Univariat

Mittelwert $\bar{x}$ und μ

$$\mu = \frac{1}{N}\sum_i^N (x_i)$$

$$\overline{x} = \frac{1}{n}\sum_i^n (x_i)$$

Varianz und Standardabweichung

$$s^2 = V = \frac{1}{n}\sum_i^n (x_i - \overline{x})^2$$

$$s = \sqrt{\frac{1}{n}\sum_i^n (x_i - \overline{x})^2}$$

$$\hat{\sigma}^2 = \frac{1}{n-1}\sum_i^n (x_i - \overline{x})^2$$

$$\hat{\sigma} = \sqrt{\frac{1}{n-1}\sum_i^n (x_i - \overline{x})^2}$$

Standardfehler

$$s_{\overline{x}} = \frac{s_x}{\sqrt{n}}$$

z-Transformation

$$z_i = \frac{x_i - \overline{x}}{s}$$

Konfidenzintervalle der Mittelwerte

$$\text{KI:}\overline{X} \pm z_1 \cdot se$$

$$\text{KI:}\overline{X} \pm z_1 \cdot \frac{s_x}{\sqrt{n}}$$

$$\text{KI}_{l.05} = \overline{x} - 1.96 \cdot \frac{s_x}{\sqrt{n}}$$

$$\text{KI}_{r.05} = \overline{x} + 1.96 \cdot \frac{s_x}{\sqrt{n}}$$

Bivariat

Kovarianz und Korrelation

$$cov = C = \frac{1}{n}\sum_i^n (x_i - \overline{x})(y_i - \overline{y})$$

$$r = \frac{\sum_i^n (x_i - \overline{x})(y_i - \overline{y})}{n \cdot s_x \cdot s_y}$$

Bivariate Regression

$$Y_i = \overline{Y} + e_i$$

$$Y_i = b_1 + b_2 X_i + e_i$$

$$\hat{Y}_i = b_1 + b_2 X_i$$

$$Y_i = \hat{Y}_i + e_i$$

Korrelation und b

$$r_{YX} = b \cdot \frac{s_X}{s_Y}$$

BETAs

$$BETA = b \cdot \frac{s_X}{s_Y} = r_{YX}$$

Multivariat

Das lineare Regressionsmodell

$$Y_i = \beta_1 + \beta_2 X_{i2} + \beta_3 X_{i3} + U_i$$

$$Y_i = b_1 + b_2 X_{i2} + b_3 X_{i3} + e_i$$

Die b's

$$b_2 = \frac{r_{y2} - r_{23} r_{y3}}{(1 - R^2_{2.3})} \frac{s_y}{s_2} = \frac{(V_3 C_{Y2} - C_{23} C_{Y3})}{(V_2 V_3 - C^2_{23})}$$

Erwartungstreue b_2

$$E(b_2) = \beta_2 + \frac{V_3 E(C_{2U})}{D} - \frac{C_{23} E(C_{3U})}{D}$$

R^2

$$R^2 = \frac{SS_M/n}{SS_T/n}$$

$$R^2_{adj.} = R^2 \cdot \frac{n-k-1}{n-1}$$

Standardfehler der b

$$s^2_{b_2} = \frac{s^2}{n} \frac{1/s^2_2}{1 - R^2_{2.3}}$$

$$s^2 = \frac{1}{n-3}\sum (e_i - \bar{e})^2$$

Toleranz (TOL)

$$TOL_{b_2} = 1 - R^2_{2.3}$$

$$TOL_{b_3} = 1 - R^2_{3.2}$$

Varianz-Inflationsfaktor (VIF)

$$VIF_{b_2} = \frac{1}{TOL_{b_2}} = \frac{1}{1 - R^2_{2.3}}$$

$$VIF_{b_3} = \frac{1}{TOL_{b_3}} = \frac{1}{1 - R^2_{3.2}}$$

Logistische Regression

$$P(y=1) = \frac{1}{1 + e^{-z}}$$

$$Y_i = b_1 + b_2 X_{i2} + b_3 X_{i3} + e_i$$

$$P(Y_i) = \frac{1}{1 + e^{-(b_1 + b_2 X_{2i})}}$$