.

Test und Entscheidung

Ein **Test** ist ein (T, K) , wobei T eine ZV der Form $T = t(X_1, \dots, X_n)$ ist und $K \subseteq \mathbb{R}$ ist eine deterministische Teilmenge von $\mathbb{R}$. Man nennt 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. Dafür berechnen wir zuerst 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 nicht verworfen bzw. angenommen, falls $T(\omega) \notin K$

Ein **Fehler 1. Art** ist, wenn die H_0 verworfen wird, obschon sie richtig ist.

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

Ein **Fehler 2. Art** ist, wenn die H_0 akzeptiert wird, obschon sie falsch ist.

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

Signifikanzniveau und Macht

Bei der Auswahl eines geeigneten Tests ist insbesondere die Minimierung von Fehlern 1. Art entscheidend.

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

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

Das Sekundäre Ziel ist es den Fehler 2. Art zu vermeiden.

Die **Macht** eines Tests wird definiert als Funktion:

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

Bem. α klein entspricht einer kleinen Wahrscheinlichkeit für Fehler 1. Art, während β gross einer kleinen Wahrscheinlichkeit für einen Fehler 2. Art entspricht.

Das obige asymmetrische Vorgehen macht es schwieriger, die Nullhypothese zu verwerfen als sie beizubehalten. **Daher wählen wir die Negation der gewünschten Aussage als Nullhypothese.**

Konstruktion von Tests

Wir nehmen an, dass $X_1, \dots, X_n$ diskret oder gemeinsam stetig sind unter $\mathbb{P}_{\theta_0}$ und $\mathbb{P}_{\theta_A}$, wobei $\theta_0 \neq \theta_A$ von der einfachen Form sind.

Der **Likelihood-Quotient** ist somit wohldefiniert:

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

Wobei wir $R(x_1, \dots, x_n) = +\infty$ definieren falls $L(x_1, \dots, x_n; \theta_0) = 0$ sein sollte. Daraus ergibt sich, dass $R \gg 1 \Rightarrow H_A > H_0$ und $R \ll 1 \Rightarrow H_A < H_0$.

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

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

Neyman-Paerson-Lemma: Der LQ-Test ist optimal, da jeder andere Test mit einem kleineren Signifikanzniveau auch eine kleinere Macht hat:

Sei $c \geq 0$ und (T, K) der LQ-Test mit Parameter c .

$$\alpha^* = \mathbb{P}_{\theta_0}[T \in K]$$

Sei (T', K') ein Test mit S-Niveau $\alpha \leq \alpha^*$ so gilt:

$$\mathbb{P}_{\theta_A}[T' \in K'] \leq \mathbb{P}_{\theta_A}[T \in K]$$

z-Test

Der z-Test wird für den Erwartungswert einer Normalverteilung mit **bekannter Varianz** gebraucht. Sei also $X_1, \dots, X_n \sim \mathcal{N}(\mu, \sigma^2)$ i.i.v.:

Hypothese: $H_0 : \mu = \mu_0$

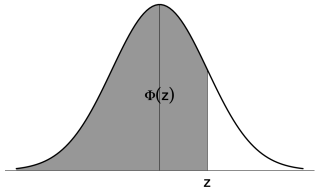
Teststatistik:

$$T = \frac{\bar{X}_n - \mu_0}{\sigma/\sqrt{n}} \sim \mathcal{N}(0, 1)$$

Kritischer Bereich:

Alternative H_A	Kritischer Bereich
$\mu < \mu_0$	$(-\infty, z_\alpha)$
$\mu > \mu_0$	$(z_{1-\alpha}, \infty)$
$\mu \neq \mu_0$	$(-\infty, z_{\alpha/2}) \cup (z_{1-\alpha/2}, \infty)$

Wobei $z_\alpha = \Phi^{-1}(\alpha)$ und $z_\alpha = -z_{1-\alpha}$.



t-Test

Der t-Test wird für den Erwartungswert einer Normalverteilung mit **unbekannter Varianz** gebraucht. Sei also $X_1, \dots, X_n \sim \mathcal{N}(\mu, \sigma^2)$ i.i.v.:

Hypothese: $H_0 : \mu = \mu_0$

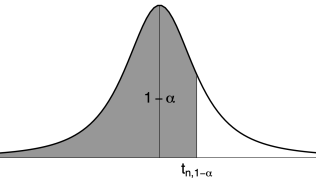
Teststatistik:

$$T = \frac{\bar{X}_n - \mu_0}{S/\sqrt{n}} \sim t_{n-1}$$

Kritischer Bereich:

Alternative H_A	Kritischer Bereich
$\mu < \mu_0$	$(-\infty, t_{n-1, \alpha})$
$\mu > \mu_0$	$(t_{n-1, 1-\alpha}, \infty)$
$\mu \neq \mu_0$	$(-\infty, t_{n-1, \alpha/2}) \cup (t_{n-1, 1-\alpha/2}, \infty)$

Tipp: $t_{n-1, \alpha} = -t_{n-1, 1-\alpha}$



p-Wert

Wir wollen eine Hypothese $H_0 : \theta = \theta_0$ gegen eine Alternativhypothese $H_A : \theta \in \Theta_A$ testen. Eine Familie von Tests $(T, (K_t)_{t \geq 0})$ heisst geordnet bzgl. T falls $K_t \subset \mathbb{R}$ und $s \leq t \Rightarrow K_t \subset K_s$ gilt. Typische Beispiele dafür sind $K_t = [t, \infty)$ (rechtsseitiger Test), $K_t = (-\infty, -t]$ (linksseitiger Test) und $K_t = (-\infty, -t] \cup [t, \infty)$ (beidseitiger Test).

Sei $H_0 : \theta = \theta_0$ eine einfache Nullhypothese und $(T, K_t)_{t \geq 0}$ 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 ist definiert als Wahrscheinlichkeit - unter der Nullhypothese - den beobachteten Wert der Prüfgrösse oder einen in Richtung der Alternativhypothese "extremere" Wert zu erhalten.

Der p -Wert hat folgende Eigenschaften:

- 1. Sei T stetig und $K_t = (t, \infty)$, so 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.

Somit ist der p -Wert das kleinste Signifikanzniveau bei dem wir H_0 verwerfen. Wir können zusammenfassend sagen:

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

Sonstiges

Nützliche Formeln

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

$$\binom{n}{k} = \frac{n!}{(n-k)!k!}$$

$$\sum_{k=0}^{\infty} q^k = \frac{1}{q} \quad \text{für } |q| < 1$$

$$\sum_{k=0}^n q^k = \frac{1 - q^{n+1}}{1 - q} \quad \text{für } q \neq 1$$

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

Exponentialfunktion / Logarithmus

Für die Exponentialfunktion gilt:

- (1) $\exp(x) \exp(y) = \exp(x + y)$
- (2) $\exp(x) > 1, \quad \forall x > 0$
- (3) $x^a = \exp(a \cdot \ln(x))$ und $x^0 = 1$
- (4) $\exp(iz) = \cos(z) + i \cdot \sin(z)$
- (5) $\exp(i \cdot \frac{\pi}{2}) = i, \exp(i\pi) = -1$ und $\exp(2i\pi) = 1$

Der natürliche Logarithmus $\ln :]0, \infty[\rightarrow \mathbb{R}$ bildet die Umkehrfunktion zu exp und ist streng monoton wachsend und stetig. Für den natürliche Logarithmus gilt:

- (1) $\ln(1) = 0$
- (2) $\ln(e) = 1$
- (3) $\ln(a \cdot b) = \ln(a) + \ln(b)$

- (4) $\ln(a/b) = \ln(a) - \ln(b)$
- (5) $\ln(x^a) = a \cdot \ln(x)$
- (6) $x^a \cdot x^b = x^{a+b}$
- (7) $(x^a)^b = x^{a \cdot b}$

Im Allgemeinen gilt $\log_b(a) = \frac{\ln(a)}{\ln(b)}$.

