9 Statistikk og modeller

9.1 Statistikk

> *restart* :

Maple siste og mest omfattende programpakke for statistikk er Statistics . Den hentes inn med

> *with*(*Statistics*) :

Her er en del statistikkommandoer

- Mean(L) beregner gjennomsnittet av tallene i listen L
- Mode(L) gir observasjonen med høyest hyppighet (frekvens) i listen L
- Range(L) gir differensen mellom høyeste og laveste verdi av tallene i en liste L
- Quartile(L, k) beregner kvartiler til tallene i listen L, k =1 (nedre kvartil), 2, 3 (øvre kvartil)
- Percentile (L, p) beregner prosentiler i listen L, der p er prosentilverdien
- Histogram (L, *valg*) fremstiller et histogram eller søylediagram
- Tally(L) grupperer tall som er like i listen L
- TallyInto(L, K) grupperer tallene i listen L i klasser med klassebredder gitt i listen K
- FrequencyTable(L) plukker ut hyppigheten av tallene i listen L
- StandardDeviation(L) beregner standardavviket for populasjonen gitt i listen L
- Sort(L) sorterer numeriske data gitt i listen L

Eksempel 9.2.1

Tallene i listen [6, 3, 6, 2, 5, 3, 5, 6, 6, 10, 4, 7, 4, 2, 3, 1, 12, 7, 2, 3, 3, 4, 10, 8, 2] angir hvor mange bokstaver som er i hvert ord i en tekst på til sammen 25 ord.

a) Lag en frekvenstabell.
b) Finn variasjonsbredden.
c) Finn hyppigste verdi og regn ut gjennomsnittet.
d) Tegn et stolpediagram.

Løsning

I all statistisk behandling av et tallmateriale i Maple kan vi skriver dataene opp på listeform.

> L := [6, 3, 6, 2, 5, 3, 5, 6, 6, 10, 4, 7, 4, 2, 3, 1, 12, 7, 2, 3, 3, 4, 10, 8, 2] :

a)

> *sort*(L)

$$[1, 2, 2, 2, 2, 3, 3, 3, 3, 3, 4, 4, 4, 5, 5, 6, 6, 6, 6, 7, 7, 8, 10, 10, 12]$$

En sjekk på at **antallet** er 25 får vi ved kommandoen nops.

> *nops*(L)

Gruppering eller opptelling av antallet av hvert tall fåes ved Tally.

> *Tally*(L)

$$[1 = 1, 2 = 4, 3 = 5, 4 = 3, 5 = 2, 6 = 4, 7 = 2, 8 = 1, 10 = 2, 12 = 1]$$

Venstre side av likhetstegnet viser de forskjellige tallene i liste. Høyre side viser antallet av hvert tall.

> *evalf*(*FrequencyTable*(L), 3)

$$\left[\begin{array}{ccccc} 1. .. 2.10 & 5. & 20.0 & 5. & 20.0 \\ 2.10 .. 3.20 & 5. & 20.0 & 10. & 40.0 \\ 3.20 .. 4.30 & 3. & 12.0 & 13. & 52.0 \\ 4.30 .. 5.40 & 2. & 8.00 & 15. & 60.0 \\ 5.40 .. 6.50 & 4. & 16.0 & 19. & 76.0 \\ 6.50 .. 7.60 & 2. & 8.00 & 21. & 84.0 \\ 7.60 .. 8.70 & 1. & 4.00 & 22. & 88.0 \\ 8.70 .. 9.80 & 0. & 0. & 22. & 88.0 \\ 9.80 .. 10.9 & 2. & 8.00 & 24. & 96.0 \\ 10.9 .. 12. & 1. & 4.00 & 25. & 100. \end{array}\right]$$

I vgs pakken fins FrekvensTabell.

> *FrekvensTabell*(L, 10, 3)

Klassebredde	*Antall*	*Prosent*	*Antall(kumulativ)*	⋯
1. ..2.10	5.	20.0	5.	⋯
2.10 ..3.20	5.	20.0	10.	⋯
3.20 ..4.30	3.	12.0	13.	⋯
4.30 ..5.40	2.	8.00	15.	⋯
5.40 ..6.50	4.	16.0	19.	⋯
6.50 ..7.60	2.	8.00	21.	⋯
7.60 ..8.70	1.	4.00	22.	⋯
8.70 ..9.80	0.	0.	22.	⋯
9.80 ..10.9	2.	8.00	24.	⋯
10.9 ..12.	1.	4.00	25.	⋯

>

Første kolonne: klassebredden, **andre kolonne:** antallet i hver klasse, **tredje kolonne:** antallet i prosent , **fjerde kolonne:** kumulativt antall, **femte kolonne:** kumulativ prosent

Variasjonsbredde, Typetall og Middelverdi fins i vgs-pakken,

b) Variasjonsbredden er

> *Range*(*L*)

$$11.$$

> *Variasjonsbredde*(*L*)

$$11., \mathit{Maksimum} = 12, \mathit{Minimum} = 1$$

c) Størst hyppighet er

> *Mode*(*L*)

$$3.$$

> *Typetall*(*L*)

$$[3], \mathit{hyppighet} = 5$$

Gjennomsnittsverdi

> *Mean*(*L*)

