Enkel lineær regresjon

$$Y_i = \alpha + \beta x_i + \epsilon$$

- Y: Respons (stok. var)
- x_i: Forklaringsvariabel (kjent, tal)
- α og β: Regresjonsparameter (param, tal, ukjent)
- ϵ: Tilfeldig støy ('feilen', stok.var)
  - $E(\epsilon) = 0$, $Var(\epsilon) = \sigma_\epsilon^2$
  - σ_ϵ^2 (param., tal, ukjent)

Dersom $\epsilon \sim N(0, \sigma^2)$

$$Y_i \sim N(\alpha + \beta x_i, \sigma^2)$$

Regresjonsparametra

Må estimere α, β, σ_ϵ^2 frå data.

Data: $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$. Antar uavh.

Parametre: α, β, σ^2

Estimatorar: A, B, S_ϵ^2

Estimat: a, b, s_ϵ^2

Minste kvardraters metode:

Finn a og b slik at

$$SSE = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (y - (a + bx_i))^2$$

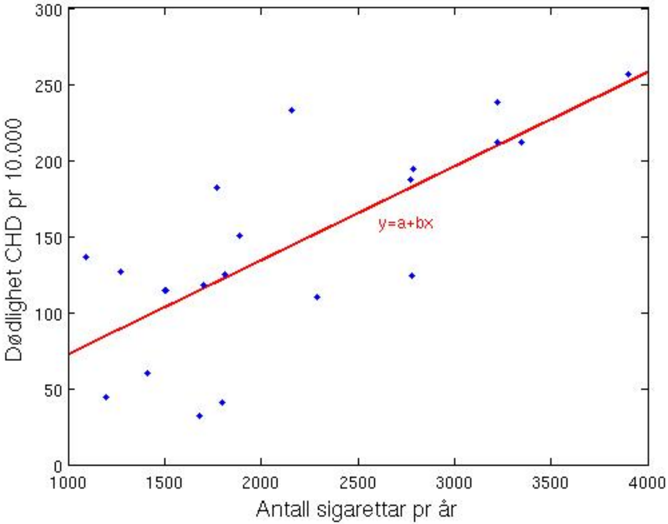
minst mogeleg.

Minste kvadraters estimatorar

- $A = \bar{Y} - B\bar{x}$
- $B = \frac{\sum_{i=1}^{n}(x_i - \bar{x})Y_i}{\sum_{i=1}^{n}(x_i - \bar{x})^2} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})Y_i}{\sum_{i=1}^{n} x_i^2 - n\bar{x}^2}$

Sigarett - hjertestans

$n = 21$, estimat: $a = 11.41$ og $b = 0.0616$



Enkel lineær regresjon

$$Y_i = \alpha + \beta x_i + \epsilon$$

Antar $\epsilon \sim N(0, \sigma^2)$

$$Y_i \sim N(\alpha + \beta x_i, \sigma^2)$$

Estimatorar

- $A \sim N(\alpha, \sigma_A^2)$
  - der $\sigma_A^2 = \frac{\sum_{i=1}^n x_i^2}{n \sum_{i=1}^n (x_i - \bar{x})^2} \sigma_\epsilon^2$
- $B \sim N(\beta, \sigma_B^2)$
  - der $\sigma_B^2 = Var(B) = \frac{\sigma_\epsilon^2}{\sum_{i=1}^n (x_i - \bar{x})^2}$
- $S^2 = \frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{n - 2}$
  - der $\hat{Y}_i = A + B x_i$