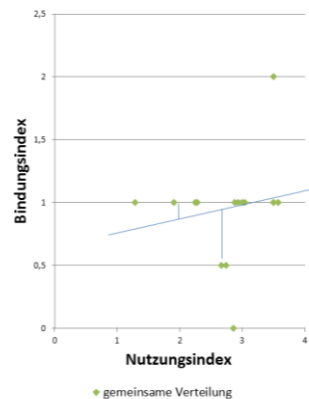


Schätzung der Regressionsgeraden

Die Fehler ε sollen minimal sein. Dies ist der Fall, wenn die Summe der quadrierten Abweichungen der Punkte zur Gerade minimal sind.



7

Schritt 1: Grundformel

$$y_i = f(x_i) = \beta x_i + \alpha + \varepsilon_i$$

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

$$\sum \varepsilon_i^2 = 0 \quad \varepsilon_i = y_i - \hat{y}_i$$

$$\sum \varepsilon_i^2 = \sum (y_i - \hat{y}_i)^2 = \text{Min}$$

$$\sum (y_i - \hat{y}_i)^2 = \sum (y_i - \beta x_i + \alpha)^2 = \text{Min}$$

$$\sum (y_i^2 - 2\alpha y_i - 2\beta y_i x_i + \alpha^2 + 2\alpha\beta x_i + \beta^2 x_i^2) = \text{Min}$$

8

Berechnung von α :

Ableiten nach α : $\frac{\partial \text{Min}}{\partial \alpha} = -\sum 2y_i + \sum 2\alpha + \sum \beta x_i$

Nach 0 umstellen: $-\sum 2y_i + \sum 2\alpha + \sum \beta x_i = 0$

$$-2 \sum y_i + 2n\alpha + 2\beta \sum x_i = 0 \quad | :2$$

$$-\sum y_i = -n\alpha - \beta \sum x_i \quad | \cdot -1$$

$$\sum y_i = n\alpha + \beta \sum x_i$$

Nach α umstellen und Ersetzen der Formeln durch die Kennwerte:

$$\alpha = \frac{1}{n} \sum y_i - \beta \frac{1}{n} \sum x_i = \bar{y} - \beta \bar{x}$$

9

Berechnung von β :

Ableiten nach β : $\frac{\partial Min}{\partial \beta} = - \sum 2x_i y_i + \sum 2\alpha x_i + \sum 2\beta x_i^2$

Nach 0 umstellen: $-2 \sum x_i y_i + 2\alpha \sum x_i + 2\beta \sum x_i^2 = 0 | : 2$

$$- \sum x_i y_i = -\alpha \sum x_i - \beta \sum x_i^2 | \cdot -1$$

$$\sum x_i y_i = \alpha \sum x_i + \beta \sum x_i^2$$

Nach β umstellen: $-\beta \sum x_i^2 = \alpha \sum x_i - \sum x_i y_i | \cdot -1$

$$\beta \sum x_i^2 = \sum x_i y_i - \alpha \sum x_i$$

Berechnung von β :

In diese Gleichung wird nun die Lösung $\alpha = \bar{y} - \beta \bar{x}$ eingesetzt:

$$\sum x_i y_i - (\bar{y} - \beta \bar{x}) \sum x_i - \beta \sum x_i^2 = 0$$

Ausmultiplizieren: $\sum x_i y_i - \bar{y} \sum x_i + \beta \bar{x} \sum x_i - \beta \sum x_i^2 = 0$

Nach β umstellen: $\beta \left(\sum x_i^2 - \bar{x} \sum x_i \right) = \sum x_i y_i - \bar{y} \sum x_i$

Dieser Term wird nun wie folgt erweitert:

auf der linken Seite wird $\sum (\bar{x}^2) - \bar{x} \sum x_i$

und auf der rechten Seite wird $\sum (\bar{x} \bar{y}) - \bar{x} \sum y_i$

hinzugefügt.

Beide Gleichungsteile ergeben insgesamt 0 und verändern somit nicht den Wert der Gleichung.

11

Erweiterter Term: $\beta \left(\sum x_i^2 + \sum (\bar{x}^2) - \bar{x} \sum x_i - \bar{x} \sum x_i \right) =$

Zusammenfassen der Summen: $\sum x_i y_i + \sum (\bar{x} \bar{y}) - \bar{x} \sum y_i - \bar{y} \sum x_i$

$$\beta \sum (x_i^2 + (\bar{x}^2) - 2\bar{x}x_i) = \sum x_i y_i + (\bar{x} \bar{y}) - \bar{x} y_i - \bar{y} x_i$$

Gemäß 2. Binomischer Formel und bzw. (a-b)(c-d) $\rightarrow$

$$\beta \sum (x_i - \bar{x})^2 = \sum (x_i - \bar{x})(y_i - \bar{y})$$

Nach β auflösen:

$$\beta = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2}$$

12

Einführung in die Statistik für
Kommunikationswissenschaftler
Deskriptive und induktive Verfahren für das
Bachelorstudium
Uhlemann, I.A.
2015, XI, 242 S., Softcover
ISBN: 978-3-658-05768-8