Multipel linear regression

tidl. $y = \beta_1 1_n + \beta_2 x + u$ (simpel lin. regr.)

nu $y = \beta_1 1_n + \beta_2 x_1 + \dots + \beta_p x_{p-1} + u$

(følger bogens indeksering
– kompliceret !)

designmatrix

$$X = [\, 1_n \;\; x_1 \;\; \dots \;\; x_{p-1} \,]$$

sæt $x_i := x_i - \bar{x} 1_n$

$$\Rightarrow 1_n^T x_i := 1_n^T (x_i - \bar{x} 1_n) = n\bar{x} - n\bar{x} = 0$$

dvs. 1_n er ortogonal på alle ændrede x_i'er.

bemærk, at $C(X) = sp\{1_n, x_1, \dots, x_{p-1}\}$ er uændret.

sæt $[x_1 \dots x_{p-1}] = \tilde{X}$

dvs. $X = [\, 1_n \;\; \tilde{X} \,]$

$$X^T X = \begin{bmatrix} 1_n^T \\ \tilde{X}^T \end{bmatrix} [\, 1_n \;\; \tilde{X} \,] = \begin{bmatrix} n & 0^T \\ 0 & \tilde{X}^T \tilde{X} \end{bmatrix}$$

$$(X^TX)^{-1} = \begin{bmatrix} \frac{1}{n} & 0^T \\ 0 & \tilde{V} \end{bmatrix}, \quad \tilde{V} = (\tilde{X}^T\tilde{X})^{-1}$$

Hypotese $H_0: \beta_2 = \dots = \beta_p = 0$

H_1: mindst et $\beta_i \neq 0$, $i = 2, \dots, p$

H_0 er ækvivalent med

$H\beta = 0$, hvor $H = [0 \; I_{p-1}]$

$$P_H = X(X^TX)^{-1}H^T K H (X^TX)^{-1}X^T$$

$$\begin{aligned} K &= (H(X^TX)^{-1}H^T)^{-1} \\ &= \left([0 \; I_{p-1}] \begin{bmatrix} \frac{1}{n} & 0^T \\ 0 & \tilde{V} \end{bmatrix} \begin{bmatrix} 0^T \\ I_{p-1} \end{bmatrix} \right)^{-1} \\ &= \left([0 \; 0 + \tilde{V}] \begin{bmatrix} 0^T \\ I_{p-1} \end{bmatrix} \right)^{-1} \\ &= (0 + \tilde{V})^{-1} \\ &= \tilde{V}^{-1} \end{aligned}$$

$$\begin{aligned} H^T K H &= \begin{bmatrix} 0^T \\ I_{p-1} \end{bmatrix} \tilde{V}^{-1} [0 \; I_{p-1}] \\ &= \begin{bmatrix} 0^T \\ \tilde{V}^{-1} \end{bmatrix} [0 \; I_{p-1}] \\ &= \begin{bmatrix} 0 & 0^T \\ 0 & \tilde{V}^{-1} \end{bmatrix} \end{aligned}$$

$$(X^TX)^{-1} H^T K H (X^TX)^{-1}$$

$$= \begin{bmatrix} \frac{1}{n} & 0^T \\ 0 & \tilde{V} \end{bmatrix} \begin{bmatrix} 0 & 0^T \\ 0 & \tilde{V}^{-1} \end{bmatrix} \begin{bmatrix} \frac{1}{n} & 0^T \\ 0 & \tilde{V} \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 0^T \\ 0 & I_{n-1} \end{bmatrix} \begin{bmatrix} \frac{1}{n} & 0^T \\ 0 & \tilde{V} \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 0^T \\ 0 & \tilde{V} \end{bmatrix}$$

$$P_H = [1_n \ \tilde{X}] \begin{bmatrix} 0 & 0^T \\ 0 & \tilde{V} \end{bmatrix} \begin{bmatrix} 1_n^T \\ \tilde{X}^T \end{bmatrix}$$

$$= [0 \ \tilde{X}\tilde{V}] \begin{bmatrix} 1_n^T \\ \tilde{X}^T \end{bmatrix}$$

