Spørsmål 1

Anta at vi er blitt tildelt data (x_i, y_i), $i = 1, \ldots, 30$. Vi antar at disse dataene følger en lineær model

$Y_i = \beta_0 + \beta_1 x_i + \varepsilon_i,$

der $\varepsilon_i, i = 1, \ldots, 30$ er uavhengige stokastiske variabler med forventningsverdi 0 og varians σ^2. Fra dataene har vi allerede regnet ut følgende

$\bar{y} = \frac{1}{30} \sum_{i=1}^{30} y_i = 63.1,$

$\bar{x} = \frac{1}{30} \sum_{i=1}^{30} x_i = 43.09,$

$\sum_{i=1}^{30} (x_i - \bar{x})^2 = 14.04,$

$\sum_{i=1}^{30} (x_i - \bar{x})(y_i - \bar{y}) = 18.89.$

Til dataene ønsker vi å tilpasse en regresjonsline ved hjelp av minste kvadraters metode. Benytt minste kvadraters metode til å finne estimater b_0 og b_1 for β_0 og β_1, basert på de gitte opplysningene om dataene. Oppgi svarene som desimaltall med tre desimaler, for eksempel 2.543 og 23.232.

$$SSE = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 = \sum_{i=1}^{n} (y_i - b_0 - b_1 x_i)^2$$

$$\frac{\partial SSE}{\partial b_0} = -2 \sum_{i=1}^{n} (y_i - b_0 - b_1 x_i) = 0 \qquad / \div -2$$

$$\Rightarrow n \cdot \bar{y} - n b_0 - \sum_{i=1}^{n} b_1 x_i = 0 \qquad \text{I}$$

$$\frac{\partial SSE}{\partial b_1} = -2 \sum_{i=1}^{n} (y_i - b_0 - b_1 x_i) x_i = 0 \qquad / \div -2$$

$$\sum_{i=1}^{n} y_i x_i = \sum b_0 x_i + \sum b_1 x_i^2$$

$$\sum_{i=1}^{n} x_i y_i = b_0 \sum x_i + b_1 \sum x_i^2 \qquad \text{II}$$

I

$$n \cdot \bar{y} - nb_0 - \sum_{i=1}^{n} b_1 x_i = 0$$

$$b_0 = \frac{n\bar{y} - b_1 \sum x_i}{n}$$

$$\boxed{\hat{b}_0 = \bar{y} - \hat{b}_1 \bar{x}}$$

II

$$\sum x_i y_i = b_0 \sum x_i + b_1 \sum x_i^2$$

$$b_1 \sum x_i^2 = \sum x_i y_i - b_0 \underbrace{\sum x_i}_{= n \cdot \bar{x}}$$

$$\begin{aligned} \sum (x_i - \bar{x})^2 &= \sum x_i^2 - 2 \cdot \sum x_i \bar{x} + \sum \bar{x}^2 \\ &= \sum x_i^2 - 2\bar{x} \sum x_i + \sum \bar{x}^2 \\ &= \sum x_i^2 - 2n\bar{x}^2 + n\bar{x}^2 \\ &= \sum x_i^2 - n\bar{x}^2 \end{aligned}$$

$$\sum (x_i - \bar{x})(y_i - \bar{y}) = \sum x_i y_i - \bar{x} \sum y_i - \bar{y} \sum x_i + \sum \bar{x}\bar{y}$$

$$= \sum x_i y_i - \cancel{n\bar{x}\bar{y}} - n\bar{x}\bar{y} + \cancel{n\bar{x}\bar{y}}$$

$$= \sum x_i y_i - n\bar{x}\bar{y}$$

$$b_1 \left(\sum x_i^2 - n\bar{x}^2 + n\bar{x}^2 \right) = \sum x_i y_i - n\bar{x}\bar{y} + n\bar{x}\bar{y} - nb_0\bar{x}$$

