Kapitel 5

Analyse af tosidede antalstabeller

Testet for *uafhængighed* i en tosidet antalstabel er, sammen med en række deraf afledte tests i flersidede antalstabeller, det vigtigste test af den type, der i kapitel 4 er beskrevet som "test for dimensionsreducerende hypotese" i polynomialfordelingen.

Eksempel 5.1. I kapitel 3 (side 14) betragtede vi tabellen

	Tv.auk.	Ikke tv.auk.	Sum
huse	656	5333	5989
lejl.	148	869	1017
Sum	804	6202	7006

hvor de 7006 låntagere i BRF Kredit, som pr. 1/1 1990 ikke havde betalt deres termin for december 1989 rettidigt, er optalt efter ejendommens kategori (parcelhus eller ejerlejlighed) og efter om de senere gik på tvangsauktion. Dette er hvad vi forstår ved en *tosidet antalstabel* (også kaldet en tosidet *kontingenstabel*), i dette tilfælde en 2×2–tabel. Som model for disse tal foreslog vi i kapitel 3 at opfatte rækkesummerne som faste, og antage at antal tvangsauktioner i de to grupper var fremkommet som værdier af binomialfordelte stokastiske variable med hver sin sandsynlighedsparameter. En næsten lige så naturlig model er imidlertid følgende, der fremkommer når totalsummen 7006 opfattes som fast: Antag at observationen

$$(x_{11}, x_{12}, x_{21}, x_{22}) = (656, 5333, 148, 869)$$

er værdien af en stokastisk variabel $(X_{11}, X_{12}, X_{21}, X_{22})$ som er polynomialfordelt af orden 4 med antalsparameter 7006 og ukendte sandsynlighedsparametre p_{11}, p_{12}, p_{21} og p_{22}. Ifølge polynomialfordelingens definition svarer dette til, at de 7006 låntagere placeres tilfældigt i tabellens celler. Eller, mere præcist: Med betegnelserne

$$S_i = \begin{cases} 1 \text{ hvis person nr. } i \text{ går på tvangsauktion,} \\ 2 \text{ ellers} \end{cases}$$

for person i's "tvangsauktionsstatus" og

$$B_i = \begin{cases} 1 \text{ hvis person nr. } i \text{ bor i hus,} \\ 2 \text{ hvis person nr. } i \text{ bor i lejlighed} \end{cases}$$

for person i's boligform, antages at de 7006 stokastiske variable (B_i, S_i), $i = 1, \ldots, 7006$ er uafhængige, identisk fordelte med fordeling givet ved

$$P((B_i, S_i) = (b, s)) = p_{bs}.$$

5.1. Hypotesen om uafhængighed.

Den hypotese vi skal interessere os for går simpelthen ud på, at B_i og S_i er stokastisk uafhængige. Altså at

$$P((B_i, S_i) = (b, s)) = P(B_i = b)\, P(S_i = s),$$

eller

$$p_{bs} = p_{b\cdot} p_{\cdot s}$$

hvor en prik i stedet for et index som sædvanlig bruges til at markere, at der er summeret over det pågældende index (altså f.eks. $p_{b\cdot} = p_{b1} + p_{b2}$). At tvangsauktionsstatus og boligform er stokastisk uafhængige kan jo også (jvf. Ssr side 21–22) udtrykkes sådan, at den betingede sandsynlighed for at en person ryger på tvangsauktion, givet hans/hendes boligform, ikke afhænger af boligformen. Kort sagt, at tvangsauktionsrisikoen er den samme for husejere og lejlighedsejere. Det var netop den hypotese vi testede i kapitel 3. Der kaldte vi det "sammenligning af to binomialfordelinger". Her kommer det til at hedde "test for uafhængighed i en 2×2 antalstabel". Men fortolkningsmæssigt er de to hypoteser svære at skelne fra hinanden, og det viser sig da også, som vi nu skal se, at kvotientteststørrelserne bliver ens og skal vurderes i samme χ^2–fordeling.

