Wahrscheinlichkeit und Statistik

David Zollikofer

Teil I
Wahrscheinlichkeit

1. Wahrscheinlichkeitsrechnung

Wahrscheinlichkeitsraum Wir definieren $(\Omega, \mathcal{F}, P)$ einen Wahrscheinlichkeitsraum, wobei Ω die Ereignismenge aus Elementarereignissen ist, $\mathcal{F} \subseteq 2^\Omega$ eine σ-Algebra und P ein Wahrscheinlichkeitsmass.

σ-Algebra Wir nennen ein Mengensystem $\mathcal{F} \subseteq 2^\Omega$ eine σ-Algebra falls:

- $\Omega \in \mathcal{F}$
- $\forall A \in \mathcal{F} : A^c \in \mathcal{F}$
- für jede Folge $(A_n)_{n\in\mathbb{N}}$ mit $A_n \in \mathcal{F}$ so ist auch $\bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$

Sicherlich ist somit zum Beispiel die Potenzmenge 2^Ω eine σ-Algebra.

Wahrscheinlichkeitsmass Wir definieren eine Abbildung $P : \mathcal{F} \to [0, 1]$. Wir nennen $P[A] \in [0, 1]$ die Wahrscheinlichkeit, dass A eintritt. Die geforderten Kolmogorov Axiome sind:

- $P[A] \geq 0 \; \forall A \in \mathcal{F}$
- $P[\Omega] = 1$
- $P\left[\bigcup_{i=1}^{\infty} A_i\right] = \sum_{i=1}^{\infty} P[A_i]$, sofern die $A_i \in \mathcal{F}$ paarweise disjunkt sind ($A_i \cap A_k = \varnothing$ wenn $i \neq k$)

Beispiele (Formale Definitionen von Wahrscheinlichkeitsräumen)

Grundlegende Prinzipien

Additivität disjunkter Ereignisse Seien $A_1, \ldots, A_n$ paarweise disjunkte Ereignisse, so gilt:

$$P[A_1 \cup \cdots \cup A_n] = P[A_1] + \cdots + P[A_n]$$

Inklusion und Exklusionsprinzip

$$\begin{aligned} P[A \cup B] &= P[A] + P[B] - P[A \cap B] \\ P[A \cup B \cup C] &= P[A] + P[B] + P[C] - P[A \cap B] - P[B \cap C] \\ &\quad - P[B \cap C] + P[A \cap B \cap C] \end{aligned}$$

Bedingte Wahrscheinlichkeit

Bedingte Wahrscheinlichkeit Seien A, B Ereignisse und $P[A] > 0$. Die bedingte Wahrscheinlichkeit von B unter der Bedingung, dass A eintritt, (gegeben A) ist definiert als:

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

Multiplikationsregel Seien $A_1, \ldots A_n$ Ereignisse mit $P[A_i] > 0$ (div by 0 Problem), dann gilt:

$$P[A_1 \cap A_2 \cap \cdots \cap A_n] = P[A_1] \cdot P[A_2|A_1] \cdot P[A_3|A_1 \cap A_2] \cdot \ldots \cdot P[A_n|A_1 \cap \cdots \cap A_{n-1}]$$

Satz der totalen Wahrscheinlichkeit Seien $B_1, \ldots B_n$ eine Zerlegung von Ω (d.h. $\bigcup_{i=1}^{n} B_i = \Omega$ und $B_i \cap B_j = \varnothing$ für $i \neq j$), so gilt für ein beliebiges Ereignis A:

$$P[A] = \sum_{i=1}^{n} P[A \cap B_i] = \sum_{i=1}^{n} P[A|B_i] \cdot P[B_i]$$

Insbesondere folgt daraus:

$$P[A] = P[A \cap B] + P[A \cap B^c] = P[A|B] \cdot P[B] + P[A|B^c] \cdot P[B^c]$$

Satz von Bayes Wenn $P[A], P[B].P[B^c] > 0$ so folgt:

$$P[B|A] = \frac{P[A \cap B]}{P[A]} = \frac{P[A|B] \cdot P[B]}{P[A|B] \cdot P[B] + P[A|B^c] \cdot P[B^c]}$$

respektive wenn $B_1, \ldots B_n$ eine Zerlegung von Ω mit $P[B_i] > 0 \; \forall i$, dann gilt für ein Ereignis A mit $P[A] > 0$:

$$P[B_k|A] = \frac{P[A \cap B_k]}{P[A]} = \frac{P[A|B_k] \cdot P[B_k]}{\sum_{i=1}^{n} P[A|B_i] \cdot P[B_i]}$$

Unabhängigkeit

Unabhängigkeit Wir nennen zwei Ereignisse A, B unabhängig falls

$$P[A \cap B] = P[A] \cdot P[B]$$

Für $P[A] \neq 0$ gilt:

$$A, B \text{ unabhängig} \iff P[B|A] = P[B]$$

Stochastische Unabhängigkeit Wir nennen $A_1, \cdots A_n$ (stochastisch) unabhängig, falls für alle Kombinationen von $A_i, \ldots A_j$ gilt dass

$$P\left[\bigcap_{i=1}^{m} A_{k_i}\right] = \prod_{i=1}^{m} P[A_{k_i}]$$

2. Diskrete Zufallsvariablen und Verteilungen

Diskrete Zufallsvariable (ZV) Sei $(\Omega, \mathcal{F}, P)$ ein diskreter Wahrscheinlichkeitsraum.
Wir nennen

$$X : \Omega \to \mathcal{W}(X) = \{x_1, \ldots x_n\} \subseteq \mathbb{R}$$

eine Zufallsvariable. mit $\mathcal{W}(X)$ der Wertebereich.
Zusätzlich definieren wir die *Gewichtsfunktion* oder *diskrete Dichte* von X als

$$p_X(x_k) = P[X = x_k] = P[\{\omega | X(\omega) = x_k\}]$$

sowie auch die Verteilfunktion

$$F_x(t) = P[X \leq t] = P[\{\omega | X(\omega) \leq t\}]$$

Verteilungen mehrerer Variablen

Gemeinsame Verteilfunktion Seien $X_1, \ldots X_n$ Zufallsvariablen. Die gemeinsame Verteilfunktion $F : \mathbb{R}^n \to [0, 1]$ ist definiert durch

$$\begin{aligned} (x_1, \ldots x_n) \mapsto F(x_1, \ldots x_n) &= P[X_1 \leq x_1, \ldots X_n \leq x_n] \\ &= \sum_{y_1 \leq x_1, \ldots, y_n \leq x_n} p(y_1, \ldots, y_n) \end{aligned}$$

Gemeinsame Verteilfunktion Seien X,Y ZV, so gilt für $F_X : \mathbb{R} \to [0,1]$:

$$F_X(x) = P[X \leq x] = P[X \leq x, Y < \infty] = \lim_{y\to\infty} F(x,y)$$

Gewichtsfunktion der Randverteilung Wir definieren $p_X(x) : \mathcal{W}(X) \to [0,1]$ als die Gewichtsfunktion der Randverteilung von X gegeben durch

$$p_X(x) = P[X = x] = \sum_{y_i \in \mathcal{W}(Y)} p(x, y_i)$$

Unabhängigkeit von Zufallsvariable Die ZV $X_1, \ldots X_n$ heissen unabhängig, gdw

$$F(x_1, \ldots, x_n) = F_{X_1}(x_1) \cdots F_{X_n}(x_n) \quad \forall x_1, \ldots, x_n$$

oder analog dazu

$$p(x_1, \ldots, x_n) = p_{X_1}(x_1) \cdots p_{X_1 n}(x_n) \quad \forall x_1, \ldots, x_n$$

Bedingte Verteilungen

Bedingte Verteilungen Seien X und Y Zufallsvariablen mit gemeinsamer Gewichtsfunktion $p(x,y)$. Die bedingte Gewichtsfunktion von X gegeben, dass $Y = y$ ist definiert durch

$$p_{X|Y}(x|y) := P[X = x | Y = y] = \frac{P[X = x, Y = y]}{P[Y = y]} = \frac{p(x,y)}{p_Y(y)}$$

Bedingte Erwartung Die Erwartung von Y gegeben, dass $X = x$ bereits eingetroffen ist, ist definiert durch:

$$\mathrm{E}[Y|X = x] = \sum_y y P(Y = y \mid X = x) = \sum_y y \frac{P(X = x, Y = y)}{P(X = x)}$$

$$\mathrm{E}[Y|X = x] = \int y \cdot f_{Y|X=x}(y)\, dy = \int y \cdot \frac{f_{X,Y}(x,y)}{f_{X=x}(y)}\, dy$$

3. Wichige diskrete Verteilungen

Diskrete Gleichverteilung

Die diskrete Gleichverteilung auf $\mathcal{W}(X) = \{x_1, \ldots, x_n\}$ hat folgende Eigenschaften:

$$p_X(x_k) = P[X = x_k] = \frac{1}{N}$$

mit

$$\mathrm{E}[X] = \frac{1}{n}\sum_{i=1}^n x_i \qquad \mathrm{Var}[X] = \frac{1}{n}\left(\sum_{i=1}^n x_i^2 - \frac{1}{n}\left(\sum_{i=1}^n x_i\right)^2\right)$$

Bernoulli Verteilung

Wir machen ein einziges 0-1 Experiment mit Erfolgswahrscheinlichkeit p. Es gilt demnach $\mathcal{W}(X) = \{0,1\}$ mit

$$p_X(1) = P[X = 1] = p \qquad p_X(0) = P[X = 0] = 1 - p$$

Woraus folgt:

$$\mathrm{E}[X] = p \qquad \mathrm{Var}[X] = p(1-p)$$

Wir schreiben dabei $X \sim Be(p)$

Binomialverteilung

Wir möchten gerne die Anzahl Erfolge bei n unabhängigen 0-1 Experimenten mit Erfolgsparameter p beschreiben. Wir schreiben $X \sim \mathrm{Bin}(n,p)$. Dabei ist $\mathcal{W}(X) = \{0, 1, \ldots, n\}$ und

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

sowie

$$\mathrm{E}[X] = np \qquad \mathrm{Var}[X] = np(1-p)$$

Geometrische Verteilung

Wir betrachten eine unendliche Folge von unabhängigen 0-1 Experimenten mit Erfolgsparameter p. X sei die Wartezeit auf den ersten Erfolg. Wir schreiben $X \sim \mathrm{Geom}(p)$ und haben:

$$p_X(k) = P[X = k] = p(1-p)^{k-1}$$

$$\mathrm{E}[X] = \frac{1}{p} \qquad \mathrm{Var}[X] = \frac{1-p}{p^2}$$

Für die geometrische Verteilung gilt auch eine Art Gedächtnislosigkeit: $P[X = n + k | X \geq n] = P[X = k]$.

