

Vorwissen

Def Geometrische Reihe

$$s_n = a_1 \cdot \frac{1 - q^n}{1 - q} \quad \lim_{n \rightarrow \infty} s_n = \frac{a_1}{1 - q}$$

Wobei $a_i = a_1 \cdot q^{i-1}$

Def Binomialkoeffizient

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

Method Partielle Integration

$a < b$, $f, g : [a, b] \rightarrow \mathbb{R}$ stetig diff.-bar

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

$$\int_a^b f(x)g'(x) \, dx = f(b)g(b) - f(a)g(a) - \int_a^b f'(x)g(x) \, dx$$

Method Substitution

$a < b$, $\phi : [a, b] \rightarrow \mathbb{R}$ stetig diff.-bar,

$I \subset \mathbb{R}$, $\phi([a, b]) \subset I$, $f : I \rightarrow \mathbb{R}$ stetig

$$\int_{\phi(a)}^{\phi(b)} f(x) \, dx = \int_a^b f(\phi(t)) \cdot \phi'(t) \, dt$$

1 Wahrscheinlichkeitsräume

Def Grundraum Ω **Elementarereignis** $\omega \in \Omega$

Def σ -Algebra $\mathcal{F} \subseteq \mathcal{P}(\Omega)$ **Ereignis** $A \in \mathcal{F}$

- (i) $\Omega \in \mathcal{F}$
- (ii) $A \in \mathcal{F} \implies A^c \in \mathcal{F}$
- (iii) $A_1, \dots, A_n \in \mathcal{F} \implies \bigcup_{i \leq n} A_i \in \mathcal{F}$

Lem. Abgeschlossenheit der σ -Algebra $\mathcal{F}$

- (i) $\emptyset \in \mathcal{F}$
- (ii) $A_1, \dots, A_n \in \mathcal{F} \implies \bigcap_{i=1}^{\infty} A_i \in \mathcal{F}$
- (iii) $A, B \in \mathcal{F} \implies A \cup B \in \mathcal{F}$
- (iv) $A, B \in \mathcal{F} \implies A \cap B \in \mathcal{F}$

Def Wahrscheinlichkeitsmass auf $(\Omega, \mathcal{F}) : \mathbb{P}$

$$\mathbb{P} : \mathcal{F} \rightarrow [0, 1] \quad \text{s.d.} \quad A \mapsto \mathbb{P}[A]$$

- (i) $\mathbb{P}[\Omega] = 1$
- (ii) $\mathbb{P}[A] = \sum_{i=1}^{\infty} \mathbb{P}[A_i] \iff A = \bigcup_{i=1}^{\infty} A_i \quad \text{s.d.} \quad \bigcap_{i=1}^{\infty} A_i = \emptyset$

Lem. Eigenschaften von $\mathbb{P}$

- (i) $\mathbb{P}[\emptyset] = 0$
- (ii) $\bigcap_{i=1}^k A_i = \emptyset \implies \mathbb{P}\left[\bigcup_{i=1}^k A_i\right] = \sum_{i=1}^k \mathbb{P}[A_i]$
- (iii) $\mathbb{P}[A^c] = 1 - \mathbb{P}[A]$
- (iv) $\mathbb{P}[A \cup B] = \mathbb{P}[A] + \mathbb{P}[B] - \mathbb{P}[A \cap B]$

Def Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \mathbb{P})$ $A \in \mathcal{F}$, $\omega \in \Omega$

- A tritt ein $\stackrel{\text{def}}{\iff} \omega \in A$ (i) $\emptyset$ tritt nie ein
- A tritt nicht ein $\stackrel{\text{def}}{\iff} \omega \notin A$ (ii) Ω tritt immer ein

Def Laplace Modell $(\Omega, \mathcal{F}, \mathbb{P})$ Ω endlich.

- (i) $\mathcal{F} = \mathcal{P}(\Omega)$
- (ii) $\forall A \in \mathcal{F} : \mathbb{P}[A] = \frac{|A|}{|\Omega|}$

Lem. Nützliche Ungleichungen

- (i) $A \subseteq B \implies \mathbb{P}[A] \leq \mathbb{P}[B]$ (Monotonie)
- (ii) $\mathbb{P}\left[\bigcup_{i=1}^{\infty} A_i\right] \leq \sum_{i=1}^{\infty} \mathbb{P}[A_i]$ (Union Bound)

$A_1, A_2, \dots$ müssen *nicht* disjunkt sein.

Lem. Stetigkeit von $\mathbb{P}$ gegen ∞

- (i) $\forall n : A_n \subseteq A_{n+1} \implies \lim_{n \rightarrow \infty} \mathbb{P}[A_n] = \mathbb{P}\left[\bigcup_{n=1}^{\infty} A_n\right]$
- (ii) $\forall n : B_n \supseteq B_{n+1} \implies \lim_{n \rightarrow \infty} \mathbb{P}[B_n] = \mathbb{P}\left[\bigcap_{n=1}^{\infty} B_n\right]$

$(A_n), (B_n)$ sind monotone Folgen von Ereignissen

1.1 Bedingte Wahrscheinlichkeit

Def Bedingte Wahrscheinlichkeit

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

$A, B \in \mathcal{F}$, $\mathbb{P}[B] > 0$

Lem. $\mathbb{P}[A \mid A] = 1$ $\mathbb{P}[A] > 0$

Lem. Totale Wahrscheinlichkeit

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

$B_1, \dots, B_n$ sind eine Partition von Ω , $\mathbb{P}[B_i] > 0$.

Lem. Bayes

$$\forall i = 1, \dots, n : \mathbb{P}[B_i \mid A] = \frac{\mathbb{P}[A \mid B_i] \cdot \mathbb{P}[B_i]}{\sum_{j=1}^n \mathbb{P}[A \mid B_j] \cdot \mathbb{P}[B_j]}$$

$B_1, \dots, B_n$ sind eine Partition von Ω , $\mathbb{P}[B_i] > 0$, $\mathbb{P}[A] > 0$.

1.2 Unabhängigkeit

Def Unabhängigkeit

$$A, B \text{ unabhängig} \stackrel{\text{def}}{\iff} \mathbb{P}[A \cap B] = \mathbb{P}[A] \cdot \mathbb{P}[B]$$

$$\text{Notation } A \perp B \stackrel{\text{def}}{\iff} A, B \text{ unabhängig}$$

Lem. Äquivalente Aussagen zur Unabhängigkeit

- (i) $\mathbb{P}[A \cap B] = \mathbb{P}[A] \cdot \mathbb{P}[B]$ (Defintion)
- (ii) $\mathbb{P}[A|B] = \mathbb{P}[A]$ (B kein Einfluss auf A)
- (iii) $\mathbb{P}[B|A] = \mathbb{P}[B]$ (A kein Einfluss auf B)

$$A, B \in \mathcal{F}, \quad \mathbb{P}[A], \mathbb{P}[B] > 0$$

Def Unabhängigkeit für Ereignismengen

$$(A_i)_{i \in I} \text{ unabhängig} \stackrel{\text{def}}{\iff} \forall J \subseteq I : \mathbb{P}\left[\bigcap_{j \in J} A_j\right] = \prod_{j \in J} \mathbb{P}[A_j]$$

I ist eine Indexmenge. Dies muss für *alle* $J \subseteq I$ (endlich) gelten.

