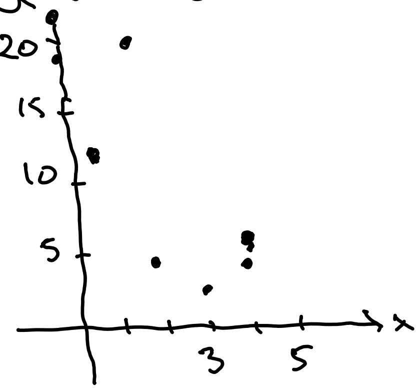
Løsning: Oppgavesett 14

1. Velger x = antall kaffekopper

y = antall abonnement

x	y
2	4
1	12
4	5
3	2
2	20
0	21
4	4

a) Spredningsdiagram



Korrelasjonskoeff:

$$r = \frac{s_{xy}}{s_x \cdot s_y} \underset{\text{kalk}}{\approx} -0.743$$

Regresjonslinje:

$$\hat{\beta} = r \cdot \frac{s_y}{s_x} \underset{\text{kalk}}{\approx} -3.98$$

$$\hat{\alpha} = \bar{y} - \hat{\beta}\bar{x} \underset{\text{kalk}}{\approx} 18.8$$

$$\Downarrow$$

$$\boxed{y = 18.8 - 3.98x}$$

b) $x = 2$:

$$\hat{y} = 18.8 - 3.98 \cdot 2 \approx 10.9$$

c) $$SSE = SS_T \cdot (1 - r^2) = (n-1)\, s_y^2 \cdot (1 - r^2)$$

$$\approx 6 \cdot 64.24 \cdot 0{,}44 \approx 172.85$$

Den kvadratiske feilfunksjonen er

$$SS_E = \varepsilon_1^2 + \varepsilon_2^2 + \dots + \varepsilon_n^2$$

hvor $\varepsilon_i = y_i - \hat{y}_i = y_i - (\hat{\alpha} + \hat{\beta} x_i)$.

Vi har at $SS_T = SS_R + SS_E$ og at

$\frac{SS_R}{SS_T} = r^2$, slik at $SS_E = SS_T \cdot (1 - r^2)$,

hvor $SS_T = \sum_{i=1}^{n} (y_i - \bar{y})^2 = (n-1) s_y^2$.

Dette gir:

σ^2 estimeres ved $S^2 = \frac{1}{n-2} \cdot SS_E$

$$\simeq \frac{1}{5} \cdot 172.85 \simeq \underline{34.6}$$

$$Var(\hat{\beta}) = \frac{\sigma^2}{\sum x_i^2 - n\bar{x}^2}$$

$$= \frac{\sigma^2}{(n-1) s_x^2} \simeq \frac{34.6}{6 \cdot 2.24} \simeq 2.58$$

$$SE(\hat{\beta}) \simeq \sqrt{2.58} \simeq \underline{\underline{1.60}}$$

d) <u>95% konfidensintervall for β:</u> $\alpha = 0.05$

$$\hat{\beta} \pm t_{\alpha/2}^{n-2} \cdot SE(\hat{\beta}) = -3.98 \pm 2.57 \cdot 1.60$$

$$\hat{=} -3.98 \pm 4.11$$

$$\underline{\underline{[-8.1, 0.1]}}$$

e) $H_0: \beta = 0$ ingen sammenheng

$H_1: \beta \neq 0$ sammenheng

Forkastingsområde: $\underline{|T| > t_{\alpha/2}^{n-2} = 2.57}$

$$T = \frac{\hat{\beta}}{SE(\hat{\beta})} \approx -2.49$$ ← ikke i forkastn.-området

($T > 2.57$ eller $T < -2.57$)

Vi <u>beholder H_0</u>, ingen sammenheng.

f) Vi antar at salgene er valgt tilfeldig, og at

$$Y = \alpha + \beta X + \varepsilon$$ med ε normalfordelt $N(0, \sigma^2)$

for en gitt verdi av x.