Negativbinomiale Verteilung

Wir betrachten eine unendliche Folge von unabhängigen 0-1 Experimenten mit Erfolgsparameter p. So sei X die Wartezeit auf den r-ten Erfolg. Wir schreiben $X \sim \mathrm{NB(r,p)}$.

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

$$\mathrm{E}[X] = \frac{r}{p} \qquad \mathrm{Var}[X] = \frac{r(1-p)}{p^2}$$

Hypergeometrische Verteilung

In einer Urne seien n Gegenstände, davon r vom Typ 1 und $n - r$ vom Typ 2. Man zieht ohne Zurücklegen m der Gegenstände; die Zufallsvariable X beschreibe die Anzahl der Gegenstände vom Typ 1 in dieser Stichprobe vom Umfang m. Dann hat X eine hypergeometrische Verteilung mit Parametern n, m, r mit $\mathcal{W}(X) = \{0, 1, \ldots, \min(m,r)\}$ sowie

$$p_X(k) = \frac{\binom{r}{k}\binom{n-r}{m-k}}{\binom{n}{m}}$$

$$\mathrm{E}[X] = m\frac{r}{n} \qquad \mathrm{Var}[X] = m\frac{r}{n}\left(1 - \frac{r}{n}\right)\frac{n-m}{n-1}$$

Poissonverteilung

Mit Parameter $\lambda > 0$ auf $\mathcal{W}(X) = \mathbb{N}_0$ schreiben wir $X \sim \mathrm{Pois}(\lambda)$

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

$$\mathrm{E}[X] = \lambda \qquad \mathrm{Var}[X] = \lambda$$

4. Allgemeine Zufallsvariable

Verteilungsfunktion Wir nennen $F_X : \mathbb{R} \to [0,1]$ mit

$$F_X(t) = P[X \leq t] := P[\{\omega | X(\omega) \leq t\}]$$

eine Verteilfunktion. Dies hat folgende Eigenschaften:

- F_X ist monoton wachsend $F_X(s) \leq F_X(t)$ für $s \leq t$ und rechtsstetig $F_X(u) \to F_X(t)$ für $u \to t$.
- $\lim_{t\to-\infty} F_X(t) = 0, \lim_{t\to+\infty} F_X(t) = 1$

umgekehrt ist jede Funktion mit diesen Eigenschaften eine Verteilungsfunktion F_X einer Zufallsvariablen X.

Wahrscheinlichkeitsmass Wir nennen $\mu_X(B) := P[X \in B]$ das Wahrscheinlichkeitsmass mit $\mu_X((-\infty, t]) = F_X(t)$

Dichtefunktion Falls $F_X(t) = \int_{-\infty}^t f_X(s)ds$ für $\forall t \in \mathbb{R}$ so nennt man $F_X(t)$ absolut stetig und $f_X(s)$ die Dichtefunktion. Mit folgenden Eigenschaften:

- $f_X \geq 0$ und $f_X = 0$ ausserhalb von $\mathcal{W}(X)$.
- $\int_{-\infty}^{\infty} f_X(s)ds = 1$ folgt aus $\lim_{t\to-\infty} F_X(t) = 1$

Umgekehrt kann man aus einer messbaren Funktion $f : \mathbb{R} \to [0,\infty)$ mit $\int_{-\infty}^{\infty} f(s)ds = 1$ eine Zufallsfariable X konstruieren.

Gemeinsame Verteilungen, unabhängige Zufallsvariablen

Gemeinsame Verteilungsfunktion Die gemeinsame Verteilfunktion von n Zufallsvariablen $X_1, \ldots, X_n$ ist die Abbildung $F : \mathbb{R}^n \to [0,1]$ mit

$$F(x_1,\ldots,x_n) := P[X_1 \le x_1, \ldots, X_n \le x_n]$$
$$F(x_1,\ldots,x_n) := \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} f(t_1,\ldots,t_n) dt_n \ldots dt_1$$

dabei ist $f : \mathbb{R}^n \to [0,\infty)$ die gemeinsame Dichte für welche gilt:

- $f(x_1,\ldots,x_n) \ge 0$ und $= 0$ ausserhalb von $\mathcal{W}(X_1,\ldots,X_n)$
- $\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(t_1,\ldots,t_n) dt_n \ldots dt_1 = 1$
- $P[(X_1,\ldots X_n) \in A] = \int_{(x_1,\ldots,x_n)\in A} f(t_1,\ldots,t_n) dt_n \ldots dt_1$

Randverteilung Haben X, Y die gemeinsame Dichtefunktion F, so ist die Funktion $F_X : \mathbb{R} \to [0,1]$ mit

$$F_X(x) := P[Y \le y] = P[x \le Y, Y < \infty] = \lim_{y\to\infty} F(x,y)$$

die sogenannte Randverteilung von X.
Falls die Dichte $f(x,y)$ existiert so gilt:

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y) dy$$
$$f_X(x) = \frac{d}{dx} F_X(x) = \frac{d}{dx} \lim_{y\to\infty} F(x,y)$$

Beispiel zusammengesetzter Zufallsvariablen

- $X = \min\{X_1,\ldots,X_n\}$, finde f_X
  Wir wissen $F_X(t) = P[X \le t] = \prod_{i=1}^n P[X_k \le t] = F(t)^n$. Da $f_X(t) = \frac{dF_X(t)}{dt}$ folgt $f_X(t) = n \cdot F(t)^{n-1} f(t)$.
- $X = \max\{X_1,\ldots,X_n\}$, finde f_X
  Wir wissen $F_X(t) = 1 - P[X > t] = 1 - P[X_1 > t, \ldots X_n > t] = 1 - (1-F(t))^n$. Da $f_X(t) = \frac{dF_X(t)}{dt}$ folgt $f_X(t) = n(1 - F(t))^{n-1} f(t)$.

Wahrscheinlichkeiten mehrerer Variablen Angenommen wir möchten $P[X > 2Y]$ ausrechnen so berechnen wir:

$$P[X > 2Y] = \int\int_A f_{X,Y}(x,y) dx dy$$

Mit $A = \{(x,y) \in \mathbb{R}^2 : x > 2y\}$

Unabhängigkeit von Zufallsvariablen Die Zufallsvariablen $X_1, \ldots, X_n$ heissen unabhängig, falls gilt

$$F(x_1,\ldots x_n) = F_{X_1}(x_1) \cdots F_{X_n}(x_n) \quad \text{respektive:}$$
$$f(x_1,\ldots x_n) = f_{X_1}(x_1) \cdots f_{X_n}(x_n) \quad \forall x_1,\ldots x_n$$

Transformation von stetigen Zufallsvariablen

- Angenommen wir haben die Dichte $f(x,y)$ gegeben und wir suchen die Verteilfunktion und Dichte von $Z = X + Y$. Wir haben $F_Z(z) = P[Z \le z] = P[X + Y \le z]$. Nun definieren wir $A_z = \{(x,y) \in \mathbb{R}^2 | x + y \le z\}$. Es gilt nun:

$$F_Z(z) = P[(X,Y) \in A_z] = \int_{A_z}\int f(x,y) dy dx$$
$$= \int_{-\infty}^{\infty} \int_{-\infty}^{z-x} f(x,y) dy dx$$

  Sowie $f_Z(z) = \frac{d}{dz} F_Z(z)$.
- Sei $U = \frac{X}{X+Y}$ mit $X, Y \sim Exp(\lambda)$; finde f_U, F_U:
  Es gilt $F_U(u) = P[U \le u] = P[\frac{X}{X+Y} \le u]$ woraus $x(u^{-1} - 1) \le y$ folgt. Somit ist das Integral $F_U = \lambda^2 \int_0^\infty e^{-\lambda x} \int_{x(u^{-1}-1)}^{\infty} e^{-\lambda y} dy dx = u$.
  Da nun $0 \le U = \frac{X}{X+Y} \le 1$ folgt $f_U(u) = 1_{u\in[0,1]}$
- Sei nun zusätzlich noch $V = X + Y$ und wir möchten $f_{U,V}$ berechnen, so gilt

$$f_{U,V} = \lambda^2 \int_0^\infty e^{-\lambda x} \left(\int_0^\infty 1_{\{\frac{x}{x+y} \le u, x+y \le v\}} e^{-\lambda y} dy \right) dx$$

Wichtige stetige Verteilungen

Gleichverteilung

Die Gleichverteilung auf $[a,b]$ mit $\mathcal{W}(X) = [a,b]$, genannt $X \sim \mathcal{U}(a,b)$:

$$f_X(t) = \begin{cases} \frac{1}{b-a} & \text{für } a \le t \le b \\ 0 & \text{sonst} \end{cases} \qquad F_X(t) = \begin{cases} 0 & \text{für } t < a \\ \frac{t-a}{b-a} & \text{für } a \le t \le b \\ 1 & \text{für } t > b \end{cases}$$

$$\mathrm{E}[X] = \int_{-\infty}^{\infty} x f(x)\, dx = \frac{1}{b-a} \int_a^b x \cdot 1\, dx = \frac{1}{2}\frac{b^2 - a^2}{b-a} = \frac{a+b}{2}$$

$$\mathrm{Var}[X] = \frac{1}{b-a} \int_a^b x^2 \cdot 1\, dx - \left(\frac{a+b}{2}\right)^2 = \frac{1}{3}\frac{b^3 - a^3}{b-a} - \left(\frac{a+b}{2}\right)^2$$

Exponentialverteilung

Die Exponentialverteilung mit Parameter $\lambda > 0$ mit $\mathcal{W}(X) = [0,\infty)$, genannt $X \sim Exp(\lambda)$:

$$f_X(t) = \begin{cases} \lambda \cdot e^{-\lambda t} & \text{für } t \ge 0 \\ 0 & \text{für } t < 0 \end{cases} \qquad F_X(t) = \begin{cases} 1 - e^{-\lambda t} & \text{für } t \ge 0 \\ 0 & \text{für } t < 0 \end{cases}$$

$$\mathrm{E}[X] = \int_0^\infty \lambda x e^{-\lambda x}\, dx = \frac{1}{\lambda}$$
$$\mathrm{Var}[X] = \frac{1}{\lambda^2}$$

Die Exponentialverteilung ist Gedächtnislos:
$$P[X > t + s | X > s] = P[X > t]$$

Beispiel (Casinogewinn) Wenn $Z \sim Exp(\lambda)$ dürfen wir c wählen so dass wenn $Z > c$ gewinnen wir c. Berechne maximal erwartete Auszahlung: $G(c) = \mathrm{E}[c1_{\{Z>c\}}] = cP[Z > c] = ce^{-\lambda c}$. Nun maximiere c.

