

Værktøjshjælp for TI-Nspire CAS

Struktur for appendiks:

Til hvert af de gennemgåede værktøjer findes der 5 afsnit. De enkelte afsnit kan læses uafhængigt af hinanden. Der forudsættes et elementært kendskab til det pågældende værktøj. Der er mange forskellige måder man kan benytte værktøjerne på – det følgende er kun et forslag – i forbindelse med den faktiske udførelse af undervisningen kan andre metoder sagtens vise sig mere hensigtsmæssige. Af samme grund er det heller ikke nødvendigt at gennemarbejde samtlige afsnit.

Det er valgfrit til såvel den skriftlige eksamen som den mundtlige eksamen om man vil benytte sig af teoretiske metoder eller eksperimentelle metoder. Til den skriftlige eksamen er de indbyggede fordelinger og rutiner et godt udgangspunkt (afsnit 4 og 5). Til den mundtlige eksamen er eksperimentel hypotesetest i forbindelse med et statistisk projekt et godt udgangspunkt (dele af afsnit 2).

Indholdsfortegnelse

| | |
|---|----------------|
| 1) Eksempler på grafisk fremstilling af data | side 1 |
| (Beskrivende statistik – Explorative Data Analysis) | |
| 1a: Uafhængighed | side 1 |
| 1b: Goodness of fit | side 1 |
| 2) Eksperimentel hypotesetest | side 3 |
| 2a: Uafhængighed | side 3 |
| Metode 1: Simulering ud fra omrøring | side 3 |
| Metode 2: Simulering ud fra produktfordeling | side 6 |
| 2b: Goodness of Fit | side 11 |
| 3) Teori: De indbyggede fordelingsfunktioner | side 15 |
| 4) Teoretiske udregninger hørende til hypotesetest | side 18 |
| 4a: Uafhængighed | side 18 |
| 4b: Goodness of Fit | side 19 |
| 5) Indbyggede testrutiner | side 21 |
| 5a: Uafhængighed | side 21 |
| 5b: Goodness of Fit | side 22 |

Følgende TI-Nspire CAS-filer følger med:

Afsnit 1: Eksempler på grafiske fremstillinger.

Afsnit 2: Eksperimentel hypotesetest

Afsnit 3: De indbyggede fordelingsfunktioner

Afsnit 4: Teoretiske udregninger

Afsnit 5: De indbyggede test

1) Eksempler på grafisk fremstilling af data (til brug for den beskrivende statistik – Explorative Data Analysis)

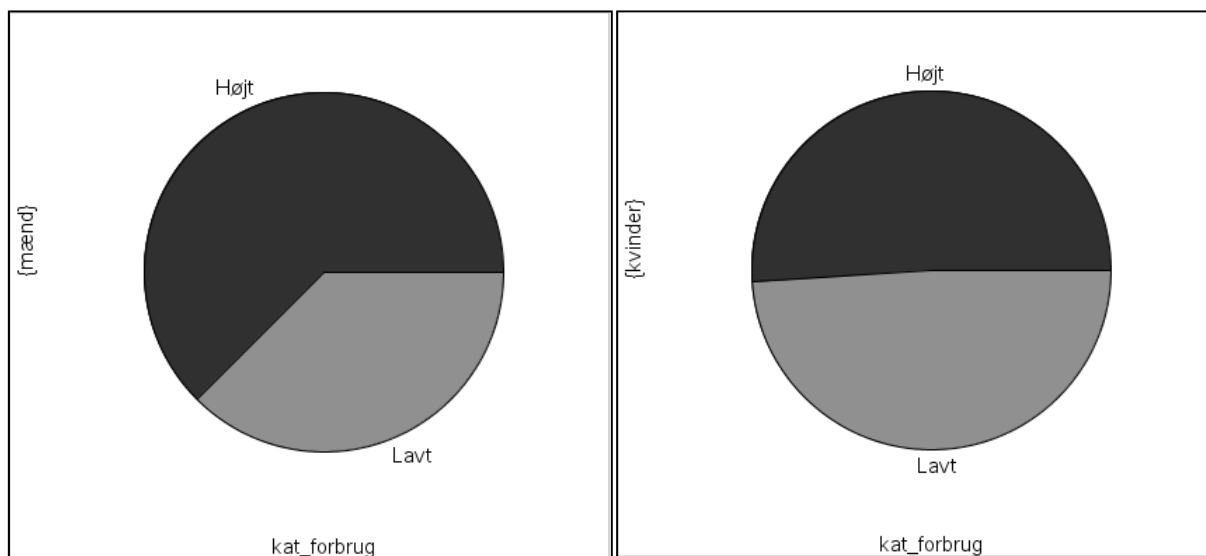
1a: Uafhængighed

Eksempel 1: (side 4 i kursusmaterialet)

| | | | |
|--------------------|----------------------------|-------------------------|--------------|
| <i>Køn\Forbrug</i> | <i>< 1500 kr./måned</i> | <i>≥ 1500 kr./måned</i> | <i>i alt</i> |
| kvinder | 98 | 102 | 200 |
| mænd | 60 | 100 | 160 |
| <i>i alt</i> | 158 | 202 | 360 |

Når vi skal vurdere om der er samme fordeling af forbruget hos henholdsvis mænd og kvinder er det formentligt nemmest at taste data ind i lister og regneark og overføre dem til grafer som frekvensplot, her illustreret som cirkeldiagrammer. Det ses da tydeligt at mænds forbrugsmønster i den pågældende stikprøve ser helt anderledes ud end kvinders forbrugsmønster.

| | A kat_forbrug | B kvinder | C mænd |
|--------|---------------|-----------|--------|
| ♦ | | | |
| 1 Lavt | | 98 | 60 |
| 2 Højt | | 102 | 100 |



1b Goodness of fit

Eksempel 2 (side 24 i kursusmaterialet)

Indkomstfordelingen i stikprøven var: $I = \text{Indkomst i 1000 kr.}$

Observeret antal

| | | | | | | | |
|----------|-------------------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------|
| $I < 50$ | $50 \leq I < 100$ | $100 \leq I < 150$ | $150 \leq I < 200$ | $200 \leq I < 300$ | $300 \leq I < 400$ | $400 \leq I < 500$ | $500 \leq I$ |
| 98 | 88 | 199 | 136 | 210 | 179 | 52 | 38 |

¹ [[Køn\Forbrug,<1500 kr./måned,≥ 1500 kr./måned,i alt][kvinder,98,102,200][mænd,60,100,160][i alt,158,202,360]]

² [[I<50 ,50≤I<100 ,100≤I<150 ,150≤I<200, 200≤I<300, 300≤I<400, 400≤I<500 ,500≤I][98,88,199,136,210,179,52,38]]

Værktøjshjælp til **TI-Nspire CAS**: Eksempler på grafisk fremstilling af data

Den forventede fordeling i stikprøven baseret på de ovenstående procenter er tilsvarende givet ved:

Forventet antal

$$\left[\begin{array}{cccccccc} I < 50 & 50 \leq I < 100 & 100 \leq I < 150 & 150 \leq I < 200 & 200 \leq I < 300 & 300 \leq I < 400 & 400 \leq I < 500 & 500 \leq I \end{array} \right]_3$$

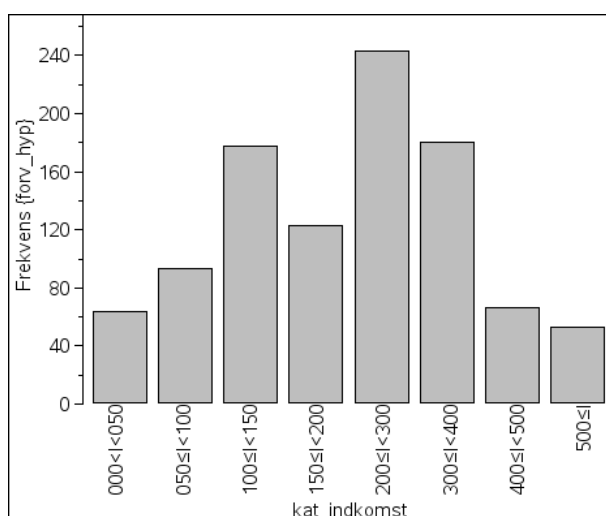
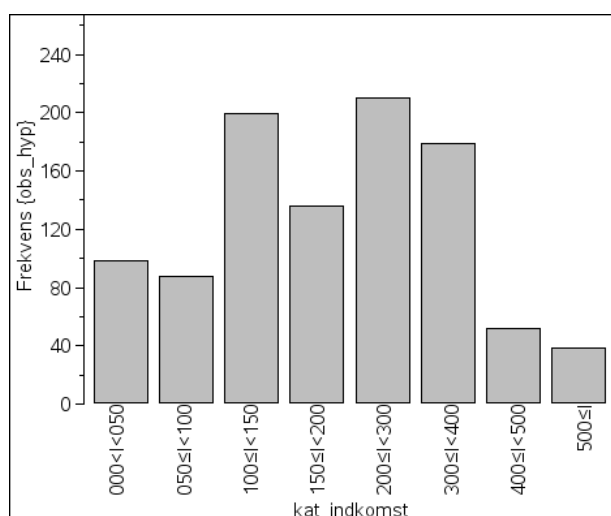
$$\left[\begin{array}{cccccccc} 64 & 93 & 178 & 123 & 243 & 180 & 66 & 53 \end{array} \right]$$

Sammenholder vi de observerede hyppigheder med de forventede følger de så nogenlunde ad.

Men man kunne måske være bekymret for, om de laveste indkomster er overrepræsenteret i stikprøven. Her ligger den observerede hyppighed et godt stykke over den forventede.

Når vi skal vurdere om der er samme fordeling af indkomster i interviewundersøgelsen (stikprøven) og landsgennemsnittet (populationen) er det formentligt nemmest at taste data ind i lister og regneark og overføre dem til grafer som frekvensplot, her illustreret som søjlediagrammer. Det ses da tydeligt at fx fordelingen for de to laveste indkomstgrupper er vendt om i stikprøven i forhold til populationen. Der synes altså at være grund til bekymringen!

| | A kat_indkomst | B obs_hyp | C forv_hyp |
|---|----------------|-----------|------------|
| 1 | 000<I<050 | 98 | 64. |
| 2 | 050≤I<100 | 88 | 93. |
| 3 | 100≤I<150 | 199 | 178. |
| 4 | 150≤I<200 | 136 | 123. |
| 5 | 200≤I<300 | 210 | 243. |
| 6 | 300≤I<400 | 179 | 180. |
| 7 | 400≤I<500 | 52 | 66. |
| 8 | 500≤I | 38 | 53. |



³ [[I<50 ,50≤I<100 ,100≤I<150 ,150≤I<200, 200≤I<300, 300≤I<400, 400≤I<500 ,500≤I][64,93,178,123,243,180,66,53]]

2) Eksperimentel hypotesetest

2a: Uafhængighed

Eksempel 1: (side 4 i kursusmaterialet)

| | | | |
|-------------------------|----------------------------|-------------------------------|-----------------|
| $Køn \setminus Forbrug$ | $< 1500 \text{ kr./måned}$ | $\geq 1500 \text{ kr./måned}$ | $i \text{ alt}$ |
| kvinder | 98 | 102 | 200 |
| mænd | 60 | 100 | 160 |
| $i \text{ alt}$ | 158 | 202 | 360 |

Simulering af nulhypotesen

For at simulere nulhypotesen, der påstår at forbruget er uafhængigt af kønnet, må vi først fastlægge en fortolkning af hvad vi mener med uafhængighed. Det kan gøres på flere måder.