<u>2.</u> x = antall år etter 1965 (dvs 1965 svarer til $x = 0$)

y = andelen skilsmisser (per 1000)

x	y
0	2.9
5	3.4
10	4.9
15	6.5
20	7.9
25	9.5
30	11.7
35	10.9
40	12.6
45	11.6

a) <u>Regresjonslinje:</u>

$$\hat{\beta} = r \cdot \frac{s_y}{s_x} \underset{\text{kalk}}{\approx} \underline{0.23}$$

$$\hat{\alpha} = \bar{y} - \hat{\beta}\bar{x} \underset{\text{kalk}}{\approx} \underline{3.01}$$

$$\underline{\underline{Y = 3.01 + 0.23X}}$$

(hvis x er årstall, blir $\hat{\alpha} = -449$)
$\hat{\beta} = 0.23$

$H_0: \beta \leq 0$

$H_1: \beta > 0$ (økt skillsmisserate)

<u>Bruker $\alpha = 5\%$:</u>

$$T = \frac{\hat{\beta}}{SE(\hat{\beta})}$$

<u>Forkastingsområde:</u>

$$T > t_{\alpha}^{n-2} = 1.860$$

$$SE(\hat{\beta}) \simeq \sqrt{\frac{s^2}{(n-1)s_x^2}} \simeq \sqrt{\frac{\frac{1}{n-2}SSE}{(n-1)s_x^2}}$$

$$\simeq \sqrt{\frac{1/8 \cdot 7.870}{7 \cdot 229.2}} \simeq 0.025$$

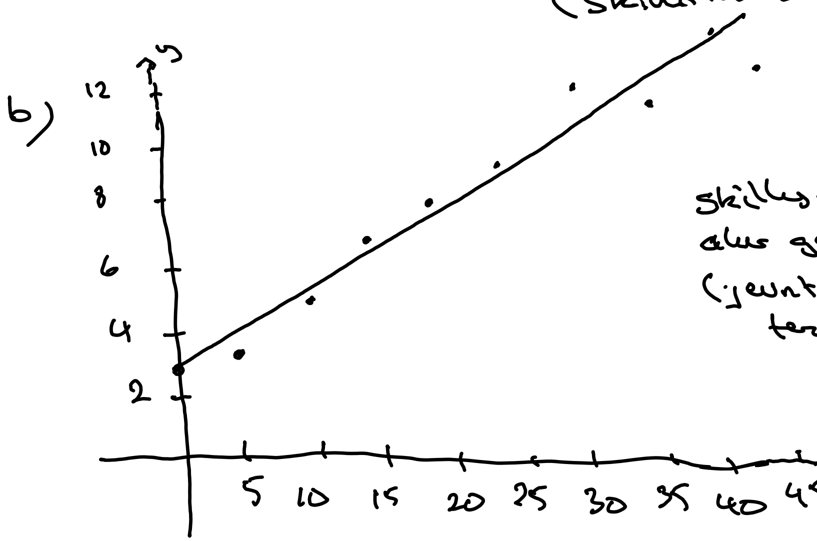
$$T \simeq \frac{0.23}{0.025} \simeq \underline{9.3}$$

T er i forkastingsomr.

<u>Forkaster H0</u>

(skillsmisseraten økt

b)



skillsmisseraten øker ganske jevnt (jevnt stigende tendens)

$$2025: \; x = 60 \Rightarrow \hat{y} = 3.01 + 60 \cdot 0.23 \simeq \underline{\underline{16.8}}$$

$\underline{3.}$ Data fra Opgs. 2

92% konfidensintervall for β : $\alpha = 0.08$

$$\hat{\beta} \pm t_{\alpha/2}^{n-2} \cdot SE(\hat{\beta}) = 0.30 \pm 2.004 \cdot 0.0176$$

$$= 0.30 \pm 0.035$$

$$t_{\alpha/2}^{n-2} = t_{0.04}^{8} \text{ fra kalk}$$

$$\underline{\underline{[0.26,\ 0.33]}}$$

Bruker data fra 1965–1995 i Opgs. 2

$\hat{\beta} \simeq \underline{0.297}$ (fra kalk.)

$$SE(\hat{\beta})^2 = \frac{\frac{1}{n-2} \cdot SSE}{(n-1) s_x^2} = \frac{\frac{1}{n-2} \cdot \cancel{(n-1)} s_y^2 \cdot (1-r^2)}{\cancel{(n-1)} s_x^2}$$

$$= \frac{1}{n-2} \frac{s_y^2}{s_x^2} (1-r^2) \underset{\text{kalk}}{\simeq} 0.00031$$

$$\Rightarrow SE(\hat{\beta}) \simeq \sqrt{0.00031} \simeq \underline{0.0176}$$

$\underline{4.}$ Se notater fra Forelesning 14

$\underline{5.}$ Se løsn. av eksamensoppgave.