Normalverteilung

Die Normalverteilung mit Parameter σ^2, der Varianz, μ, dem Erwartungswert sowie $\mathcal{W}(X) = \mathbb{R}$. Man schreibt $X \sim \mathcal{N}(\mu, \sigma^2)$

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

Problem: Das Integral ist mühsam. Trick Wir plotten die Dichte $\varphi(t)$ sowie Verteilfunktion $\Phi(t)$ von $\mathcal{N}(0,1)$ und benutzen dass $\frac{X-\mu}{\sigma} \sim \mathcal{N}(0,1)$.

Es gilt $\Phi(-x) = 1 - \Phi(x)$ sowie $\Phi^{-1}(x) = -\Phi^{-1}(1-x)$

Zudem gilt für $X \sim N(\mu_X, \sigma_X^2)$, $Y \sim N(\mu_Y, \sigma_Y^2)$ mit $Z = X + Y$ dass $Z \sim N(\mu_X + \mu_Y, \sigma_X^2 + \sigma_Y^2)$.

Ist $X \sim \mathcal{N}(\mu, \sigma^2)$ so gilt für $Z = \frac{X-\mu}{\sigma}$ dass $Z \sim \mathcal{N}(0,1)$

Paretoverteilung

Die Paretoverteilte Variable $X \sim \mathrm{Par}(x_0, \alpha)$ mit

$$f_X(t) = \begin{cases} \frac{\alpha x_0^\alpha}{x^{\alpha+1}} & \text{für } x \ge x_0 \\ 0 & \text{sonst} \end{cases} \qquad F_X(t) = \begin{cases} 1 - \left(\frac{x}{x_0}\right)^{-\alpha} & \text{für } x \ge x_0 \\ 0 & \text{sonst} \end{cases}$$

$$\mathrm{E}[X] = \begin{cases} x_0 \frac{\alpha}{\alpha-1} & \text{für } \alpha > 1 \\ \infty & \text{sonst} \end{cases} \qquad \mathrm{Var}[X] = \begin{cases} \frac{x_0^2 \alpha}{(\alpha-1)^2(\alpha-2)} & \text{für } \alpha > 2 \\ \infty & \text{für } 1 < \alpha \le 2 \end{cases}$$

Diese Verteilung modelliert zum Beispiel die Krankenkosten eines Versicherten pro Jahr.

Bekannte Verteilungen

$$X_i \overset{iid}{\sim} \mathcal{N}(\mu,\sigma^2) \Rightarrow S_n \sim \mathcal{N}(n\cdot\mu, n\cdot\sigma^2)$$

Erwartungswerte und Varianzen

$$\begin{aligned} \mathrm{E}(S_n) &= n\cdot \mathrm{E}(X_1) & \mathrm{E}(\overline{X}_n) &= \mathrm{E}(X_1)\\ \mathrm{Var}(S_n) &= n\cdot \mathrm{Var}(X_1) & \mathrm{Var}(\overline{X}_n) &= \tfrac{1}{n}\mathrm{Var}(X_1)\\ \sigma_{S_n} &= \sqrt{n}\cdot\sigma_{X_1} & \sigma_{\overline{X}_n} &= \tfrac{1}{\sqrt{n}}\cdot\sigma_{X_1} \end{aligned}$$

5. Kennzahlen von Zufallsvariablen

Erwartungswert

- Diskret

$$\mu = \mathrm{E}[X] = \sum_{x_i\in\mathcal{W}(X)} x_i\cdot P(X=x_i) = \sum_{x_i\in\mathcal{W}(X)} x_i\cdot p(x_i)$$

$$\mathrm{E}(g(X)) = \sum_{x_i\in\mathcal{W}(X)} g(x_i)\cdot P(X=x_i)$$

- Stetig:

$$\mu = \mathrm{E}[X] = \int_{-\infty}^{\infty} x\cdot f_X(x)\,\mathrm{d}x$$

- Allgemein:

$$\mathrm{E}(aX+b) = a\,\mathrm{E}[X]+b$$

$$\mathrm{E}\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n \mathrm{E}(X_i)$$

Sind X_i, X und Y unabhängig, so gilt

$$\mathrm{E}(XY) = \mathrm{E}[X]\,\mathrm{E}(Y)$$

$$\mathrm{E}\left(\prod X_i\right) = \prod \mathrm{E}(X_i)$$

Bedingte Erwartung Die Erwartung von Y gegeben, dass $X=x$ bereits eingetroffen ist, ist definiert durch:

$$\mathrm{E}[Y|X=x] = \sum_y yP(Y=y\mid X=x) = \sum_y y\frac{P(X=x,Y=y)}{P(X=x)}$$

$$\mathrm{E}[Y|X=x] = \int y\cdot f_{Y|X=x}(y)\,dy = \int y\cdot\frac{f_{X,Y}(x,y)}{f_{X=x}(y)}\,dy$$

Beispiel (nicht konvergenter Erwartungswert): Sei X eine Cauchy-verteilte Zufallsvariable, mit $\mathcal{W}(X)=\mathbb{R}$ und

$$f_X(t) = \frac{1}{\pi}\frac{1}{1+t^2} \qquad F_X(t) = \frac{1}{2}+\frac{1}{\pi}\arctan(t)$$

Dann gilt:

$$\int_{-\infty}^{\infty}|x|f_X(x)dx = \lim_{b\to\infty}\frac{1}{\pi}\log(1+b^2) = +\infty$$

Varianz

- Diskret

$$\mathrm{Var}[X] = \sum_{x_i\in\mathcal{W}(X)}(x_i-\mu)^2\cdot P(X=x_i)$$

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

- Stetig:

$$\mathrm{Var}[X] = \mathrm{E}[X^2]-\mathrm{E}[X]^2 = \int_{-\infty}^{\infty}(x-\mathrm{E}[X])^2\cdot f_X(x)\,\mathrm{d}x$$

- Allgemein:

$$\begin{aligned} \mathrm{Var}(aX+b) &= a^2\cdot\mathrm{Var}[X]\\ \mathrm{Var}[X] &= E(X^2)-E[X]^2\\ \mathrm{Var}[X] &= E((X-\mu)^2)\\ \mathrm{Var}(aX+bY+c) &= a^2\,\mathrm{Var}[X]+b^2\,\mathrm{Var}(Y)+2ab\,\mathrm{Cov}(X,Y) \end{aligned}$$

Für unkorrelierte Zufallsvariablen X_i gilt

$$\mathrm{Var}\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n \mathrm{Var}(X_i)$$

$$\mathrm{Var}(\overline{X}_n) = \frac{\sigma^2}{n}$$

Standardabweichung

$$\sigma = \sqrt{\mathrm{Var}[X]}$$

$$\sigma_{aX+b} = |a|\cdot\sigma_X$$

α-Quantile, Median

Sei $0<\alpha<1$

$$P(X\le q(\alpha)) = \alpha \quad\Longrightarrow\quad F_X(q(\alpha)) = \alpha$$

$$\Longrightarrow\quad q(\alpha) = F_X^{-1}(\alpha)$$

Median: $\frac{1}{2}$-Quantil

Lineare Transformation

$$Y = aX+b \quad\Longrightarrow\quad f_Y(x) = \frac{1}{|a|}f_X\left(\frac{x-b}{|a|}\right)$$

Kovarianz und Korrelation

Kovarianz

$$\begin{aligned} \mathrm{Cov}(X,Y) &= \mathrm{E}\left((X-\mathrm{E}[X])\,(Y-E(Y))\right)\\ &= \mathrm{E}[XY]-\mathrm{E}[X]\,\mathrm{E}[Y]\\ \mathrm{Cov}(X,X) &= \mathrm{Var}(X) \end{aligned}$$

Korrelation

$$\mathrm{Corr}(X,Y) = \rho_{XY} = \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)}\cdot\sqrt{\mathrm{Var}(Y)}} = \frac{\mathrm{Cov}(X,Y)}{\sigma_X\cdot\sigma_Y}$$

$\rho_{XY}\in[-1,1]$ als dimensionsloses Mass des linearen Zusammenhanges

$$\rho_{XY} = \pm 1 \iff Y = a\pm bX \quad \text{mit } b>0, a\in\mathbb{R}$$

Rechenregeln

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

$\mathrm{Cov}(\cdot,\cdot)$ ist bilinear:

$$\begin{aligned} \mathrm{Cov}(a+bX,c+dY) &= b\cdot d\cdot\mathrm{Cov}(X,Y)\\ \mathrm{Cov}(X+Y,Z) &= \mathrm{Cov}(X,Z)+\mathrm{Cov}(Y,Z)\\ \mathrm{Var}(X+Y) &= \mathrm{Var}(X)+\mathrm{Var}(Y)+2\cdot\mathrm{Cov}(X,Y) \end{aligned}$$

Cov ist symmetrisch

Sind die Zufallsvariablen X und Y unabhängig, so gilt $\mathrm{Cov}(X,Y)=\mathrm{Corr}(X,Y)=0$. Die Umkehrung gilt jedoch im Allgemeinen nicht.

unabhängig $\Longrightarrow$ paarweise unabhängig $\Longrightarrow$ unkorreliert

Beispiel Unkorreliert aber nicht unabhängig: Sei $X\sim\mathcal{N}(0,1)$ und $Y=X^2$. Dann gilt $\mathrm{E}[XY]=\mathrm{E}[X^3]=0=\mathrm{E}[X]\,\mathrm{E}[Y]$. Somit sind X,Y unkorreliert $\to$ offensichtlich aber nicht unabhängig.

