1 Statistische Maßzahlen

empirische Kovarianz

$$\tilde{s}_{xy} = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})$$

empirische Varianz:

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

Standardabweichung

$$\sigma = \tilde{s} = \tilde{s}_x = \sqrt{\tilde{s}^2}$$

empirischer Korrelationskoeffizient (nach Bravais Pearson)

$$r_{xy} = \frac{\tilde{s}_{xy}}{\tilde{s}_x \cdot \tilde{s}_y}$$

Spearman'sche Rangkorrelationskoeffizient von X und Y

$$r_{SP}(x,y) = \frac{\sum\limits_{i=1}^{n}\Big(rg(x_i) - \overline{rg_x})(rg(y_i) - \overline{rg_y}\Big)}{\sqrt{\sum\limits_{i=1}^{n}(rg(x_i) - r\bar{g}_x)^2 \cdot \sum\limits_{i=1}^{n}(rg(y_i) - r\bar{g}_y)^2}}$$

1.1 Lineare Regression

optimale Ausgleichgerade:

$$y = f(x) = \hat{a}x + \hat{b}$$

$$\hat{a} = \frac{\tilde{s}_{xy}}{\tilde{s}_x^2}$$

$$\hat{b} = \bar{y} - \hat{a}\bar{x}$$

Bestimmtheitsmaß, Modellgüte:

$$R^2 = r_{xy}^2$$

2 Wahrscheinlichkeit

2.1 Allgemein gilt

$P(A \cap B) = P(A) + P(B) - P(A \cup B)$
$P(A \cup B) = P(A) + P(B) - P(A \cap B)$
$P(A \cap \overline{B}) = P(A) - P(A \cap B)$
$P(\overline{A} \cap \overline{B}) = P(\overline{A \cup B}) = 1 - P(A \cup B)$

2.2 für die bedingte Wahrscheinlichkeit gilt

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

2.3 totale Wahrscheinlichkeit

Setzt sich das Ereignis A zusammen aus den sich gegenseitig ausschließenden Ereignissen $A \cap B1$ und $A \cap B2$, so gilt:
$P(A) = P(A|B_1) \cdot P(B_1) + P(A|B_2) \cdot P(B_2)$

3 Zufallsvariablen

3.1 Dichtefuntion

$\mathbb{P}(X = x) = f(x)$ ist einfach die Wahrscheinlichkeit für x.

3.2 Verteilungsfunktion

$\mathbb{P}(X \leq x) = F(x)$ gibt die Wahrscheinlichkeit an, dass X kleiner oder gleich x ist.

3.3 Erwartungswert bei diskreter ZV

Dieser Wert ist als Mittel zu erwarten.

$$\mathbb{E}(X) = \sum_i x_i \cdot f(x_i)$$

3.4 Randdichte

bei diskret: siehe Ränder der Kontingenztabelle.
bei stetig: $f_X(x) = \int_{unten}^{oben} f(x,y)dy$ d.h. die Gemeinsame dichte nach y integrieren um die Randdichte von x zu bekommen.