

Reeksamen august 2010, Opgave 3

(a)

Hvis de fem sandsynlighedsparametre betegnes $p_1, \dots, p_5$ får vi for eksempel

$$\hat{p}_1 = \frac{21}{490} = 0.0429.$$

95% sikkerhedsgrænser fås ved hjælp af formelen

$$\begin{aligned} p_1 &= 0.0429 \pm 1.96 \sqrt{\frac{0.0429 \times (1 - 0.0429)}{490}} \\ &= 0.0429 \pm 1.96 \times 0.009154 = 0.0429 \pm 0.0179 \end{aligned}$$

svarende til sikkerhedsintervallet $[0.0250, 0.0608]$. Ved tilsvarende beregninger for de fire andre sandsynlighedsparametre fås estimater og konfidensgrænser som følger:

	Nedre	Estimat	Øvre
16-24	0.0250	0.0429	0.0608
25-34	0.0209	0.0361	0.0512
35-54	0.0831	0.1012	0.1193
55-64	0.0819	0.1144	0.1470
65-	0.0912	0.1171	0.1431

(b) Kvotienttestet for $p_1 = p_2 = p_3 = p_4 = p_5$, som beregningsmæssigt er det samme som testet for uafhængighed i antalstabellen, giver

$$\begin{aligned} -2 \log q &= 2(21 \log 21 + \dots + 520 \log 520 \\ &\quad - 261 \log 261 - \dots - 589 \log 589 \\ &\quad + 3095 \log 3095) = 50.1203. \end{aligned}$$

Denne størrelse skal vurderes i en χ^2 -fordeling med $(5 - 1)(2 - 1) = 4$ frihedsgrader, hvor den er ekstremt signifikant (99.99%-fraktilen er kun 23.513). Så der er ingen tvivl om at tilbøjeligheden til at være medlem af et politisk parti afhænger af alderen. Ser vi på estimaterne er der en ret markant stigning omkring de 35 år. Kort beskrevet kan vi sige, at andelen af befolkningen som er partimedlemmer ligger nogenlunde konstant omkring 2% for dem der er under 35, og omkring 8-9% for dem som er over 35. Forklaringen på dette kan være, at ældre mennesker interesserer sig mere for politik end yngre. Eller de interesserer sig for det på en sådan måde, at de får behov for at melde sig ind i et parti. Men ved fortolkningen skal man huske, at der er en indbygget stigning med alderen, som ligger i at når folk først har meldt sig ind bliver de i mange tilfælde ved med at være medlemmer resten af livet. Man kan således *ikke* konkludere ud fra tallene, at de yngre årgange er mindre "politiske" end de ældre, når man ser på det samlede livsforløb, for det vil jo først vise sig senere.

(c)

Her kan man foretage testet for uafhængighed i den 2×2 -tabel der udgøres af første og anden række. Hvis dette bliver godkendt kan vi antage, at $p_1 = p_2$, hvilket netop betyder at de to aldersgrupper kan slås sammen. Tabellen ser sådan ud:

Partimedlem	Ja	Nej	SUM
Ald.gr.	----- -----		
16-24	21	469	490
25-34	21	561	582
	----- -----		
SUM	42	1030	1072

Vi får

$$-2 \log q = 2(21 \log 21 + \dots) = 0.3232$$

som er insignifikant i en hvilken som helst χ^2 -fordeling (også den med 1 frihedsgrad). Vi kan altså slå de første to grupper sammen.

Dette fremgår også af listen af estimater og sikkerhedsgrænser fra spørgsmål (a). Hvis man gør opmærksom på det næsten totale overlap af de to sikkerhedsintervaller, er man tæt på at have besvaret spørgsmålet korrekt.