

χ^2 -test

Jens Carstensen
2012

Populationer og stikprøver

En gruppe personer eller genstande, som en undersøgelse beskæftiger sig med, kaldes en *population*.

- Hvis man vil undersøge danskernes tv-vaner, er samtlige danskere populationen.
- Hvis man vil gennemføre en undersøgelse af tilfredsheden med togdriften på kystbanen, kan pendlerne på Østsjælland være populationen.
- Hvis man vil undersøge hvilken helbredstilstand de arktiske isbjørne har, er populationen samtlige isbjørne omkring nordpolen
- Hvis man ønsker at danne sig et overblik over udbredelsen af tobaksrygning, kan en population være alle personer mellem 15 og 20 år fordelt over hele landet.
- Hvis man ønsker at foretage en spørgeskemaundersøgelse af holdningen til en betalingsring om København, kan alle bilister på Sjælland være populationen.

Da man normalt ikke kan undersøge alle medlemmer i en hel population, må man nøjes med et vist udvalg. Et sådant udvalg af populationen kaldes en *stikprøve*. Den egenskab, man undersøger, skal tilstræbes at være til stede i stikprøven på samme måde som i hele populationen. Dette udtrykker vi ved at sige, at stikprøven er *repræsentativ*, fordi den i denne forstand repræsenterer hele populationen.

I almindelighed kan man ikke være sikker på, at en stikprøve er repræsentativ – man kan dog prøve at undgå de mere indlysende *systematiske fejl* ved udvælgelsen af stikprøven.

Systematiske fejl

Ved udtagelsen af en stikprøve fra en population kan man begå såkaldt *systematiske fejl*, som gør, at stikprøven ikke er repræsentativ. I eksemplerne oven for kan det være:

- Hvis man vil undersøge tv-vaner for personer over 70 år eller kun for personer med arbejde i landbruget, begår man systematiske fejl.
- Hvis man kun undersøger pendlere under 25 år begår man en systematisk fejl.
- Hvis man i en stikprøve kun udtager grønlandske isbjørne og ikke dem fra Alaska eller Canada, begår man sikkert en systematisk fejl.
- Hvis man kun undersøger tobaksrygning kun hos piger mellem 15 og 20 år, begår man en systematisk fejl.
- Hvis man kun spørger personer over 60 år eller kun personer på Nørrebro, begår man en systematisk fejl.

Skjulte variable

I en stikprøve kan optræde *skjulte variable*, dvs. forhold som kan ødelægge resultaterne, selvom stikprøven tilsyneladende er repræsentativ.

- En undersøgelse af universitetsstuderendes skønnumre viser ved en stikprøve, at de franskstuderende ved Københavns Universitet gennemgående har mindre skønnumre end de ingeniørstuderende på DTU. Får man så store fødder af at læse til ingeniør?

Test af en hypotese

Ordet *hypotese* betyder antagelse eller teori. Vi kan fx kaste med en sædvanlig terning og opstille hypotesen

H : Terningen er ærlig

dvs. der er ikke noget af øjentalene 1, 2, 3, 4, 5 eller 6, der optræder oftere end det 'bør' (ca. $\frac{1}{6}$ af kastene).

Vi kan teste hypotesen ved at kaste terningen fx 60 gange og notere det antal gange, hvert af øjentalene optræder. Hvis vi får følgende fordeling

Øjental	1	2	3	4	5	6
Antal gange	11	9	12	8	10	10

vil vi sikkert godtage hypotesen. Hvis vi imidlertid får fordelingen:

Øjental	1	2	3	4	5	6
Antal gange	4	4	20	16	13	3

vil vi formentlig forkaste hypotesen om, at terningen er ærlig.

Dette ligger i, at den første fordeling af øjental ved 60 kast er forholdsvis *sandsynlig*, mens den anden fordeling er særdeles *usandsynlig*. Vi går ikke her nærmere ind på sandsynlighedsbegrebet.

I statistikken er der tradition for at kalde den hypotese, man vil teste, for *nulhypotesen*:

H_0 : Terningen er ærlig .

Den *alternative hypotese* er

H : Terningen er uærlig .

Signifikansniveau

I et bestemt spil påstås sandsynligheden for gevinst at være 25%, dvs. hvis man spiller mange gange, vil man i gennemsnit vinde 25% af spillene. Vi ønsker at teste nulhypotesen

H_0 : Spillet giver gevinst ca. hver 4. gang.

Vi vil teste hypotesen ved at spille 30 gange. Hvis hypotesen er sand, kan man beregne, at sandsynligheden for at vinde 4 eller færre gange er ca. 10% og at sandsynligheden for at vinde 3 eller færre gange er ca. 3,7%.

Hvis vi vinder 4 gange, vil vi ikke forkaste hypotesen, fordi sandsynligheden for et antal gevinster på 4 eller derunder ikke er lav nok (10% er ikke *meget* usandsynligt). Vinder vi kun 3 gange, vil vi forkaste hypotesen (sandsynligheden 3,7% er så lav, at vi ikke godtager påstanden om gevinst hver 4. gang).

Grænsen for at forkaste nulhypotesen sættes ofte til 5% og man siger, at testens *signifikansniveau* er 5% (eng. *significant*: af betydning). Signifikansniveauet kan også sættes også ofte til 1%.