Metode 1: Vi diskuterer først **omrøring**.

Vi konstruerer først en krydsliste for kombinationen af køn og forbrug der er i overensstemmelse med de oplyste hyppigheder. Derefter konstruerer vi to lister for køn og forbrug, der er i overensstemmelse med de oplyste hyppigheder (benyt som vist FrequencyTable►List-kommandoen):

køn:=freqtable►list('kat_køn','obs_hyp)

Vi får også brug for de forventede hyppigheder til udregning af teststørrelsen. De udregnes ved hjælp af celleformlen D1: $\frac{\text{sumif}(\text{kat_køn},a1,\text{obs_hyp}) \cdot \text{sumif}(\text{kat_forbrug},b1,\text{obs_hyp})}{\text{sum}(\text{obs_hyp})} \cdot 1$.

Endelig kan vi udregne teststørrelsen:

chi2_obs:=sum $\left(\frac{(\text{obs_hyp}-\text{forv_hyp})^2}{\text{forv_hyp}} \right) \rightarrow 4.77353$

Så er vi klar til **omrøringen**! I TI-Nspire CAS udføres omrøringen ved hjælp af RandSamp-kommandoen: **omrørt_køn:=randsamp(køn,dim(køn),1)**

hvor slutparameteren 1 angiver at der er tale om en stikprøve UDEN tilbagelæsning. Derved udføres en tilfældig permutation af elementerne i en liste. For at få optalt hyppighederne for omrøringen benyttes celleformlen: **=sum(iffn('omrørt_køn=a1 and 'forbrug=b1,1,0))**

Endelig har vi oprettet et frekvensplot for **kat_mix** og **sim_hyp**, så vi kan visualisere simuleringen, når vi taster CTRL-R i Lister og Regnearksapplikationen! I de observerede hyppigheder ligger KvindeHøjt, KvindeLavt og MændHøjt ca. lige højt. Men i simuleringen er det altid KvindeHøjt, der ligger et stykke over de andre!

| | A kat_køn | B kat_forbrug | C kat_mix | D obs_hyp | E forv_hyp | F sim_hyp |
|---|-----------|---------------|------------|-----------|------------|-----------|
| ♦ | | | | | | |
| 1 | Kvinde | Lavt | KvindeLavt | 98 | 87.7778 | 87 |
| 2 | Kvinde | Højt | KvindeHøjt | 102 | 112.222 | 113 |
| 3 | Mand | Lavt | MandLavt | 60 | 70.2222 | 71 |
| 4 | Mand | Højt | MandHøjt | 100 | 89.7778 | 89 |

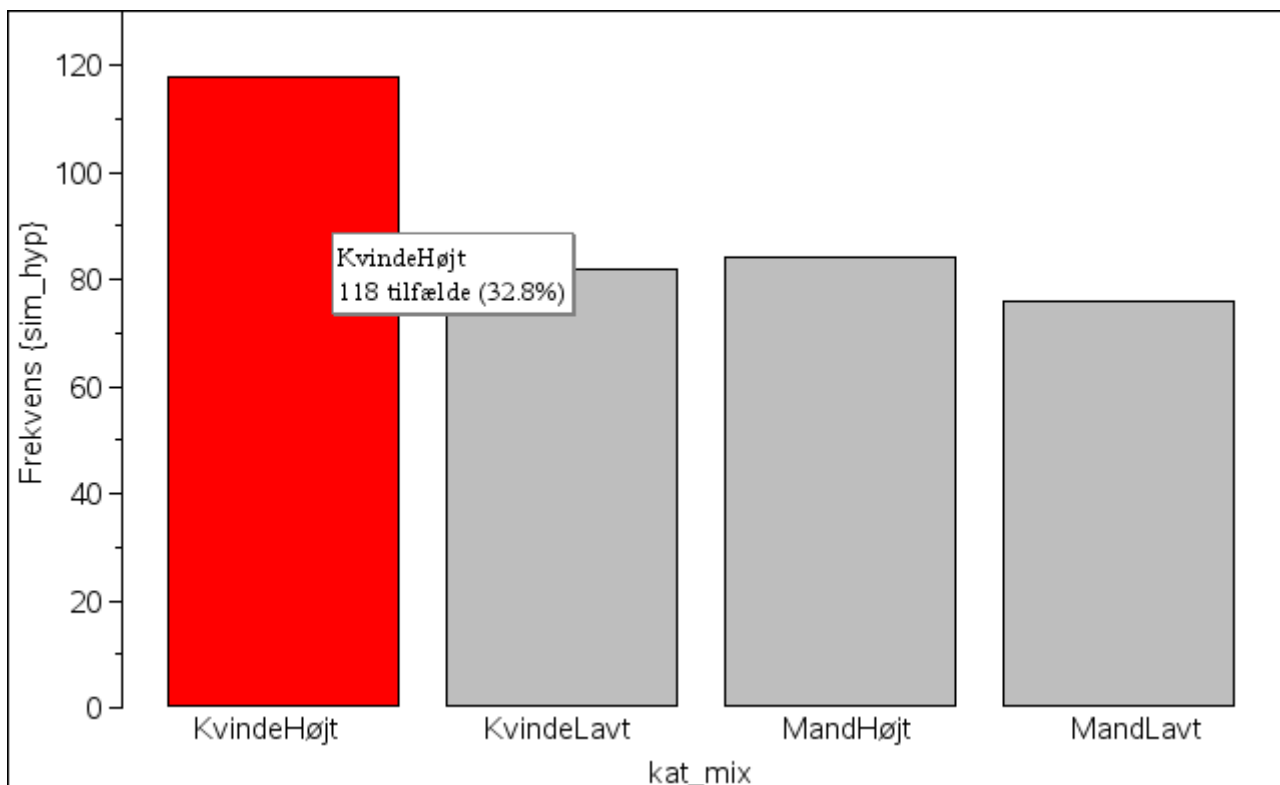
⁴ [[Køn\Forbrug,<1500 kr./måned,≥ 1500 kr./måned,i alt][kvinder,98,102,200][mænd,60,100,160][i alt,158,202,360]]

⁵ ((sumif('kat_køn,a1,'obs_hyp)*sumif('kat_forbrug,b1,'obs_hyp))/(sum('obs_hyp')))*1.

⁶ chi2_obs:=sum((((obs_hyp-forv_hyp)^(2))/(forv_hyp)))

Værktøjshjælp til **TI-Nspire CAS**: Eksperimentel hypotesetest

| | G køn | H forbrug | I omrørt_køn |
|----|----------|------------|---------------------------|
| ♦ | =freqtbl | =freqtable | =randsamp(køn,dim(køn),1) |
| 1 | Kvinde | Lavt | Mand |
| 2 | Kvinde | Lavt | Kvinde |
| 3 | Kvinde | Lavt | Kvinde |
| 4 | Kvinde | Lavt | Mand |
| 5 | Kvinde | Lavt | Kvinde |
| 6 | Kvinde | Lavt | Kvinde |
| 7 | Kvinde | Lavt | Mand |
| 8 | Kvinde | Lavt | Kvinde |
| 9 | Kvinde | Lavt | Mand |
| 10 | Kvinde | Lavt | Kvinde |
| 11 | Kvinde | Lavt | Mand |
| 12 | Kvinde | Lavt | Mand |



Vi skal nu også have udregnet teststørrelsen **Chi2_Sim** for simuleringen, hvilket skal ske i et regneark, så den opdateres løbende:

$$\text{chi2_sim} := \text{sum} \left(\frac{(\text{sim_hyp} - \text{forv_hyp})^2}{\text{forv_hyp}} \right)^7$$

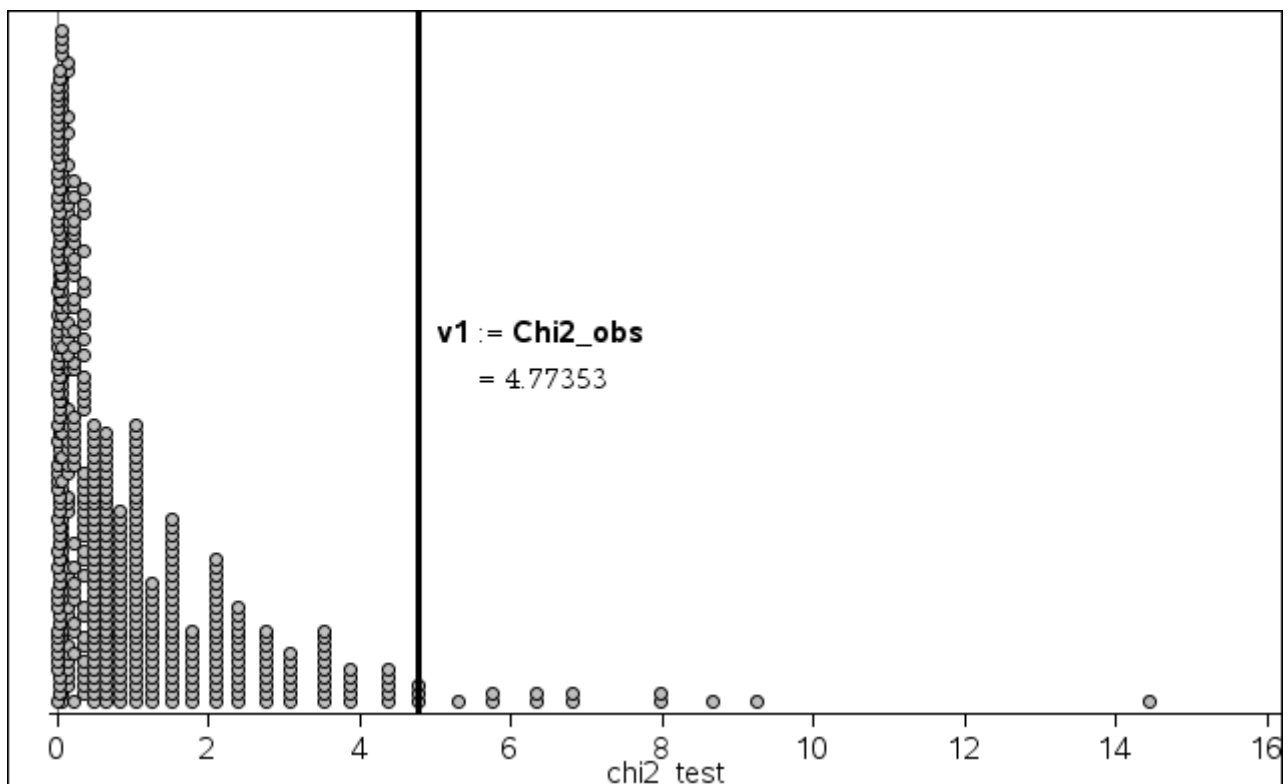
⁷ chi2_sim:=sum((((sim_hyp-'forv_hyp')^(2))/('forv_hyp')))

