

Eksamen maj 2010, Opgave 4

(a)

Beregning af mellemregningsstørrelser:

$$\bar{x} = 3298/47 = 70.170$$

$$\bar{y} = 938/47 = 19.957$$

$$\text{SSD}_x = 238560 - 3298^2/47 = 7138.638$$

$$\text{SSD}_y = 19126 - 938^2/47 = 405.915$$

$$\text{SPD}_{xy} = 66505 - 3298 \times 938/47 = 685.340$$

$$\text{SSD}_{\text{res}} = 405.915 - 685.340^2/7138.638 = 340.119$$

Vi får så følgende estimater for parametrene:

$$\hat{\beta} = 685.340/7138.638 = 0.09600$$

$$\hat{\alpha} = 19.957 - 70.170 \times 0.09600 = 13.221$$

$$\hat{\sigma}^2 = 340.119/(47 - 2) = 7.5582$$

Til udregning af 95% sikkerhedsgrænser for β får vi brug for, at 97.5%-fraktilen i T-fordelingen med 45 frihedsgrader er 2.014. Så fås

$$\begin{aligned}\beta &= 0.09600 \pm 2.014 \sqrt{\frac{7.5582}{7138.638}} = 0.09600 \pm 2.014 \times 0.03254 \\ &= 0.09600 \pm 0.06554\end{aligned}$$

svarende til 95% sikkerhedsintervallet [0.03047, 0.16154]

(b) Estimatet for den forventede værdi af y_i for et amt med $x_i = 80$ er jo

$$13.221 + 0.09600 \times 80 = 20.901.$$

95% sikkerhedsgrænser for denne størrelse fås således:

$$20.901 \pm 2.014 \sqrt{\left(\frac{1}{47} + \frac{(80 - 70.170)^2}{7138.638}\right) \times 7.5582} = 20.901 \pm 1.033$$

svarende til sikkerhedsintervallet [19.87, 21.93].

(c) Listen af parameterestimater viser at ingen af de andre variable er signifikante, dvs. de kan hver for sig fjernes fra modellen. Variansanalysekemaet viser, at de faktisk alle sammen kan fjernes ved succesiv testning. Samtidig får vi bekræftet (som det også fremgår indirekte af at sikkerhedsintervallet for hældningen ikke indeholder 0) at variablen FER ikke kan fjernes fra den simple regressionsmodel ($P=0.0050$).