Ableitungs und Integrations Regeln

$F(x)$	$F'(x) = f(x)$
Summenregel	$(f(x) + g(x))' = f'(x) + g'(x)$
Produktregel	$(f(x) \cdot g(x))' = f'(x) \cdot g(x) + f(x) \cdot g'(x)$
Quotientenregel	$\left(\frac{f(x)}{g(x)}\right)' = \frac{f'(x) \cdot g(x) - f(x) \cdot g'(x)}{g^2(x)}$ wenn $g(x) \neq 0$
Kettenregel	$(f(g(x)))' = f'(g(x)) \cdot g'(x)$
Part. Integration	$\int f'(x) \cdot g(x) dx = f(x) \cdot g(x) - \int f(x) \cdot g'(x) dx$
Substitution	$\int_{\varphi(a)}^{\varphi(b)} f(x) \, dx = \int_a^b f(\varphi(t)) \cdot \varphi'(t) \, dt$
Logarithmus	$\int \frac{f'(t)}{f(t)} \, dt = \log(f(x))$

Bsp. Substitution: Wir wollen $\int \cos(x^2) 2x \, dx$ berechnen dabei gehen wir wie folgt vor:

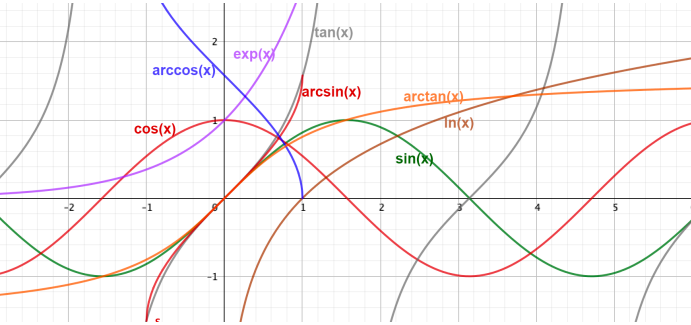
- 1. Zuerst bestimmen wir die Substitution: $u = x^2$
- 2. Nun berechnen wir die Umkehrfunktion: $x = \sqrt{u}$
- 3. Dann brauchen wir noch: $\frac{du}{dx} = \frac{dx^2}{dx} = \frac{2x}{1} = 2x \implies dx = \frac{du}{2x}$
- 4. Zuletzt können wir dies im Integral einsetzen und erhalten:

$$\int \cos(x^2) 2x \, dx = \int \cos(u) 2x \frac{du}{2x} = \int \cos(u) \, du = \sin(u)$$

Typische Ableitungen und Stammfunktionen

$F(x)$	$F'(x) = f(x)$
c	0
x^a	$a \cdot x^{a-1}$
$\frac{1}{a+1} x^{a+1}$	x^a
$\frac{1}{a \cdot (n+1)} (ax + b)^{n+1}$	$(ax + b)^n$
$\frac{x^{\alpha+1}}{\alpha+1}$	$x^\alpha, \alpha \neq -1$
$\sqrt{x}$	$\frac{1}{2\sqrt{x}}$
$\sqrt[n]{x}$	$\frac{1}{n} x^{\frac{1}{n}-1}$
$\frac{2}{3} x^{\frac{3}{2}}$	$\sqrt{x}$
$\frac{n}{n+1} x^{\frac{1}{n}+1}$	$\sqrt[n]{x}$
e^x	e^x
$\ln(x)$	$\frac{1}{x}$
$\log_a x $	$\frac{1}{x \ln a} = \log_a(e) \frac{1}{x}$
a^{cx}	$a^{cx} \cdot c \ln a$
x^x	$x^x \cdot (1 + \ln x) \quad x > 0$
$x \cdot (\ln x - 1)$	$\ln x $

Funktionen



Aufgaben

Multiple Choice

Seien $A_1, A_2, A_3 \in \mathcal{F}$ paarweise unabhängige Ereignisse, welche Aussage ist korrekt?

- ☒ Die Ereignisse A_1, A_2, A_3 sind nicht zwangsläufig unabhängig
- ☐ Die Ereignisse A_1, A_2, A_3 sind zwangsläufig unabhängig

Seien A, B zwei Ereignisse. $\mathbb{P}[A \cup B] = \mathbb{P}[A] + \mathbb{P}[B]$ gilt, falls:

- ☐ A, B unabhängig sind
- ☒ A, B disjunkt sind
- ☐ A eine Teilmenge von B ist

Seien A, B, C Ereignisse. Welche der folgenden Aussagen ist wahr?

- ☐ Falls A, B sowie A, C unabhängig sind, so sind auch $A, B \cap C$ unabhängig
- ☐ Falls A, B sowie B, C unabhängig sind, so sind auch A, C unabhängig
- ☒ Falls A, B und C unabhängig sind, so sind auch $A, B \cap C$ unabhängig

Seien X, Y reelle ZV. Welche Aussage ist im allgemeinen falsch?

- ☐ $\text{Var}[X] \geq 0$
- ☒ $\text{Var}[X + Y] = \text{Var}[X] + \text{Var}[Y]$
- ☐ $\text{Var}[aX + b] = a^2 \text{Var}[X]$

Für stetige ZV gilt immer:

- ☒ Die Verteilungsfunktion ist stetig
- ☐ Die Dichtefunktion ist stetig
- ☐ Weder noch

Für die Dichte ist eine Gleichverteilte ZV ein Gegenbeispiel.

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.

Sei $Z = (X, Y)$ eine $\mathbb{R}^2$ -wertige ZV mit Dichte:

$$f_Z(x, y) = \frac{1}{2\pi} e^{-\frac{1}{2}(x^2 + y^2)}$$

Welche Aussage ist korrekt?

- ☐ X und Y sind korreliert und nicht unabhängig
- ☐ X und Y sind unkorreliert und nicht unabhängig
- ☒ X und Y sind unkorreliert und unabhängig

Sei X eine reellwertige ZV mit Dichte f_X und Verteilungsfunktion F_X . Welche der Aussagen ist im Allgemeinen falsch?

- ☐ F_X ist stetig
- ☒ F_X ist strikt monoton wachsend
- ☐ $\mathbb{P}[a \leq X \leq b] = F(b) - F(a)$ für $-\infty < a \leq b < \infty$

F_X kann auch nur monoton wachsend sein, nicht strikt.

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

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

Sei X eine ZV mit F_X . Was gilt für $a < b$?

- ☐ $\mathbb{P}[a < X < b] = F_X(b) - F_X(a)$
- ☒ $\mathbb{P}[a < X \leq b] = F_X(b) - F_X(a)$
- ☐ $\mathbb{P}[a \leq X \leq b] = F_X(b) - F_X(a)$

Sei $X \sim \text{Poisson}(\lambda)$, $\lambda > 0$, welche der Aussagen ist korrekt?

- ☐ $\mathbb{P}[X > 5] = 1 - \mathbb{P}[X < 5]$
- ☒ $\mathbb{P}[X \leq 1 | X \geq 1] = \frac{\lambda}{1+\lambda}$
- ☐ $2X \sim \text{Poisson}(2\lambda)$

Sei $X \sim \text{Poisson}(\lambda)$, $\lambda > 0$, was ist $\mathbb{E}[X^2]$?