Krydstabel - test af uafhængighed

En tilfældigt udvalgt gruppe personer får stillet et spørgsmål, de skal besvare med ja eller nej. Man får følgende tabel over svarene (en såkaldt *krydstabel*):

	Ja	Nej	I alt
Mænd			60
Kvinder			90
I alt	100	50	150

Forventede værdier

Hvis mænd og kvinder har samme holdning til spørgsmålet, må man gå ud fra, at den brøkdelen af mændene, der svarer ja (eller nej) og den brøkdelen af kvinderne, der svarer ja (eller nej), er den samme, nemlig begge lig med brøkdelen af ja-svar fra hele gruppen. Denne brøkdelen af ja-svar er

$$\frac{100}{150} = \frac{2}{3}.$$

Det *forventede* antal ja-svar fra mændene er altså

$$\text{antal ja-svar fra mænd} : \frac{2}{3} \cdot 60 = 40.$$

Derefter kan resten af tabellen udfyldes:

$$\text{Antal nej-svar fra mænd} : 60 - 40 = 20$$

$$\text{Antal ja-svar fra kvinder} : 100 - 40 = 60$$

$$\text{Antal nej-svar fra kvinder} : 90 - 60 = 30 \text{ eller } 50 - 20 = 30.$$

Tabellen ser nu sådan ud:

Forventet	Ja	Nej	I alt
Mænd	40	20	60
Kvinder	60	30	90
I alt	100	50	150

Læg mærke til, at disse forventede værdier netop er udregnet

under forudsætning af, at der ikke er forskel på de to køns svar!

Vores nulhypotese, som vi ønsker at teste, er

H_0 : Der er ingen forskel på mænds og kvinders holdning til spørgsmålet.

Teststørrelsen χ^2

Vi ser igen på tabellen over ja- og nej-svar fra mænd og kvinder. Nu viser det sig, at de faktiske svar fordeler sig sådan:

Faktisk	Ja	Nej	I alt
Mænd	46	14	60
Kvinder	54	36	90
I alt	100	50	150

Vi vil udregne et bestemt tal, der angiver, hvor meget mænds og kvinders holdning til spørgsmålet adskiller sig fra hinanden.

Vi udregner for hvert af de fire felter i krydstabellen tallet

$$\frac{(\text{faktisk} - \text{forventet})^2}{\text{forventet}}$$

og derefter lægge de fire tal sammen. χ er det græske bogstav chi (udtales 'ki'). Vi får ved hjælp af tabellerne ovenfor, at

$$\chi^2 = \frac{(46-40)^2}{40} + \frac{(14-20)^2}{20} + \frac{(54-60)^2}{60} + \frac{(36-30)^2}{30} = 4,5 .$$

Tallet 4,5 kaldes χ^2 -teststørrelsen.

Antal frihedsgrader

Ovenstående krydstabel indeholder 2 rækker (mand/kvinde) og 2 søjler (ja/nej). Hvis vi kender et af de 4 tal i tabellen (fx de 46 mandlige ja-sigere), kan vi beregne, hvilke tal, de øvrige tre felter skal indeholde, fordi vi i forvejen kender de totale antal i nederste linje og højre søjle. Da man altså kun frit kan vælge ét tal (og de øvrige tre tal følger af dette), siger man, at *antallet af frihedsgrader* er 1.

Hvis vi foretager en undersøgelse af mænds og kvinders holdning til et bestemt spørgsmål med tre svarmuligheder (ja/nej/ved ikke), kunne vi få en tabel som denne;

	Ja	Nej	Ved ikke	I alt
Mænd	25	47		90
Kvinder				135
I alt	50	95	80	225

Vi kan skrive tal i én rækkes felter (bortset fra et enkelt felt), og derefter er resten af felternes indhold fastlagt, dvs. vi har frihed til at skrive tal i 2 felter i hver række. I tabellen kan vi så udregne de resterende 4 tal. I en krydstabel med 2 rækker og 3 søjler er antallet af

frihedsgrader derfor 2, hvilket vi kan skrive sådan:

$$(2 - 1)(3 - 1) = 1 \cdot 2 = 2 ,$$

I en krydstabel med p rækker og q søjler er antallet af frihedsgrader

$$(p - 1)(q - 1) .$$

p-værdien

Vi vender tilbage til nulhypotesen:

H_0 : Der er ingen forskel på mænds og kvinders holdning til spørgsmålet i eksemplet uden svarmuligheden 'ved ikke'.

Vi vil bruge et signifikansniveau på 5%. Vi fandt, at χ^2 -teststørrelsen var 4,5 og at antallet af frihedsgrader er 1.

Vi ønsker at finde ud af, hvor sandsynligt det er at få en teststørrelse på 4,5 eller derover. Denne sandsynlighed kaldes *p-værdien*. Ved hjælp af cas finder vi i dette tilfælde *p*-værdien til $p = 0,034$.

Sandsynligheden for, at teststørrelsen χ^2 overstiger 4,5 er altså $0,034 = 3,4\%$. Dette regner vi for temmelig usandsynligt (nemlig under 5%), hvis nulhypotesen gælder. Det er derfor, vi forkaster den. Vi kan altså konkludere:

På signifikansniveau 5% er der forskel på mænds og kvinders holdning.