Værktøjshjælp til **TI-Nspire CAS**: Eksperimentel hypotesetest

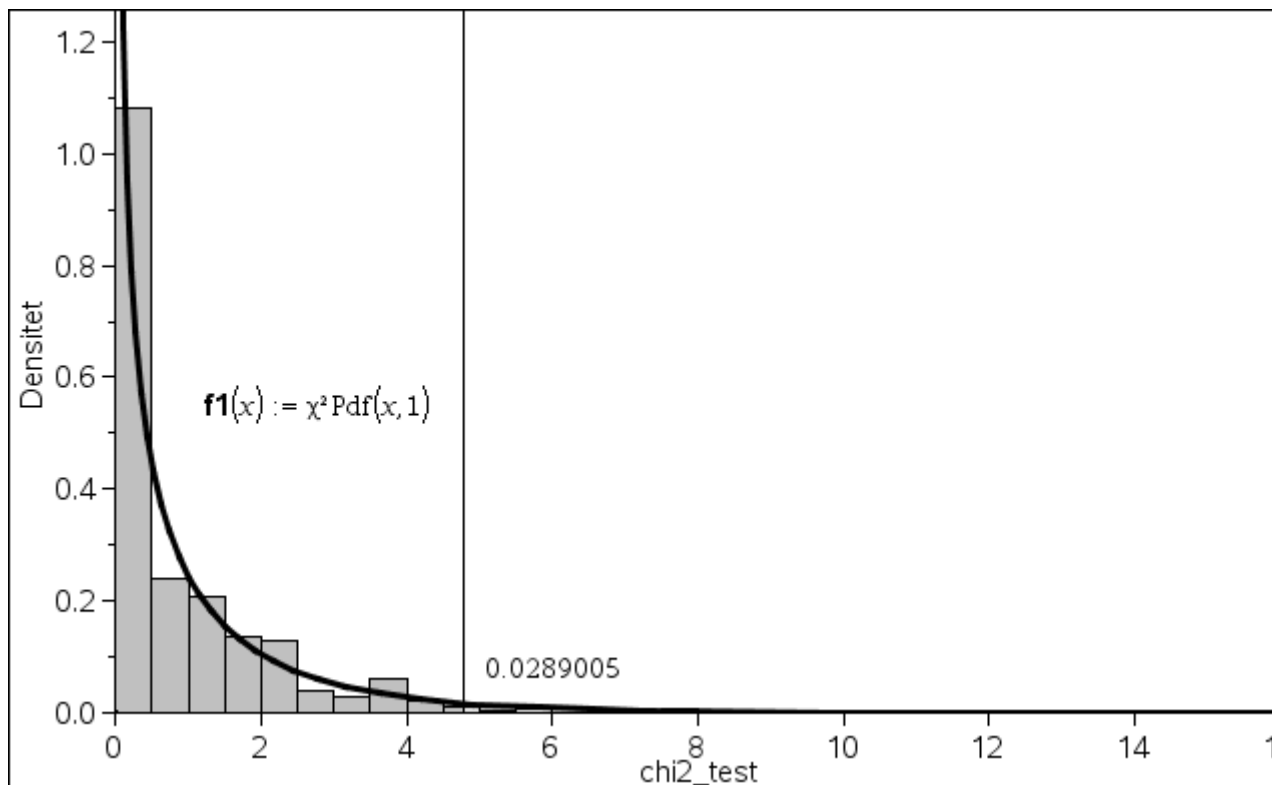
Vi kan så gå på datafangst efter den simulerede teststørrelse. For at holde styr på simuleringen benytter vi manuel datafangst (**CTRL R** efterfulgt af **CTRL .**)

Så skal vi blot have udført simuleringen systematisk rigtig mange gange. Her har vi udført simuleringen 500 gange og set prikdiagrammet blive bygget systematisk op. Prikdiagrammet understreger den grynede natur af chi2-testet. Vi kan da umiddelbart tælle, at der er 15 skæve målinger og dermed estimere p -værdien til 15/500, dvs. ca. 3.0%. Den observerede fordeling er derfor forskellig fra den forventede fordeling på 5%-niveauet.

| A | B | C | D |
|----|---|---------------|----------|
| | | chi2_test | |
| | | =capture(chi2 | |
| 1 | Chi2_sim= | 1.525 | 0.225592 |
| 2 | | | 0.352488 |
| 3 | | | 0.352488 |
| 4 | | | 0.352488 |
| 5 | | | 0.068242 |
| 6 | | | 1.0428 |
| 7 | | | 1.24583 |
| 8 | | | 5.75316 |
| 9 | | | 0.651961 |
| 10 | | | 0.352488 |
| 11 | | | 0.814388 |
| | | | 0.474206 |
| B1 | $\text{chi2_sim} := \text{sum} \left(\frac{(\text{sim_hyp} - \text{forv_hyp})^2}{\text{forv_hyp}} \right)$ | | |



Vi kan også illustrere testfordelingen med et histogram overlejret med den teoretiske fordeling. Som det ses stemmer den empiriske simulerede fordeling og den teoretiske fordeling fint overens!



Metode 2: Denne gang lægger vi os tættere op af sandsynlighedsregningen og udnytter at sandsynlighedsfordelingen for et mix af to uafhængige variable er givet ved produktfordelingen, dvs. vi ganger de respektive sandsynligheder sammen.

Da vi ikke har fået oplyst sandsynlighedsfordelingen for de enkelte variable køn og forbrug, estimerer vi dem ud fra den observerede stikprøve:

| | | | |
|--------------------|----------------------------|-------------------------|--------------|
| <i>Køn\Forbrug</i> | <i>< 1500 kr./måned</i> | <i>≥ 1500 kr./måned</i> | <i>i alt</i> |
| kvinder | 98 | 102 | 200 |
| mænd | 60 | 100 | 160 |
| <i>i alt</i> | 158 | 202 | 360 |

De observerede hyppigheder skrives ind i krydslister og de observerede sandsynligheder for simultanfordelingen udregnes ved hjælp af celleformlen:

$$= \left(\frac{\text{sumif}(\text{/kat_køn}, a1, \text{/obs_hyp})}{\text{sum}(\text{/obs_hyp})} \right) \cdot \left(\frac{\text{sumif}(\text{/kat_forbrug}, b1, \text{/obs_hyp})}{\text{sum}(\text{/obs_hyp})} \right)$$

Hvis nulhypotesen er korrekt er de to variable uafhængige og de tilhørende sandsynligheder for den simultane fordeling er derfor netop givet ved produktfordelingen:

{"KvindeLavt", "KvindeHøjt", "MandLavt", "MandHøjt"}

{0.243827, 0.311728, 0.195062, 0.249383}

⁸ [[Køn\Forbrug,<1500 kr./måned,≥ 1500 kr./måned,i alt][kvinder,98,102,200][mænd,60,100,160][i alt,158,202,360]]

Værktøjshjælp til **TI-Nspire CAS**: Eksperimentel hypotesetest

Her skal vi nu denne gang bruge den kumulerede fordeling til at konstruere stikprøven! Vi trækker derfor 360 tilfældige tal mellem 0 og 1 (**Roulette**) og afgør i hvert enkelt tilfælde, hvor det tilfældige tal falder indenfor den kumulerede fordeling. Derved simulerer vi netop produktfordelingen for de to uafhængige variable (dvs. i det væsentlige nulhypotesen).

Vi skal nu have udregnet teststørrelsen hørende til observationerne:

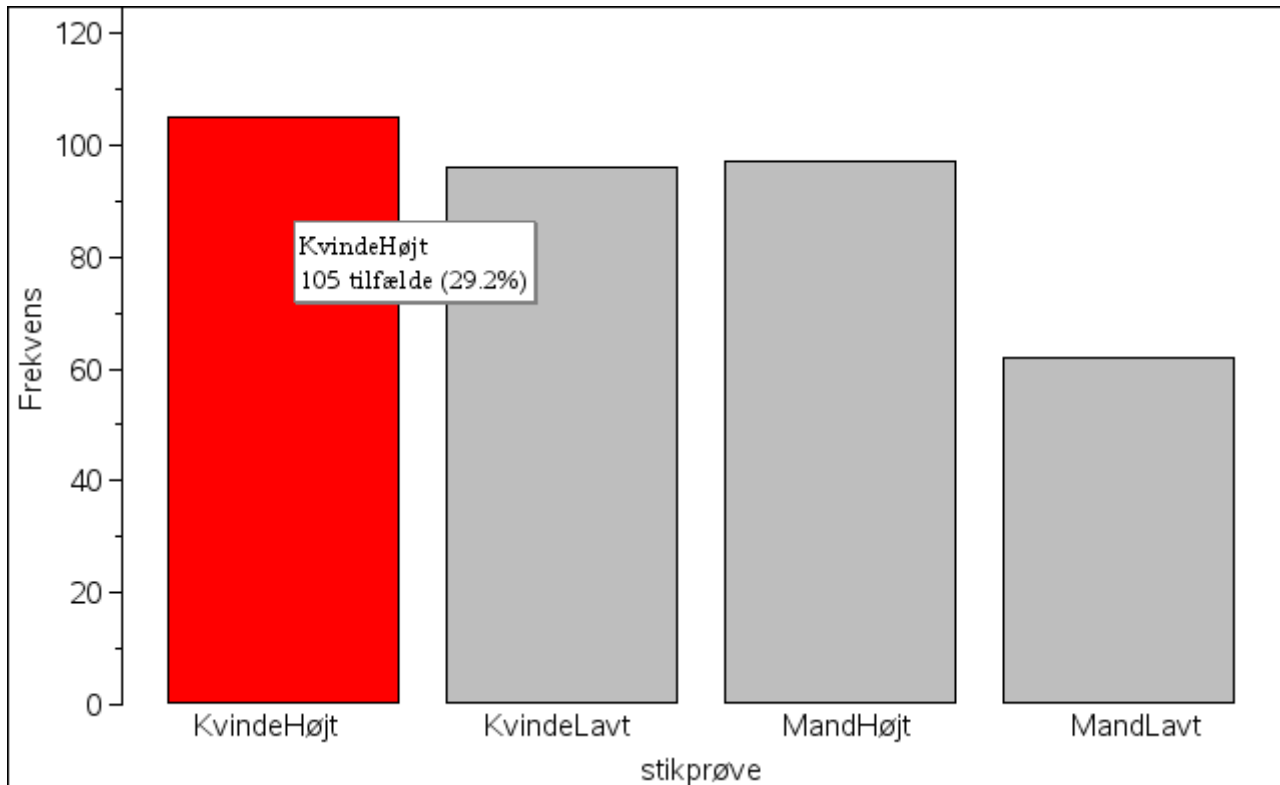
$$\text{chi2_obs} := \text{sum} \left(\frac{(\text{obs_hyp} - \text{forv_hyp})^2}{\text{forv_hyp}} \right)^9 \rightarrow 4.77353$$

Tilsvarende skal vi have udregnet teststørrelsen for simuleringen, men den er mere subtil: VI kan ikke bare bruge de forventede værdier hørende til produktfordelingen, for så er vi jo reelt i gang med at teste om de observerede hyppigheder passer med produktfordelingen, hvilket er en goodness-of-fit test med 3 frihedsgrader for en kendt fordeling. I stedet må vi til hver af de simulerede hyppigheder udregne de tilhørende forventede hyppigheder – ud fra antagelsen om uafhængighed (nulhypotesen). Det var ikke noget problem ved omrøringen, for der holder vi jo fast i marginalhyppighederne. Men denne gang ændres antallet af kvinder, osv. sig i hver simulering. Så nu er der forskel – OG FORSKELLEN ER AFGØRENDE!