$$b_1 \left(\sum (x_i - \bar{x})^2 + n\bar{x}^2 \right) = \sum (x_i - \bar{x})(y_i - \bar{y}) + n\bar{x}\bar{y} - nb_0\bar{x} \qquad / b_0 = \bar{y} - b_1\bar{x}$$

$$b_1 \left(\sum (x_i - \bar{x})^2 + \cancel{n\bar{x}^2} \right) = \sum (x_i - \bar{x})(y_i - \bar{y}) + \cancel{n\bar{x}\bar{y}} - \cancel{n\bar{x}\bar{y}} + \cancel{b_1 n\bar{x}^2}$$

$$\Rightarrow \boxed{\hat{b}_1 = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2}}$$

Får at

$$\hat{b}_1 = \frac{18{,}89}{14{,}04} = 1{,}3454$$

$$\hat{b}_0 = 63{,}1 - b_1 \cdot 43{,}09 = 5{,}1249$$

Spørsmål 2

Lederne av en fransk butikkjede innen detaljhandel ønsker å undersøke sammenhengen mellom ukentlig omsetning Y (i tusen euro) og butikkareal x (i kvadratmeter) for butikkene sine. Fra seks butikker har de samlet inn dataene $\{(x_1, y_1), \ldots, (x_6, y_6)\}$ der x_i og y_i betegner henholdsvis areal og ukentlig omsetning for butikk nr. $i = 1, \ldots, 6$. Dataene er som gjengitt i tabellen under.

x_i	25	32	39	28	48	17
y_i	4.7	4.3	6.73	3.97	7.0	2.49

Det oppgis at:

$\sum_{i=1}^{6} x_i = 189$

$\sum_{i=1}^{6} y_i = 29.19$

$\sum_{i=1}^{6} x_i y_i = 1007.06$

$\sum_{i=1}^{6} x_i^2 = 6547.0$

Lederne antar at dataene følger en enkel lineær regresjonsmodell og ønsker å benytte dem til å tilpasse en lineær regresjonslinje, $\hat{y} = b_0 + b_1 x$, ved hjelp av minste kvadraters metode.

(a) Hva blir verdiene til b_0 og b_1? Oppgi svarene som desimaltall med fire desimaler, for eksempel 13.7461 og 0.0530.

Har koeffisientene fra oppgave 1

$$\Rightarrow \quad \hat{b}_1 = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2} = \frac{\sum x_i y_i - \bar{x}\sum y_i - \bar{y}\sum x_i + \sum \bar{x}\bar{y}}{\sum x_i^2 - 2\bar{x}\sum x_i + \sum \bar{x}^2}$$

$$= \frac{\sum x_i y_i - \frac{1}{n}\sum x_i \sum y_i - \frac{1}{n}\sum x_i \sum y_i + \cancel{n} \cdot \left(\frac{1}{\cancel{n}}\sum x_i\right)\cdot\left(\frac{1}{n}\sum y_i\right)}{\sum x_i^2 - 2\cdot\frac{1}{n}\left(\sum x_i\right)^2 + n\cdot\left(\frac{1}{n}\sum x_i\right)^2}$$

$$= \frac{\sum x_i y_i - \frac{1}{n}\sum x_i \cdot \sum y_i}{\sum x_i^2 - \frac{1}{n}\left(\sum x_i\right)^2}$$

$$= \frac{1007{,}06 - \frac{1}{6}\cdot 189 \cdot 29{,}19}{6547{,}0 - \frac{1}{6}\cdot 189^2}$$

$$\underline{\underline{= 0{,}14756}}$$

$$b_0 = \bar{y} - b_1\bar{x} = \frac{1}{n}\sum y_i - b_1 \frac{1}{n}\sum x_i = \frac{1}{6}\left(29{,}19 - 0{,}14756 \cdot 189\right)$$

$$\underline{\underline{= 0{,}21696}}$$

(b) For en butikk med areal $x_0 = 52$ kvadratmeter, hva blir predikert ukentlig omsetning $\hat{y}_0$? Oppgi svaret som et desimaltall med to desimaler, for eksempel 13.74 og 0.05.

$$\hat{y}_0 = b_0 + b_1 \bar{x}$$

$$= 0{,}21696 + 0{,}14756 \cdot 52$$

$$= \underline{\underline{7{,}89}}$$

Spørsmål 3

Ein klesbutikk i Trondheim har over ein periode på ni veker kjøpt annonseplass i Adressavisa. Samtidig har dei målt salet i kvar veke. Klesbutikken ynskjer å undersøke samanhengen mellom reklamekostnadene og salet.

