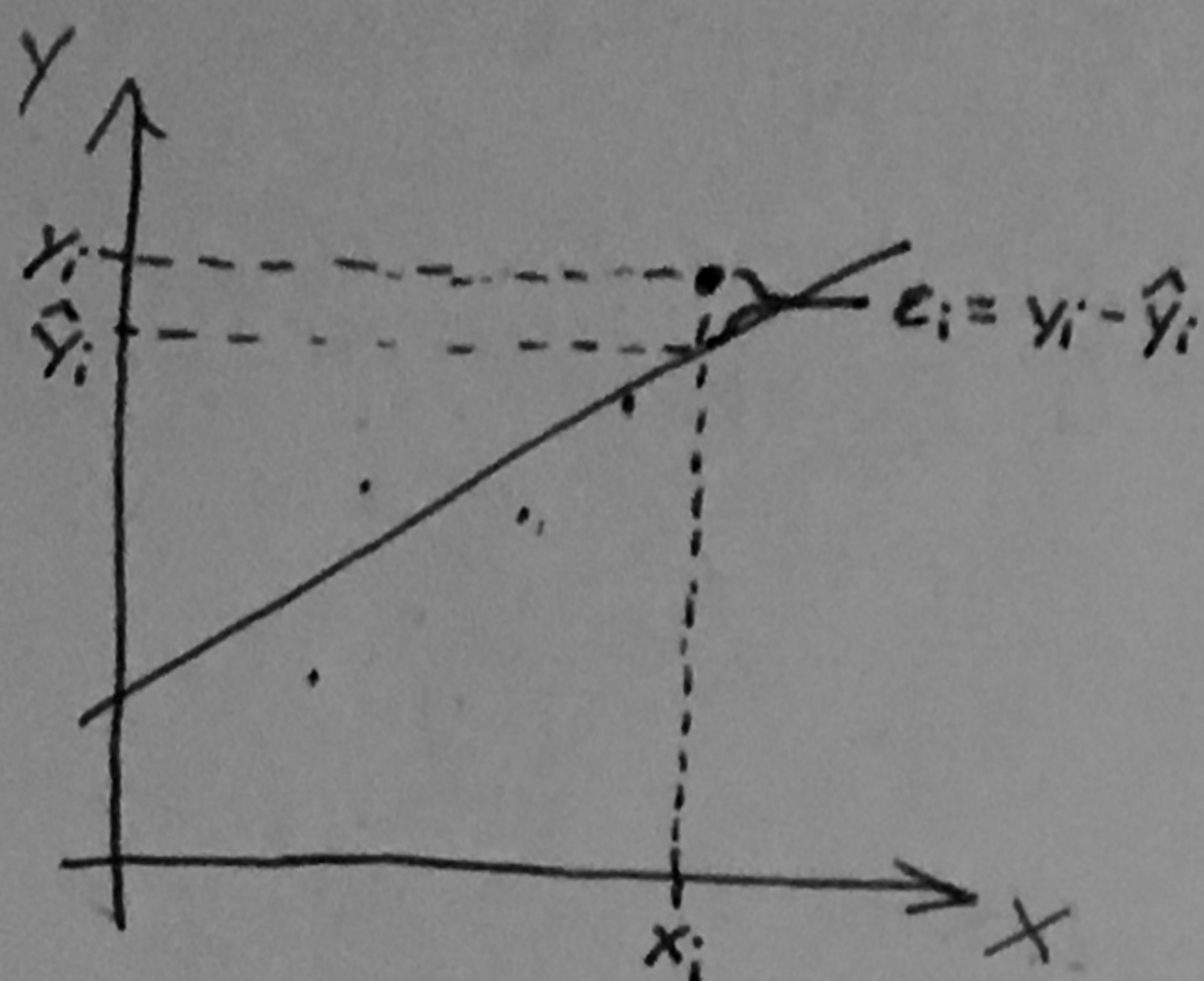


Blatt 10



Bei einer linearen Regression sucht man die lineare Funktion, die den Zusammenhang von X und Y am besten wiedergibt. Diese Funktion, die Regressionsgerade $\hat{Y} = aX + b$, besitzt die Steigung

a und den Achsenabschnitt b . Die echten Y -Werte weisen eine Varianz σ^2 um den Wert der Regressionsgerade auf. Dieser Wert muss minimiert werden, um die optimale Regressionsgerade zu finden.

Nach der Maximum-Likelihood-Methode können folgende Schätzer für die Regressionsgerade hergeleitet werden:

$$A = \frac{\text{Cov}(X, Y)}{\text{Var}(X)} \quad (\sim \mathcal{N}(a, \frac{\sigma^2}{n \text{Var}(X)}) \quad \text{für } a)$$

$$B = \bar{Y} - A\bar{X} \quad (\sim \mathcal{N}(b, \sigma^2 \cdot (\frac{1}{n} + \frac{\bar{X}^2}{n \text{Var}(X)})) \quad \text{für } b)$$

$$S^2 = \frac{\text{SSR}}{n-2} \quad \text{für } \sigma^2$$

wobei $\bar{Y} = \frac{1}{n}(Y_1 + \dots + Y_n)$, $\bar{X} = \frac{1}{n}(X_1 + \dots + X_n)$ für die bekannten Datenpunkte $(x_1, y_1), \dots, (x_n, y_n)$ und

$$\text{Cov}(X, Y) = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{X})(y_i - \bar{Y}) = E(XY) - E(X)E(Y)$$

$$\text{Var}(X) = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{X})^2 = E(X^2) - (E(X))^2$$

SSR (sum of squares of residue) ist die Varianz der Abweichungen des Werts der Regressionsgerade vom eig. Y -Wert.

$$SSR = \sum_{i=1}^n (Y_i - (AX_i + B))^2 = \sum_{i=1}^n (Y_i - AX_i - B)^2 = n \text{Var}(Y) \left(1 - \frac{\text{Cov}(X,Y)^2}{\text{Var}(X)\text{Var}(Y)}\right)$$

$(1-\alpha)$ -Konfidenzintervall

für a : $\left[A - S \cdot \sqrt{\frac{1}{n \text{Var}(X)}} \cdot t_{n-2; 1-\alpha/2} ; A + S \cdot \sqrt{\frac{1}{n \text{Var}(X)}} \cdot t_{n-2; 1-\alpha/2} \right]$

für b : $\left[B - S \cdot \sqrt{\left(\frac{1}{n} + \frac{\bar{x}^2}{n \text{Var}(X)}\right)} \cdot t_{n-2; 1-\alpha/2} ; B + S \cdot \sqrt{\left(\frac{1}{n} + \frac{\bar{x}^2}{n \text{Var}(X)}\right)} \cdot t_{n-2; 1-\alpha/2} \right]$

für σ^2 : $\left[\frac{SSR}{\chi^2_{n-2; 1-\alpha/2}} ; \frac{SSR}{\chi^2_{n-2; \alpha/2}} \right]$

$(1-\alpha)$ -Prognose-Intervall für Y an der Stelle x

$$\left[Ax + B \pm \left(S \cdot \sqrt{\left(1 + \frac{1}{n} + \frac{(x - \bar{x})^2}{\text{Var}(X) \cdot n}\right)} \cdot t_{n-2; 1-\alpha/2} \right) \right]$$