| | A kat_køn | B kat_forbrug | C kat_mix | D obs_hyp | E forv_hyp | F sim_hyp | G sim_forv_hyp |
|---|-----------|---------------|------------|-----------|------------|-----------|----------------|
| ♦ | | | | | =360*obs_ | | |
| 1 | Kvinde | Lavt | KvindeLavt | 98. | 87.7778 | 96 | 88.2167 |
| 2 | Kvinde | Højt | KvindeHøjt | 102. | 112.222 | 105 | 112.783 |
| 3 | Mand | Lavt | MandLavt | 60. | 70.2222 | 62 | 69.7833 |
| 4 | Mand | Højt | MandHøjt | 100. | 89.7778 | 97 | 89.2167 |

| | H obs_prob | I kum_prob | J roulette | K stikprøve | L | M |
|----|------------|--------------|------------|---------------|----------------|-----|
| ♦ | | =cumulatives | =rand(360) | =seq(piecewis | | |
| 1 | 0.243827 | 0.243827 | 0.276133 | KvindeHøjt | Antal sim="" | 500 |
| 2 | 0.311728 | 0.555556 | 0.931517 | KvindeHøjt | Antal skæve="" | 11 |
| 3 | 0.195062 | 0.750617 | 0.359599 | KvindeHøjt | | |
| 4 | 0.249383 | 1. | 0.478685 | MandHøjt | | |
| 5 | | | 0.956837 | KvindeHøjt | | |
| 6 | | | 0.763855 | KvindeHøjt | | |
| 7 | | | 0.647918 | KvindeLavt | | |
| 8 | | | 0.179289 | KvindeLavt | | |
| 9 | | | 0.680705 | MandHøjt | | |
| 10 | | | 0.599886 | KvindeLavt | | |
| 11 | | | 0.149443 | MandHøjt | | |
| 12 | | | 0.735549 | KvindeLavt | | |

⁹ chi2_obs:=sum((((obs_hyp-forv_hyp)^(2)))/(forv_hyp))



Vi skal nu også have udregnet teststørrelsen **Chi2_Sim** for simuleringen, hvilket skal ske i et regneark, så den opdateres løbende:

$$\mathbf{chi2_sim} := \text{sum} \left(\frac{(\text{sim_hyp} - \text{sim_forv_hyp})^2}{\text{sim_forv_hyp}} \right)_{10}$$

Vi kan så gå på datafangst efter den simulerede teststørrelse. For at holde styr på simuleringen benytter vi manuel datafangst (**CTRL R** efterfulgt af **CTRL .**)

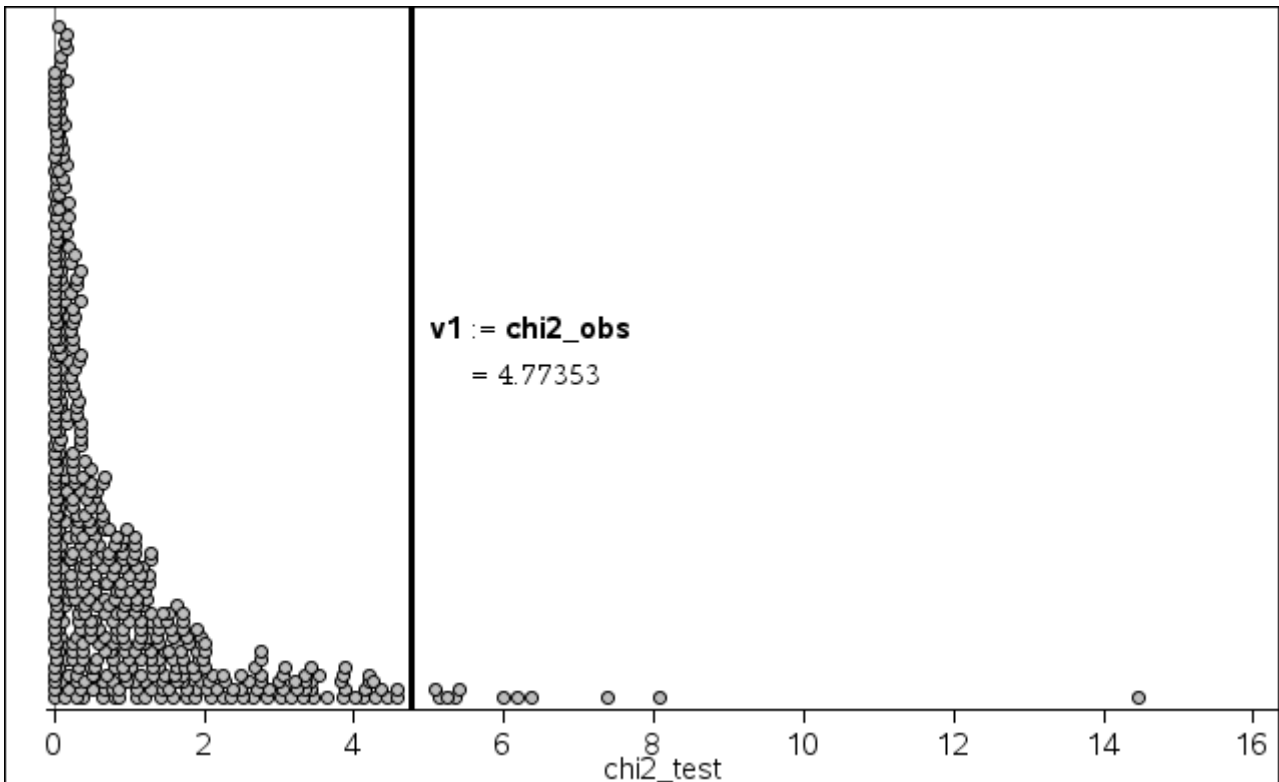
Så skal vi blot have udført simuleringen systematisk rigtigt mange gange. Her har vi udført simuleringen 500 gange og set prikdiagrammet blive bygget systematisk op. Prikdiagrammet understreger den grynedede natur af chi2-testet! Vi kan da umiddelbart tælle, at der er 11 skæve målinger og dermed estimere p -værdien til 11/500, dvs. 2.2%. Den observerede fordeling er derfor forskellig fra den forventede fordeling på 5%-niveauet.

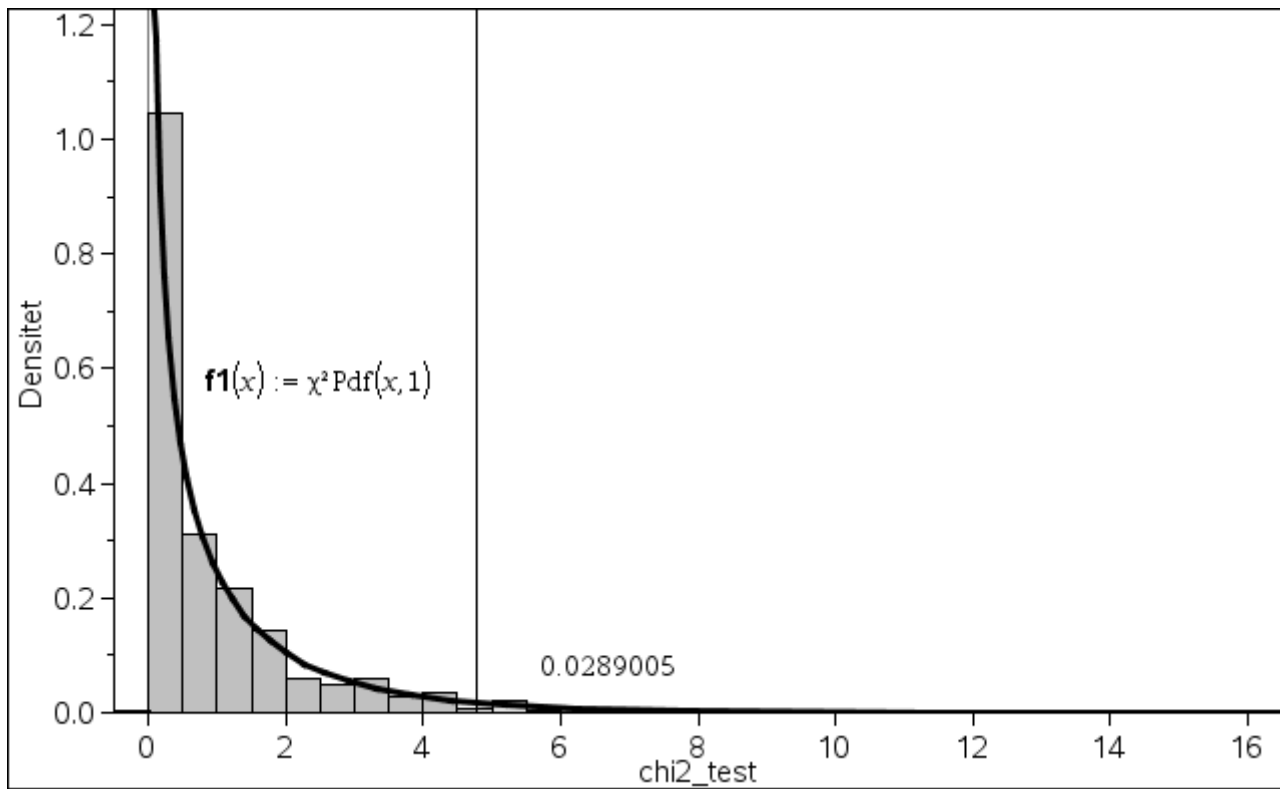
Vi kan også illustrere testfordelingen med et histogram overlejret med den teoretiske fordeling. Som det ses stemmer den empiriske simulerede fordeling og den teoretiske fordeling fint overens!

¹⁰ `chi2_sim:=sum((((sim_hyp-sim_forv_hyp)^2))/(sim_forv_hyp))`

Værktøjshjælp til **TI-Nspire CAS**: Eksperimentel hypotesetest

| A | B | C | D | E | F | G | H |
|----|--|---------------|----------|---|---|---|---|
| | | chi2_test | | | | | |
| | | =capture(chi2 | | | | | |
| 1 | chi2_sim""= | 2.771 | 0.000199 | | | | |
| 2 | | | 0.370094 | | | | |
| 3 | | | 1.58577 | | | | |
| 4 | | | 3.86816 | | | | |
| 5 | | | 0.148087 | | | | |
| 6 | | | 0.196436 | | | | |
| 7 | | | 0.082286 | | | | |
| 8 | | | 0.641366 | | | | |
| 9 | | | 0.000013 | | | | |
| 10 | | | 0.813741 | | | | |
| 11 | | | 0.200048 | | | | |
| 12 | | | 0.674764 | | | | |
| B1 | $\text{chi2_sim} := \text{sum} \left(\frac{(\text{'sim_hyp}' - \text{'sim_forv_hyp})^2}{\text{'sim_forv_hyp'}} \right)$ | | | | | | |





2b: Goodness of Fit

Eksempel 2: (side 24 i kursusmaterialet)

Indkomstfordelingen i stikprøven var: $I = \text{Indkomst i 1000 kr.}$

Observeret antal

$$\left[\begin{array}{cccccccc} I < 50 & 50 \leq I < 100 & 100 \leq I < 150 & 150 \leq I < 200 & 200 \leq I < 300 & 300 \leq I < 400 & 400 \leq I < 500 & 500 \leq I \end{array} \right]^{11}$$

$$\left[\begin{array}{cccccccc} 98 & 88 & 199 & 136 & 210 & 179 & 52 & 38 \end{array} \right]$$

Den forventede fordeling i stikprøven baseret på de ovenstående procenter er tilsvarende givet ved:

Forventet antal

$$\left[\begin{array}{cccccccc} I < 50 & 50 \leq I < 100 & 100 \leq I < 150 & 150 \leq I < 200 & 200 \leq I < 300 & 300 \leq I < 400 & 400 \leq I < 500 & 500 \leq I \end{array} \right]^{12}$$

$$\left[\begin{array}{cccccccc} 64 & 93 & 178 & 123 & 243 & 180 & 66 & 53 \end{array} \right]$$

Sammenholder vi de observerede hyppigheder med de forventede følger de så nogenlunde ad. Men man kunne måske være bekymret for, om de laveste indkomster er overrepræsenteret i stikprøven. Her ligger den observerede hyppighed et godt stykke over den forventede.