2 Zufallsvariablen

Def Zufallsvariable

$$X : \Omega \rightarrow \mathbb{R} \quad \text{s.d.} \quad \forall a \in \mathbb{R} : \left\{ \omega \in \Omega \mid X(\omega) \leq a \right\} \in \mathcal{F}$$

$(\Omega, \mathcal{F}, \mathbb{P})$ ein Wahrscheinlichkeitsraum

Def Indikatorfunktion

$$\forall \omega \in \Omega : \quad \mathbb{I}_A(\omega) := \begin{cases} 0 & \omega \notin A \\ 1 & \omega \in A \end{cases} \quad A \in \mathcal{F}$$

Notation Ereignisse bezüglich Zufallsvariablen

$$\begin{aligned} \{X \leq a\} &= \left\{ \omega \in \Omega \mid X(\omega) \leq a \right\} \\ \{a < X \leq b\} &= \left\{ \omega \in \Omega \mid a < X(\omega) \leq b \right\} \\ \mathbb{P}[X \leq a] &= \mathbb{P}\left[\{X \leq a\}\right] \end{aligned}$$

Def Verteilungsfunktion $F_X : \mathbb{R} \rightarrow [0, 1]$

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

$$\text{Notation } \lim_{x \rightarrow a^-} F_X(a) = F(a^-) = \mathbb{P}[X < a]$$

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

Th. Eigenschaftern der Verteilungsfunktion

- (i) F_X monoton
- (ii) F_X rechtsstetig
- (iii) $\lim_{a \rightarrow -\infty} F_X(a) = 0 \quad \wedge \quad \lim_{a \rightarrow \infty} F_X(a) = 1$

Beispiel: Zeige, dass F eine Verteilungsfunktion ist:

$$F(t) = \begin{cases} 0 & t \leq 0 \\ 1 - \exp(-\frac{t}{4}) & t > 0 \end{cases}$$

- (i) F ist monoton wachsend, da $F'(t) = \frac{1}{4} \exp(-\frac{t}{4}) > 0 \forall t \in (-\infty, 0]$.
- (ii) F ist rechtsstetig, da F eine Komp. stetiger Funktionen ist.
- (iii) Es gilt $\lim_{t \rightarrow -\infty} F(t) = 0$ und $\lim_{t \rightarrow \infty} F(t) = 1$

Def Unabhängigkeit

$$X_1, \dots, X_n \text{ unabhängig} \stackrel{\text{def}}{\iff} \forall x_1, \dots, x_n \in \mathbb{R} :$$

$$\mathbb{P}[X_1 \leq x_1, \dots, X_n \leq x_n] = \mathbb{P}[X_1 \leq x_1] \cdots \mathbb{P}[X_n \leq x_n]$$

Th. Unabhängigkeit von Gruppierungen

$X_1, \dots, X_n$ sind unabhängig, dann sind auch $Y_1, \dots, Y_k$ unabhängig:

$$Y_1 = \phi_1(X_1, \dots, X_{i_1}), \dots, Y_k = \phi_k(X_{i_{k-1}+1}, \dots, X_{i_k})$$

$1 \leq i_1 < i_2 < \dots < i_k \leq n$ sind Indizes, $\phi_1, \dots, \phi_k$ sind Abbildungen

Def Folgen von Zufallsvariablen

- (i) unabhängig $\stackrel{\text{def}}{\iff} \forall n \geq 1 : X_1, \dots, X_n$ unabhängig
- (ii) i.i.d $\stackrel{\text{def}}{\iff}$ unabhängig, und $\forall i, j : F_{X_i} = F_{X_j}$

i.i.d = Independent & identically distributed

2.1 Diskrete Zufallsvariablen

Def Diskrete Zufallsvariable

$$X : \Omega \rightarrow \mathbb{R} \text{ diskret} \stackrel{\text{def}}{\iff} \exists (W \subset \mathbb{R}) \preceq \mathbb{N} : \quad \mathbb{P}[X \in W] = 1$$

bzw. die Werte von X liegen f.s. in W

Lem. Variablen diskreter Grundräume sind diskret

$$\Omega \preceq \mathbb{N} \implies X : \Omega \rightarrow \mathbb{R} \text{ ist diskret}$$

Def Verteilung diskreter Variablen

$$\left(p(x) \right)_{x \in W} \text{ s.d. } p(x) := \mathbb{P}[X = x] \text{ heisst Verteilung}$$

X ist diskret mit $W \preceq \mathbb{N}$ $p(x)$ ist *nicht* F_X

$$\text{Th. } \forall p(x) : \quad \sum_{x \in W} p(x) = 1 \quad p(x) \text{ ist eine diskrete Verteilung}$$

Lem. Zur diskreten Verteilung existiert eine Variable

$$\forall \left(p(x) \right)_{x \in W} \in [0, 1] : \quad \exists (\Omega, \mathcal{F}, \mathbb{P}), X \text{ mit Verteilung } p(x)$$

D.h man kann sagen: "Sei X eine Variable mit Verteilung $(p(x))_{x \in W}$ "

Th. Diskrete Verteilungsfunktion

$$F_X(x) = \mathbb{P}[X \leq x] = \sum_{y \leq x} p(y) \quad y \in W$$

2.2 Diskrete Verteilungen

Def Bernoulli-Verteilung $X \sim \text{Ber}(p)$

Intuitiv: Münzwurf

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

$$0 \leq p \leq 1, \quad W = \{0, 1\}$$

Def Binomial-Verteilung $X \sim \text{Bin}(n, p)$

Intuitiv: Anzahl Erfolge bei wiederholtem Bernoulli-Experiment

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

$$0 \leq p \leq 1, \quad n \in \mathbb{N}, \quad W = \{0, \dots, n\}$$

Th. Bernoulli-Summen sind Binomial-verteilt

$$S_n := X_1 + \dots + X_n \sim \text{Bin}(p, n)$$

$$0 \leq p \leq 1, \quad n \in \mathbb{N}, \quad X_1, \dots, X_n \sim \text{Ber}(p) \text{ unabhängig}$$

Lem. $\text{Bin}(1, p) = \text{Ber}(p)$

Lem. $X_1 \sim \text{Bin}(n, p), X_2 \sim \text{Bin}(m, p) \implies X_1 + X_2 \sim \text{Bin}(n + m, p)$

Def Geometrische Verteilung $X \sim \text{Geom}(p)$

Intuitiv: Bernoulli Experiment erfolgreich beim k -ten Versuch

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

$$0 < p \leq 1, \quad k \in \mathbb{N}_{\neq 0}$$

Def Negativbinomiale Verteilung $X \sim \text{NBin}(r, p)$

Intuitiv: Wartezeit zum r -ten Erfolg

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

$$r \in \mathbb{N}, \quad p \in [0, 1], \quad k \in \{r, r+1, \dots\}, \quad \text{Nur auf Folien}$$

Th. Interpretation

Für $X_1, X_2, \dots$ unabhängig s.d. $X_i \sim \text{Ber}(p)$:

$$T_r := \inf \left\{ n \geq 1 \mid \sum_{i=1}^n X_i = r \right\} \sim \text{NBin}(r, p)$$

Lem. $X_1, X_2, \dots, X_r \sim \text{Geom}(p) \implies X := \sum_{i=1}^r X_i \sim \text{NBin}(r, p)$

Def Hypergeometrische Verteilung $X \sim \text{H}(n, r, m)$

$$\mathbb{P}[X = k] = \frac{\binom{r}{k} \binom{n-r}{m-k}}{\binom{n}{m}}$$

$$n \in \mathbb{N}, \quad m, r \in \{0, 1, \dots, n\}, \quad k \in \{0, 1, \dots, \min(m, r)\}$$

Th. Interpretation

n Objekte in einer Urne, r von Typ 1, $n - r$ von Typ 2

Ziehe m , Sei X die Anzahl Objekte Typ 1, dann $X \sim \text{H}(n, r, m)$

Das modelliert z.B. Lotto

Def Poisson-Verteilung

Intuitiv: Sehr seltene Ereignisse

$$\mathbb{P}[X = k] = \frac{\lambda^k}{k!} e^{-\lambda}$$

$$\lambda \in \mathbb{R}_{>0}, \quad k \in \mathbb{N}_0$$

Th. Poisson-Approximation der Binomialverteilung

$$X_n \sim \text{Bin}(\frac{\lambda}{n}), \quad N \sim \text{Poisson}(\lambda), \quad n \geq 1$$

$$\lim_{n \rightarrow \infty} \mathbb{P}[X_n = k] = \mathbb{P}[N = k]$$

Intuitiv: Für grosse n haben X_n und N sehr ähnliche Eigenschaften

Remark Auch genannt: *Konvergenz in Verteilung, Schwache Konvergenz*
Convergence in distribution/law, weak convergence

2.3 Stetige Verteilungen

Def Stetig verteilte Zufallsvariable $X : \Omega \rightarrow \mathbb{R}$

Intuitiv: $f_X(t)dt$ ist Die W.-keit, dass X in $[t, t + dt]$ ist

$$\exists f_X \mathbb{R} \rightarrow \mathbb{R}_+ \quad \text{s.d.} \quad F_X(x) = \int_{-\infty}^x f_X(t) dt$$