$$= 0 + \tilde{X}\tilde{V}\tilde{X}^T$$

$$= \tilde{X}\tilde{V}\tilde{X}^T$$

$$\hat{\mu} - \hat{\mu}_0 = (P - P_0) y = P_H y$$

Testvariabel

$$F = \frac{\frac{1}{p-1} \|\hat{\mu} - \hat{\mu}_0\|^2}{s^2} \sim F(p-1, n-p)$$

under H_0

$$s^2 = \frac{1}{n-p} \|y - \hat{\mu}\|^2$$

Bemærk $(P-P_0)y = \tilde{X}\tilde{V}\tilde{X}^T y$

$$= \tilde{X} (\tilde{X}^T\tilde{X})^{-1} \tilde{X}^T y$$

$$= \tilde{X} \hat{\tilde{\beta}},$$

hvor $\hat{\tilde{\beta}} = (\hat{\beta}_2, \dots, \hat{\beta}_p)$

Kontrol $\hat{\beta} = (X^T X)^{-1} X^T y$

$$= \begin{bmatrix} \frac{1}{n} & 0^T \\ 0 & \tilde{V} \end{bmatrix} \begin{bmatrix} 1_n^T \\ \tilde{X}^T \end{bmatrix} y$$

$$= \begin{bmatrix} \frac{1}{n} & 0^T \\ 0 & \tilde{V} \end{bmatrix} \begin{bmatrix} n\bar{y} \\ \tilde{X}^T y \end{bmatrix} = \begin{bmatrix} \bar{y} \\ \tilde{V} \tilde{X}^T y \end{bmatrix}$$

$$= \begin{bmatrix} \bar{y} \\ (\tilde{X}^T \tilde{X})^{-1} \tilde{X}^T y \end{bmatrix},$$

dvs. $\hat{\beta}_1 = \bar{y}$ og $\hat{\tilde{\beta}} = (\tilde{X}^T \tilde{X})^{-1} \tilde{X}^T y$

$p = 2$, dvs. simpel lin. regr.:

$\tilde{X} = x := x - \bar{x}$ (tilbagetransformation)

$\| \tilde{X} \hat{\beta}_2 \|^2 = \hat{\beta}_2^2 S_{xx}$

$$F = \frac{\hat{\beta}_2^2 S_{xx}}{S^2} \sim F(1, n-2) \text{ under } H_0$$

$$s^2 = \frac{1}{n-2} (S_{yy} - \hat{\beta}_2^2 S_{xx}) \quad (1x)$$

kan ækvivaleres med

$$T = \frac{\hat{\beta}_2 \sqrt{S_{xx}}}{S} \sim t(n-2) \text{ under } H_0$$

konfidensinterval for β_2:

$$\beta_2 = \hat{\beta}_2 \pm t_{1-\frac{\alpha}{2}}(n-2) \frac{s}{\sqrt{S_{xx}}}$$

Determinationskoefficienten

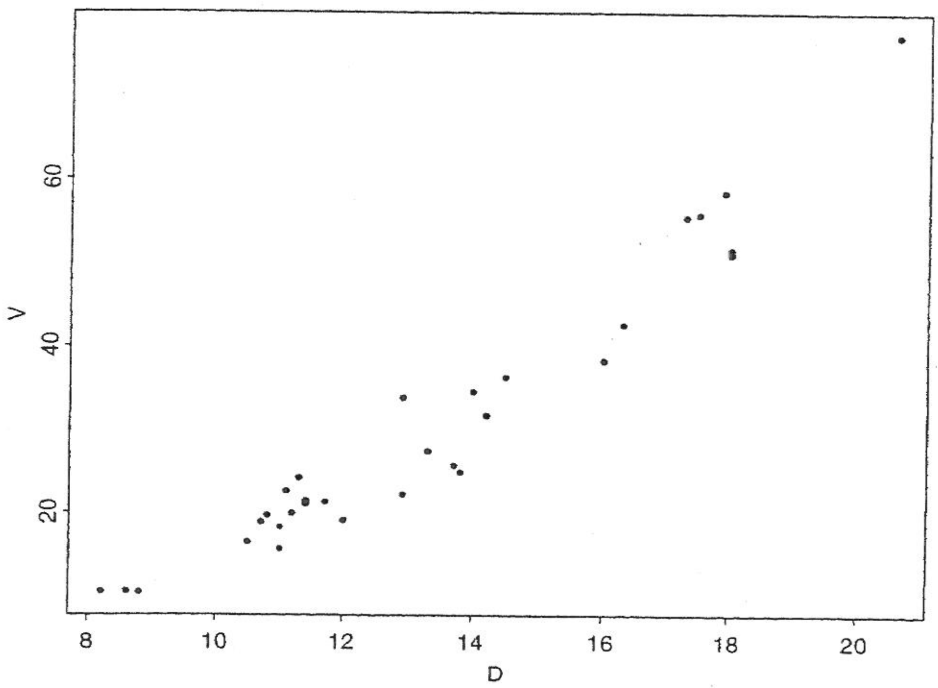
$$R^2 = \frac{S_{yy} - \|y - \hat{\mu}\|^2}{S_{yy}} \qquad \left(\|y - \hat{\mu}\|^2 = SS_{res} \right)$$

