Beschreibende Statistik

Lageparameter

Arithmetisches Mittel

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{1}{n}(x_1 + \cdots + a_n)$$

Geometrisches Mittel

$$\bar{x}_{geom} = \sqrt[n]{x_1 \cdot x_2 \cdot \cdots \cdot x_n}$$

Median

$$\tilde{x} = \begin{cases} x_{\frac{n+1}{2}} \\ \frac{1}{2}(x_{\frac{n}{2}} + x_{\frac{n}{2}+1}) \end{cases}$$

Der Median $\tilde{x}$ minimiert die Funktion

$$\sum_{i=0}^{n} |x_i - \tilde{x}|$$

Streuungsmaße

(empirische) Varianz

$$var = \sigma^2 = {s_n}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \tilde{x})^2$$

Standardabweichung

$$\sigma = \sqrt{\sigma^2}$$

$$\sigma = \sqrt{{s_n}^2}$$

Kovarianz und Korrelationskoeffizient

Kovarianz

$$cov(x,y) = S_{xy} = \frac{1}{n-1}\sum_{i=1}^{1}(x_i - \bar{x}) \cdot (y_i - \bar{y})$$

Korrelationskoeffizent

$$r_{xy} = \frac{S_{xy}}{S_x \cdot S_y}$$

Der Korrelationskoeffizent liegt immer zwischen $-1 \leq r \leq 1$

Regressionsrechnung

Regressionsgerade

Variante 1

$$y = \bar{y} + \frac{S_{xy}}{{\sigma_x}^2} \cdot (x - \bar{x})$$

Variante 2

$$y = b + a \cdot x$$

$$a = \frac{S_{xy}}{{\sigma_x}^2} \text{ und } b = \bar{y} - a \cdot \bar{x}$$

Wahrscheinlichkeitstheorie

Wahrscheinlichkeitsräume

Der Wahrscheinlichkeitsbegriff

$$Ergebnismenge = \Omega$$

Beispiel Würfel $\Omega = \{1,2,3,4,5,6\}$

Ein Ereignis ist eine Teilmenge der Ergebnismenge

$\varnothing \subseteq \Omega \mathrel{\widehat{=}}$ unmögliches Ereignis

$\Omega \subseteq \Omega \mathrel{\widehat{=}}$ sicheres Ereignis

$A = \{1,2,3\}$ Ereignis

$\bar{A} = \{4,5,6\}$ Gegenereignis

Elementarereignis

einelementige Teilmenge von Ω

Ereignis, eine 3 werfen

$$B = \{3\}$$

$$P(\{3\}) = \frac{1}{6}$$

Laplace-Versuch

Jedes Elementarereignis ist gleich wahrscheinlich

$$P(\{\omega_i\}) = \frac{1}{|\Omega|}$$

$$P(A) = \frac{|A|}{|\Omega|} = \frac{3}{6} = \frac{1}{2}$$

Bedingte Wahrscheinlichkeit

Bedingte Wahrscheinlichkeit

Wahrscheinlichkeit für A unter der Bedingung B

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

$$P(\bar{A}|B) = 1 - P(A|B)$$

Formel von Bayes

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}$$

Satz der totalen Wahrscheinlichkeit

$$P(A) = \sum_{i}^{n}(P(A|B_i) \cdot P(B_i)$$

Allgemeine Regeln

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

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

$$P(\bar{A}) = 1 - P(A)$$

$$P(A \mathbin{\bar{\cup}} B) = P(\bar{A} \cap \bar{B})$$

$$P(A \mathbin{\bar{\cap}} B) = P(\bar{A} \cup \bar{B})$$

Wenn A und B unabhängig, dann gilt

$$P(A \cap B) = P(A) \cdot P(B)$$

Zufallsvariablen

Eine Zufallsvariable ist eine Zuordnungsvorschrift die jedem möglichen Ergebnis eines Zufallsexperiments eine Größe zuordnet

$$X = k \mathrel{\widehat{=}} \{\omega \in \Omega | X(\omega = k\}$$

$$X = 3 \mathrel{\widehat{=}} \{\omega \in \Omega | X(\omega = 3\}$$

$$X \leq k \mathrel{\widehat{=}} \{\omega \in \Omega | X(\omega \leq k\}$$

Diskrete Verteilungen

Binomialverteilung

Mit zurücklegen, Wahrscheinlichkeit für jedes Ereignis gleich

$$X \sim B(n,p)$$

$n =$: Stichprobenumfang

$p =$: Wahrscheinlichkeit

(p muss bei Binomialverteilung fest bleiben)

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

$$P(X \leq k) = \sum_{i=0}^{k} \binom{n}{i} \cdot p^i \cdot (1-p)^{n-i}$$

$$P(X > k) = 1 - P(X \leq k)$$

Eingabe Taschenrechner

$$\binom{n}{k} \mathrel{\widehat{=}} n\,|nCr|\,k$$

Binomialverteilung kann mit der Normalverteilung approximiert werden, bedingung ist

$$X \sim B(n,p) \overset{a}{=} N(n \cdot p, n \cdot p \cdot (1-p))$$

falls gilt

$$n \cdot p \cdot (1-p) > 9$$

Hypergeometrische Verteilung

Ohne zurücklegen, Wahrscheinlichkeit ändert sich nach jedem Ereignis

$$X \sim H(N, M, n)$$

$n =$: Stichprobenumfang

$N =$: Gesamtzahl

$M =$: Anzahl der Elemente mit der Eigenschaft

$$P(X = k) = \frac{\binom{M}{k} \cdot \binom{N-M}{n-k}}{\binom{N}{n}}$$

$$P(X \leq k) = \sum_{i=0}^{k} \frac{\binom{M}{i} \cdot \binom{N-M}{n-i}}{\binom{N}{n}}$$

$$P(X > k) = 1 - P(X \leq k)$$

Poisson Verteilung

$$X \sim Pois(\lambda)$$

$$P(X = k) = \pi_\lambda(k) = \frac{\lambda^k}{k!} \cdot e^{-\lambda}$$

Geometrische Verteilung

$$X \sim Geom(n,p)$$

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

Beispiel: Ein Würfel wird so lange gewürfelt bis eine 6 Auftritt. Die Zufallsvariable X ist gleich Anzahl der Würfe

Stetige Verteilungen

Normalverteilung

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

Ist $X \sim N(0,1)$ dann heißt sie Standardnormalverteilt

Jede Normalverteilung kann standisiert werden, das heißt die Kurve wird in die Mitte des Graphen verschoben

Wenn $X \sim N(\mu, \sigma^2)$
dann ist $Z = \frac{x-\mu}{\sigma} \sim N(0,1)$