Stetig, da F_X somit eine stetige Funktion ist

Remark f_X heisst *Dichtefunktion* von X

Th. Dichtefunktion ist die Ableitung der Verteilung

F_X stetig, stückweise diff.-bar $-\infty = x_0 < x_1 < \dots < x_{n-1} < x_n = \infty$

$$f_X = \begin{cases} F'_X(x) & \exists k \in \{0, \dots, n-1\} \text{ s.d. } x \in (x_k, x_{k+1}) \\ a_k & x \in \{x_1, \dots, x_{n-1}\} \end{cases}$$

Remark f_X (Stetig) ist Intuitiv das Analogon zu p_x (Diskret)

Def Gleichverteilung $X \sim \mathcal{U}([a, b])$

Intuitiv: Zufälliger Punkt in $[a, b]$

$$f_X(x) = \frac{1}{b-a} \cdot \mathbb{I}_{\{a \leq x \leq b\}} = \begin{cases} \frac{1}{b-a} & x \in [a, b] \\ 0 & \text{sonst} \end{cases}$$

Lem. Intervalle $\mathbb{P}[X \in [c, c+l]] = \frac{l}{b-a}$

$X \sim \mathcal{U}([a, b])$

Def Exponentialverteilung $T \sim \text{Exp}(\lambda)$

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

$$\lambda > 0$$

Lem. Eigenschaften von $T \sim \text{Exp}(\lambda)$

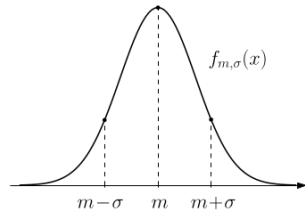
$$1. \quad \forall t \geq 0 : \quad \mathbb{P}[T > t] = e^{-\lambda t}$$

$$2. \quad \forall t, s \geq 0 : \quad \mathbb{P}[T > t + s \mid T > t] = \mathbb{P}[T > s]$$

(2) wird *Gedächtnislosigkeit* genannt.

Def Normalverteilung $X \sim \mathcal{N}(m, \sigma^2)$

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



Lem. Summe von Normalverteilten ist normalverteilt

$X_1, \dots, X_n$ unabh. s.d. $X_i \sim \mathcal{N}(m_i, \sigma_i^2)$

$$Z = m_0 + \lambda_1 X_1 + \dots + \lambda_n X_n$$

$$Z \sim \mathcal{N}\left(\underbrace{m_0 + \sum_{i=1}^n \lambda_i X_i}_{m_Z}, \underbrace{\sum_{i=1}^n \lambda_i^2 \sigma_i^2}_{\sigma_Z^2}\right)$$

Def Standardnormalverteilung

$$X \sim \mathcal{N}(0, 1) \quad \Phi(x) := \mathbb{P}[X \leq x]$$

Lem. Symmetrie $\Phi(-x) = 1 - \Phi(x)$

Lem. Normalverteilung via $\mathcal{N}(0, 1)$

$Z \sim \mathcal{N}(m, \sigma^2)$ arbiträr, $X \sim \mathcal{N}(0, 1)$

$$Z = m + \sigma \cdot X$$

Lem. Standardisierung

$Z \sim \mathcal{N}(m, \rho^2)$, $x \in \mathbb{R}$

$$\mathbb{P}[x \leq Z] = \Phi\left(\frac{x-m}{\rho}\right)$$

3 Gemeinsame Verteilungen

3.1 Gemeinsame Diskrete Verteilung

Def Gemeinsame diskrete Verteilung

$X_1, \dots, X_n$ diskret, $W_i \preceq \mathbb{N}$, $X_i \in W_i$ fast sicher

$$p := \left(p(x_1, \dots, x_n)\right)_{x_1 \in W_1, \dots, x_n \in W_n}$$

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

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

Th. Randverteilung

Die einzelnen Verteilungen p_X lassen sich extrahieren:

$\forall z \in W_i$:

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

$X_1, \dots, X_n$ diskret, gem. Vert. $p = \left(p(x_1, \dots, x_n)\right)_{x_1 \in W_1, \dots, x_n \in W_n}$

Remark Nicht umgekehrt: Aus den Randverteilungen lässt sich nichts über die gem. Vert. schließen.

3.2 Gemeinsame Stetige Verteilung

Def Gemeinsame Stetige Verteilung

$X_1, \dots, X_n : \Omega \rightarrow \mathbb{R}$, $f : \mathbb{R}^n \rightarrow \mathbb{R}_+$, $a_1, \dots, a_n \in \mathbb{R}$

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

f heißt *gemeinsame Dichte* von $(X_1, \dots, X_n)$

Remark Seien $X_1, \dots, X_n \stackrel{\text{i.i.d.}}{\sim} \mathcal{X}$:

$$\mathbb{P}[X_1 \leq x, \dots, X_n \leq x] = \prod_{i=1}^n \mathbb{P}[X_i \leq x] = \left(\mathbb{P}[X \leq x]\right)^n$$

Th.
$$\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f(x_1, \dots, x_n) dx_n \dots dx_1 = 1$$

Umgekehrt existiert für jedes solches f ein Raum $(\Omega, \mathcal{F}, \mathbb{P})$

Beispiel: Finde c s.d. $f_{X,Y}$ eine Dichte ist, wobei $f_{X,Y} = \begin{cases} ce^{-x} & 0 < x < y \\ 0 & \text{sonst} \end{cases}$

$$\begin{aligned} \int_0^\infty \int_0^x ce^{-x} dy dx &= c \cdot \int_0^\infty xe^{-x} dx \\ &= c \cdot \left([-xe^{-x}]_0^\infty + \int_0^\infty e^{-x} dx \right) \\ &= c \cdot \left(0 + [-e^{-x}]_0^\infty \right) \\ &= c \end{aligned}$$

Also gilt wegen dem vorherigen Satz: $c = 1$.

Remark Randdichten: Integration über übrige Variablen.

Th. Erwartungswert

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

Th. Unabhängigkeit

$$X_1, \dots, X_n \text{ stetig, } f(x_1, \dots, x_n) = f_1(x_1) \dots f_n(x_n)$$

Remark Unabhängige stetige Variablen sind automatisch gemeinsam stetig

4 Charakterisierung

4.1 Erwartungswert

Def Allg. Erwartungswert (nicht-negativ)

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

$X : \Omega \rightarrow \mathbb{R}, \quad X(\omega) \geq 0 \quad \forall \omega \in \Omega$

Remark $\mathbb{E}[X]$ kann unendlich sein

Th. $\forall \omega \in \Omega : X(\omega) \geq 0 \implies \mathbb{E}[X] \geq 0$

Gleichheit: $\mathbb{E}[X] = 0 \iff X = 0$, fast sicher

Th. Diskreter Erwartungswert

$X : \Omega \rightarrow \mathbb{R}, \quad W \preceq \mathbb{N}, \quad \phi : \mathbb{R} \rightarrow \mathbb{R}$

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

Def Erwartungswert (stetig)

$X : \Omega \rightarrow \mathbb{R}, \quad f(x)$ Dichtefunktion

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

Th. Eigenschaften

$X, Y : \Omega \rightarrow \mathbb{R}, \quad \lambda \in \mathbb{R}$

- (i) $\mathbb{E}[\lambda X] = \lambda \mathbb{E}[X]$
- (ii) $\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y]$
- (iii) $X \leq Y \implies \mathbb{E}[X] \leq \mathbb{E}[Y]$
- (iv) $X \perp Y \implies \mathbb{E}[X \cdot Y] = \mathbb{E}[X] \cdot \mathbb{E}[Y]$

Th. Dichtefunktion bei Abbildungen

$\phi : \mathbb{R} \rightarrow \mathbb{R}$ stückweise stetig, beschränkt

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

$f : \mathbb{R} \rightarrow \mathbb{R}_+$ s.d. $\int_{-\infty}^{\infty} f(x) dx = 1$

Th. Unabhängigkeit (durch Abbildungen)

$\forall \phi, \psi : \mathbb{R} \rightarrow \mathbb{R}$ stückweise stetig, beschränkt

$$X, Y \text{ unabh.} \iff \mathbb{E}[\phi(X)\psi(Y)] = \mathbb{E}[\phi(X)] \cdot \mathbb{E}[\psi(Y)]$$