Eksempel. I en virksomhed spørges de ansatte om deres mening om en bestemt politik, som firmaet skal følge. Man har spurgt personalet i produktionen, i administrationen samt de overordnede ledere og fået følgende svar:

Faktisk	Produktionen	Administrationen	Ledere	I alt
For	75	35	12	122
Imod	25	15	8	48
I alt	100	50	20	170

Vi ønsker at teste hypotesen

H_0 : Meningen for/imod afhænger ikke af personalegruppen

på signifikansniveau 5%.

Indtastes disse data i cas fås en teststørrelse på $\chi^2 = 1,959$ og $p = 37,5\%$. Der altså 37,5% sandsynlighed for, at teststørrelsen er 1,959 eller derover. Dette er ikke 'meget usandsynligt' og da p -værdien altså er over 5%, kan vi ikke forkaste hypotesen.

Hvis vi er interesseret, kan vi opstille en tabel over de forventede svar:

Forventet	Produktionen	Administrationen	Ledere	I alt
For	a	b		122
Imod				48
I alt	100	50	20	170

Personalet i produktionen udgør $\frac{100}{170}$ af de adspurgte. Hvis nulhypotesen er korrekt, må det forventede antal 'for'-svar i denne gruppe medarbejdere også være $\frac{100}{170}$ af de i alt 122 'for'-svar, dvs.

$$a = \frac{100}{170} \cdot 122 = 71,76$$

På samme måde udgør det administrative personale $\frac{50}{170}$ af de adspurgte, så det forventede antal 'for'-svar må være samme brøkdelen af samtlige 'for'-svar, så vi får

$$b = \frac{50}{170} \cdot 122 = 35,88$$

Derefter kan resten af tabellen udfyldes:

Forventet	Produktionen	Administrationen	Ledere	I alt
For	71,76	35,88	14,36	122
Imod	28,24	14,12	5,64	48
I alt	100	50	20	170

χ^2 -testen har vist, at disse forventede tal ikke afviger signifikant fra de faktiske. Derfor forkastes nulhypotesen ikke.

Ved brug af cas-værktøj (håndholdt symbolregner eller matematikprogram til pc), får vi de forventede værdier beregnet automatisk. Den hånberegning, vi her har vist, er medtaget for at forklare princippet.

Eksempel. I en undersøgelse så man på personers ægteskabelig stand sammenlignet med deres religiøse tilknytning. Vi vil teste om civilstand (dvs. ægteskabelig stilling) og religiøst tilhørsforhold er uafhængige på et signifikansniveau på 5%.

De faktiske data er følgende:

Faktisk	Katolsk	Protestantisk	Anden	I alt
Gift	135	130	60	325
Fraskilt	40	50	25	115
Enlig	25	20	15	60
I alt	200	200	100	500

Nulhypotesen er

H_0 : Ægteskabelig stand og religiøst tilhørsforhold er uafhængige.

Ved indtastning i cas får vi, at teststørrelsen er $\chi^2 = 3,34$ og $p = 0,503$. Da 50,3% er større end 5%, forkastes nulhypotesen ikke.

Med de data, der er angivet i tabellen kan vi altså ikke påstå, at ægteskabelig stilling og religiøst tilhørsforhold har noget med hinanden at gøre.

De forventede værdier fås ved cas til

Forventet	Katolsk	Protestantisk	Anden	I alt
Gift	130	130	65	325
Fraskilt	46	46	23	115
Enlig	24	24	12	60
I alt	200	200	100	500

Test for fordeling, goodness-of-fit

Vi bruger χ^2 -test også til at teste, om en observeret fordeling af en række observationer afviger fra en forud given fordeling. Da vi undersøger, hvor godt en fordeling passer ('fits') til en teoretisk fordeling, går metoden også under betegnelsen *goodness-of-fit*.

Fremgangsmåden minder meget om metoden ved kryds-tabeller, og vi viser ved et par eksempler hvordan metoden fungerer.

Eksempel. Vi kaster en bestemt terning 60 gange. Vi får følgende antal af 1'ere, 2'ere, 3'ere,....:

Antal øjne	1	2	3	4	5	6
Antal udfald, faktisk	12	9	10	6	11	12

Nulhypotesen er

H_0 : Terningen er ærlig .

Hvis terningen er ærlig, forventer vi hvert øjental i $\frac{1}{6}$ af kastene, så vi får tabellen til venstre over det forventede antal udfald. Vi udregner igen summen af størrelsen for hvert øjental:

Antal øjne	1	2	3	4	5	6
Antal udfald, forventet	10	10	10	10	10	10

$$\frac{(\text{observeret} - \text{forventet})^2}{\text{forventet}},$$

så vi får teststørrelsen χ^2 til

$$\begin{aligned}\chi^2 &= \frac{(12-10)^2}{10} + \frac{(9-10)^2}{10} + \frac{(10-10)^2}{10} + \frac{(6-10)^2}{10} + \frac{(11-10)^2}{10} + \frac{(12-10)^2}{10} \\ &= \frac{4}{10} + \frac{1}{10} + 0 + \frac{16}{10} + \frac{1}{10} + \frac{4}{10} = 2,6.\end{aligned}$$

Antallet af frihedsgrader er $6 - 1 = 5$, fordi vi i tabellen ovenfor med observerede antal øjne har frihed til at udfylde 5 af felterne, hvorefter det sjette felt er fastlagt.