Transformation von Zufallsvariablen Sei X eine Zufallsvariable und $Y=g(X)$ eine weitere. Ist X stetig mit Dichte $f_X(x)$ so ist:

$$\mathrm{E}[Y] = \mathrm{E}[g(X)] = \int_{-\infty}^{+\infty} g(x)f_X(x)dx$$

Falls X,Y zwei stetige Zufallsvariablen sind $g(x,y):\mathbb{R}^2\to\mathbb{R}$ eine Funktion, dann gilt:

$$\mathrm{E}[g(X,Y)] = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g(x,y)f_{X,Y}(x,y)dxdy$$

Mehrere i.i.d Zufallsvariablen

Seien $X_1,\dots,X_n \overset{iid}{\sim} F$. Dann ist $Y=g(X_1,\dots,X_n)$ mit $g:\mathbb{R}^n\longrightarrow\mathbb{R}$ eine neue Zufallsvariable. Wichtige Vertreter sind die Summe von Zufallsvariablen und die relative Häufigkeit.

$$S_n = X_1+\ldots+X_n \qquad \overline{X}_n = \frac{1}{n}S_n$$

Die Verteilungen von S_n und $\overline{X}_n$ sind im allgemeinen aber schwierig zu bestimmen.

Bekannte Verteilungen

$$X_i \overset{iid}{\sim} \text{Exponential}(\lambda) \Rightarrow S_n \sim \text{Exponential}(n\cdot\lambda)$$
$$X_i \overset{iid}{\sim} \text{Poisson}(\lambda_i) \Rightarrow S_n \sim \text{Poisson}(\textstyle\sum\lambda_i)$$
$$X_i \overset{iid}{\sim} \text{Bernoulli}(p) \Rightarrow S_n \sim \text{Binomial}(n,p)$$
$$X_i \overset{iid}{\sim} \mathcal{N}(\mu,\sigma^2) \Rightarrow S_n \sim \mathcal{N}(n\cdot\mu, n\cdot\sigma^2)$$

Erwartungswerte und Varianzen

$$\mathrm{E}(S_n) = n\cdot\mathrm{E}(X_1) \qquad \mathrm{E}(\overline{X}_n) = \mathrm{E}(X_1)$$
$$\mathrm{Var}(S_n) = n\cdot\mathrm{Var}(X_1) \qquad \mathrm{Var}(\overline{X}_n) = \tfrac{1}{n}\mathrm{Var}(X_1)$$
$$\sigma_{S_n} = \sqrt{n}\cdot\sigma_{X_1} \qquad \sigma_{\overline{X}_n} = \frac{1}{\sqrt{n}}\cdot\sigma_{X_1}$$

6. Ungleichungen und Grenzwertsätze

Wir bezeichnen in der folgenden Diskussion $X_i,\dots,X_n$ Zufallsvariablen. (welche meistens i.i.d sind). Dann definieren wir

$$S_n = \sum_{i=1}^n X_i \qquad \overline{X_n} = \frac{1}{n}S_n = \frac{1}{n}\sum_{i=1}^n X_i$$

i.i.d. Zufallsvariablen Dies sind Zufallsvariablen die unabhängig und identisch verteilt sind.

Allgemeine Ungleichungen

Markov Ungleichung Sei X eine Zufallsvariable und $g : W(X) \to [0,\infty)$ eine wachsende Funktion. Für jedes $c \in \mathbb{R}$ gilt dann

$$P[X \geq c] \leq \frac{\mathrm{E}[g(X)]}{g(c)}$$

Chebyshev Ungleichung Sei X eine Zufallsvariable mit endlicher Varianz. Für jedes $b > 0$ gilt:

$$P[|X - \mathrm{E}[X]| \geq b] \leq \frac{\mathrm{Var}[X]}{b^2}$$

Schwache Gesetz der grossen Zahlen Sei $X_1, X_2, \dots$ eine Folge von *unabhängigen* Zufallsvariablen mit gleichem Erwartungswert $\mathrm{E}[X_i] = \mu$ sowie gleicher Varianz $\mathrm{Var}[X_i] = \sigma^2$.

Sei $\overline{X}_n = \frac{1}{n}S_n = \frac{1}{n}\sum_{i=1}^n X_i$. Dann konvergiert $\overline{X}_n$ für $n \to \infty$ stochastisch gegen $\mu = \mathrm{E}[X_i]$, respektive:

$$P[|\overline{X}_n - \mu| > \epsilon] \overset{n\to\infty}{\longrightarrow} 0 \ \text{ für jedes } \epsilon > 0$$

Bemerkung: Es reicht wenn die X_i nur paarweise unkorreliert sind (d.h. $\mathrm{Cov}(X_i, X_k) = 0$ für $i \neq k$)

Starke Gesetz der grossen Zahlen Sei $X_1, X_2, \dots$ eine Folge von *unabhängigen* Zufallsvariablen die alle die gleiche Verteilung haben mit Erwartungswert $\mathrm{E}[X_i] = \mu$ endlich. Für $\overline{X}_n = \frac{1}{n}S_n = \frac{1}{n}\sum_{i=1}^n X_i$ gilt dann

$$\overline{X_n} \overset{n\to\infty}{\longrightarrow} \mu \quad \text{P-fastsicher.}$$

respektive

$$P\left[\{\omega \in \Omega | \overline{X_n}(\omega) \overset{n\to\infty}{\longrightarrow} \mu\}\right] = 1$$

Unterschied zwischen starkem und schwachem Gesetz der grossen Zahlen: Beim ersten Gesetz ist die Wahrscheinlichkeit, dass $|\overline{X}_n - \mu| > \epsilon$ nie 0 sondern nur asymptotisch mit $n \to \infty$ 0. Somit gibt es eine sehr kleine Wahrscheinlichkeit, dass $|\overline{X}_n - \mu| > \epsilon$. Beim starken Gesetz ist jedoch die Wahrscheinlichkeit, dass dies passiert 0. In anderen Worten wenn wir unendlich viele $\overline{X_n}$ anschauen werden nur endlich viele nicht nach μ konvergieren.

Zusammengefasst: Beim schwachen ist die Chance von Nichtkonvergenz sehr klein, beim zweiten 0.

Der zentrale Grenzwertsatz

Zentraler Grenzwertsatz Sei $X_1, X_2, \dots$ eine Folge von *i.i.d Zufallsvariablen (independent and identically distributed random variables)* mit $\mathrm{E}[X_i] = \mu$ und $\mathrm{Var}[X_i] = \sigma^2$. Für die Summe $S_n = \sum_{i=1}^n X_i$, respektive $\overline{X}_n = \frac{X_1+\dots+X_n}{n}$ gilt dann:

$$\lim_{n\to\infty} P\left[\frac{S_n - n\mu}{\sigma\sqrt{n}} \leq x\right] = \lim_{n\to\infty} P\left[\frac{\overline{X}_n - \mu}{\sigma/\sqrt{n}} \leq x\right] = \Phi(x)$$

Wobei Φ die Verteilungsfunktion von $\mathcal{N}(0,1)$ ist.

In Praxis definiert man $S_n^* = \frac{S_n - n\mu}{\sigma\sqrt{n}} = \frac{S_n - \mathrm{E}[S_n]}{\sqrt{\mathrm{Var}[S_n]}}$ da $\mathrm{E}[S_n] = n\mu$ sowie $\mathrm{Var}[S_n] = n\sigma^2$. Als auch definieren wir $\overline{X_n} = \frac{1}{n}S_n$.

Nun gilt

$$P[S_n^* \leq x] \approx \Phi(x) \quad \text{für } n \text{ gross}$$
$$S_n^* \overset{\text{approx}}{\sim} \mathcal{N}(0,1)$$
$$S_n \overset{\text{approx}}{\sim} \mathcal{N}(n\mu, n\sigma^2)$$
$$\overline{X_n} \overset{\text{approx}}{\sim} \mathcal{N}(\mu, \frac{1}{n}\sigma^2)$$

Kontinuitätskorrektur
Angenommen $S_n \overset{\text{approx.}}{\sim} \mathcal{N}(np, np(1-p))$ sowie

$$P[a < S_n \leq b] = P\left[\frac{a - np}{\sqrt{np(1-p)}} < S_n^* \leq \frac{b - np}{\sqrt{np(1-p)}}\right]$$
$$\approx \Phi\left(\frac{b + \frac{1}{2} - np}{\sqrt{np(1-p)}}\right) - \Phi\left(\frac{a + \frac{1}{2} - np}{\sqrt{np(1-p)}}\right)$$

wobei $S_n^* \sim \mathcal{N}(0,1)$, so haben wir die Untere Schranke um $\frac{1}{2}$ nach oben verschoben, da a nicht drin ist, und b auch, da b drin ist. Intuitiv wollen wir damit die Stäbe aus dem Histogramm der Binomialverteilung zentriert über die Werte von a bis b setzten.