5.2. Kvotienttestet for uafhængighed.

I den fulde model har vi (jvf. kapitel 4) maksimaliseringsestimaterne

$$\hat{p}_{11} = \frac{656}{7006} \qquad \hat{p}_{12} = \frac{5333}{7006} \qquad \hat{p}_{21} = \frac{148}{7006} \qquad \hat{p}_{22} = \frac{869}{7006}$$

Hypotesen $p_{bs} = p_{b\cdot} p_{\cdot s}$ om uafhængighed kan parametriseres på formen

$$p_{bs} = \rho_b \sigma_s$$

hvor de fire parametre

$$\rho_1 = p_{1\cdot}\,, \quad \rho_2 = p_{2\cdot}\,, \quad \sigma_1 = p_{\cdot 1}\,, \quad \sigma_2 = p_{\cdot 2}\,,$$

i virkeligheden kun repræsenterer en hypotese af dimension 2, fordi vi har

$$\rho_1 + \rho_2 = 1 \text{ og } \sigma_1 + \sigma_2 = 1.$$

Likelihoodfunktionen under uafhængighedsmodellen bliver som funktion af disse parametre

$$L(\rho_1, \rho_2, \sigma_1, \sigma_2) = (\rho_1\sigma_1)^{656} (\rho_1\sigma_2)^{5333} (\rho_2\sigma_1)^{148} (\rho_2\sigma_2)^{869}$$

$$= \left(\rho_1^{5989}\rho_2^{1017}\right) \left(\sigma_1^{804}\sigma_2^{6202}\right).$$

Denne funktion genkender vi som produktet af to likelihoodfunktioner for polynomialfordelingsmodeller af orden 2. Modellerne kan oven i købet fortolkes som modeller for henholdsvis låntagernes fordeling på boligform og deres fordeling på tvangsauktionsstatus. Da maksimeringen af likelihooden kan foretages ved maksimering af de to faktorer hver for sig følger det umiddelbart, at ML–estimatorerne under uafhængighedsmodellen er givet ved

$$\hat{\rho}_1 = \frac{5989}{7006} \qquad \hat{\rho}_2 = \frac{1017}{7006} \qquad \hat{\sigma}_1 = \frac{804}{7006} \qquad \hat{\sigma}_2 = \frac{6202}{7006}$$

Kvotientteststørrelsen bliver således

$$q = \frac{\left(\hat{\rho}_1^{5989}\ \hat{\rho}_2^{1017}\right)\ \left(\hat{\sigma}_1^{804}\ \hat{\sigma}_2^{6202}\right).}{\hat{p}_{11}^{656}\ \hat{p}_{12}^{5333}\ \hat{p}_{21}^{148}\ \hat{p}_{22}^{869}}$$

$$= \frac{\left(\left(\frac{5989}{7006}\right)^{5989}\ \left(\frac{1017}{7006}\right)^{1017}\right)\ \left(\left(\frac{804}{7006}\right)^{804}\ \left(\frac{6202}{7006}\right)^{6202}\right).}{\left(\frac{656}{7006}\right)^{656}\ \left(\frac{5333}{7006}\right)^{5333}\ \left(\frac{148}{7006}\right)^{148}\ \left(\frac{869}{7006}\right)^{869}.}$$

Dette er præcis den samme kvotientteststørrelse, som vi i kapitel 3 (side 16–17) beregnede for de samme tal i forbindelse med sammenligning af to binomialfordelinger. Som i kapitel 3 fås

$$-2\log q = 2(656\log 656 + \cdots + 869\log 869 + 7006\log 7006$$

$$-5989\log 5989 - \cdots - 6202\log 6202) = 10.46,$$

og antallet af frihedsgrader som skal benyttes er også det samme, nemlig $(4-1) - 2(2-1) = 1$. Testets konklusion er derfor den samme — nemlig at vi, med en P–værdi på ca. 0.001 og i hvert fald højst 0.005, må forkaste hypotesen om uafhængighed.