kan tolkes som den del af den totale variation i de observerede værdier, som regressionsmodellen kan forklare.

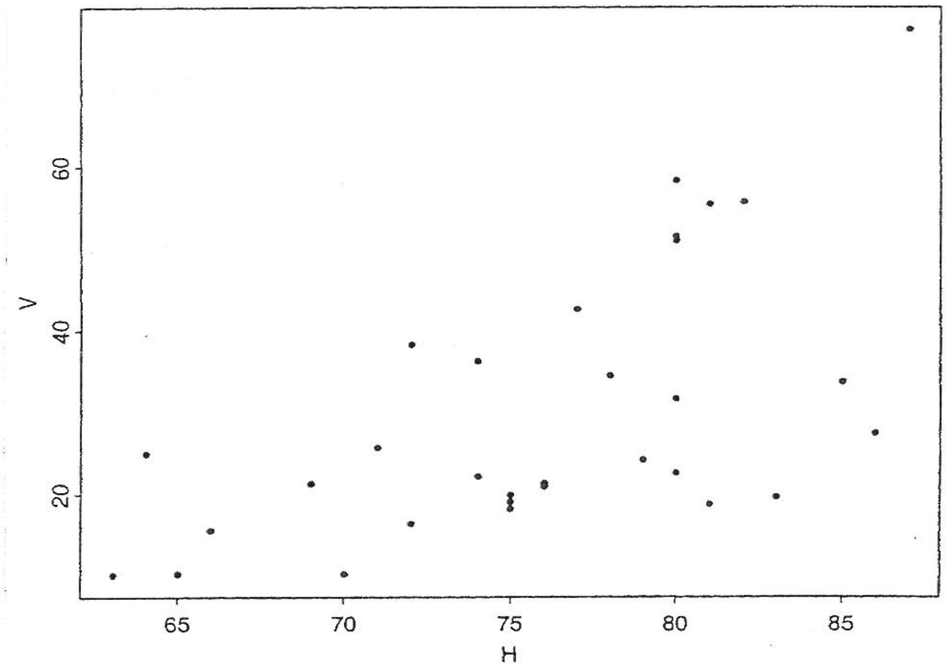
eks. Datamateriale vedr. træer, bog s. 194

D : diameter
H : højde
V : Volumen

Kan der findes en 'formel' $V = f(D, H)$?



Cherry tree data: scatter plots of (D, V)



Cherry tree data: scatter plots of (H, V)

Forsøg med $V = \gamma_0 D^{\gamma_1} H^{\gamma_2}$ *

$$\gamma_0 = \begin{cases} \frac{\pi}{4} \sim \text{cylinderform} \\ \frac{\pi}{12} \sim \text{kegleform} \end{cases}$$

transformation:

$$\ln V = \ln \gamma_0 + \gamma_1 \ln D + \gamma_2 \ln H$$

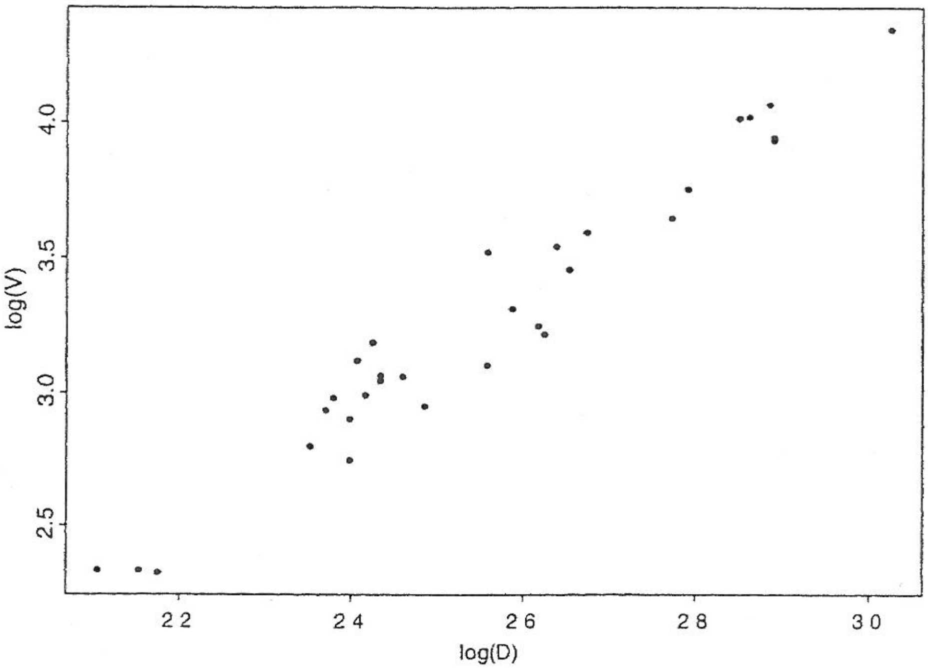
eller

$$y = \alpha + \gamma_1 x_1 + \gamma_2 x_2$$

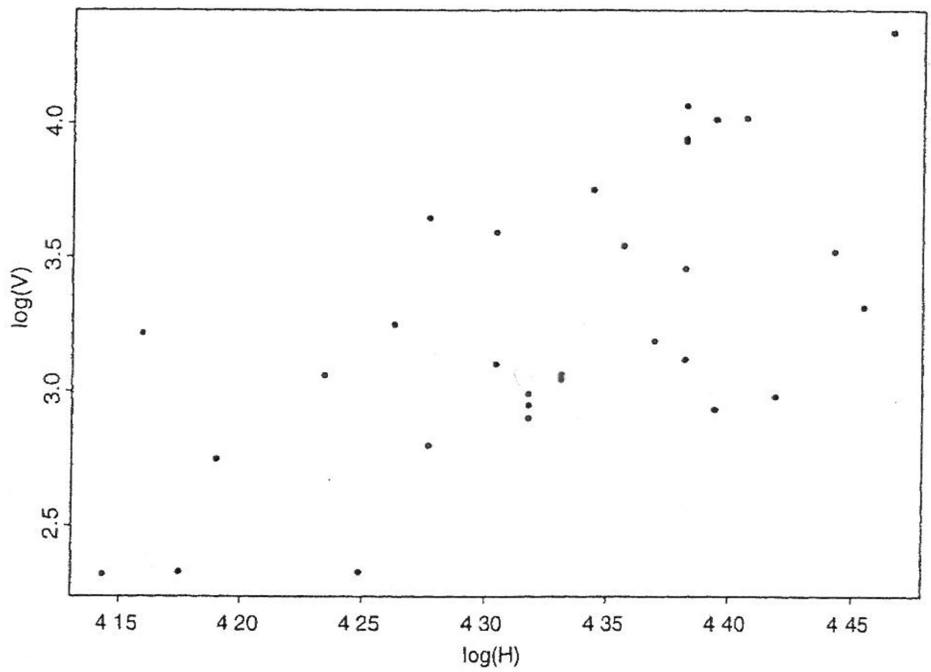
Nye plot, se næste side

I god overensstemmelse med en lineær regressionsmodel!