Auch verallgemeinert für $X_1, \dots, X_n, \phi_1, \dots, \phi_n$

4.2 Varianz

Def Varianz

$\mathbb{E}[X^2] < \infty$

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

Def Standardabweichung

$$\rho(X) := \sqrt{\mathbb{V}[X]}$$

Notation Auch ρ, ρ_X

Lem. Varianz (Alternativ)

$$\mathbb{V}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$$

Beispiel: Bestimme $\mathbb{E}[X^2]$, wobei $X \sim \text{Poisson}(\lambda)$, $\lambda > 0$.

$$\begin{aligned} \mathbb{V}[X] &= \mathbb{E}[X^2] - \mathbb{E}[X]^2 & (X \sim \text{Poisson}(\lambda)) \\ \lambda &= \mathbb{E}[X^2] - \lambda^2 \\ \lambda + \lambda^2 &= \mathbb{E}[X^2] \end{aligned}$$

Lem. Eigenschaften

- (i) $\mathbb{V}[X] \geq 0$
- (ii) $\mathbb{V}[aX] = a^2 \mathbb{V}[X]$
- (iii) $\mathbb{V}[X + a] = \mathbb{V}[X]$

Lem. Addition (Unabhängigkeit)

$X_1, \dots, X_n$ paarweise unabhängig

$$\mathbb{V}\left[\sum_{k=1}^n X_k\right] = \sum_{k=1}^n \mathbb{V}[X_k]$$

Beispiel: $\mathbb{E}[X^2]$ allgemein.

$$\begin{aligned} \mathbb{E}[X^2] &= \mathbb{E}[\phi(X)] \text{ s.d. } \phi(x) = x^2 \\ \implies \mathbb{E}[X^2] &= \int_{-\infty}^{+\infty} \phi(x) f_X(x) dx = \boxed{\int_{-\infty}^{+\infty} x^2 \cdot f_X(x) dx} \end{aligned}$$

4.3 Kovarianz

Def Kovarianz

X, Y s.d. $\mathbb{E}[X^2], \mathbb{E}[Y^2] < \infty$

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

Lem. Kovarianz (Alternativ)

$$\text{cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$$

Remark $\text{cov}(X, X) = \mathbb{V}[X]$

Lem. Eigenschaften von cov

- (i) $\text{cov}(X, Y) \geq 0$
- (ii) $\text{cov}(X, Y) = \text{cov}(Y, X)$
- (iii) X, Y unabh. $\implies \text{cov}(X, Y) = 0$ Nicht umgekehrt!
- (iv) $\mathbb{V}[X \pm Y] = \mathbb{V}[X] + \mathbb{V}[Y] \pm 2\text{cov}(X, Y)$
- (v) $\text{cov}\left(\sum_{i=1}^n X_i, \sum_{j=1}^n Y_j\right) = \sum_{i=1}^n \sum_{j=1}^n \text{cov}(X_i, Y_j)$
- (vi) $\text{cov}(aX + b, cY + d) = ac \cdot \text{cov}(X, Y)$
- (vii) $\text{cov}\left(X, (eY + f) + (gZ + h)\right) = e\text{cov}(X, Y) + g\text{cov}(X, Z)$

Def Kovarianzmatrix $\mathbf{X} = (X_1, \dots, X_n)^\top$

$$\Sigma = \text{cov}(\mathbf{X}) = \begin{bmatrix} \mathbb{V}[X_1] & \text{cov}(X_1, X_2) & \cdots & \text{cov}(X_1, X_n) \\ \text{cov}(X_2, X_1) & \text{cov}(X_2, X_2) & \cdots & \text{cov}(X_2, X_n) \\ \vdots & \vdots & \ddots & \vdots \\ \text{cov}(X_n, X_1) & \text{cov}(X_n, X_2) & \cdots & \text{cov}(X_n, X_n) \end{bmatrix}$$

Beispiel: $\mathbb{E}[\max(X_1, \dots, X_n)]$ s.d. $X_1, \dots, X_n \stackrel{\text{i.i.d.}}{\sim} \mathcal{X}$

$$\begin{aligned} F_{\max(X_1, \dots, X_n)}(x) &= \mathbb{P}[X_1 \leq x, \dots, X_n \leq x] \\ &= \mathbb{P}[X_1 \leq x] \cdots \mathbb{P}[X_n \leq x] \\ &= (\mathbb{P}[X \leq x])^n = F_X(x)^n \\ \implies f_{\max(X_1, \dots, X_n)}(x) &= \partial_x (F_X(x))^n \\ \implies \mathbb{E}[\max(X_1, \dots, X_n)] &= \boxed{\int_{-\infty}^{+\infty} x \cdot \partial_x (F_X(x))^n} \end{aligned}$$

5 Ungleichungen

Th. Markov

$X \geq 0$, $g : X(\Omega) \rightarrow [0, \infty)$ wachsend

$$\forall c \in \mathbb{R} \text{ s.d. } g(c) > 0 : \quad \mathbb{P}[X \geq c] \leq \frac{\mathbb{E}[g(X)]}{g(c)}$$

Th. Jensen

$\phi : \mathbb{R} \rightarrow \mathbb{R}$ konvex, $\mathbb{E}[\phi(X)], \mathbb{E}[X]$ wohldefiniert

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

Lem. Dreiecksungleichung

Jensen mit $\phi(X) = |X|$ und $\phi(X) = X^2$

$$|\mathbb{E}[X]| \leq \mathbb{E}[|X|] \quad \mathbb{E}[|X|] \leq \sqrt{\mathbb{E}[X^2]}$$

Th. Chebychev

Y s.d. $\mathbb{V}[Y] < \infty$, $c > 0$

$$\mathbb{P}\left[|Y - \mathbb{E}[Y]| \geq c\right] \leq \frac{\mathbb{V}[Y]}{c^2}$$

Th. Chernoff

6 Grenzwertsätze

div

Notation $\bar{X}_n = \frac{1}{n} S_n = \frac{1}{n} \sum_{i=1}^n X_i$

Th. Schwaches Gesetz der grossen Zahlen

$X_1, X_2, \dots$ unabh., $\forall k : \mathbb{E}[X_k] = \mu, \mathbb{V}[X_k] = \sigma^2$

Es konvergiert $\bar{X}_n$ für $n \rightarrow \infty$ gegen $\mu = \mathbb{E}[X_i]$

$$\mathbb{P}\left[|\bar{X}_n - \mu| < \epsilon\right] \xrightarrow{n \rightarrow \infty} 1$$

Remark $\mathbb{V}\left[\frac{1}{n} \sum_{i=1}^n X_i\right] = \frac{1}{n^2} \sum_{i=1}^n \mathbb{V}[X_i] = \frac{n}{n^2} \mathbb{V}[X_i] = \frac{1}{n} \mathbb{V}[X_i]$

Th. Starkes Gesetz der grossen Zahlen

$X_1, X_2, \dots$ unabh., $\forall k : \mathbb{E}[X_k] = \mu$

$$\mathbb{P}\left[\left\{\omega \in \Omega \mid \bar{X}_n \xrightarrow{n \rightarrow \infty} \mu\right\}\right] = 1$$

Def Konvergenz in Verteilung

$(X_n)_{n \in \mathbb{N}}$, X Zufallsvariablen

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

Falls:

$$\forall x \in \mathbb{R} : \quad \lim_{n \rightarrow \infty} \mathbb{P}[X_n \leq x] = \mathbb{P}[X \leq x]$$

Remark Diskrete Zufallsvariablen können zu stetigen konvergieren.

Th. Zentraler Grenzwertsatz

$(X_k)_{k \geq 1}$ i.i.d s.d. $\mathbb{E}[X_k] = \mu, \mathbb{V}[X_k] = \sigma^2, S_n = \sum_{i=1}^n X_i$

$$\lim_{n \rightarrow \infty} \mathbb{P}\left[\frac{S_n - n\mu}{\sigma\sqrt{n}} \leq x\right] = \Phi(x)$$

$\Phi(x)$ ist die Verteilung von $\mathcal{N}(0, 1)$.

$$S_n \overset{\text{approx.}}{\sim} \mathcal{N}(n\mu, n\sigma^2)$$

ZGS wird häufig zur Approximation von Summen verwendet

Def Standardisierung von S_n

$$S_n^* := \frac{S_n - n\mu}{\sigma\sqrt{n}} = \frac{S_n - \mathbb{E}[S_n]}{\sqrt{\mathbb{V}[S_n]}}$$

Im Skript auch Z_n

7 Statistik

Wir betrachten Daten $\{x_i\}_{i=1}^n$ als Realisierung von unbekannten Zufallsvariablen $X_1(\omega), \dots, X_n(\omega)$. Ziel ist es, die Verteilung von $X_1, \dots, X_n$ zu finden durch Modellierung.