Det er naturligvis også muligt at føre et rent matematisk bevis for det faktum, at de to kvotientteststørrelser er ens. Det afgørende argument i denne forbindelse er, at den model vi betragtede i kapitel 3 kan fortolkes som en *betinget model* i den polynomialfordelingsmodel, vi ser på i øjeblikket. Af opgave 3.4.3 (b2) i Ssr., samt lidt yderligere overvejelser vedrørende uafhængighed, ser man, at hvis vi i den fulde polynomialfordelingsmodel med parametre p_{bs} betinger med antal låntagere i de to ejendomstyper, så får vi netop en betinget fordeling, der for hver boligform beskriver antallet af tvangsauktioner som binomialfordelt. De to binomialfordelte størrelser (antal tvangsauktioner blandt husejere og antal tvangsauktioner blandt lejlighedsejere) bliver uafhængige og får hver sin sandsynlighedsparameter, nemlig (med notation som i kapitel 3)

$$p_1 = \frac{p_{11}}{p_{11} + p_{12}} \quad \text{og} \quad p_2 = \frac{p_{21}}{p_{21} + p_{22}}.$$

Hvis man nu ved brug af kædereglen (Ssr. sætning 2.3.1) omskriver likelihoodfunktionen for polynomialfordelingsmodellen til

$$L(p_{11}, \ldots, p_{22}) = P\big((X_{bs}) = (x_{bs})\big)$$

$$= P\big((X_{b\cdot}) = (x_{b\cdot})\big) P\big((X_{bs}) = (x_{bs}) \mid (X_{b\cdot}) = (x_{b\cdot})\big),$$

så vil man se, at værdien af den første (ikke–betingede) sandsynlighed på højre side bliver den samme uanset om man indsætter maksimaliseringsestimaterne under den fulde model eller under uafhængighedsmodellen. Både i den fulde model og i uafhængighedsmodellen vil rækkesummernes fordeling nemlig blive estimeret ved en polynomialfordeling, der som sandsynlighedsparametre har de tilsvarende relative hyppigheder. Derfor forkorter denne faktor ud, når kvotientteststørrelsen dannes, og det der bliver tilbage er netop kvotientteststørrelsen for sammenligning af de to binomialfordelinger.

5.3. Uafhængighed i en vilkårlig tosidet antalstabel.

Betragt en antalstabel

$$(x_{rs} \mid r = 1, \ldots, R,\ s = 1, \ldots, S)$$

med R rækker og S søjler. Vi anvender betegnelserne

$$x_{r\cdot} = x_{r1} + \cdots + x_{rS}, \quad x_{\cdot s} = x_{1s} + \cdots + x_{Rs}$$

for række– og søjlesummer, ligesom vi med $x_{\cdot\cdot}$ betegner summen af alle de RS ikke–negative heltal i tabellen. I analogi med den model, vi betragtede for 2×2–tabellen fra BRF Kredit, betragter vi modellen der fortolker tabellens indhold som udfaldet af en polynomialfordelt variabel med antalsparameter $n = x_{\cdot\cdot}$, orden RS og sandsynlighedsparametre $p_{rs}, r = 1, \ldots, R,\ s = 1, \ldots, S$. Parameterområdet er altså Δ_{RS}, så modellens dimension er $RS - 1$. Maksimaliseringsestimatorerne under denne model har vi tidligere udregnet til

$$\hat{p}_{rs} = \frac{x_{rs}}{n}.$$

Ved hypotesen om *uafhængighed* i en sådan tabel forstår vi hypotesen om, at parametrene p_{rs} kan skrives som produkter af række– og søjleparametre på formen

$$p_{rs} = \rho_r \sigma_s$$

hvor $\rho_1, \ldots, \rho_R$ og $\sigma_1, \ldots, \sigma_S$ er positive reelle tal med

$$\sum_{r=1}^{R} \rho_r = \sum_{s=1}^{S} \sigma_s = 1,$$

dvs.

$$((\rho_1, \ldots, \rho_R), (\sigma_1, \ldots, \sigma_S)) \in \Delta_R \times \Delta_S.$$

Denne parametrisering ses at være injektiv, idet ρ'er og σ'er jo kan rekonstrueres ud fra sandsynlighedsparametrene $p_{rs} = \rho_r \sigma_s$ som henholdsvis række- og søjlesummer. Dimensionen af uafhængighedsmodellen er således $(R-1)+(S-1) = R+S-2$.

Fortolkningen af hypotesen om uafhængighed er præcis den samme som i det specialtilfælde ($R = S = 2$) vi har betragtet. Under den sædvanlige fortolkning af polynomialfordelingen som fordelingen af antallene i tabellens celler, når n individer placeres tilfældigt i tabellen efter en sandsynlighedsfordeling på $\{1, \ldots, R\} \times \{1, \ldots, S\}$, svarer uafhængighedshypotesen til den antagelse, at rækkenummeret og søjlenummeret for et individ er uafhængige.

