

### Reeksamen juli 2006, Opgave 3

(a)

Selve estimaterne for de fire sandsynlighedsparemetre bliver

$$\hat{p}_1 = \frac{167}{1411} = 0.1184,$$

$$\hat{p}_2 = \frac{261}{2272} = 0.1149,$$

$$\hat{p}_3 = \frac{262}{2874} = 0.09116,$$

$$\hat{p}_4 = \frac{93}{1464} = 0.06352.$$

Udregning af 95% sikkerhedsgrænser:

$$p_1 = 0.1184 \pm 1.96 \times \sqrt{\frac{0.1184 \times (1 - 0.1184)}{1411}} = 0.1184 \pm 0.0169$$

svarende til sikkerhedsintervallet  $[0.1015, 0.1353]$ .

$$p_2 = 0.1149 \pm 1.96 \times \sqrt{\frac{0.1149 \times (1 - 0.1149)}{2272}} = 0.1149 \pm 0.0131$$

svarende til sikkerhedsintervallet  $[0.1018, 0.1280]$ .

$$p_3 = 0.09116 \pm 1.96 \times \sqrt{\frac{0.09116 \times (1 - 0.09116)}{2874}} = 0.09116 \pm 0.01053$$

svarende til sikkerhedsintervallet  $[0.08063, 0.10169]$ .

$$p_4 = 0.06352 \pm 1.96 \times \sqrt{\frac{0.06352 \times (1 - 0.06352)}{1464}} = 0.06352 \pm 0.01249$$

svarende til sikkerhedsintervallet  $[0.05103, 0.07601]$ .

(b)

Testet for homogenitet er ækvivalent med det sædvanlige test for uafhængighed i  $4 \times 2$ -tabellen. Vi får således teststørrelsen

$$-2 \log q = 2(167 \log 167 + \dots + 1371 \log 1371)$$

$$-783 \log 783 - \dots - 1464 \log 1464$$

$$+8021 \log 8021) = 36.97$$

som skal vurderes i en  $\chi^2$ -fordeling med  $(4-1)(2-1)=3$  frihedsgrader. Her er den større end 99.99%-fraktilen (21.108), så vi kan med meget stor sikkerhed forkaste hypotesen om homogenitet. Dvs. sandsynligheden for at gå på tvangsauktion afhænger af alderen, tydeligvis sådan at jo ældre man er jo mindre sandsynligt er det at man går på tvangsauktion.

(Pearsons teststørrelse, som lige så godt kan bruges, bliver 35.25, uændret konklusion).

(c)

Den model vi betragtede i spørgsmål (a) og (b) svarer til at  $\text{logit}(p_i)$  som funktion af alderen er stykkevis konstant på de fire intervaller, som svarer til aldersgrupperne. I model 2 er  $\text{logit}(p_i)$  en lineær funktion af alderen, og i model 1 er denne lineære funktion erstattet med en funktion som er stykket sammen af lineære funktioner på de fire intervaller med hver sin hældning og afskæring. Da model 1 således svarer til en ret fleksibel klasse af funktioner, må det antages at næsten enhver afvigelse fra linearitet vil føre til at model 1 beskriver afhængigheden af alder bedre end model 2. Testet for reduktion af model 1 til model 2 kan således fortolkes som et test for, om afhængigheden af alder er logit-lineær. Model 1 har  $4+4=8$  parametre, medens model 2 har  $1+1=2$ . Faldet i dimension er således 6, så teststørrelsen 7.45 skal vurderes i en  $\chi^2$ -fordeling med 6 frihedsgrader. Her er den klart insignifikant, dvs. reduktionen af model 1 til model 2 kan godkendes. Det ser således ud til at en simpel logit-lineær model fuldtud forklarer alderens effekt på sandsynligheden for tvangsauktion.