

GRUNDLAGEN DER STATISTIK

(1)

Lageparameter

Mittelwert:

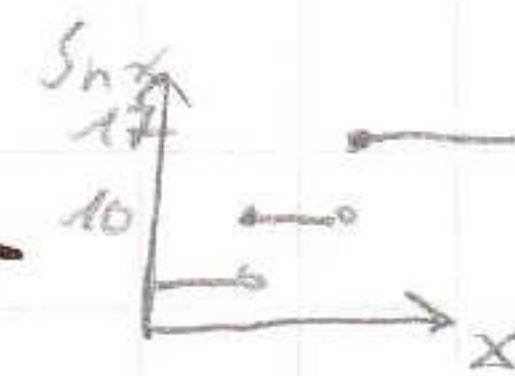
$$\bar{x}_n = \frac{1}{n} \sum_{i=1}^n x_i$$

Median:

$\tilde{x}$, 0,5 - Fraktiles

empirisches Fraktile:

$S_n(x_p) = p$, Wert für den Summenpolygon p ergibt



Modalwert:

häufigstes Merkmal

Variationskoeffizient

Streuungsparameter

Spannweite:

$$x_n - x_1$$

Quartilabstand:

$$x_{0,75} - x_{0,25}$$

Bezeichnet in den 50% der Werte liegen mind. Hälfte der Beobachtungswerte

Mittlere absolute Abweichung:

$$\frac{1}{n} \sum_{i=1}^n |x_i - \tilde{x}|$$

Mittlere quad. Abw.:

$$\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 = s^2, \text{ empirische Varianz}$$

empirische Streuung:

$$\sqrt{s^2}$$

Ziel der schließenden Statistik

stat. Versuch:

wiederholbare Handlung, mehr oder weniger

ungewisser Ausgang

Wahrscheinlichkeit:

Grad des Vertrauens in Ereignis

Ereignis:

Menge von möglichen Versuchsausgängen

WAHRSCHEINLICHKEITEN BZW WAHRSCHEINLICHKEITSBEGRIFFE

klassische

Bedingung:

$$M = \{a_1, \dots, a_m\}$$

$$A = \{a_{i_1}, \dots, a_{i_j}\}, \quad i_j \in \{1, \dots, m\}, \text{ Ereignis}$$

Formel:

$$W(A) := \frac{g}{m} = \frac{\text{günstige}}{\text{mögliche}}$$

geometrische

Bedingung:

$$M = \infty, \quad M \in \mathbb{R}^k$$

$$\text{endlich viele } I(M) < \infty$$

Formel:

$$\frac{I(A)}{I(M)}$$

axiomatische

Bedingung:

$\mathcal{A}$ Ereignissystem, σ Unmögliches, e Sicheres

Formel:

$$A \wedge A^\perp = \sigma, \quad A \vee A^\perp = e$$

$$W(\sigma) = 0, \quad W(e) = 1$$

$$W\left(\bigvee_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} W(A_k)$$

subjektive Wahrscheinlichkeit

Bedingung:

$$W(A|H)$$

Formel:

$$W(A \vee B|H) = W(A|H) + W(B|H)$$

disjunkt

$$W(A \wedge B|H) = W(B|A \wedge H) \cdot W(A|H)$$

A eingetreten

WAHRSCHEINLICHKEITSRÄUME

diskret

Bedingungen:

$$M = \{a_1, \dots\} \text{ abzählbar unendlich}$$

$$p_i = W(a_i)$$

$$\sum_{i=1}^{\infty} p_i = 1$$

Formel:

$$W(A) = \sum_{a \in A} W(a)$$

kontinuierliche

Bedingung:

$$M = \mathbb{R}^k$$

$\mathcal{E}$ = kleinstes Ereignisfeld, das alle Intervalle umfasst

Formel:

Wahrscheinlichkeit durch Dichtefunktion

Additionstheorem

Formel:

$$W(A \cup B) = W(A) + W(B) - W(A \cap B)$$



Multiplikationstheorem

Formel:

$$W(A \cap B) = W(A) \cdot W(B|A) = W(B) \cdot W(A|B)$$

Bedingte Wahrscheinlichkeiten

Bedingung:

A, B Ereignisse

$$W(A) > 0$$

Formel:

$$\frac{W(A \cap B)}{W(A)} = W(B|A)$$

Bayes'sche Formel

Bedingung:

Satz der vollständigen Wahrscheinlichkeit

Formel:

$$W(H_i|A) = \frac{W(H_i) \cdot W(A|H_i)}{\sum_{k=1}^n W(H_k) \cdot W(A|H_k)}$$

Multiplikationstheorem

Satz der vollständigen Wahrscheinlichkeit

stochastische Unabhängigkeit

Formel:

$$W(A \cap B) = W(A) \cdot W(B) \Rightarrow A \perp B$$

$$W(A|B) = W(A|B^c) = W(A)$$

Produktwahrscheinlichkeitsräume

Formel:

$$(M_i, \mathcal{E}_i, W_i), M_1 \times \dots \times M_n, \bigotimes_{i=1}^n \mathcal{E}_i$$

$$W(A_1 \times \dots \times A_n) = W_1(A_1) \dots W_n(A_n)$$

Beispiel:

n -maliges Ziehen mit Zurücklegen

STOCHASTISCHE GRUNDBEGRIFFE

(2)

Ereignisfeld $\mathcal{E}$

Beschreibung:

System von Teilmengen von M

Formel:

$$\underline{M \in \mathcal{E}}, \quad \underline{A \in \mathcal{E} \Rightarrow A^c \in \mathcal{E}}$$

$$\underline{\text{Folge } A_1, A_2, \dots \quad A_i \in \mathcal{E}}$$

$$\text{Morgansche Regeln} \quad \underline{\bigcup_{i=1}^{\infty} A_i \in \mathcal{E}}, \quad \underline{\bigcap_{i=1}^{\infty} A_i \in \mathcal{E}}$$

Wahrscheinlichkeitsverteilung W

Beschreibung:

Fügt jedem Element aus $\mathcal{E}$ einen Wert zwischen

$(0, 1)$ zu.

Formel:

$$\underline{W(M) = 1}$$

$$W\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} W(A_k) \quad \text{disjunkt paarweise}$$

VERTEILUNGEN

3

Stochastische Größen

Beschreibung:

Größen, die variieren

Funktionen von Ausgängen von statistischen Experimenten

Beispiele:

$X(\omega)$ ^{zufälliges Ereignis, Versuchsausgang} z.B. maximale Temperatur

$\underline{X}(\omega)$ ^{Versuchsausgang} z.B. Temperaturverlauf

Formel:

$$X: M \rightarrow \mathbb{R}^m$$

$$X^{-1}(B) := \{\omega \in M : X(\omega) \in B\} \in \mathcal{E}$$

m-dimensionale SG

messbare Funktion:

stetig (keine Sprünge), Punkte eindeutig zuordenbar

$$W_X \{X \in B\} : W(X^{-1}(B))$$

Verteilung stochastischer Größen

Beschreibung:

Bestimmt die Verteilung von SG X

Verteilungsfunktionen

Beschreibung:

$F(\cdot)$ einer SG X ist durch die Funktionswerte

$$F(x) = W\{X \leq x\} \quad \text{bestimmt}$$

$$\{X \leq x\} := X^{-1}((-\infty, x]) \in \mathcal{E}$$

Formel:

$$1) \quad 0 \leq F(x) \leq 1, \quad \text{zwischen 0 und 1}$$

$$2) \quad x_1 < x_2 \Rightarrow F(x_1) \leq F(x_2)$$

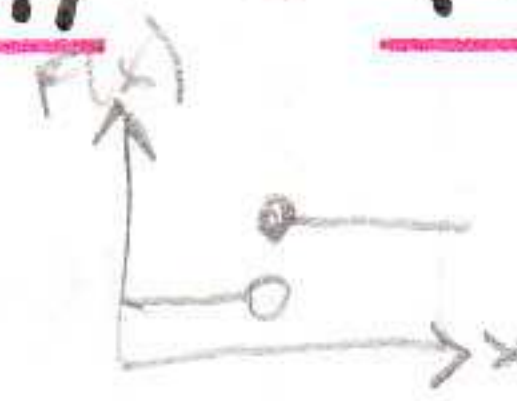
monoton wachsend

$$3) \quad \lim_{x \rightarrow -\infty} F(x) = 0, \quad \lim_{x \rightarrow \infty} F(x) = 1$$

konvergiert negativ gegen 0 und positiv gegen 1

$$4) \quad \lim_{h \downarrow 0} F(x+h) = F(x)$$

rechtsstetig

 von zwei gilt immer die Bed.

$$5) \quad W\{X = x\} = F(x) - \lim_{h \downarrow 0} F(x-h) \quad \text{Sprunghöhe ist Wahrscheinlichkeit}$$

Wahrscheinlichkeit ist der linksseitige Grenzwert

$$6) \quad a < b \Rightarrow W\{a < X \leq b\} = F(b) - F(a)$$

Typen von Verteilungen

diskrete

Beschreibung:

Verteilung ist treppenförmig

Bedingungen:

abzählbar viele $a_i \in \mathbb{R}$ ohne Häufungspunkt

$$\underline{W(\{a_i\}) = p_i > 0}, \quad \sum_{i=1}^{\infty} p_i = 1$$

Borelmenge ist eine Teilmenge der reellen Zahlen, $\mathcal{E}$

ist ein System von Borelmengen, jede Borelmenge ist

ein Element von $\mathcal{E}$, jedes Intervall $(a, b]$ mit

$a, b \in \mathbb{R}$, Vereinigen, schneiden, Differenz $\Rightarrow$ Borelmenge

$$\underline{W(B) = \sum_{a_i \in B} W(\{a_i\})}$$

kontinuierliche

Beschreibung:

SG X alle Werte eines Intervalls

Wahrscheinlichkeit durch integrierbare Dichtefunktion

Bedingungen:

$$f: \mathbb{R} \rightarrow [0, \infty) \quad \text{mit} \quad \int_{-\infty}^{\infty} f(x) dx = 1$$

$$\underline{W([a, b]) = \int_a^b f(x) dx}$$

$$F(x) = \int_{-\infty}^x f(\xi) d\xi$$

$$F'(x) = f(x)$$

$$W\{X=x\} = 0$$

gemischte

Bedingung:

Konvexkombination einer diskreten VF und

einer kontinuierlichen VF

$$F(x) = p \cdot F_1(x) + (1-p) \cdot F_2(x)$$

Mischverteilung

Bedingung:

Kombination von $F_i(\cdot)$

egal ob diskret oder kontinuierlich

ERWARTUNGSWERT

(4)

Beschreibung:

- Mittlerer Wert einer Verteilung
- Zentrum der Wahrscheinlichkeitsverteilung
- Wert einer SG, den sie am häufigsten im Durchschnitt annimmt

diskrete Verteilung

$$\underline{EX := \sum_{x: p(x) > 0} x \cdot p(x)}$$

EX ist das Mittel μ der Verteilung

kontinuierliche Verteilung

$$\underline{EX := \int_{-\infty}^{\infty} x \cdot f(x) dx}$$

Beispiel: $X \sim N(\mu, \sigma^2) \Rightarrow EX = \mu$

gemischte Verteilung

$$\underline{EX := \sum_{i=1}^m x_i \cdot p(x_i) + \int_a^b x \cdot f^*(x) dx}$$

von Funktionen von SG (Satz vom unbewussten Statistiker)

Bedingungen:

X eine SG

$\psi: \mathbb{R} \rightarrow \mathbb{R}$, messbare Funktion

$Y := \psi(X)$, $\exists EX[\psi(X)]$

diskret:

$$EX[\psi(X)] = \underline{\sum_{x: p(x) > 0} \psi(x) \cdot p(x)}$$

kontinuierlich:

$$EX[\psi(X)] = \underline{\int_{-\infty}^{\infty} \psi(x) f(x) dx}$$

allgemein:

$$EX[\psi(X)] = \underline{\int_{-\infty}^{\infty} \psi(x) dF(x)}$$

VARIANZ

$$\sqrt{\text{Varianz}} = \text{Streuung}$$

Beschreibung:

- Durchschnittliche Quadratische Abweichung
- Maß für die Abweichung einer Zufallsvariable vom Erwartungswert

Berechnung:

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

$c, d \in \mathbb{R}$ und konstant

$$\text{Var}(c \cdot X + d) = c^2 \text{Var } X$$

Verschiebungssatz

Bedingungen: X eine SG

$$\exists \text{Var } X$$

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

gemischte Verteilung (Satz vom unbewussten Statistiker)

$$\text{Var } X = \sum_{i=1}^m (x_i - \mathbb{E}X)^2 \cdot p(x_i) + \int_a^b (x - \mathbb{E}X)^2 \cdot f^*(x) dx$$