Som i det specialtilfælde vi har betragtet splitter likelihoodfunktion op i et produkt af to funktioner af hvert sit sæt parametre:

$$L\left((\rho_1, \ldots, \rho_R), (\sigma_1, \ldots, \sigma_S)\right) = \prod_{r=1}^{R} \prod_{s=1}^{S} (\rho_r \sigma_s)^{x_{rs}}$$

$$= \left(\prod_{r=1}^{R} \rho_r^{x_{r\cdot}}\right) \left(\prod_{s=1}^{S} \sigma_s^{x_{\cdot s}}\right)$$

Da de to faktorer i sidste udtryk hver for sig er likelihoodfunktioner for polynomialfordelingsmodeller (nemlig modellerne for henholdsvis række– og søjlemarginaler i tabellen) får vi maksimaliseringsestimatorerne

$$\hat{\rho}_r = \frac{x_{r\cdot}}{n} \text{ og } \hat{\sigma}_s = \frac{x_{\cdot s}}{n}.$$

Maksimaliseringsestimatorerne for selve sandsynlighedsparametrene under uafhængighedshypotesen bliver således

$$\hat{p}_{rs} = \frac{x_{r\cdot} x_{\cdot s}}{n^2}.$$

Kvotientteststørrelsen for test af uafhængighed mod den fulde model bliver, ifølge udregninger som vi overspringer fordi vi allerede har gennemført dem i 2×2–eksemplet,

$-2 \log q =$

$$2 \left(\sum_{r=1}^{R} \sum_{s=1}^{S} x_{rs} \log x_{rs} - \sum_{r=1}^{R} x_{r\cdot} \log x_{r\cdot} - \sum_{s=1}^{S} x_{\cdot s} \log x_{\cdot s} + n \log n \right).$$

Denne størrelse er, under hypotesen om uafhængighed, approksimativt χ^2–fordelt med

$$(RS - 1) - (R + S - 2) = (R - 1)(S - 1)$$

frihedsgrader.

5.4. De fittede værdier.

Det forventede antal i celle (r, s) er som bekendt

$$\mathrm{E}X_{rs} = np_{rs}.$$

Estimatoren for denne forventede værdi under uafhængighedsmodellen bliver således

$$\hat{\mathrm{E}}X_{rs} = n\hat{p}_{rs} = \frac{x_{r\cdot}x_{\cdot s}}{n}.$$

Disse størrelser kaldes i statistikerjargon de *fittede* (altså tilpassede) *værdier*. Ifølge en tommelfingerregel, som også blev nævnt i kapitel 4, kan χ^2–approksimationen til fordelingen af $-2 \log Q$ bruges når alle disse størrelser er mindst 5. Ofte vil man dog bruge approksimationen i tilfælde hvor nogle af de fittede værdier er helt nede i nærheden af 1. Så må man blot tage et vist forbehold vedrørende nøjagtigheden af den opnåede P–værdi.

Eksempel 5.1 fortsat. Betragt igen 2×2–tabellen vedrørende tvangsauktioner fra starten af dette kapitel. Følgende tabel viser de fittede værdier under hypotesen om uafhængighed sammen med de oprindelige antal:

	Tv.auk.		Ikke tv.auk.		Sum	
	Obs.	Forv.	Obs.	Forv.	Obs.	Forv.
huse	656	687.3	5333	5301.7	5989	5989.0
lejl.	148	116.7	869	900.3	1017	1017.0
Sum	804	804.0	6202	6202.0	7006	7006.0

Bemærk at marginalsummerne i 2×2–tabellen af fittede værdier er de samme som i den oprindelige tabel. Det kan man bruge til kontrol af at de fittede værdier er udregnet korrekt.

5.5. Pearson's test for uafhængighed.

Et intuitivt rimeligt kriterium for uafhængighed går ud på at de observerede værdier ligger tæt på de fittede værdier. Spørgsmålet er hvor store afvigelserne skal være, før de ikke længere kan bortforklares som tilfældigheder. En teststørrelse, der umiddelbart afspejler denne tankegang, er Pearson's teststørrelse, som i BRF–eksemplet bliver

$$\frac{(656-687.3)^2}{687.3}+\frac{(5333-5301.7)^2}{5301.7}+\frac{(148-116.7)^2}{116.7}+\frac{(869-900.3)^2}{900.3}$$

$= 11.09.$

I ord: "Summen af observerede minus fittede kvadreret, divideret med fittede". Denne teststørrelse er ifølge sætning 4.1 en approksimation til kvotientteststørrelsen, og kan således bruges i stedet for denne.