$$4.96000000000000$$

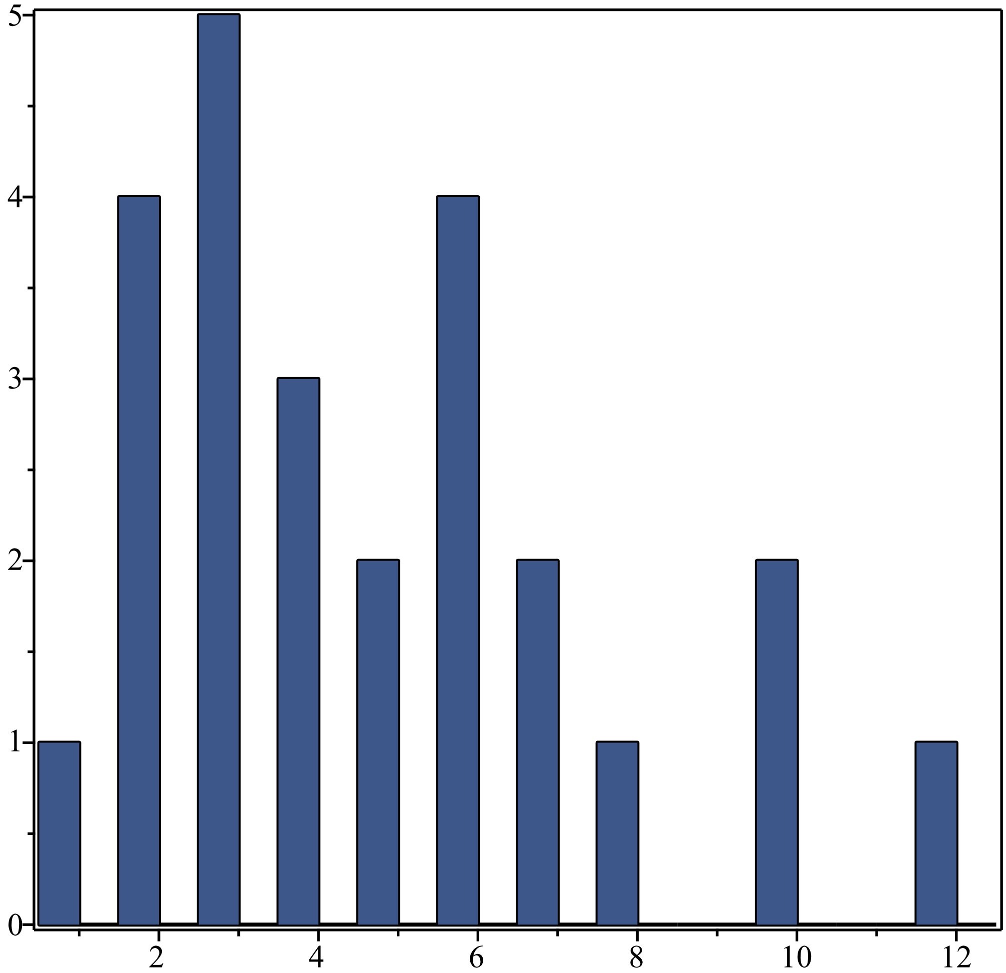
> $Middelverdi(L)$

$$4.96000000000000$$

d)

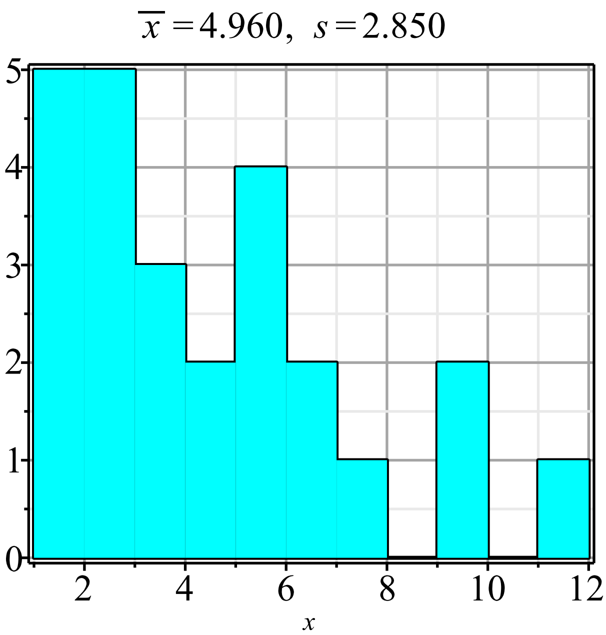
Et stolpediagram med 10 stolper, like mange som det er tall i listen gir følgende histogram.
StolpeDiagram er i vgs - pakken.

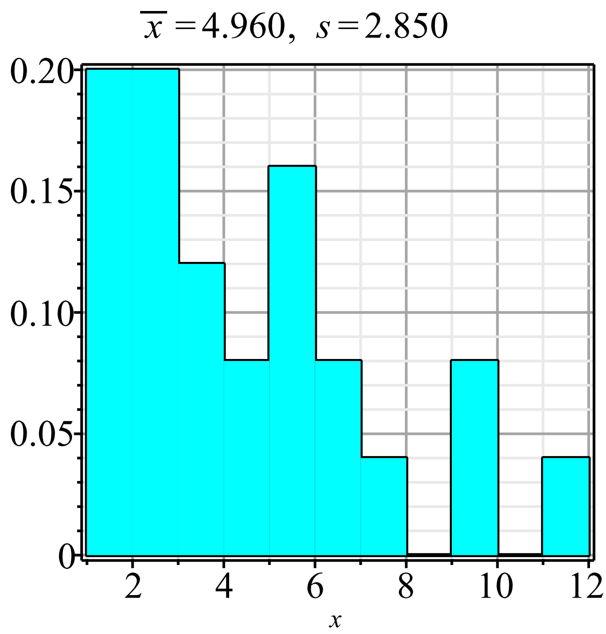
> $Histogram(L, frequencyscale = absolute, binwidth = 0.5)$



