

Opgave 3, eksamen juni 2008.

(a)

$$\hat{p}_1 = \frac{216}{395} = 0.5468,$$

$$\hat{p}_2 = \frac{1899}{4823} = 0.3937.$$

99% sikkerhedsgrænser for p_1 :

$$\begin{aligned} p_1 &= 0.5468 \pm 2.576 \times \sqrt{0.5468 \times (1 - 0.5468)/395} \\ &= 0.5468 \pm 2.576 \times 0.0250 = 0.5468 \pm 0.0644 \end{aligned}$$

svarende til 99% konfidensintervallet **[0.4824, 0.6112]**.

99% sikkerhedsgrænser for p_2 :

$$\begin{aligned} p_2 &= 0.3937 \pm 2.576 \times \sqrt{0.3937 \times (1 - 0.3937)/4823} \\ &= 0.3937 \pm 2.576 \times 0.007035 = 0.3937 \pm 0.0181 \end{aligned}$$

svarende til 99% konfidensintervallet **[0.3756, 0.4118]**.

(b)

Det test vi skal foretage er identisk med testet for uafhængighed i 2×2 tabellen:

$$\begin{aligned} -2 \log q &= 2(216 \log 216 + \cdots + 2924 \log 2924 \\ &\quad - 2115 \log 2115 - \cdots - 4823 \log 4823 \\ &\quad + 5218 \log 5218) = 34.77 \end{aligned}$$

som er større end 99.99% fraktilen i χ^2 -fordelingen med 1 frihedsgrad (15.137). Vi får altså en markant forkastelse af hypotesen $p_1 = p_2$, helt i overensstemmelse med at konfidensintervallerne ikke overlapper. Det er åbenbart de ledende medarbejdere, som er mest tilfredse med deres arbejde i virksomheden.

(c)

Reduktionen bliver klart forkastet, idet reduktionen i dimension er $(2 + 4 - 1) - 2 = 3$, og 99.99% fraktilen i χ^2 -fordelingen med 3 frihedsgrader er 21.108. Forskellen i tilfredshed kan altså ikke kun forklares ved at lederne er mere tilfredse end de menige medarbejdere. Ancienniteten spiller også ind. Hvordan – det kan vi ikke vide, uden at få oplyst estimaterne.

Modellen, hvor kun status indgår, er den samme som vi betragtede i spørgsmål (a) og (b), karakteriseret ved at sandsynligheden for at svare "tilfreds" alene afhænger af om man er leder eller ej. Blot var datasættet der reduceret til en 2×2 -tabel ved addition af binære variable med samme fordeling. Så konklusionen af det test, vi lige har foretaget, er at den model vi arbejdede med i første del af opgaven er utilstrækkelig.