- ☐ λ
- ☐ $1/\lambda^2$
- ☒ $\lambda \cdot (\lambda + 1)$
- ☐ λ^2

Es gilt $\mathbb{P}[X > t + s | X > s] = \mathbb{P}[X > t]$ für alle $t, s \geq 0$, falls

- ☐ $X \sim \mathcal{U}([a, b])$
- ☐ $X \sim \text{Poisson}(\lambda)$
- ☒ $X \sim \text{Exp}(\lambda)$ (Gedächtnislosigkeit)

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

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

Sei $X \sim \mathcal{N}(0, 1)$ und $Y \sim U[0, 1]$ ZV mit $\text{Cov}(X, Y) = 0$. Welche Aussage ist korrekt?

- ☐ X, Y unabhängig sind
- ☒ $\mathbb{E}[XY] = 0$

- ☐ Keine der beiden anderen Aussagen ist wahr

Dies gilt weil $\mathbb{E}[X] = 0$ und $\text{Cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] = 0 \Rightarrow \mathbb{E}[XY] = 0$.

Sei $X_1, X_2, \dots$ eine Folge von ZV, welche alle $\mathbb{E}[X_i] = \mu$ und $\text{Var}[X_i] = \sigma^2$ haben. Sei $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$. Dann gilt:

- ☐ $\bar{X} \xrightarrow{n \rightarrow \infty} \mu$ in Wahrscheinlichkeit
- ☐ $\bar{X} \xrightarrow{n \rightarrow \infty} \mu$ P-fastsicher
- ☒ Im Allgemeinen sind 1. und 2. falsch

Falls zu einem gegebenen Test die Hypothese auf dem 2.5% Niveau abgelehnt wird, dann

- ☒ wird sie auch auf dem 5% Niveau abgelehnt
- ☐ wird sie auch auf dem 1% Niveau abgelehnt
- ☐ Im Allgemeinen kann nicht behauptet werden, dass sie auf einem höheren oder tieferen Niveau auch abgelehnt wird

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

- ☐ Die ZV X, Y sind nicht notwendigerweise stetig, dies hängt von $f_{X,Y}$ ab
- ☒ Die ZV X, Y sind immer stetig

Wir betrachten die gemeinsame Verteilung von zwei diskreten ZV X, Y . Welche Aussagen sind korrekt?

- ☐ Aus den einzelnen Gewichtsfunktionen kann man immer die gemeinsame Gewichtsfunktion berechnen
- ☐ Aus den einzelnen Gewichtsfunktionen und $\text{Cov}(X, Y)$ kann man immer die gemeinsame Gewichtsfunktion berechnen
- ☒ Aus der gemeinsamen Gewichtsfunktion kann man immer die einzelnen Gewichtsfunktionen berechnen

Sei $X_1, \dots, X_n$ uiv. $\sim \mathcal{N}(0, 1)$. Welche Aussagen sind korrekt?

- ☒ $X_1^2 + \dots + X_n^2 \sim \chi_n^2$
- ☒ $\frac{1}{n}(X_1 + \dots + X_n)^2 \sim \chi_n^2$
- ☐ $(X_1 + \dots + X_n)^2 \sim \chi_n^2$
- ☒ $X_1^2 + X_2^2 \sim \text{Exp}(1/2)$

Wenn das Signifikanzniveau α eines Test kleiner wird, dann

- ☐ wird der Verwerfungsbereich für die Nullhypothese grösser
- ☐ wird die Macht des Tests grösser
- ☒ wird die Wahrscheinlichkeit für den Fehler 2. Art grösser

Sonstige Aufgaben

Aufgabe Es werden ein blauer und ein grüner Würfel geworfen. Wir wählen $\Omega = \{1, \dots, 6\}^2$. Wir betrachten die Algebra:

$$\mathcal{F}_{\text{sym}} = \{A \subseteq \Omega \mid \forall (\omega_1, \omega_2) \in \Omega, (\omega_1, \omega_2) \in A \Leftrightarrow (\omega_2, \omega_1) \in A\}$$

Zeige, dass $\mathcal{F}_{\text{sym}}$ eine σ -Algebra ist.

Wir müssen 3 Eigenschaften überprüfen.

- $\forall (\omega_1, \omega_2) \in \Omega \Leftrightarrow (\omega_2, \omega_1) \in \Omega$ daher gilt $\Omega \in \mathcal{F}_{\text{sym}}$.

- Sei $A \in \mathcal{F}_{\text{sym}}$. Somit gilt für jedes $(\omega_1, \omega_2) \in \Omega$

$$(\omega_1, \omega_2) \in A \Leftrightarrow (\omega_2, \omega_1) \in \Omega$$

was äquivalent ist zu

$$(\omega_1, \omega_2) \in A^c \Leftrightarrow (\omega_2, \omega_1) \in A^c$$

somit ist $A^c \in \mathcal{F}_{\text{sym}}$.

- Seien $A_1, A_2, \dots \in \mathcal{F}_{\text{sym}}$ gilt:

$$\begin{aligned} (\omega_1, \omega_2) \in \bigcup_{i=1}^{\infty} A_i &\Leftrightarrow \exists i. (\omega_1, \omega_2) \in A_i \\ &\Leftrightarrow \exists i. (\omega_2, \omega_1) \in A_i \\ &\Leftrightarrow (\omega_2, \omega_1) \in \bigcup_{i=1}^{\infty} A_i \end{aligned}$$

Somit folgt $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}_{\text{sym}}$.

Aufgabe Seien $\mathbb{P}[A] = 0.1, \mathbb{P}[B] = 0.3$ zudem wissen wir, dass $\mathbb{P}[\text{nicht } A, B] = 0.7$. Was ist $\mathbb{P}[A \cap B]$?

$$\mathbb{P}[A \cap B] = \mathbb{P}[A] + \mathbb{P}[B] - \mathbb{P}[A \cup B] = 0.1 + 0.3 - (1 - \mathbb{P}[(A \cup B)^c]) = 0.1$$

Aufgabe Sei $(X_i)_{i \geq 1}$ eine unendliche Folge von unabhängig Ber(1/2)-verteilten ZV. Wir betrachten folgenden Algorithmus:

```
i ← 1
while Xi = Xi+1 = 1 do
    i = i + 2
end while
return Z = Xi + 2 · Xi+1
```

1) Zeige, dass der Algorithmus immer nach endlich vielen Schritten terminiert.

Wir definieren $A_j := \{\text{While-Schleife wird j-Mal durchlaufen}\}$ und berechnen:

$$\begin{aligned} \mathbb{P}[A_j] &= \mathbb{P}\left[\bigcap_{i=1}^{2j} \{X_i = 1\} \cap (\{X_{2j+1} = 0\} \cup \{X_{2j+2} = 0\})\right] \\ &= \left(\prod_{i=1}^{2j} \mathbb{P}[X_i = 1]\right) \cdot \mathbb{P}[X_{2j+1} = 0] \cdot \mathbb{P}[X_{2j+2} = 0] \\ &= \left(\frac{1}{2}\right)^{2j} \cdot \frac{3}{4} \end{aligned}$$

Wenn wir nun über alle A_j summieren, sehen wir, dass der Algorithmus immer in endlich Schritten terminieren wird.

2) Zeige, dass Z eine gleichverteilte ZV in $\{0, 1, 2\}$ ist.

Wir wissen, dass alle A_j disjunkt sind.

$$\begin{aligned} \mathbb{P}[Z = 0] &= \mathbb{P}[\{Z = 0\} \cap A] + \mathbb{P}[\{Z = 0\} \cap A^c] = \sum_{j=0}^{\infty} \mathbb{P}[\{Z = 0\} \cap A_j] \\ &= \sum_{j=0}^{\infty} \left(\frac{1}{2}\right)^{2j+2} = \frac{1}{4} \sum_{j=0}^{\infty} \left(\frac{1}{4}\right)^j = \frac{1}{4} \cdot \frac{1}{1 - \frac{1}{4}} = \frac{1}{3} \end{aligned}$$

Dies können wir nun auch für 1, 2 machen uns sehen, dass Z gleichverteilt sein muss.