Man wählt hierzu eine Familie $(\Omega, \mathcal{F}, \mathbb{P}_\vartheta)_{\vartheta \in \Theta}$ wobei $(\Omega, \mathcal{F})$ fix ist und $\mathbb{P}_\vartheta$ variabel z.B. $\mathbb{P}_\vartheta = \mathcal{N}(\vartheta_1, \vartheta_2)$ (Modell) und versucht ein optimales ϑ bzw. $\mathbb{P}_\vartheta$ zu finden.

Statistisches Vorgehen

1. Datenbeschreibung, intuitives Verständnis von $\{x_i\}_{i=1}^n$
2. Definition von Θ und $(\mathbb{P}_\vartheta)_{\vartheta \in \Theta}$
3. Schätzer nutzen für $(x_1, \dots, x_n) \mapsto \vartheta \in \Theta$
4. Überprüfung von ϑ via statistischem Test
5. Alternativ: Konfidenzbereich über Θ finden

7.1 Schätzer

Def Schätzer (Estimator)

$$T_l = t_l(X_1, \dots, X_n) \mapsto \vartheta_l \in (\vartheta_1, \dots, \vartheta_m)$$

t_l heisst *Schätzfunktion*, eine Auswertung $T_l(\omega)$ heisst *Schätzwert*.
(Estimator function, sample estimate)

Notation $T = (T_1, \dots, T_m), \vartheta = (\vartheta_1, \dots, \vartheta_m)$

Def Erwartungstreu (unbiased)

$$T \text{ Erwartungstreu} \stackrel{\text{def}}{\iff} \forall \vartheta \in \Theta : \mathbb{E}_\vartheta[T] = \vartheta$$

Def Erwarteter Schätzfehler $\mathbb{E}_\vartheta[T] - \vartheta$ (Bias)

Def MSE $\text{MSE}_\vartheta[T] = \mathbb{E}_\vartheta[(T - \vartheta)^2]$ (Mean Sq. Error)

Def Konsistent

$$T^{(n)} \text{ konsistent für } \vartheta \stackrel{\text{def}}{\iff} \lim_{n \rightarrow \infty} \mathbb{P}_\vartheta[|T^{(n)} - \vartheta| > \epsilon] = 0$$

Intuitiv: Bei beliebig hohem n konvergiert $T^{(n)}$ zu ϑ unter $\mathbb{P}_\vartheta$

Lem. Zerlegung des MSE

$$\text{MSE}_\vartheta[T] = \mathbb{E}_\vartheta[(T - \vartheta)^2] = \mathbb{V}_\vartheta[T] + \underbrace{(\mathbb{E}_\vartheta[T] - \vartheta)^2}_{=0 \text{ (Erwartungstreu)}}$$

Remark T erwartungstreu $\implies \text{MSE}_\vartheta[T] = \mathbb{V}_\vartheta[T]$

Beispiel: $T = X_n \sim \text{Ber}(\vartheta)$ ist erwartungstreu, da $\mathbb{E}_\vartheta[T] = \vartheta$ ($T \sim \text{Ber}(\vartheta)$).
Trotzdem ist $T^{(n)}$ nicht konsistent für $\vartheta \in (0, 1)$, da $X_n(\omega) \in \{0, 1\}$.

7.2 Maximum Likelihood Methode (MLE)

Motivation: Wie findet man einen guten Schätzer T ?

Abhängig ob $(X_1, \dots, X_n)$ stetig/diskret:

- Gemeinsame Gewichtsfunktion $p_X(x_1, \dots, x_n; \vartheta)$
- Gemeinsame Dichtefunktion $f_X(x_1, \dots, x_n; \vartheta)$

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

Remark Häufig ist X_k i.i.d. d.h. $p_X(\dots) = \prod_{k=1}^n p_X(x_k; \vartheta)$

Def Likelihood Funktion

$$L(x_1, \dots, x_n; \vartheta) = \begin{cases} p_X(x_1, \dots, x_n; \vartheta), & \text{diskret} \\ f_X(x_1, \dots, x_n; \vartheta), & \text{stetig} \end{cases}$$

Def Maximum Likelihood Schätzer

Maximierung von $\vartheta \mapsto L(X_1, \dots, X_n; \vartheta)$

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

Remark Äquivalent kann man $\log(L)$ maximieren.

Sinnvoll falls X_k i.i.d. weil dann $\log(L) = \sum_{k=1}^n \log(p_X(x_k; \vartheta))$

Anwendung:

1. Gemeinsame Dichte/Verteilung finden
2. $f(\vartheta) := \ln(L(x_1, \dots, x_n; \vartheta))$
3. ϑ^* finden, s.d. $f'(\vartheta^*) = 0$
4. Argumentieren, dass ϑ^* das Maximum ist (z.B. via f'')

7.3 Momentenschätzer

Annahme: $\mathbb{P}_\vartheta$ s.d. $X_i, \dots, X_n$ i.i.d.

Def Momentenschätzer

$$T = (T_1, T_2) \text{ für } \left(\mathbb{E}_\vartheta[X], \mathbb{V}_\vartheta[X] \right)$$

$$T_1 = \frac{1}{n} \sum_{i=1}^n X_i \quad T_2 = \frac{1}{n} \sum_{i=1}^n X_i^2 - \underbrace{\left(\frac{1}{n} \sum_{i=1}^n X_i \right)^2}_{T_1^2}$$

Remark Allgemein nicht erwartungstreu

Def Empirische Varianz (Unbiased Sample Variance)

$$T'_1 = T_1 \quad T'_2 = \frac{n}{n-1} T_2$$

Remark Erwartungstreu

7.4 Verteilungen von Schätzern

Motivation: Die Verteilung eines Schätzers T .

Es gibt wenig exakte Aussagen, daher approximativ via ZGS

Beispiel: $T = \frac{1}{n} \sum_{i=1}^n X_i$
ZGS $\implies T \sim \mathcal{N}(n\mathbb{E}_\vartheta[X_i], n\mathbb{V}_\vartheta[X_i])$ (für grosse n)

Def χ^2 -Verteilung

Parameter m : *Freiheitsgrade*

$$X \sim \chi_m^2 \stackrel{\text{def}}{\iff} f_X(x) = \frac{1}{2^{\frac{m}{2}} \Gamma(\frac{m}{2})} x^{\frac{m}{2}-1} e^{-\frac{x}{2}} \quad (x \geq 0)$$

Def Euler'sche Gammafunktion

$$\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt \quad \forall n \in \mathbb{N}_0 : \Gamma(n+1) = n!$$

Remark Entstehung der χ^2 -Verteilung

$X_1, \dots, X_m$ i.i.d. s.d. $X_i \sim \mathcal{N}(0, 1)$

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

Def Student'sche t -Verteilung

$$X \sim t_m \stackrel{\text{def}}{\iff} f_X(x) = \frac{\Gamma(\frac{m+1}{2})}{\sqrt{m\pi}\Gamma(\frac{m}{2})} \left(1 + \frac{x^2}{m}\right)^{-\frac{m+1}{2}}$$

Remark Entstehung der t -Verteilung

$X \sim \mathcal{N}(0, 1)$ und $Y \sim \chi_m^2$

$$\frac{X}{\sqrt{\frac{1}{m}Y}} \sim t_m$$

Remark $m = 1 \implies t_1$ ist Cauchy Verteilung

Remark $\lim_{m \rightarrow \infty} t_m = \mathcal{N}(0, 1)$

Remark t_m ist symmetrisch um 0 (wie $\mathcal{N}$)

Th. Aussagen für Normalverteilungen

$X_1, \dots, X_n \sim \mathcal{N}(\mu, \sigma^2)$

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

$$1. \bar{X}_n \sim \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right) \text{ also } \frac{\bar{X}_n - \mu}{\frac{\sigma}{\sqrt{n}}} \sim \mathcal{N}(0, 1)$$

$$2. \frac{n-1}{\sigma^2} S^2 = \frac{1}{\sigma^2} \sum_{i=1}^n (X_i - \bar{X}_n)^2 \sim \chi_{n-1}^2$$

$$3. \bar{X}_n \perp S^2$$

$$4. \frac{\bar{X}_n - \mu}{\frac{S}{\sqrt{n}}} = \frac{\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}}}{\frac{S/\sigma}{\sqrt{\frac{1}{n-1} \frac{n-1}{\sigma^2} S^2}}} \sim t_{n-1}$$

7.5 Tests

Motivation: Modelliert $\vartheta \in \Theta$ die Daten $\{x_i\}_{i=1}^n$ gut?

