Vejledning til GYM17

Indhold

Vejledning i brug af Gym17-pakken

Installation

Gym-pakken vil automatisk være installeret på din pc eller mac, hvis du benytter cd'en

Maple 17 - Til danske Gymnasier

eller en af de tilsvarende installere. Det eneste, du behøver, for at indlæse Gym-pakken, er at benytte Maple kommandoen

$with(Gym)$

$[ChiKvadratGOFtest, ChiKvadratUtest, Cos, ExpReg, LinReg, LogistReg, PolyReg, PowReg, PropReg, Sin, Tan, antalobs, antalstabel, arealP, arealT, bidrag, binomialTest, boksplot, cart2pol, chicdf, chipdf, det, dotP, ev, forventet, fraktil, frekvens, frekvensTabel, gennemsnit, grupperData, hat, hyppighed, invCos, invSin, invTan, invchi, invnorm, invt, kumuleretFrekvens, kvartiler, len, median, middel, normalcdf, normalpdf, pindediagramBIN, plotHistogram, plotPindediagram, plotResidualer, plotSumkurve, plotTrappekurve, pol2cart, proj, residualer, spredning, sumkurve, tInterval, tTest, tcdf, testLin, tpdf, trappekurve, trappekurveBIN, typeinterval, typetal, varians, vinkel, visMatrix, vsolve, zInterval, zInterval1Andel, zTest]$

Gym-pakken består af en række rutiner, der skal gøre arbejdet med Maple mere bekvemt inden for

- Deskriptiv statistik
- Regressioner
- Trigonometri
- Vektorregning
- Statistiske test

Nedenfor vil nogle af Gym-pakkens rutiner blive behandlet, opdelt efter de 5 ovennævnte områder. Beskrivelserne, der ledsages af små instruktive eksempler, vil ikke omfatte alle detaljer. For en mere detaljeret beskrivelse henvises til on-line hjælpen i Gym-pakken. Du taster blot

$?Gym$

efterfulgt af Enter, hvorefter du kan navigere i on-line hjælpen via hyperlink.

Hvis du **ikke** har benyttet cd'en 'Maple 17 - Til danske Gymnasier' eller en af de tilsvarende installere til din installation, behøver du ikke at geninstallere Maple.

Du kan downloade Gym-pakken fra MapleGym.dk, og manuelt placere de 3 filer (Gym.lib, Gym.ind og gymhelp.hdb) i lib-mappen i Maple 17 installationen.

Du finder lib-mappen her:

PC	C:\Program Files\Maple 17\lib
Mac	Macintosh HD / Library / Frameworks / Maple.framework / Versions / 17 / lib

1 Deskriptiv statistik

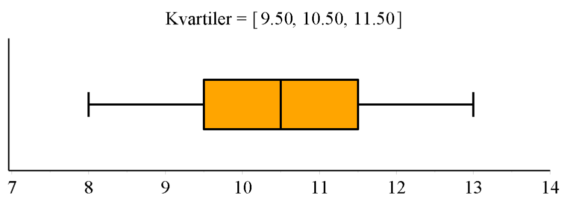
1.1 Ikke-grupperede observationssæt

Du kan indtaste data i en liste, en vektor, en matrix eller indlæse data fra en ekstern fil (fx i Excel format). Her er data i en liste (listen behøver ikke at være sorteret):

$obs := [8, 8, 9, 9, 9, 10, 10, 10, 10, 10, 11, 11, 11, 11, 11, 12, 12, 13, 13, 13]:$

Du kan tegne et boksplot direkte på baggrund af de rå data, hvor du tillige vil få kvartilsættet oplyst:

$boksplot(obs)$



Vil du bestemme hyppigheder, frekvenser og kvartiler (ud fra en trappekurve), kan du gå frem på denne måde:

Hyppigheder og frekvenser findes (du kan også benytte hyppighedstabellen som input til *frekvenser*):

$$H := hyppighed(obs) = \begin{bmatrix} 8 & 2 \\ 9 & 3 \\ 10 & 5 \\ 11 & 5 \\ 12 & 2 \\ 13 & 3 \end{bmatrix}$$

$$F := frekvens(obs) = \begin{bmatrix} 8 & 0.100 \\ 9 & 0.150 \\ 10 & 0.250 \\ 11 & 0.250 \\ 12 & 0.100 \\ 13 & 0.150 \end{bmatrix}$$

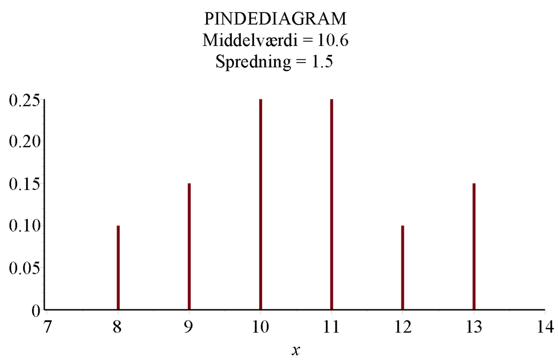
$$kumuleretFrekvens(F) = \begin{bmatrix} 8 & 0.100 \\ 9 & 0.250 \\ 10 & 0.500 \\ 11 & 0.750 \\ 12 & 0.850 \\ 13 & 1. \end{bmatrix}$$

Du kan klare det hele i én arbejdsgang ved at få udskrevet en frekvenstabel - du kan dog ikke benytte tabellen i senere beregninger, da det kun er tekst der udskrives *frekvensTabel*(*obs*)

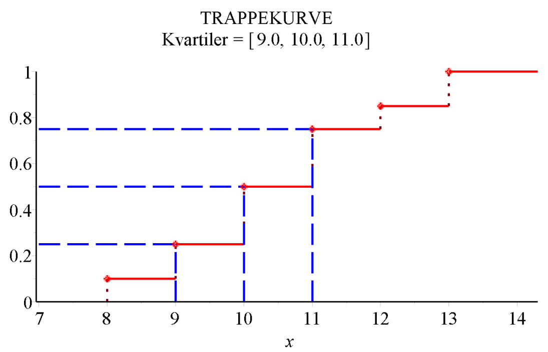
observation	hyppighed	frekvens	kumuleret
8	2	0.1	0.1
9	3	0.15	0.25
10	5	0.25	0.5
11	5	0.25	0.75
12	2	0.1	0.85
13	3	0.15	1

Pindediagrammet og trappekurven tegnes sådan her (i stedet for hyppigheds- eller frekvenstabellen kan du benytte de rå data):

plotPindediagram(*H*)



plotTrappekurve(*F*)



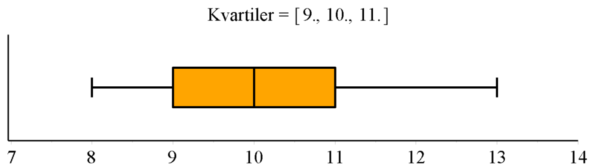
Læg mærke til, at du ikke får samme kvartilsæt som i boksplottet ovenforDet skyldes, at to forskellige metoder benyttes (læs mere om dette i online hjælpen):

Ved direkte bestemmelse af kvartilsættet, skal du oplyse, hvilken metode du vil benytte. Boksplot-metoden er standard ved rå data, så	*kvartiler*(*obs*) = [9.50, 10.50, 11.50]
Er der sket en optælling af data i enten en hyppighedstabel, så er trappekurve metoden standard:	*kvartiler*(*H*) = [9., 10., 11.]
Hvis du vil benytte trappekurvemetoden til bestemmelse af kvartilsættet på baggrund af de rå data, skal du tilføje en parameter:	*kvartiler*(*obs*, *metode* = 1) = [9., 10., 11.]