3) Sei T die Laufzeit (Anzahl Schleifendurchläufe), berechne $\mathbb{E}[T]$. Wir verwenden die Tailsum Formel: $\mathbb{E}[T] = \sum_{j=1}^{\infty} \mathbb{P}[T \geq j]$. Für $\mathbb{P}[T \geq j]$ gilt:

$$\mathbb{P}[T \geq j] = \mathbb{P}\left[\bigcap_{i=1}^{2j} \{X_i = 1\}\right] = \prod_{i=1}^{2j} \mathbb{P}[X_i = 1] = (1/4)^j$$

Daher gilt:

$$\mathbb{E}[T] = \sum_{j=1}^{\infty} \mathbb{P}[T \geq j] = \left(\sum_{j=0}^{\infty} (1/4)^j\right) - 1 = \frac{1}{1 - 1/4} - 1 = \frac{1}{3}$$

4) Konstruiert einen Algorithmus, der eine Ber(1/5)-verteilten ZV ausgibt. Wir betrachten folgenden Algorithmus:

```

i ← 1
while  $X_i = X_{i+2} = 1$  or  $X_{i+1} = X_{i+2} = 1$  do
  i = i + 3
end while
return  $Z = X_i + 2 \cdot X_{i+1} + 4 \cdot X_{i+2} = 4 ? 1 : 0$ 

```

Es ist leicht wie in (1), (2) zu zeigen, dass er alle Eigenschaften erfüllt.

5) Zeige, dass die ZV $S = T + 1 \sim \text{Geom}(3/4)$ ist.

$$\mathbb{P}[T = j] = \mathbb{P}[T \geq j] - \mathbb{P}[T \geq j + 1] = \left(\frac{1}{4}\right)^j - \left(\frac{1}{4}\right)^{j+1} = \frac{3}{4} \left(\frac{1}{4}\right)^j$$

Somit ist $S \sim \text{Geom}(3/4)$.

Aufgabe Konstruiere aus $U \sim \mathcal{U}[0, 1]$ ein $\text{Ber}(1/2)$ verteilte ZV Z .

Die verallgemeinerte Inverse F^{-1} einer $\text{Ber}(1/2)$ verteilte ZV ist definiert als:

$$F^{-1}(\alpha) = \begin{cases} 0, & 0 < \alpha < 2/3 \\ 1, & 2/3\alpha < 1 \end{cases}$$

Aus Theorem 2.12 folgt, dass $Z := F^{-1}(U)$ $\text{Ber}(1/2)$ verteilt ist.

Aufgabe Seien X, Y ZV mit gemeinsamer Dichte $f_{X,Y}(x, y) = 4 \frac{y}{x^3} \mathbb{I}_{0 < x \leq 1, 0 < y \leq x^2}$. Berechne $\mathbb{E}[X/Y]$.

$$\begin{aligned} \mathbb{E}[X/Y] &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{x}{y} f_{X,Y}(x, y) dy dx \\ &= \int_0^1 \int_0^{x^2} \frac{x}{y} \cdot 4 \frac{y}{x^3} dy dx = \int_0^1 \int_0^{x^2} \frac{4}{x^2} dy dx \\ &= \int_0^1 \frac{4}{x^2} \cdot x^2 dx = 4 \end{aligned}$$

Aufgabe Sei $T \sim \text{Exp}(\lambda)$. Berechne die Dichte von $T' = c \cdot T^2$ und den Erwartungswert von T' .

Sei $\phi: \mathbb{R} \mapsto \mathbb{R}$ messbar und beschränkt. Wir definieren $\psi(x) = \phi(c \cdot x^2)$. Somit erhalten wir:

$$\begin{aligned} \mathbb{E}[\phi(T')] &= \mathbb{E}[\phi(c \cdot T^2)] = \mathbb{E}[\psi(T)] = \int_{-\infty}^{\infty} \psi(x) \lambda e^{-\lambda x} \mathbb{1}_{x \geq 0} dx \\ &= \int_0^{\infty} \phi(c \cdot x^2) \lambda e^{-\lambda x} dx = \int_0^{\infty} \phi(y) \lambda e^{-\lambda \sqrt{y/c}} \frac{dy}{2\sqrt{cy}} \end{aligned}$$

Wobei wir die Dichte der Exponentialverteilung verwendet haben. daraus folgt:

$$f_{T'}(y) = \frac{\lambda}{2\sqrt{cy}} e^{-\lambda \sqrt{y/c}}$$

Für den Erwartungswert gilt $\mathbb{E}[c \cdot T^2] = c \cdot \mathbb{E}[T^2]$:

$$\mathbb{E}[T^2] = \int_0^{\infty} x^2 \lambda e^{-\lambda x} dx = \frac{2}{\lambda^2}$$

Somit erhalten wir $\mathbb{E}[T'] = \frac{2c}{\lambda^2}$.

Aufgabe Seien $P \sim \mathcal{U}[0, 1]$, $H \sim \mathcal{U}[0, P]$ ZV. Bestimme die gemeinsame Dichte von P und H .

Wir wissen:

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

Somit ist:

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

Aufgabe Sei X eine ZV mit Dichte f_X und sei $Y = e^X$. Was ist die Dichte von $f_Y(y)$, $y > 0$, von Y ?

$$f_Y = \mathbb{P}[Y \leq y] = \mathbb{P}[e^X \leq y] = \mathbb{P}[X \leq \log y] = F_X(\log y)$$

$$\Rightarrow f_y(y) = \frac{d}{dy} F_Y(y) = \frac{f_X(\log y)}{y}$$

Aufgabe Seien X_1, X_2 unabhängige ZV, beide gleichverteilt auf dem Intervall $[0, 1]$ und sei $X = \max(X_1, X_2)$. Berechne die Dichtefunktion von X und $\mathbb{P}[X_1 \leq x | X \geq y]$.

$$F_X(t) = \mathbb{P}[\max(X_1, X_2) \leq t] = \mathbb{P}[X_1 \leq t] \cdot \mathbb{P}[X_2 \leq t] = F_{X_1}(t) \cdot F_{X_2}(t)$$

$$\Rightarrow f_X(t) = \frac{d}{dt} F_{X_1}(t) \cdot F_{X_2}(t) = \frac{d}{dt} t^2 \cdot \mathbb{I}_{0 \leq t \leq 1} = 2t \cdot \mathbb{I}_{0 \leq t \leq 1}$$

Für die Wahrscheinlichkeit brauchen wir eine Fallunterscheidung:

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

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

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

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

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

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

Aufgabe Seien X, Y diskrete ZV mit Gewichtsfunktion:

$$p(j, k) = \begin{cases} C \cdot \left(\frac{1}{2}\right)^k & \text{für } k = 2, 3, \dots \text{ und } j = 1, \dots, k-1 \\ 0 & \text{sonst} \end{cases}$$

Bestimme die Konstante C .

$$\begin{aligned} 1 &\stackrel{!}{=} \sum_{k=2}^{\infty} \sum_{j=1}^{k-1} C \cdot \left(\frac{1}{2}\right)^k = C \cdot \sum_{j=1}^{\infty} \left(\frac{1}{2}\right)^{j+1} \sum_{k=0}^{\infty} \left(\frac{1}{2}\right)^k \\ &= C \cdot \sum_{j=1}^{\infty} \frac{\left(\frac{1}{2}\right)^{j+1}}{1 - \frac{1}{2}} = C \cdot \sum_{j=1}^{\infty} \left(\frac{1}{2}\right)^j = C \cdot \frac{\frac{1}{2}}{1 - \frac{1}{2}} = C \end{aligned}$$

Beweise, dass bedingt auf das Ereignis $Y = k$ die ZV X gleichverteilt auf $\{1, \dots, k-1\}$ ist.

$$p_Y(k) = \sum_{j=1}^{k-1} \left(\frac{1}{2}\right)^k = (k-1) \cdot \left(\frac{1}{2}\right)^k$$

Nun verwenden wir den Satz der bedingten Wahrscheinlichkeit:

$$p_{X|Y}(j, k) = \frac{p(j, k)}{p_Y(k)} = \frac{\left(\frac{1}{2}\right)^2}{(k-1) \cdot \left(\frac{1}{2}\right)^2} = \frac{1}{1-k} \quad \text{für } j = 1, \dots, k-1$$

Somit ist X gegeben $Y = k$ gleichverteilt auf $\{1, \dots, k-1\}$.