Vi vælger at bruge signifikansniveau 5%. På cas finder vi, at p -værdien er 0,761, så sandsynligheden for at få en teststørrelse, der er 2,6 eller derover, er 76,1%. Dette er langt over 5%, så vi godtager hypotesen om, at terningen er ærlig.

Ved indtastning af de observerede værdier og de forventede værdier udregner cas selv teststørrelsen χ^2 , så vi ikke selv behøver at foretage regningen ovenfor - den er kun vist for at forklare, hvordan tallet 2,6 er fremkommet.

Eksempel. En undersøgelse påstår, at internetbrugere i erhvervslivet sender et antal e-mails om ugen, der fordeler sig sådan:

Antal e-mails, Forventet	Højst 5	6-20	Over 21
Antal procent	30%	40%	30%

I virksomheden *X-tronics* foretager man en undersøgelse af 100 ansattes brug af e-mails og finder tallene i tabellen herunder.

Antal e-mails, Faktisk	Højst 5	6-20	Over 21
Antal ansatte	25%	35%	40%

Nulhypotesen er:

H_0 : *X-tronics'* medarbejderes e-mail-antal afviger ikke signifikant fra undersøgelsen,

og denne vil vi teste på et signifikansniveau på

5%. Vi finder på cas, at $\chi^2 = 4,79$ og p -værdien er 0,091 = 9,1%. Da p -værdien overstiger 5%, godtager vi nulhypotesen. Vi kan altså ikke påstå, at de ansatte i *X-tronics* har et e-mailforbrug, der afviger signifikant fra det almindelige erhvervsliv.

Eksempel. Lad os igen se på 60 terningkast fra eksemplet side 2, hvor antallet af øjne har denne fordeling:

Øjental	1	2	3	4	5	6
Antal gange	4	4	20	16	13	3

Nulhypotesen er

H_0 : Terningen er ærlig,

som vi vil teste på et signifikansniveau på 5%.

Ved hjælp af cas får vi $\chi^2 = 26,6$ og p -værdien er 0,00068 $\approx$ 0,07%. Dette betyder, at sandsynligheden for at få en teststørrelse på 26,6 eller derover, er 0,07%. Dette er yderst usandsynligt (dvs. under 5%), så vi forkaster nulhypotesen og mener altså, at terningen er uærlig.

Kritisk værdi

Vi foretager en χ^2 -test med 5% signifikansniveau, enten på en krydstabel eller på test for fordeling. Vi kan være interesseret i den præcise grænse mellem forkastelse og accept af nulhypotesen, eller med andre ord: Hvilken værdi af teststørrelsen χ^2 giver en p -værdi, der er præcis 5%? Denne værdi kan findes på cas, og vi henviser til afsnittet om symbolregnere.

χ^2 -test på symbolregnere

Vi angiver neden for, hvordan indtastningerne af data på de gængse symbolregnere kan foregå, når du skal foretage χ^2 -test.

TI-nspire CAS (sort model, 2010)

Vi gennemgår de to tilfælde: *krydstabel* og *test for fordeling* (goodness of fit) separat.

Krydstabel

Indstil maskinen til regning. Vi viser, hvordan du indtaster tilfældet med spørgsmålet til mænd og kvinder side 3.

Indtast de faktiske tal

Du indtaster først tabellen rækkevis sådan:

$[[46,14] [54,36]]$.

De kantede parenteser sidder på en tast nederst til venstre på hovedtastaturet. Bemærk, at der fremkommer begge kantede parenteser på en gang: $[[$. Tast to gange, så du nu har $[[]]$. Udfyld første række 46,14 (husk kommaet, der sidder nederst til venstre på en lille grå tast), så du nu har

$[[46,14]]$.

Flyt markøren lidt til højre, så den står mellem de to sidste parenteser og tast igen kantet parentes, så der nu står

$[[46,14] []]$

I den sidste kantede parentes indtaster du 2. række i skemaet:

$[[46,14] [54,36]]$.

Bemærk, at der *ikke* skal sættes komma mellem de kantede parenteser med de to rækkers indhold.

Skemaet er en såkaldt *matrix* (dvs. et talskema), og du giver det et navn, fx m.

$[[46,14] [54,36]] \rightarrow m$.

Pilen fås frem ved hjælp af tasterne (ctrl) (var). Ved (enter) angiver maskinen tabellen, som vi herefter kan kalde frem ved blot at skrive navnet m:

$\begin{bmatrix} 46 & 14 \\ 54 & 36 \end{bmatrix}$.

Hvis krydstabellen indeholder to rækker med hver tre tal, skal indtastningen have følgende format:

$[[x , x , x] [x , x , x]]$.

Foretag testen

Tast (menu), vælg *Statistik*, og *Stat tests...* og χ^2 2-vejstest med (enter). Du får vinduet χ^2 2-vejstest frem.

