

Eksamen juni 2004, Opgave 3

(a)

Procentvis fordeling af alder inden for hver af de to respondentgrupper:

Alder:	–30	31–40	41–50	51–60	61–
Version:					
Net	35.5	28.4	17.3	13.1	5.7
Tlf	13.9	20.6	15.6	20.3	29.6

Som det umiddelbart ses er der en markant tendens til at “net-responderne” er yngre end “tlf-responderne”. Hvilket vel er helt i overensstemmelse med hvad man ville vente. Vi ved ikke noget om hvordan responderne er udvalgt, men hvis “net-responderne” er udvalgt fra en eller anden fortegnelse over e-mail adresser, vil der helt klart være overvægt af unge i denne gruppe. Noget lignende — bare omvendt — gælder hvis “tlf-responderne” er udvalgt fra en eller anden liste af fastnet telefonnumre. Og også selvom de adspurgte skulle være valgt på en måde der korrigerer for aldersskævhed, vil deres tilbøjelighed til at svare på spørgsmålene sikkert afspejle aldersgruppernes forskellige vaner omkring kommunikation.

(b)

Testet vi skal udføre er det sædvanlige test for uafhængighed:

$$-2 \log q = 2 \times (125 \log 125 + \dots + 89 \log 89 - \dots + 653 \log 653) = 99.28$$

som ligger ekstremt langt ude til højre i χ^2 -fordelingen med 4 frihedsgrader (99.99%-fraktilen er 23.513). Vi konkluderer, at der er en tydelig forskel mellem de to aldersfordelinger.

Pearson's test giver 94.35 (altså samme konklusion).

(c)

Estimatet for p_{11} = sandsynligheden for at en net-responent er under 31 bliver

$$125/352 = 0.3551.$$

Approksimative 95%-sikkerhedsgrenser udregnes ved hjælp af den sædvanlige formel:

$$p_{11} = 0.3551 \pm 1.96 \times \sqrt{\frac{1}{352} \times \frac{125}{352} \times \frac{352 - 125}{352}} = 0.3551 \pm 0.0500$$

svarende til sikkerhedsintervallet [0.3051,0.4051].