LØSNINGSFORSLAG EKSAMEN 27. mai 2008

Oppgave 11

a) La Y_i = pris på bil i.

Modell 1: $Y_i = \alpha + \beta x_i + e_i$ der x_i = års modell i antall år etter år 2000 for bil i og $e_i \sim N(0,\sigma)$ og e_i-ene er feilledd som er uavhengige, i = 1, ...,10.

Modell 2: $Y_i = \alpha + \beta x_i + e_i$ der x_i = antall tusen kjørte km. for bil i og $e_i \sim N(0,\sigma)$ og e_i-ene er feilledd som er uavhengige, i = 1, ...,10.

Modell 1 har best tilpasningsevne til dataene fordi: R^2 = 71.7% som er bra, (over 70%) og residual plottet ser greit ut. (Siste påstand kan kanskje diskuteres)

Modell 2 har R^2 = 52.6% som er mindre enn for modell 1. Residual plottet i modell 2 har trakt-form, det er ikke så bra.

<u>Velger derfor modell 1.</u>

b) Nedre grense = $\hat{\beta} - t_{0.01,8}\hat{SE}(\hat{\beta})$ = $13.271 - 2.896 \cdot 2.948$ = 13.271 – 8.537 = 4.734

Øvre grense = $\hat{\beta} + t_{0.01,8}\hat{SE}(\hat{\beta})$ = $13.271 + 2.896 \cdot 2.948$ = 13.271 + 8.537 =21.808

<u>[4.73, 21.81] er et 98% konfidensintervall for β.</u>

Hvis modell 2 velges i 11a:

Nedre grense = $\hat{\beta} - t_{0.01,8}\hat{SE}(\hat{\beta})$ = $-0.615 - 2.896 \cdot 0.2064$ = -0.615 – 0.598 = -1.213

Øvre grense = $\hat{\beta} + t_{0.01,8}\hat{SE}(\hat{\beta})$ = $-0.615 + 2.896 \cdot 0.2064$ = -0.615 + 0.598 = -0.017

<u>[-1.213, -0.017] er et 98% konfidensintervall for β i modell 2.</u>

c) Vi bruker tilpasset modell 1: $\hat{Y} = 74.1 + 13.3x$

Bilen til damen har x = 1. $\hat{Y}_D = 74.1 + 13.3 \cdot 1 = 87.4$

<u>Hun bør sette prisen til 87400 kr.</u>

Bilen hun kan bytte til har x = 5: $\hat{Y}_B = 74.1 + 13.3 \cdot 5 = 140.6$

Mellomlegget = $\hat{Y}_B - \hat{Y}_D = 140.6 - 87.4 = 53.2$

<u>Hun må betale 53200 kr. i mellomlegg.</u>

Hvis modell 2 velges i 11a:

Da bruker vi tilpasset modell 2: $\hat{Y} = 158 - 0.615x$

Bilen til damen har x = 85. $\hat{Y}_D = 158 - 0.615 \cdot 85 = 105.725$

<u>Hun bør sette prisen til 105725 kr hvis hun bruker modell 2.</u>

Bilen hun kan bytte til har x = 37: $\hat{Y}_B = 158 - 0.615 \cdot 37 = 135.245$

Mellomlegget = $\hat{Y}_B - \hat{Y}_D = 135.245 - 105.725 = 29.52$

<u>Hun må betale 29520 kr. i mellomlegg hvis hun bruker modell 2.</u>

Oppgave 12

a) La Y_{ij} = gjennomsnittspoeng for tur j til sted i. Antar $Y_{ij} \sim N(\mu_i,\sigma)$ j = 1,…10 og i = 1, 2, 3. Alle observasjonene antas å være uavhengige.

Tester hypotesene H_0: $\mu_1 = \mu_2 = \mu_3$ mot H_1: minst 2 steder blir vurdert forskjellig.

p-verdien = 0.006 < 0.05 $\Rightarrow$ Forkaster H_0. Vi påstår at minst 2 steder blir vurdert som forskjellige.

b) Vi tester H_0: $\mu_1 - \mu_2 = 0$ mot H_1: $\mu_1 - \mu_2 > 0$

$\bar{y}_{1\bullet} = 4.51$ $S_1 = 0.3601$ $\bar{y}_{2\bullet} = 3.93$ $S_p = 0{,}4715$ fra data-utskriften.

$T = \dfrac{\bar{y}_{1\bullet} - \bar{y}_{2\bullet}}{S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} = \dfrac{4.51-3.93}{0.4715\sqrt{\frac{1}{10}+\frac{1}{10}}} = 2{,}75 > t_{0.05,27} = 1.703$ og vi kan forkaste

H_0 på 5% nivå. Alternativt kan vi (fra kalkulator) bruke p-verdi = 0.005 < 0.05 som også gir at vi skal forkaste H_0 på 5% nivå.

Påstår at sted 1 blir vurdert som bedre enn sted 2.

Oppgave 13

a) Etter ombyggingen antar vi at $X_i \sim N(\mu,90)$ i = 1,…7. og at det er uavhengige X-er.

Nedre grense = $\bar{X} - t_{0.025,8} \cdot \dfrac{\sigma}{\sqrt{n}} = 364.57 - 1.96 \cdot \dfrac{90}{\sqrt{7}} = 364.57 - 66.67 = 297.9$

Øvre grense = $\bar{X} + t_{0.025,8} \cdot \dfrac{\sigma}{\sqrt{n}} = 364.57 + 1.96 \cdot \dfrac{90}{\sqrt{7}} = 364.57 + 66.67 =$ 431.24

[297.9 , 431.24] er et 95% konfidensintervall for μ, forventet betalt beløp.

Tester: H_0: $\mu = 305$ mot H_1: $\mu \neq 305$, Finner 305 i intervallet, så H_0 må aksepteres. Vi kan ikke påstå at forventet betalt beløp har forandret seg etter ombyggingen. Testen har 5% nivå, siden det er et 95% konfidensintervall. ($\alpha = 0.05$)

b) Bruker formel 6.11 i læreboka:

$n \geq \left(\dfrac{2 \cdot z_{\alpha/2}\sigma}{L}\right)^2 = \left(\dfrac{2 \cdot 1.96 \cdot 90}{20}\right)^2 = 311.1696$ Dvs: Vi må ta 312 observasjoner.

Oppgave 14

La. $p = P(\,HIV\,)$ for tilfeldig valgt gruvearbeider.

$\hat{p} = \dfrac{X}{n} = \dfrac{12}{32} = 0.375$

Vi tester: H_0: p = 0.5 mot H_1: p < 0.5 (= p_0)

$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = -1.414 > -1.645 = -z_{0.05}$ og vi kan ikke forkaste H_0 på 5% nivå. Alternativt kan vi bruke (fra kalkulator) p-verdien = 0.079 > 0.05 og vi aksepterer H_0. <u>Det kan altså ikke påstås at mindre enn 50% av gruvearbeiderne i selskapet har HIV.</u>