Beispiele:

$$X \sim N(\mu, \sigma^2) \Rightarrow \text{Var } X = \sigma^2$$

$$X \sim \text{Exp}_\tau \Rightarrow \text{Var } X = \tau^2$$

$$X \sim A_p \Rightarrow \text{Var } X = p \cdot (1-p)$$

$$X \sim B_{n,p} \Rightarrow \text{Var } X = n \cdot p \cdot (1-p)$$

$$X \sim U_{a,b} \Rightarrow \text{Var } X = \frac{(b-a)^2}{12}$$

Standardisierung

Bedingungen:

X eine SG

$$\exists \text{Var } X = \sigma^2 \Rightarrow \exists \mathbb{E}X = \mu$$

allgemein:

$$\frac{X - \mu}{\sigma}$$

$$\mathbb{E}\left(\frac{X - \mu}{\sigma}\right) = 0, \quad \text{Var}\left(\frac{X - \mu}{\sigma}\right) = 1$$

VERTEILUNG VON FUNKTIONEN SG

9

Bedingungen:

X eine SG

$$\underline{X \sim F(\cdot)}$$

$\psi: \mathbb{R} \rightarrow \mathbb{R}$ messbar

$$\underline{Y = \psi(X)}$$

$Y \sim G(\cdot)$ ist gesucht

1. Satz

Bedingung:

$\psi: \mathbb{R} \rightarrow \mathbb{R}$ ^{jede stetige ist messbar und auch jede monoton}
messbar, monotonsteigend, invertierbar

Satz:

$$\underline{G(x) = F(\psi^{-1}(x))} \quad \forall x \in \mathbb{R}$$

2. Satz

Bedingung:

Bedingungen 1. Satz

$X \sim f(\cdot)$, kontinuierlich verteilt

$\psi^{-1}(\cdot)$ differenzierbar $\Rightarrow Y \sim g(\cdot)$, kont. verteilt

Satz:

$$\underline{g(x) = f(\psi^{-1}(x)) \cdot \left| \frac{d\psi^{-1}(x)}{dx} \right|}$$
 noch $\times$ differenziert

3. Satz

Bedingung:

$$\underline{X \sim N(\mu, \sigma^2)}$$

$$c, d \in \mathbb{R}, \quad c \neq 0$$

Satz:

$$\underline{cX + d \sim N(c\mu + d, c^2\sigma^2)}$$

4. Satz

Bedingung:

$$X \sim F(\cdot)$$

$$\underline{\exists F^{-1}(\cdot)}$$

Satz:

$$\text{daraus folgt } \underline{F(X) \sim U_{0,1}}$$

STOCHASTISCHE VEKTOREN

Beschreibung

- Versuchsausgänge, die durch $x_1, \dots, x_m$ reelle Zahlen beschrieben werden
- Stochastisches Modell $(X_1, \dots, X_m)$, X_i : SG

Randverteilungen allgemein

$$\underline{X} = (X_1, \dots, X_m)$$

Verteilung jedes Teilvektors $(X_{i_1}, \dots, X_{i_k})$

Satz:

$$\underline{X} = (X, Y) \quad \underline{X} \sim F(x, y)$$

$$F_1(x) = \lim_{y \uparrow \infty} F(x, y)$$

Randverteilung von X

$$F_2(y) = \lim_{x \uparrow \infty} F(x, y)$$

Randverteilung von Y

Randdichten

Satz:

$$\underline{X} = (X, Y) \quad \underline{X} \sim f(x, y)$$

$$f_1(x) = \int_{-\infty}^{\infty} f(x, y) dy$$

Randdichte von X

$$f_2(y) = \int_{-\infty}^{\infty} f(x, y) dx$$

Randdichte von Y

Randverteilungen

Satz:

$$\underline{X} = (X, Y) \quad \underline{X} \sim p(x, y)$$

$$p_1(x) = \sum_{y \in M_Y} p(x, y)$$

Punktwahrscheinlichkeit X

$$p_2(y) = \sum_{x \in M_X} p(x, y)$$

Punktwahrscheinlichkeit Y

MEHRDIMENSIONALE VERTEILUNGEN

⑥

diskrete Verteilung

Bedingung:

M eine abzählbare Menge von m -Tupel

kein Häufungspunkt

$$\sum_{(x_1, \dots, x_m) \in M} p(x_1, \dots, x_m) = 1$$

Satz:

$$W(A) := \sum_{(x_1, \dots, x_m) \in A} p(x_1, \dots, x_m)$$

kontinuierliche Verteilung

Bedingung:

$$(X_1, \dots, X_m) \sim f(x_1, \dots, x_m)$$

$$f: \mathbb{R}^m \rightarrow [0, \infty)$$

$$\int_{\mathbb{R}^m} f(x_1, \dots, x_m) dx_1 \dots dx_m = 1$$

I am wenigsten umfassende Intervall mit m Dimensionen

$$I_m := \prod_{i=1}^m (a_i, b_i] \quad , \quad a_i < b_i \in \mathbb{R}$$

Satz:

$$W(B) := \int_B f(x_1, \dots, x_m) dx_1 \dots dx_m$$

Spezialfall $B = I_m = \prod_{i=1}^m (a_i, b_i]$ gilt

$$W(I_m) := \int_{a_m}^{b_m} \dots \int_{a_1}^{b_1} f(x_1, \dots, x_m) dx_1 \dots dx_m$$

Funktionen von Stochastischen Vektoren

Bedingung:

$$\underline{X} = (X_1, \dots, X_m)$$

$$\psi: \mathbb{R}^m \rightarrow \mathbb{R} \quad , \quad \text{messbar}$$

$Z := \psi(X_1, \dots, X_m)$ eine SG

$$\exists \mathbb{E}\psi(\underline{X})$$

diskret:

$$\underline{X} \sim (M_X, p(x_1, \dots, x_m))$$

$$\mathbb{E}\psi(X_1, \dots, X_m) = \sum_{(x_1, \dots, x_m) \in M_X} \psi(x_1, \dots, x_m) p(x_1, \dots, x_m)$$

kontinuierlich:

$$\underline{X} \sim f(x_1, \dots, x_m)$$

$$\mathbb{E}\psi(X_1, \dots, X_m) = \int_{\mathbb{R}^m} \psi(x_1, \dots, x_m) f(x_1, \dots, x_m) dx_1 \dots dx_m$$

Linearität der Erwartungsbildung

Bedingung:

$X_1, \dots, X_n$ 1-dim SG mit $\exists \mathbb{E}X_i$

$a_1, \dots, a_n$ reelle Konstanten

Satz:

$$\mathbb{E}\left(\sum_{i=1}^n a_i X_i\right) = \sum_{i=1}^n a_i \mathbb{E}X_i \quad \Rightarrow \quad \mathbb{E}(X+Y) = \mathbb{E}X + \mathbb{E}Y$$

KOVARIANZ, KORRELATION & UNABHÄNGIGKEIT

Beschreiben Zusammenhänge zwischen SG

Kovarianz

Beschreibung:

$\text{Cov}(X, Y)$ ist das Maß für die Abhängigkeit

Kovarianz qualifiziert den (linearen) Zusammenhang

positiver und negativer Wert möglich

Formel:

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

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

Satz:

$X_1, \dots, X_n$ sind SG, $\exists \text{Var} X_i$, $c_1, \dots, c_n$ reelle Konstanten

$$\text{Var}\left(\sum_{i=1}^n c_i X_i\right) = \sum_{i=1}^n c_i^2 \text{Var} X_i + 2 \cdot \sum_{i=1}^{n-1} \sum_{j=2}^n \underset{i < j}{c_i c_j} \text{Cov}(X_i, X_j)$$

$$\text{Sonderfall: } \text{Var}(X + Y) = \text{Var} X + \text{Var} Y + 2 \cdot \text{Cov}(X, Y)$$

Korrelationskoeffizient

Beschreibung:

Maß für die Stärke einer näherungsweise linearen Beziehung

Maß für den linearen Zusammenhang

Bedingungen:

X, Y sind SG

$\exists \text{Varianz}$

Formel:

$$\rho_{X,Y} := \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var} X} \cdot \sqrt{\text{Var} Y}}$$

$$\underline{-1 \leq \rho_{X,Y} \leq 1}$$

$$\underline{\rho_{X,Y} = 0} \Rightarrow \underline{\text{unkorreliert}} \quad \text{stoch. Unab.}$$

Satz:

paarweise unkorrelierte SG:

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

Stochastische Unabhängigkeit

Bedingungen:

X, Y sind SG, $X \sim F_1(\cdot)$, $Y \sim F_2(\cdot)$, gemeinsam $G(\cdot, \cdot)$

Formel:

$$\text{Es gilt } \underline{G(x, y) = F_1(x) \cdot F_2(y)} \quad (X \perp\!\!\!\perp Y)$$

diskret:

$$\underline{p(x, y) = p_1(x) \cdot p_2(y)}$$

kontinuierlich:

$$\underline{f(x, y) = f_1(x) \cdot f_2(y)}$$

Anmerkung

$$\perp\!\!\!\perp \text{ wenn } \rho = 0$$

BEINGTE VERTEILUNGEN

7

(X, Y) ein 2-dimensionaler Stochastischer Vektor z.B. Geschwindigkeit & Bremsweg

diskrete Verteilungen

Bedingung:

$$(X, Y) \sim p(x, y)$$

$p_1(\cdot)$ = Randverteilung X , $p_2(\cdot)$ = Randverteilung Y

Formel:

$$p(y|x) = \frac{W\{X=x \wedge Y=y\}}{W\{X=x\}} = \frac{p(x, y)}{p_1(x)}$$

$$p(x|y) = \frac{p(x, y)}{p_2(y)}$$

kontinuierliche Verteilung

Bedingung:

$$(X, Y) \sim f(x, y)$$

$f_1(\cdot)$ = Randdichte X , $f_2(\cdot)$ = Randdichte Y

Problem bei $W\{X=x\} = 0$

Formel:

$$f(y|x) := \frac{f(x, y)}{f_1(x)} \quad \forall x \text{ mit } f_1(x) > 0$$

$$f(x|y) := \frac{f(x, y)}{f_2(y)} \quad \forall y \text{ mit } f_2(y) > 0$$

Regressionsfunktion

Beschreibung:

bedingte Erwartung von $Y|X$

Formel:

(X, Y) stochastischer Vektor

Funktion $x \rightarrow E(Y|X=x)$

$y \rightarrow E(X|Y=y)$

Erwartungswert von Y unter der Bedingung X

FUNKTIONEN VON STOCHASTISCHEN VEKTOREN

Bedingungen: $\underline{X} = (X_1, \dots, X_m)$ ein stochastischer Vektor

$$\underline{\psi}: \mathbb{R}^m \rightarrow \mathbb{R}^k, \text{ messbar}$$

$$\underline{Z} = \underline{\psi}(\underline{X})$$

Beispiel: X, Y sind SG $Z = X + Y$ $\psi: \mathbb{R}^2 \rightarrow \mathbb{R}$

$$(X_1, \dots, X_m) \text{ SV} \quad Z = \min(X_1, \dots, X_m) \quad \psi: \mathbb{R}^m \rightarrow \mathbb{R}$$

Lineartransformation von $N(\mu, \Sigma)$

Bedingungen: $\underline{X} = (X_1, \dots, X_m)^T$ stochastischer Vektor

$$\underline{X} \sim N(\mu, \Sigma)$$

reguläre Matrix $\underline{A} = (a_{ij})_{\substack{i=1(1)m \\ j=1(1)m}}$

Formel: $\underline{Z} = \underline{A} \cdot \underline{X} \sim N(\underline{A} \cdot \mu, \underline{A} \Sigma \underline{A}^T)$

Satz: unabhängige SG

Bedingung: $X_1, \dots, X_n$ sind unabhängige SG

$$X_i \sim F_i(\cdot)$$

Formel: $W\{\max(X_1, \dots, X_n) \leq x\} = \prod_{i=1}^n F_i(x)$

$$W\{\min(X_1, \dots, X_n) \leq x\} = 1 - \prod_{i=1}^n [1 - F_i(x)]$$

Ordnungsstatistiken

Beschreibung: Der Größe nach geordnete SG

Bemerkung: $X_{(1)} = \min(X_1, \dots, X_n)$ $X_{(n)} = \max(X_1, \dots, X_n)$

Faltung von Wahrscheinlichkeitsverteilungen

Beschreibung: Vereinigung von zwei SG

Bedingung: X, Y unabhängige SG

$$X \sim F(\cdot), Y \sim G(\cdot)$$

$H(\cdot)$ VF von $X + Y$

Formel: $H(x) = \int_{-\infty}^{\infty} F(x-t) dG(t)$

diskret: $\sum_{x \in M_X} p_X(x) \cdot p_Y(a-x)$

$$a = x + y \rightarrow y = a - x$$

kontinuierlich: $\int_{-\infty}^{\infty} f_Y(y) \cdot f_X(a-y) dy$

Beispiel: Taschenlampe + Reservebatterie

Additionstheorem für $N(\mu, \sigma^2)$

(8)

Bedingung: $X_i \sim N(\mu_i, \sigma_i^2)$

$X_1, \dots, X_n$ sind unabhängig

Formel:

$$\sum_{i=1}^n X_i \sim N\left(\sum_{i=1}^n \mu_i, \sum_{i=1}^n \sigma_i^2\right)$$