Dei registrerte salsinntektene $y_1, \ldots, y_{10}$ samt tilhøyrande reklamekostnader $x_1, \ldots, x_{10}$ frå dei ni vekene er oppsummert i følgande tabell:

Veke nr	Reklamekostnad (i kr)	Salsinntekt (i kr)
1	2500	12000
2	3000	13000
3	3000	12900
4	3500	12700
5	4500	13600
6	5000	14600
7	5000	14700
8	5500	14700
9	6500	15800
10	7000	16100

Du kan bruke følgande (avrunda) summer i utrekningane dine:

$\bar{x} = 4550.0$

$\sum_{i=1}^{10}(x_i - \bar{x})^2 = 21.225 \cdot 10^6$

$\bar{y} = 14010.0$

$\sum_{i=1}^{10}(x_i - \bar{x})(y_i - \bar{y}) = 18.645 \cdot 10^6$

Anta at salsinntekta er ein normalfordelt stokastisk variabel med $\mathrm{E}(Y) = \beta_0 + \beta_1 x$ og $\mathrm{Var}(Y) = \sigma^2$. Anta vidare at salsinntekta for forskjellige veker er uavhengige av kvarandre, og at variansen σ^2 er ukjend.

Gitt at klesbutikken bruker 6000 kr på reklamekostnader i veke 12, kva er forventa salsinntekt i kr? Oppgi svaret som næraste multippel av 10, for eksempel 10240 eller 75660.

$$\boxed{\hat{b}_1 = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sum(x_i - \bar{x})^2}} = \frac{18{,}645 \cdot 10^6}{21{,}225 \cdot 10^6} = 0{,}878$$

$$\boxed{\hat{b}_0 = \bar{y} - \hat{b}_1 \bar{x}} = 14010 - 0{,}878 \cdot 4550 = 10015$$

$$y_0 = 10015 + 0{,}878 \cdot 6000 = 15283 \approx \underline{\underline{15280}}$$

Spørsmål 4

Mosjonisten Eva startet med løpetrening for 5 år siden, og har hvert år siden deltatt i samme distanse under Trondheim maraton. Tiden i minutter hun har brukt på hvert løp, samt alderen hennes er oppgitt i tabellen nedenfor.

Alder, x_i	51	52	53	54	55
Tid, y_i	49.62	50.74	49.06	47.5	46.99

Eva ønsker å undersøke tiden sin mot alderen sin, og antar at tidene hennes kan sees på som realisasjoner av uavhengige normalfordelte variabler $Y_1, \ldots, Y_5$, hvor $\mathrm{E}(Y_i) = \beta_0 + \beta_1 x_i$ og $\mathrm{Var}(Y_i) = \sigma^2$.

Fra dataene har Eva regnet ut følgende:

$\bar{x} = \frac{1}{5}\sum_{i=1}^{5} x_i = 53$,

$\bar{y} = \frac{1}{5}\sum_{i=1}^{5} y_i = 48.782$,

$S_{xx} = \sum_{i=1}^{5}(x_i - \bar{x})^2 = 10$,

$S_{yy} = \sum_{i=1}^{5}(y_i - \bar{y})^2 = 9.468$ og

$S_{xy} = \sum_{i=1}^{5}(x_i - \bar{x})(y_i - \bar{y}) = -8.5$.