> $StolpeDiagram(L, \text{abs}, 11)$

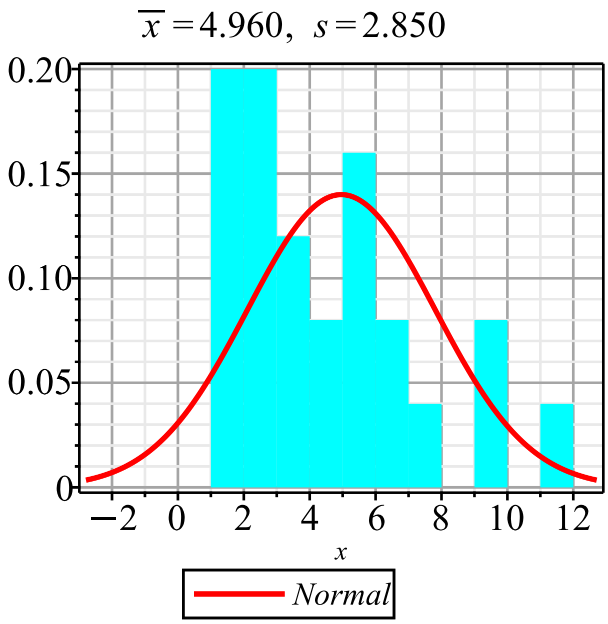
> $StolpeDiagram(L, \text{rel}, 11, \text{Ingen})$





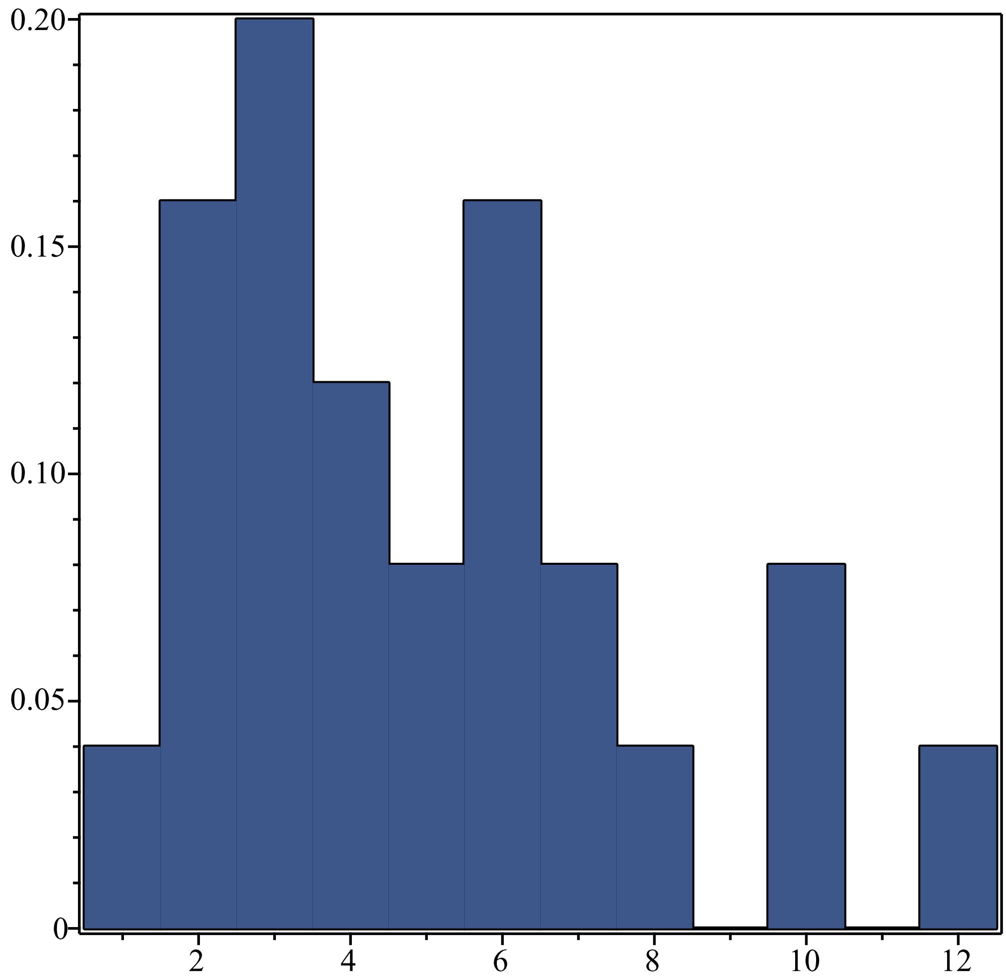
> *StolpeDiagram*(*L*, rel, 11, Normal)

>



Høyden på hver av søylene angir tallets hyppighet. Bredden på søylene angis ved **binwidth**.

> *Histogram*(*L*, *frequencyscale* = *relative*, *binwidth* = 1)



Med angivelsen **relative** er summen av søylenes areal lik 1

Eksempel 9.2.2
På en skole ble førti tolvåringer veid hos skolelegen. Resultatet fremgår av følgende liste:
[53, 52, 54, 56, 50, 49, 52, 50, 57, 52, 55, 60, 51, 43, 53, 64, 57, 58, 57, 50, 56, 50, 55, 50, 59, 42, 55, 45, 45, 52, 46, 47, 46, 58, 48, 50, 55, 47, 53, 51]

a) Klassedel materialet fra 42 til 65 og med klassebredde på 3.
b) Beregn gjennomsnittet, medianen og standardavviket.
c) Fremstill et histogram
d) Beregn kvartiler og noen prosentiler.