Additionstheorem für Gammaverteilungen $\text{Gam}(\alpha_i, \lambda)$

Bedingung: $X_i \sim \text{Gam}(\alpha_i, \lambda)$

$X_1, \dots, X_n$ sind unabhängig

Formel:

$$\sum_{i=1}^n X_i \sim \text{Gam}\left(\sum_{i=1}^n \alpha_i, \lambda\right)$$

Erlang-Verteilung $\text{Er}_{n, \lambda}$

Bedingung: $X_1, \dots, X_n$ unabhängig exponentialverteilt

Dichte:

$$f(x) = \lambda e^{-\lambda x} I_{(0, \infty)}(x)$$

Formel:

$$\sum_{i=1}^n X_i \Rightarrow f(x|n, \lambda) = \frac{\lambda^n}{(n-1)!} x^{n-1} e^{-\lambda x} I_{(0, \infty)}(x)$$

Beispiel:

Warteschlangen

FOLGEN VON SG

unendliche Folge ist unabhängig, wenn je endlich unabhängig sind

UIV-Folge: beliebig oft wiederholbare statistische Versuche

Konvergenzen

fast sicher $\Rightarrow$ Gesetz der großen Zahlen

Verteilungskonvergenz $\Rightarrow$ Zentrale Grenzwertsatz

fast sicher gleichmäßige $\Rightarrow$ Fundamentalsatz

Gesetz der großen Zahlen

Beschreibung: Konvergenz von Folgen SG

Bedingung: UIV-Folgen

Satz: $h_n(B)$, $X_i = 1$ falls B eintritt sonst 0

$$h_n(B) = \frac{1}{n} \sum_{i=1}^n X_i := \bar{X}_n$$

$\bar{X}_n$ (vor Durchführung) $\xrightarrow{f.s.}$ gegen X_0 , wenn

$$W(\{\omega : \lim_{n \rightarrow \infty} X_n(\omega) = X_0(\omega)\}) = 1$$

$$\bar{X}_n \xrightarrow[n \rightarrow \infty]{f.s.} X_0$$

Zentraler Grenzwertsatz

Beschreibung: Folge von FV, deren standardisierte Summe
konvergiert gegen $N(0,1)$

Bedingung: unendliche Folge $X_1, \dots$, von SG, unabhängig
 $\exists \text{Var } X_k$

Lindeberg-Bedingung: keine $F_k(\cdot)$ hat dominierenden Einfluss

Satz:

standardisierte Summe $Z_n := \frac{\sum_{k=1}^n X_k - E(\sum_{k=1}^n X_k)}{\sqrt{\text{Var}(\sum_{k=1}^n X_k)}}$

$$\lim_{n \rightarrow \infty} W\{Z_n \leq x\} = \Phi(x) \text{ Standard-Normalverteilung}$$

Lindeberg-Bedingung durch $s^2 := \text{Var}(\sum_{k=1}^n X_k) = \sum_{k=1}^n \text{Var } X_k$

Normalapproximation

9

Formel:

$$\sum_{k=1}^n X_k = S$$

$X_1, \dots, X_n$ für großes n

$$Z = \frac{S - \mathbb{E}S}{\sqrt{\text{Var} S}} \stackrel{!}{\sim} N(0,1) \quad \text{"approximativ"}$$

$$S = \sqrt{\text{Var} S} \cdot Z + \mathbb{E}S \stackrel{!}{\sim} N(\mathbb{E}S, \text{Var} S)$$

Fundamentalsatz der Statistik

Beschreibung:

Folge der Werte der empirischen Verteilungsfunktion

konvergiert gegen die zugrunde liegende

Verteilungsfunktion

Bedingungen:

$X_1, X_2, \dots$ eine UIV-Folge

$$X_i \sim F(\cdot)$$

$X_1 = x_1, X_2 = x_2, \dots, X_n = x_n$ empirische VF $F_n^*(\cdot)$

$$F_n^*(x; X_1, \dots, X_n) = \frac{1}{n} \sum_{i=1}^n I_{(-\infty, x)}(X_i)$$

Satz:

$$\mathbb{E}(F_n^*(x; X_1, \dots, X_n)) = F(x)$$

$$F_n^*(x; X_1, \dots, X_n) \xrightarrow[n \rightarrow \infty]{f.s.} F(x)$$

$$W\left\{ \lim_{n \rightarrow \infty} \sup_{x \in \mathbb{R}} |F_n^*(x) - F(x)| = 0 \right\} = 1$$

fast sichere gleichmäßige Konvergenz

STICHPROBEN UND STATISTIKEN

Gesucht ist die Verteilung einer SG anhand von Beobachtungen

Stichprobe

Bedingungen:

$X_1, \dots, X_n$ eine UIV-Folge

$X_1, \dots, X_n$ sind mathematische Stichproben

$X_1 = x_1, \dots, X_n = x_n$ sind konkrete Stichproben

$\psi: \mathbb{R}^n \rightarrow \mathbb{R}^k$, messbar

$\psi(X_1, \dots, X_n)$ ergibt k-dimensionale Statistik

Stichprobenmittel:

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

unverzerrt, effizient, konsistent (bzgl.) SF für $\mathbb{E}X$

Stichprobenvarianz:

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

Stichprobenraum:

sofern X einen M_X^n hat

diskret:

$$X \sim (M_X, p(x))$$

Schätzfunktion

$$w(x_1, \dots, x_n) = \prod_{i=1}^n p(x_i)$$

Wahrscheinlichkeitsverteilung von SP

kontinuierlich:

$$X \sim f(\cdot)$$

Schätzfunktion

$g(x_1, \dots, x_n)$ ist die Dichte der SP

$$g(x_1, \dots, x_n) = \prod_{i=1}^n f(x_i)$$

Schätzfunktion

Statistik zur Schätzung einer charakteristischen Größe der

Wahrscheinlichkeitsverteilung von X

Parameterraum

Beschreibung:

W_θ von X hängt von Parameter θ ab

$\theta \in \Theta$, Θ Parameterraum

geprüfte Parameter $\tau(\theta)$ $\tau: \Theta \rightarrow \mathbb{R}$

Beispiele:

$\mathbb{E}X_\tau$, $\theta = \tau$, $\Theta = (0, \infty)$

$N(\mu, \sigma^2)$, $\tau(\theta) = \mu$, $\tau(\theta) = \sigma$

KLASSISCHE PUNKTSCHÄTZUNG

(10)

Bedingungen:

$$SG \quad X \sim (M_X, W_X)$$

$X_1, \dots, X_n$ sind Stichproben von X

$v: M_X^n \rightarrow \mathbb{R}^k$, $k=1 \Rightarrow X$ ist SG, $k>1 \Rightarrow X$ ist SV

$v(X_1, \dots, X_n)$ ist Schätzfunktion z.B. Stichprobenmittel

$v(x_1, \dots, x_n)$ ist Schätzwert

bei Geraden $\tau(\theta)$

$$A(X_1, \dots, X_n), \quad A: M_X^n \rightarrow \mathbb{R}$$

Unverzerrtheit

Beschreibung:

Güteeigenschaft von Schätzfunktionen

$v(X_1, \dots, X_n)$ ist unverzerrt, wenn

$\mathbb{E}v(X_1, \dots, X_n) = \xi$, charakteristischer Wert

$\mathbb{E}_\theta v(X_1, \dots, X_n) = \xi$, für parametrisches Model

Satz:

falls $\exists \mathbb{E}X \Rightarrow \bar{X}_n$ unverzerrte SF für $\mathbb{E}X$

falls $\exists \text{Var}X \Rightarrow S_n^2$ unverzerrte SF für $\text{Var}X$

Beispiel:

$$X \sim N(\mu, \sigma^2) \Rightarrow \mathbb{E}_\theta \bar{X}_n = \mu$$

Effizienz

Beschreibung:

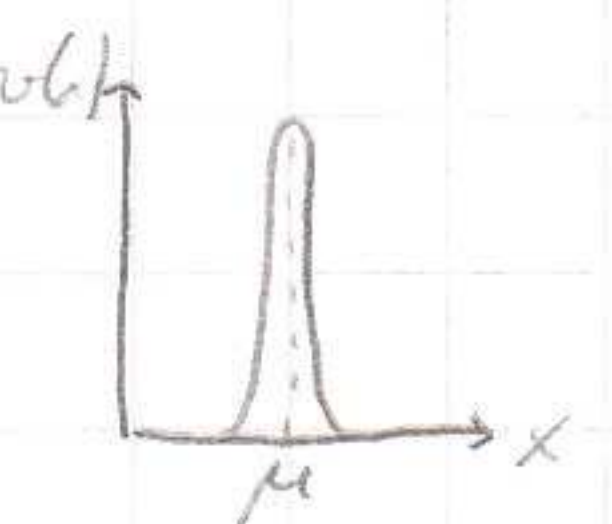
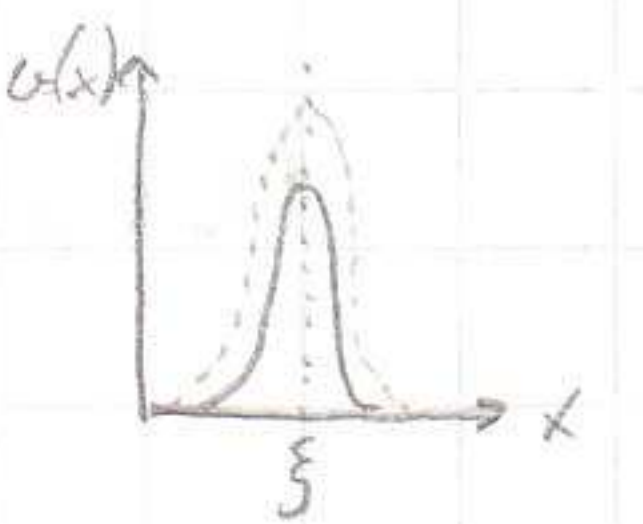
unverzerrte Schätzfunktion mit kleinster Varianz

SF $A^*(X_1, \dots, X_n)$ ist effizient, wenn

$$\text{Var}A^*(X_1, \dots, X_n) = \min_{A \in T} \text{Var}A(X_1, \dots, X_n)$$

Satz:

$\bar{X}_n$ ist die effiziente SF für $\mathbb{E}X$



Konsistenz

Beschreibung:

Stichprobenumfang $n \rightarrow \infty$

Schätzfunktion gegen gesuchten Wert

erwartungstreu auch bei $n=1$, konsistent bei $n \rightarrow \infty$

Bedingungen:

$X_1, X_2, \dots$ eine UIV-Folge

$A_n(X_1, \dots, X_n)$ eine Folge von Schätzfunktionen

$$\underline{A_n(X_1, \dots, X_n) \xrightarrow[n \rightarrow \infty]{W} \xi}$$

$$\underline{\lim_{n \rightarrow \infty} W\{\xi - \varepsilon < A_n(X_1, \dots, X_n) < \xi + \varepsilon\} = 1}$$

$$\lim_{n \rightarrow \infty} W\{|A_n(X_1, \dots, X_n)| \leq \xi + \varepsilon\} = 1$$

Plausibilität

diskret:

$$\underline{X \sim p(\cdot | \theta)}$$

Plausible Schätzwerte sind solche, die einer

konkreten Stichprobe für ein diskretes

stochastisches Modell maximale Wahrscheinlichkeit

erteilen

$$\underline{w(x_1, \dots, x_n | \theta) = \prod_{i=1}^n p(x_i | \theta)}$$

$w(\overbrace{x_1, \dots, x_n}^{\text{konkret}} | \theta)$ ist eine Funktion von θ

$$\underline{\text{Plausibilitätsfunktion } f(\theta; \overbrace{x_1, \dots, x_n}^{\text{konstant}}) = \prod_{i=1}^n p(x_i | \theta)}$$

plausibler Schätzwert $\hat{\theta}$ für θ_0 ist jener Wert

für den die Wahrscheinlichkeit der enthaltenen Stichproben

maximal ist

$$\underline{f(\hat{\theta}; x_1, \dots, x_n) = \max_{\theta \in \Theta} f(\theta; x_1, \dots, x_n)}$$

kontinuierlich:

$$\underline{X \sim f(\cdot | \theta)}$$

gemeinsame Dichtefunktion normaler Stichproben

$$\underline{g(x_1, \dots, x_n | \theta) = \prod_{i=1}^n f(x_i | \theta)}$$

$g(\overbrace{x_1, \dots, x_n}^{\text{konkret}} | \theta)$ Funktion von $\theta = f(\theta; x_1, \dots, x_n)$

$$\underline{f(\hat{\theta}; x_1, \dots, x_n) = \max_{\theta} f(\theta; x_1, \dots, x_n)}$$

plausible Schätzfunktion: ordnet jeder Stichprobe den plausiblen Schätzwert

$$v(x_1, \dots, x_n) \text{ zu } \Rightarrow v(X_1, \dots, X_n)$$

KONFIDENZBEREICHE

(11)

