Formelsammlung - Stochastik

Antoine Hauck

February 19, 2008

1 Wahrscheinlichkeit

Bedingte W'keit (Bayes Theorem)

$$P\,(A|B) = \frac{P(B|A)\,P(A)}{P(B)} \tag{1}$$

$$P\,(A|B) = \frac{P(B|A)\,P(A)}{P(B|A)\,P(A) + P(B|\bar{A})\,P(\bar{A})} \tag{2}$$

Erwartungswert

$$E[X] = \sum_{i=1}^{n} p(w_i)\,X(w_i) \tag{3}$$

Varianz

$$Var[X] = \sum_{i=1}^{n} \left(X(w_i) - E[X]\right)^2\, p(w_i) \tag{4}$$

Standardabweichung

$$\sigma = \sqrt{Var[X]} \tag{5}$$

Binomial-Verteilung

$$P(X = k) = p_k = \binom{n}{k} p^k\, (1-p)^{n-k} \tag{6}$$

$$E[X] = n\,p \quad , \quad Var[X] = n\,p\,(1-p) \tag{7}$$

Poisson-Vertilung

$$P(X = k) = p_k = \frac{\lambda^k}{k!}\,e^{-\lambda} \tag{8}$$

$$E[X] = \lambda \quad , \quad Var[X] = \lambda \tag{9}$$

2 Regression und Korrelation

Lineare Regression

$$S(a,b) = \sum_{i=1}^{n} (y(x_i) - y_i)^2 = \sum_{i=1}^{n} (a\,x_i + b - y_i)^2 = MIN \tag{10}$$

Minimum definiert durch

$$\frac{\partial S(a,b)}{\partial a} = 0 \quad \wedge \quad \frac{\partial S(a,b)}{\partial b} = 0 \tag{11}$$

daraus resultierende Funktionsparameter

$$a = \frac{\sum_{i=1}^{n} x_i\, y_i - n\,\bar{x}\,\bar{y}}{\sum_{i=1}^{n} x_i^2 - n\,\bar{x}^2} \quad , \quad b = \bar{y} - a\,\bar{x} \tag{12}$$

Wobei

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i \quad , \quad \bar{y} = \frac{1}{n} \sum_{i=1}^{n} y_i \tag{13}$$

Varianz

$$\sigma_x^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2 \tag{14}$$

$$\sigma_y^2 = \frac{1}{n-1} \sum_{i=1}^{n} (y_i - \bar{y})^2 \tag{15}$$

Kovarianz

$$c_{xy} = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})\ (y_i - \bar{y}) \tag{16}$$

Korrelationskoeffizient

$$r_{xy} = \frac{c_{xy}}{\sigma_x\,\sigma_y} \tag{17}$$

Vereinfachte Darstellung

$$a = \frac{c_{xy}}{\sigma_x^2} \tag{18}$$

$$y(x) = \frac{c_{xy}}{\sigma_x^2}\ (x - \bar{x}) + \bar{y} \tag{19}$$

3 Vierfeldertest

	Erfolge	Misserfolge	Summe
Probe A	E_A	M_A	$E_A + M_A$
Probe B	E_B	M_B	$E_B + M_B$
Summe	$E_A + E_B$	$M_A + M_B$	$N_A + N_B = N$

Prüfgrösse

$$\chi^2 = \frac{N \cdot (E_A \cdot M_B - E_B \cdot M_A)^2}{(E_A + E_B) \cdot (M_A + M_B) \cdot (E_A + M_A) \cdot (E_B + M_B)} \tag{20}$$

Ergebnisse sind statistisch signifikant wenn $\chi^2 \geq 3.85$ ($p \leq 0.05$)

$$p = \frac{1}{2} \cdot 10^{\frac{-\chi^2}{3.84}} \tag{21}$$