Aufgabe Sei X eine ZV mit Verteilungsfunktion F_X . Zeige, dass X diskret ist.

$$F_X(a) = \begin{cases} 0, & a < 1 \\ 1/5, & 1 \leq a < 4 \\ 3/4, & 4 \leq a < 6 \\ 1, & 6 \leq a \end{cases}$$

Wir stellen fest, dass $\mathbb{P}[X = x] = 0$ für alle $x \notin \{1, 4, 6\}$. Da die Menge $\{1, 4, 6\}$ endlich ist, ist die ZV diskret.

Aufgabe Sei T eine ZV mit Verteilungsfunktion F_T . Zeige, dass t stetig ist.

$$F_T(a) = \begin{cases} 0, & a < 0 \\ 1 - e^{-2a}, & a \geq 0 \end{cases}$$

Wir stellen fest, dass F_T stückweise stetig differenzierbar ist (auf $(-\infty, 0)$ und $(0, \infty)$). Somit folgt, dass T eine stetige ZV ist.

Aufgabe Seien X, Y ZV mit gemeinsamer Dichtefunktion:

$$f(x, y) = \frac{1}{x^2 y^2} \quad \text{für } x \geq 1, y \geq 1$$

Berechne die Verteilungsfunktion F_U von $U = \frac{X}{Y}$.

Für $u \leq 0$, $F(u) = 0$. Für $u > 1$, $(X > Y)$:

$$\begin{aligned} F(u) &= \mathbb{P}[X/Y \leq u] = \mathbb{P}[X \leq Y u] = \int_1^{\infty} \int_1^{yu} \frac{1}{x^2 y^2} dx dy \\ &= \int_1^{\infty} \frac{1}{y^2} - \frac{1}{uy^3} dy = 1 - \frac{1}{2u} \end{aligned}$$

Für $0 < u \leq 1$, $(X \leq Y)$ gehen wir gleich vor, aber mit der unteren Integralgrenze $1/u$ für das äussere Integral. Wir erhalten dann $F(u) = \frac{u}{2}$.

Aufgabe Die ZV X hat Verteilungsfunktion $F_\alpha(x) = \exp(-\exp(-(x-\alpha)))$. Zeige, dass F_α eine Verteilungsfunktion ist.

Für F_α muss folgendes gelten:

- (rechts)stetigkeit
- monoton wachsend
- $\lim_{x \rightarrow \infty} F_\alpha(x) = 1$
- $\lim_{x \rightarrow -\infty} F_\alpha(x) = 0$

Die Funktion ist stetig da sowohl $x - \alpha$ als auch e^{-x} stetige Funktionen sind und f, g stetig $\Rightarrow g \circ f$ stetig. Für die Monotonie verwenden wir: x monoton wachsend $\Rightarrow e^x$ monoton wachsend. $x - \alpha$ ist monoton wachsend $\Rightarrow e^{-(x-\alpha)}$ ist monoton fallend $\Rightarrow \exp(-\exp(-(x-\alpha)))$ ist monoton wachsend. Zuletzt gilt noch:

$$\lim_{x \rightarrow -\infty} e^{-e^{-(x-\alpha)}} = \lim_{x \rightarrow \infty} e^{-x} = 0, \quad \lim_{x \rightarrow \infty} e^{-e^{-(x-\alpha)}} = \lim_{x \rightarrow \infty} e^0 = 1$$

Aufgabe Zeige, dass für eine diskrete ZV Z mit Werten in $\mathbb{N}$ gilt:

$$\exists q \in (0, 1) : Z \sim \text{Geom}(q) \Leftrightarrow \mathbb{P}[Z > n] = \mathbb{P}[Z > n + k | Z > k] \forall n, k \geq 1$$

" $\Rightarrow$ " Sei $Z \sim \text{Geom}(q)$. Dann gilt $\mathbb{P}[Z > k] = (1-q)^k$ und somit:

$$\begin{aligned} \mathbb{P}[Z > n + k | Z > k] &= \frac{\mathbb{P}[Z > n + k, Z > k]}{\mathbb{P}[Z > k]} = \frac{\mathbb{P}[Z > n + k]}{\mathbb{P}[Z > k]} \\ &= \frac{(1-q)^{n+k}}{(1-q)^k} = (1-q)^n = \mathbb{P}[Z > n] \end{aligned}$$

" $\Leftarrow$ " Sei $\mathbb{P}[Z > n] = \mathbb{P}[Z > n + k | Z > k]$. Dann gilt:

$$\mathbb{P}[Z > n] = \mathbb{P}[Z > n + k | Z > k] = \frac{\mathbb{P}[Z > n + k, Z > k]}{\mathbb{P}[Z > k]} = \frac{\mathbb{P}[Z > n + k]}{\mathbb{P}[Z > k]}$$

Definiere nun $f(n) = \mathbb{P}[Z > n]$:

$$f(n) = \frac{f(n+k)}{f(k)} \Rightarrow f(n)f(k) = f(n+k)$$

Durch Iteration folgt, dass $f(n) = a^n$ mit $a = f(1)$ und damit:

$$\mathbb{P}[Z = n] = \mathbb{P}[Z > n-1] - \mathbb{P}[Z > n] = f(n-1) - f(n) = (1-a) \cdot a^{n-1}$$

Da $a = f(1) = \mathbb{P}[Z > 1] \in [0, 1]$ folgt $q = 1 - a \in [0, 1]$ und damit $Z \sim \text{Geom}(q)$.

Aufgabe Seien $X \sim \mathcal{P}(\lambda)$, $Y \sim \mathcal{P}(\mu)$, zeige, dass $X + Y \sim \mathcal{P}(\lambda + \mu)$.

Für $k \in \mathbb{N}_0$:

$$\begin{aligned}\mathbb{P}[X+Y=k] &= \sum_{l=0}^k \mathbb{P}[X+Y=k, Y=l] = \sum_{l=0}^k \mathbb{P}[X=k-l, Y=l] \\ &= \sum_{l=0}^k \frac{\lambda^{k-l}}{(k-l)!} e^{-\lambda} \frac{\mu^l}{l!} e^{-\mu} = e^{-(\lambda+\mu)} \sum_{l=0}^k \frac{\lambda^{k-l}}{(k-l)!} \frac{\mu^l}{l!} \\ &= e^{-(\lambda+\mu)} \frac{1}{k!} \sum_{l=0}^k \frac{k!}{(k-l)! \cdot l!} \lambda^{k-l} \mu^l = e^{-(\lambda+\mu)} \frac{(\lambda+\mu)^k}{k!}\end{aligned}$$

Aufgabe Seien W_1, W_2 endlich oder abzählbar und $p: W_1 \times W_2 \mapsto [0, 1]$ eine Funktion mit:

$$\sum_{x_1 \in W_1, x_2 \in W_2} p(x_1, x_2) = 1$$

Seien weiterhin U_1, U_2 zwei unabhängige $\mathcal{U}[0, 1]$ verteilte ZV. Ziel ist es mithilfe dieser ZV zwei weitere ZV X_1, X_2 zu konstruieren, sodass (X_1, X_2) die gemeinsame Gewichtsfunktion p haben.