Løsning

a)

> $L := [53, 52, 54, 56, 50, 49, 52, 50, 57, 52, 55, 60, 51, 43, 53, 64, 57, 58, 57, 50, 56, 50, 55, 50, 59, 42, 55, 45, 45, 52, 46, 47, 46, 58, 48, 50, 55, 47, 53, 51]:$

> $sort(L)$

$[42, 43, 45, 45, 46, 46, 47, 47, 48, 49, 50, 50, 50, 50, 50, 50, 51, 51, 52, 52, 52, 52, 53, 53, 53, 54, 55, 55, 55, 55, 56, 56, 57, 57, 57, 58, 58, 59, 60, 64]$

> $Range(L)$

$$22.$$

> $Variasjonsbredde(L)$

$$22., Maksimum = 64, Minimum = 42$$

Med klassebredde på 3 får vi 7 klasser.

> $TallyInto(L, default, bins = 7)$

$[42.0000000000000 .. 45.1428571428571 = 4, 45.1428571428571 .. 48.2857142857143 = 5, 48.2857142857143 .. 51.4285714285714 = 9, 51.4285714285714 .. 54.5714285714286 = 8, 54.5714285714286 .. 57.7142857142857 = 9, 57.7142857142857 .. 60.8571428571429 = 4, 60.8571428571429 .. 64.0000000000000 = 1]$

> $evalf(\%, 3)$

$[42.0 .. 45.1 = 4., 45.1 .. 48.3 = 5., 48.3 .. 51.4 = 9., 51.4 .. 54.6 = 8., 54.6 .. 57.7 = 9., 57.7 .. 60.9 = 4., 60.9 .. 64.0 = 1.]$

b)

> $\mu = Mean(L), M = Median(L), \sigma = StandardDeviation(L)$

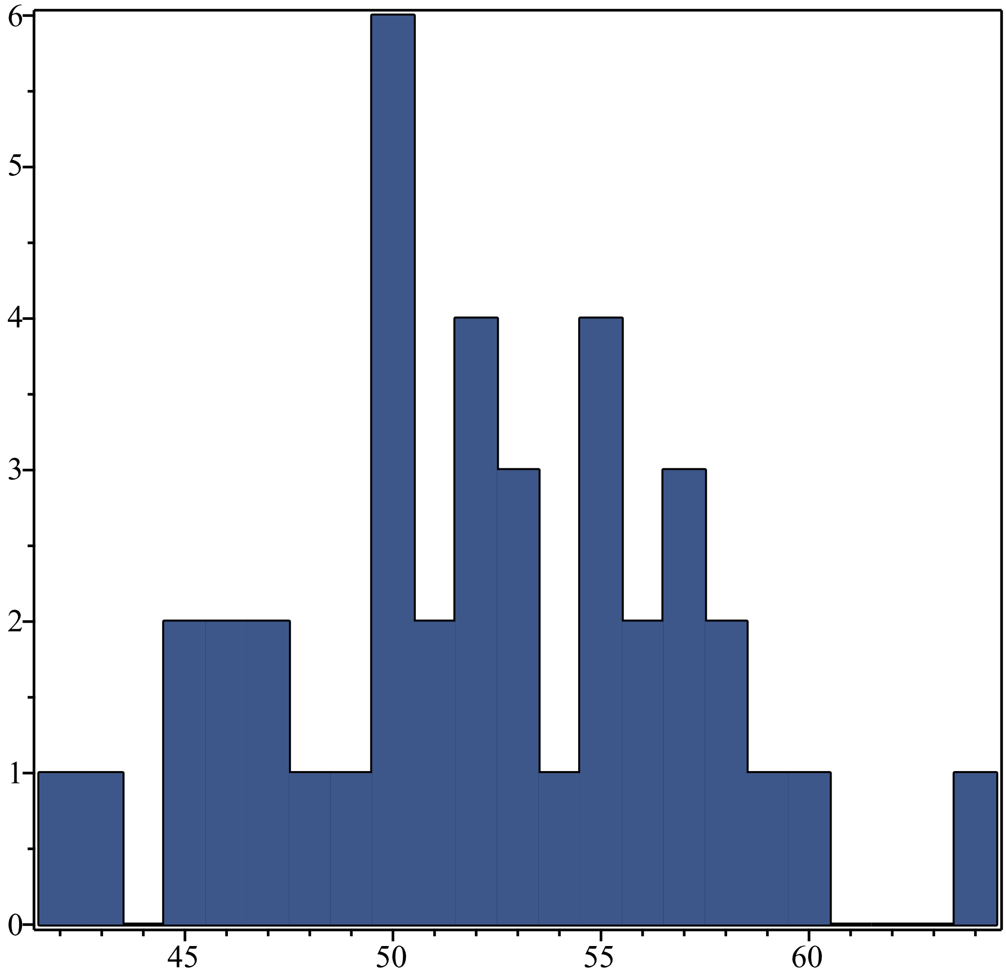
$$\mu = 52.0750000000000, M = 52., \sigma = 4.88528664880840$$