(a) Regn ut minste kvadratsums estimater, b_0 og b_1, for henholdsvis β_0 og β_1. Oppgi svarene som desimaltall med tre desimaler, for eksempel 4.453 og 1.234.

$$\boxed{\hat{b}_1 = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sum(x_i - \bar{x})^2}} = \frac{-8{,}5}{10} = \underline{\underline{-0{,}85}}$$

$$\boxed{\hat{b}_0 = \bar{y} - \hat{b}_1 \bar{x}} = 48{,}782 + 0{,}85 \cdot 53 = \underline{\underline{93{,}832}}$$

(b) Eva vil gjennomføre en hypotesesetest av $H_0 : \beta_1 = 0$ mot $H_1 : \beta_1 \neq 0$. Hva blir verdien av observatoren hun da baserer testen sin på? Du kan bruke at $S^2 = (S_{yy} - B_1 S_{xy})/(5-2)$ er en forventningsrett estimator for σ^2. Oppgi svaret som et desimaltall med tre desimaler, for eksempel 1.345 og 9.203.

Fra foreløsning vet vi at

$$\hat{\beta}_1 \sim N\left(\beta_1, \frac{\sigma^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}\right)$$

Standardiserer:

$$Z = \frac{\hat{\beta}_1 - \beta_1}{\sqrt{\frac{\sigma^2}{\sum(x_i - \bar{x})^2}}} \sim N(0,1)$$

Kjenner ikke σ^2

$$\hat{\sigma}^2 = S^2 = \frac{S_{yy} - \hat{\beta}_1 S_{xy}}{(n-2)}$$

La $V = \frac{(n-2)S^2}{\sigma^2}$, denne er χ^2-fordelt med n-2 frihetsgrader

(vist i forelesning)

Bruker isteden at $T = \frac{Z}{\sqrt{V/\nu}}$ er student t-fordelt med ν frihetsgrader

$$T = \frac{\frac{\hat{\beta}_1 - \beta_1}{\sqrt{\frac{\sigma^2}{\sum(x_i-\bar{x})^2}}}}{\sqrt{\frac{\frac{(n-2)S^2}{\sigma^2}}{(n-2)}}} \sim t_{n-2}$$

$$T = \frac{\frac{\hat{\beta}_1 - \beta_1}{\sqrt{\frac{\sigma^2}{\sum(x_i-\bar{x})^2}}}}{\sqrt{\frac{\frac{\cancel{(n-2)}S^2}{\sigma^2}}{\cancel{(n-2)}}}} = \frac{\hat{\beta}_1 - \beta_1}{\sqrt{\frac{\cancel{\sigma^2}}{\sum(x_i-\bar{x})^2}}} \cdot \frac{1}{\sqrt{\frac{S^2}{\cancel{\sigma^2}}}} = \frac{\hat{\beta}_1 - \beta_1}{\sqrt{\frac{S^2}{\sum(x_i-\bar{x})}}} \sim t_{n-2}$$

T er testobservator, under H_0 er $\beta_1 = 0$

$$\Rightarrow \quad T = \frac{\hat{\beta}_1 - \beta_1}{\sqrt{\frac{S^2}{\sum(x_i-\bar{x})}}} = \frac{-0,85-0}{\sqrt{\frac{0,74767}{10}}} = -3,1086$$

$$S^2 = \frac{9,468 - (-0,85 \cdot (-8,5))}{5-2} = 0,74767$$

(c) Regn ut "coefficient of determination" $R^2 = 1 - \frac{\text{SSE}}{\text{SST}}$ for den tilpassede modellen. Husk at R^2 er et tall i [0,1], $\text{SSE} = (5-2) \cdot S^2$ og $\text{SST} = S_{yy}$. Oppgi svaret som et desimaltall med tre desimaler, for eksempel 0.345 og 0.900.

$$R^2 = 1 - \frac{SSE}{SST} = 1 - \frac{3 \cdot S^2}{S_{yy}} = 1 - \frac{(9{,}468 - 0{,}85 \cdot 8{,}5)}{9{,}468} = \frac{0{,}85 \cdot 8{,}5}{9{,}468} = \underline{\underline{0{,}763}}$$

Spørsmål 5

Oddvar er ein ivrig syklist, og han vil no undersøke kor mange desiliter væske han inntar i løpet av kvar sykkeltur. La x være lengda, i km, på turen og Y tilhøyrande væskeforbruk, i desiliter. Oddvar kan sjølve velge lengda på sykkelturen, slik at x ikkje blir sett på som ein stokastisk variabel.