- og hvis du vil tvinge Maple til at bruge boksplotmetoden, hvor trappekurvemetoden er standard, skal du tilføje en parameter:	$kvartiler(H, metode = 2) = [9.50, 10.50, 11.50]$

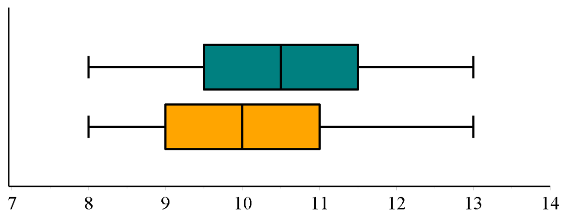
Vil du have tegnet et boksplot på basis af kvartilsættet bestemt ved trappekurvemetoden, skal du blot tilføje en parameter:

$boksplot([obs, metode = 1])$



De to metoder kan sammenlignes i et kombineret boksplot

$boksplot([obs, metode = 1], [obs, metode = 2])$



Pindediagrammet og trappekurven giver oplysning om middelværdi, spredning og kvartilsæt. Disse (og øvrige deskriptorer) kan også findes direkte:

$gennemsnit(obs) = 10.550$

$spredning(obs) = 1.499166435$

$typetal(obs) = [10, 11]$

Fraktiler kan også bestemmes. Fx findes 0.6-fraktilen således - her skal benyttes en hyppigheds- eller en frekvenstabel)

$fraktil(F, 0.6) = 11$

Forskriften for trappekurven kan fås således:

$trappekurve(F, x)$

$$\begin{cases} 0 & x < 8 \\ 0.100000000000000 & 8 \leq x \textbf{ and } x < 9 \\ 0.250000000000000 & 9 \leq x \textbf{ and } x < 10 \\ 0.500000000000000 & 10 \leq x \textbf{ and } x < 11 \\ 0.750000000000000 & 11 \leq x \textbf{ and } x < 12 \\ 0.850000000000000 & 12 \leq x \textbf{ and } x < 13 \\ 1 & 13 \leq x \end{cases} \tag{1.1}$$

1.2 Grupperede observationssæt

Data kan grupperes med funktionen *grupperData*:

$obs := [21.3, 13.7, 7.4, 13.4, 12.8, 9.2, 8.9, 4.2, 15.5, 11.9, 18.2, 14.1, 10.9, 21.7, 10.1, 10.2, 4.2, 23.3, 21.7, 9.9, 16.1, 22.4, 8.5, 13.1, 15.3, 19.0, 14.4, 15.6, 18.2, 14.9, 10.8, 13.7, 11.5, 24.8, 13.7, 14.6, 21.1, 10.1, 24.7, 15.6, 17.2, 12.4, 16.1, 12.9, 15.2, 24.9, 26.1, 19.4, 19.4, 10.7]:$

Start med at finde den mindste og den største observation:

$\min(obs) = 4.2$

$\max(obs) = 26.1$

Alle observationer ligger i al fald i intervallet $[0, 30]$. Dette interval opdeles i delintervaller af længde 5, og observationerne grupperes efter denne opdeling i 6 grupper:

$G := grupperData(obs, 0..30, 6)$

$$\begin{bmatrix} 0..5. & 2 \\ 5...10. & 5 \\ 10...15. & 20 \\ 15...20. & 13 \\ 20...25. & 9 \\ 25...30 & 1 \end{bmatrix} \tag{1.2}$$

Hvis du har de grupperede data givet, skal du indtaste data i en matrix med observationsintervallerne i 1. søjle og hyppigheder (eller frekvenser) i 2. søjle.

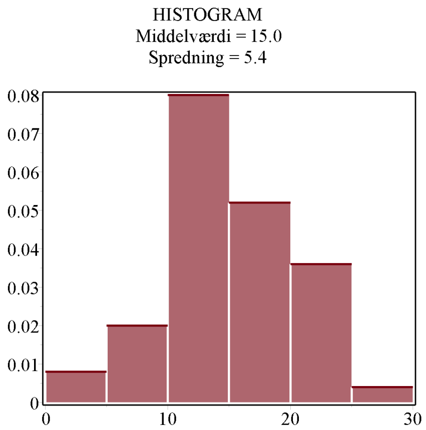
Intervalfrekvenserne og de kumulerede frekvenser finder du sådan her:

$$frekvens(G) = \begin{bmatrix} 0..5. & 0.0400 \\ 5...10. & 0.100 \\ 10...15. & 0.400 \\ 15...20. & 0.260 \\ 20...25. & 0.180 \\ 25...30 & 0.0200 \end{bmatrix}$$

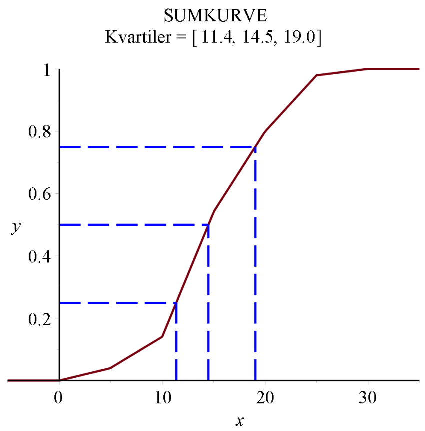
$$kumuleretFrekvens(G) = \begin{bmatrix} 0..5. & 0.0400 \\ 5...10. & 0.140 \\ 10...15. & 0.540 \\ 15...20. & 0.800 \\ 20...25. & 0.980 \\ 25...30 & 1. \end{bmatrix}$$

Histogram og sumkurve tegnes:

plotHistogram(*G*)



plotSumkurve(*G*)



Histogrammet og sumkurven giver oplysning om middelværdi, spredning og kvartilsæt. Disse (og øvrige deskriptorer) kan findes direkte:

middel(*G*) = 15.0000000000000

spredning(*G*) = 5.40832691319598

typeinterval(*G*) = [10. .15.]

Fraktiler kan også bestemmes. Fx findes 0.6-fraktilen således:

fraktil(*G*, 0.6) = 16.15384615

Forskriften for sumkurven fås således:

$sumkurve(G, x)$

$$\begin{cases} 0 & x < 0 \\ 0.00800000000000000\,x & 0 \leq x \textbf{ and } x < 5. \\ 0.0200000000000000\,x - 0.0600000000000000 & 5. \leq x \textbf{ and } x < 10. \\ 0.0800000000000000\,x - 0.660000000000000 & 10. \leq x \textbf{ and } x < 15. \\ 0.0520000000000000\,x - 0.240000000000000 & 15. \leq x \textbf{ and } x < 20. \\ 0.0360000000000000\,x + 0.0800000000000003 & 20. \leq x \textbf{ and } x < 25. \\ 0.00400000000000000\,x + 0.880000000000000 & 25. \leq x \textbf{ and } x < 30 \\ 1 & 30 \leq x \end{cases} \tag{1.3}$$

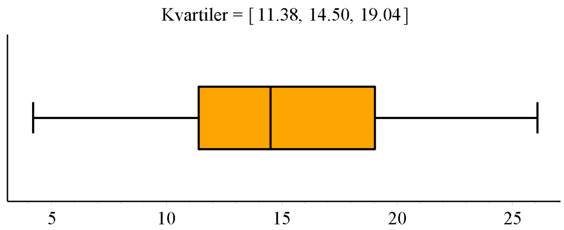
Skal du aflæse nogle værdier på din sumkurve - fx værden i 22 - så indsætter du blot i *sumkurve*:

$sumkurve(G, 22) = 0.872000000000000$

der fortæller, at 87.2% af alle observationer er mindre end eller lig med 22.