Som vi har set er kvotienttestet for uafhængighed i en 2×2–tabel formelt ækvivalent med kvotienttestet for at sandsynlighedsparametrene er ens i de to binomialfordelinger, som man får når rækkesummerne opfattes som givne. Det samme gælder om Pearson's teststørrelse, sådan forstået at den teststørrelse vi har beregnet ovenfor er den samme som den vi (ligeledes under navnet Pearson's teststørrelse) beregnede i kapitel 3 (side 17–18). Forskellen skyldes kun afrundingsfejl. Beviset for, at de to teststørrelser er ens, er i princippet bare et spørgsmål om elementær algebra. Men det er umådeligt besværligt, så vi vil ikke gennemføre det her.

Den generelle formel for Pearson's teststørrelse ved test for uafhængighed i en $R \times S$ antalstabel (x_{rs}) er

$$\text{“Pearson”} = \sum_{r=1}^{R}\sum_{s=1}^{S}\frac{\left(x_{rs}-\frac{x_{r\cdot}x_{\cdot s}}{x_{\cdot\cdot}}\right)^2}{\frac{x_{r\cdot}x_{\cdot s}}{x_{\cdot\cdot}}}$$

Opgave 5.5.1. Betragt tabellen

	Tv.auk.	Ikke tv.auk.
–30	139	985
30–40	240	1587
40–50	256	1982
50–	105	1074

som klassificerer 6368 låntagere i BRF Kredit efter låntagers alder og tvangsauktion/ikke tvangsauktion. Udregn fittede værdier under uafhængighedsmodellen, og foretag testet for uafhængighed både ved hjælp af kvotienttestet og Pearson's teststørrelse.

5.6. Sammenligning af polynomialfordelinger.

Den formelle ækvivalens mellem testet til sammenligning af to binomialfordelinger og testet for uafhængighed i en 2×2–tabel kan generaliseres til vilkårlige $R \times S$ antalstabeller i følgende forstand. Betragt følgende to testsituationer, knyttet til en $R \times S$ antalstabel (x_{rs}):

(1) Hele tabellen opfattes som udfaldet af en polynomialfordelt stokastisk variabel (orden RS, antalsparameter $n = x_{..}$). Vi ønsker at teste hypotesen om uafhængighed.

(2) Tabellens rækkesummer opfattes som givne, og rækkernes indhold antages at være fremkommet som værdierne af uafhængige polynomialfordelte variable (orden S, antalsparametre $x_{1\cdot}, \ldots, x_{R\cdot}$) med hvert sit sæt af sandsynlighedsparametre. Vi ønsker at teste hypotesen om at disse R polynomialfordelinger har de samme sandsynlighedsparametre.

Disse to testproblemer viser sig at være ækvivalente, i den forstand at kvotientteststørrelserne bliver ens (og deres frihedsgradsantal ligeså). Et bevis for dette kan føres på samme måde som i tilfældet $R = S = 2$. Sammenhængen mellem parameterestimaterne under hypotesen i modellerne (1) og (2) kan kort beskrives ved, at de fittede værdier (altså de estimerede middelværdier af observationerne) bliver de samme.

Den praktiske konsekvens af dette er, at man ved statistisk analyse af tosidede antalstabeller i visse henseender kan tillade sig at tage let på, om det er hele tabellen der opfattes som polynomialfordelt, svarende til at totalsummen $x_{..}$ er fast, eller det er rækkerne (eller søjlerne) der opfattes som uafhængige polynomialfordelte variable, svarende til at rækkesummerne $x_{r\cdot}$ (eller søjlesummerne $x_{\cdot s}$) opfattes som faste.