Def Hypothese $H_0 : \Theta_0 \subseteq \Theta$, **Alternative** $H_A : \Theta_A \subseteq \Theta$

Remark Häufig: $\Theta_A = \Theta_0^c = \Theta \setminus \Theta_0$

Def Θ_x **Einfach** $\stackrel{\text{def}}{\iff} |\Theta_x| = 1$

Ansonsten: *Zusammengesetzt*

Def Test (T, K)

Eine *Entscheidungsregel*, um H_0 durch $\{x_i\}_{i=1}^n$ zu bewerten

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

Kritischer Bereich $K \subseteq \mathbb{R}$

Entscheidungsregel: H_0 verworfen $\iff T(\omega) \in K$

Remark Da T zufällig: $\{T \in K\} = \{\omega \in \Omega \mid T(\omega) \in K\}$ und $\mathbb{P}_\vartheta[T \in K]$ ist berechenbar.

Def Fehler 1. Art: H_0 wird abgelehnt, obwohl korrekt

$\vartheta \in \Theta_0$ s.d. $T \in K$. $\mathbb{P}_\vartheta[T \in K]$ für $\vartheta \in \Theta_0$ ist W.-keit 1. Art

Def Fehler 2. Art: H_0 wird angenommen, obwohl falsch

$\vartheta \in \Theta_A$ s.d. $T \notin K$. $\mathbb{P}_\vartheta[T \notin K]$ für $\vartheta \in \Theta_A$ ist W.-keit 2. Art

7.5.1 Statistisches Vorgehen

Def Signifikanzniveau α s.d. $\sup_{\vartheta \in \Theta_0} \mathbb{P}_\vartheta[T \in K] \leq \alpha$

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

Method Asymmetrisches Verfahren:

1. α auswählen. (W.-keit Fehler 1. Art auswählen)
2. β maximieren. (W.-keit Fehler 2. Art minimieren)

Remark Die Wahl von H_0, H_A ist signifikant. Gemäss Verfahren ist es schwieriger, die Hypothese zu verwerfen als beizubehalten. Man nutzt d.h. die *Negation* der gewünschten Aussage als Hypothese H_0 .

7.6 Konstruktion von Tests

$X_1, \dots, X_n$ diskrete/stetig unter $\mathbb{P}_{\vartheta_0}$ und $\mathbb{P}_{\vartheta_A}$

Def Likelihood Quotient

$x_1, \dots, x_n \in \mathbb{R}, \quad \vartheta_0 \in \Theta_0, \quad \vartheta_A \in \Theta_A$

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

Intuitiv, wenn R gross ist, ist $\{x_i\}_{i=1}^n$ deutlich wahrscheinlicher in $\mathbb{P}_{\vartheta_0}$ als $\mathbb{P}_{\vartheta_A}$.

Remark $R(x_1, \dots, x_n; \vartheta_0, \vartheta_A) := +\infty \iff L(x_1, \dots, x_n; \vartheta_A) = 0$

Def Likelihood Quotient Test $c \geq 0$

$$T = R(X_1, \dots, X_n; \vartheta_0, \vartheta_A) \quad K = (c, \infty]$$

Lem. Neyman-Pearson

$\forall (T', K')$ s.d. $\mathbb{P}_{\vartheta_0}[T' \in K'] = \alpha \leq \alpha^* :$

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

$\Theta_0 = \{\vartheta_0\}, \Theta_A = \{\vartheta_A\}, (T, K)$ ist L.Q.-Test, $\alpha^* = \mathbb{P}_{\vartheta_0}[T \in K]$

D.h. jeder solcher Test mit kleinerem α hat ein kleineres β .

Remark Die Situation $|\Theta_0| = |\Theta_A| = 1$ ist selten. Man wendet trotzdem eine Verallgemeinerung des oberen Prinzips für komplexere Fälle an.

Def Verallgemeinerter Likelihood Quotient

$$R(x_1, \dots, x_n) := \frac{\sup_{\vartheta \in \Theta_0} L(x_1, \dots, x_n; \vartheta)}{\sup_{\vartheta \in \Theta_A} L(x_1, \dots, x_n; \vartheta)}$$

$$\tilde{R}(x_1, \dots, x_n) := \frac{\sup_{\vartheta \in \Theta_0 \cup \Theta_A} L(x_1, \dots, x_n; \vartheta)}{\sup_{\vartheta \in \Theta_A} L(x_1, \dots, x_n; \vartheta)}$$

Def Gauss-Z-Test

$X_1, \dots, X_n \stackrel{\text{i.i.d.}}{\sim} \mathcal{N}(\vartheta, \sigma^2)$, σ^2 bekannt

$$H_0 : \vartheta = \vartheta_0$$

$$H_A : \vartheta > \vartheta_0 \vee \vartheta < \vartheta_0$$

$$T : T = \frac{\bar{X}_n - \vartheta_0}{\frac{\sigma}{\sqrt{n}}} \sim \mathcal{N}(0, 1) \text{ unter } \mathbb{P}_{\vartheta_0}$$

$$K : K = \begin{cases} (c_>, \infty) \vee (-\infty, c_<) & \text{Einseitig} \\ (-\infty, -c_\neq) \cup (c_\neq, \infty) & \text{Beidseitig} \end{cases}$$

Remark Im beidseitigen Fall verwirft man H_0 falls $|T| > c_\neq$.

Remark Man bestimmt $c_>, c_<, c_\neq$ durch Verteilung T unter $\mathbb{P}_{\vartheta_0}$

Beispiel: Gauss-z-Test mit Quantilen

$$\begin{aligned} \alpha &= \mathbb{P}_{\vartheta_0}[T \in K] \\ &= \mathbb{P}_{\vartheta_0}[T > c_>] \\ &= 1 - \mathbb{P}_{\vartheta_0}[T \leq c_>] = \boxed{1 - \Phi(c_>)} \end{aligned}$$

$c_> = \Phi^{-1}(1 - \alpha) =: z_{1-\alpha}$ ist das $(1 - \alpha)$ -Quantil von $\mathcal{N}(1, 0)$.
Man verwirft für $\vartheta > \vartheta_0$ also H_0 falls:

$$\frac{1}{n} \sum_{i=1}^n X_i > \vartheta_0 + z_{1-\alpha} \frac{\sigma}{\sqrt{n}}$$

Wegen der Symmetrie von $\mathcal{N}(0, 1)$: $c_< = z_\alpha = -z_{1-\alpha}$ und $c_\neq = z_{1-\frac{\alpha}{2}}$

Def t-Test

$X_1, \dots, X_n \stackrel{\text{i.i.d.}}{\sim} \mathcal{N}(\mu, \sigma^2)$, $\vartheta = (\mu, \sigma^2)$

$$H_0 : \mu = \mu_0$$

$$T : T = \frac{\bar{X}_n - \mu_0}{\frac{S}{\sqrt{n}}} \sim t_{n-1} \text{ unter } \mathbb{P}_{\vartheta_0}$$

$$S : S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_k - \bar{X}_n)^2$$

$$K : K = \begin{cases} (c_>, \infty) \vee (-\infty, c_<) & \text{Einseitig} \\ (-\infty, -c_\neq) \cup (c_\neq, \infty) & \text{Beidseitig} \end{cases}$$

$$c : c = \begin{cases} c_> = t_{n-1, 1-\alpha} \\ c_< = t_{n-1, \alpha} = -t_{n-1, 1-\alpha} \\ c_\neq = t_{n-1, 1-\frac{\alpha}{2}} \end{cases}$$

Def γ -Quantil $t_{m, \gamma}$ s.d. $\mathbb{P}[X \leq t_{m, \gamma}] = \gamma$ für $X \sim t_m$

