

Einfaches Regressionsmodell

Eine Gerade in einem cartesischen Koordinatensystem können wir durch die Gleichung

$$y_i = b_0 + b_1x_i \quad (1)$$

beschreiben. Im linearen Regressionsmodell schreiben wir

$$\hat{y}_i = b_0 + b_1x_i \quad (2)$$

um zu verdeutlichen, dass die \hat{y}_i -Werte (sprich: y-Dach) nicht die empirisch beobachteten, sondern die aus dem Modell vorhergesagten Werte sind.

Die gesuchte Gerade soll so gewählt werden, dass die quadrierten Abstände zwischen den durch das Modell vorhergesagten und den empirisch beobachteten Werten minimiert werden; es soll also gelten:

$$\sum_{i=1}^n e_i^2 = \sum_{i=1}^n (y_i - \hat{y}_i)^2 = \text{Minimum}$$

. Die Werte $e_i = y_i - \hat{y}_i$ werden als Residuen bezeichnet.