Oddvar antar at væskeinntaket Y er ein normalfordelt stokastisk variabel med forventning $E(Y) = \beta_0 + \beta_1 x$ og varians $\text{Var}(Y) = \sigma^2$. Både β_0, β_1 og σ^2 er ukjende parametre. Han antar videre at væskeinntaka $Y_1, \ldots, Y_n$ for n turar med tilhøyrande lengder $x_1, \ldots, x_n$ er uavhengige.

Etter $n = 20$ sykkelturar har Oddvar målt følgande:

$S_{xx} = \sum_{i=1}^{20}(x_i - \bar{x})^2 = 80$

$S_{yy} = \sum_{i=1}^{20}(y_i - \bar{y})^2 = 1.5$

$S_{xy} = \sum_{i=1}^{20}(x_i - \bar{x})(y_i - \bar{y}) = 6.0$

(a) Rekn ut verdien til eit forventningsrett estimat for variansen σ^2. Oppgi svaret som eit desimaltal med tre desimalar, for eksempel 0.125 eller 0.053.

Bruker estimatoren fra forrige oppgave

$$\hat{\sigma}^2 = S^2 = \frac{S_{yy} - \hat{\beta}_1 S_{xy}}{n-2} = \frac{1{,}5 - 0{,}075 \cdot 6}{20-2} = \underline{\underline{0{,}05833}} \left(= \frac{7}{120}\right)$$

$$\boxed{\hat{b}_1 = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sum(x_i - \bar{x})^2}} = \frac{S_{xy}}{S_{xx}} = \frac{6}{80} = \frac{3}{40} = 0{,}075$$

Bruker testobservatoren fra tidligere:

$$T = \frac{\hat{\beta}_1 - \beta_1}{\sqrt{\frac{S^2}{\sum(x_i - \bar{x})}}} \sim t_{n-2}$$

Får da:

$$P(-t_{\alpha/2, n-2} \leq T \leq t_{\alpha/2, n-2}) = 0{,}95 = 1 - \alpha \Rightarrow \alpha = 0{,}05$$

$$P\left(-t_{0{,}025,18} \leq \frac{\hat{\beta}_1 - \beta_1}{\sqrt{\frac{S^2}{S_{xx}}}} \leq t_{0{,}025,18}\right) = 0{,}95$$

Løser ulikhetene, får da intervallet:

$$\left[\hat{\beta}_1 - t_{0{,}025,18}\sqrt{\frac{S^2}{S_{xx}}}\ ,\ \hat{\beta}_1 + t_{0{,}025,18}\sqrt{\frac{S^2}{S_{xx}}}\right]$$

Med $S^2 = 0{,}058$, $S_{xx} = 80$, $t_{0{,}025,18} = 2{,}101$, $\hat{\beta}_1 = 0{,}075$

$$\underline{\underline{[0{,}018,\ 0{,}132]}}$$

(c) Utfallet av hypotesetesten $H_0 : \beta_1 = 0$ vs. $H_1 : \beta_1 \neq 0$ for 5% signifikansnivå blir å forkaste H_0. True eller False?