Def Gepaarter Zweistichproben-Test

$m = n$, σ^2 identisch

$$X_1, \dots, X_n \stackrel{\text{i.i.d.}}{\sim} \mathcal{N}(\mu_X, \sigma^2) \text{ unter } \mathbb{P}_\vartheta$$

$$Y_1, \dots, Y_n \stackrel{\text{i.i.d.}}{\sim} \mathcal{N}(\mu_Y, \sigma^2) \text{ unter } \mathbb{P}_\vartheta$$

Remark $(X_k - Y_k) \stackrel{\text{i.i.d.}}{\sim} \mathcal{N}(\mu_X - \mu_Y, 2(1 - \rho)\sigma^2)$
 $\rho \in (-1, 1)$ ist Korrelation $\bar{X}, \bar{Y}$. Bei Unabhängigkeit: $\rho = 0$

Def Ungepaarter Zweistichproben-Test $\stackrel{\text{def}}{\iff} m \neq n$

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

Bei unbekanntem σ^2 wird S^2 genutzt:

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

7.7 p-Wert

Motivation: Bisherige Tests nutzen alle $\mathcal{N}(\mu, \sigma^2)$.

p-Wert Intuitiv: Die W.-keit, unter H_0 einen min. so extremen Wert für T zu bekommen wie $T(\omega)$ aus der Stichprobe.

Def Geordnete Testsammlung $(T, K_t)_{t \geq 0}$

$$\forall t : K_t \subseteq \mathbb{R} \quad \wedge \quad s \leq t \implies K_s \supset K_t$$

z.B. $(T, K_t)_{t \geq 0}$ s.d. $K_t = (t, \infty)$ oder $K_t = (-\infty, -t)$

Def p-Wert

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

$$p\text{-Wert} = G(T) : \Omega \rightarrow [0, 1]$$

$H_0 : \vartheta = \vartheta_0$ einfach, $(T, K_t)_{t \geq 0}$ geordnet

Remark Interpretation

$$\alpha > p \implies (T, K_t) \text{ lehnt } H_0 \text{ ab}$$

$$\alpha \leq p \implies (T, K_t) \text{ akzeptiert } H_0$$

Remark p-Wert ist unabhängig von H_A

7.8 Konfidenzintervall

Motivation: T liefert nur ein ϑ , wieso nicht einen Intervall?

Def Konfidenzbereich

$$C(x_1, \dots, x_n) \subseteq \Theta \quad \tilde{C}(X_1, \dots, X_n) \subseteq \Theta$$

Def Konfidenzbereich, Niveau $1 - \alpha$

$$\forall \vartheta \in \Theta : \mathbb{P}[\vartheta \in C(X_1, \dots, X_n)] \geq 1 - \alpha$$

Def Konfidenzintervall

A, B sind Zufallsvariablen

$$I = [A, B] \text{ s.d. } \mathbb{P}[\vartheta \in [A, B]] = \mathbb{P}[A \leq \vartheta \leq B] \geq 1 - \alpha$$

$$A = a(X_1, \dots, X_n) \quad B = b(X_1, \dots, X_n)$$

8 Tabellen

8.1 Integrale & Ableitungen

$F(x)$	$f(x)$	$f'(x)$
x	c	0
$\frac{1}{a+1} \cdot x^{a+1}$	x^a	$a \cdot x^{a-1}$
$\frac{1}{a \cdot (n+1)} \cdot (ax+b)^{n+1}$	$(ax+b)^n$	$n \cdot (ax+b)^{n-1} \cdot a$
$\ln x $	$\frac{1}{x} = x^{-1}$	$-\frac{1}{x^2} = -x^{-2}$
$\frac{2}{3} \cdot x^{\frac{3}{2}}$	$\sqrt{x} = x^{\frac{1}{2}}$	$\frac{1}{2\sqrt{x}} = \frac{1}{2} \cdot x^{-\frac{1}{2}}$
$\frac{n}{n+1} \cdot x^{\frac{1}{n}+1}$	$\sqrt[n]{x} = x^{\frac{1}{n}}$	$\frac{1}{n} \cdot x^{\frac{1}{n}-1}$
e^x	e^x	e^x
$\frac{1}{\ln(a)} \cdot a^x$	a^x	$a^x \cdot \ln(a)$
$x \cdot (\ln x - 1)$	$\ln x $	$\frac{1}{x} = x^{-1}$
$\frac{x}{\ln(a)} \cdot (\ln x - 1)$	$\log_a(x)$	$\frac{1}{x \cdot \ln(a)}$
$-\cos(x)$	$\sin(x)$	$\cos(x)$
$\sin(x)$	$\cos(x)$	$-\sin(x)$
$-\ln \cos(x) $	$\tan(x)$	$\frac{1}{\cos(x)^2} = 1 + \tan(x)^2$
$\ln \sin(x) $	$\cot(x)$	$-\frac{1}{\sin(x)^2}$
$x \cdot \arcsin(x) + \sqrt{1-x^2}$	$\arcsin(x)$	$\frac{1}{\sqrt{1-x^2}}$
$x \cdot \arccos(x) - \sqrt{1-x^2}$	$\arccos(x)$	$-\frac{1}{\sqrt{1-x^2}}$
$x \cdot \arctan(x) - \frac{\ln(x^2+1)}{2}$	$\arctan(x)$	$\frac{1}{x^2+1}$
$\sinh(x)$	$\cosh(x)$	$\sinh(x)$
$\cosh(x)$	$\sinh(x)$	$\cosh(x)$
$\ln \cosh(x) $	$\tanh(x)$	$\frac{1}{\cosh(x)^2} = 1 - \tanh(x)^2$
	$\operatorname{arcsinh}(x)$	$\frac{1}{\sqrt{x^2+1}}$
	$\operatorname{arccosh}(x)$	$\frac{1}{\sqrt{x^2-1}}$
	$\operatorname{arctanh}()$	$\frac{1}{1-x^2}$

8.2 Weitere Integrale & Ableitungen

$F(x)$	$f(x)$
$\frac{1}{a} \ln ax+b $	$\frac{1}{ax+b}$
$\frac{ax}{c} - \frac{ad-bc}{c^2} \ln cx+d $	$\frac{a(cx+d)-c(ax+b)}{(cx+d)^2}$
$\frac{x}{2}f(x) + \frac{a^2}{2} \ln x+f(x) $	$\sqrt{a^2+x^2}$
$\frac{x}{2}f(x) - \frac{a^2}{2} \arcsin(\frac{x}{ a })$	$\sqrt{a^2-x^2}$
$\frac{x}{2}f(x) - \frac{a^2}{2} \ln x+f(x) $	$\sqrt{x^2-a^2}$
$\ln(x+\sqrt{x^2\pm a^2})$	$\frac{1}{\sqrt{x^2\pm a^2}}$
$\arcsin(\frac{x}{ a })$	$\frac{1}{\sqrt{a^2-x^2}}$
$\frac{1}{a} \cdot \arctan(\frac{x}{a})$	$\frac{1}{x^2+a^2}$
$-\frac{1}{a} \cos(ax+b)$	$\sin(ax+b)$
$\frac{1}{a} \sin(ax+b)$	$\cos(ax+b)$
x^x	$x^x \cdot (1+\ln x)$
$(x^x)^x$	$(x^x)^x (x+2x\ln x)$
$x^{(x^x)}$	$x^{(x^x)} (x^{x-1} + \ln x \cdot x^x (1+\ln x))$
$\frac{1}{2}(x - \frac{1}{2} \sin(2x))$	$\sin(x)^2$
$\frac{1}{2}(x + \frac{1}{2} \sin(2x))$	$\cos(x)^2$

8.3 Werte der trigonometrischen Funktionen

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	0	$\frac{1}{2}$	$\frac{\sqrt{2}}{2}$	$\frac{\sqrt{3}}{2}$	1	$\frac{\sqrt{3}}{2}$	$\frac{1}{2}$	0	-1
cos	1	$\frac{\sqrt{3}}{2}$	$\frac{\sqrt{2}}{2}$	$\frac{1}{2}$	0	$-\frac{1}{2}$	$-\frac{\sqrt{3}}{2}$	-1	0
tan	0	$\frac{\sqrt{3}}{3}$	1	$\sqrt{3}$		$-\sqrt{3}$	$-\frac{\sqrt{3}}{3}$	0	

