Statistik 1

Formelsammlung

Deskriptive Statistik und Häufigkeiten

Formel	Beschreibung
$h_j = \frac{n_j}{n}$	Relative Häufigkeit
$j = \lfloor (n-1) * p + 1 \rfloor$ $g = (n-1) * p + 1 - j$ $(1-g) * x_j + g * x_{j+1}$	Quantil
$\overline{x} = \hat{\mu} = \frac{\sum_{i=1}^{n} x_i}{n}$	Mittelwert
$s_x^2 = \hat{\sigma}_x^2 = \frac{\sum_{i=1}^{n} (x_i - \overline{x})^2}{n-1}$	Varianz
$s_x = \hat{\sigma}_x = \sqrt{s_x^2}$	Standardabweichung
$\hat{\sigma}_{\overline{x}} = \frac{\hat{\sigma}}{\sqrt{n}}$	(geschätzter) Standardfehler des Mittelwerts
$\hat{\sigma}_{\tilde{x}} = \sqrt{\frac{\pi}{2n}} \hat{\sigma}$	(geschätzter) Standardfehler des Medians
$\hat{\sigma}_{\hat{\sigma}^2} = \hat{\sigma}^2 \sqrt{\frac{2}{n-1}}$	(geschätzter) Standardfehler der Varianz
$H = -\frac{1}{\ln k} * \sum_{j=1}^{k} h_j * \ln h_j$	Relativer Informationsgehalt

Formel	Beschreibung
$R = x_{max} - x_{min}$	Spannweite, Range
$IQA = IQR = Q_3 - Q_1$	Interquartilsabstand, Interquartilsrange
$g = \frac{1}{n} * \sum_{i=1}^{n} (\frac{x_i - \overline{x}}{s_x})^3$	Schiefe
$w = \frac{1}{n} * \sum_{i=1}^{n} (\frac{x_i - \overline{x}}{s_x})^4$ $Exzess = w - 3$	Kurtosis, Wölbung

Wahrscheinlichkeit

Formel	Beschreibung
$p(A) = \frac{n_A}{n_\Omega}$	Wahrscheinlichkeit Laplace-Experiment
$p(A \cap B) = p(A\|B) * p(B) = p(B\|A) * p(A)$	Multiplikationstheorem für abhängige Ereignisse
$p(A \cap B) = p(A) * p(B)$	Multiplikationstheorem für unabhängige Ereignisse
$p(A\|B) = \frac{p(B\|A) * p(A)}{p(B)}$	Satz von Bayes

Verteilungen

Formel	Beschreibung
$X \sim \mathcal{B}(n, \pi)$ $f_{x,n,\pi} = p(X = x) = \binom{n}{x} * \pi^x * (1-\pi)^{n-x}$ $F_{x,n,\pi} = p(X \leq x) = \sum_{i=0}^{x} p(X = x_i)$	Binomialverteilung
$X \sim \mathcal{P}ois(\lambda)$ $f_{x,\lambda} = p(X = x) = \frac{\lambda^x}{x!} * e^{-\lambda}$ $F_{x,\lambda} = p(X \leq x) = \sum_{i=0}^{x} p(X = x_i)$	Poissonverteilung
$X \sim \mathcal{E}xp(\lambda)$ $f_{x,\lambda} = \lambda e^{-\lambda x}, \quad x \geq 0$ $F_{x,\lambda} = p(X \leq x) = 1 - e^{-\lambda x}$	Exponentialverteilung
$X \sim \mathcal{N}(\mu, \sigma^2)$ $f_{x,\mu,\sigma^2} = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$ $F_{x,\mu,\sigma^2} = p(X \leq x) = \int_{-\infty}^{x} f_{x,\mu,\sigma^2}(t)\,\mathrm{d}t$	Normalverteilung
$X \sim t(df)$ $f_{df} = \frac{\Gamma\left(\frac{df+1}{2}\right)}{\sqrt{df\pi}\,\Gamma\left(\frac{df}{2}\right)} \left(1 + \frac{x^2}{df}\right)^{-\frac{df+1}{2}}$ $F_{df} = p(X \leq x) = \int_{-\infty}^{x} f_{df}(t)\,\mathrm{d}t$	t-Verteilung
$X \sim \chi^2(df)$ $f_{df} = \frac{1}{2^{\frac{df}{2}}\Gamma(\frac{df}{2})} x^{\frac{df}{2}-1} \exp(-\frac{x}{2})$ $F_{df} = p(X \leq x) = \int_{0}^{x} f_{df}(t)\,\mathrm{d}t$	χ^2-Verteilung

Konfidenzintervalle und Signifikanztests

Formel	Beschreibung
$KI = (-\infty, \bar{x} + z_{1-\alpha} * \hat{\sigma}_{\bar{x}}]$ $KI = [\bar{x} - z_{1-\frac{\alpha}{2}} * \hat{\sigma}_{\bar{x}}, \bar{x} + z_{1-\frac{\alpha}{2}} * \hat{\sigma}_{\bar{x}}]$ $KI = [\bar{x} - z_{1-\alpha} * \hat{\sigma}_{\bar{x}}, \infty)$	Konfidenzintervall des Mittelwerts (auf Normalverteilung basierend
$t_{emp} = \frac{\bar{x} - \mu_0}{\hat{\sigma}_{\bar{X}}}$ $df = n - 1$	Einstichproben-t-Test