Løsning: Vi skal have simuleret nulhypotesen og benytter derfor RandSamp-kommandoen til at udtrække en stikprøve fra en **ideel population**, der repræsenterer den forventede fordeling, sådan som den fremgår af tallene fra Danmarks statistik. Vi indskriver derfor lister med indkomst-kategorier, de observerede hyppigheder og de forventede hyppigheder. Den ideelle population konstrueres derefter med kommandoen:

ideel:=freqtable▶list(kat_indkomst,forv_hyp).

Da stikprøven også består af 1000 individer skal vi nu have trukket 1000 individer fra populationen MED tilbagelægning, så hver indkomstkategori hver gang har samme sandsynlighed for at blive udtrukket! Vi kan gentage stikprøven ved at taste **CTRL R** (mens vi har aktiveret lister og regneark vinduet).

De simulerede hyppigheder kan så findes ved hjælp af celleformlen =countif(**stikprøve,a1**).

Med dette på plads kan vi nu både udregne den observerede teststørrelse og den simulerede teststørrelse:

$$\text{chi2_obs} := \text{sum} \left(\frac{(\text{obs_hyp} - \text{forv_hyp})^2}{\text{forv_hyp}} \right)^{13} \rightarrow 33.8848$$

¹¹ [[I<50 ,50≤I<100 ,100≤I<150 ,150≤I<200, 200≤I<300, 300≤I<400, 400≤I<500 ,500≤I][98,88,199,136,210,179,52,38]]

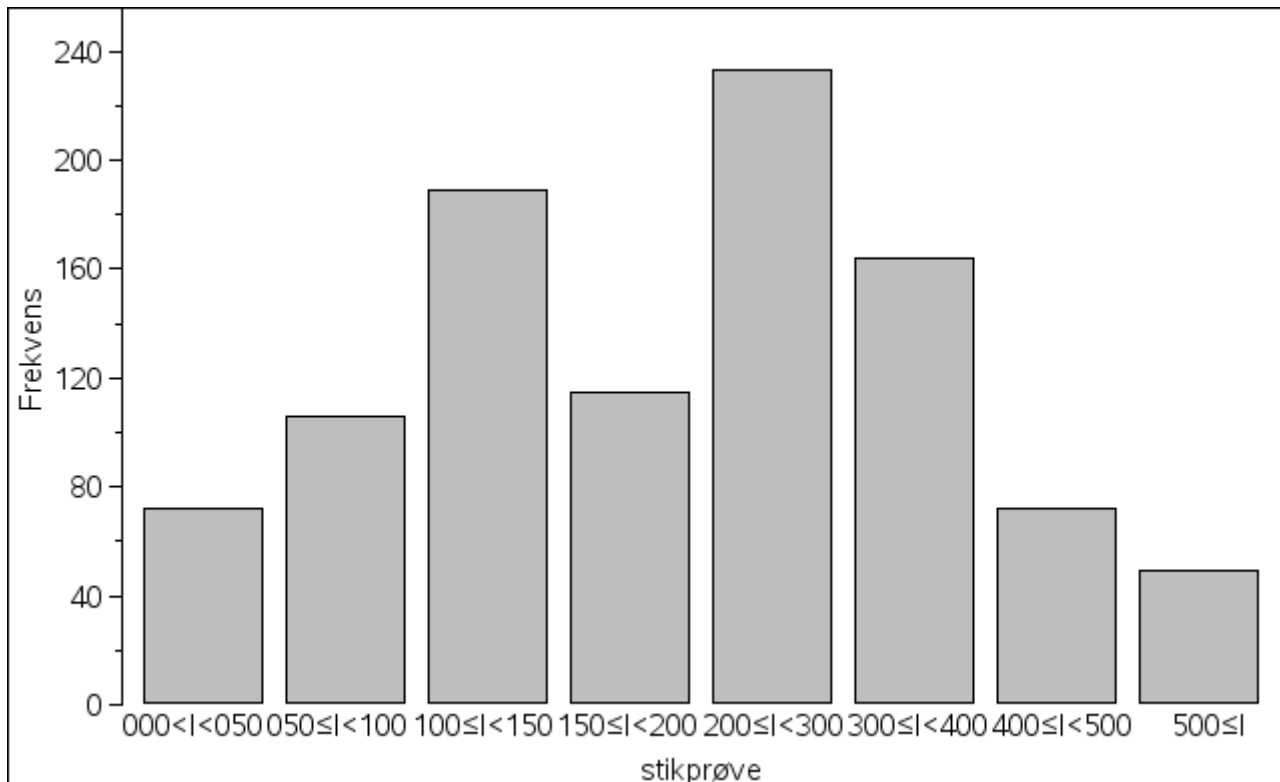
¹² [[I<50 ,50≤I<100 ,100≤I<150 ,150≤I<200, 200≤I<300, 300≤I<400, 400≤I<500 ,500≤I][64,93,178,123,243,180,66,53]]

¹³ chi2_obs:=sum((((obs_hyp-forv_hyp)^(2))/(forv_hyp)))

Værktøjshjælp til **TI-Nspire CAS**: Eksperimentel hypotesetest

| | A kat_indkomst | B obs_hyp | C forv_hyp | D sim_hyp |
|---|----------------|-----------|------------|-----------|
| ♦ | | | | |
| 1 | 000<I<050 | 98 | 64. | 72 |
| 2 | 050≤I<100 | 88 | 93. | 106 |
| 3 | 100≤I<150 | 199 | 178. | 189 |
| 4 | 150≤I<200 | 136 | 123. | 115 |
| 5 | 200≤I<300 | 210 | 243. | 233 |
| 6 | 300≤I<400 | 179 | 180. | 164 |
| 7 | 400≤I<500 | 52 | 66. | 72 |
| 8 | 500≤I | 38 | 53. | 49 |

| | E ideel | F stikprøve | G | H |
|---|-------------|----------------|--------------|-----|
| ♦ | =freqtable▶ | lis=randsamp(i | | |
| 1 | 000<I<050 | 100≤I<150 | Antal sim= | 500 |
| 2 | 000<I<050 | 100≤I<150 | Antal skæve= | 1 |
| 3 | 000<I<050 | 150≤I<200 | | |
| 4 | 000<I<050 | 100≤I<150 | | |
| 5 | 000<I<050 | 000<I<050 | | |
| 6 | 000<I<050 | 500≤I | | |
| 7 | 000<I<050 | 100≤I<150 | | |
| 8 | 000<I<050 | 000<I<050 | | |
| 9 | 000<I<050 | 000<I<050 | | |



Værktøjshjælp til **TI-Nspire CAS**: Eksperimentel hypotesetest

Den simulerede teststørrelse skal udregnes i et regneark, da den skal være dynamisk og opdateres løbende:

$$\text{chi2_sim} := \text{sum} \left(\frac{(\text{sim_hyp} - \text{forv_hyp})^2}{\text{forv_hyp}} \right)^{14}$$

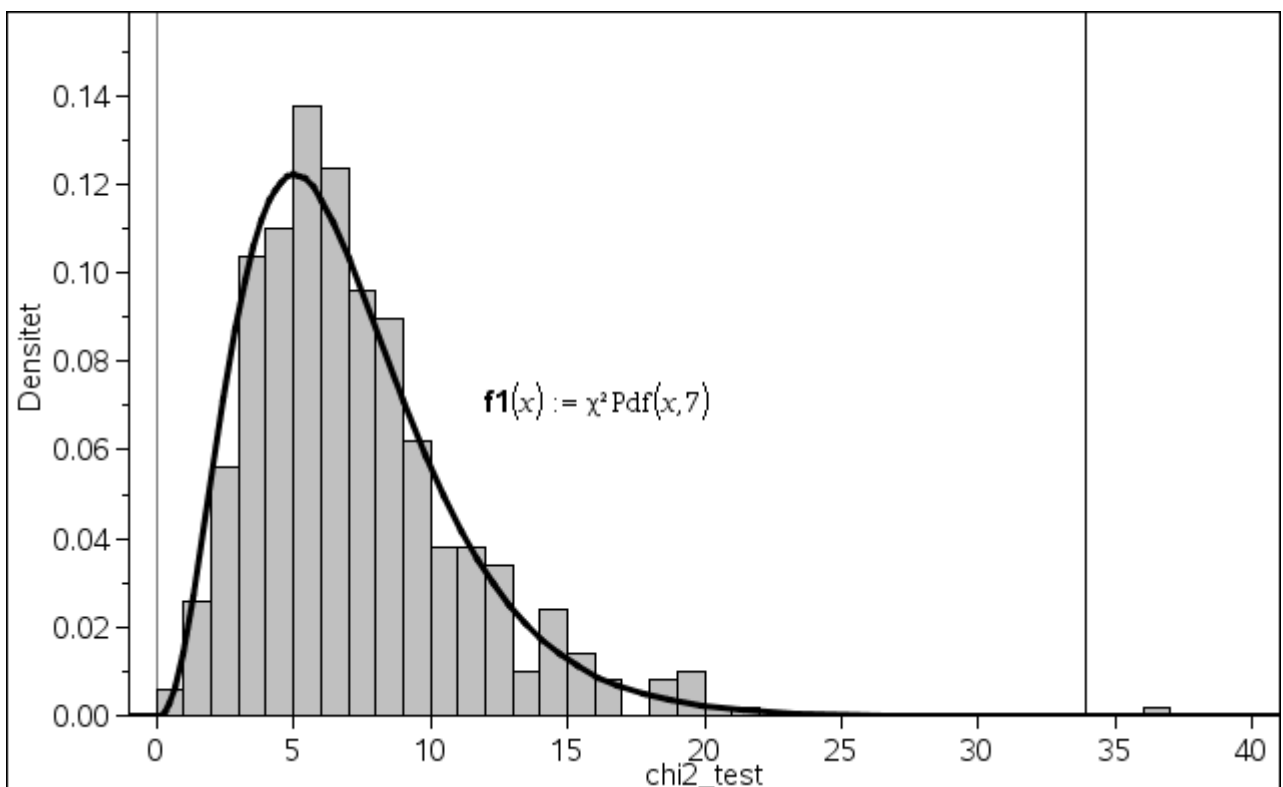
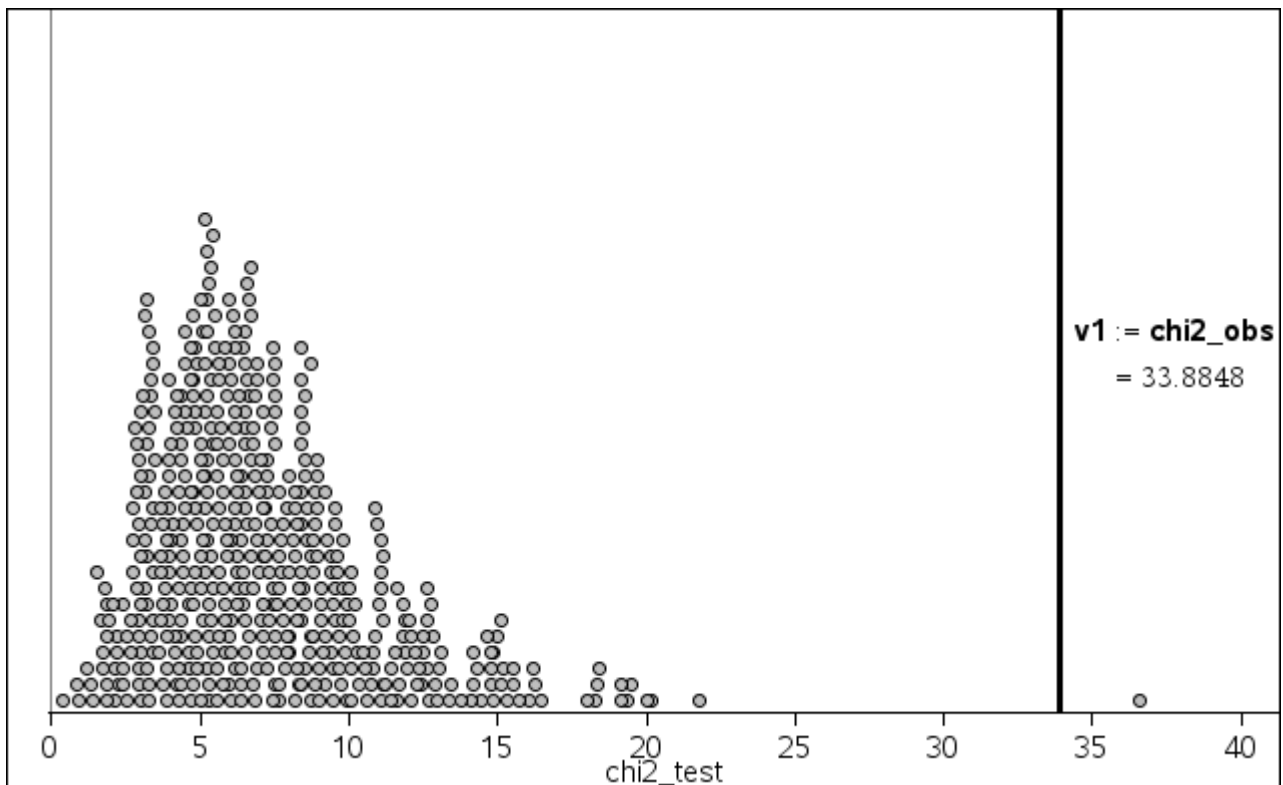
Vi kan så gå på datafangst efter den simulerede teststørrelse. For at holde styr på simuleringen benytter vi manuel datafangst (**CTRL R** efterfulgt af **CTRL .**)