Güteangaben für Parameterschätzungen

Überdeckungswahrscheinlichkeit $1 - \alpha$

Teilmenge $C_{1-\alpha} \subset \Theta$, θ_0 bei $X \sim f(\cdot | \theta)$ gesucht

$$\underline{W\{\theta_0 \in C_{1-\alpha}\} = 1 - \alpha}$$

Pivot - Größen

Beschreibung:

Funktion $g(X_1, \dots, X_n; \theta)$ einer Stichprobe und

des Parameters θ , die nicht von θ abhängt

$$X \sim W_\theta, \theta \in \Theta$$

Satz:

Stichprobe von $N(\mu, \sigma^2)$

$$\frac{\bar{X}_n - \mu}{S_n / \sqrt{n}} \sim t_{n-1}$$

$$\frac{(n-1)S_n^2}{\sigma^2} \sim \chi_{n-1}^2$$

Formel:

Konfidenzbereich für μ einer $N(\mu, \sigma^2)$

$$\left[\bar{X}_n - \frac{S_n}{\sqrt{n}} \cdot t_{n-1; 1-\frac{\alpha}{2}} ; \bar{X}_n + \frac{S_n}{\sqrt{n}} \cdot t_{n-1; 1-\frac{\alpha}{2}} \right]$$

Formel:

Konfidenzbereich für σ^2 einer $N(\mu, \sigma^2)$

$$\left[\frac{(n-1)S_n^2}{\chi_{n-1; 1-\frac{\alpha}{2}}^2} ; \frac{(n-1)S_n^2}{\chi_{n-1; \frac{\alpha}{2}}^2} \right]$$

STATISCHE HYPOTHESEN UND TESTS

statische Hypothese ist eine Aussage über die Verteilung = $\mathcal{H}_0$

statistischer Test ist ein Verfahren zur Entscheidung ob Hypothese angenommen wird

Wahrscheinlichkeitspapier

Durchführung:

Transformation der $(x, F(x))$ -Ebene auf (t, z) -Ebene

Bilder der $F(\cdot)$ aus $\mathcal{H}_0$ Gerade

Zusammenhang zwischen u_p und x_p

$$u_p = \frac{x_p - \bar{x}}{\sigma}$$

Fehlerarten und Fehlerwahrscheinlichkeiten

Verwerfung einer richtigen Hypothese $\mathcal{H}_0$ = Fehler erster Art

Fehlerwahrscheinlichkeit α = $W\{\mathcal{H}_0 \text{ verworfen} \mid \mathcal{H}_0 \text{ richtig}\}$

Annahme einer falschen Hypothese $\mathcal{H}_1$ = Fehler zweiter Art

Fehlerwahrscheinlichkeit β = $W\{\mathcal{H}_1 \text{ angenommen} \mid \mathcal{H}_1 \text{ falsch}\}$

Ideal: beide α und β klein halten bei Tests

Verwerfungsräume, Teststatistiken und kritischer Bereich

Bedingungen:

Stichproben $X_1, \dots, X_n$

Hypothese $\mathcal{H}$

M_x^n Stichprobenraum

Formel:

M_x^n wird geteilt, $M_x^n = V \cup A$

konkrete Stichprobe x_i in $V \Rightarrow \mathcal{H}$ verworfen

Konstruktion eines Tests: α möglichst klein

falls $\mathcal{H}_0$ richtig

$$\underline{W\{u_{\frac{\alpha}{2}} \leq T \leq u_{1-\frac{\alpha}{2}}\} = 1 - \alpha}$$

u_p = p-Fraktilen

Komplement der Teststatistik $(-u_{1-\frac{\alpha}{2}}, u_{1-\frac{\alpha}{2}})$

heißt kritischer Bereich

$$\underline{V = \{(x_1, \dots, x_n) : | \text{Verteilung} | \geq u_{1-\frac{\alpha}{2}}\}}$$

TESTS FÜR NORMALVERTEILUNG

(12)

t-Test für den Erwartungswert

Bedingungen:

$$X \sim N(\mu, \sigma^2)$$

μ und σ^2 sind unbekannt

$x_1, \dots, x_n$ eine konkrete Stichprobe

Test für $H_0: \mu = \mu_0$

Verwerfungsraum durch α gegeben

Formel:

$$V = \{(x_1, \dots, x_n) : \frac{\sqrt{n} \cdot |\bar{X}_n - \mu_0|}{S_n} \geq t_{n-1; 1-\frac{\alpha}{2}}\}$$

$$W \{1. \text{ Art}\} = W \left\{ \frac{\sqrt{n} \cdot |\bar{X}_n - \mu_0|}{S_n} \geq t_{n-1; 1-\frac{\alpha}{2}} \right\} = \alpha$$

Test für Varianz

Bedingung:

$$X \sim N(\mu, \sigma^2)$$

μ und σ^2 sind unbekannt

$x_1, \dots, x_n$ eine konkrete Stichprobe

Test für $H_0: \sigma^2 = \sigma_0^2$

Verwerfungsraum durch α gegeben

Formel:

$$V = \{(x_1, \dots, x_n) : \frac{(n-1)S_n^2}{\sigma_0^2} \notin (\chi_{n-1}^2; \frac{\alpha}{2}, \chi_{n-1}^2; 1-\frac{\alpha}{2})\}$$

Zwei Stichproben selbe Normalverteilung

Bedingung:

$$\text{Stichprobe } X_1, \dots, X_m \sim N(\mu_X, \sigma^2)$$

$$\text{Stichprobe } Y_1, \dots, Y_n \sim N(\mu_Y, \sigma^2)$$

Stichproben unabhängig

Formel EX:

$H_0: \mu_X = \mu_Y$ mit Wahrscheinlichkeit α eines Fehlers

$$\text{verwerfen falls } t = \frac{\bar{X}_m - \bar{Y}_n}{\sqrt{(\frac{1}{m} + \frac{1}{n}) \cdot \frac{(m-1)S_x^2 + (n-1)S_y^2}{m+n-2}}} \geq t_{m+n-2; 1-\frac{\alpha}{2}}$$

Formel VarX:

$$\frac{S_m^2 / \sigma_x^2}{S_n^2 / \sigma_y^2} \sim F_{m-1, n-1}$$

Chiquadrat - Anpassungstest

Überprüfen ob X einer Verteilung folgt

Satz:

Ein fache Hypothese

$X_1, X_2, \dots$ unbeschränkte SP auf $(M_X, \mathbb{E})$

$n \cdot w_j$ sind die zu Erwartenden absoluten Häufigkeiten von A_j

Zerlegung von M_X in Teilintervalle $A_1, \dots, A_r$

$$\lim_{n \rightarrow \infty} W \{ Z_n = \sum_{j=1}^r \frac{[Y_j^{(n)} - n \cdot w_j]^2}{n \cdot w_j} \leq \chi_{r-1, p}^2 \} = p$$

Z_n ist asymptotisch nach χ_{r-1}^2 verteilt

Test:

$H_0: X \sim W$

$$V = \{ (x_1, \dots, x_n) : z_n \geq \chi_{r-1, 1-\alpha}^2 \}$$

Wahrscheinlichkeit eines Fehlers 1. Art $n \rightarrow \infty$ gegen α

Satz:

Zusammengesetzte Parameterhypothese

$X_1, \dots, X_n$ SP von X

$X \sim W_\theta, \theta \in \Theta$

$\hat{\theta}^{(n)}$ plausible Schätzwert

$$\lim_{n \rightarrow \infty} W \{ Z_n = \sum_{j=1}^r \frac{[Y_j^{(n)} - n \cdot w_j(\hat{\theta}^{(n)})]^2}{n \cdot w_j(\hat{\theta}^{(n)})} \leq \chi_{r-s-1, p}^2 \} = p$$

Z_n asymptotisch nach χ_{r-s-1}^2 verteilt

$$V = \{ (x_1, \dots, x_n) : z_n(x_1, \dots, x_n) \geq \chi_{r-s-1, 1-\alpha}^2 \}$$

Kurzform χ^2 -Anpassungstest

Überprüfen ob X einer Verteilung folgt

Satz:

$X_1, X_2, \dots$ unbeschränkte SP

$A_1, \dots, A_r$ sind Teilintervalle von M_X

$$Z_n = \sum_{j=1}^r \frac{[Y_j^{(n)} - n \cdot w_j]^2}{n \cdot w_j}$$

$$\lim_{n \rightarrow \infty} W \{ Z_n \leq \chi_{r-1, p}^2 \} = p$$

$$V = \{ (x_1, \dots, x_n) : z_n \geq \chi_{r-1, 1-\alpha}^2 \}$$

Beschreibung:

beschreibt kausale, nicht deterministische Zusammenhänge
mit stochastischen Modellen

statistische Zusammenhänge von Merkmalen

Beispiel:

X Geschwindigkeit, Y_x Bremsweg

$$Y_x = \psi(x) + U_x \quad \text{Restgröße}$$

lineare Regressionsfunktion: $Y_x \sim W_\theta$

$$\psi(x) = \mathbb{E}_\theta Y_x = \psi(x, \theta)$$

Regressionsfunktion ist linear, wenn sie als Funktion der unbekannten Parameter linear ist

Beispiele:

$$\psi(x, \theta) = \alpha + \beta x, \quad \theta = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \quad \text{linear}$$

$$\psi(x, \theta) = \theta_0 + \theta_1 x + \theta_2 x^2 + \dots + \theta_k x^k \quad \text{polynomisch}$$

$$\psi(\underline{x}, \theta) = \theta_1 x_1 + \theta_2 x_2 + \dots + \theta_k x_k = \underline{x} \theta \quad \text{multiple}$$

Regressionsparameter:

müssen aus Beobachtungen (x_i, y_i) geschätzt werden

Regressionsgeraden:

$$x \in \mathbb{R}$$

$$\psi(x, \alpha, \beta) = \alpha + \beta x$$

$$Y_x = \alpha + \beta x + U_x \quad \text{mit konstanter Varianz}$$

$$\text{Var } U_x \equiv \sigma^2$$

optimale Schätzwerte $\hat{\alpha}$ und $\hat{\beta}$:

$$\sum_{i=1}^n (y_i - \hat{\alpha} - \hat{\beta} x_i)^2 = \min_{\alpha, \beta} \sum_{i=1}^n (y_i - \alpha - \beta x_i)^2$$

Schätzfunktionen:

$$\text{linear: } A = \sum_{i=1}^n c_i Y_i \quad B = \sum_{i=1}^n d_i Y_i \quad Y_i = Y_{x_i}$$

Satz von Gauß-Markoff: $x_1, \dots, x_n$ sind nicht alle gleich

$Y_1, \dots, Y_n$ sind unkorreliert

$\text{Var } Y_i \equiv \sigma^2$ alle Varianzen sind gleich

$$\mathbb{E} Y_i = \alpha + \beta x_i$$

α, β reelle Konstanten, α, β, σ unbekannt

$$A = \frac{(\sum_{i=1}^n x_i^2) \cdot (\sum_{i=1}^n Y_i) - (\sum_{i=1}^n x_i) \cdot (\sum_{i=1}^n x_i Y_i)}{n(\sum_{i=1}^n x_i^2) - (\sum_{i=1}^n x_i)^2}$$

BAYES - STATISTIK

unbekannte Größen werden durch SG beschrieben

θ durch stochastisches Modell beschrieben

$$X \sim W_\theta$$

$\tilde{\theta} \sim \pi(\cdot)$ A-priori - Verteilung von θ , vor Erhebung

$\pi(\cdot | D)$ A-posteriori - Verteilung , nach Beobachtung D

Bayes'sches Theorem

Bedingungen:

$$X \sim f(\cdot | \theta) , \quad \tilde{\theta} \sim \pi(\cdot)$$

$x_1, \dots, x_n$ sind Stichproben

durch SP bedingte Verteilung von $\tilde{\theta} = \pi(\cdot | x_1, \dots, x_n)$

diskret:

$$X \sim p(\cdot | \theta) , \quad \theta \in \{\theta_1, \dots, \theta_k\} = \Theta$$

Zerlegung von Θ

$$p(y|x) = \frac{p(x|y) \cdot p_1(y)}{\sum_{y \in Y} p_1(y) \cdot p(x|y)}$$

$$\pi(\theta_j | x_1, \dots, x_n) = \frac{w(x_1, \dots, x_n | \theta_j) \pi(\theta_j)}{\sum_{i=1}^k w(x_1, \dots, x_n | \theta_i) \pi(\theta_i)}$$

bei einfachen SP gilt wegen Unabhängigkeit

$$w(x_1, \dots, x_n | \theta_j) = \prod_{i=1}^n p(x_i | \theta_j)$$

Plausibilitätsfunktion:

$$\text{Konstante } c = \sum_{j=1}^k \pi(\theta_j) l(\theta_j, x_1, \dots, x_n)$$

$$\pi(\theta_j | x_1, \dots, x_n) = \frac{1}{c} \pi(\theta_j) l(\theta_j; x_1, \dots, x_n)$$

kontinuierlich:

$$X \sim f(\cdot | \theta) , \quad \tilde{\theta} \sim \pi \quad \text{A-priori-Dichte}$$