I feltet *Observeret matrix* taster du bogstavet m, som er navnet på skemaet. Flyt så markøren ned på OK ved hjælp af tasten (tab) og tast (enter). Nu kommer de ønskede resultater frem, nemlig teststørrelsen $\chi^2 = 4,5$ og p -værdien 0,034. Maskinen oplyser desuden, at antallet af frihedsgrader er 1: $df = 1$ (degrees of freedom).

De forventede tal

Du kan få en krydstabel med de forventede tal frem ved at taste (var), flytte markøren ned til *stat.expmatrix* (expected matrix) og taste (enter). Maskinen skriver på linjen *stat.expmatrix* og du taster (enter) og får de forventede værdier:

$$\begin{bmatrix} 40 & 20 \\ 60 & 30 \end{bmatrix}.$$

Kritisk værdi

Den kritiske værdi (niveau) fås ved at taste (menu), vælge *Statistik* og *Fordelinger...* og *Invers χ^2* med (enter). I vinduet *Invers χ^2* taster du i feltet *Areal* værdien 0,95 og i feltet *Frihedsgrad*, df tallet 1. Gå med (tab) til feltet OK og tast (enter). Vi får 3,84. Sandsynligheden for, at χ^2 er over 3,84 er altså 5% og under 3,84 95%.

Test for fordeling (goodness of fit)

Vi viser indtastningen af eksemplet med en virksomheds ansattes brug af email. Indstil regnemaskinen til regning, tast (menu), vælg *Statistik*, og *Stat tests...* og χ^2 GOF med (enter). Du får vinduet χ^2 GOF frem.

I feltet *Observeret liste* indtaster du de observerede data mellem et par krøllede parenteser {}, så der nu står

$$\{25,35,40\}.$$

De enkelte tal er adskilt af komma. Skift til feltet *Forventet liste* med (tab) og indtast

$$\{30,40,30\}.$$

Tasten med krøllede parenteser {} sidder nederst til venstre på tastaturet.

Skift med (tab) til feltet *Frihedsgrader*, df og tast 2 (fordelingen består af 3 tal). Skift til OK med (tab) og tast (enter).

Du får nu teststørrelsen $\chi^2 = 4,79$ og p -værdien $p = 0,091$.

TI-nspire CAS (grå model, 2008)

Hvis du i denne maskine har et program, der er magen til programmet for den sorte maskine for 2010, der er beskrevet oven for, skal du selvfølgelig bruge metoderne oven for.

Vi gennemgår de to tilfælde: *krydstabel* og *test for fordeling* (goodness of fit) separat.

Krydstabel

Indstil maskinen til regning. Vi viser, hvordan du indtaster tilfældet med spørgsmålet til mænd og kvinder side 3.

Indtast de faktiske tal

Du indtaster først tabellen rækkevis sådan:

$[[46,14] [54,36]]$.

De kantede parenteser sidder på tasten mellem de grønne A- og B-tester. Læg mærke til, at der fremkommer begge kantede parenteser på en gang: $[[]]$. Tast to gange, så du nu har $[[[]]]$. Udfyld første række 46,14 (husk kommaet, der sidder nederst til venstre på en lille grå tast), så du nu har

$[[46,14]]$.

Flyt markøren lidt til højre, så den står mellem de to sidste parenteser og tast igen kantet parentes, så der nu står

$[[46,14] []]$

I den sidste kantede parentes indtaster du 2. række i skemaet:

$[[46,14] [54,36]]$.

Bemærk, at der ikke skal sættes komma mellem de kantede parenteser med de to rækkers indhold.

Skemaet er en såkaldt *matrix* (dvs. et talskema), og du giver det et navn, fx m.

$[[46,14] [54,36]] \rightarrow m$.

Pilen fås frem ved hjælp af tastene (ctrl) (var). Ved (enter) angiver maskinen tabellen, som vi herefter kan kalde frem ved blot at skrive navnet m:

$\begin{bmatrix} 46 & 14 \\ 54 & 36 \end{bmatrix}$.

Hvis krydstabellen indeholder to rækker med hver tre tal, skal indtastningen have følgende format:

$[[x , x , x] [x , x , x]]$.

Foretag testen

Tast (menu), vælg *Statistik*, og *Stat tests...* og χ^2 2-vejstest med (enter). Du får vinduet χ^2 2-vejstest frem.

I feltet *Observeret matrix* taster du bogstavet m, som er navnet på skemaet. Flyt så markøren ned på OK ved hjælp af tasten (tab) og tast (enter). Nu kommer de ønskede resultater frem, nemlig teststørrelsen $\chi^2 = 4,5$ og p-værdien 0,034. Maskinen oplyser desuden, at antallet af frihedsgrader er 1: $df = 1$ (degrees of freedom).

De forventede tal

Du kan få en krydstabel med de forventede tal frem ved at taste (var), flytte markøren ned til *stat.expmatrix* (expected matrix) og taste (enter). Maskinen skriver på linjen *stat.expmatrix* og du taster (enter) og får de forventede værdier:

$$\begin{bmatrix} 40 & 20 \\ 60 & 30 \end{bmatrix}.$$

Kritisk værdi

Den kritiske værdi (niveau) fås ved at taste (menu), vælge *Statistik* og *Fordelinger...* og *Invers χ^2* med (enter). I vinduet *Invers χ^2* taster du i feltet *Areal* værdien 0,95 og i feltet *Frihedsgrad, df* tallet 1. Gå med (tab) til feltet OK og tasat (enter). Vi får 3,84. Sandsynligheden for, at χ^2 er over 3,84 er altså 5% og under 3,84 95%.