Så skal vi blot have udført simuleringen systematisk rigtigt mange gange. Her har vi udført simuleringen 500 gange og set prikdiagrammet blive bygget systematisk op. Vi kan da umiddelbart tælle, at der er 1 skæv måling (det er faktisk uhyre sjældent man fanger en skæv i dette tilfælde!) og dermed estimere p -værdien til $1/500$, dvs. ca. 0.2%. Den observerede fordeling er derfor forskellig fra den forventede fordeling på 5%-niveauet.

Vi kan også illustrere testfordelingen med et histogram overlejret med den teoretiske fordeling. Som det ses stemmer den empiriske simulerede fordeling og den teoretiske fordeling rimeligt overens!

| | A | B | C chi2_test |
|----|-----------|---------|----------------------|
| ◆ | | | =capture(chi2_sim,0) |
| 1 | chi2_sim= | 6.69839 | 5.30278 |
| 2 | | | 8.72983 |
| 3 | | | 4.70912 |
| 4 | | | 12.1087 |
| 5 | | | 3.00772 |
| 6 | | | 7.61069 |
| 7 | | | 5.53781 |
| 8 | | | 4.28517 |
| 9 | | | 2.91595 |
| 10 | | | 11.0414 |
| 11 | | | 4.17446 |
| 12 | | | 11.6282 |

¹⁴ $\text{chi2_sim} := \text{sum}(\frac{(\text{sim_hyp} - \text{forv_hyp})^2}{\text{forv_hyp}})$



3) Teori: De indbyggede fordelingsfunktioner

De indbyggede fordelingsfunktioner:

Chi-kvadrat (χ^2) fordelingen hedder chi2. Når man skal arbejde med chi-kvadratfordelingen kan man benytte de følgende operatører:

$$y = \text{chi2Pdf}(x,df): \quad (\text{Point Distribution Function})$$

$$p = \text{chi2Cdf}(x\text{lav},x\text{høj},df) \quad (\text{Cumulative Distribution Function})$$

$$x = \text{invChi2}(p,df) \quad (\text{den inverse fordeling})$$

Vi ser først på **tæthedsfunktionen**:

$$\text{chi2Pdf}(x,df) \rightarrow \chi^2\text{Pdf}(x,df)$$

Som det ses kan vi ikke uden videre få oplyst forskriften. Vil man arbejde med forskriften skal den derfor indføres som en brugerdefineret funktion:

$$\text{chi2tæt}(x,df) := \begin{cases} 0, & x < 0 \\ \frac{x^{\frac{df}{2}-1} \cdot e^{-\frac{x}{2}}}{\int_0^\infty (x^{\frac{df}{2}-1} \cdot e^{-\frac{x}{2}}) dx}, & x \geq 0 \end{cases} \xrightarrow{15} \text{chi2tæt}(x,df) := \begin{cases} 0, & x < 0 \\ \frac{x^{\frac{df}{2}-1} \cdot e^{-\frac{x}{2}}}{\int_0^\infty (x^{\frac{df}{2}-1} \cdot e^{-\frac{x}{2}}) dx}, & x \geq 0 \end{cases}$$

Det kan godt se lidt uoverskueligt ud, men for konkrete værdier af antallet af frihedsgrader forenkles udtrykket – ikke mindst for de lige frihedsgrader:

$$\text{chi2tæt}(x,4) \rightarrow \begin{cases} 0, & x < 0 \\ \frac{x \cdot e^{-\frac{x}{2}}}{4}, & x \geq 0 \end{cases}$$

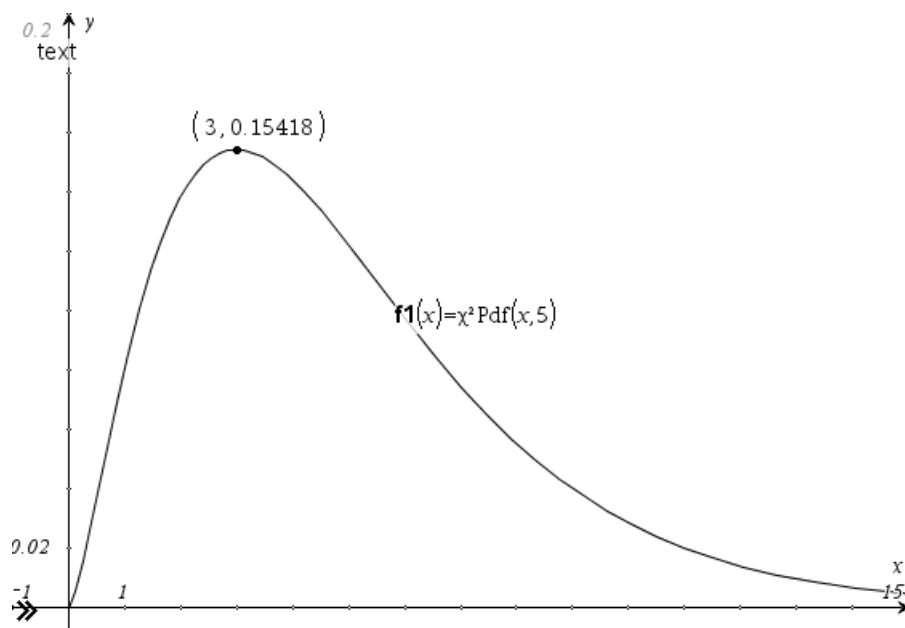
$$\text{chi2tæt}(x,5) \rightarrow \begin{cases} 0, & x < 0 \\ \frac{\frac{3}{2} \cdot x \cdot e^{-\frac{x}{2}}}{7.51988}, & x \geq 0 \end{cases}$$

Her har vi successivt forenklet udtrykkene ved at markere deludtryk og evalueret dem! For de ulige frihedsgrader fører integrationen kun til en numerisk approksimation.

Det er også nemt at afbilde tæthedsfunktionen (som har maksimum i $x = df - 2$, dvs. i dette tilfælde 3):

¹⁵ $\text{chi2tæt}(x,df) := \text{piecewise}(0,x < 0, ((x^{((df)/(2)-1)} \cdot e^{((-x)/(2))}) / (\int_0^\infty ((x^{((df)/(2)-1)} \cdot e^{((-x)/(2))}), x, 0, \infty))), x \geq 0)$

Værktøjshjælp til TI-Nspire CAS: De indbyggede fordelingsfunktioner



Vi ser dernæst på den **kumulerede fordeling**:

$$\text{chi2Cdf}(x_{\text{lav}}, x_{\text{høj}}, df) \rightarrow \chi^2 \text{Cdf}(x_{\text{lav}}, x_{\text{høj}}, df)$$

Som det ses kan vi igen ikke få oplyst forskriften. Vil man arbejde med forskriften skal den derfor indføres som en brugerdefineret funktion:

$$\text{chi2kumuleret}(x_{\text{lav}}, x_{\text{høj}}, df) := \frac{\int_{x_{\text{lav}}}^{x_{\text{høj}}} \left(x^{\frac{df}{2}-1} \cdot e^{-\frac{x}{2}} \right) dx}{\int_0^{\infty} \left(x^{\frac{df}{2}-1} \cdot e^{-\frac{x}{2}} \right) dx} \quad 16 \rightarrow$$

$$\text{chi2kumuleret}(x_{\text{lav}}, x_{\text{høj}}, df) := \frac{\int_{x_{\text{lav}}}^{x_{\text{høj}}} \left(x^{\frac{df}{2}-1} \cdot e^{-\frac{x}{2}} \right) dx}{\int_0^{\infty} \left(x^{\frac{df}{2}-1} \cdot e^{-\frac{x}{2}} \right) dx}$$

Igen forenkles det betydeligt for et konkret antal frihedsgrader, ikke mindst for de lige frihedsgrader:

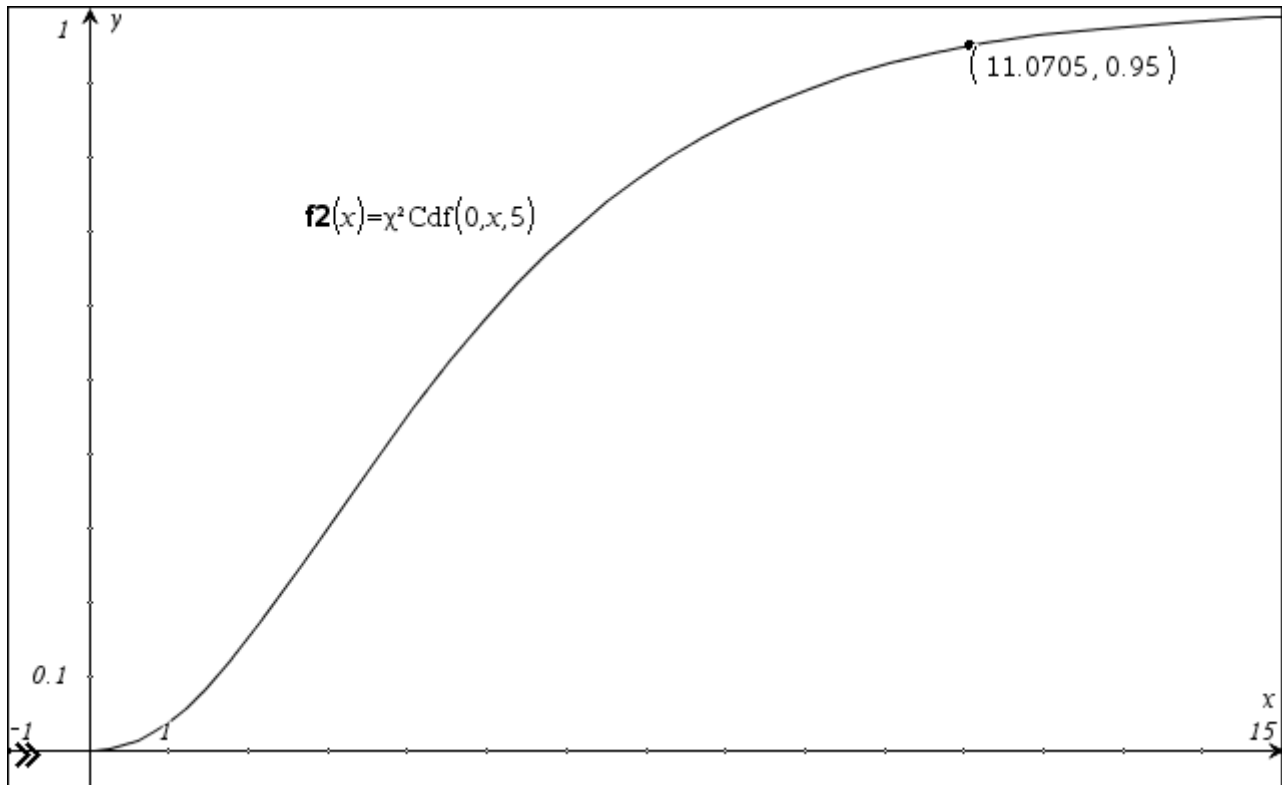
$$\text{chi2kumuleret}(x_{\text{lav}}, x_{\text{høj}}, 4) \rightarrow \frac{\left(e^{\frac{x_{\text{høj}}}{2}} \cdot (x_{\text{lav}}+2) - (x_{\text{høj}}+2) \cdot e^{\frac{x_{\text{lav}}}{2}} \right) \cdot e^{-\frac{x_{\text{høj}}}{2}} \cdot \frac{x_{\text{lav}}}{2}}{2}$$

$$\text{chi2kumuleret}(0, x_{\text{høj}}, 4) \rightarrow \frac{e^{-\frac{x_{\text{høj}}}{2}} \cdot \left(2 \cdot e^{\frac{x_{\text{høj}}}{2}} - x_{\text{høj}} - 2 \right)}{2} \rightarrow \frac{-x_{\text{høj}}}{2 \cdot \sqrt{e^{x_{\text{høj}}}}} - \frac{1}{\sqrt{e^{x_{\text{høj}}}}} + 1$$

(hvor det sidste udtryk er fremkommet ved at anvende en expand-kommando)

$$\text{chi2kumuleret}(x_{\text{lav}}, x_{\text{høj}}, 5) \rightarrow 0.132981 \cdot \int_{x_{\text{lav}}}^{x_{\text{høj}}} \left(x^{\frac{3}{2}} \cdot e^{-\frac{x}{2}} \right) dx$$

¹⁶ ((∫(x^(((df)/(2))-1)*e^((-x)/(2))),x,x_lav,x_høj))/(∫(x^(((df)/(2))-1)*e^((-x)/(2))),x,0,∞))



Endelig kan man finde **fraktilerne** (den inverse kumulerede fordeling).

$$\text{invChi2}(0.95, 5) \rightarrow 11.0705$$

Vi ser altså at 95% af observationerne ligger under 11.0705, hvis vi har en stokastisk variabel, der er chi-kvadrat fordelt med 5 frihedsgrader:

$$\text{chi2Cdf}(0, 11.0705, 5) \rightarrow 0.95$$

4) Teoretiske udregninger hørende til hypotesetest

4a: Uafhængighed

$$\left[\begin{array}{l|cc|c} \text{Køn} \backslash \text{Forbrug} & < 1500 \text{ kr./måned} & \geq 1500 \text{ kr./måned} & \text{i alt} \\ \hline \text{kvinder} & 98 & 102 & 200 \\ \text{mænd} & 60 & 100 & 160 \\ \hline \text{i alt} & 158 & 202 & 360 \end{array} \right]^{17}$$

Løsning: Vi skal nu undersøge om der er afhængighed mellem køn og tøjforbrug! Vi får da først og fremmest brug for at beregne de forventede værdier og teststørrelsen.

De forventede værdier udregnes nemmest ved at indskrive tabellen som krydslister i regnearket og så anvende den viste celleformel, der trækkes ned gennem søjlen.

| | A kat_køn | B kat_forbrug | C obs_hyp | D forv_hyp | E x_var | F y_var |
|---|-----------|---------------|-----------|------------|---------|---------|
| 1 | Kvinde | Lavt | 98 | 87.7778 | | |
| 2 | Kvinde | Højt | 102 | 112.222 | | |
| 3 | Mand | Lavt | 60 | 70.2222 | | |
| 4 | Mand | Højt | 100 | 89.7778 | | |

| | |
|----|---|
| D1 | = $\frac{\text{sumif}(\text{'kat_køn'}, a1, \text{'obs_hyp'}) \cdot \text{sumif}(\text{'kat_forbrug'}, b1, \text{'obs_hyp'})}{\text{sum}(\text{'obs_hyp'})} \cdot 1$ |
|----|---|

Vi finder da:

$$\text{forv_hyp} \rightarrow \{87.7778, 112.222, 70.2222, 89.7778\}$$

$$\text{chi2_obs} := \text{sum} \left(\frac{(\text{obs_hyp} - \text{forv_hyp})^2}{\text{forv_hyp}} \right)^{18} \rightarrow 4.77353$$

Herefter er vejen banet for udregning af p-værdien, dvs. sandsynligheden for at vi rammer mindst lige så skævt som det observerede:

$$\text{p_værdi} := \text{chi2CDF}(\text{chi2_obs}, \infty, 1) \rightarrow 0.028901$$

Den kritiske sandsynlighed er altså 2.89%, hvorfor afvigelsen er signifikant på 5%-niveau, dvs. vi forkaster nulhypotesen om uafhængighed på 5%-niveauet.

Vi kunne også have fundet den kritiske grænse for teststørrelsen:

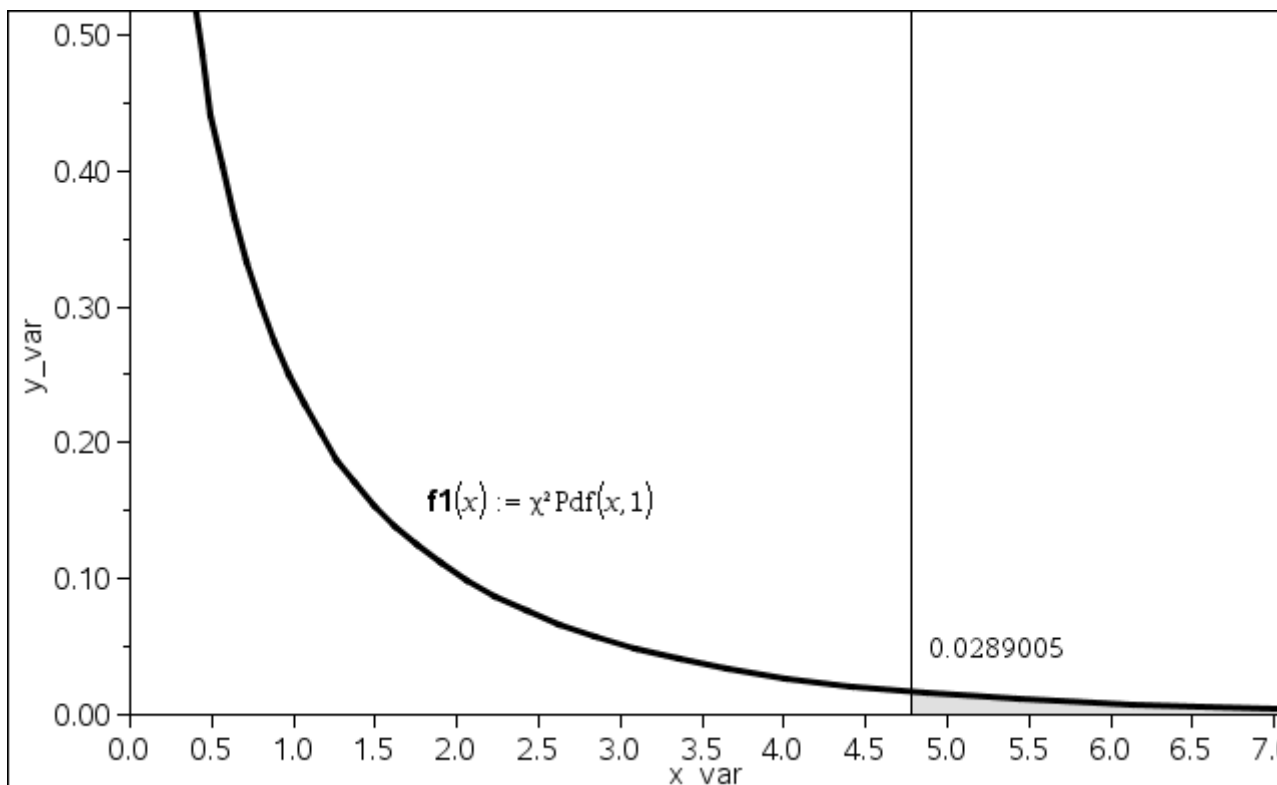
$$\text{chi2_krit} := \text{invChi2}(0.95, 1) \rightarrow 3.84146$$

¹⁷ [[Køn\Forbrug,<1500 kr./måned,≥ 1500 kr./måned,i alt][kvinder,98,102,200][mænd,60,100,160][i alt,158,202,360]]

¹⁸ $\text{chi2_obs} := \text{sum}(\frac{(\text{obs_hyp} - \text{forv_hyp})^2}{\text{forv_hyp}})$

Hvis teststørrelsen ligger over 3.84 er afvigelsen altså kritisk, dvs. vi må forkaste nulhypotesen (på 5%-niveauet).

Endelig kan vi illustrere testen grafisk. Det gøres nemmest i data og statistik-applikationen. For at få lov til at tegne en graf, må vi dog først indføre to tomme variable **x_var** og **y_var**!



4b Goodness of fit

Indkomstfordelingen i stikprøven var: $I = \text{Indkomst i 1000 kr.}$

| | | | | | | | |
|----------|-------------------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------|
| $I < 50$ | $50 \leq I < 100$ | $100 \leq I < 150$ | $150 \leq I < 200$ | $200 \leq I < 300$ | $300 \leq I < 400$ | $400 \leq I < 500$ | $500 \leq I$ |
| 98 | 88 | 199 | 136 | 210 | 179 | 52 | 38 |

Den forventede fordeling i stikprøven baseret på de ovenstående procenter er tilsvarende givet ved:

| | | | | | | | |
|----------|-------------------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------|
| $I < 50$ | $50 \leq I < 100$ | $100 \leq I < 150$ | $150 \leq I < 200$ | $200 \leq I < 300$ | $300 \leq I < 400$ | $400 \leq I < 500$ | $500 \leq I$ |
| 64 | 93 | 178 | 123 | 243 | 180 | 66 | 53 |

Sammenholder vi de observerede hyppigheder med de forventede følges de så nogenlunde ad.