$x_1, \dots, x_n$ sind Stichproben

$$f(y|x) = \frac{f_y(y) \cdot f(x|y)}{\int_{-\infty}^{\infty} f_y(y) \cdot f(x|y) dy}$$

$$\pi(\theta | x_1, \dots, x_n) = \frac{\pi(\theta) l(\theta; x_1, \dots, x_n)}{\int_{\Theta} \pi(\theta) l(\theta; x_1, \dots, x_n) d\theta} \quad \text{Randdichte der SP}$$

c ist Konstant w.o.

$$\pi(\theta | x_1, \dots, x_n) = \frac{1}{c} \pi(\theta) l(\theta; x_1, \dots, x_n)$$

Kurzform:

$$\pi(\theta | D) \propto \pi(\theta) \cdot l(\theta; D)$$

Verwendung der A-posteriori-Verteilung

(19)

Bedingungen:

$$X \sim f(\cdot | \theta)$$

$$\pi(\cdot), D \text{ Daten}, \pi(\cdot | D)$$

Prädikatverteilung:

Prognose für $X|D$ durch die Randverteilung

von X aus der gemeinsamen Verteilung von

$$(X, \tilde{\theta}) \sim f(x|\theta)\pi(\theta|D)$$

Randverteilung von $(X, \tilde{\theta})$ nach θ

$$f(x|D) = \int_{\Theta} f(x|\theta)\pi(\theta|D) d\theta$$

$f(\cdot|D)$ heißt Prädikatsdichte

A-posteriori-Bayes-Schätzer: $\exists[\tilde{\theta}|D]$ falls er existiert ist der Schätzer $\hat{\theta}$

$$\hat{\theta} = \int_{\Theta} \theta \pi(\theta|D) d\theta$$

$$\hat{\tau}(\theta) := \int_{\Theta} \tau(\theta) \pi(\theta|D) d\theta$$

HPD-Bereiche

Bedingungen:

$$\theta \in \Theta \subseteq \mathbb{R}^k, \text{ Daten } D \Rightarrow \pi(\cdot|D)$$

$$\text{Sicherheit: } 1 - \alpha$$

Formel:

HPD-Vertrauensbereich ist eine Teilmenge von Θ

$$\Theta^* \subseteq \Theta$$

$$\int_{\Theta^*} \pi(\theta|D) d\theta = 1 - \alpha$$

Θ^* , $\pi(\theta|D)$ ist darin größtmöglich

$$\pi(\theta|D) \geq c \quad ; \quad c \text{ größtmögliche Konstante}$$

A-posteriori-Wahrscheinlichkeiten von statistischen Hypothesen

Bedingungen:

$$X \sim W_{\theta}, \pi(\cdot), D, \pi(\cdot|D)$$

$$\text{Parameterhypothese: } \mathcal{H}_0: \theta \in \Theta_0 \subset \Theta$$

$$\mathcal{H}_1: \theta \in \Theta_1 \subset \Theta \setminus \Theta_0$$

$$\alpha_0 := W\{\tilde{\theta} \in \Theta_0 | D\}, \alpha_1 := W\{\tilde{\theta} \in \Theta_1 | D\}$$

$$W\{\Theta_j | D\} = \int_{\Theta_j} \pi(\theta|D) d\theta$$

relative Plausibilität $\frac{\alpha_1}{\alpha_0}$

$$\frac{\int_{\Theta_1} \pi(\theta) l(\theta; D) d\theta}{\int_{\Theta_0} \pi(\theta) l(\theta; D) d\theta}$$

Bayes'sche Entscheidungen

Verbindung mit Nutzen

θ ist die Unbekannte, SP $D = (x_1, \dots, x_n) \Rightarrow \pi(\cdot|D)$ von $\tilde{\theta}$

Entscheidungskriterium ist der zu erwartende Verlust $\mathbb{E}_{\pi(\cdot|D)} L(\tilde{\theta}, d_j)$

Die Bayes'sche Entscheidung ist jene mit dem kleinsten a-posteriori zu erwartenden

Verlust

Wirtschaftsentscheidungen sind jene, die den Nutzen Maximieren

Nutzenfunktion: $U(\theta, d) \Rightarrow \bar{U}(d) = \mathbb{E}_{\pi(\theta|D)} U(\tilde{\theta}, d)$

optimale Entscheidung: $\bar{U}(d) = \mathbb{E}_{\pi(\theta|D)} U(\tilde{\theta}, d) \Rightarrow \bar{U}(d_{\text{opt}}) = \max \{\bar{U}(d) : d \in \mathcal{D}\}$

Bayes-Schätzer bezüglich Verlust

Bedingungen:

θ ist optimaler Schätzwert

$$X \sim W_\theta$$

Formel:

Verlustfunktion definiert aus $L(\theta, \hat{\theta})$

$\hat{\theta} = \psi(x_1, \dots, x_n)$ Schätzwert für θ

Bayes-Schätzer für 1-dim ist θ bezüglich $L(\cdot, \cdot)$,

der $\mathbb{E}_{\pi(\cdot|D)} L(\theta, \hat{\theta})$ des Verlusts minimiert

Wahrscheinlichkeit eines Verlusts wird minimiert

Satz:

Bayes-Schätzer $\hat{\theta}_B$ für 1-dim Parameter θ ist

unter quadratischer Verlustfunktion gleich dem

Erwartungswert der A-posteriori-Verteilung

$$\hat{\theta}_B = \mathbb{E}_{\pi(\cdot|D)} \tilde{\theta}$$

Satz

Bayes-Schätzer bei $L(\theta, \hat{\theta}) = |\theta - \hat{\theta}|$ (Verlustfunktion)

gleich dem Median von $\pi(\cdot|D)$ von $\tilde{\theta}$

Bayes-Tests

Wahrscheinlichkeiten von Parameterhypothesen sind mit $\pi(\cdot|D)$ von $\tilde{\theta}$ berechenbar

$$\mathcal{H} : \theta \in \Theta_0, \quad W \{ \tilde{\theta} \in \Theta_0 | D \}$$

UNSCHARFE INFORMATIONEN

15

Unschärfe Zahlen

Beschreibung:

reelle Zahlen x_0 durch Indikatorfunktion charakterisiert $I_{\{x_0\}}(\cdot)$

unscharfe Zahl x^* durch charakterisierende Funktion $\xi(\cdot)$

$$\xi: \mathbb{R} \rightarrow [0, 1]$$

$$\exists x_0 \in \mathbb{R} : \xi(x_0) = 1$$

$$\forall \delta \in (0, 1] \text{ ist } B_\delta(x^*) := \{x \in \mathbb{R} : \xi(x) \geq \delta\}$$

δ -Schnitt ist ein endliches abgeschlossenes Intervall

$$[a_\delta, b_\delta] \text{ bei } \delta \in (0, 1]$$

unscharfe Zahlen sind Spezialfälle von fuzzy sets

$\int_{\mathbb{R}} \xi(x) dx$ ist das Maß der Unschärfe

Unschärfe Stichproben

Beschreibung:

konkrete SP von kontinuierlicher SG sind eine

endliche Folge von unscharfen Zahlen $x_1^*, \dots, x_n^*$

Kombination von $x_1^*, \dots, x_n^*$ zu unscharfen Vektor

Formel:

X hat M_X , SP hat M_X^n

unscharfer Vektor durch $\xi: \mathbb{R}^n \rightarrow [0, 1]$ definiert

Kombination von $\xi_1(\cdot), \dots, \xi_n(\cdot)$:

$$\xi(x_1, \dots, x_n) := \min_{i=1, \dots, n} \xi_i(x_i) \quad \text{geringste Unschärfe}$$

$$\delta\text{-Schnitt von } \xi(\cdot, \dots, \cdot) \Rightarrow \underline{C}_\delta(x^*) := \bigcap_{i=1}^n C_\delta(x_i^*)$$

Verallgemeinerte Schätzungen

Beschreibung:

Schätzfunktion $v: M_x^n \rightarrow \Theta$

$$\hat{\theta} = v(x_1, \dots, x_n)$$

$$\hat{\theta}^* = v(x_1^*, \dots, x_n^*)$$

Durch das Erwartungsprinzip erhält man $\psi(\cdot)$ von $\hat{\theta}^*$
 charakterisierende Funktion $\psi(\theta)$ ist definiert durch die Obergrenze der Schätzfunktion

Formel:

$$\psi(\theta) = \begin{cases} \sup \{ \xi(x) : v(x) = \theta \} & \text{falls } v^{-1}(\theta) \neq \emptyset \\ 0 & \text{falls } v^{-1}(\theta) = \emptyset \end{cases}$$

für $\theta \in \mathbb{R}$ und stetige Funktion $v(\cdot, \dots, \cdot)$ ist abiges

$\psi(\cdot)$ die charakterisierende Funktion einer unscharfen Zahl

Verallgemeinerte Addition

Bedingung:

x^* hat $\xi(\cdot)$, y^* hat $\eta(\cdot)$

$x^* \oplus y^*$ hat $\psi(\cdot)$.

Formel:

$$\psi(z) := \begin{cases} \sup_{(x,y)} \{ \min [\xi(x), \eta(y)] : x+y=z \} \\ 0 & \text{falls } \nexists (x,y) : x+y=z \end{cases}$$

1 Multiplikation mit Konstanten

Bedingung:

x^* hat $\xi(\cdot)$, $c \in \mathbb{R} \Rightarrow c \cdot x^*$ hat $\eta(\cdot)$

Formel:

$$\eta(x) := \xi\left(\frac{x}{c}\right)$$

Satz, wenn charakterisierende Funktionen von selber Gestalt

Bedingung:

$\xi(\cdot)$ einer SP von selber Gestalt und symmetrisch

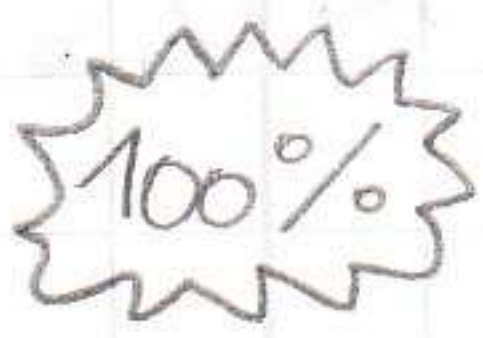
unterschiedliche Lageparameter m_i

charakterisierende Funktion des Stichprobenmittels

$$\bar{X}_n^* := \text{Lageparameter } m = \frac{1}{n} \sum_{i=1}^n m_i \text{ und}$$