Du kan også få tegnet et boksplot for et grupperet observationssæt. Her er du dog nødt til at angive minimums- og maksimumsværdien, hvis disse kendes, ellers benyttes (her) 0 og 30.

$boksplot([G, \{4.2, 26.1\}])$



2 Regressioner

Gym-pakken indeholder 6 rutiner til regression: *LinReg, PropReg, ExpReg, PowReg, PolyReg* og *LogistReg*. Disse er meget fleksible og flere dataformater tillades.

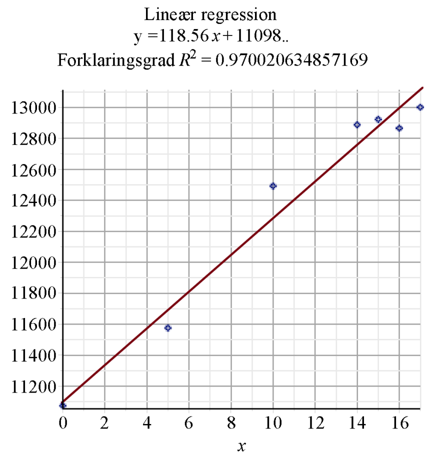
I nedenstående eksemper repræsenteres data i to lister, men det kunne også være som vektorer, matricer og arrays (hvis data importeres fra en ekstern fil).

2.1 Lineær regression

$X := [0, 5, 10, 14, 15, 16, 17]:$

$Y := [11073, 11575, 12492, 12887, 12922, 12865, 13000]:$

$LinReg(X, Y)$



LinReg giver på én gang regressionsligningen, forklaringsgraden og et plot af datapunkterne sammen med grafen for regressionsligningen.

Er der behov for at definere regressionsudtrykket som en funktion, klares dette ved at tilføje det ønskede navn på den uafhængige variabel som en tredje parameter i *LinReg*. Fx

$LinReg(X, Y, x)$

$$118.561475409836\,x + 11097.8237704918 \tag{2.1}$$

Hvis du skal finde værdien i fx $x = 20$ af regressionsudtrykket, skal du blot sætte 20 ind i $LinReg(X, Y, x)$ i stedet for x:

$LinReg(X, Y, 20) = 13469.0532786885$

Vil du have et mere bekvemt navn til regressionsfunktionen, kan du definere

$f := x \rightarrow LinReg(X, Y, x)$

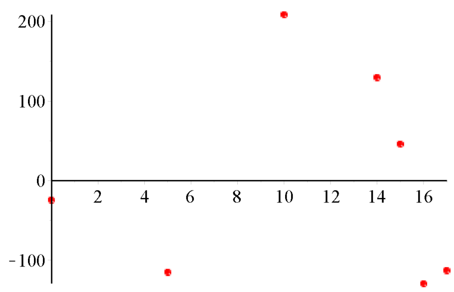
$$x \rightarrow Gym\text{:-}LinReg(X, Y, x) \tag{2.2}$$

Du kan få fat i regressionskoefficienterne med *reg_koeff*

$$\begin{bmatrix} b \\ a \end{bmatrix} = \begin{bmatrix} 11097.8237704918 \\ 118.561475409836 \end{bmatrix} \tag{2.3}$$

Du kan få fat i residualerne med kommandoen $residualer(X, Y, LinReg)$ og du kan plotte residualerne med

$plotResidualer(X, Y, LinReg)$

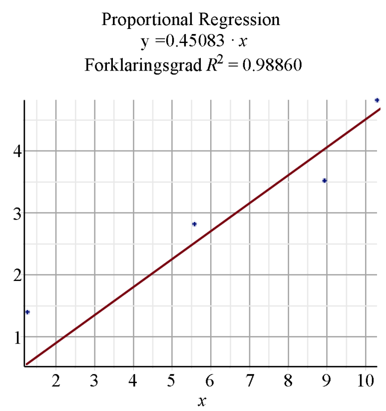


2.2 Proportional regression

$X := [1.26, 5.58, 8.94, 10.3]$:

$Y := [1.4, 2.82, 3.52, 4.82]$:

$PropReg(X, Y)$



PropReg giver på én gang regressionsligningen, forklaringsgraden og et plot af datapunkterne sammen med grafen for regressionsligningen.

Er der behov for at definere regressionsudtrykket som en funktion, klares dette ved at tilføje det ønskede navn på den uafhængige variabel som en tredje parameter i *PropReg*. Fx

PropReg(*X*, *Y*, *x*)

$$0.4508342416\,x \tag{2.4}$$

Hvis du skal finde værdien i fx $x = 5$ af regressionsudtrykket, skal du blot sætte 5 ind i *PropReg*(*X*, *Y*, *x*) i stedet for *x*:

LinReg(*X*, *Y*, 5) = 2.61800098570725

Vil du have et mere bekvemt navn til regressionsfunktionen, kan du definere

$f := x \rightarrow PropReg(X, Y, x)$

$$x \rightarrow Gym\text{:-}PropReg(X, Y, x) \tag{2.5}$$

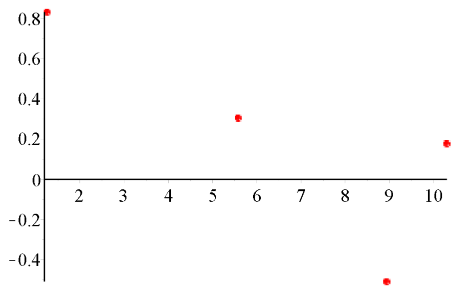
Du kan få fat i regressionskoefficienten med

reg_koeff

$$a = 0.4508342416 \tag{2.6}$$

Du kan få fat i residualerne med kommandoen *residualer*(*X*, *Y*, *PropReg*) og du kan plotte residualerne med

plotResidualer(*X*, *Y*, *PropReg*)

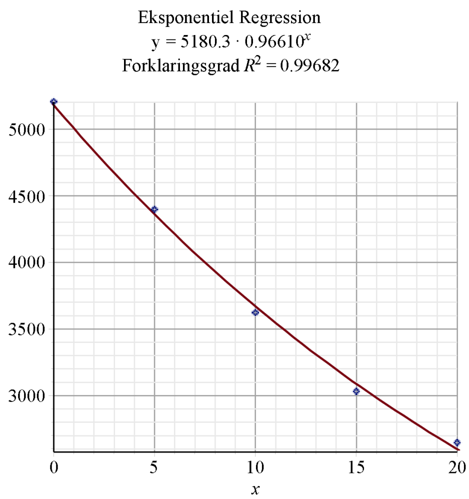


2.3 Eksponentiel regression

$X := [0, 5, 10, 15, 20]$:

$Y := [5205, 4397, 3622, 3031, 2647]$:

$ExpReg(X, Y)$



ExpReg giver på én gang regressionsligningen, forklaringsgraden og et plot af datapunkterne sammen med grafen for regressionsligningen.

Hvis ovenstående graf ønskes tegnet med en logaritmisk y-akse, klares dette ved et højre-klik i grafen, og y-aksen indstilles til logaritmisk under 'axes' i kontekstmenuen.