> $\mathrm{mu} = Middelverdi(L), \mathrm{sigma} = Standardavvik(L)$

$$\mu = 52.0750000000000, \sigma = 4.88528664880840$$

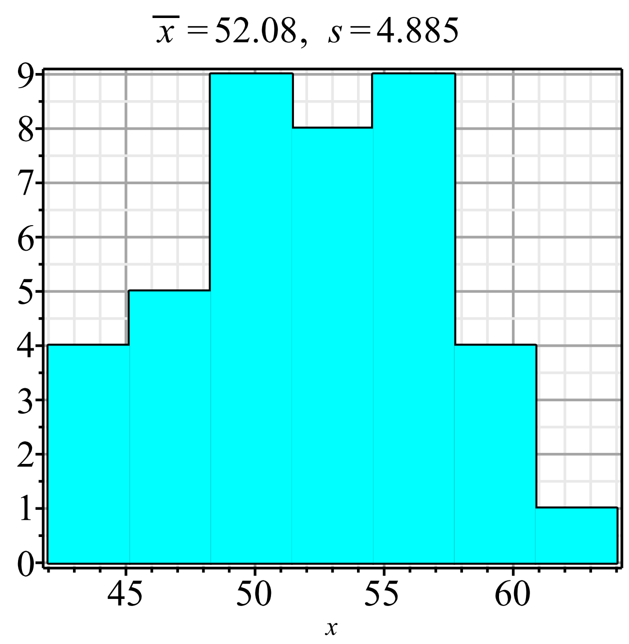
c)

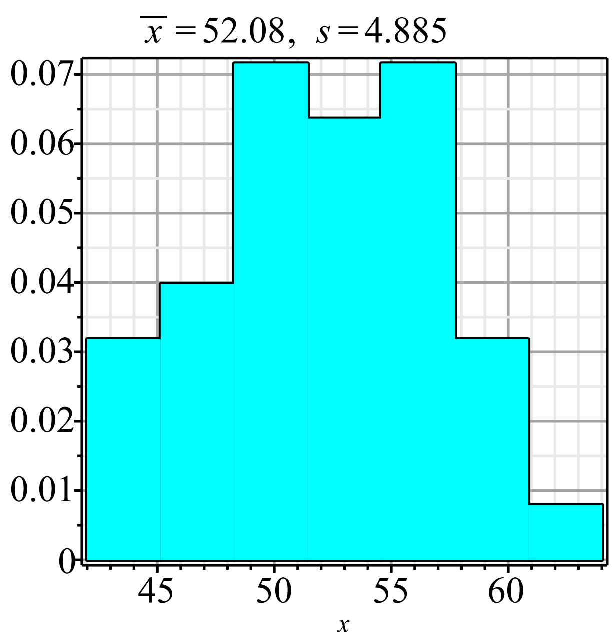
> $Histogram(L, frequencyscale = absolute, binwidth = 1)$



> $StolpeDiagram(L, \mathrm{abs}, 7)$

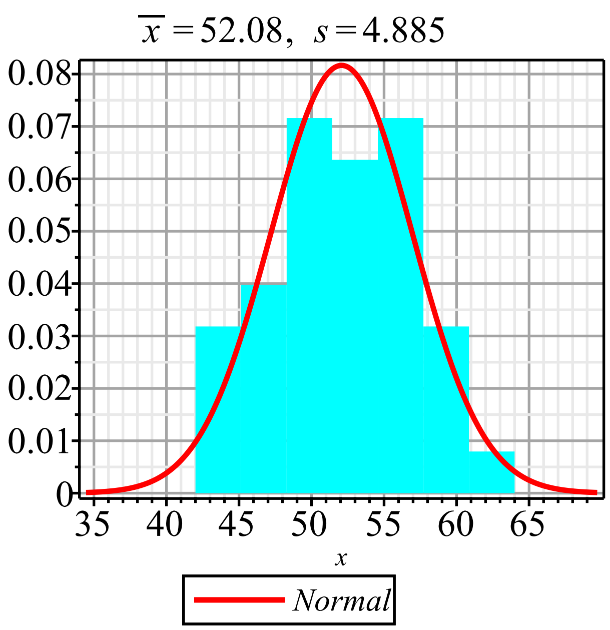
> $StolpeDiagram(L, \mathrm{rel}, 7, \mathrm{Ingen})$





> *StolpeDiagram*(*L*, rel, 7, Normal)

>



> *FrekvensTabell*(*L*, 7, 3)

$$\begin{bmatrix} Klassebredde & Antall & Prosent & Antall(kumulativ) & \cdots \\ 42...45.1 & 4. & 10.0 & 4. & \cdots \\ 45.1..48.3 & 5. & 12.5 & 9. & \cdots \\ 48.3..51.4 & 9. & 22.5 & 18. & \cdots \\ 51.4..54.6 & 8. & 20.0 & 26. & \cdots \\ 54.6..57.7 & 9. & 22.5 & 35. & \cdots \\ 57.7..60.9 & 4. & 10.0 & 39. & \cdots \\ 60.9..64. & 1. & 2.50 & 40. & \cdots \end{bmatrix}$$

Kvartiler

> $seq(Q_i = Quartile(L, i), i = 1..3)$

$$Q_1 = 49.4166666666667, Q_2 = 52., Q_3 = 55.5833333333333$$

> $seq(Kvartil(L, \mathrm{i}), i = 1..3)$

$$49.4166666666667, 52., 55.5833333333333$$

>

Prosentiler

> $seq(P_{5\,i} = Percentile(L, 5\,i), i = 1..19)$

$$P_5 = 43.7000000000000, P_{10} = 45.3666666666667, P_{15} = 46.3833333333333, P_{20} = 47.4000000000000, P_{25} = 49.4166666666667, P_{30} = 50., P_{35} = 50., P_{40} = 50.4666666666667, P_{45} = 51.4833333333333, P_{50} = 52., P_{55} = 52.5166666666667, P_{60} = 53., P_{65} = 54.5500000000000, P_{70} = 55., P_{75} = 55.5833333333333, P_{80} = 56.6000000000000, P_{85} = 57., P_{90} = 58., P_{95} = 59.6500000000000$$

> $Prosentil(L, 95)$

$$59.6500000000000$$

>

9.2 Sannsynlighetsfordelinger

Statistikkpakken Statistics inneholder 38 forskjellige sannsynlighetsfordelinger.
Kommandoen RandomVariable genererer en tilfeldig variable med spesifisert fordeling.
PDF beregner sannsynlighets tetthetsfunksjonen og CDF den kumulative fordelingsfunksjonen.