der selben Gestalt

Diskrete Verteilungen



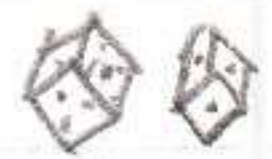
Dirac - Verteilung

δ_μ

$\mu \in \mathbb{R}, X \sim \delta_\mu$

$$W(\{\mu\}) = 1$$

$$W(\{x\}) = 0 \quad \forall x \neq \mu$$



Diskrete Gleichverteilung

D_m

$m \in \mathbb{N}, X \sim D_m$

$$M_x = \{1, \dots, m\}$$

$$W(\{x\}) = \frac{1}{m}$$

Beispiel: Würfel



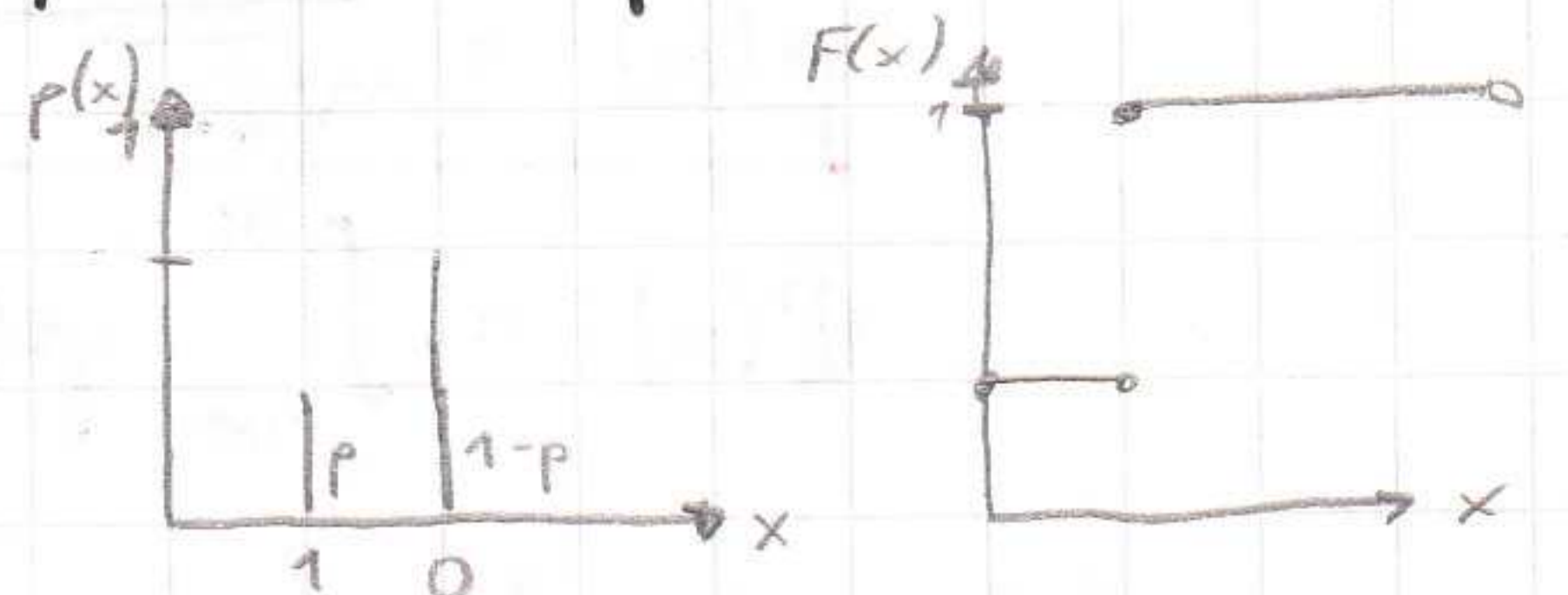
Alternativverteilung

A_p

$p \in (0,1), X \sim A_p$

$$M_x = \{0,1\}$$

$$W(\{1\}) = p, \quad W(\{0\}) = 1-p$$



Binomialverteilung

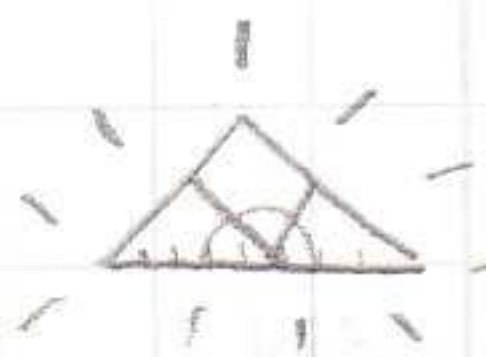
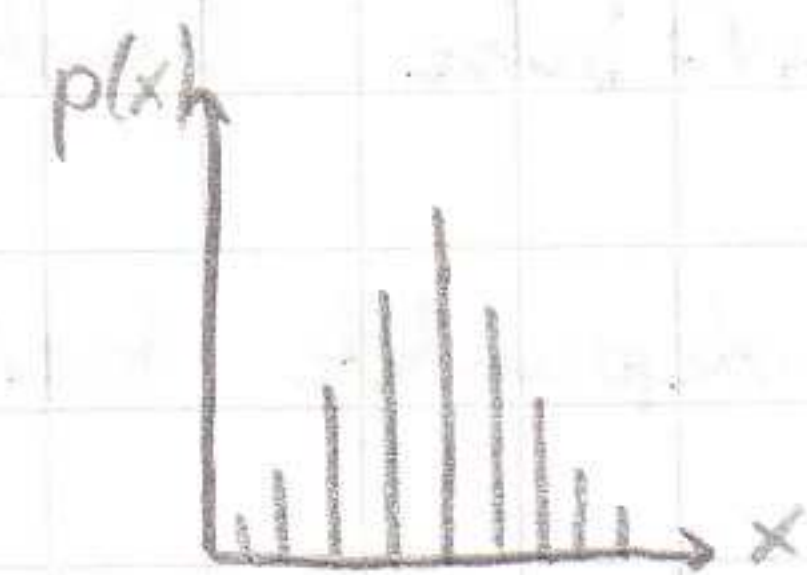
$B_{n,p}$

$n \in \mathbb{N}, p \in (0,1), X \sim B_{n,p}$

n-malige Durchführung eines Alternativversuchs

$$M_x = \{1, \dots, n\}$$

$$W(\{k\}) = \binom{n}{k} p^k (1-p)^{n-k}$$



Hypergeometrische Verteilung

$H_{N,A,n}$

$N \in \mathbb{N}, A \in \mathbb{N}, n \in \mathbb{N}, X \sim H_{N,A,n}$

Ziehen ohne Zurücklegen

Beispiel: Stichprobe, zerstörende Prüfung

X ist die Anzahl der Ausschussstücke in der Stichprobe

$$M_x = \{a_1, a_1+1, \dots, a_2\}$$

$$a_1 = \max\{0, n - (N-A)\}$$

$$a_2 = \min\{n, A\}$$

mögliche Auswahl von n Elementen

$$W(\{a\}) = \frac{\binom{A}{a} \binom{N-A}{n-a}}{\binom{N}{n}}$$

Anzahl der Ausschüsse, Beginn

Anzahl, Ende



Poisson - Verteilung

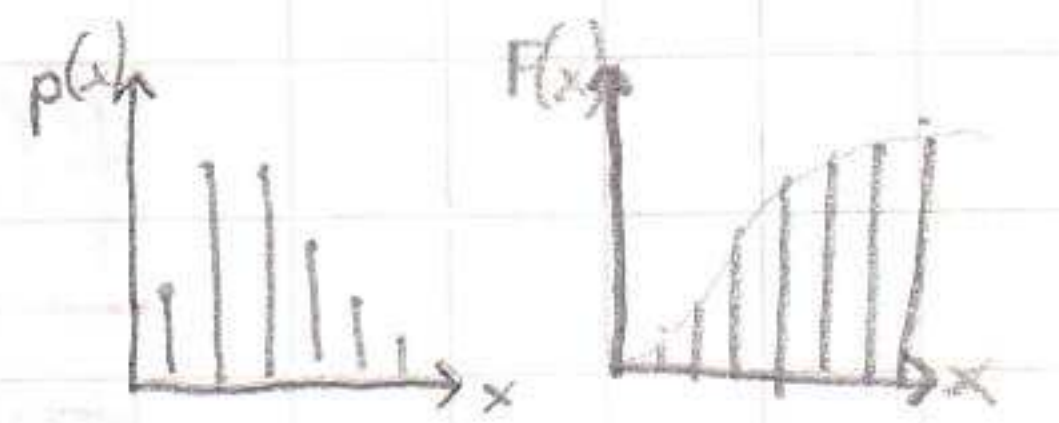
P_μ

$\mu > 0, X \sim P_\mu$

$$M_x = \mathbb{N}$$

$$W(\{k\}) = \frac{\mu^k e^{-\mu}}{k!}$$

Beispiel: Versicherung



Geometrische - Verteilung

G_p

$p \in (0,1), X \sim G_p$

X ist die Anzahl der Durchführungen einer unabhängigen Alternativverteilung bis 1 auftritt

$$M_x = \mathbb{N}$$

$$W\{X=n\} = p \cdot (1-p)^{n-1}$$

Kontinuierliche Verteilungen



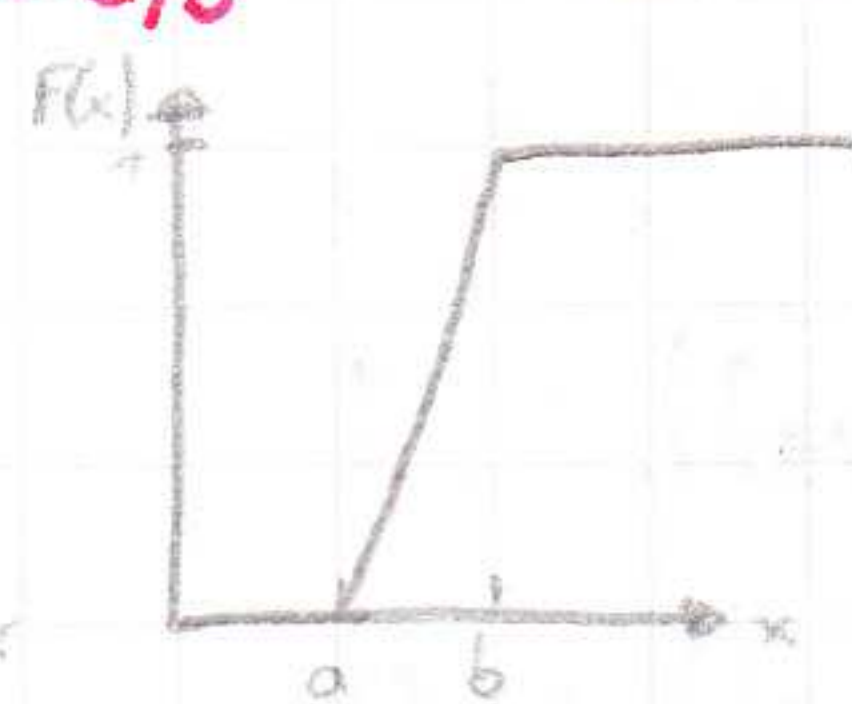
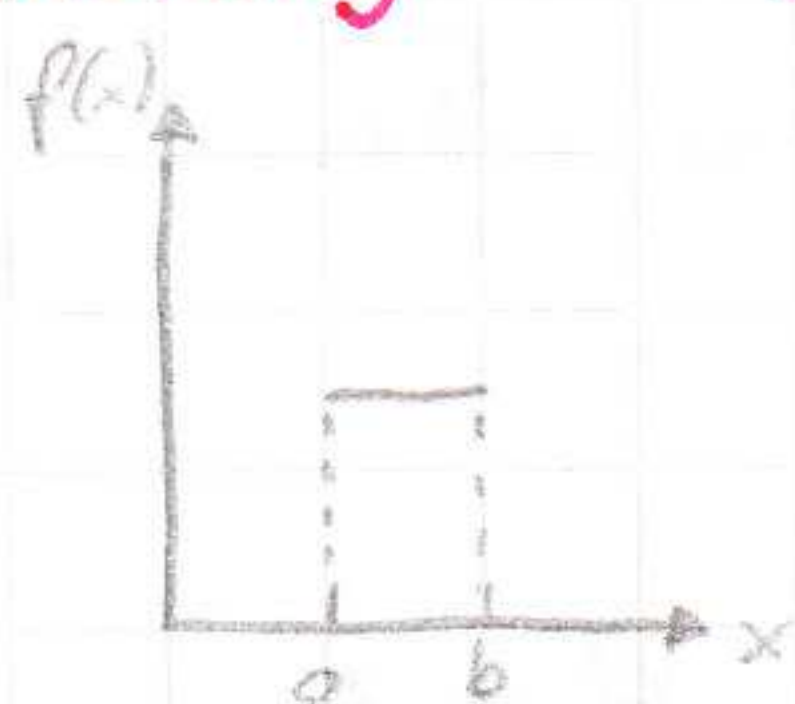
Uniforme / kontinuierliche Gleichverteilung

$U_{a,b}$

$a < b \in \mathbb{R}, X \sim U_{a,b}$

$$f(x) = \frac{1}{b-a} I_{(a,b)}(x)$$

Sonderfall $U_{0,1}$



(EX)

Exponentialverteilung

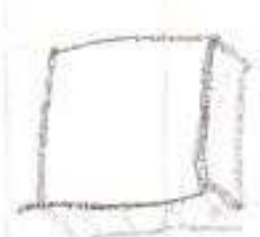
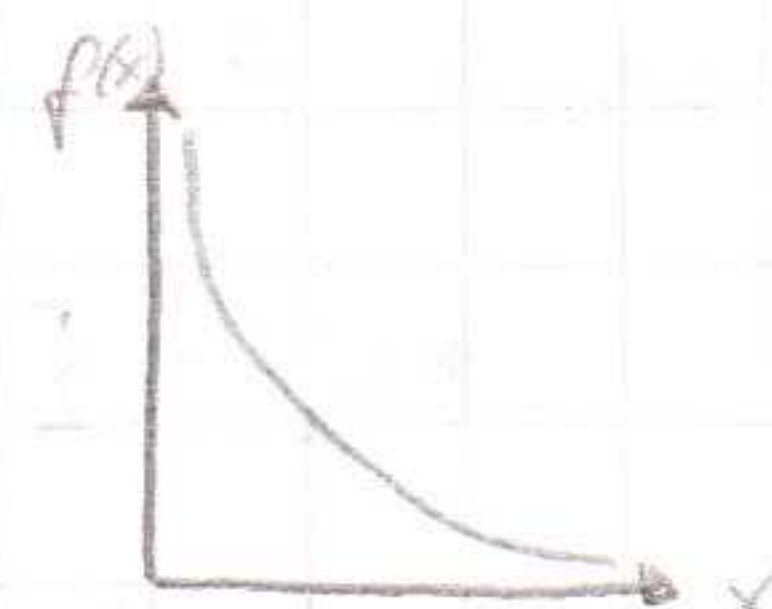
Ex_τ