Er der behov for at definere regressionsudtrykket som en funktion, klares dette ved at tilføje det ønskede navn på den uafhængige variabel som en tredje parameter i *ExpReg*. Fx

$ExpReg(X, Y, t) = 5180.28826541937\ 0.966099634291293^t$

Hvis du skal finde værdien i fx $t = 7$ af regressionsudtrykket, skal du blot sætte 7 ind i $ExpReg(X, Y, t)$ i stedet for t:

$ExpReg(X, Y, 7) = 4069.18421755868$

Vil du have et mere bekvemt navn til regressionsfunktionen, kan du definere

$f := x \rightarrow ExpReg(X, Y, t)$

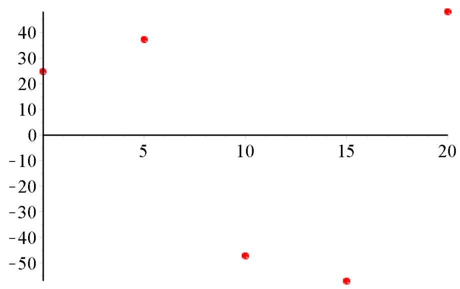
$$x \rightarrow Gym\text{:-}ExpReg(X, Y, t) \tag{2.7}$$

Du kan få fat i regressionskoefficienterne med

reg_koeff

$$\begin{bmatrix} b \\ a \end{bmatrix} = \begin{bmatrix} 5180.28826541937 \\ 0.966099634291293 \end{bmatrix} \tag{2.8}$$

Du kan få fat i residualerne med kommandoen $residualer(X, Y, ExpReg)$ og du kan plotte residualerne med *plotResidualer*$(X, Y, ExpReg)$

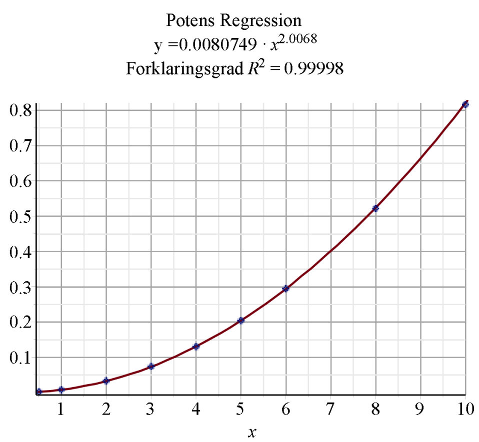


2.4 Potensregression

$X := [0.5, 1, 2, 3, 4, 5, 6, 8, 10]:$

$Y := [0.002, 0.008, 0.033, 0.074, 0.130, 0.204, 0.294, 0.522, 0.816]:$

PowReg(X, Y)



PowReg giver på én gang regressionsligningen, forklaringsgraden og et plot af datapunkterne sammen med grafen for regressionsligningen.

Hvis ovenstående graf ønskes tegnet med en logaritmiske akser, klares dette ved et højre-klik i grafen, og både x-aksen og y-aksen indstilles til logaritmisk under 'axes' i kontekstmenuen.

Er der behov for at definere regressionsudtrykket som en funktion, klares dette ved at tilføje det ønskede navn på den uafhængige variabel som en tredje parameter i *PowReg*. Fx

$PowReg(X, Y, x) = 0.00807487866766576\,x^{2.00675834882888}$

Hvis du skal finde værdien i fx $x = 9$ af regressionsudtrykket, skal du blot sætte 9 ind i $PowReg(X, Y, x)$ i stedet for x:

$PowReg(X, Y, 9) = 0.663850257440628$

Vil du have et mere bekvemt navn til regressionsfunktionen, kan du definere

$f := x \rightarrow PowReg(X, Y, x)$

$$x \rightarrow Gym\text{:-}PowReg(X, Y, x) \tag{2.9}$$

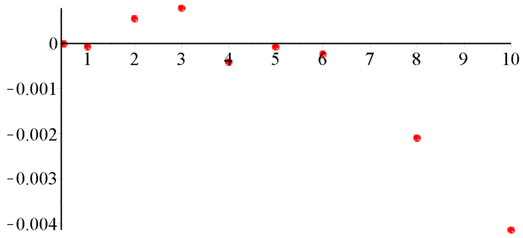
Du kan få fat i regressionskoefficienterne med

reg_koeff

$$\begin{bmatrix} b \\ a \end{bmatrix} = \begin{bmatrix} 0.00807487866766576 \\ 2.00675834882888 \end{bmatrix} \tag{2.10}$$

Du kan få fat i residualerne med kommandoen $residualer(X, Y, PowReg)$ og du kan plotte residualerne med

$plotResidualer(X, Y, PowReg)$



2.5 Polynomiel regression

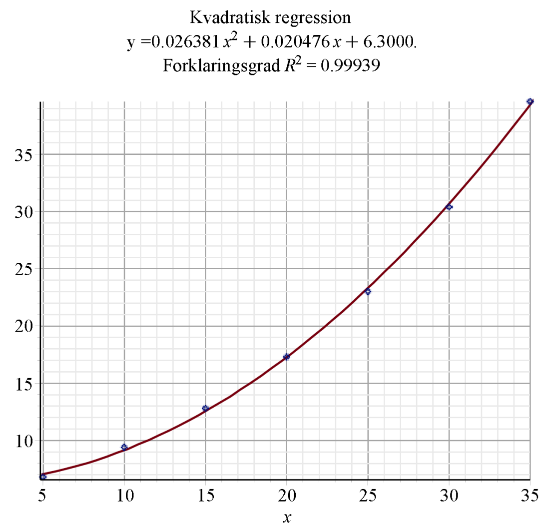
Med *PolyReg* kan du tilpasse data til et andengradspolynomium, et tredjegradspolynomium, osv. Bruger du *PolyReg* til at tilpasse til et førstegradspolynomium, svarer dette helt til lineær regression.

Lad os som eksempel se på en tilpasning til et andengradspolynomium - eller kvadratisk regression, som det også kaldes.

$X := [5, 10, 15, 20, 25, 30, 35]:$

$Y := [6.8, 9.4, 12.8, 17.3, 23.0, 30.4, 39.6]$:

$PolyReg(X, Y, 2)$



Bemærk, at denne regression kræver 3 parametre, hvor den sidste angiver graden (2 for andengrads, 3 for tredjegrads, osv.)

Skal du bruge regressionsudtrykket (uden graf), så skal du bruge 4-parameter versionen:
$PolyReg(X, Y, 2, x)$

$$0.00815745840940880\,x^2 + 0.0000111684228955428\,x + 0.0000408431437811916 \tag{2.11}$$

$\xrightarrow{\text{at 5 digits}}$

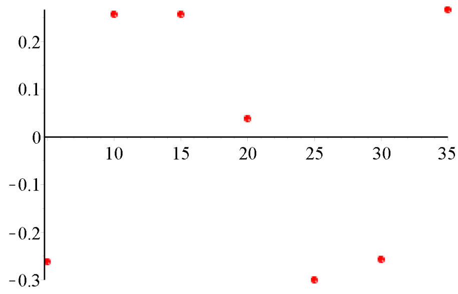
$$0.0081575\,x^2 + 0.000011168\,x + 0.000040843$$

Du kan få fat i regressionskoefficienterne med

reg_koeff

$$\begin{bmatrix} c \\ b \\ a \end{bmatrix} = \begin{bmatrix} 0.0000408431437811916 \\ 0.0000111684228955428 \\ 0.00815745840940880 \end{bmatrix} \tag{2.12}$$

Du kan få fat i residualerne med kommandoen $residualer(X, Y, LinReg)$ og du kan plotte residualerne med *plotResidualer*$(X, Y, [PolyReg, 2])$



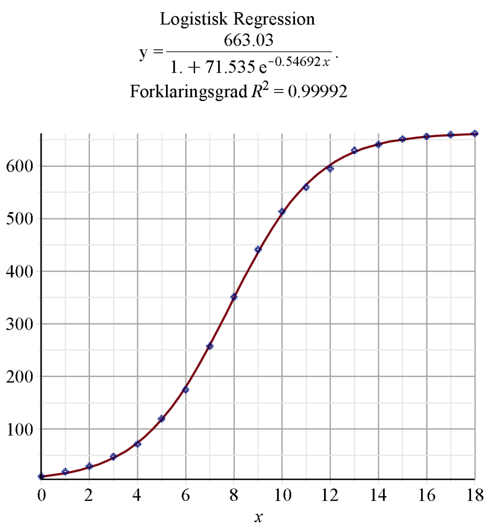
2.6 Logistisk regression

Tilpasning af data til en logistisk funktion kan ske med *LogistReg*:

$X := [0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18]$:

$Y := [9.6, 18.6, 29.0, 47.2, 71.1, 119.1, 174.6, 257.3, 350.7, 441.0, 513.3, 559.7, 594.8, 629.4, 640.8, 651.1, 655.9, 659.6, 661.8]$:

LogistReg(X, Y)



Hvis du kun vil have funktionsudtrykket benyttes:

LogistReg(X, Y, x)

$$\frac{663.032899928510}{1 + 71.5350647080719\, e^{-0.546919002044827\, x}} \tag{2.13}$$

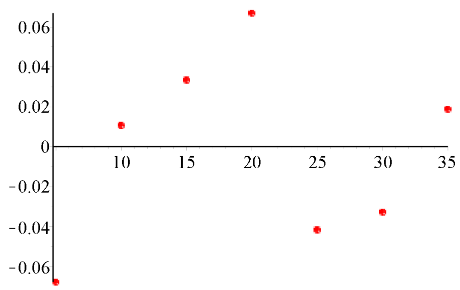
Du kan få fat i regressionskoefficienterne med

reg_koeff

$$\begin{bmatrix} 71.5350647080719 \\ 0.546919002044827 \\ 663.032899928510 \end{bmatrix} \tag{2.14}$$

Du kan få fat i residualerne med kommandoen $residualer(X, Y, LogistReg)$ og du kan plotte residualerne med

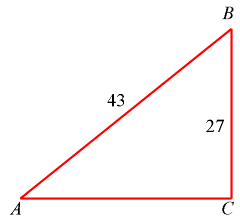
$plotResidualer(X, Y, LogistReg)$



3 Trigonometri

Maple funktionerne sin, cos og tan regner i radianer, hvilket ikke er hensigtsmæssigt i forbindelse med trigonometri. For at slippe for at konvertere gradmål for vinkler til radianer er Gym-pakken udstyret med funktionerne Sin, Cos og Tan, der regner i grader. Desuden er Gym-pakken udstyret med de inverse funktioner invSin, invCos og invTan, der virker som på en lommeregner.

3.1 Retvinklet trekant

I den retvinklede trekant ABC er $C = 90°$, $c = 43$ og $a = 27$ Beregn A, B og b.	

Det nemmeste er at opskrive udtrykket for $Sin(A)$, og med det samme indsætte de givne sider:

$$Sin(A) = \frac{27}{43} \xrightarrow{\text{solutions for A}} 38.89587133 \xrightarrow{\text{assign to a name}} A$$

Dvs., at $\#A = 38.9\,°$

I stedet kan vi naturligvis benytte *invSin*, men det er nemmere at lade Maple tage sig af omskrivningerne:

$$invSin\left(\frac{27}{43}\right) = 38.89587132$$

Tilsvarende bestemmes vinkel B:

$$Cos(B) = \frac{27}{43} \xrightarrow{\text{solutions for B}} 51.10412867 \xrightarrow{\text{assign to a name}} B$$

Dvs., at $\#B = 51.1\,°$

Tilbage er blot at bestemme b:

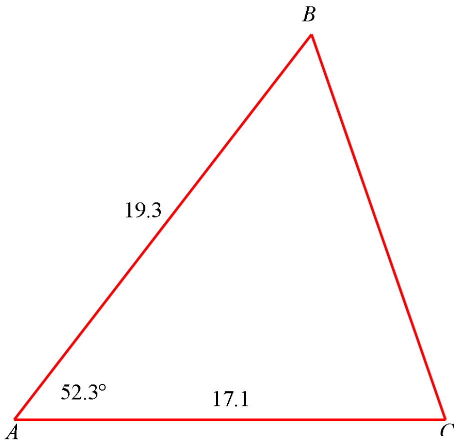
$$Cos(A) = \frac{b}{43} \xrightarrow{\text{solutions for b}} 33.46640106 \xrightarrow{\text{assign to a name}} b$$

Længden af siden b er således $b = 33.46$

3.2 Vilkårlig trekant

Beregn de ubekendte stykker i trekant ABC når

$A = 52.3°$, $b = 17.1$ og $c = 19.3$.



Start med at skrive cosinusrelationen op, hvor du straks indsætter de kendte værdier (måske skal du rense variablen a inden du starter)

$$a^2 = 19.3^2 + 17.1^2 - 2 \cdot 19.3 \cdot 17.1 \cdot Cos(52.3) \xrightarrow{\text{solutions for a}} 16.16339884, -16.16339884 \xrightarrow{\text{select entry 1}} 16.16339884 \xrightarrow{\text{assign to a name}} a$$

Til bestemmelse af vinklerne kan du bruge såvel sinusrelationen som cosinusrelationen, men det sikreste er at bruge cosinusrelationen, da den kun giver én brugbar løsning:

$$Cos(C) = \frac{a^2 + 17.1^2 - 19.3^2}{2\,a \cdot 17.1} \xrightarrow{\text{solutions for C}} 70.86783053 \xrightarrow{\text{assign to a name}} C$$

Tilbage er blot at beregne vinkel B:

$$B := 180 - 52.3 - C = 56.83216947$$

Hvis du foretrækker at definere de kendte sider og vinkler ved deres navne, og derved bruge de generelle udgaver af sinus- og cosinusrelationerne, er der intet til hinder for det. Husk blot altid at tjekke, at de variabler, du vil løse med hensyn til, ikke allerede er tildelt værdier. Kig i paletten 'Variables', højreklik på den variabel, du vil slette, og vælg 'unasign'.

Du finder mange flere eksempler i online hjælpen.

4 Vektorregning

4.1 2D-vektorer

To vektorer $\vec{a}$ og $\vec{b}$ er defineret ved

$\vec{a} := \langle 2, 3 \rangle :$

$\vec{b} := \langle -5, 4 \rangle :$

Længden af $\vec{a}$	$len(\vec{a}) = \sqrt{13}$
Vinklen mellem $\vec{a}$ og $\vec{b}$	$vinkel(\vec{a}, \vec{b}) = 85.03025926$
Projektionen af $\vec{a}$ på $\vec{b}$	$proj(\vec{a}, \vec{b}) = \begin{bmatrix} -\frac{10}{41} \\ \frac{8}{41} \end{bmatrix}$
Arealet af det parallelogram, der udspændes af $\vec{a}$ og $\vec{b}$:	$arealP(\vec{a}, \vec{b}) = 23$
Arealet af den trekant, der udspændes af $\vec{a}$ og $\vec{b}$:	$arealT(\vec{a}, \vec{b}) = \frac{23}{2}$
$\vec{a}$ skrevet på **polær form**	$cart2pol(\vec{a}) = [\sqrt{13}, 56.309]$
$\vec{c} := [2, 30] = [2, 30]$ skrevet med **cartesiske koordinater**	$pol2cart(\vec{c}) = \begin{bmatrix} \sqrt{3} \\ 1 \end{bmatrix}$
Tværvektoren til $\vec{a}$	$hat(\vec{a}) = \begin{bmatrix} -3 \\ 2 \end{bmatrix}$
Determinanten af $\vec{a}$ og $\vec{b}$	$det(\vec{a}, \vec{b}) = 23$
Enhedsvektor ensrettet med $\vec{a}$	$ev(\vec{a}) = \begin{bmatrix} \frac{2}{13}\sqrt{13} \\ \frac{3}{13}\sqrt{13} \end{bmatrix}$

Skalarproduktet mellem to vektorer kan udregnes som $\vec{a}.\vec{b}$, men denne implementation benytter det komplekse skalarprodukt, og kan give utilsigtede resultater i symbolske beregninger.