Normalfordeling

> $X := (\text{mu}, \text{sigma}, t) \rightarrow RandomVariable(Normal(\text{mu}, \text{sigma}))$:

> $N := (\text{mu}, \text{sigma}, t) \rightarrow PDF(X(\text{mu}, \text{sigma}), t)$:

> $'N(\mu, \sigma, t)' = N(\mu, \sigma, t)$

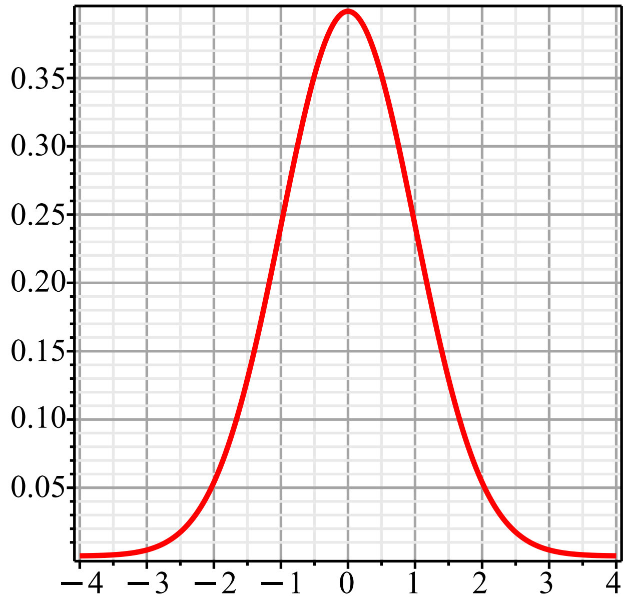
$$N(\mu, \sigma, t) = \frac{\sqrt{2}\, \mathrm{e}^{-\frac{(t-\mu)^2}{2\sigma^2}}}{2\sqrt{\pi}\,\sigma}$$

> $Mean(X(\mu, \sigma, t)),$
$StandardDeviation(X(\mu, \sigma, t))$

$$\mu, \sigma$$

> $p1 := DensityPlot(X(0, 1, t), range = -4 ..4, thickness = 2, color = red, gridlines)$:

%



> $G := (\text{mu}, \text{sigma}, t) \rightarrow CDF(X(\text{mu}, \text{sigma}), t)$:

> $p2 := plot(G(0, 1, t), t = -4 ..4, thickness = 2, color = blue, legend = \text{"Kumulativ fordeling"})$:

> $'G(0, 1, x)' = G(0, 1, x)$

$$G(0, 1, x) = \frac{1}{2} + \frac{\mathrm{erf}\left(\frac{x\sqrt{2}}{2}\right)}{2}$$

som fremkommer av

> $\int_{-\infty}^{x} N(0, 1, t)\, \mathrm{d}t$

$$\int_{-\infty}^{x} \frac{\sqrt{2}\, \mathrm{e}^{-\frac{t^2}{2}}}{2\sqrt{\pi}}\, \mathrm{d}t$$