1) Was ist die Gewichtsfunktion p_{X_1} der Randverteilung von X_1 ? Es muss gelten:

$$p_{X_1}(x_1) = \sum_{x_2 \in W_2} p(x_1, x_2)$$

Wir definieren $q_0 = 0, q_j = \sum_{i=1}^j p_{X_1}(w_i)$ und $X_1(\omega) = \sum_{j \geq 1} w_j \mathbb{I}_{U_1(\omega) \in (q_{j-1}, q_j]}$. Um zu überprüfen, dass X_1 die gewünschte Gewichtsfunktion hat, berechnen wir:

$$\mathbb{P}[X_1 = x_1] = \sum_{j \geq 1} \mathbb{P}[U_1(\omega) \in (q_{j-1}, q_j]] \mathbb{I}_{x_1 = w_j} = p_{X_1}(x_1)$$

was genau zu zeigen war.

Aufgabe Seien U_1, U_2, U_3 unabhängige $\mathcal{U}[0, 1]$ verteilte ZV. Wir betrachten die stetigen ZV:

$$L = \min(U_1, U_2, U_3), \quad M = \max(U_1, U_2, U_3)$$

1) Berechne die Dichte von M und L . Die drei ZV haben die gemeinsame Dichte:

$$f(u_1, u_2, u_3) = \mathbb{I}_{u_1 \in [0, 1]} \mathbb{I}_{u_2 \in [0, 1]} \mathbb{I}_{u_3 \in [0, 1]}$$

Sei nun also ϕ stückweise stetig und beschränkt. Wir berechnen:

$$\mathbb{E}[\phi(M)] = \int_0^1 \int_0^1 \int_0^1 \phi(\max(u_1, u_2, u_3)) du_1 du_2 du_3$$

Wir unterscheiden 6 verschiedene Fälle, je nachdem welche Variable das Maximum annimmt. Wir berechnen den Fall $u_3 < u_2 < u_1$ und erhalten:

$$\int_{-\infty}^{\infty} \phi(u_1) \frac{1}{2} u_1^2 \mathbb{I}_{u_1 \in [0, 1]} du_1$$

Da die sechs Fälle symmetrisch sind haben wir

$$\mathbb{E}[\phi(M)] = 6 \cdot \int_{-\infty}^{\infty} \phi(m) \frac{1}{2} m^2 \mathbb{I}_{m \in [0, 1]} dm$$

somit ist die Dicht $f_M(m) = 3m^2 \mathbb{I}_{m \in [0, 1]}$. Da $(1 - U_1, 1 - U_2, 1 - U_3)$ aus Symmetriegründen die gleiche gemeinsame Dichte hat wie (U_1, U_2, U_3) , hat L die gleiche Dichte wie $1 - M$ und auf diese Weise erhält man ebenfalls $f_L(l) = 3(1 - l)^2 \mathbb{I}_{l \in [0, 1]}$.

2) Zeige, dass für φ, ϕ stückweise stetig, beschränkt

$$\mathbb{E}[\phi(M)\varphi(L)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \phi(m)\varphi(l)6(m-l)\mathbb{I}_{0 \leq l \leq m \leq 1} dldm$$

Wir verwenden das Resultat aus der ersten Teilaufgabe und die Symmetrie:

$$\begin{aligned}&\int_0^1 \int_0^1 \int_0^1 \phi(\max(u_1, u_2, u_3)) \varphi(\max(u_1, u_2, u_3)) du_1 du_2 du_3 \\ &= \int_0^1 \int_0^1 \phi(u_1) \varphi(u_3) (u_1 - u_3) \mathbb{I}_{u_3 \leq u_1} du_1 du_3 \\ &\Rightarrow \mathbb{E}[\phi(M)\varphi(L)] = 6 \cdot \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \phi(m)\varphi(l)(m-l)\mathbb{I}_{0 \leq l \leq m \leq 1} dldm\end{aligned}$$

3) Bestimme die gemeinsame Dicht und Verteilungsfunktion von (M, L) . Wir bestimmen $\phi(x) = \mathbb{I}_{x \leq a}, \varphi(x) = \mathbb{I}_{x \leq b}$ und erhalten::

$$\begin{aligned}F_{M,L}(a, b) &= \mathbb{P}[M \leq a, L \leq b] = \mathbb{E}[I_{M \leq a} I_{L \leq b}] \\ &= \int_{-\infty}^a \int_{-\infty}^b 6(m-l)\mathbb{I}_{0 \leq l \leq m \leq 1} dmdl\end{aligned}$$

Somit ist $f_{M,L}(m, l) = 6(m-l)\mathbb{I}_{0 \leq l \leq m \leq 1}$ die gemeinsame Dicht. Für die Verteilungsfunktion berechnen wir das obige Integral, hierbei unterscheiden wir verschiedene Fälle:
 $a \leq 0$ oder $b \leq 0$:

$$F_{M,L}(a, b) = 0$$

$a \geq 1$:

$$F_{M,L}(a, b) = F_L(b) = 1 - (1 - b)^3$$

$b \geq 1$:

$$F_{M,L}(a, b) = F_M(a) = a^3$$

$0 \leq a \leq b \leq 1$:

$$F_{M,L}(a, b) = \mathbb{P}[M \leq a, L \leq b] = \mathbb{P}[M \leq a] = a^3$$

$0 \leq b \leq a \leq 1$:

$$\begin{aligned}\mathbb{P}[M \leq a, L \leq b] &= \int_{-\infty}^a \int_{-\infty}^b 6(m-l)\mathbb{I}_{0 \leq l \leq m \leq 1} dmdl \\ &= \int_0^a \int_0^{\min(b, m)} 6(m-l) dldm \\ &= \int_0^a [6ml - 3l^2]_0^{\min(b, m)} dm \\ &= \int_0^b 3m^2 dm + \int_b^a 3b(2m-b) dm \\ &= b^3 + 3ab(a-b)\end{aligned}$$

Aufgabe Wir haben $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$ mit $X_1, \dots, X_n$ 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.

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

Aufgabe Sei $Y \sim \chi_v^2$:

$$Y = \sum_{k=1}^v Z_k^2$$

wobei $Z_1, \dots, Z_v$ uiv. $\mathcal{N}(0, 1)$ sind. Bestimme die Wahrscheinlichkeit:

$$\mathbb{P}\left[\left|\frac{Y}{v} - 1\right| \leq \frac{3}{4}\right]$$

Hierfür verwenden wir die Chebyshev-Ungleichung:

$$\begin{aligned}\mathbb{P}\left[\left|\frac{Y}{v} - 1\right| \leq \frac{3}{4}\right] &= \mathbb{P}\left[\left|\frac{Y - v}{v}\right| \leq \frac{3}{4}\right] = 1 - \mathbb{P}\left[|Y - v| > \frac{3v}{4}\right] \\ &\geq 1 - \frac{2v}{9v^2/16} = 1 - \frac{32}{9v}\end{aligned}$$

Aufgabe Ein Gerät hat eine Lebensdauer $H \sim \text{Exp}(\lambda)$ mit Erwartungswert 60.

