Enkel lineær regresjon

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

- Y: Respons (stok. var)
- x_i: Forlkaringsvariabel (kjent, tal)
- α og β: Regresjonsparameter (param, tal, ukjent)
- ϵ: Residual ('feilen', stok.var)
  - $E(\epsilon) = 0$, $Var(\epsilon) = \sigma^2$
  - σ^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^2 frå data.
Data: $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$.

Estimatorar

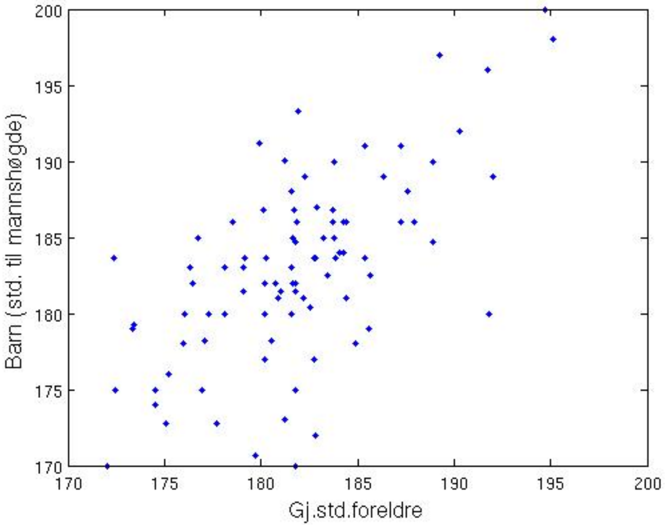
- $A = \bar{Y} - B\bar{x}$
- $B = \dfrac{\sum_{i=1}^{n}(x_i - \bar{x})Y_i}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$
- $S^2 = \dfrac{\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2}{n-2}$
  - Der $\hat{Y}_i = A + Bx_i$

Estimat

- Stok.var $\longrightarrow$ data
- a, b, s^2, y^*

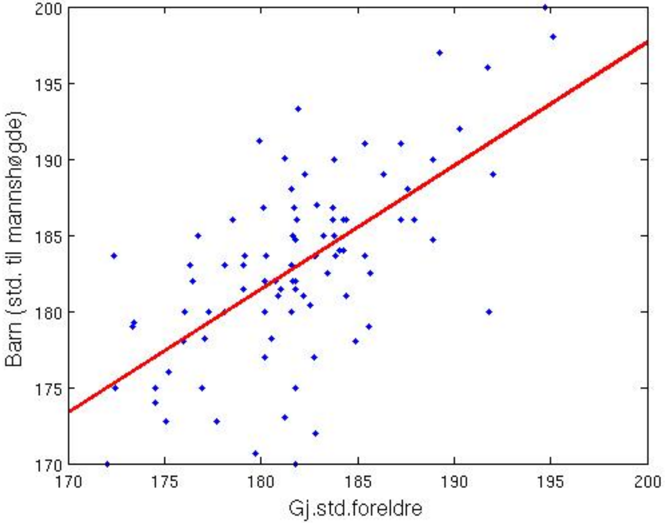
Høgde foreldre, høgde barn

$$x_i = 0.5(h_{far} + 1.08 \cdot h_{mor})$$

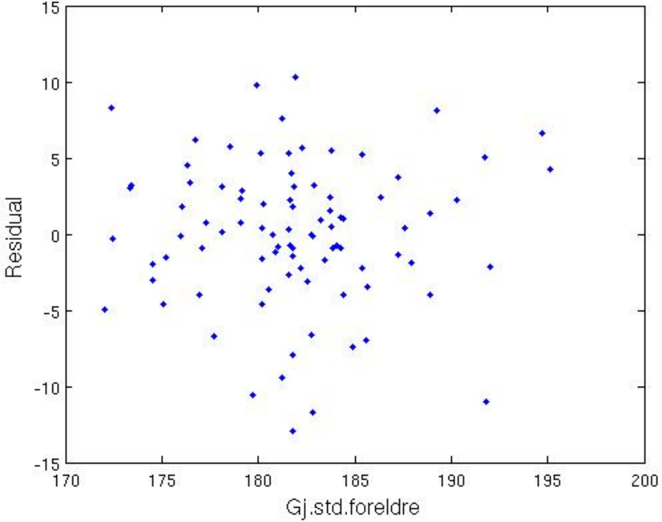


Regresjonslinje

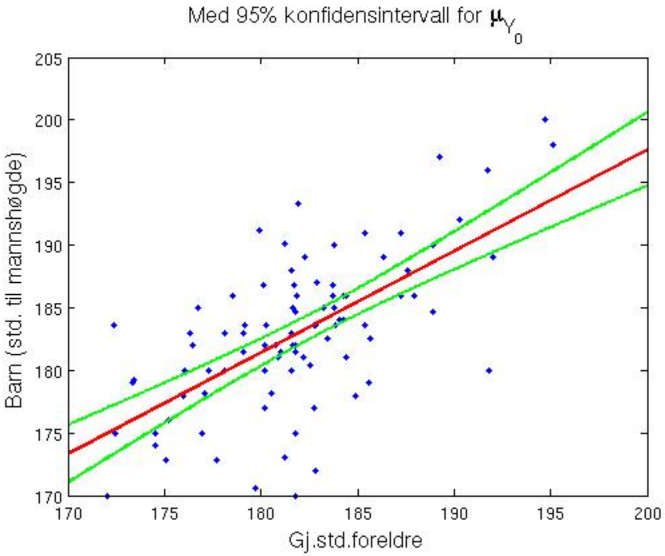
$$a = 35.4,\ b = 0.81,\ s^2 = 22.5 = 4.74^2$$



Residual


15
10
5
0
-5
-10
-15
170
175
180
185
190
195
200
Residual
Gj.std.foreldre

Konfidensintervall



Prediksjonsintervall