Spickzettel: Stochastik und Statistik

Daten auswerten

Mittelwert

$$\bar{x} = \frac{x_1 + x_2 + \ldots + x_n}{n}$$

Median

mittlerer Wert der sortierten Liste

bei gerader Anzahl: Mittelwert der beiden mittleren

Modus, Spannweite

häufigster Wert $\qquad x_{\max} - x_{\min}$

Quartile, Boxplot

Q_1, Median, $Q_3 \qquad \text{IQR} = Q_3 - Q_1$

Varianz

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

Standardabweichung s = Wurzel daraus. Bei Stichproben (Schätzung) teilt man meist durch n – 1 statt n.

Relative Häufigkeit

$$h = \frac{\text{absolute Häufigkeit}}{\text{Anzahl}}$$

Wahrscheinlichkeit

Laplace

$$P(E) = \frac{\text{günstige}}{\text{mögliche Ergebnisse}}$$

Gegenereignis

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

„mindestens einmal“ = 1 – „keinmal“

Pfadregeln

entlang des Pfades multiplizieren, Pfade addieren

Bedingte Wahrscheinlichkeit

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

Unabhängigkeit

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

Satz von Bayes

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

Ohne Zurücklegen

$$P(\text{2-mal rot}) = \frac{3}{5} \cdot \frac{2}{4} = \frac{3}{10}$$

3 rote, 2 blaue Kugeln

Vierfeldertafel

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

Zeilen und Spalten addieren sich

Kombinatorik

Fakultät

$$n! = 1 \cdot 2 \cdot \ldots \cdot n \qquad 0! = 1$$

Mit Reihenfolge, mit Zurücklegen

$$n^k$$

Mit Reihenfolge, ohne Zurücklegen

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

Ohne Reihenfolge (Lotto)

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

6 aus 49: 13 983 816

Zufallsgrößen

Erwartungswert

$$E(X) = \mu = \sum x_i \cdot P(X = x_i)$$

Varianz

$$V(X) = \sum (x_i - \mu)^2 \cdot P(X = x_i) \qquad \sigma = \sqrt{V(X)}$$

Binomialverteilung

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

Kenngrößen

$$\mu = n \cdot p \qquad \sigma = \sqrt{n \cdot p \cdot (1-p)}$$

Normalverteilung

$$P(X \le x) = \Phi\left(\frac{x - \mu}{\sigma}\right)$$

Kumuliert

$P(X \le k) = P(X = 0) + \ldots + P(X = k) \qquad P(X \ge k) = 1 - P(X \le k - 1)$

Beurteilende Statistik

Sigma-Regeln

$$P(|X - \mu| \le \sigma) \approx 68{,}3\,\%$$

2sigma: 95,4 %, 3sigma: 99,7 % (wenn sigma > 3)

95-%-Prognoseintervall

$$[\mu - 1{,}96\sigma;\ \mu + 1{,}96\sigma]$$

Konfidenzintervall (Näherung)

$$h \pm 1{,}96 \cdot \sqrt{\frac{h(1-h)}{n}}$$

Hypothesentest

$P(X \in$ Ablehnungsbereich$) \le \alpha$

alpha = Signifikanzniveau

Fehler

1. Art: H_0 fälschlich abgelehnt 2. Art: H_0 fälschlich behalten