Test for fordeling (goodness of fit)

Vi viser indtastningen af eksemplet med en virksomheds ansattes brug af email. Indstil regnemaskinen til regning, tast (menu), vælg *Statistik*, og *Stat tests...* og χ^2 GOF med (enter). Du får vinduet χ^2 GOF frem.

I feltet *Observeret liste* indtaster du de observerede data mellem et par krøllede parenteser {}, så der nu står

$$\{25,35,40\}.$$

De enkelte tal er adskilt af komma. Skift til feltet *Forventet liste* med (tab) og indtast

$$\{30,40,30\}.$$

Tasten med krøllede parenteser {} sidder nederst til venstre på tastaturet.

Skift med (tab) til feltet *Frihedsgrader, df* og tast 2 (tabellen indeholder 3 tal i hver række). Skift til OK med (tab) og tast (enter).

Du får nu teststørrelsen $\chi^2 = 4,79$ og p -værdien $p = 0,091$.

TI-89

Vi gennemgår de to tilfælde: *krydstabel* og *test for fordeling* (goodness of fit) separat.

Krydstabel

Vi viser, hvordan du indtaster tilfældet med spørgsmålet til mænd og kvinder side 3.

Indtast de faktiske tal

Du indtaster først tabellen på indtastningslinjen rækkevis sådan:

$$[[46,14] [54,36]] \rightarrow m.$$

De kantede parenteser sidder på tasten med komma og på tasten med $\div$. Læg mærke til, at der er komma mellem tallene i hver række, men ikke komma mellem de to inderste kantede parenteser (rækkerne). Tabellen kaldes m og pilen fås med tasten (STO ►). Tast (enter), så angiver maskinen tabellen sådan:

$$\begin{bmatrix} 46 & 14 \\ 54 & 36 \end{bmatrix} \rightarrow m.$$

Foretag testen

Tast (APPS) og vælg *Stats/List Editor*. Opfordringen *Select Current Folder*, besvarer du med (enter).

Der fremkommer en skærm med fire spalter, hvori man kan skrive tal. Du taster F6, idet der øverst på skærmen står: *F6: Tests* (tast (2nd) (F1)). I den fremkomne menu vælger du med bjælken *Chi2 2-way* og kommer ind i vinduet *Chi-square 2-Way*.

I feltet *Observed Mat* taster du m, fordi skemaet er navngivet med m. I feltet *Store Expected to* angiver du et andet bogstav, fx z. Så bliver de forventede værdier opbevaret i et skema ved navn z, som du senere kan kalde frem.

Tast derefter (enter) og du får de ønskede oplysninger: $\chi^2 = 4,5$ og p -værdien $p = 0,034$. Desuden anføres antallet af frihedsgrader $df = 1$ (degrees of freedom).

De forventede tal

Vend tilbage til hovedskærmen med (home). På indtastningslinjen taster du z og maskinen skriver så de forventede værdier:

$$\begin{bmatrix} 40 & 20 \\ 60 & 30 \end{bmatrix}.$$

Kritisk værdi

Tast (APPS) og vælg *Stats/List Editor*. Du får et skærbillede med 4 kolonner. Vælg *F5 Distr* (dvs. distributions: fordelinger). I menuen vælger du *Inverse* med ► og vælger *Inverse Chi-square* med (enter). I feltet *Area* skriver du 0,95 og i feltet *Deg of Freedom, df* skriver du 1. Tast (enter) (enter) og du får resultatet 3,84.

Sandsynligheden for, at χ^2 er over 3,84 er altså 5% og under 3,84 95%.

Test for fordeling (goodness of fit)

Vi viser indtastningen af eksemplet med en virksomheds ansattes brug af email.

Tast (APPS) og vælg *Stats/List Editor* med (enter). Der fremkommer et skærbillede med titlen *Folder Selection for Statistics Application*. Du taster blot (enter) og får et skærbillede med fire spalter. Vælg *F6 Tests* og i den fremkomne menu flytter du bjælken ned på *Chi2 GOF*, taster (enter) og kommer ind i vinduet *Chi-square Goodness of Fit*.

I feltet *Observed list* indtaster du i krøllede parenteser de faktiske tal:

$$\{25,35,40\},$$

og i feltet *Expected List* taster du de forventede værdier:

$$\{30,40,30\}.$$

I feltet *Deg of Freedom, df* indtaster du antallet af frihedsgrader, som her er 2. Tast derefter (enter) og de ønskede oplysninger fremkommer: $\chi^2 = 4,79$ og p -værdien $p = 0,091$.

OPGAVER

χ^2 -**Opgave 1.** I nedenstående tabel ses resultaterne af samtlige hjerteklapoperationer i 2007-08 ved Odense Universitetshospital (OUH) sammenlignet med resultaterne ved landets øvrige hjertecentre.