Men man kunne måske være bekymret for, om de laveste indkomster er overrepræsenteret i stikprøven. Her ligger den observerede hyppighed et godt stykke over den forventede.

Løsning: Vi skal undersøge om den observerede fordeling følger den forventede. Vi overfører derfor data til lister og udregner teststørrelsen:

$$\text{chi2_obs} := \text{sum} \left(\frac{(\text{obs_hyp} - \text{forv_hyp})^2}{\text{forv_hyp}} \right) \quad 21 \rightarrow 33.8848$$

¹⁹ $[[I < 50, 50 \leq I < 100, 100 \leq I < 150, 150 \leq I < 200, 200 \leq I < 300, 300 \leq I < 400, 400 \leq I < 500, 500 \leq I] [98, 88, 199, 136, 210, 179, 52, 38]]$

²⁰ $[[I < 50, 50 \leq I < 100, 100 \leq I < 150, 150 \leq I < 200, 200 \leq I < 300, 300 \leq I < 400, 400 \leq I < 500, 500 \leq I] [64, 93, 178, 123, 243, 180, 66, 53]]$

²¹ $\text{chi2_obs} := \text{sum}(\frac{(\text{obs_hyp} - \text{forv_hyp})^2}{\text{forv_hyp}})$

Herefter er vejen banet for udregning af p -værdien, dvs. sandsynligheden for at vi rammer mindst lige så skævt som det observerede:

$$p_værdi := \text{chi2CDF}(\text{chi2_obs}, \infty, 7) \rightarrow 0.000018$$

Den kritiske sandsynlighed er altså 0.0018%, hvorfor afvigelsen er signifikant på 1%-niveau, dvs. vi forkaster nulhypotesen om at stikprøven er repræsentativ for landsfordelingen på 1%-niveauet.

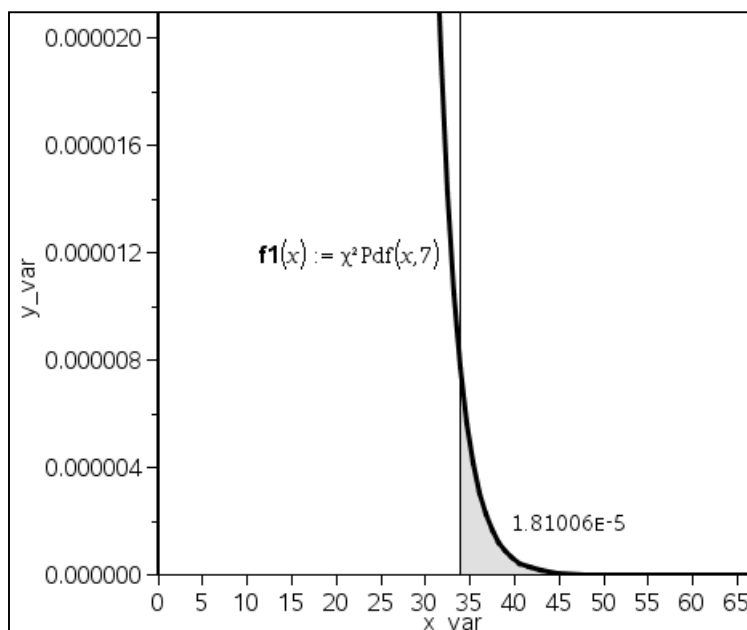
| | A kat_indkomst | E obs_hyp | C forv_hyp | D x_var | E y_var |
|---|----------------|-----------|------------|---------|---------|
| 1 | <50 | 98 | 64. | | |
| 2 | 50≤ <100 | 88 | 93. | | |
| 3 | 100≤ <150 | 199 | 178. | | |
| 4 | 150≤ <200 | 136 | 123. | | |
| 5 | 200≤ <300 | 210 | 243. | | |
| 6 | 300≤ <400 | 179 | 180. | | |
| 7 | 400≤ <500 | 52 | 66. | | |
| 8 | 500≤ | 38 | 53. | | |

Vi kunne også have fundet den kritiske grænse for teststørrelsen:

$$\text{chi2_krit} := \text{invChi2}(0.99, 7) \rightarrow 18.4753$$

Hvis teststørrelsen ligger over 18.47 er afvigelsen altså kritisk, dvs. vi må forkaste nulhypotesen (på 1%-niveauet).

Endelig kan vi illustrere testen grafisk. Det gøres nemmest i data og statistik-applikationen. For at få lov til at tegne en graf, må vi dog først indføre to tomme variable **x_var** og **y_var**! I dette tilfælde er det dog meget svært at illustrere det kritiske område uden at grafen for fordelingen bliver meget uoverskuelig.



5) Indbyggede testrutiner

5a: Uafhængighed

Eksempel 1: (side 4 i kursusmaterialet)

| <i>Køn\Forbrug</i> | <i>< 1500 kr./måned</i> | <i>≥ 1500 kr./måned</i> | <i>i alt</i> |
|--------------------|----------------------------|-------------------------|--------------|
| kvinder | 98 | 102 | 200 |
| mænd | 60 | 100 | 160 |
| i alt | 158 | 202 | 360 |

Løsning: Vi skal afgøre om de oplyste data er i rimelig overensstemmelse med nulhypotesen om uafhængighed mellem **Køn** og **Forbrug**. Vi benytter det indbyggede test for uafhængighed af to variable (2 way test), der forventer at få oplyst matricen for de observerede hyppigheder:

$$\mathbf{obs} := \begin{bmatrix} 98 & 102 \\ 60 & 100 \end{bmatrix}^{23} \rightarrow \begin{bmatrix} 98 & 102 \\ 60 & 100 \end{bmatrix}$$

$$\mathbf{chi2way\ obs : Stat.results} \rightarrow \begin{bmatrix} \text{"Title"} & \text{"}\chi^2 \text{ 2-way Test"} \\ \text{"}\chi^2\text{"} & 4.77353 \\ \text{"PVal"} & 0.028901 \\ \text{"df"} & 1. \\ \text{"ExpMatrix"} & \text{"[...]"} \\ \text{"CompMatrix"} & \text{"[...]"} \end{bmatrix}$$

Vi får da som vist en række koncentrerede oplysninger:

- 1) Teststørrelsen har værdien 4.77
- 2) p -værdien er 2.89%, dvs. sandsynligheden for at finde en teststørrelse, der er mindst lige så skæv som den observerede er 2.89%. Nulhypotesen forkastes altså på signifikansniveauet 5%, men den forkastes ikke på 1% niveau!
- 3) Teststørrelsen er chi2 fordelt med 1 frihedsgrad.
- 4) Vi får adgang til matricen for de forventede værdier:

$$\mathbf{stat.expmatrix} \rightarrow \begin{bmatrix} 87.7778 & 112.222 \\ 70.2222 & 89.7778 \end{bmatrix}$$

- 5) Vi får også adgang til matricen for de enkelte bidrag til chi2-teststørrelsen:

$$\mathbf{stat.compmatrix} \rightarrow \begin{bmatrix} 1.19044 & 0.931133 \\ 1.48805 & 1.16392 \end{bmatrix}$$

²² [[Køn\Forbrug,<1500 kr./måned,≥ 1500 kr./måned,i alt][kvinder,98,102,200][mænd,60,100,160][i alt,158,202,360]]

²³ $\mathbf{obs}:=$ [[98,102][60,100]]

5b: Goodness of fit

Eksempel 2: (side 24 i kursusmaterialet)

Indkomstfordelingen i stikprøven var: $I = \text{Indkomst i 1000 kr.}$

Observeret antal

$$\left[\begin{array}{cccccccc} I < 50 & 50 \leq I < 100 & 100 \leq I < 150 & 150 \leq I < 200 & 200 \leq I < 300 & 300 \leq I < 400 & 400 \leq I < 500 & 500 \leq I \end{array} \right]_{24}$$

$$\left[\begin{array}{cccccccc} 98 & 88 & 199 & 136 & 210 & 179 & 52 & 38 \end{array} \right]$$

Den forventede fordeling i stikprøven baseret på de ovenstående procenter er tilsvarende givet ved:

Forventet antal

$$\left[\begin{array}{cccccccc} I < 50 & 50 \leq I < 100 & 100 \leq I < 150 & 150 \leq I < 200 & 200 \leq I < 300 & 300 \leq I < 400 & 400 \leq I < 500 & 500 \leq I \end{array} \right]_{25}$$

$$\left[\begin{array}{cccccccc} 64 & 93 & 178 & 123 & 243 & 180 & 66 & 53 \end{array} \right]$$

Sammenholder vi de observerede hyppigheder med de forventede følger de så nogenlunde ad. Men man kunne måske være bekymret for, om de laveste indkomster er overrepræsenteret i stikprøven. Her ligger den observerede hyppighed et godt stykke over den forventede.

Løsning: Vi skal afgøre om de oplyste data er i rimelig overensstemmelse med nulhypotesen om samme fordeling for stikprøve og population. Vi benytter det indbyggede test for Goodness of fit, der forventer at få oplyst lister for de observerede hyppigheder, de forventede hyppigheder, samt antallet af frihedsgrader (som vi altså selv skal oplyse i denne test!). Udføres testet fra **Lister** og **Regneark** får vi ydermere mulighed for at vælge en grafisk illustration af testet, hvor vi dog ikke kan få det kritiske område at se, fordi de anvendte data er så ekstreme:

```
stat.results → [ "Title"      "χ² GOF"
                  "χ²"        33.8848
                  "PVal"     0.000018
                  "df"       7.
                  "CompList"  "{...}" ]
```

| | A kat_indkomst | E obs_hyp | C forv_hyp | D | E |
|---|----------------|-----------|------------|------------|-----------------|
| ♦ | | | | | =χ²GOF('obs, |
| 1 | I<50 | 98 | 64. | Title | χ² GOF |
| 2 | 50≤I<100 | 88 | 93. | χ² | 33.8848 |
| 3 | 100≤I<150 | 199 | 178. | PVal | 0.000018 |
| 4 | 150≤I<200 | 136 | 123. | df | 7. |
| 5 | 200≤I<300 | 210 | 243. | CompLis... | {18.0625,0.2... |
| 6 | 300≤I<400 | 179 | 180. | | |
| 7 | 400≤I<500 | 52 | 66. | | |
| 8 | 500≤ | 38 | 53. | | |

²⁴ [[I<50 ,50≤I<100 ,100≤I<150 ,150≤I<200 ,200≤I<300 ,300≤I<400 ,400≤I<500 ,500≤I][98,88,199,136,210,179,52,38]]

²⁵ [[I<50 ,50≤I<100 ,100≤I<150 ,150≤I<200 ,200≤I<300 ,300≤I<400 ,400≤I<500 ,500≤I][64,93,178,123,243,180,66,53]]

Vi får da som vist en række oplysninger:

- 1) Teststørrelsen har værdien 33.88
- 2) p -værdien er 0.0018%, dvs. sandsynligheden for at finde en teststørrelse mindst lige så stor som den observerede er 0.0018%. Da den ligger under signifikansniveauet 1% forkastes nulhypotesen på signifikansniveauet 1%.
- 3) Teststørrelsen er χ^2 fordelt med 7 frihedsgrader (som vi selv har måttet oplyse).
- 4) Vi får adgang til de enkelte bidrag til teststørrelsen:

stat.complist \rightarrow {18.0625,0.268817,2.47753,1.37398,4.48148,0.005556,2.9697,4.24528}

Hovedbidraget kommer altså fra den første kategori, dvs. for de laveste indkomster.

