Reeksamen august 2002, Opgave 4

(a)

For aldersgruppen "–34" estimeres sandsynligheden for at svare "bedre" ved

$$\hat{p}_{13} = \frac{34}{71} = 0.4789$$

Approksimative 95% sikkerhedsgrænser fås efter formlen

$$p_{13} = \frac{34}{71} \pm 1.96\sqrt{\frac{34 \times (71 - 34)}{71^3}} = 0.4789 \pm 0.1162$$

svarende til sikkerhedsintervallet [0.3627,0.5951]. For aldersgruppe 2 (35–54) fås tilsvarende

$p_{23} = 0.3939 \pm 0.1179$ svarende til [0.2761,0.5118]

og for aldersgruppe 3 (55–)

$p_{33} = 0.1571 \pm 0.0853$ svarende til [0.0719,0.2424].

(b)

Hypotesen om at der ikke er forskel mellem de to yngste aldersgrupper svarer til uafhængighed i 2×3–tabellen som udgøres af de første to rækker. Vi udfører kvotienttestet:

$$2(21\log 21 + \cdots + 26\log 26 - 71\log 71 - \cdots - 60\log 60 + 137\log 137)$$

= 2.5724 (Pearson's test giver 2.5651). Denne størrelse skal vurderes i en χ^2–fordeling med (3-1)(2-1) = 2 frihedsgrader, hvor den er klart insignifikant. Der ser altså ikke ud til at være forskel på de to yngste grupper.

(c)

Hvis der overhovedet er forskel mellem aldersgrupperne skulle det således vise sig når vi tester for uafhængighed i hele tabellen. Kvotienttestet giver her

$$2(21\log 21 + \cdots + 11\log 11 - 71\log 71 - \cdots - 71\log 71 + 207\log 207)$$

= 22.0637 (Pearson: 20.6793). Denne størrelse skal vurderes i en χ^2–fordeling med (3-1)(3-1) = 4 frihedsgrader. Af tabellen ses at dette svarer til en P–værdi som er mindre end 0.0005, en halv promille. Vi må konkludere, at der er tydelig forskel mellem den ældste gruppe og de to andre (vi kunne måske have gjort dette endnu mere overbevisende ved at slå de to første grupper sammen, men det lægger opgaveteksten ikke op til). Forskellen skyldes først og fremmest, at der er meget færre af de ældre, som har svaret at deres situation er bedre end sidste år, end der er i de to yngste grupper.