- a. Undersøg på 5% signifikansniveau, om dødelighed indenfor 30 dage efter en hjerteklapoperation er ens ved OUH og de øvrige hjertecentre.

	Død indenfor 30 dage efter operationen	Overlevet de første 30 dage
OUH	11	206
Øvrige hjertecentre	32	1374

- b. Hvad ville resultatet være blevet, hvis vi havde valgt et signifikansniveau på 1%? Inddrag det statistiske begreb skjulte variable i din kommentar til resultatet.

χ^2 -**Opgave 2.** I en opinionsundersøgelse i X-købing er 500 personer blevet spurgt, om man er for eller imod de ændrede åbningstider i svømmehallen.

Det oplyses, at man har en hypotese om, at de to køn har samme indstilling.

	For	Imod	Ved ikke	Sum
Kvinder	151	79	34	264
Mænd	115	70	51	236
Sum	266	149	85	500

- a. Udfyld under antagelse af at hypotesen er sand tabellen til højre over den forventede fordeling.

- b. Bestem χ^2 -teststørrelsen for den nævnte hypotese, bestem den tilsvarende p -værdi og kommenter resultatet.

	For	Imod	Ved ikke	Sum
Kvinder				264
Mænd				236
Sum	266	149	85	500

χ^2 -**Opgave 3.** I et forsøg har nogle forskere undersøgt sammenhængen mellem indtagelse af pesticidet Avadex og forekomsten af svulster hos mus. I forsøget indgik i alt 145 mus, hvoraf 25 fik Avadex, mens de sidste 120 mus fungerede som kontrolgruppe. Resultatet blev, at 6 af de mus, der fik Avadex, udviklede svulster og 8 af musene i kontrolgruppen udviklede svulster i den samme periode.

- a. Opstil en 2×2 -tabel over forsøget, og angiv antallet af frihedsgrader.
- b. Undersøg på et 5% signifikansniveau om indtagelse af Avadex har betydning for udvikling af svulster hos mus.

χ^2 -Opgave 4. I en klinisk afprøvning af en bestemt type medicin på en bestemt type alvorlig sygdom har man opdelt de lige behandlingskrævende patienter i to grupper A og B.

Gruppe A får medicinen, mens gruppe B er en kontrolgruppe, som derfor får placebo (fx kalktabeletter). Ingen af grupperne véd, hvilken type behandling, de får.

Efter 6 måneders behandling undersøger man virkningen af behandlingen i hver af de to grupper, og resultatet af denne undersøgelse ses i tabellen.

- Opstil en tabel over de forventede værdier for 'Død' og 'Overlevende' i de to grupper A og B.
- Undersøg, om medicineringen kan siges at have en betydning for patienternes overlevelseschancer, når signifikansniveauet fastlægges til 5%.

Status efter 6 måneder	Gruppe A	Gruppe B	I alt
Død	8	17	25
Overlevende	75	48	123
I alt	83	65	148

χ^2 -Opgave 5. Hjerter- og lungeoperationer kan medføre forstoppelse (postoperativ obstipation). På et hospital er man interesseret i at undersøge, om problemet har samme omfang ved de to operationstyper.

Man opstiller dette som en hypotese, man vil teste, og beder alle hjemsendte hjerter- og lungepatienter svare på følgende spørgsmål:

Har du efter operationen haft forstoppelse i en grad, der har påvirket dine daglige gøremål?

Patienternes svar fremgår af tabellen, hvor tallene i parentes angiver de forventede værdier for patienternes svar.

- Gør rede for, hvordan de forventede værdier er beregnet.
- Bestem χ^2 -teststørrelsen og den tilsvarende p -værdi og kommenter resultatet.

Problemer Operationstype	Ja	Nej	Total
Hjerteroperation	9 (12,97)	51 (47,03)	60
Lungeoperation	15 (11,03)	36 (39,97)	51
Total	24	87	111

χ^2 -Opgave 6. I en undersøgelse på en skole har en gruppe elever bedt 179 elever om at angive køn og rygevaner. En optælling viser, at der er 36 rygere og 143 ikke-rygere blandt de adspurgte. Endvidere viser optællingen, at der er ud af 179 var 78 piger og 101 drenge. Det antages, at det er uafhængigt af køn, hvorvidt en elev ryger eller ej.

	Ryger	Ikke-ryger	
Piger			78
Drenge			101
Sum	36	143	179

Skema 1. Forventet fordeling

a. Bestem under denne antagelse det forventede antal rygere blandt drenge og udfyld de tomme pladser i skema 1 over den forventede fordeling af eleverne på rygevaner og køn.

I skema 2 ses stikprøvens faktiske fordeling af eleverne på køn og rygevaner.

	Ryger	Ikke-ryger	I alt
Piger	21	57	78
Drenge	15	86	101
I alt	36	143	179

Skema 2. Faktisk fordeling

b. Bestem χ^2 -teststørrelsen og den tilsvarende p -værdi.

I en lignende undersøgelse kommer man frem til, at χ^2 -teststørrelsen er 6,34.

c. Giver denne undersøgelse belæg for at hævde, at rygevaner er uafhængige af køn, når der benyttes et signifikansniveau på 5%?

