Zusammenfassung "Einfache lineare Regression"

Datengenerierender Prozess und Schätzgleichung:

$$y_i = \alpha + \beta x_i + u_i \qquad u_i \sim IN(0, \sigma^2)$$

$$y_i = \hat{\alpha} + \hat{\beta} x_i + \hat{u}_i \qquad i = 1, \dots n$$

Werte auf der Regressionsgeraden ("fitted values"):

$$\hat{y}_i = \hat{\alpha} + \hat{\beta} x_i$$

Summe der quadrierten Residuen (Residual Sum of Squares):

$$RSS = \sum_{i=1}^{n} \hat{u}_i^{\,2} = \sum_{i=1}^{n} \left(y_i - \hat{\alpha} - \hat{\beta} x_i \right)^2 \Rightarrow \min$$

Geschätzte Residualvarianz und Standardfehler der Gleichung:

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^{n} \hat{u}_i^2}{n-2} \qquad \hat{\sigma} = \sqrt{\frac{\sum_{i=1}^{n} \hat{u}_i^2}{n-2}}$$

Geschätzter Standardfehler von $\hat{\beta}$:

$$SE(\hat{\beta}) = \sqrt{\frac{\hat{\sigma}^2}{S_{xx}}} \qquad S_{xx} = \sum (x_i - \bar{x})^2$$

$$\frac{\hat{\beta} - \beta}{SE(\hat{\beta})} \sim t_{n-2} \qquad t\text{ - Verteilung}$$

Varianzanalyse:

$$\underbrace{TSS = \sum (y_i - \bar{y})^2 \qquad ESS = \sum (\hat{y}_i - \bar{y})^2 \qquad RSS = \sum \hat{u}_i^{\,2}}_{\text{total} \qquad \text{explained} \qquad \text{residual}}$$

sum of squares

$TSS = ESS + RSS \qquad R^2 = ESS/TSS = 1 - RSS/TSS$