> $value(\%)$

$$\frac{1}{2} + \frac{\mathrm{erf}\left(\frac{x\sqrt{2}}{2}\right)}{2}$$

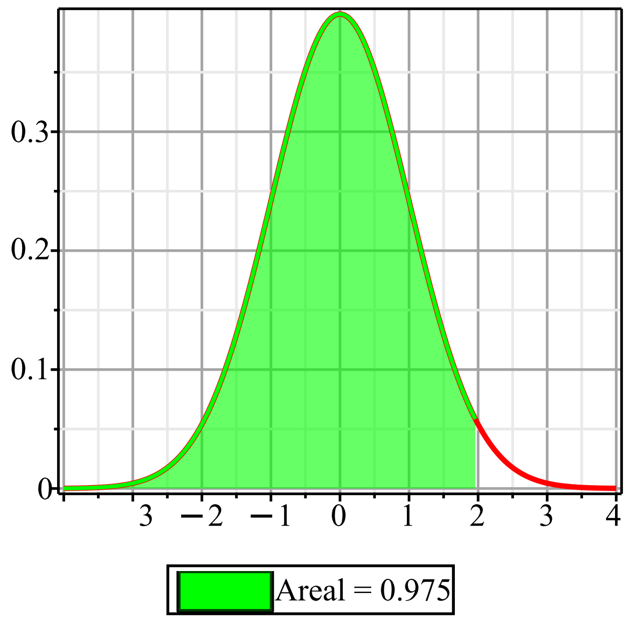
>

> $'G(0, 1, 1.96)' = G(0, 1, 1.96)$

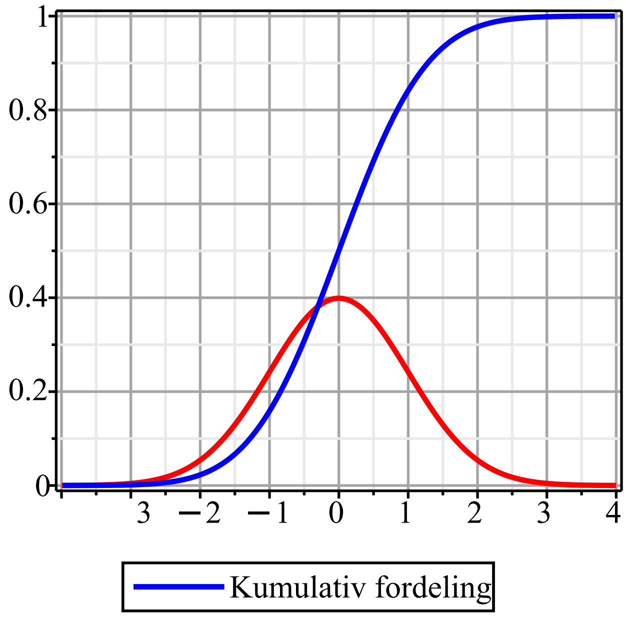
$$G(0, 1, 1.96) = 0.975002104851780$$

> $p3 := plot(N(0, 1, x), x = -4 ..1.96, color = green, filled = true, legend = \text{"Areal = 0.975"})$:

> $plt := display(p1, p3)$: %



> $display(p1, p2)$



> $G(0, 1, 1.0) - G(0, 1, -1.0)$

0.682689492137086

> $G(0, 1, 1.96) - G(0, 1, -1.96)$

0.950004209703559

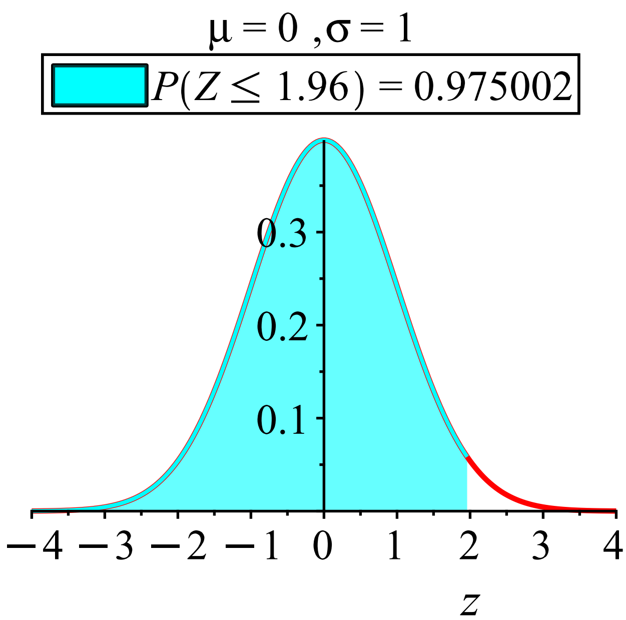
> $1 - G(0, 1, 1.96)$

0.0249978951482205

> $-Quantile(X(0, 1), 0.0249978951)$

1.96000000082450

> $NormalFordeling(0, 1, 1.96, venstre)$



> $K := \text{alpha} \rightarrow -Quantile(X(0, 1), \text{alpha}) :$

Noen kvantilverdier

> $z_{0.001} = K(0.001), z_{0.005} = K(0.005), z_{0.01} = K(0.01), z_{0.025} = K(0.025), z_{0.05} = K(0.05), z_{0.1} = K(0.1)$

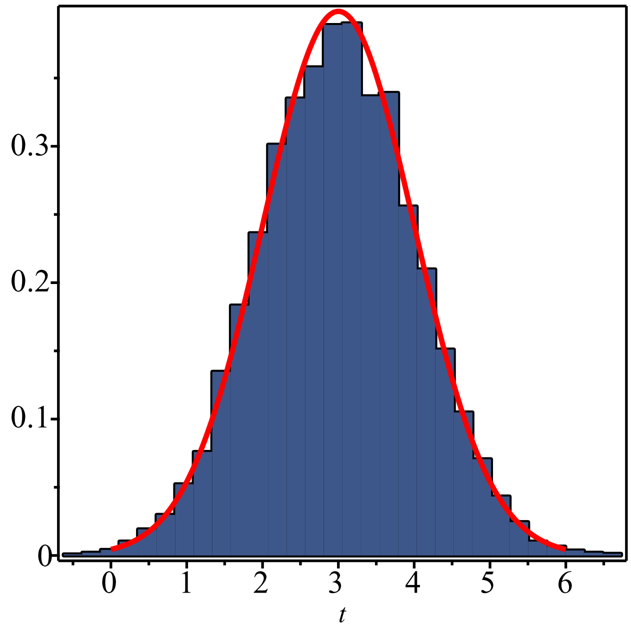
$z_{0.001} = 3.09023230616741, z_{0.005} = 2.57582930355009, z_{0.01} = 2.32634787404074, z_{0.025} = 1.95996398453944, z_{0.05} = 1.64485362695213, z_{0.1} = 1.28155156554473$

> $S := Sample(Normal(3, 1), 10000) :$
$p4 := plot(N(3, 1, t), t = 0..6, color = red, thickness = 2) :$