8.4 Grenzwerte: Folgen

Credits: D. Camenisch

$\lim_{x \rightarrow \infty} \frac{1}{x} = 0$	$\lim_{x \rightarrow \infty} 1 + \frac{1}{x} = 1$
$\lim_{x \rightarrow \infty} e^x = \infty$	$\lim_{x \rightarrow -\infty} e^x = 0$
$\lim_{x \rightarrow \infty} e^{-x} = 0$	$\lim_{x \rightarrow -\infty} e^{-x} = \infty$
$\lim_{x \rightarrow \infty} \frac{e^x}{x^m} = \infty$	$\lim_{x \rightarrow -\infty} xe^x = 0$
$\lim_{x \rightarrow \infty} \ln(x) = \infty$	$\lim_{x \rightarrow 0} \ln(x) = -\infty$
$\lim_{x \rightarrow \infty} (1+x)^{\frac{1}{x}} = 1$	$\lim_{x \rightarrow 0} (1+x)^{\frac{1}{x}} = e$
$\lim_{x \rightarrow \infty} (1+\frac{1}{x})^b = 1$	$\lim_{x \rightarrow \infty} (1+\frac{1}{x})^b = 1$
$\lim_{x \rightarrow \infty} x^a q^x = 0, \forall 0 \leq q < 1$	$\lim_{x \rightarrow \infty} n^{\frac{1}{n}} = 1$
$\lim_{x \rightarrow \pm \infty} (1+\frac{1}{x})^x = e$	$\lim_{x \rightarrow \infty} (1-\frac{1}{x})^x = \frac{1}{e}$
$\lim_{x \rightarrow \pm \infty} (1+\frac{k}{x})^{mx} = e^{km}$	$\lim_{x \rightarrow 0} \frac{\sin x}{x} = 1$
$\lim_{x \rightarrow 0} \frac{1}{\cos(x)} = 1$	$\lim_{x \rightarrow 0} \frac{\cos x - 1}{x} = 0$
$\lim_{x \rightarrow 0} \frac{\log 1-x}{x} = -1$	$\lim_{x \rightarrow 0} x \log x = 0$
$\lim_{x \rightarrow 0} \frac{1-\cos x}{x^2} = \frac{1}{2}$	$\lim_{x \rightarrow 0} \frac{e^x-1}{x} = 1$
$\lim_{x \rightarrow 0} \frac{x}{\arctan x} = 1$	$\lim_{x \rightarrow \infty} \arctan x = \frac{\pi}{2}$
$\lim_{x \rightarrow \infty} \left(\frac{x}{x+k}\right)^x = e^{-k}$	$\lim_{x \rightarrow 0} \frac{e^x-1}{x} = 1$
$\lim_{x \rightarrow 0} \frac{a^x-1}{x} = \ln(a) \forall a > 0$	$\lim_{x \rightarrow 0} \frac{e^{ax}-1}{x} = a$
$\lim_{x \rightarrow 0} \frac{\ln(x+1)}{x} = 1$	$\lim_{x \rightarrow 1} \frac{\ln(x)}{x-1} = 1$
$\lim_{x \rightarrow \infty} \frac{\ln(x)}{x} = 0$	$\lim_{x \rightarrow \infty} \frac{\log(x)}{x^a} = 0$
$\lim_{x \rightarrow \infty} \sqrt[x]{x} = 1$	$\lim_{x \rightarrow \infty} \frac{2x}{2^x} = 0$
$\lim_{x \rightarrow \frac{\pi}{2}^-} \tan x = +\infty$	$\lim_{x \rightarrow \frac{\pi}{2}^+} \tan x = -\infty$
$\lim_{x \rightarrow \infty} \frac{\sin x}{x} = 0$	$\lim_{x \rightarrow 0^+} x \ln x = 0$

8.5 Grenzwerte: Reihen

Credits: D. Camenisch

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

8.6 Verteilungen

Credits: N. Wehrli & S. Metzker

Verteilung	Notation	Parameter	$\mathbb{E}[X]$	$\mathbb{V}(X)$	$p_X(t)/f_X(t)$	$F_X(t)$
Gleichverteilung	unbekannt	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	$\text{Ber}(p)$	p : ErfolgsW'keit	p	$p \cdot (1 - p)$	$p^t(1 - p)^{1-t}$	$1 - p$ für $0 \leq t < 1$
Binomial	$\text{Bin}(n, p)$	n : Anzahl Versuche p : ErfolgsW'keit	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	$\text{Geo}(p)$	p : ErfolgsW'keit (t : Anzahl Versuche)	$\frac{1}{p}$	$\frac{1-p}{p^2}$	$p(1 - p)^{t-1}$	$1 - (1 - p)^t$
Poisson	$\text{Poisson}(\lambda)$	λ : Erwartungswert und Varianz	λ	λ	$\frac{\lambda^t}{t!} e^{-\lambda}$	$e^{-\lambda} \sum_{k=0}^t \frac{\lambda^k}{k!}$
Gleichverteilung (im Intervall)	$U \sim \mathcal{U}([0, 1])$	$[a, b]$: Intervall	$\frac{a+b}{2}$	$\frac{1}{12}(b - a)^2$	$\begin{cases} \frac{1}{b-a} & a \leq x \leq b \\ 0 & \text{sonst} \end{cases}$	$\begin{cases} 0 & x \leq a \\ \frac{t-a}{b-a} & a < x < b \\ 1 & x \geq b \end{cases}$
Exponentialv.	$\text{Exp}(\lambda)$	$\lambda : \frac{1}{\mathbb{E}[X]}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$	$\begin{cases} \lambda e^{-\lambda t} & t \geq 0 \\ 0 & t < 0 \end{cases}$	$\begin{cases} 1 - e^{-\lambda t} & t > 0 \\ 0 & t \leq 0 \end{cases}$
Normalverteilung	$\mathcal{N}(\mu, \sigma^2)$	$\mu : \mathbb{E}[X]$ σ^2 : Varianz	μ	σ^2	$\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(t-\mu)^2}{2\sigma^2}}$	$\frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^t e^{-\frac{1}{2}(\frac{y-\mu}{\sigma})^2} dy$
χ^2 -Verteilung	χ_m^2	n : Freiheitsgrad	n	$2n$	$\frac{1}{2^{\frac{n}{2}} \Gamma(\frac{n}{2})} t^{\frac{n}{2}-1} e^{-\frac{t}{2}}$ für $t > 0$	$P(\frac{n}{2}, \frac{t}{2})$
t-Verteilung	t_m	n : Freiheitsgrad	$\begin{cases} 0 & n > 1 \\ \text{undef.} & \text{sonst} \end{cases}$	$\begin{cases} \frac{n}{n-2} & n > 2 \\ \infty & 1 < n \leq 2 \\ \text{undef.} & \text{sonst} \end{cases}$	$\frac{\Gamma(\frac{n+1}{2})}{\sqrt{n\pi} \cdot \Gamma(\frac{n}{2})} \left(1 + \frac{t^2}{n}\right)^{-\frac{n+1}{2}}$	zu kompliziert
Negativbinomial	$\text{NBin}(r, p)$	$r \in \mathbb{N}, p \in [0, 1]$	$\frac{r}{p}$	$\frac{r(1-p)}{p^2}$	$\binom{k-1}{r-1} p^r (1-p)^{k-r}$	zu kompliziert
Cauchy-Verteilung	Cauchy(x_0, γ)	$x_0 \in \mathbb{R}, \gamma > 0$	Existiert nicht	Existiert nicht	$\frac{1}{\pi} \frac{\gamma}{\gamma^2 + (x-x_0)^2}$	$\frac{1}{2} + \frac{1}{\pi} \arctan\left(\frac{x-x_0}{\gamma}\right)$
Hypergeometrisch	$H(n, r, m)$	$n \in \mathbb{N}, m, r \in \{1, \dots, n\}$	$m \frac{r}{n}$	$m \frac{r}{n} \left(1 - \frac{r}{n}\right) \frac{n-m}{n-1}$	$\frac{\binom{r}{k} \binom{n-r}{m-k}}{\binom{n}{m}}$	$\sum_{y=0}^k \frac{\binom{r}{y} \binom{n-r}{m-y}}{\binom{n}{m}}$