Grosse Abweichungen und Chernoff-Schranken

Momenterzeugende Funktion Wir definieren die momenterzeugende Funktion einer Zufallsvariable X als

$$M_X(t) = \mathrm{E}\left[e^{tX}\right]$$

welche auf $[0,\infty)$ wohldefiniert ist aber unendlich gross werden kann.
Aufgrund der Definition gilt:

- Diskret: $M_X(t) = \sum_{i=1}^\infty e^{tx_i}p_i$
- Stetig: $M_X(t) = \int_{-\infty}^\infty e^{tx}f(x)\,dx$

Beispiel Sei $X \sim Exp(\lambda)$. Dann gilt $M_X(t) = \mathrm{E}[e^{tX}] = \int_0^\infty e^{tx}\lambda e^{-\lambda x}dx = \lambda\int_0^\infty e^{-(\lambda-t)x}dx = -\frac{\lambda}{\lambda-t}e^{-(\lambda-t)x}\,|_0^\infty$

Abschätzung mit momenterzeugender Funktion Seien $X_1,\dots,X_n$ i.i.d. Zufallsvariablen für welche die momenterzeugende Funktion $M_X(t)$ für alle $t \in \mathbb{R}$ endlich ist. Für jedes b gilt dann:

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

Chernoff Schranke Seien $X_1,\dots,X_n$ unabhängig mit $X_i \sim Be(p_i)$ und $S_n = \sum_{i=1}^n X_i$. Sei ferner $\mu_n = \mathrm{E}[S_n] = \sum_{i=1}^n p_i$ und $\delta > 0$. Dann gilt:

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

Alternativ gibt es auch folgende Abschätzungen: (nicht im Skript!)

$$P[X \geq (1+\delta)\mu] \leq e^{-\frac{\delta^2}{2+\delta}\mu} \ \text{ für alle } \delta > 0$$
$$P[X \leq (1-\delta)\mu] \leq e^{-\mu\delta^2/2} \ \text{ für } 0 < \delta < 1$$

Teil II
Statistik

7. Schätzer

Grundbegriffe

- $\vartheta = (\vartheta_1, \dots, \vartheta_n)$ ist ein Vektor an Parameter die wir gerne Schätzen möchten. Sie müssen Teil von der Dichtefunktion sein.
- $T = (T_1, \dots, T_n)$ sind Schätzer, das sind Zufallsvariablen die ϑ bestmöglichst schätzen. Dabei gilt $T_i = t_i(X_1, \dots, X_n)$ für ein geeignetes $t_i : \mathbb{R}^n \to \mathbb{R}$
- Wir nennen $x_1, \dots, x_n$ Daten/Realisationen von Zufallsvariablen mit $x_i = X_i(\omega)$.
- $T_i(\omega) = t_i(X_1(\omega), \dots, X_n(\omega)) = t_i(x_1, \dots, x_n)$ heisst Schätzwert.

Gütebegriffe

Erwartungstreue & Bias

$$\mathrm{E}_\vartheta[T] = \vartheta \iff \text{Erwartungstreu}$$
$$bias = \mathrm{E}_\vartheta[T] - \vartheta$$

Konsistenz Wir nennen eine Folge von Schätzern $T^{(n)}$ konsistent für ϑ falls

$$\lim_{n\to\infty} P_\vartheta[|T^{(n)} - \vartheta| > \epsilon] = 0 \quad \forall \epsilon > 0$$

$\rightarrow$ versuche es mit Chebychev

Mean Squared Error (MSE)

$$MSE_\vartheta[T] = \mathrm{E}_\vartheta[(T-\vartheta)^2] = \mathrm{Var}_\vartheta[T] + (E_\vartheta[T] - \vartheta)^2$$
$$= \mathrm{Var}_\vartheta[T] + (bias^2)$$

Maximum-Likelihood-Methode

Maxmium-Likelihood-Schätzer Seien $X_1, \dots X_n$ Zufallsvariablen von n Stichproben und $x_1, \dots, x_n$ die Realisationen dieser Zufallsvariablen.
Wir definieren die Likelihood-Funktion als

$$L(x_1, \dots x_n; \vartheta) = \begin{cases} p(x_1, \dots, x_n; \vartheta) & \text{diskreter Fall} \\ f(x_1, \dots, x_n; \vartheta) & \text{stetiger Fall} \end{cases}$$

Wenn die $X_1, \dots, X_n$ i.i.d. sind, so gilt $p(x_1, \dots, x_n; \vartheta) = \prod_{i=1}^n p_X(x_i; \vartheta)$ respektive $f(x_1, \dots, x_n; \vartheta) = \prod_{i=1}^n f_X(x_i; \vartheta)$
Wenn die $X_1, \dots X_n$ i.i.d. sind, dann verwenden wir oft die log-likelihood-Funktion $\log(L)$ welche das $\prod$ in ein $\sum$ verwandelt.
Nun maximiere die Funktion (meistens reicht die Nullstelle): Die gibt dann z.B. eine Funktion wie $\vartheta = \frac{1}{n}\sum_{i=1}^n x_i$, wir ersetzten nun x_i durch X_i und ϑ durch T zu $T = \frac{1}{n}\sum_{i=1}^n X_i$

Beispiel 1:
Seien $X_1, \dots, X_n$ i.i.d und das Modell $P_\theta \sim Poisson(\theta)$. Wir bauen log-Likelihood Funktion:

$$\log(L(x_1, \dots, x_n)) = \log\left(\prod_{i=1}^n e^{-\theta}\frac{\theta^{x_i}}{x_i!}\right)$$
$$= -\theta n + \log(\theta)\sum_{i=1}^n x_i - \sum_{i=1}^n \log(x_i!)$$

nun leiten wir ab:

$$\frac{\partial}{\partial\theta}\log(L(x_1, \dots, x_n)) = -n + \frac{\sum_{i=1}^n x_i}{\theta} \overset{!}{=} 0$$
$$\implies \theta = \frac{\sum_{i=1}^n x_i}{n}$$

Beispiel 2 (nicht ableitbar):
Sei $f_X(x) = \begin{cases} e^{\alpha - x} & x \geq \alpha \\ 0 & \text{sonst} \end{cases}$ Wir möchten den MLS berechnen:

$$L(\vec{x}, \theta) = \prod_{i=1}^n f_\alpha(x_i) = \prod_{i=1}^n e^{\alpha - x_i} 1_{x_i \geq \alpha}$$
$$= \exp\left(n\alpha - \sum_{i=1}^n x_i\right) \cdot \prod_{i=1}^n 1_{x_i \geq \alpha}$$
$$= \exp\left(n\alpha - \sum_{i=1}^n x_i\right) \cdot 1_{\min x_i \geq \alpha}$$
$$= \begin{cases} \exp(n\alpha - \sum_{i=1}^n x_i) & \alpha \leq \min x_i \\ 0 & \text{sonst} \end{cases}$$

Somit ist der MLS $\widehat{\alpha} = \min x_i$, da dann α den obigen Ausdruck maximiert.

Momentenschätzer

Momentenschätzer
Die Idee ist, dass wir ein Gleichungssystem auflösen bei welchen:

- Das j-te Moment $m_j(\vartheta) = \mathrm{E}_\vartheta(X_1^j)$
- Das j-te Stichprobenmoment $m_j(x) = \frac{1}{n}\sum_{i=1}^n x_i^j$

Dann setzen wir die Momente gleich und lösen nach ϑ auf.

Es gilt bekanntlicherweise dass $\mathrm{E}[X] = \mu$ sowie $\mathrm{E}[X^2] = \mu^2 + \sigma^2$. Wir können somit das folgende System auflösen:

$$\frac{x_1 + \ldots + x_n}{n} = \mu$$
$$\frac{x_1^2 + \ldots + x_n^2}{n} = \mu^2 + \sigma^2$$

Zentrale Lemmata Seien $X_1, \dots, X_n$ i.i.d. $\sim \mathcal{N}(\mu, \sigma^2)$ dann gilt:

- $\overline{X_n}$ ist Normalverteilt $\sim \mathcal{N}(\mu, \frac{1}{n}\sigma^2)$ und $\frac{\overline{X_n} - \mu}{\sigma/\sqrt{n}} \sim \mathcal{N}(0,1)$
- $\frac{n-1}{\sigma^2}S^2 = \frac{1}{\sigma^2}\sum_{i=1}^n (X_i - \overline{X_n})$ ist χ^2 Verteilt mit $n-1$ Freiheitsgraden.
- $\overline{X_n}$ und S^2 sind unabhängig.
- Der Quotient
$$\frac{\overline{X_n} - \mu}{S/\sqrt{n}} = \frac{\frac{\overline{X_n} - \mu}{\sigma/\sqrt{n}}}{S/\sigma} = \frac{\frac{\overline{X_n} - \mu}{\sigma/\sqrt{n}}}{\sqrt{\frac{1}{n-1}\frac{n-1}{\sigma^2}S^2}}$$
ist t-verteilt mit $n-1$ Freiheitsgraden.

Dabei galt $\overline{X_n} = \frac{1}{n}\sum_{i=1}^n X_i$ sowie $S^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \overline{X_n})^2$ die empirische Stichprobenvarianz.

8. Tests

Für einen Test definieren wir:

- Hypothese $H_0 : \vartheta \in \Theta_0$
- Alternative $H_A : \vartheta \in \Theta_A$

Wir bauen eine Teststatistik T. Wir verwerfen die Hypothese wenn $T(\omega) \in K$ liegt, wobei K ein Intervall, der Verwerfungsbereich ist.

- Signifikanzniveau $\alpha \in (0,1)$ wird gewählt mit $\sup_{\vartheta\in\Theta_0} P_\vartheta[T \in K] \leq \alpha$. Es gilt $\alpha = P[\text{Typ 1 Fehler} | H_0 \text{ ist wahr}]$
- Die Macht des Testes ist $\beta : \Theta_A \to [0,1]$ mit $\beta(\vartheta) = P_\vartheta[T \in K]$ mit $P[\text{Typ 2 Fehler}] = \beta$

mit

- **Typ 1 Fehler:** verwerfe H_0 obwohl sie stimmte
- **Typ 2 Fehler:** bestätige H_0 obwohl sie eigentlich falsch ist

Dabei gilt:
Zuerst wählt man das Signifikanzniveau dann maximiert man die Macht des Testes.

