

Opgave 4, reeksamen august 2007.

(a)

Mellemregninger:

$$SSD_x = 302521 - 5361^2/100 = 15117.79$$

$$SSD_y = 67247.9691 - 2562.55^2/100 = 1581.3441$$

$$SPD_{xy} = 137934.91 - 5361 \times 2562.55/100 = 556.6045$$

$$SSD_{res} = 1581.3441 - 556.6045^2/15117.79 = 1560.8511.$$

Herefter fås parameterestimerne

$$\hat{\beta} = 556.6045/15117.79 = \mathbf{0.03682},$$

$$\hat{\alpha} = 2562.55/100 - (5361/100) \times 0.03682 = \mathbf{23.652},$$

$$\hat{\sigma}^2 = 1560.8511/(100 - 2) = \mathbf{15.927}.$$

(b)

Det forventede bmi for en person på 50 år er ifølge modellen $\alpha + \beta \times 50$. Estimatet for denne parameterfunktion er

$$\hat{\alpha} + \hat{\beta} \times 50 = 23.652 + 0.03682 \times 50 = \mathbf{25.49}$$

Med angivelse af 95% sikkerhedsgrænser fås (idet 97.5%-fraktilen i T-fordelingen med 98 frihedsgrader er 1.984)

$$\begin{aligned} \alpha + \beta \times 50 &= 25.49 \pm 1.984 \sqrt{\left(\frac{1}{100} + \frac{(53.61 - 50)^2}{15117.79} \right) 15.927} \\ &= 25.49 \pm 1.984 \times 0.4159 = 25.49 \pm 0.83 \end{aligned}$$

svarende til 95% sikkerhedsintervallet [**24.66,26.32**].

(c)

Den specificerede model svarer til at hvert køn har sin egen regressionslinie (hvor de i spørgsmål (a) og (b) antoges at have en fælles regressionslinie). Det første test (nederst i tabellen, med $P=0.424491$) viser, at de to linier kan antages at have samme hældning, altså at være parallelle. Det næste test ($P=0.225125$) viser at den fælles hældning kan sættes lig med 0, svarende til at bmi overhovedet ikke afhænger af alderen men kun af kønnet. Effekten af køn er derimod signifikant ($P=0.015139$), så vi ender med en ensidet variansanalysemodel, som går ud på at bmi kun afhænger af kønnet. Hvilken vej denne forskel går kan vi ikke sige noget om på basis af de oplysninger der er givet (med mindre man beregner de to gruppegennemsnit, men det kræves naturligvis ikke).