5.7. Residualer og normerede residualer.

Pearson's teststørrelse er lig med kvadratsummen $\sum_{r=1}^{R} \sum_{s=1}^{S} u_{rs}^2$ af størrelserne

$$u_{rs} = \frac{x_{rs} - \frac{x_{r\cdot} x_{\cdot s}}{n}}{\sqrt{\frac{x_{r\cdot} x_{\cdot s}}{n}}}.$$

Disse størrelser kaldes de *normerede residualer.* Begrundelsen for denne betegnelse er, at man generelt kalder forskellene mellem observationer og fittede værdier i en model for *residualerne.* Residualerne i uafhængighedsmodellen er altså

$$r_{rs} = x_{rs} - \frac{x_{r\cdot} x_{\cdot s}}{n}.$$

Ved de *normerede* residualer forstår man normalt residualerne divideret med deres estimerede standardafvigelser. Det er nu ikke helt det u_{rs}'erne ovenfor er. Variansen på r_{rs} er ikke er så let at beregne. Men man kan godt indse, at den er noget i nærheden af $\frac{x_{r\cdot} x_{\cdot s}}{n}$. Nærmere kan vi ikke i øjeblikket komme en begrundelse for, at disse størrelser u_{rs} kaldes normerede residualer.

Betydningen af disse størrelser ligger i, at de måler den enkelte observations "afvigelse fra modellen" på en skala, der er sammenlignelig med den skala som defineres ved den normerede normale fordeling; forstået sådan, at normerede residualer større end ca. 2 eller mindre end ca. –2 kan tages som udtryk for en markant afvigelse fra uafhængighedsmodellen.

Dette er først og fremmest interessant hvis vi har forkastet modellen. I så fald er kvadratsummen af de normerede residualer (altså Pearson's teststørrelse) jo for stor til at kunne beskrives som χ^2–fordelt med $(R-1)(S-1)$ frihedsgrader. Ved at se på hvilke normerede residualer der er "skyld" i, at kvadratsummen er blevet for stor, kan man i konkrete tilfælde få en forestilling om, hvordan afvigelsen fra uafhængighedsmodellen kan beskrives.

5.8. Et lidt større eksempel.

Eksempel 5.2.

Tilfredshed:	++	+	–	– –	Sum
Anciennitet:					
0–5	546	488	40	3	1077
6–10	495	790	93	7	1385
11–20	527	880	85	9	1501
Over 20	563	666	55	6	1290
Sum	2131	2824	273	25	5253

Ovenstående tabel er udtrukket fra resultatet af en spørgeskemaundersøgelse vedrørende trivsel på arbejdspladsen, som Falck gennemførte i 1994. 5253 medarbejdere (nemlig dem der havde besvaret begge spørgsmål) er optalt efter deres svar på følgende to spørsgsmål:

Hvor mange år har du været ansat i Falck?
Svarmuligheder: 0–5 år, 6–10 år, 11–20 år, over 20 år.

Hvor tilfreds er du med at være ansat i Falck?
Svarmuligheder: Meget tilfreds, overvejende tilfreds, overvejende utilfreds, meget utilfreds.

Baggrunden for trivselsundersøgelsen var en betydelig omstrukturering af organisationen. Man kunne forestille sig, at det især var de "gamle" medarbejdere, som i meget lang tid havde arbejdet under en bestemt struktur, der var utilfredse på grund af ændringerne. Omvendt kan man sige, at hvis man har været meget længe i Falck, så er det nok fordi man er rimeligt tilfreds med det.

Det første man bør gøre, før man kaster sig ud i en diskussion af forskellene mellem holdningerne i de fire anciennitetsgrupper, er at teste for uafhængighed. Hvis hypotesen om uafhængighed kan godkendes, er der nemlig ikke grundlag for at sige så meget mere. At uafhængighedshypotesen kan godkendes betyder jo, at de tendenser man eventuelt kan få øje på og tillægge en eller anden vidtløftig arbejdssociologisk betydning, lige så godt kan skyldes rene tilfældigheder.

Vi udregner kvotienttestet:

$$
\begin{aligned}
-2\log q = {} & \\
& 2(546\log 546 + \cdots + 6\log 6 \\
& -1077\log 1077 - \cdots - 25\log 25 \\
& +5253\log 5253) = 88.4635.
\end{aligned}
$$

Denne testørrelse skulle under hypotesen om uafhængighed være approksimativt χ^2–fordelt med $(4-1)(4-1) = 9$ frihedsgrader. Da 99.99% fraktilen i denne fordeling er 33.72, kan vi med stor sikkerhed konkludere, at der ikke er uafhængighed i denne tabel.