* Formodning: $\gamma_1 \simeq 2$, $\gamma_2 \simeq 1$



Cherry tree data: scatter plots of $(\log D, \log V)$



Cherry tree data: scatter plots of $(\log H, \log V)$

Beregninger

$$(X^\top X)^{-1} = \begin{pmatrix} 96.613 & 3.1340 & -24.171 \\ 3.1340 & 0.8437 & -1.223 \\ -24.171 & -1.2228 & 6.308 \end{pmatrix}$$

$$X^\top y = \begin{pmatrix} 101.5 \\ 263.0 \\ 439.9 \end{pmatrix}$$

Stat. 21.4.2010

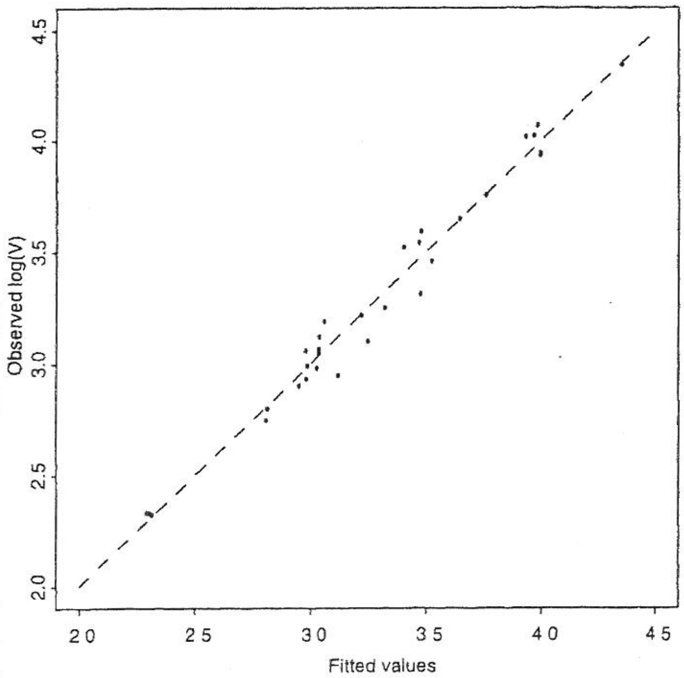
$$s^2 = \frac{\sum_i (y_i - \hat{y}_i)^2}{28} = (0.08172)^2$$

$\hat{\alpha} = -6.620$ (standard error 0.803)
$\hat{\gamma}_1 = 1.976$ (standard error 0.075)
$\hat{\gamma}_2 = 1.119$ (standard error 0.205)

$H_0: \gamma_2 = 0 \qquad H_1: \gamma_2 \neq 0$

$$t_{obs} = \frac{1{,}119}{0{,}205} \approx 5{,}45 \qquad (\text{værdi i } t(\overset{31-3}{28}))$$

p-værdi $< 1‰$ $\qquad H_0$ forkastes



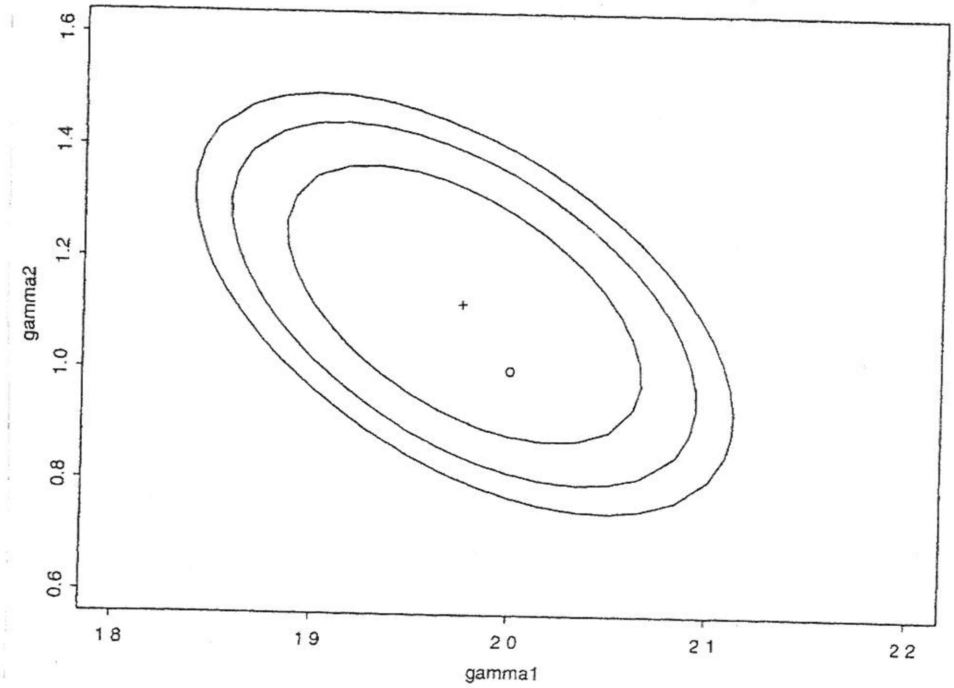
Cherry tree data: observed values of $\log V$ *versus fitted values*

