Fasit eksamen STK1000 1.juni 2007

Oppgave 1

a). $\mu_X = 2.0$, $\sigma_X = 1.0$. Formler i kap. 4.4.

b). $Y \sim B(7, 0.1)$. $P(Y \geq 2) = 1 - P(Y < 2) = 1 - P(Y = 0) - P(Y = 1) = 0.1497(0.15)$.

Oppgave 2

a). $H_0 : \mu_1 - \mu_2 = 0$ 'Ingen forskjell i forventet benmasse'

$H_a : \mu_1 - \mu_2 \neq 0$ 'Forskjell i forventet benmasse'

$$t = \frac{\bar{x_1} - \bar{x_2}}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{12.6 - 11.6}{\sqrt{\frac{1.3^2}{60} + \frac{2.5^2}{60}}} = 2.7489$$

b). $P = 2P(T \geq 2.7489)$ der $T \sim t(59)$. Fra tabell ser vi at $P < 2 \cdot 0.005 = 0.01$. Forkaster H_0 og konkluderer med signifikant forskjell i forventet benmasse. Det er godtatt å benytte pooled-t-prosedyre i a)-c). I så fall har man $T \sim t(118)$.

c). Bruker $\bar{x_1} - \bar{x_2} \pm t^* \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$ der $t^* = 1.676$ finnes fra tabell (90%, 50 frihetsgrader (evt. 60)). Innsatt får vi intervallet (0.3903, 1.6097).

d). Pooled-t med $s_P = 1.2510$. Testobservator

$$t = \frac{\bar{x_1} - \bar{x_2}}{s_P \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = \frac{22.5 - 22.7}{1.251 \sqrt{\frac{1}{60} + \frac{1}{60}}} = -0.8757$$

$P = 2P(T \geq 0.8757)$ der $T \sim t(118)$. Fra tabell ser vi at $P > 2 \cdot 0.2 = 0.4$. Ingen grunn til å forkaste H_0: forventet vekt er like i de to gruppene.

e). 2-utvalgs-t-prosedyrer (og pooled-t) robust mot ikkenormalitet spesielt når $n_1 = n_2$. Har også at n_1 og n_2 er store.

Oppgave 3

a). Y_i PCB, responsvariabel, $p = 5$ PCB-typer forklaringsvariable $x_{i1}, x_{i2}, x_{i3}, x_{i4}, x_{i5}$, $i = 1, ..., 69$. Multippel lineær regresjons-modell:

$$Y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i4} + \beta_5 x_{i5} + \epsilon_i$$

der vi antar at ϵ_i er uavhengige og normalfordelte $N(0, \sigma)$ for alle i. Antakelsene OK fra normalfordelingsplott og residualplott.

Parametre i modellen er β_0, β_1 β_2, β_3, β_4, β_5 og σ, estimatene for disse finner vi fra utskriften som $b_0 = 1.741$, $b_1 = 3.4875$, $b_2 = 0.8653$, osv.... og $S = 5.93473$.

b). $H_0 : \beta_2 = 0 \quad$ mot $\quad H_a : \beta_2 \neq 0$

$t = \frac{b_2}{SE_{b_2}} = \frac{0.8653}{0.2607} = 3.3191 (= T)$.

$P = 2P(T \geq 3.3191)$ der $T \sim t(63)$, dvs. $(n - p - 1)$ frihetsgrader. Fra tabell ser vi at $P < 2 \cdot 0.001 = 0.002$.

c). Alle fem signifikante. $R^2 = 99.1\%$ meget høyt. Kan anta at de fem typene har god evne til å predikere total PCB.