Fx fås:

$\vec{a} := \langle 2, t \rangle : \vec{b} := \langle -5\,t, 4 \rangle$

$\vec{a}.\vec{b} = -10\,t + 4\,\overline{t}$

der med det almindelige skalarprodukt skulle give $-6\,t$. Til at håndtere den slags situationer indeholder Gym-pakken funktionen *dotP*:

$dotP(\vec{a}, \vec{b}) = -6\,t$

4.2 3D-vektorer

To vektorer $\vec{a}$ og $\vec{b}$ er defineret ved

$\vec{a} := \langle 2, 3, -7 \rangle$:

$\vec{b} := \langle -5, 4, 3 \rangle$:

Længden af $\vec{a}$	$len(\vec{a}) = \sqrt{62}$
Vinklen mellem $\vec{a}$ og $\vec{b}$	$vinkel(\vec{a}, \vec{b}) = 109.9530536$
Projektionen af $\vec{a}$ på $\vec{b}$	$proj(\vec{a}, \vec{b}) = \begin{bmatrix} \frac{19}{10} \\ -\frac{38}{25} \\ -\frac{57}{50} \end{bmatrix}$
Arealet af det parallelogram, der udspændes af $\vec{a}$ og $\vec{b}$:	$arealP(\vec{a}, \vec{b}) = \sqrt{2739}$
Arealet af den trekant, der udspændes af $\vec{a}$ og $\vec{b}$:	$arealT(\vec{a}, \vec{b}) = \frac{1}{2}\sqrt{2739}$
Enhedsvektor ensrettet med $\vec{a}$	$ev(\vec{a}) = \begin{bmatrix} \frac{1}{31}\sqrt{62} \\ \frac{3}{62}\sqrt{62} \\ -\frac{7}{62}\sqrt{62} \end{bmatrix}$

4.3 Løsning af vektorligninger (vsolve)

$\vec{a} := \langle 1, 2 \rangle : \vec{b} := \langle -1, 3 \rangle : \vec{c} := \langle -1, 13 \rangle$:

Løsning af ligningen

$s \cdot \vec{a} + t \cdot \vec{b} = \vec{c}$

$$\begin{bmatrix} s - t \\ 2s + 3t \end{bmatrix} = \begin{bmatrix} -1 \\ 13 \end{bmatrix} \tag{4.1}$$

sker med kommandoen

$vsolve\left(s \cdot \vec{a} + t \cdot \vec{b} = \vec{c}, \{s, t\}\right)$

$$\{s = 2, t = 3\} \tag{4.2}$$

Når elementerne er tal, kan parameterlisten udelades uden problemer:

$vsolve\left(s \cdot \vec{a} + t \cdot \vec{b} = \vec{c}\right)$

$$\{s = 2, t = 3\} \tag{4.3}$$

Er vektorerne parametriserede, fx

$\vec{a} := \langle x, 2\rangle : \vec{b} := \langle -1, 3\rangle : \vec{c} := \langle -1, 13\rangle :$

giver *vsolve* uden parameterliste

$vsolve(s\cdot\vec{a} + t\cdot\vec{b} = \vec{c})$

$$\left\{s = \frac{10}{3x+2}, t = \frac{13x+2}{3x+2}, x = x\right\} \tag{4.4}$$

og med parameterlisten

$vsolve(s\cdot\vec{a} + t\cdot\vec{b} = \vec{c}, \{s, t\})$

$$\left\{s = \frac{10}{3x+2}, t = \frac{13x+2}{3x+2}\right\} \tag{4.5}$$

For 3D-vektorer forholder det sig tilsvarende

$\vec{a} := \langle 1, 2, -3\rangle : \vec{b} := \langle -1, 3, 7\rangle : \vec{c} := \langle -1, 13, 15\rangle :$

$vsolve(s\cdot\vec{a} + t\cdot\vec{b} = \vec{c}, \{s, t\})$

$$\{s = 2, t = 3\} \tag{4.6}$$

$vsolve(s\cdot\vec{a} + t\cdot\vec{b} = \vec{c})$

$$\{s = 2, t = 3\} \tag{4.7}$$

Er vektorerne parametriserede, fx

$\vec{a} := \langle x, 2, -3\rangle : \vec{b} := \langle -1, 3, 7\rangle : \vec{c} := \langle -1, 13, 15\rangle :$

giver *vsolve* uden parameterlisten

$vsolve(s\cdot\vec{a} + t\cdot\vec{b} = \vec{c})$

$$\{s = 2, t = 3, x = 1\} \tag{4.8}$$

og med parameterlisten får vi intet resultat, idet løsningen ikke kan parametriseres mht. x

$vsolve(s\cdot\vec{a} + t\cdot\vec{b} = \vec{c}, \{s, t\})$

Vektorerne kan naturligvis også indtastes direkte:

$vsolve(s\cdot\langle 1, 2, -3\rangle + t\cdot\langle -1, 3, 7\rangle = \langle -1, 13, 15\rangle)$

$$\{s = 2, t = 3\} \tag{4.9}$$

Skæring mellem linjer i rummet

I forbindelse med skæring mellem linjer givet ved parameterfremstillinger er *vsolve* meget nytig:

Hvis to parameterfremstillinger er givet ved

$l := t \rightarrow \langle 3, 0, -2\rangle + t\cdot\langle 3, 2, 3\rangle$

$$t \rightarrow \langle 3, 0, -2\rangle + t\,\langle 3, 2, 3\rangle \tag{4.10}$$

$m := s \rightarrow \langle 4, 2, 1\rangle + s\cdot\langle 1, 2, 3\rangle$

$$s \rightarrow \langle 4, 2, 1\rangle + s\,\langle 1, 2, 3\rangle \tag{4.11}$$

så findes skæringspunktet ved

$vsolve(l(t) = m(s))$

$$\{s = -1, t = 0\} \tag{4.12}$$

Skæring mellem planer (givet ved parameterfremstillinger)

$p := (s, t) \to \langle 0, -2, 0\rangle + s \cdot \langle -4, 0, 4\rangle + t \cdot \langle 0, -2, 4\rangle$

$$(s, t) \to \langle 0, -2, 0\rangle + s\langle -4, 0, 4\rangle + t\langle 0, -2, 4\rangle \tag{4.13}$$

$q := (u, v) \to \langle 0, 0, 3\rangle + u \cdot \langle -2, 6, 0\rangle + v \cdot \langle -2, 0, 3\rangle$

$$(u, v) \to \langle 0, 0, 3\rangle + u\langle -2, 6, 0\rangle + v\langle -2, 0, 3\rangle \tag{4.14}$$

$vsolve(p(s, t) = q(u, v))$

$$\left\{s = s, t = \frac{4}{3} + \frac{2}{3}\, s, u = -\frac{2}{9}\, s - \frac{7}{9}, v = \frac{20}{9}\, s + \frac{7}{9}\right\} \tag{4.15}$$

Parameterfremstillingen for skæringslinjen findes ved at indsætte s og t i p - eller u og v i q:

$p\left(s, \frac{4}{3} + \frac{2}{3}\, s\right)$

$$\begin{bmatrix} -4\, s \\ -\frac{14}{3} - \frac{4}{3}\, s \\ \frac{20}{3}\, s + \frac{16}{3} \end{bmatrix} \tag{4.16}$$

$q\left(-\frac{2}{9}\, s - \frac{7}{9}, \frac{20}{9}\, s + \frac{7}{9}\right)$

$$\begin{bmatrix} -4\, s \\ -\frac{14}{3} - \frac{4}{3}\, s \\ \frac{20}{3}\, s + \frac{16}{3} \end{bmatrix} \tag{4.17}$$

5 Statistik

5.1 Fordelinger

Normalfordeling

Gym-pakken indeholder hertil fordelingsfunktionen *normalcdf*, tæthedsfunktionen *normalpdf* og *invnorm* til bestemmelse af fraktiler i normalfordelingen.