Forkaster H_0 dersom $|\hat{\beta}_1|$ er for langt unna 0, altså at $|\hat{\beta}_1| > k$

Ønskes dermed at:

$$P(\hat{\beta}_1 \geq k \mid \beta_1 = 0) \leq 0{,}05$$

$$\Rightarrow P\left(|T| \geq t_{\alpha/2} \mid \beta_1 = 0\right) \leq 0{,}05 = \alpha$$

$$P\left(T \geq t_{\alpha/2} \cup T \leq -t_{\alpha/2}\right) \leq 0{,}05 \qquad t_{\alpha/2} = t_{0{,}025,18} = 2{,}101$$

$$T_{obs} = \frac{\hat{\beta}_1}{\sqrt{\frac{S^2}{S_{xx}}}} = \frac{S_{xy}/S_{xx}}{\sqrt{S^2/S_{xx}}} = \frac{6/80}{\sqrt{\frac{7}{120}/80}} = 2{,}777$$

$$\Rightarrow |T_{obs}| = 2{,}777 > t_{\alpha/2} = 2{,}101 \Rightarrow \text{Forkast } H_0 \Rightarrow \underline{\underline{\text{True}}}$$

Spørsmål 6

Regresjonsmodellen $\widehat{Y}_i = B_0 + B_1 x_i$ har blitt tilpasset et datasett bestående av 26 observasjoner (x_i, y_i) for i = 1,...,26. Ved hjelp av minste kvadraters metode har man funnet estimatene $b_0 = 24.74$ og $b_1 = 0.726$.

Det oppgis følgende:

$\bar{x} = \frac{1}{26}\sum_{i=1}^{26} x_i = 17.3$

$\bar{y} = \frac{1}{26}\sum_{i=1}^{26} y_i = 37.3$

$S_{xx} = \sum_{i=1}^{26}(x_i - \bar{x})^2 = 519.9$

$S_{yy} = \sum_{i=1}^{26}(y_i - \bar{y})^2 = 348.1$

$S_{xy} = \sum_{i=1}^{26}(x_i - \bar{x})(y_i - \bar{y}) = 377.3$

og at

$S^2 = \frac{1}{n-2}\sum_{i=1}^{26}(Y_i - \widehat{Y}_i)^2 = \frac{1}{n-2}(S_{yy} - B_1 S_{xy})$

er en forventningsrett estimator for $\sigma^2 = \mathrm{Var}[Y_i], i = 1, 2, \ldots, 26$

(a) Regn først ut et forventningsrett estimat for $\mathrm{Var}[\widehat{Y}_0] = \mathrm{Var}[B_0 + B_1 x_0] = \mathrm{Var}[\bar{Y} - B_1\bar{x} + B_1 x_0]$, der $x_0 = 6.5$. Du kan benytte at $\bar{Y} = \frac{1}{26}\sum_{i=1}^{26} Y_i$ og B_1 er uavhengige stokastiske variabler. Oppgi svaret som et desimaltall med tre desimaler, for eksempel 0.442 og 10.234.

Ettersom $\bar{Y}$ og $\hat{\beta}_1$ er uavhengige, er $\mathrm{cov}[\bar{Y}, \hat{\beta}_1] = 0$

$\Rightarrow$ Vi har at

$$\mathrm{Var}[\bar{Y} \underbrace{- \hat{\beta}_1\bar{x} + \hat{\beta}_1 x_0}_{\hat{\beta}_1(x_0 - \bar{x})}] = \mathrm{Var}[\bar{Y}] + (x_0 - \bar{x})^2\,\mathrm{Var}[\hat{\beta}_1]$$

Må finne $\mathrm{Var}[\hat{\beta}_1]$

Kan vise at $\sum(x_i - \bar{x})(Y_i - \bar{Y}) = \sum(x_i - \bar{x})Y_i$

$$\Rightarrow \hat{\beta}_1 = \frac{\sum(x_i - \bar{x})Y_i}{\sum(x_i - \bar{x})^2} = \frac{\sum(x_i - \bar{x})Y_i}{S_{xx}}$$

$$\mathrm{Var}[\hat{\beta}_1] = \mathrm{Var}\left[\frac{\sum(x_i - \bar{x})Y_i}{S_{xx}}\right] = \frac{1}{S_{xx}^2}\sum \mathrm{Var}[(x_i - \bar{x})Y_i]$$