Macht Berechnen: Beispiel

Angenommen wir haben einen Test $T = \frac{\overline{X}-\mu_0}{\sigma/\sqrt{9}}$ sowie ein Verwerfungsbereich $K = [1.28, \infty)$ (einseitiger Test) mit $H_0 : \mu = 84$. Nun wollen wir die Macht des Testes zugunsten der konkreten Alternativhypothese $H_A : \mu = 85$ testen.
Nun rechnen wir

$$\begin{aligned} P_{\mu=85}[T \in K] &= P_{\mu=85}[T \geq 1.28] \\ &= P_{\mu=85}\left[\frac{\overline{X}-85}{\sigma/\sqrt{9}} - \frac{\mu_0 - 85}{\sigma/\sqrt{9}} \geq 1.28\right] \\ &= P_{\mu=85}\left[\frac{\overline{X}-85}{\sigma/\sqrt{9}} \geq 0.28\right] \\ &= 1 - \Phi(0.28) \approx 38.97\%; \end{aligned}$$

Kommentar: Mit $\frac{\overline{X}-85}{\sigma/\sqrt{9}} - \frac{\mu_0-85}{\sigma/\sqrt{9}}$ verschieben wir die Verteilung wie gewollt, so dass sie nun über μ_2 zentriert ist. Der Fehler 2ter Art ist (1 - Macht) oder $P_\theta[T \notin K]$

Neyman-Pearson-Lemma Wenn $\Theta_0 = \{\vartheta_0\}$, $\Theta_A = \{\vartheta_A\}$, $K = [0, c)$ sowie $\alpha^* = P_{\vartheta_0}[T \in K] = P_{\vartheta_0}[T < c]$.
Jeder andere Test als

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

der Likelihood Koeffizient, mit Signifikanzniveau $\alpha \leq \alpha'$ hat kleinere Macht.

Statistische Testverfahren

***z*-Test** Ein Test für den Erwartungswert bei bekannter Varianz. Seien $X_1, \ldots, X_n \sim \mathcal{N}(\mu, \sigma^2)$ unter P_ϑ mit bekanntem σ^2. Es ist $H_0 : \vartheta = \vartheta_0$ Möglich Variationen sind:

- $H_A : \vartheta > \vartheta_0$: $K = (c_>, \infty)$, $c_> = z_{(1-\alpha)}$
- $H_A : \vartheta < \vartheta_0$: $K = (-\infty, -c_<)$, $c_< = z_\alpha$.
- $H_A : \vartheta \neq \vartheta_0$: $K = (-\infty, -c_{\neq}) \cup (c_{\neq}, \infty)$, $c_{\neq} = z_{1-\frac{\alpha}{2}}$.

Mit $T = \frac{\overline{X_n}-\vartheta_0}{\sigma/\sqrt{n}} \sim \mathcal{N}(0,1)$ sowie $z_x = \Phi^{-1}(x)$.

Beachte dass $\overline{X_n} = \frac{1}{n}\sum_{i=1}^n X_i$

***t*-Test** Ein Test für den Erwartungswert bei unbekannter Varianz. Seien $X_1, \ldots, X_n \sim \mathcal{N}(\mu, \sigma^2)$ unter $P_{\vec{\vartheta}}$ mit $\vec{\vartheta} = (\mu, \sigma^2)$. Wir haben $\Theta_0 = \{\sigma\} \times (0, \infty)$. Die Teststatistik ist gegeben durch $T = \frac{\overline{X_n}-\mu_0}{S/\sqrt{n}} \sim t_{n-1}$ mit $S^2 = \frac{1}{n-1}\sum_{i=1}^n \left(X_i - \overline{X_n}\right)^2$ für σ^2

- $H_A : \vartheta > \vartheta_0$: $K = (c_>, \infty)$, $c_> = t_{n-1,1-\alpha}$
- $H_A : \vartheta < \vartheta_0$: $K = (-\infty, c_<)$, $c_< = t_{n-1,\alpha} = -t_{n-1,1-\alpha}$.
- $H_A : \vartheta > \vartheta_0$: $K = (-\infty, -c_{\neq}) \cup (c_{\neq}, \infty)$, $c_{\neq} = t_{n-1,1-\frac{\alpha}{2}}$.

wobei $t_{m,\gamma}$ das γ-Quantil der t_m Verteilung ist.

Gepaarte Zweistichproben-Test bei Normalverteilung Wenn $X_1, \ldots, X_n$ i.i.d $\sim \mathcal{N}(\mu_X, \sigma^2)$ sowie $Y_1, \ldots, Y_n$ i.i.d $\sim \mathcal{N}(\mu_Y, \sigma^2)$ mit gleicher Anzahl Proben und gleicher Varianz, so definiert man $Z_i = X_i - Y_i$ die unter $P\vartheta$ i.i.d zu $\sim \mathcal{N}(\mu_X - \mu_Y, 2\sigma^2)$ sind. Dann verwende entweder z-Test oder t-Test.

Ungepaarter Zweistichproben-Test bei Normalverteilung Seien unter $P\vartheta$ $X_1, \ldots, X_n$ i.i.d. $\sim \mathcal{N}(\mu_X, \sigma^2)$ sowie $Y_1, \ldots, Y_m$ i.i.d. $\sim \mathcal{N}(\mu_Y, \sigma^2)$ wobei $m \neq n$ sein kann.

- *Ungepaarter Zweistichproben-z-Test wenn σ^2 bekannt ist:* Es ist $T = \frac{(\overline{X_n}-\overline{Y_m})-(\mu_X-\mu_Y)}{\sigma\sqrt{\frac{1}{n}+\frac{1}{m}}} \sim \mathcal{N}(0,1)$
- *Ungepaarter Zweistichproben-t-Test wenn σ^2 unbekannt ist:* Wir haben die beiden empirischen Varianzen $S_X^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \overline{X_n})^2$ und $S_Y^2 = \frac{1}{n-1}\sum_{i=1}^n (Y_i - \overline{Y_n})^2$. Es gilt $S^2 = \frac{1}{m+n-2}\left((n-1)\,S_X^2 + (m-1)\,S_Y^2\right)$ mit der Teststatistik $T = \frac{(\overline{X_n}-\overline{Y_m})-(\mu_X-\mu_Y)}{S\sqrt{\frac{1}{n}+\frac{1}{m}}} \sim t_{n+m-2}$

9. Konfidenzbereiche

Konfidenzbereich Ein Konfidenzbereich ist eine Menge $C(X_1, \ldots, X_n) \subseteq \Omega$ zum Niveau $1-\alpha$ falls gilt

$$P_\vartheta\left[\vartheta \in C(X_1, \ldots, X_n)\right] \geq 1 - \alpha \quad \text{für alle } \vartheta \in \Omega$$

Bekannte Konfidenzbereiche

- μ unbekannt, σ^2 bekannt:
  - μ zum Niveau $1-\alpha$:
    $$\left[\overline{X_n} - z_{(1-\frac{\alpha}{2})}\frac{\sigma}{\sqrt{n}}\,;\ \overline{X_n} + z_{(1-\frac{\alpha}{2})}\frac{\sigma}{\sqrt{n}}\right]$$
- μ, σ^2 unbekannt:
  - μ zum Niveau $1-\alpha$:
    $$\left[\overline{X_n} - t_{n-1,1-\frac{\alpha}{2}}\frac{S}{\sqrt{n}}, \overline{X_n} + t_{n-1,1-\frac{\alpha}{2}}\frac{S}{\sqrt{n}}\right]$$
  - σ^2 zum Niveau $1-\alpha$:
    $$\left[\frac{(n-1)S^2}{\chi^2_{n-1,1-\frac{\alpha}{2}}}, \frac{(n-1)S^2}{\chi^2_{n-1,\frac{\alpha}{2}}}\right]$$

Konstruktion von Konfidenzintervallen
Angenommen wir haben 8 Gewichte gegeben die wir als Realisationen von $X_1, \ldots, X_8$ i.i.d. $\sim \mathcal{N}(\mu, \sigma^2)$. Als Schätzer verwenden wir $\mu = \overline{X_n} = \frac{1}{n}\sum_{i=1}^n X_i$ und als Stichprobenvarianz $S^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \overline{X_n})^2$.
Wir machen den Ansatz $C(X_1, \ldots, X_n) = [\overline{X_n} - \cdots, \overline{X_n} + \cdots]$ und wollen erreichen dass $1 - \alpha \leq P_\vartheta[\vartheta \in C(X_1, \ldots, X_n)] = P_\vartheta\left[\mu \in [\overline{X_n} - \cdots, \overline{X_n} + \cdots]\right] = P_\vartheta\left[|\overline{X_n} - \mu| \leq \cdots\right]$
Da $\frac{\overline{X_n}-\mu}{S/\sqrt{n}} \sim t_{n-1}$ brauchen wir $\frac{\cdots}{S/\sqrt{n}} = t_{n-1,1-\frac{\alpha}{2}}$.
Dies gibt uns dann das Konfidentintervall für μ zum Niveau $1-\alpha$:

$$C(X_1, \ldots, X_n) = \left[\overline{X_n} - t_{n-1,1-\frac{\alpha}{2}}\frac{S}{\sqrt{n}}, \overline{X_n} + t_{n-1,1-\frac{\alpha}{2}}\frac{S}{\sqrt{n}}\right]$$

Um ein Konfidenzintervall für σ^2 zu konstruieren verwenden wir $\frac{1}{\sigma^2}(n-1)S^2 = \frac{1}{\sigma^2}\sum_{i=1}^n (X_i - \overline{X_n})^2 \sim \chi^2_{n-1}$
Wir rechnen wieder $1 - \alpha = P_\vartheta\left[\chi^2_{n-1,\frac{\alpha}{2}} \leq \frac{1}{\sigma^2}(n-1)S^2 \leq \chi^2_{n-1,1-\frac{\alpha}{2}}\right] = P_\vartheta\left[\frac{(n-1)S^2}{\chi^2_{n-1,1-\frac{\alpha}{2}}} \leq \sigma^2 \leq \frac{(n-1)S^2}{\chi^2_{n-1,\frac{\alpha}{2}}}\right]$
Dies impliziert

$$C(X_1, \ldots, X_n) = \left[\frac{(n-1)S^2}{\chi^2_{n-1,1-\frac{\alpha}{2}}}, \frac{(n-1)S^2}{\chi^2_{n-1,\frac{\alpha}{2}}}\right]$$

10. Kombinatorik

Wir definieren die

- **Auf wie viele Arten kann man n Objekte (z.B. nebeneinander) anordnen?**
  Dies ist die Anzahl Permutationen von n Elementen und ist $n!$.