Funktionerne har middelværdi og spredning som valgfrie parametre, og udelades disse, er standardværdierne $\mu = 0$ og $\sigma = 1$. Fx er

$normalcdf(1) = 0.8413447460$

og er der tale om en normalfordeling med $\mu = 100$ og $\sigma = 10$, findes værdien i fx 105 ved

$normalcdf(100, 10, 105) = 0.6914624612$

- hvilket naturligvis også kan bestemmes som

$$\int_{-\infty}^{105} normalpdf(100, 10, t)\, dt = 0.6914624609$$

For at besteme 95%-fraktilen i denne fordeling benyttes

$invnorm(100, 10, 0.95) = 116.448536269521$

t-fordelingen

Gym-pakken indeholder hertil fordelingsfunktionen *tcdf*, tæthedsfunktionen *tpdf* og *invt* til bestemmelse af fraktiler i t-fordelingen.

Alle funktioner kræver antallet af frihedsgrader som parameter. Fx findes værdien i 0.7 af t-fordelingen med 5 frihedsgrader således

$tcdf(5, 0.7) = 0.742425525842592$

- hvilket naturligvis også kan bestemmes som

$$\int_{-\infty}^{0.7} tpdf(5, x)\, dx = 0.7424255256$$

For at besteme 95%-fraktilen i denne fordeling benyttes

$invt(5, 0.95) = 2.01504256032808$

χ^2-fordelingen

Gym-pakken indeholder hertil fordelingsfunktionen *chicdf*, tæthedsfunktionen *chipdf* og *invchi* til bestemmelse af fraktiler i χ^2-fordelingen.

Alle funktioner kræver antallet af frihedsgrader som parameter. Fx findes værdien i 4 af χ^2-fordelingen med 5 frihedsgrader således

$chicdf(5, 4) = 0.4505840491$

- hvilket naturligvis også kan bestemmes som

$$\int_{-\infty}^{4} chipdf(5, x)\, dx = 0.4505840485$$

For at besteme 95%-fraktilen i denne fordeling benyttes

$invchi(5, 0.95) = 11.0704974062099$

5.2 Konfidensintervaller

Konfidensinterval for μ i normalfordelingen (σ kendt)

obs := [229.4, 229.7, 230.2, 230.2, 232, 231.2, 230, 230.6, 230, 229.4, 230.9, 228.5, 231.5, 230.9, 231.2, 227.9, 230.6, 232, 230.3, 232.3] :

Det vides, at observationerne er normalfordelte med spredning $\sigma = 1.5$.

95%-konfidensintervallet for middelværdien kan bestemmes ved:

zInterval(*obs*, 1.5, 0.95)

$$[229.782608094715, 231.097391905285] \tag{5.1}$$

For at kunne benytte 4-parameter versionen, skal antallet af observationer, et estimat for middelværdien og spredningen være kendt. Hvis fx $\overline{x} = 267$, $n = 400$ og $\sigma = 50$, kan 95%-konfidensintervallet for middelværdien bestemmes ved

zInterval(267, 50, 400, 0.95)

$$[262.100090038651, 271.899909961349] \tag{5.2}$$

Konfidensinterval for μ i normalfordelingen (σ ukendt)

obs := [5, 4.4, 5.7, 5.6, 5.5, 5.2, 5.0, 4.8, 3.6, 4.1, 4.6, 4.9, 4.0, 6.7, 5.5, 5.4, 6.7, 5.8, 5.4, 4.8, 5.9, 5.1, 3.8, 4.1, 6.7] :

Det vides, at observationerne er normalfordelte. 95%-konfidensintervallet for middelværdien kan bestemmes ved:

tInterval(*obs*, 0.95)

$$[4.77418085650167, 5.48981914349834] \tag{5.3}$$

For at kunne benytte 4-parameter versionen, skal antallet af observationer og estimater for middelværdi og spredning være kendte. Hvis fx $\overline{x} = 120$, $s = 15$ og $n = 68$, kan 95%-konfidensintervallet for middelværdien bestemmes ved

tInterval(120, 15, 68, 0.95)

$$[116.369226496765, 123.630773503235] \tag{5.4}$$

Konfidensinterval for andelen p i binomialfordelingen

Antag at vi en stikprøve på n = 180 personer har x = 40 personer der er imod højere skat. Et konfidensinterval for andelen p af personer, der er imod højere skat beregnes *zInterval1Andel*(40, 180, 0.95)

$$[0.161488017783562, 0.282956426616438] \tag{5.5}$$

I stedet for at benytte antallet af succes'er som input, kan andelen p benyttes.

$$zInterval1Andel\left(\frac{40}{180}, 180, 0.95\right)$$

$$[0.161488017783562, 0.282956426616438] \tag{5.6}$$

5.3 z-test for én middelværdi (spredning kendt)

obs := [229.4, 229.7, 230.2, 230.2, 232, 231.2, 230, 230.6, 230, 229.4, 230.9, 228.5, 231.5, 230.9, 231.2, 227.9, 230.6, 232, 230.3, 232.3] :

Det vides, at observationerne er normalfordelte med med spredning $\sigma = 1.5$. Test af hypotesen

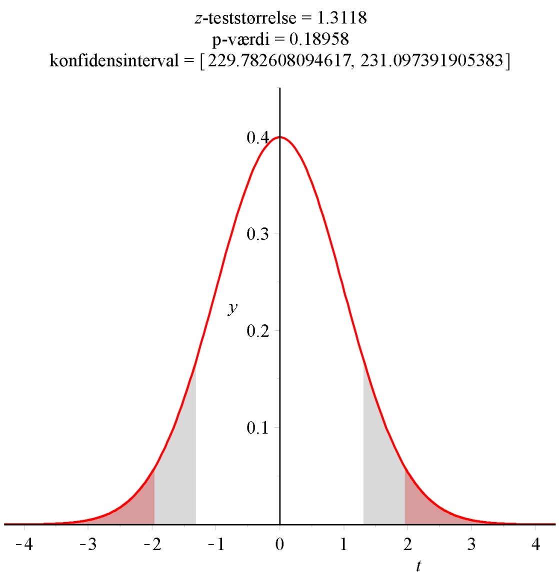
H_0: $\mu = 230$

mod alternativet

H_0 : $\mu \neq 230$

på niveau 5%, sker med kommandoen

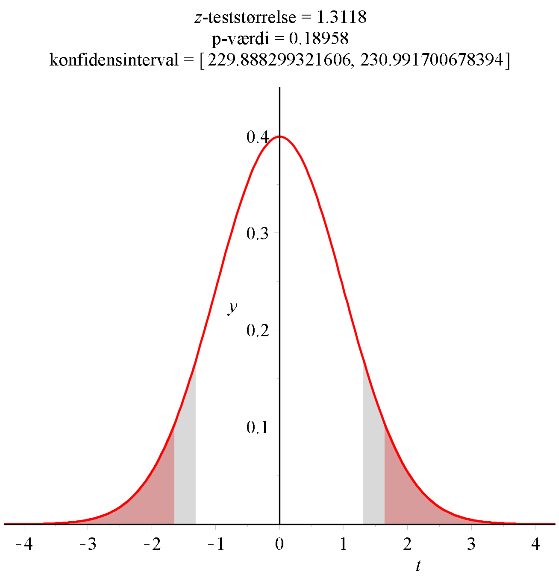
zTest(*obs*, 230, 1.5)



Den kritiske mængde og p-værdien vises som skaverede områder under standard normalfordelingsgrafen.

