

Opgave 4, eksamen juni 2008.

(a)

Udregninger:

$$\bar{x} = 8198/53 = 154.68.$$

$$\bar{y} = 124205/53 = 2343.49.$$

$$SSD_x = 1429832 - 8198^2/53 = 161771.55.$$

$$SSD_y = 318811200 - 124205^2/53 = 27737954.2.$$

$$SPD_{xy} = 20892204 - 8198 \times 124205/53 = 1680268.34.$$

$$SSD_{\text{res}} = 27737954.2 - 1680268.34^2/161771.55 = 10285554.6.$$

Herefter fås estimaterne

$$\hat{\beta} = \frac{1680268.34}{161771.55} = \mathbf{10.3867}$$

$$\hat{\alpha} = 2343.49 - 154.68 \times 10.3867 = \mathbf{736.888}$$

$$\hat{\sigma}^2 = \frac{10285554.6}{53 - 2} = \mathbf{201677.5}$$

95% sikkerhedsgrenser for β : Ved opslag ses at 97.5%-fraktilen for T-fordelingen med 51 frihedsgrader er ca. 2.008. Vi får så

$$\begin{aligned} \beta &= \hat{\beta} \pm 2.008 \sqrt{\text{var}(\hat{\beta})} \approx 10.3867 \pm 2.008 \sqrt{\frac{201677.5}{161771.55}} \\ &= 10.3867 \pm 2.008 \times 1.1165 = 10.3867 \pm 2.2416 \end{aligned}$$

svarende til sikkerhedsintervallet [**8.1451,12.6282**].

(b)

Den forventede pris for et hus på 120 m² estimeres (i enhed 1000 kr.) ved

$$736.888 + 10.3867 \times 120 = \mathbf{1983.29}.$$

95% sikkerhedsgrenserne for dette estimat er givet ved

$$1983.29 \pm 2.008 \sqrt{\left(\frac{1}{53} + \frac{(154.68 - 120)^2}{161771.55} \right) 201677.5} = 1983.29 \pm 146.2$$

svarende til sikkerhedsintervallet [**1837.1,2129.5**].

(c)

At de to sidste tests i variansanalysekemaet fører til godkendelse (med P-værdier på 0.21 og 0.33) betyder netop, at vi kan reducere den nye model til den gamle ved fjernelse af de to nye led SOVEV og GRAREAL fra modelformlen. Det vil sige, at når man kender boligarealet, får man ingen yderligere information om prisen ved også at kende antallet af soveværelser og grundens areal. Især når det gælder grundens areal kan dette forekomme lidt besynderligt — men der er jo også kun tale om lille datasæt.