Konfidensellipse for (γ_1, γ_2)

$$H = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$K = (H(X^TX)^{-1}H^T)^{-1} = \begin{pmatrix} 1.648 & 0.320 \\ 0.320 & 0.221 \end{pmatrix}$$

$$\left\{ (\gamma_1,\gamma_2) \;\middle|\; \frac{[\hat\gamma_1-\gamma_1 \quad \hat\gamma_2-\gamma_2]\, K \begin{bmatrix} \hat\gamma_1-\gamma_1 \\ \hat\gamma_2-\gamma_2 \end{bmatrix}}{2 s^2} \le f_{1-\alpha}(2,\underset{\substack{\uparrow \\ 31-3}}{28}) \right\}$$



Cherry tree data: confidence regions for (γ_1, γ_2), at confidence levels 75%, 90% and 95%; the + sign marks the MLE point, ∘ shows the (2,1) point

Bemærk (2,1) ligger 'pænt' placeret, i overensstemmelse med formodningen. □

Polynomial regression

$$y = \beta_1 + \beta_2 x + \beta_3 x^2 + \ldots + \beta_{m+1} x^m + u$$

Designmatrix
$$X = \begin{bmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^m \\ 1 & x_2 & x_2^2 & \cdots & x_2^m \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^m \end{bmatrix}$$

Ortogonalt design

X kan erstattes af

$$Z = [1_n \; z_1 \; z_2 \; \ldots \; z_m] \, ,$$

hvor z_i'erne er søjlevektorer, der opfylder $z_j^T z_k = 0$ for $j \neq k$, altså ortogonale.

$y = X\beta + u$ ændres til $y = Z\gamma + u$

Bemærk, at $(Z^T Z)^{-1}$ er diagonalmatrix dvs

- $\hat{\gamma}_j = \dfrac{z_j^T y}{\| z_j \|^2}$, $j = 2, \ldots, m+1$, $\hat{\gamma}_1 = \bar{y}$
- $\hat{\gamma}_i$'erne er uafhængige
- SS_{tot} kan dekomponeres i

komponent	kvadrat-sum	frihedsgrader
1	$n\bar{y}^2$	1
2	$\dfrac{(z_1^T y)^2}{\|z_1\|^2}$	1
$\vdots$	$\vdots$	$\vdots$
$m+1$	$\dfrac{(z_m^T y)^2}{\|z_m\|^2}$	1
residual	$S_{yy} - \sum_{j=2}^{m} \dfrac{(z_i^T y)^2}{\|z_i\|^2}$	$n-1-m$
	S_{yy}	$n-1$
total	SS_{tot}	n

(frihedsgraderne for komponent 2 til $m+1$: i alt m)

$$S_{yy} = SS_{tot} - n\bar{y}^2 = \sum_i y_i^2 - n\bar{y}^2 = \sum_i (y_i - \bar{y})^2$$

Stat. 21.4.2010

To observationsrækker – igen!

$z_1, z_2, \dots, z_n \qquad Z_i \sim N(\beta_1, \sigma^2)$

$x_1, x_2, \dots, x_m \qquad X_j \sim N(\beta_2, \sigma^2)$

sæt $y = \begin{bmatrix} z_1 \\ \vdots \\ z_n \\ x_1 \\ \vdots \\ x_m \end{bmatrix}, \quad X = \begin{bmatrix} 1 & 0 \\ \vdots & \vdots \\ 1 & 0 \\ 0 & 1 \\ \vdots & \vdots \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1_n & 0 \\ 0 & 1_m \end{bmatrix},$