$\tau > 0, X \sim Ex_\tau$

$$f(x) = \frac{1}{\tau} e^{-\frac{x}{\tau}} I_{(0,\infty)}(x)$$

$$F(x) = (1 - e^{-\frac{x}{\tau}}) I_{(0,\infty)}(x)$$

Beispiel: Wartezeit



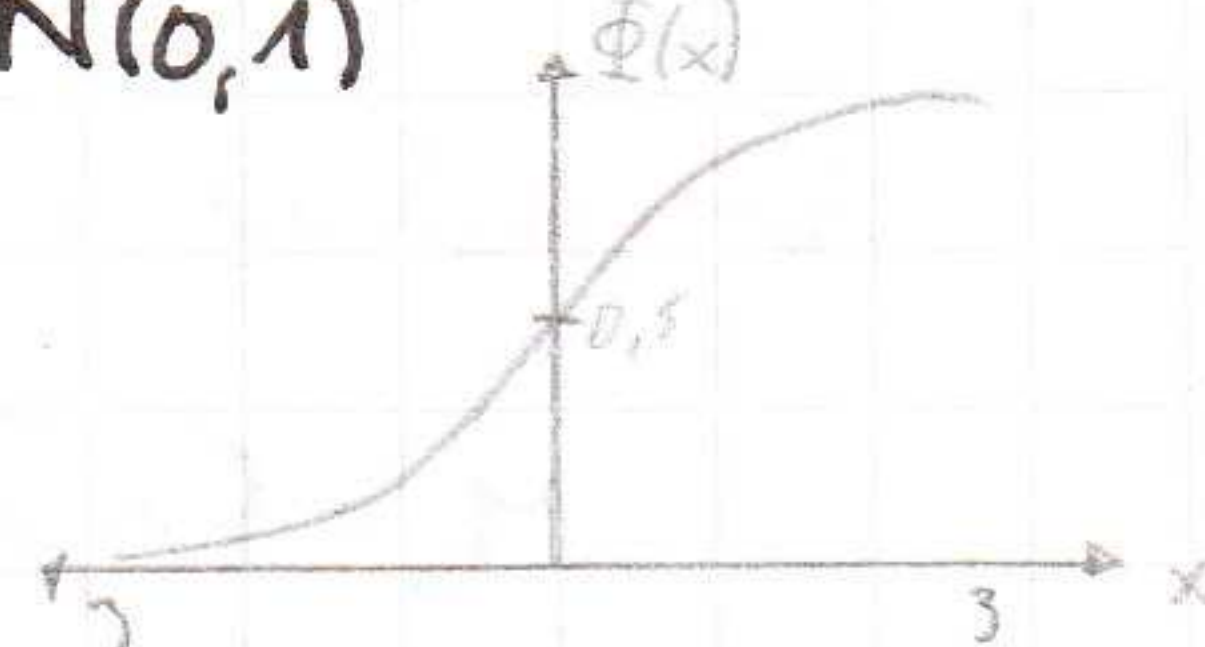
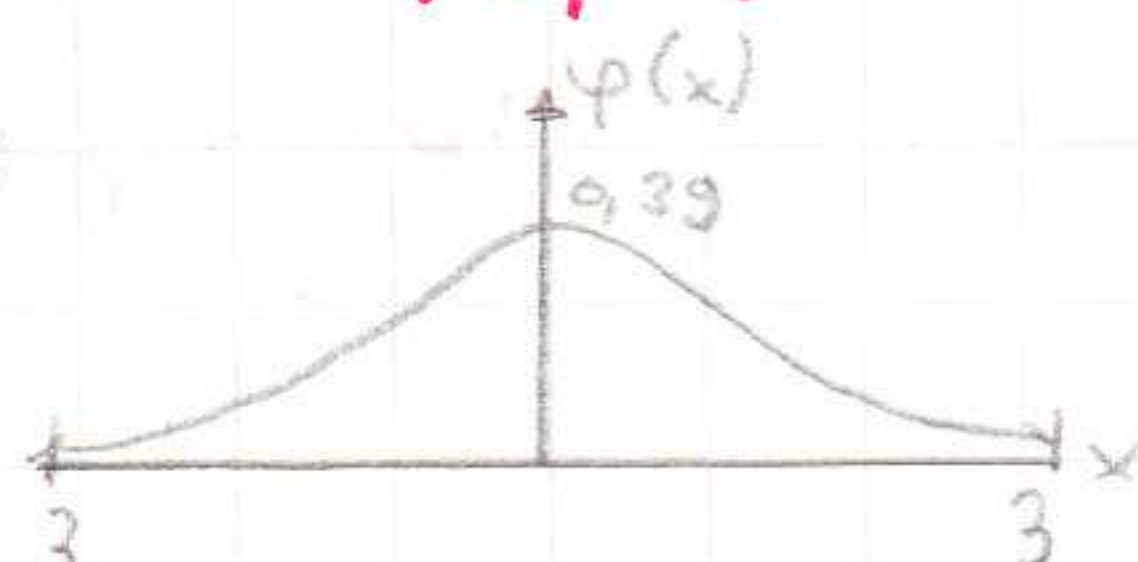
Standard - Normalverteilung

$N(0,1)$

$X \sim N(0,1)$

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

$$F(x) = \Phi(x) = \int_{-\infty}^x \varphi(x) dx$$



Allgemeine Normalverteilung

$N(\mu, \sigma^2)$

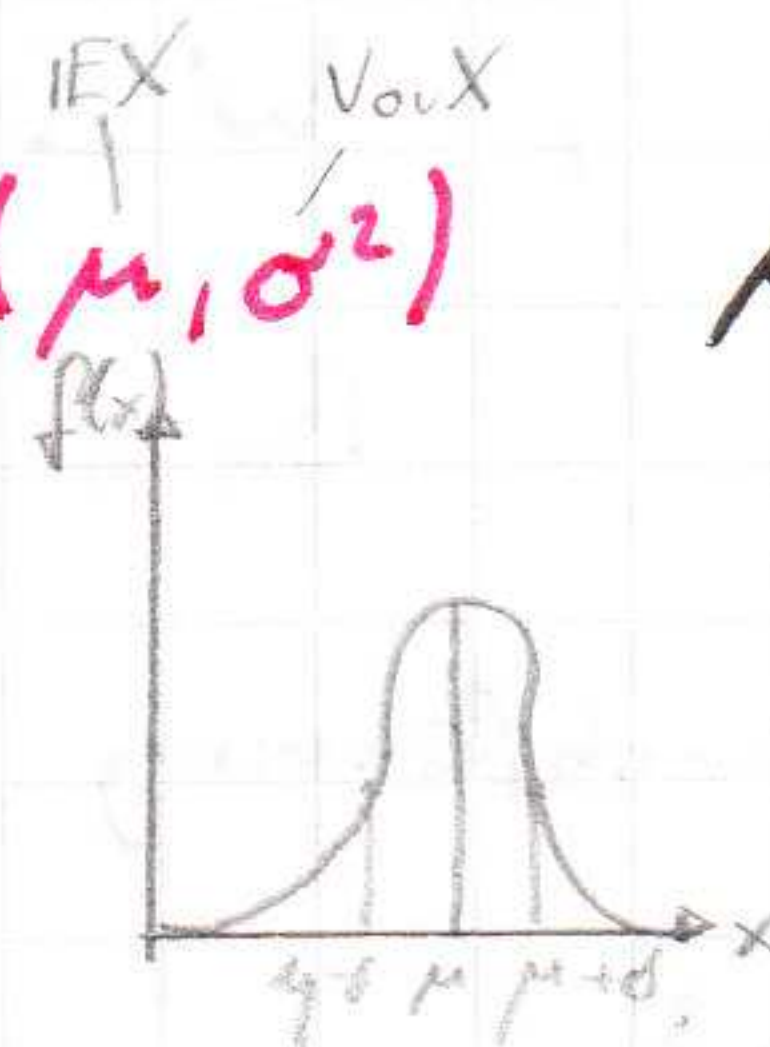
$\mu \in \mathbb{R}, \sigma^2 > 0, X \sim N(\mu, \sigma^2)$

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

Gauß'sche Glockenkurve

Wendepunkte bei $\mu \pm \sigma$

$$F(x | \mu, \sigma^2) = \Phi\left(\frac{x-\mu}{\sigma}\right)$$



log

Logarithmische Normalverteilung

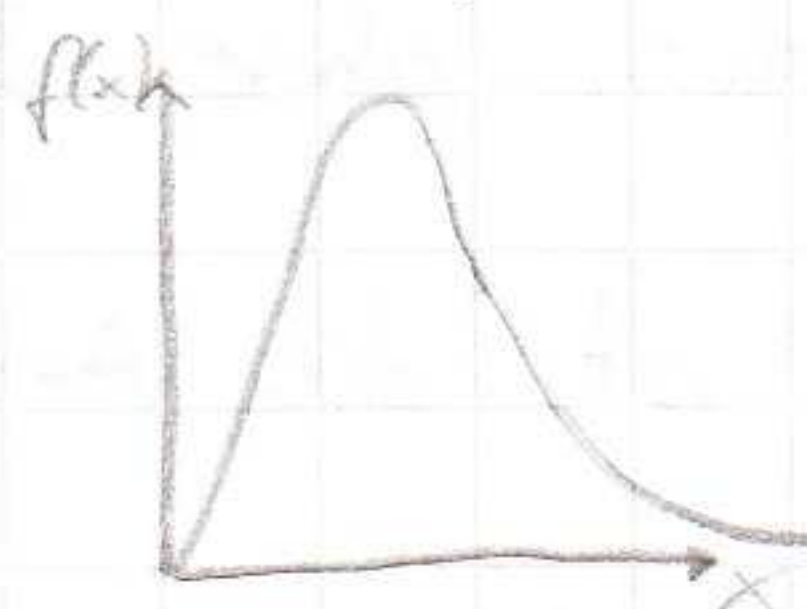
$LN(\mu, \sigma^2)$

$\mu \in \mathbb{R}, \sigma^2 > 0, X \sim LN(\mu, \sigma^2)$

$$\ln X \sim N(\mu, \sigma^2)$$

$$F_{\mu, \sigma^2}(x) = \Phi\left(\frac{\ln x - \mu}{\sigma}\right) \cdot I_{(0,\infty)}(x)$$

$$f(x) = \frac{1}{x\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{\ln x - \mu}{\sigma}\right)^2} I_{(0,\infty)}(x)$$



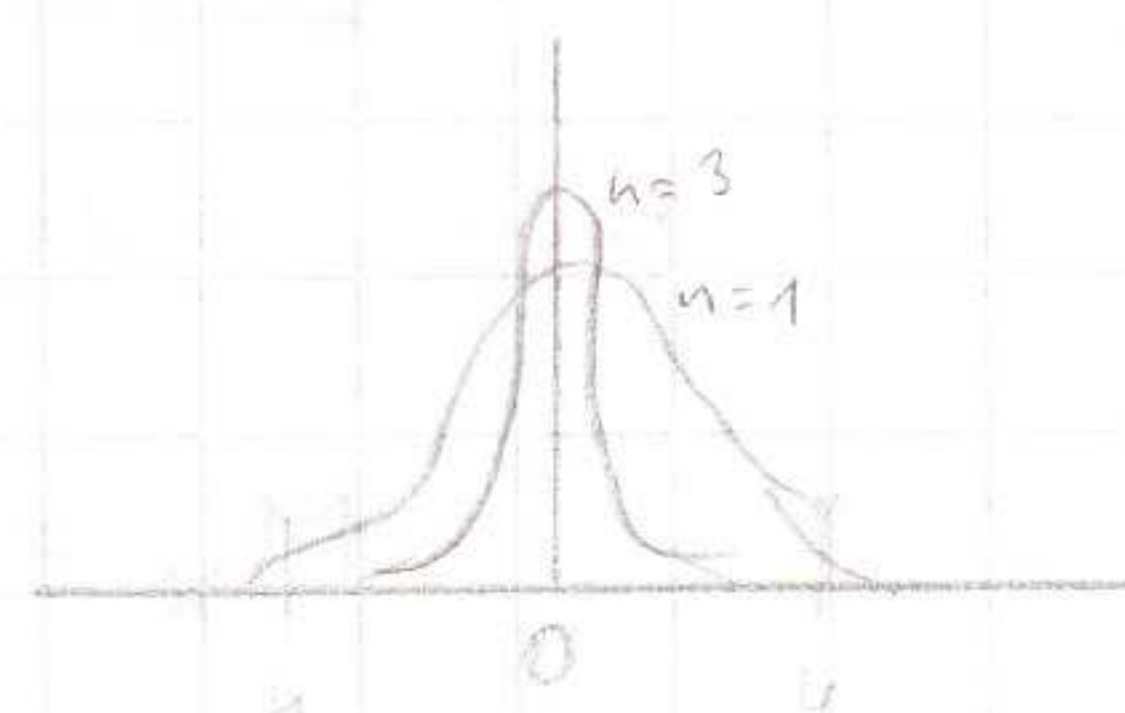
t_n - Verteilung

t_n

$n \in \mathbb{N}, X \sim t_n$

$$f(x) = \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})\sqrt{n\pi}} \cdot \frac{1}{(1+\frac{x^2}{n})^{\frac{n+1}{2}}} \quad \forall x \in \mathbb{R}$$