Som standard er konfidensniveauet sat til til 95%. Ønskes dette ændret til fx 90%, tilføjes paramereteren $\mathit{konfidens} = 0.9$:

$\mathit{zTest}(\mathit{obs}, 230, 1.5, \mathit{konfidens} = 0.9)$



5.4 t-test for én middelværdi (spredning ukendt)

$\mathit{obs} := [5, 4.4, 5.7, 5.6, 5.5, 5.2, 5.0, 4.8, 3.6, 4.1, 4.6, 4.9, 4.0, 6.7, 5.5, 5.4, 6.7, 5.8, 5.4, 4.8, 5.9, 5.1, 3.8, 4.1, 6.7]:$

Det vides, at observationerne er normalfordelte. Test af hypotesen

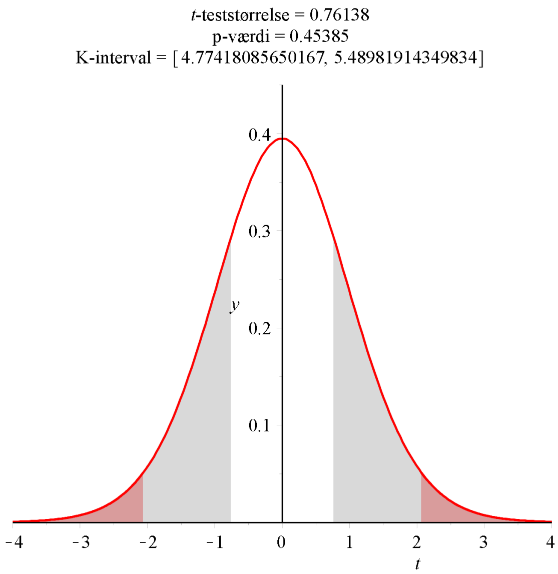
H_0: $\mu = 5$

mod alternativet

H_0 : $\mu \neq 5$

på niveau 5%, sker med kommandoen

$tTest(obs, 5)$

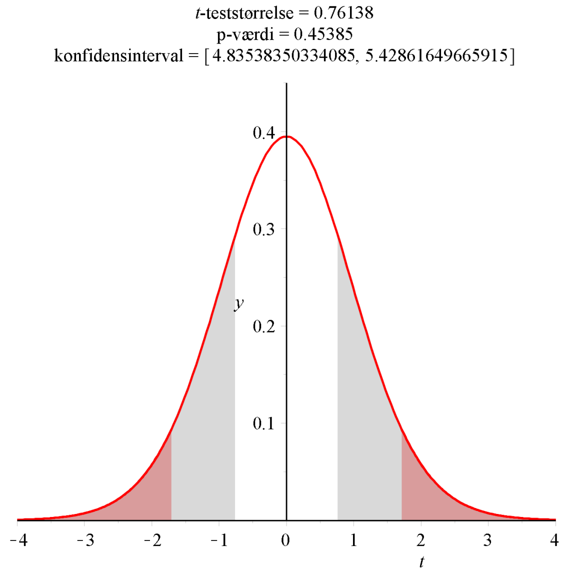


Den kritiske mængde og p-værdien vises som skaverede områder under standard normalfordelingsgrafen.

Som standard er konfidensniveauet sat til til 95%. Ønskes dette ændret til fx 90%, tilføjes paramereteren *konfidens* = 0.9:

(5.7)

$tTest(obs, 5, konfidens = 0.9)$



5.5 Chi^{2}test

Gym-pakken indeholder en række funktioner til anvendelse i forbindelse med χ^2 - test:

antalstabel

- en hjælpefunktion til brug i forbindelse med optælling af data efter inddelingskriterier. Typisk vil data hentes fra fx Excel.

$A := antalstabel(M)$

$$\begin{bmatrix} \text{"Observeret"} & \text{"Aviser eller blade"} & \text{"Internet eller mobil"} & \text{"Radio eller TV"} & \text{"i alt"} \\ \text{"Kvinde"} & 171 & 103 & 369 & 643 \\ \text{"Mand"} & 97 & 80 & 180 & 357 \\ \text{"i alt"} & 268 & 183 & 549 & 1000 \end{bmatrix} \quad (5.8)$$

forventet

- en hjælpefunktion til beregning af forventede værdier i en tabel under forudsætning af uafhængighed mellem inddelingskriterierne. Her benyttes antalstabellen som input. Har du ikke antalstabellen til rådighed, kan du istedet benytte matricen af observerede værdier

$$\begin{bmatrix} 171 & 103 & 369 \\ 97 & 80 & 180 \end{bmatrix}$$

forventet(A)

"Forventet"	"Aviser eller blade"	"Internet eller mobil"	"Radio eller TV"	"i alt"
"Kvinde"	172.32	117.67	353.01	643
"Mand"	95.676	65.331	195.99	357
"i alt"	268	183	549	1000

(5.9)

bidrag

- beregner de enkelte 'cellers' bidrag til χ^2-teststørrelsen. Her benyttes antalstabellen som input. Har du ikke antalstabellen til rådighed, kan du istedet benytte matricen af observerede værdier

$$\begin{bmatrix} 171 & 103 & 369 \\ 97 & 80 & 180 \end{bmatrix}$$

bidrag(A)

"Bidrag"	"Aviser eller blade"	"Internet eller mobil"	"Radio eller TV"	"i alt"
"Kvinde"	0.01017255867	1.828685219	0.7245636744	2.563421452
"Mand"	0.01832200343	3.293682341	1.305026450	4.617030794
"i alt"	0.02849456210	5.122367560	2.029590124	7.180452246

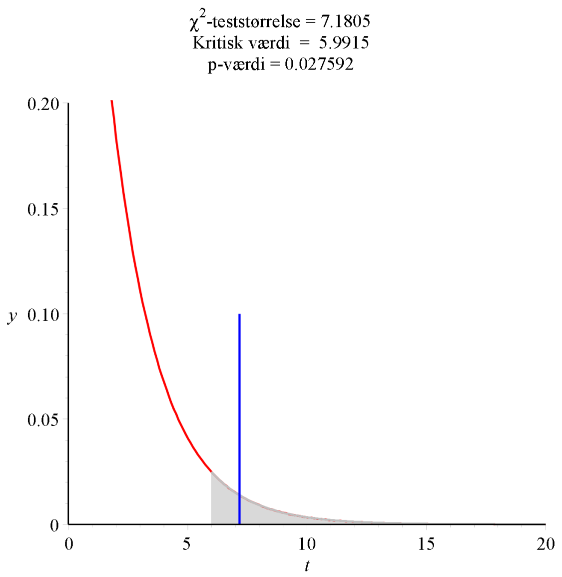
(5.10)

ChiKvadratUtest

beregner en χ^2 - test for uafhængighed i en matrix. Her benyttes antalstabellen som input. Har du ikke antalstabellen til rådighed, kan du istedet benytte matricen af observerede værdier

$$\begin{bmatrix} 171 & 103 & 369 \\ 97 & 80 & 180 \end{bmatrix}$$

ChiKvadratUtest(A)



ChiKvadratGOFtest

beregner en χ^2 - test for Godness of Fit mellem en observeret liste og en forventet liste:

$obs := [5, 4, 5, 6, 5, 13]:$

$forv := [6.33, 6.33, 6.33, 6.33, 6.33, 6.33]:$

$ChiKvadratGOFtest(obs, forv)$