1) Nun haben wir zwei Geräte und wollen wissen wie die Verteilung für die Zeit bis zum ersten Defekt verteilt ist. Aus dem Erwartungswert ergibt sich $H_1, H_2 \sim \text{Exp}(1/60)$. Nun ist $T = \min(H_1, H_2)$ mit Verteilungsfunktion:

$$\begin{aligned}F_T(t) &= \mathbb{P}[T \leq t] = \mathbb{P}[\min(H_1, H_2) \leq t] \\ &= 1 - \mathbb{P}[\min(H_1, H_2) > t] = 1 - \mathbb{P}[H_1 > t, H_2 > t] \\ &= 1 - \mathbb{P}[H_1 > t] \cdot \mathbb{P}[H_2 > t] \\ &= 1 - \exp(-2\lambda t)\end{aligned}$$

D.h. $T \sim \text{Exp}(1/30)$.

2) Wir ersetzen beide Geräte wenn eines Defekt ist und wollen wissen wie gross die Wahrscheinlichkeit ist, mehr als 35 Ersatzteile in drei Jahren zu benötigen.

Sei T_i die Zeit, bis das i -te Teil ersetzt wird, $S = T_1 + \dots + T_{36}$. Nun wollen wir mit dem Grenzwertsatz die Wahrscheinlichkeit berechnen.

$$\begin{aligned}\mathbb{P}[S \leq 1095] &= \mathbb{P}\left[\frac{S - n\mathbb{E}[T_i]}{\sqrt{\sigma_T^2 n}} \leq \frac{1095 - n\mathbb{E}[T_i]}{\sqrt{\sigma_T^2 n}}\right] \\ &= \mathbb{P}\left[\frac{S - 36 \cdot 30}{\sqrt{30^2 \cdot 36}} \leq \frac{1095 - 36 \cdot 30}{\sqrt{30^2 \cdot 36}}\right] \\ &= \mathbb{P}\left[\frac{S - 1080}{180} \leq \frac{15}{180}\right] \\ &\approx \Phi(1/12) \approx \Phi(0.08) = 0.5319\end{aligned}$$

Aufgabe Sei $n \geq 1$ und $(X_i)_{1 \leq i \leq n}$ eine Folge uiv. ZV mit $X_i \sim \mathcal{N}(1, 4)$. Wir definieren $S_n = X_1 + \dots + X_n$ und $\bar{X}_n = S_n/n$. Berechne die Wahrscheinlichkeit $\mathbb{P}[\mathbb{E}[X_1] - 1 \leq X_1 \leq \mathbb{E}[X_1] + 1]$.

Durch Standardisieren erhalten wir $Z = (X_1 - 1)/2 \sim \mathcal{N}(0, 1)$.

$$\begin{aligned}\mathbb{P}[\mathbb{E}[X_1] - 1 \leq X_1 \leq \mathbb{E}[X_1] + 1] &= \mathbb{P}[0 \leq X_1 \leq 2] = \mathbb{P}[-1/2 \leq Z \leq 1/2] \\ &= \Phi(1/2) - \Phi(-1/2) \\ &= \Phi(1/2) - (1 - \Phi(1/2)) = 2\Phi(1/2) - 1\end{aligned}$$

Aufgabe Um die Anzahl Fische N in einem See zu bestimmen gehen wir wie folgt vor, zuerst werden 500 Fische gefangen und markiert. Danach werden wieder 200 Fische gefangen und die Anzahl X der markierten Fische gezählt.

1) $X \sim \text{Bin}(n, \theta)$, wie gross ist n ? Wie gross ist θ , wenn die Gesamtzahl der Fische $N = 2000$ ist?

$n = 200$, da wir 200 Fische herausziehen. $\theta = \frac{500}{N} = \frac{500}{2000} = \frac{1}{4}$

2) Die Beobachtung gibt einen Wert für X von 40. Gebe eine Schätzung für θ und eine Schätzung für N ab.

Wir schätzen θ mit $T = X/n$, der realisierte Schätzwert ist also $\theta = 1/5$. Wenn wir nun $\theta = 500/N$ nach N auflösen erhalten wir $N = 2500$.

3) Bestimme ein approximatives Konfidenzintervall für θ mit $\alpha = 0.05$.

$$T = \frac{X - n\theta}{\sqrt{n\theta(1-\theta)}} \sim \mathcal{N}(0, 1)$$

Aus dem zentralen Grenzwertsatz folgt daher $\mathbb{P}_\theta[-1.96 \leq T \leq 1.96] \geq 0.95$. Unter verwendung von $\theta(1 - \theta) \leq 1/4$ ergibt sich ein approximatives Vertrauensintervall für θ von:

$$[T - \frac{1.96}{2\sqrt{n}}, T + \frac{1.96}{2\sqrt{n}}] = [0.13, 0.27]$$

Aufgabe Wir haben eine Münze und vermuten, dass Sie gezinkt ist und häufiger auf Kopf landet. Wir bezeichnen den i -ten Wurf mit X_i und das Resultat ist 1 wenn Kopf und 0 wenn Zahl. Wir beobachten folgende Ergebnisse:

$$[0, 1, 1, 1, 1, 1, 0, 1, 1, 1]$$

Führe einen Test mit $\alpha = 0.01$ durch.

1) Modell: Wir wählen X_i uiv. Ber(θ) unter $\mathbb{P}_\theta$ wobei $\theta \in [0, 1]$.

2) Nullhypothese und Alternativhypothese:

$$H_0 : \theta = \frac{1}{2} \qquad H_A : \theta > \theta_0$$

3) Teststatistik:

Für $X \sim \text{Ber}(\theta)$ ist die Gewichtsfunktion $p_X(k) = p^k(1 - p)^{1-k}$. Somit ist der Likelihood-Quotient:

$$\begin{aligned} R(x_1, ..., x_10; \theta_A, \theta_0) &= \frac{L(x_1, ..., x_n; \theta_A)}{L(x_1, ..., x_n; \theta_0)} \\ &= \left(\underbrace{\frac{\theta_A}{\theta_0}}_{>1}\right)^{\sum x_i} \left(\underbrace{\frac{1 - \theta_A}{1 - \theta_0}}_{<1}\right)^{\sum x_i} \end{aligned}$$

Somit ist R genau dann gross, wenn $\sum_{i=1}^{10} X_i$ gross ist. Daher wählen wir die Teststatistik:

$$T = \sum_{i=1}^{10} X_i$$

4) Verteilung der Teststatistik unter H_0 :

Da $X_i \sim \text{Ber}(\theta)$ ist $T \sim \text{Bin}(10, \theta)$ unter $\mathbb{P}_\theta$.

5) Verwerfungsbereich:

Wir wählen $K = (c, 10]$. Um c zu bestimmen rechnen wir:

$$\mathbb{P}_\theta[T \in K] \leq \alpha \Rightarrow \mathbb{P}_\theta[T \leq c] \geq 1 - \alpha$$

Mit der gegebenen Tabelle sehen wir, dass $c = 9$ der kleinste Wert für c ist, welcher die Bedingung erfüllt. Somit ist $K = (9, 10] = \{10\}$.

6) beobachte Wert der Teststatistik: $T(\omega) = 8$

7) Testentscheid: $T(\omega) \notin K$, wir verwerfen die Nullhypothese nicht.

Aufgabe Wir stellen fest, dass im Durchschnitt $\approx 1.1\%$ der Ware kaputt ist. Eigentlich sollte aber höchstens 1.0% der Ware defekt sein. Daher überprüfen wir 16 Pakete an 1000 Produkten je. Die dazugehörigen empirischen Werte sind $\bar{x}_{16} = 11, s_{16} = 6$. Wir nehmen an, dass die Daten näherungsweise aus einer $\mathcal{N}(\mu, \sigma^2)$ -verteilten Stichprobe stammen.

1) Lässt sich mit Signifikanzniveau 10% zeigen, dass mehr als 1% der Ware defekt ist?