$$EX = 0 \quad \text{Var} X = \frac{n}{n-2}$$



$t_n \xrightarrow{n \rightarrow \infty} N(0,1)$



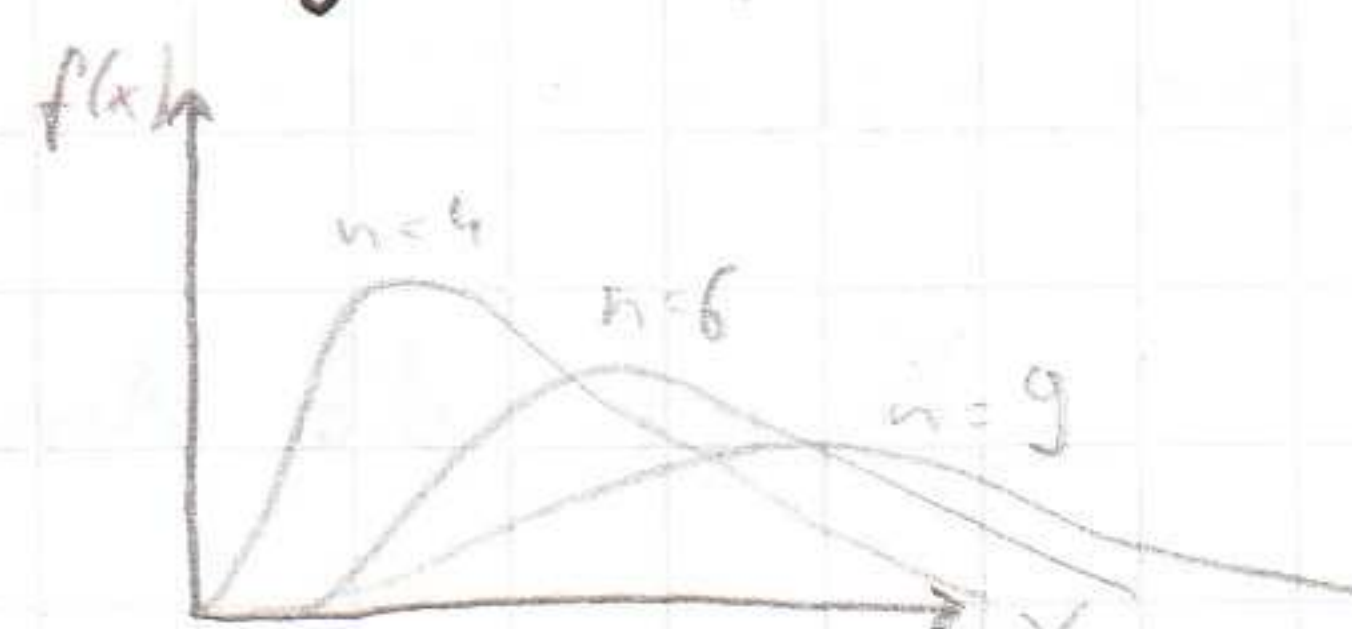
χ_n^2 - Verteilung

χ_n^2

$n \in \mathbb{N}$ (Freiheitsgrade), $X \sim \chi_n^2$

$$f(x) = \frac{x^{\frac{n}{2}-1} e^{-\frac{x}{2}}}{\Gamma(\frac{n}{2}) \cdot 2^{\frac{n}{2}}} I_{(0,\infty)}(x)$$

$$EX = n \quad \text{Var} X = 2n$$



Sätze & Gesetze

Gesetz der großen Zahlen

Die relative Häufigkeit des Ereignisses A konvergiert fast sicher gegen die tatsächliche Wahrscheinlichkeit des Ereignisses A bei unendlicher Versuchsdurchführung

$$\underline{h_n(A) \xrightarrow[n \rightarrow \infty]{f.s.} p = W(A)}$$

$\bar{X}_n$ = SG vor Durchführung des Versuchs

X_n = SG nach Durchführung

X_0 = echte bzw. tatsächliche SG

Folge $X_1, X_2, \dots$ von SG X_n konvergiert fast sicher gegen SG X_0

$$\underline{\bar{X}_n \xrightarrow[n \rightarrow \infty]{f.s.} X_0}$$

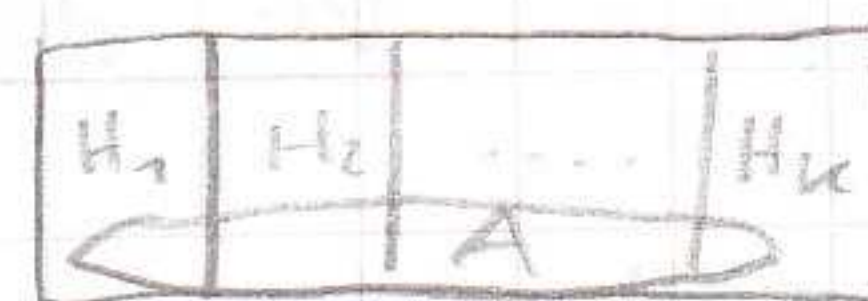
Satz der vollständigen Wahrscheinlichkeit

In jedem Wahrscheinlichkeitsraum $(M, \mathcal{E}, W)$ gilt für

eine endliche oder abzählbare Folge von paarweisen disjunkten

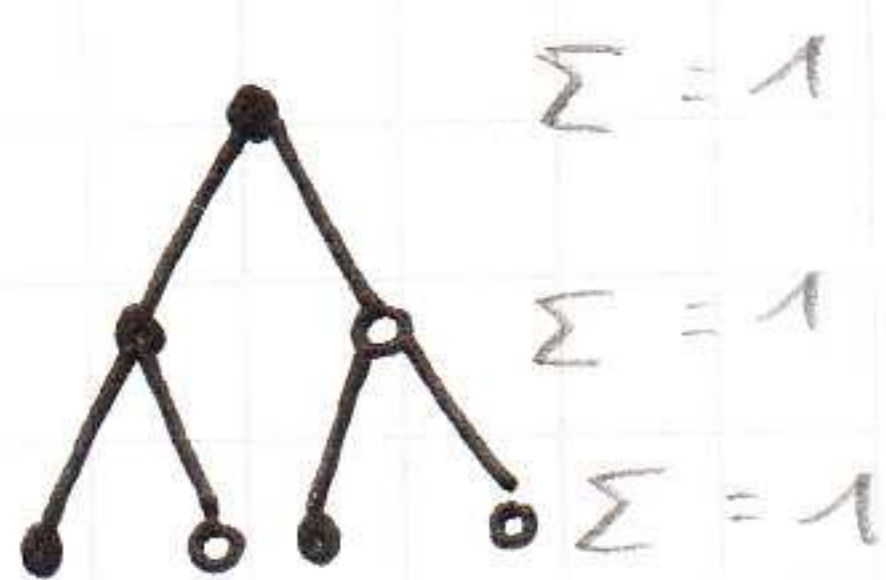
Ereignissen mit $W\left(\bigcup_{k=1}^N H_k\right) = 1$ und $W(H_k) > 0$:

$$\underline{W(A) = \sum_{k=1}^N \left(W(H_k) \cdot W(A | H_k) \right)}$$



Wahrscheinlichkeit von A ist die Summe aus dem Produkten von der Wahrscheinlichkeit eines Ereignisses mal der Wahrscheinlichkeit von A unter der Vorbedingung des Ereignisses.

$$\text{Beispiel: } W(A) = W(A|B) \cdot W(B) + W(A|\bar{B}) \cdot W(\bar{B})$$



Satz vom unbewussten Statistiker

Bei einer diskret verteilten SG lässt sich bei einer gegebenen messbaren Funktion ψ (stetige, monotone Funktionen sind messbar)

der Satz vom unbewussten Statistiker anwenden:

$$\underline{\mathbb{E}[\psi(X)] = \sum_{x: p(x) > 0} \psi(x) p(x)}$$

Erwartungswert ist die Summe der Einzelwahrscheinlichkeiten mal dem Funktionswert von $\psi(x)$

Bei einer kontinuierlich verteilten SG:

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

Bei stochastischen Vektor ist der Satz auch anwendbar

sofern $\psi: \mathbb{R}^m \rightarrow \mathbb{R}$

$$\underline{\mathbb{E} \psi(X_1, \dots, X_m) = \sum_{(x_1, \dots, x_m) \in M_x} \psi(x_1, \dots, x_m) p(x_1, \dots, x_m)}$$

$$\underline{\mathbb{E} \psi(X_1, \dots, X_m) = \int_{\mathbb{R}^m} \dots \int \psi(x_1, \dots, x_m) f(x_1, \dots, x_m) dx_1 \dots dx_m}$$

Zentraler Grenzwertsatz

Der Satz besagt, dass eine ∞ Folge von SG, deren Varianz existiert und deren VF die Lindeberg-Bedingungen erfüllen, bei der Bildung ihrer standardisierten Summen eine VF ergeben, die gegen die Standard-Normalverteilung konvergiert.

Lindeberg-Bedingung: • SG sind unabhängig

- keine VF hat dominierenden Einfluss auf Gesamtverteilung
statistische Versuche beliebig oft wiederholbar
- UIV-Folgen erfüllen Bedingungen

standardisierte Summen

$$\underline{Z_n := \frac{\sum_{k=1}^n X_k - \mathbb{E}(\sum_{k=1}^n X_k)}{\sqrt{\text{Var}(\sum_{k=1}^n X_k)}}$$

Fundamentalsatz der Statistik

Der Fundamentalsatz besagt, dass die empirische VF $F_n^*(\cdot)$, bei der die Folge (UIV-Folge) konkrete Werte annimmt, fast sicher gegen die tatsächliche VF $F(\cdot)$ der SGn X_i konvergiert.

$$\mathbb{E}(F_n^*(x; X_1, \dots, X_n)) = F(x)$$

Erwartungswert der empirischen VF_s ist die tatsächliche VF

$$F_n^*(x; X_1, \dots, X_n) \xrightarrow[n \rightarrow \infty]{\text{f.s.}} F(x)$$

Fundamentalsatz bei UIV-Folgen mit $X_i \sim F(\cdot)$

$$W \left\{ \lim_{n \rightarrow \infty} \sup_{x \in \mathbb{R}} |F_n^*(x) - F(x)| = 0 \right\} = 1$$

Folge der $F_n^*(\cdot)$ konvergiert mit einer Wahrscheinlichkeit von 1 gegen die tatsächliche Verteilungsfunktion $F(\cdot)$

Abweichung muss bei $\lim_{n \rightarrow \infty}$ zwischen $F_n^*(\cdot)$ und $F(\cdot)$ nicht mehr angenommen werden

Es handelt sich hierbei um die fast sichere gleichmäßige Konvergenz

Satz von Gauß-Markoff

Regressionsgerade: $\overset{\text{Regressionsfunktion}}{\eta(x, \alpha, \beta)} = \alpha + \beta x$ $\theta = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}, x \in [a, b]$

$\overset{\text{Regressionsmodell}}{Y_x} = \alpha + \beta x + U_x$ mit konstanter Varianz

$$\text{Var } U_x \equiv \sigma^2, \quad \mathbb{E}Y_x = \alpha + \beta x, \quad \mathbb{E}U_x = 0$$

Daten (x_i, y_i) $i = 1(1)n$, mind. zwei verschiedene x_i

Methode der kleinsten Abstandsquadratsumme:

$$\sum_{i=1}^n (y_i - \hat{\alpha} - \hat{\beta}x_i)^2 = \underset{\alpha, \beta}{\text{Min}} \sum_{i=1}^n (y_i - \alpha - \beta x_i)^2$$

Satz: SP (x_i, y_i) mit Voraussetzungen

1) $x_1, \dots, x_n$ sind nicht alle gleich

2) $y_1, \dots, y_n$ sind unkorreliert

3) $\text{Var } Y_i \equiv \sigma^2$ Varianzen sind gleich

4) $\mathbb{E}Y_i = \alpha + \beta x_i$

α, β reelle Konstanten, α, β, σ^2 unbekannt

$$A = \frac{\left(\sum_{i=1}^n x_i^2 \right) \cdot \left(\sum_{i=1}^n y_i \right) - \left(\sum_{i=1}^n x_i \right) \left(\sum_{i=1}^n x_i y_i \right)}{n \left(\sum_{i=1}^n x_i^2 \right) - \left(\sum_{i=1}^n x_i \right)^2}$$

$$B = \frac{n \left(\sum_{i=1}^n x_i y_i \right) - \left(\sum_{i=1}^n x_i \right) \cdot \left(\sum_{i=1}^n y_i \right)}{n \left(\sum_{i=1}^n x_i^2 \right) - \left(\sum_{i=1}^n x_i \right)^2}$$

$$\text{Var A} = \frac{\sum_{i=1}^n x_i^2}{\text{Nenner}} \sigma^2$$

$$\text{Var B} = \frac{n}{\text{Nenner}} \sigma^2$$

$$\text{Cov(A,B)} = - \frac{\sum_{i=1}^n x_i}{\text{Nenner}} \sigma^2$$