- **Auf wie viele Arten kann man k aus n Objekten auswählen (mit $k \leq n$ ohne Zurücklegen)?**
  Dies ist die Anzahl Kombinationen ist $\binom{n}{k} = \frac{n!}{k!(n-k)!}$.
- **Wie viele Sequenzen der Länge m kann ma mit den n Symbolen bilden?**
  Dies ist die Anzahl der Variationen (mit Wiederholung) und ist n^m

Teil III

Analysis

11. Ableitung

Ableitung Sei $D \subset \mathbb{R}$, $f : D \to \mathbb{R}$ und x_0 ein Häufungspunkt von D. f ist in x_0 differenzierbar, falls der Grenzwert

$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}$$

existiert. Ist dies der Fall, wird der Grenzwert mit $f'(x_0)$ bezeichnet.
Alternativ nutzt man auch $x = x_0 + h$

$$f'(x_0) = \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h}$$

Konvexität

Konvex $f : I \to \mathbb{R}$ ist konvex (auf I) falls für alle $x \leq y$ $x, y \in I$ und $\lambda \in [0, 1]$

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y)$$

Zudem gilt für $x_0 < x < x_1$ in I:

$$\frac{f(x) - f(x_0)}{x - x_0} \leq \frac{f(x_1) - f(x)}{x_1 - x}$$

Man beweist dies indem man $x = (1 - \lambda)x_0 + \lambda x_1$ wählt und somit $\lambda = \frac{x - x_0}{x_1 - x_0}$

Wichtige Taylorapproximationen um $x = 0$

- $\boxed{\frac{1}{1-x}}$ Für alle $x \in (1, 0)$ gilt:

$$\frac{1}{1-x} = 1 + x + x^2 + x^3 + x^4 + \cdots = \sum_{n=0}^{\infty} x^n$$

- $\boxed{e^x}$ Für alle $x \in \mathbb{R}$ gilt:

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} = \sum_{n=0}^{\infty} \frac{x^n}{n!}$$

- $\boxed{\cos(x)}$ Für alle $x \in R$ gilt:

$$\cos(x) = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \frac{x^8}{8!} - \cdots = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!}$$

- $\boxed{\sin(x)}$ Für alle $x \in R$ gilt:

$$\sin(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \frac{x^9}{9!} - \cdots = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!} = \sum_{n=1}^{\infty} (-1)^{(n-1)} \frac{x^{2n-1}}{(2n-1)!}$$

- $\boxed{\ln(1+x)}$ Für alle $x \in (-1, 1]$ gilt:

$$\ln(x+1) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \frac{x^5}{5} - \cdots = \sum_{n=1}^{\infty} (-1)^{(n+1)} \frac{x^n}{n}$$

- $\boxed{\arctan(x)}$ Für alle $x \in [-1, 1]$ gilt:

$$\arctan(x) = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \frac{x^9}{9} - \cdots = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{2n+1}$$

- $\boxed{(1+x)^\alpha}$ Für alle $x \in \mathbb{R}$ gilt:

$$(1+x)^\alpha = 1 + \alpha x + \frac{\alpha(\alpha-1)}{2!}x^2 + \cdots = \sum_{k=0}^{\infty} \binom{\alpha}{k} x^k$$

- $\boxed{\sinh(x)}$ Für alle $x \in \mathbb{R}$ gilt:

$$\sinh(x) = x + \frac{x^3}{6} + \frac{x^5}{120} + \mathcal{O}(x^7) = \sum_{k=0}^{\infty} \frac{x^{1+2k}}{(1+2k)!}$$

- $\boxed{\cosh(x)}$ Für alle $x \in \mathbb{R}$ gilt:

$$\cosh(x) = 1 + \frac{x^2}{2} + \frac{x^4}{24} + \frac{x^6}{720} + \mathcal{O}(x^7) = \sum_{k=0}^{\infty} \frac{x^{2k}}{(2k)!}$$

Fundamentalsatz

Fundamentalsatz der Differentialrechnung Sei $f : [a, b] \to \mathbb{R}$ stetig. Dann gibt es eine Stammfunktion F von f, die bis auf eine additive Konstante eindeutig bestimmt ist und es gilt:

$$\int_a^b f(x)dx = F(b) - F(a)$$

Beweis: Existenz folgt aus Stammfunktionssatz. Seien F_1, F_2 Stammfkt., dann gilt $F_1' - F_2' = 0$. Somit ist $F_1 - F_2 = C$ mit $F(x) = C + \int_a^x f(t)dt$. Es folgt auch $F(a) = \int_a^a f(t)dt + C$ und somit $F(a) = C$. Es folgt daraus $F(b) - F(a) = \int_a^b f(t)dt$ □

Ableitung des Integrals

Mit der Kettenregel folgt aus dem Fundamentalsatz:

$$\frac{d}{dx}\left(\int_{u(x)}^{v(x)} f(t)dt\right) = f(v)\frac{dv}{dx} - f(u)\frac{du}{dx}$$

Integrale Ausrechnen

Integrationskonstante C nicht vergessen!

Direkte Integrale

Diese sind vom Typ $\int f(g(x))g'(x)dx = F(g(x))$.

Partielle Integration

Partielle Integration

$$\int f' \cdot g \, dx = f \cdot g - \int f \cdot g' \, dx$$

Integrale rationaler Funktionen

Partielle Integration

$$\int \frac{p(x)}{q(x)}dx$$

Wenn nun $\deg(p) \geq \deg(q)$ dann machen wir eine Polynomdivision $p : q$, sonst mache eine Parzialbruchzerlegung

Substitutionsregel

Substitutionsregel Ist f stetig und g erfüllt:

$$y = g(x) \iff x = g^{-1}(y)$$

Dann gilt:

$$\int_a^b f(g(x))g'(x)dx = \int_{g(a)}^{g(b)} f(y)dy$$

Als Merksatz gilt $dy = g'(x)dx$ respektive $dx = \frac{1}{t}dt$

Integrale der Form $\int F(e^x, \sinh(x), \cosh(x))dx$

Substituiere mit $e^x = t$, $(dx = \frac{1}{t}dt)$
Beispiel:

$$\int \frac{e^{2x}}{e^x+1}dx = \int \frac{t^2}{t+1}\frac{1}{t}dt = \int \frac{t+1-1}{t+1}dt$$
$$\int \frac{1}{\cosh(x)}dx = \int \frac{1}{\frac{1}{2}(e^x+e^{-x})}dx = \int \frac{2}{t+\frac{1}{t}}\frac{1}{t}dt = \frac{2}{t^2+1}dt$$

Integrale der Form $\int F(\log(x))dx$

Substituiere mit $\log(x) = t$, $(dx = e^t dt)$
Beispiel:

$$\int (\log(x))^2 dx = \int t^2 e^t dt = t^2 e^t - \int 2te^t dt$$
$$= x(\log(x))^2 - 2x\log(x) + 2x + C$$

Integrale der Form $\int F(\sqrt[a]{Ax+B})dx$

Substituiere mit $t = \sqrt[a]{Ax+B}$
Beispiel:

$$\int \frac{1}{\sqrt{x}\sqrt{1-x}} = \int \frac{1}{t\sqrt{1-t^2}}2tdt = \int \frac{2}{\sqrt{1-t^2}}$$

Integrale die $\sin, \cos, \tan$ **in geraden Potenzen enthalten**

Substituiere mit $\tan(x) = t$, $(dx = \frac{1}{1+t^2}dt)$. Es gilt zudem:

$$\sin^2(x) = \frac{t^2}{1+t^2} \qquad \cos^2(x) = \frac{1}{1+t^2}$$

Integrale die $\sin, \cos, \tan$ **in ungeraden Potenzen enthalten**

Substituiere mit $\tan(\frac{x}{2}) = t$, $(dx = \frac{2}{1+t^2}dt)$. Es gilt zudem:

$$\sin(x) = \frac{2t}{1+t^2} \qquad \cos(x) = \frac{1-t^2}{1+t^2}$$

Integrale mit $\sqrt{Ax^2+Bx+C}$ **im Nenner**

Mithilfe quadratischer Ergänzung auf einen der folgenden Fälle zurückführen:

$$\int \frac{1}{\sqrt{1-x^2}}dx = \operatorname{arcsin}(x) + C$$
$$\int \frac{1}{\sqrt{x^2-1}}dx = \operatorname{arcosh}(x) + C$$
$$\int \frac{1}{\sqrt{1+x^2}}dx = \operatorname{arcsinh}(x) + C$$

Integrale mit $\sqrt{Ax^2+Bx+C}$ **im Zähler**

Mithilfe quadratischer Ergänzung auf einen der folgenden Fälle zurückführen, dann substituieren

$$\int \sqrt{1-x^2}dx \quad \text{subsitution: } x = \sin(t) \Leftarrow dx = \cos(t)dt$$
$$\int \sqrt{x^2-1}dx \quad \text{subsitution: } x = \cosh(t) \Leftarrow dx = \sinh(t)dt$$
$$\int \sqrt{1+x^2}dx \quad \text{subsitution: } x = \sinh(t) \Leftarrow dx = \cosh(t)dt$$

Sonstiges

Binomialsatz $\forall x, y \in \mathbb{C}$, $n \geq 1$ gilt:

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

Binomialkoeffizienten

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

ABC / Mitternachtsformel

Gegeben: $ax^2 + bx + c = 0$

Lösung: $x_{1,2} = \dfrac{-b \pm \sqrt{b^2-4ac}}{2a}$

Logarithmus Regeln

$$\log_b(x \cdot y) = \log_b(x) + \log_b(y)$$
$$\log_b(M^k) = k \cdot \log_b(M)$$

Summenformeln

$$\sum_{k=1}^{n} k = \frac{n(n+1)}{2}$$
$$\sum_{k=1}^{n} (2k-1) = n^2$$
$$\sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}$$

Gerade & Ungerade Funktion Eine Funktion heisst:

- GERADE wenn $f(-x) = f(x)$
- UNGERADE wenn $f(-x) = -f(x)$

Dabei sind $f(x) = 1$, $f(x) = |x|$, $f(x) = x^2$, $f(x) = \cos(x)$ alles gerade Funktionen.
Im Gegenzug sind $f(x) = sgn(x)$, $f(x) = x$, $f(x) = \tan(x)$, $f(x) = \sin(x)$ ungerade Funktionen.