χ^2 -Opgave 7. En skoleklasse vil undersøge, om drikkevaner er uafhængige af køn. Drikkevanerne inddeles i tre grupper: Drikker ofte, drikker en gang imellem, drikker aldrig. Resultatet af elevernes undersøgelser fremgår af tabellen.

	Drikker ofte	Drikker af og til	Drikker ikke	I alt
Piger	18	43	10	71
Drenge	48	175	42	264
I alt	66	218	51	335

a. Giver resultatet belæg for at antage, at drikkevaner er uafhængige af køn, når der vælges et signifikansniveau på 5%? Dit svar skal indeholde en opstilling af en nulhypotese og bestemmelse af den tilsvarende teststørrelse.

b. Undersøgelsen kritiseres for, at inddelingen mellem 'drikker ofte' og 'drikke af og til' er uklar. Opstil selv ud fra tallene en ny tabel over de 335 elever, hvor du kun anvender to inddelingskriterier, og undersøg, om man når til samme konklusion nu, som man gjorde i spørgsmål a.

χ^2 -Opgave 8. For en bestemt storproducent af legetøjsbolde gælder, at diameteren af boldene opfylder målene i tabellen.

Mindre end 20 cm	Mellem 20 cm og 22 cm	Over 22 cm
10%	85%	5%

En forsendelse på 100 legetøjsbolde stoppes i tolden. Fordelingen af diameteren af disse bolde ses i tabellen til højre.

Mindre end 20 cm	Mellem 20 cm og 22 cm	Over 22 cm
28 stk.	160 stk.	12 stk.

a. Gennemfør en hypotesetest til at undersøge om forsendelsen kan stamme fra storproducenten.

χ^2 -Opgave 9. Tabellen viser sammenhørende værdier for to størrelser x og y .

x	1,0	2,1	4,2	6,7	9,3
y	10,2	27,5	34,4	70,0	65,1

I en model antages, at sammenhængen mellem x og y er $y = ax + b$.

a. Bestem tallene a og b ved hjælp af lineær regression, og bestem ved hjælp af modellen de forventede værdier af y .

b. Opstil en tabel over de faktiske og de forventede værdier af y , angiv antallet af frihedsgrader og benyt statistiske metoder til at vurdere, om modellen kan accepteres.

χ^2 -Opgave 10. I et eksperiment krydsede studerende lyserøde blomster med hinanden. Ifølge arvelighedslovene skulle det forventede resultat være 25% røde, 50% lyserøde og 25% hvide. Resultatet af eksperimentet blev 236 blomster, der fordelte sig på farverne som vist i tabellen.

Farve	Rød	Lyserød	Hvid
Hyppighed	66	115	55

a. Beregn ud fra arvelighedslovene de forventede hyppigheder for de tre farver

b. Bestem χ^2 -teststørrelsen for den hypotese, at forsøget er i overensstemmelse med arvelighedslovene. Skriv et forslag til hvilken konklusion den studerende kan skrive i sin rapport, når der anvendes et signifikansniveau på 5%.

χ^2 -Opgave 11. I klageafdelingen i et firma modtager man klager pr. telefon. Firmaet har erfaring for, at klagebehandlingstiden pr. klage fordeles sig som givet i tabellen.

Antal minutter	0-5	5-10	10-15	>15	I alt
Andel af det samlede antal klager	30%	40%	20%	10%	100%

Ved en kontroloptælling i en bestemt måned viste det sig, at klagebehandlingstiden pr. klage fordelte sig som vist i tabellen til højre.

Klagebehandlingstid	0-5	5-10	10-15	>15	I alt
Observeret	37	53	25	5	120

a. Hjælp firmaet med at opstille en nulhypotese, der kan testes. Undersøg dernæst, om man på 5% signifikansniveau kan sige, at mønstret i klagebehandlingstiden har ændret sig.

χ^2 -Opgave 12. Ved folketingsvalget i 2007 fordelte stemmerne sig således:

Parti	S	Rad	Kons	SF	DF	V	EnhLis	LibAll	KrDem
Stemmer i %	25,5	5,1	10,4	13,0	13,8	26,3	2,2	2,8	0,9

Ved en meningsmåling med 968 deltagere foretaget af Greens Analyseinstitut fordelte stemmerne sig således:

Parti	S	Rad	Kons	SF	DF	V	EnhLis	LibAll	KrDem
Stemmer i %	26,3	5,4	9,6	16,9	14,2	24,0	2,1	1,1	0,5

Vi vil anvende meningsmålingen til at undersøge følgende nulhypotese:

Stemmefordelingen er uændret i forhold til valget.

- Gør rede for, hvad der er populationen og hvad der er stikprøven. Opstil en tabel over de observerede antal stemmer og de forventede antal stemmer under nulhypotesen.
- Bestem teststørrelsen, der tester om hypotesen skal accepteres eller forkastes, og afgør på et 5% signifikansniveau, om man kan sige, at stemmefordelingen har ændret sig.
- Det ser ud som om partiet SF har haft betydelig fremgang. Opstil en 2×2 -tabel inddelt efter SF og øvrige partier med observerede og forventede værdier, og afgør med statistiske metoder, om SFs fremgang er signifikant på 5% niveau.