Vi kan også forsøge med Pearson's teststørrelse. De fittede værdier under uafhængighedshypotesen ser sådan ud:

Tilfredshed:	++	+	−	−−	Sum
Anciennitet:					
0–5	436.9	579.0	56.0	5.1	1077
6–10	561.9	744.6	72.0	6.6	1385
11–20	608.9	806.9	78.0	7.1	1501
Over 20	523.3	693.5	67.0	6.1	1290
Sum	2131	2824	273	25	5253

(med nød og næppe alle sammen ≥ 5) og vi får

$$
\text{“Pearson”} = \frac{(546-436.9)^2}{436.9} + \cdots + \frac{(6-6.1)^2}{6.1} = 88.89;
$$

altså samme konklusion igen: Der er forskel på anciennitetsgruppernes generelle tilfredshed med Falck som arbejdsplads.

Tilbage står så det mest interessante, nemlig at beskrive, hvad denne forskel går ud på. Det kan man blandt andet bruge de normerede residualer til. De ser sådan ud:

Tilfredshed:	++	+	−	−−
Anciennitet:				
0–5	+5.2	−3.8	−2.1	−0.9
6–10	−2.8	+1.7	+2.5	+0.2
11–20	−3.3	+2.6	+0.8	+0.7
Over 20	+1.7	−1.0	−1.5	−0.1

Den mest markante afvigelse fra uafhængighedsmodellen synes at ligge i, at der er uforholdsmæssigt mange af de helt "nye" (anciennitetsgruppen 0–5) der er "meget tilfredse". Herved står de først og fremmest i modsætning til de to midtergrupper, idet de helt "gamle" også har en vis overvægt af "meget tilfredse".

Disse konklusioner kan man også aflæse af følgende tabel, der for hver anciennitetsgruppe angiver fordelingen på de fire svarkategorier i procent:

Tilfredshed:	++	+	−	−−	Sum
Anciennitet:					
0–5	50.7	45.3	3.7	0.3	100.0
6–10	35.7	57.0	6.7	0.5	100.0
11–20	35.1	58.6	5.7	0.6	100.0
Over 20	43.6	51.6	4.3	0.5	100.0

5.9. Test for række– eller søjlehomogenitet i en tosidet tabel.

Antag at vi for en tosidet antalstabel

$$(x_{rs} \mid r = 1, \dots, R,\ s = 1, \dots, S)$$

med R rækker og S søjler har opstillet en sædvanlig polynomialfordelingsmodel, og i denne har fået godkendt hypotesen om uafhængighed. Antag at vi herefter ønsker at teste hypotesen om *rækkehomogenitet,*

$$p_{1\cdot} = \cdots = p_{R\cdot} = \frac{1}{R}.$$

For de enkelte cellers sandsynlighedsparametre betyder dette åbenbart at

$$p_{rs} = p_{r\cdot}p_{\cdot s} = \frac{1}{R}p_{\cdot s},$$

altså at sandsynlighedsparameteren p_{rs} ikke afhænger af rækkenummeret r.

Testet vil ikke blive udledt i detaljer her, da det er intuitivt indlysende og nemt at regne ud, at det kan føres tilbage til et test af den type vi har omtalt som "test for homogenitet" i kapitel 4. Man skal bare teste for homogenitet i den polynomialfordeling, der beskriver rækkesummernes fordeling.

Eksempel 5.3. Antag, at vi i data vedrørende antal fødsler per måned i 1988 (kapitel 4, side 26) havde fået oplyst fødselstallene måned for måned opdelt på drenge– og pigefødsler (altså i form af en MÅNED × KØN tabel, 12×2). I så fald ville det være naturligt at begynde med at teste for uafhængighed. Hypotesen om uafhængighed svarer her til den meget nærliggende formodning, at kønsproportionen for nyfødte er konstant året rundt. Hvis vi kunne få dette godkendt, ville vi kunne fortsætte med test for rækkehomogenitet, svarende til hypotesen om "lige mange fødsler i årets 12 måneder". Dette test er, ifølge ovenstående, identisk med det vi udførte i kapitel 4 på basis af de totale antal fødsler per måned.