$\beta = \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}$

$$X^T X = \begin{bmatrix} 1_n^T & 0^T \\ 0^T & 1_m^T \end{bmatrix} \begin{bmatrix} 1_n & 0 \\ 0 & 1_m \end{bmatrix} = \begin{bmatrix} n & 0 \\ 0 & m \end{bmatrix}$$

$$X^T y = \begin{bmatrix} 1_n^T & 0^T \\ 0^T & 1_m^T \end{bmatrix} y = \begin{bmatrix} n\bar{z} \\ m\bar{x} \end{bmatrix}$$

$$\hat{\beta} = \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix} = \begin{bmatrix} \bar{z} \\ \bar{x} \end{bmatrix} \qquad \left(= \begin{bmatrix} \hat{\mu} \\ \hat{\eta} \end{bmatrix} \text{ jf. tidl. eks.} \right)$$

$$\hat{y} = X\hat{\beta} = \begin{bmatrix} \bar{z} \\ \vdots \\ \bar{z} \\ \bar{x} \\ \vdots \\ \bar{x} \end{bmatrix}, \qquad y - \hat{y} = \begin{bmatrix} z_1 - \bar{z} \\ \vdots \\ z_n - \bar{z} \\ x_1 - \bar{x} \\ \vdots \\ x_m - \bar{x} \end{bmatrix}$$

$$\| y - \hat{y} \|^2 = \sum_i (z_i - \bar{z})^2 + \sum_j (x_j - \bar{x})^2$$

$$\hat{\sigma}^2 = \frac{1}{m+n} \| y - \hat{y} \|^2 \qquad \text{(som tidl.)}$$

$$s^2 = \frac{1}{m+n-2} \| y - \hat{y} \|^2$$

Hypotese $H_0 : H\beta = 0 \quad , \quad H = [1 \;\; -1]$

$H_1 : H\beta \neq 0$

H_0 er ekvivalent med $\beta_1 - \beta_2 = 0$

eller $\beta_1 = \beta_2$

$$\hat{\beta}_0 = \hat{\beta} - (X^T X)^{-1} H^T K H \hat{\beta}$$

$$K = (H (X^T X)^{-1} H^T)^{-1}$$

$$= \left([1 \;\; -1] \begin{bmatrix} \frac{1}{n} & 0 \\ 0 & \frac{1}{m} \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right)^{-1}$$

$$= \left(\left[\tfrac{1}{n} \;\; -\tfrac{1}{m} \right] \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right)^{-1}$$

$$= \left(\frac{1}{n} + \frac{1}{m} \right)^{-1}$$

$$\hat{\beta}_0 = \begin{bmatrix} \bar{z} \\ \bar{x} \end{bmatrix} - \begin{bmatrix} \frac{1}{n} & 0 \\ 0 & \frac{1}{m} \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \left(\frac{1}{n} + \frac{1}{m} \right)^{-1} [1 \;\; -1] \begin{bmatrix} \bar{z} \\ \bar{x} \end{bmatrix}$$

$$= \begin{bmatrix} \bar{z} - \frac{1}{n} \dfrac{\bar{z} - \bar{x}}{\frac{1}{n} + \frac{1}{m}} \\[2ex] \bar{x} - \frac{1}{m} \dfrac{\bar{z} - \bar{x}}{\frac{1}{n} + \frac{1}{m}} \end{bmatrix} = \begin{bmatrix} \dfrac{n\bar{z} + m\bar{x}}{m+n} \\[2ex] \dfrac{n\bar{z} + m\bar{x}}{n+m} \end{bmatrix}$$

$$= \frac{n\bar{z} + m\bar{x}}{m+n} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

(alternativt: $\hat{\beta}_{10} = \hat{\beta}_{20} = \bar{y} = \dfrac{n\bar{z} + m\bar{x}}{n+m}$)

$$F = \frac{(H\hat{\beta})^T K H\hat{\beta}}{S^2}$$

$$= \frac{(\bar{z}-\bar{x})\left(\frac{1}{n}+\frac{1}{m}\right)^{-1}(\bar{z}-\bar{x})}{S^2}$$

$$= \frac{(\bar{z}-\bar{x})^2}{\left(\frac{1}{n}+\frac{1}{m}\right)S^2} \sim F(1, n+m-2)$$

under H_0

I stedet kan benyttes

$$T = \frac{\bar{z}-\bar{x}}{S\sqrt{\frac{1}{n}+\frac{1}{m}}} \sim t(n+m-2)$$

under H_0