Wir führen einen nach oben einseitigen t-Test durch, der Verwerfungsbereich entspricht:

$$K = [t_{n-1, 1-\alpha}, \infty) = [1.341, \infty)$$

Wir haben $H_0 : \mu = \mu_0 = 10$ und $H_A : \mu > \mu_0$. Die Realisierung der Teststatistik entspricht:

$$t = \sqrt{16} \frac{\bar{x}_{16} - \mu_0}{s_{16}} = 1.33 \notin K$$

Somit wird H_0 nicht verworfen.

2) Welcher Test kann man durchführen, wenn $\sigma = 3$ gegeben ist?

Wir führen einen nach oben einseitigen z-Test durch, der Verwerfungsbereich entspricht:

$$K = [z_{1-\alpha}, \infty) = [\Phi^{-1}(1 - \alpha), \infty) = [1.29, \infty)$$

Die Hypothese bleibt und die Realisierte Teststatistik ist:

$$z = \sqrt{16} \frac{\bar{x}_{16} - \mu_0}{\sigma} = 1.33 \in K$$

Somit wird H_0 verworfen.

Aufgabe Angenommen wir haben 16 unabhängige Messungen X_i mit Mittelwert $\bar{x} = 204.2$. Wir wollen nun wissen ob der Grenzwert von 200 mit Niveau 5% überschritten wurde.

1) Bestimme einen geeigneten Test und führe diesen durch.

Wir verwenden einen einseitigen z-Test, mit $H_0 : \mu = \mu_0 = 200$, $H_A : \mu > \mu_0$ und Verwerfungsbereich:

$$K = [z_{1-\alpha}, \infty) = [\Phi^{-1}(1 - \alpha), \infty) = [1.64, \infty)$$

Die realisierte Teststatistik ist:

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{204.2 - 200}{10/4} = 1.68 \in K$$

Somit wird die Nullhypothese verworfen.

2) Wie wahrscheinlich ist es, dass man mit 16 unabhängigen Messungen eine Grenzwertüberschreitung nachweisen kann, wenn der wahre Wert tatsächlich über dem Grenzwert bei 205 lag?

Die Nullhypothese kann verworfen werden, falls $\bar{X} > 204.1$. Somit haben wir $(\mu_A = 205, \sigma = 10)$:

$$\begin{aligned} \mathbb{P}[\bar{X} > 204.1] &= \mathbb{P}\left[\frac{\bar{X} - \mu_A}{\sigma/\sqrt{n}} > \frac{204.1 - \mu_A}{\sigma/\sqrt{n}}\right] \\ &= \mathbb{P}\left[\frac{\bar{X} - \mu_A}{\sigma/\sqrt{n}} > -0.36\right] \\ &= \mathbb{P}[Z > -0.36] = \mathbb{P}[Z \leq 0.36] = 0.6406 \end{aligned}$$

Im letzten Schritt verwenden wir die Symmetrie der Normalverteilung.

3) Wie wahrscheinlich ist es, dass man mit 16 unabhängigen Messungen fälschlicherweise eine Grenzwertüberschreitung nachweist, obwohl der wahre Wert bei 200 lag?

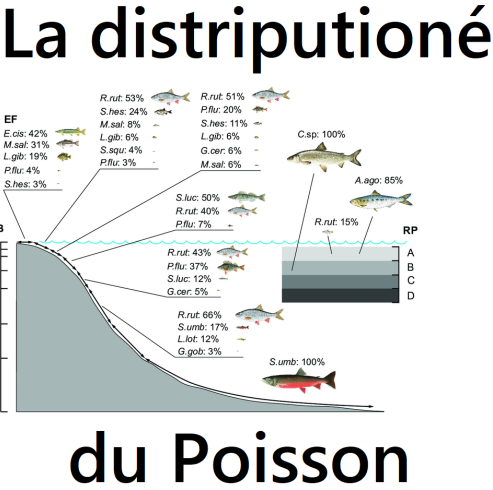
5% , dies entspricht dem Niveau des Tests.

Aufgabe Sei H_0 eine einfache Nullhypothese und T die Teststatistik einer stetigen Zufallsvariable mit Dicht f_T unter $\mathbb{P}_\theta$. Betrachten wir die geordnete Familie von Tests $(T, K_t)_{t \geq 0}$ wobei $K = (t, \infty)$. Zeige, dass der p -Wert $G(T)$ gleichverteilt auf $[0, 1]$ ist.

Wir wollen zeigen, dass $\mathbb{P}_{\theta_0}[G(T) \leq a] = a \cdot \mathbb{I}_{a \in [0, 1]}$. Da $K_t = (t, \infty)$ ist die Funktion G gegeben durch $G(t) = \mathbb{P}_{\theta_0}[T > t] = \mathbb{P}_{\theta_0}[T \leq 1]$ monoton fallend und stetig. Ausserdem nimmt G Werte in $(0, 1)$ an. Wir betrachten die Inverse von G , bezeichnet mit H :

$$\mathbb{P}_{\theta_0}[G(T) \leq a] = \mathbb{P}_{\theta_0}[H(G(T)) \geq H(a)] = \mathbb{P}_{\theta_0}[T > H(a)] = G(H(a)) = a$$

Im ersten Schritt dreht sich die Ungleichung, da H ebenfalls monoton fallend ist.



Diskrete Verteilungen

Verteilung	Parameter	$\mathbb{E}[X]$	$\text{Var}(X)$	$p_X(t)$	$F_X(t)$
Gleichverteilung	n : Anzahl Ereignisse x_i : Ereignisse	$\frac{1}{n} \sum_{i=1}^n x_i$	$\frac{1}{n} \sum_{i=1}^n x_i^2 - \frac{1}{n^2} (\sum_{i=1}^n x_i)^2$	$\frac{1}{n}$	$\frac{ \{k: x_k \leq t\} }{n}$
Bernoulli	p : ErfolgsWK	p	$p \cdot (1 - p)$	$p^t (1 - p)^{1-t}$	$1 - p$ für $0 \leq t < 1$
Binomial	p : ErfolgsWK n : Anzahl Versuche	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	$\frac{\lambda^t}{t!} e^{-\lambda}$	$e^{-\lambda} \sum_{k=0}^t \frac{\lambda^k}{k!}$

Stetige Verteilungen

Verteilung	Parameter	$\mathbb{E}[X]$	$\text{Var}(X)$	$f_X(t)$	$F_X(t)$
Gleichverteilung	$[a, b]$: Intervall	$\frac{a+b}{2}$	$\frac{1}{12}(b - a)^2$	$\frac{1}{b-a}$	$\frac{t-a}{b-a}$
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	$\frac{1}{\sqrt{2\pi}\sigma^2} e^{-\frac{(t-\mu)^2}{2\sigma^2}} \quad -\infty < t < \infty$	$\frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^t e^{-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2} dy$
χ^2 -Verteilung	n : Freiheitsgrad	n	$2n$	$\frac{1}{2^{\frac{n}{2}} \Gamma(\frac{n}{2})} t^{\frac{n}{2}-1} e^{-\frac{t}{2}} \quad t > 0$	$\text{Gamma}(\frac{n}{2}, \frac{t}{2})$
t-Verteilung	n : Freiheitsgrad	$\begin{cases} 0 & n > 1 \\ \text{undef.} & \text{sonst} \end{cases}$	$\begin{cases} \frac{n}{n-2} & n > 2 \\ \infty & 1 < n \leq 2 \\ \text{undef.} & \text{sonst} \end{cases}$	$\frac{\Gamma(\frac{n+1}{2})}{\sqrt{n\pi} \cdot \Gamma(\frac{n}{2})} \left(1 + \frac{t^2}{n}\right)^{-\frac{n+1}{2}}$	oof