Hvis vi i dette tilfælde *ikke* kunne få godkendt hypotesen om uafhængighed, ville der være tale om en noget mere kompliceret situation. Det ville jo betyde, at drenge– og pigefødsler har hver sin årstidsvariation. I så fald ville det måske slet ikke være så interessant om det *samlede* antal fødsler er konstant fra måned til måned.

5.10. Test for uafhængighed i en deltabel.

Betragt tabellen vedrørende Falck–undersøgelsen på side 40. Her er tallene i sidste søjle meget små, og det er derfor fristende at slå de to sidste søjler sammen til én. En betingelse for, at vi kan gøre dette uden at kassere væsentlig information, må være at frekvensen af "meget utilfredse" blandt de "overvejende eller meget utilfredse" er omtrent den samme i de fire anciennitetsgrupper. Hvis dette er tilfældet, kan vi med rimelighed påstå, at der ikke er nogen grund til at skelne mellem disse to holdninger "$-$" og "$--$".

Som hypotese om parametrene i den fulde polynomialfordelingsmodel kan dette formuleres på følgende måde. De otte sandsynlighedsparametre

$$p_{rs},\ r=1,2,3,4,\ s=1,2$$

hørende til de første to søjler gør vi ingen antagelser om. Men om sandsynlighedsparametrene i de to sidste søjler antager vi at der gælder

$$p_{r3}=(1-p_{r1}-p_{r2})(1-\pi),\ p_{r4}=(1-p_{r1}-p_{r2})\pi$$

hvor π betegner den betingede sandsynlighed for at en person (uanset anciennitetsgruppe) er "meget utilfreds", givet at vedkommende er enten "overvejende utilfreds" eller "meget utilfreds".

Sagt på en anden måde, ved hjælp af stokastiske variable R_i og S_i der beskriver person nr. i's anciennitet og holdning: Hypotesen går ud på, at den betingede fordeling af S_i, givet $R_i=r$ og $S_i\in\{3,4\}$, er uafhængig af r. Eller at S_i og R_i er uafhængige i den betingede fordeling, givet $S_i\in\{3,4\}$.

Maksimaliseringsestimatorerne i denne model viser sig at blive som følger. Estimatorerne for sandsynlighedsparametrene hørende til de første to søjler bliver de samme som i den fulde model. Det er ikke så underligt, da hypotesen jo overhovedet ikke lægger nye bånd på disse parametre. Maksimaliseringsestimatoren for π bliver (heller ikke forbavsende)

$$\hat{\pi}=\frac{x_{\cdot 4}}{x_{\cdot 3}+x_{\cdot 4}}=\frac{25}{298}.$$

Af elementære udregninger, som vi også springer over, følger det herefter at kvotientteststørrelsen for hypotesen bliver nøjagtigt den samme som vi ville have fået ved at teste for uafhængighed i den 4×2–tabel, der udgøres af de to sidste søjler. Bidraget fra de to første søjler forkorter ud, da estimatorerne for de tilsvarende sandsynlighedsparametre er ens i de to modeller. Vi får altså

$$-2\log q = 2(40\log 40 + 3\log 3 + \cdots + 6\log 6$$

$$-43\log 43 - \cdots - 25\log 25 + 298\log 298) = 0.706$$

Denne teststørrelse skal vurderes i en χ^2–fordeling med 3 ($=(4-1)(2-1)$) frihedsgrader, hvor den så ganske afgjort ikke er signifikant. Vi kan altså godkende hypotesen om, at fordelingen mellem de to negative svarkategorier ikke afhænger af ancienniteten. Det er derfor tilladeligt at slå de to sidste søjler sammen, og således reducere data til tabellen

Tilfredshed:	++	+	− / − −	Sum
Anciennitet:				
0–5	546	488	43	1077
6–10	495	790	100	1385
11–20	527	880	94	1501
Over 20	563	666	61	1290
Sum	2131	2824	298	5253

På tilsvarende måde kan man for en hvilken som helst antalstabel teste for uafhængighed i en deltabel, der er fremkommet ved at nogle af rækkerne eller nogle af søjlerne er udtaget.

OPGAVE 5.10.1. I tabellen vedrørende Falck undersøgelsen ser det ud som om der ikke er forskel på de to midterste anciennitetsgrupper, når det gælder holdningen til Falck som arbejdsplads. Foretag et test for dette.