$$= \frac{1}{S_{xx}^2}\sum(x_i - \bar{x})^2\,\mathrm{Var}[Y_i]$$

$$= \frac{1}{S_{xx}^{\cancel{2}}}\cancel{S_{xx}}\,\sigma^2 = \frac{\sigma^2}{S_{xx}}$$

$$Var[\bar{Y} - \hat{\beta}_1\bar{x} + \hat{\beta}_1 x_0] = Var[\bar{Y}] + (x_0-\bar{x})^2 \, Var[\hat{\beta}_1] \qquad Var[\bar{Y}] = \frac{\sigma^2}{n}$$

$$= \frac{\sigma^2}{n} + \frac{\sigma^2}{S_{xx}}(x_0-\bar{x})^2$$

$$= \sigma^2\left(\frac{1}{n} + \frac{(x_0-\bar{x})^2}{S_{xx}}\right)$$

Estimerer σ^2 med s^2: $$= s^2\left(\frac{1}{n} + \frac{(x_0-\bar{x})^2}{S_{xx}}\right)$$

$$= \frac{1}{n-2}\left(S_{yy} - \frac{S_{xy}}{S_{xx}} \cdot S_{xy}\right)\left(\frac{1}{n} + \frac{(x_0-\bar{x})^2}{S_{xx}}\right)$$

$$= \frac{1}{n-2}\left(S_{yy} - \frac{S_{xy}^2}{S_{xx}}\right)\left(\frac{1}{n} + \frac{(x_0-\bar{x})^2}{S_{xx}}\right)$$

$$= \frac{1}{24}\left(348{,}1 - \frac{377{,}3^2}{519{,}9}\right)\left(\frac{1}{26} + \frac{(6{,}5-17{,}3)^2}{519{,}9}\right)$$

$$= \underline{\underline{0{,}813}}$$

(b) Regn ut et 95% konfidensintervall for forventningsverdien $E(\widehat{Y}_0)$ i punktet $x_0 = 6.5$. Oppgi grensene til intervallet med to desimaler, for eksempel 0.45 og 23.93.

$$E[\hat{Y}_0] = E[\hat{\beta}_0 + \hat{\beta}_1 x_0] = \beta_0 - \beta_1 x_0$$

Normaliserer $\hat{Y}_0$:

$$Z = \frac{\hat{Y}_0 - E[\hat{Y}_0]}{\sqrt{Var[\hat{Y}_0]}} \sim N(0,1)$$

$$T = \frac{\hat{Y}_0 - E[\hat{Y}_0]}{\sqrt{est.\ Var[\hat{Y}_0]}} \sim t_{\alpha, n-2}$$

$Var[\hat{Y}_0]$ erstattes av estimatoren fra a), den er χ^2-fordelt med $\nu = n-2$ (den er alt delt på $\nu = n-2$)

$0{,}95 = 1-\alpha \Rightarrow \alpha = 0{,}05$, $\alpha/2 = 0{,}025$

$$\Rightarrow P(-t_{0,025,24} \leq T \leq t_{0,025,24}) = 0{,}95$$

$$P\left(-t_{0,025,24} \leq \frac{\hat{Y}_0 - E[\hat{Y}_0]}{\sqrt{est.\ Var[\hat{Y}_0]}} \leq t_{0,0025,24}\right) = 0{,}95$$

Løses ulikhetene:

$$\left[\hat{Y}_0 - t_{0{,}025,24}\sqrt{est.Var[\hat{Y}_0]},\ \hat{Y}_0 - t_{0{,}025,24}\sqrt{est.Var[\hat{Y}_0]}\right]$$

Har at $est.Var[\hat{Y}_0] = 0{,}813$, $\hat{Y}_0 = b_0 + b_1 x_0 = 24{,}74 + 0{,}726 \cdot 6.5 = 29{,}459$, $t_{0{,}025,24} = 2{,}064$

$$\underline{\underline{[27.598,\ 31.320]}}$$