Reeksamen juli 2006, Opgave 4

(a)

Estimaterne for de tre middelværdiparametre bliver

$$\hat{\mu}_1 = \frac{814}{5} = 162.80$$

$$\hat{\mu}_2 = \frac{4187}{26} = 161.04$$

$$\hat{\mu}_3 = \frac{1712}{11} = 155.64$$

Bidragene til residualkvadratsummen fra de tre grupper er

$132684 - 814^2/5 = 164.80$
$675015 - 4187^2/26 = 746.96$
$266936 - 1712^2/11 = 486.55$

så vi får

$\mathrm{SSD}_{\mathrm{res}} = 164.80 + 746.96 + 486.55 = 1398.31.$

Nu kan vi beregne estimatet for variansen:

$$\hat{\sigma}^2 = \frac{1398.31}{42 - 3} = 35.854\,.$$

95% Sikkerhedsgrænser for middelværdiestimaterne: Da 97.5%–fraktilen i T–fordelingen med 39 frihedsgrader ifølge tabellen er ca. 2.0225 får vi

$$\mu_1 = 162.80 \pm 2.0225\sqrt{\frac{35.854}{5}} = 162.80 \pm 5.42$$

svarende til sikkerhedsintervallet $[157.38, 168.22]$,

$$\mu_2 = 161.04 \pm 2.0225\sqrt{\frac{35.854}{26}} = 161.04 \pm 2.38$$

svarende til sikkerhedsintervallet $[158.66, 163.42]$, og

$$\mu_3 = 155.64 \pm 2.0225\sqrt{\frac{35.854}{11}} = 155.64 \pm 3.65$$

svarende til sikkerhedsintervallet $[151.99, 159.29]$.

(b)

Bartlett's test for varianshomogenitet:

$$\frac{39\log\frac{1398.31}{39} - 4\log\frac{164.80}{4} - 25\log\frac{746.96}{25} - 10\log\frac{486.55}{10}}{1 + \frac{1}{6}\left(\frac{1}{4} + \frac{1}{25} + \frac{1}{10} - \frac{1}{39}\right)} = 0.895\,.$$

Denne størrelse skal vurderes i en χ^2–fordeling med 2 frihedsgrader, hvor den er klart insignifikant. Varianserne i de tre aldersgrupper kan således antages at være ens.

Test for ens middelværdier: Vi udregner først den totale kvadratafvigelsessum SSD_y:

$\mathrm{S}_y = 814 + 4187 + 1712 = 6713$
$\mathrm{SS}_y = 132684 + 675015 + 266936 = 1074635.00$
$\mathrm{SSD}_y = 1074635 - 6713^2/42 = 1673.83.$

Herefter kan vi udregne F–teststørrelsen for homogenitet:

$$f = \frac{(1673.83 - 1398.31)/2}{35.85} = 3.84\,.$$

Denne størrelse skal vurderes i en F–fordeling med (2,39) frihedsgrader. Da 95%–fraktilen er ca. 3.23 udviser testet svag signifikans (99%–fraktilen er 5.18, så på basis af tabellen kan vi kun sige at $0.01 < P < 0.05$). Der er således en tendens til at højden afhænger af alderen. Estimaterne tyder på at den falder med alderen. Især ser det ud som om kvinder under 60 er højere end dem over 60.

(c)

Vi sammenligner de to yngste aldersgrupper:

$$t = \frac{162.80 - 161.04}{\sqrt{\left(\frac{1}{5} + \frac{1}{26}\right) \times 35.854}} = 0.602$$

Denne størrelse skal vurderes i en T–fordeling med 39 frihedsgrader, hvor den er klart insignifikant. Der er således ingen påviselig forskel mellem de to yngste grupper.