

# Anwendung der partiellen Differentiation in der Statistik: Regressionsgerade

Gesucht: Eine Gerade, die zu allen Punkten den kleinsten Abstand hat;  
d.h. die Differenzen  $e_i$  müssen insgesamt möglichst klein sein (Optimierungsproblem)

Ansatz: 
$$\sum_{i=1}^n (e_i)^2 \rightarrow \min.$$

Gesucht ist Gerade:  $\tilde{y} = b_0 + b_1 x$

Es gilt:

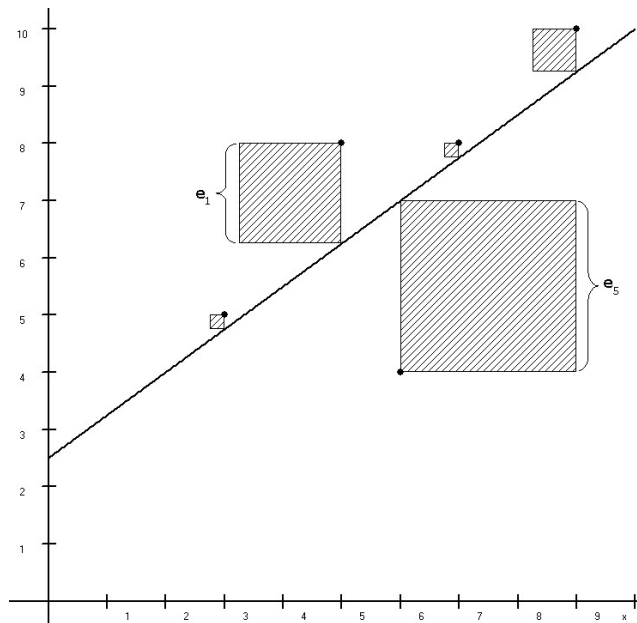
$(x_i | y_i)$  sind die Punkte des Streudiagramms

ebenso gilt:

$(x_i | \tilde{y}_i)$  sind die Punkte auf der Geraden

daraus folgt:  $\xrightarrow{y_i \leftrightarrow \tilde{y}_i} \tilde{y}_i = b_0 + b_1 x_i$

und  $\Delta y = y_i - \tilde{y}_i$  [= Seitenlänge des Quadrats]



$$\rightarrow \sum_{i=1}^n (e_i)^2 = \sum_{i=1}^n (y_i - \tilde{y}_i)^2 = \sum_{i=1}^n (y_i - b_0 - b_1 x_i)^2$$

$$E(b_1; b_0) = \sum_{i=1}^n (y_i^2 - y_i b_0 - y_i b_1 x_i - y_i b_0 + b_0^2 + b_0 b_1 x_i - y_i b_1 x_i + b_0 b_1 x_i + b_1^2 x_i^2)$$

$$E(b_1; b_0) = \sum_{i=1}^n (y_i^2 - 2y_i b_0 - 2y_i b_1 x_i + b_0^2 + 2b_0 b_1 x_i + b_1^2 x_i^2)$$

$$E(b_1; b_0) = \sum_{i=1}^n y_i^2 - 2b_0 \sum_{i=1}^n y_i - 2b_1 \sum_{i=1}^n x_i y_i + b_0^2 \sum_{i=1}^n 1 + 2b_0 b_1 \sum_{i=1}^n x_i + b_1^2 \sum_{i=1}^n x_i^2$$

$$E(b_1; b_0) = \sum_{i=1}^n y_i^2 - 2b_0 \sum_{i=1}^n y_i - 2b_1 \sum_{i=1}^n x_i y_i + n b_0^2 + 2b_0 b_1 \sum_{i=1}^n x_i + b_1^2 \sum_{i=1}^n x_i^2$$

$$\frac{\partial E(b_1; b_0)}{\partial b_1} = -2 \sum_{i=1}^n x_i y_i + 2b_0 \sum_{i=1}^n x_i + 2b_1 \sum_{i=1}^n x_i^2 = 0$$

$$\frac{\partial E(b_1; b_0)}{\partial b_0} = -2 \sum_{i=1}^n y_i + 2n b_0 + 2b_1 \sum_{i=1}^n x_i = 0$$

Lösung der Gleichungen der partiellen Ableitungen:

$$\text{I) } -2 \sum_{i=1}^n x_i y_i + 2b_0 \sum_{i=1}^n x_i + 2b_1 \sum_{i=1}^n x_i^2 = 0$$

$$\text{II) } -2 \sum_{i=1}^n y_i + 2nb_0 + 2b_1 \sum_{i=1}^n x_i = 0$$

Es gilt:

$$\sum_{i=1}^n x_i = n \cdot \bar{x} \quad \text{und} \quad \sum_{i=1}^n y_i = n \cdot \bar{y} \quad (\text{arithmetisches Mittel})$$

eingesetzt:

$$\text{I) } -2 \sum_{i=1}^n x_i y_i + 2b_0 n \cdot \bar{x} + 2b_1 \sum_{i=1}^n x_i^2 = 0$$

$$\text{II) } -2n \cdot \bar{y} + 2nb_0 + 2b_1 n \cdot \bar{x} = 0 \xrightarrow{:(2n)} -\bar{y} + b_0 + b_1 \bar{x} = 0$$

$$\rightarrow b_0 = \bar{y} - b_1 \bar{x}$$

$$\xrightarrow{\text{II) eingesetzt in I)}} \text{I) } -2 \sum_{i=1}^n x_i y_i + 2(\bar{y} - b_1 \bar{x}) n \cdot \bar{x} + 2b_1 \sum_{i=1}^n x_i^2 = 0$$

$$\xrightarrow{:2} \text{I) } -\sum_{i=1}^n x_i y_i + n \cdot \bar{x} \cdot \bar{y} - b_1 n \cdot (\bar{x})^2 + b_1 \sum_{i=1}^n x_i^2 = 0$$

$$\rightarrow -b_1 n \cdot (\bar{x})^2 + b_1 \sum_{i=1}^n x_i^2 = \sum_{i=1}^n x_i y_i - n \cdot \bar{x} \cdot \bar{y}$$

$$\rightarrow b_1 \left[ \sum_{i=1}^n x_i^2 - n \cdot (\bar{x})^2 \right] = \sum_{i=1}^n x_i y_i - n \cdot \bar{x} \cdot \bar{y}$$

$$\rightarrow b_1 = \frac{\sum_{i=1}^n x_i y_i - n \cdot \bar{x} \cdot \bar{y}}{\sum_{i=1}^n x_i^2 - n \cdot (\bar{x})^2} \stackrel{\cdot \frac{1}{n}}{=} \frac{\frac{1}{n} \cdot \sum_{i=1}^n x_i y_i - \bar{x} \cdot \bar{y}}{\frac{1}{n} \cdot \sum_{i=1}^n x_i^2 - (\bar{x})^2} = \frac{\text{Cov}(X, Y)}{V(X)}$$