Injektiv

$$\forall x_1, x_2 \in M : f(x_1) = f(x_2) \implies x_1 = x_2$$
$$\text{or } x_1 \neq x_2 \implies f(x_1) \neq f(x_2)$$

Surjektiv

$$\forall y \in N \; \exists x \in M : y = f(x)$$

Umkehrsatz - Beispiel Zeige dass $x + e^x$ bijektiv von $\mathbb{R}$ auf $\mathbb{R}$ abbildet. Es gilt $f'(x) = 1 + e^x > 0$, somit ist f streng monoton wachsend in $\mathbb{R}$ und Umkehrbar. Weil $\lim_{x\to-\infty} f(x) = -\infty$ und $\lim_{x\to\infty} f(x) = \infty$ ist f bijektiv von $\mathbb{R}$ nach $\mathbb{R}$

Kreuzprodukt

$$\vec{a} \times \vec{b} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \times \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} a_2 b_3 - a_3 b_2 \\ a_3 b_1 - a_1 b_3 \\ a_1 b_2 - a_2 b_1 \end{pmatrix}$$

Wichtige Integrale

- $\int \sin^2 ax\,dx = \frac{x}{2} - \frac{1}{4a}\sin 2ax + C = \frac{x}{2} - \frac{1}{2a}\sin ax\cos ax + C$
- $\int \sin^n ax\,dx = -\frac{\sin^{n-1} ax\cos ax}{na} + \frac{n-1}{n}\int \sin^{n-2} ax\,dx \qquad (\text{for } n > 0)$
- $\int x^n \sin ax\,dx = -\frac{x^n}{a}\cos ax + \frac{n}{a}\int x^{n-1}\cos ax\,dx$
- $\int \cos^2 ax\,dx = \frac{x}{2} + \frac{1}{4a}\sin 2ax + C = \frac{x}{2} + \frac{1}{2a}\sin ax\cos ax + C$
- $\int \cos^n ax\,dx = \frac{\cos^{n-1} ax\sin ax}{na} + \frac{n-1}{n}\int \cos^{n-2} ax\,dx \qquad (\text{for } n > 0)$
- $\int x^n \cos ax\,dx = \frac{x^n \sin ax}{a} - \frac{n}{a}\int x^{n-1}\sin ax\,dx$
- $\int (\sin ax)(\cos ax)\,dx = \frac{1}{2a}\sin^2 ax + C$
- $\int (\sin^n ax)(\cos ax)\,dx = \frac{1}{a(n+1)}\sin^{n+1} ax + C \qquad (\text{for } n \neq -1)$
- $\int (\sin ax)(\cos^n ax)\,dx = -\frac{1}{a(n+1)}\cos^{n+1} ax + C \qquad (\text{for } n \neq -1)$
- $$\int (\sin^n ax)(\cos^m ax)\,dx = -\frac{(\sin^{n-1} ax)(\cos^{m+1} ax)}{a(n+m)} + \frac{n-1}{n+m}\int (\sin^{n-2} ax)(\cos^m ax)\,dx \qquad (\text{for } m, n > 0)$$
- $\int \sin^2(x)\cos^2(x)dx = \frac{1}{4}\int sin^2(2x)dx = \frac{1}{4}\int \frac{1-\cos(4x)}{2}dx = \frac{x}{8} - \frac{1}{8}\frac{\sin(4x)}{4} + C$

Typische Integrale

- $\int \frac{1}{x}\,dx = \ln|x|$
- $\int \frac{1}{x^2}\,dx = -\frac{1}{x}$
- $\int \frac{1}{x+a}\,dx = \ln|x+a|$
- $\int \ln(x)\,dx = x(\ln(x)-1)$
- $\int \ln(ax+b)\,dx = \frac{(ax+b)\ln(ax+b)-ax}{a}$
- $\int \frac{1}{(x+a)^2}\,dx = -\frac{1}{x+a}$
- $\int \frac{1}{\sqrt{x}}\,dx = 2\sqrt{x}$
- $\int \sqrt{1-x^2}dx = \frac{\arcsin(x)+x\sqrt{1-x^2}}{2} + C$
- $\int \frac{1}{ax+b}\,dx = \frac{1}{a}\ln|ax+b|$
- $\int \frac{1}{1+x^2}\,dx = \frac{1}{2}\ln|1+x^2|$
- $\int (ax+b)^n\,dx = \frac{(ax+b)^{n+1}}{(n+1)a}, (n \neq -1)$
- $\int x(ax+b)^n\,dx = \frac{(ax+b)^{n+2}}{(n+2)a^2} - \frac{b(ax+b)^{n+1}}{(n+1)a^2}$
- $\int \frac{ax+b}{px+q}\,dx = \frac{ax}{p} + \frac{bp-aq}{p^2}\ln|pq+q|$

- $\int \frac{1}{a^2+x^2}\,dx = \frac{1}{a}\arctan(\frac{x}{a})$
- $\int \frac{1}{a^2-x^2}\,dx = \frac{1}{2a}\ln\left|\frac{a+x}{a-x}\right|$
- $\int \sqrt{x}\,dx = \frac{2}{3}\sqrt{x^3}$
- $\int \sqrt{1-x^2}\,dx = \frac{1}{2}\left(x\sqrt{1-x^2} + \frac{1}{\sin(x)}\right)$
- $\int a^{xb+c}\,dx = \frac{a^{bx+c}}{b\log(a)}$

Trionometrische Funktionen

- $\int \sin(ax)\,dx = -\frac{1}{a}\cos(ax)$
- $\int \cos(ax)\,dx = \frac{1}{a}\sin(ax)$
- $\int \sin(ax)^2\,dx = \frac{x}{2} - \frac{sin(2ax)}{4a}$
- $\int \frac{1}{\sin^2 x}\,dx = -\cot x$
- $\int x\sin(ax)\,dx = \frac{\sin(ax)}{a^2} - \frac{x\cos(ax)}{a}$
- $\int \cos^2(ax)\,dx = \frac{x}{2} + \frac{\sin(2ax)}{4a}$
- $\int \frac{1}{\cos^2(x)}\,dx = \tan x$
- $\int \cos(ax)\,dx = \frac{\cos(ax)}{a^2} + \frac{x\sin(ax)}{a}$
- $\int \sin(ax)\cos(ax)\,dx = -\frac{\cos^2(ax)}{2a}$
- $\int \tan(ax)\,dx = -\frac{1}{a}\ln|\cos(ax)|$

Exponentialfunktion

- $\int e^{ax}\,dx = \frac{1}{a}e^{ax}$
- $\int xe^{ax}\,dx = e^{ax}\cdot\left(\frac{ax-1}{a^2}\right)$
- $\int x\ln(x)\,dx = \frac{1}{2}x^2(\ln(x) - \frac{1}{2})$
- $\int_{-\infty}^{\infty} e^{-\frac{1}{a}x^2}\,dx = \sqrt{a\pi}$

Vektoranalysis

$$\Delta f = \operatorname{div}(\operatorname{grad} f),$$
$$\Delta f = \nabla\cdot(\nabla f) = (\nabla\cdot\nabla)f = \nabla^2 f.$$
$$\nabla f = \begin{pmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \\ \frac{\partial}{\partial z} \end{pmatrix}\cdot f$$

Integral der Normalverteilung

$$\int_{-\infty}^{\infty} e^{-a(x+b)^2}\,dx = \sqrt{\frac{\pi}{a}}$$

Funktion	Ableitung	Bemerkung / Regel
x	1	
x^2	$2x$	
x^n	$n\cdot x^{n-1}$	$n \in \mathbb{R}$
$\frac{1}{x} = x^{-1}$	$-\frac{1}{x^2}$	
$\sqrt{x} = x^{\frac{1}{2}}$	$\frac{1}{2\sqrt{x}}$	
$\sqrt[n]{x} = x^{\frac{1}{n}}$	$\frac{x^{\frac{1}{n}-1}}{n}$	$\int x^{1/n}dx = \frac{nx^{1/n+1}}{n+1} + C$
e^x	e^x	
a^x	$\ln(a)\cdot a^x$	
$x^x = e^{x\log(x)}$	$x^x\cdot(\log(x)+1)$	Kettenregel $e^{x\log(x)}$
$\ln(x)$	$\frac{1}{x}$	
$x\ln(x) - x$	$\ln(x)$	
$\sin(x)$	$\cos(x)$	
$\cos(x)$	$-\sin(x)$	
$\tan(x) = \frac{\sin(x)}{\cos(x)}$	$\frac{1}{\cos^2(x)}$ $1+\tan^2(x)$	$=$
$\cot(x) = \frac{\cos(x)}{\sin(x)}$	- $\frac{1}{\sin^2(x)}$	
$\arcsin(x)$	$\frac{1}{\sqrt{1-x^2}}$	$\arcsin: [-1,1] \to [-\frac{\pi}{2}, \frac{\pi}{2}]$
$\arccos(x)$	$-\frac{1}{\sqrt{1-x^2}}$	$\arccos: [-1,1] \to [0,\pi]$
$\arctan(x)$	$\frac{1}{1+x^2}$	$\arctan: (-\infty,\infty) \to (-\frac{\pi}{2}, \frac{\pi}{2})$
$\operatorname{arccot}(x)$	$-\frac{1}{1+x^2}$	$\operatorname{arccot}: (-\infty,\infty) \to (0,\pi)$
$\cosh(x)$	$\sinh(x)$	
$\sinh(x)$	$\cosh(x)$	
$\tanh(x)$	$\frac{1}{\cosh^2(x)}$	
$\operatorname{arsinh}(x)$	$\frac{1}{\sqrt{1+x^2}}$	$\forall x \in R$
$\operatorname{arcosh}(x)$	$\frac{1}{\sqrt{x^2-1}}$	$\forall x \in (1,\infty)$
$\operatorname{artanh}(x)$	$\frac{1}{1-x^2}$	$\forall x \in (-1,1)$
$g(x)\cdot h(x)$	$g(x)\cdot h'(x) + g'(x)\cdot h(x)$	Produktregel
$(g(x))^n$	$n\cdot(g(x))^{n-1}\cdot g'(x)$	Potenzregel
$\frac{g(x)}{h(x)}$	$\frac{g'(x)\cdot h(x)-g(x)\cdot h'(x)}{(h(x))^2}$	Quotientenregel
$h(g(x))$	$h'(g(x))\cdot g'(x)$	Kettenregel