> $p5 := Histogram(S) :$

> $display(p4, p5)$

>



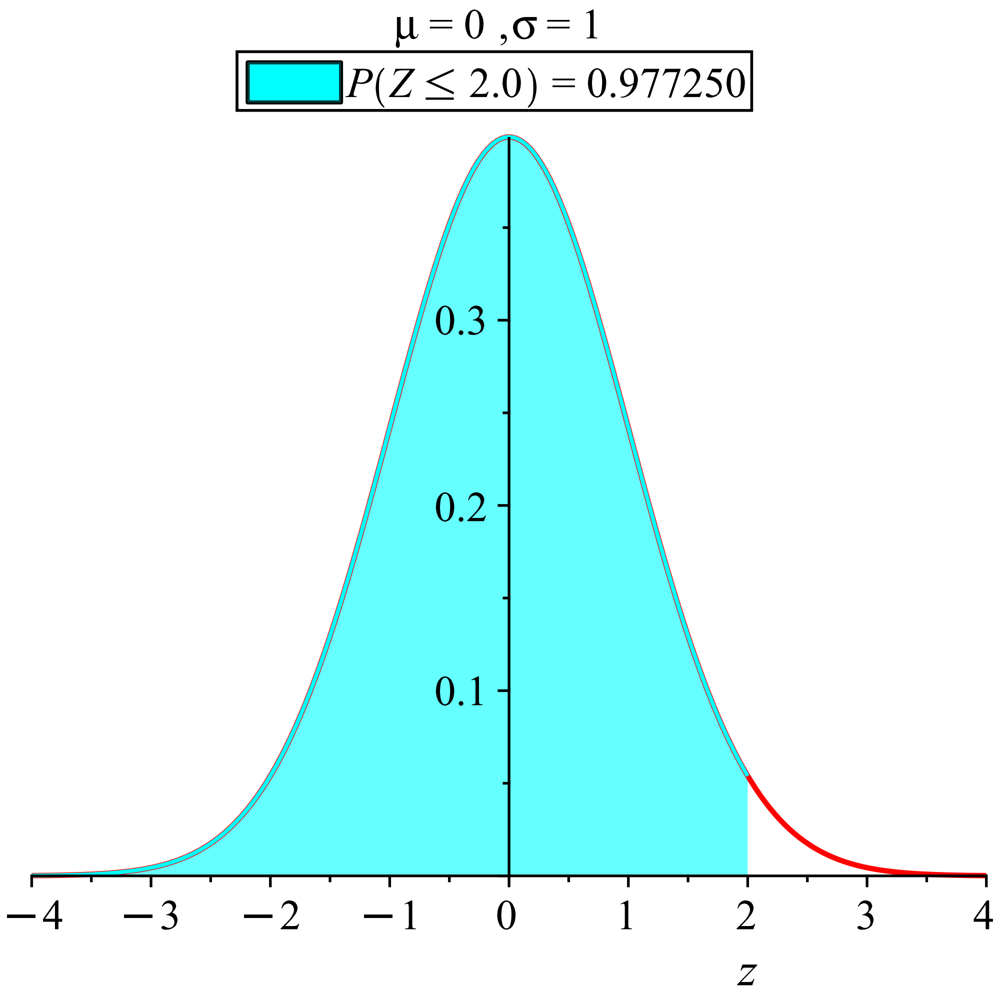
Tabellverdier

Normalfordelings-tabellen under er kopiert inn fra www.nkhansen.com/normalfordelingstabell

z	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
0,6	0,7257	0,7291	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549
0,7	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852
0,8	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133
0,9	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,8340	0,8365	0,8389
1,0	0,8413	0,8438	0,8461	0,8485	0,8508	0,8531	0,8554	0,8577	0,8599	0,8621
1,1	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749	0,8770	0,8790	0,8810	0,8830
1,2	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,8980	0,8997	0,9015
1,3	0,9032	0,9049	0,9066	0,9082	0,9099	0,9115	0,9131	0,9147	0,9162	0,9177
1,4	0,9192	0,9207	0,9222	0,9236	0,9251	0,9265	0,9279	0,9292	0,9306	0,9319
1,5	0,9332	0,9345	0,9357	0,9370	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441
1,6	0,9452	0,9463	0,9474	0,9484	0,9495	0,9505	0,9515	0,9525	0,9535	0,9545
1,7	0,9554	0,9564	0,9573	0,9582	0,9591	0,9599	0,9608	0,9616	0,9625	0,9633
1,8	0,9641	0,9649	0,9656	0,9664	0,9671	0,9678	0,9686	0,9693	0,9699	0,9706
1,9	0,9713	0,9719	0,9726	0,9732	0,9738	0,9744	0,9750	0,9756	0,9761	0,9767
2,0	0,9772	0,9778	0,9783	0,9788	0,9793	0,9798	0,9803	0,9808	0,9812	0,9817
2,1	0,9821	0,9826	0,9830	0,9834	0,9838	0,9842	0,9846	0,9850	0,9854	0,9857
2,2	0,9861	0,9864	0,9868	0,9871	0,9875	0,9878	0,9881	0,9884	0,9887	0,9890
2,3	0,9893	0,9896	0,9898	0,9901	0,9904	0,9906	0,9909	0,9911	0,9913	0,9916
2,4	0,9918	0,9920	0,9922	0,9925	0,9927	0,9929	0,9931	0,9932	0,9934	0,9936
2,5	0,9938	0,9940	0,9941	0,9943	0,9945	0,9946	0,9948	0,9949	0,9951	0,9952
2,6	0,9953	0,9955	0,9956	0,9957	0,9959	0,9960	0,9961	0,9962	0,9963	0,9964
2,7	0,9965	0,9966	0,9967	0,9968	0,9969	0,9970	0,9971	0,9972	0,9973	0,9974
2,8	0,9974	0,9975	0,9976	0,9977	0,9977	0,9978	0,9979	0,9979	0,9980	0,9981
2,9	0,9981	0,9982	0,9982	0,9983	0,9984	0,9984	0,9985	0,9985	0,9986	0,9986
3,0	0,9987	0,9987	0,9987	0,9988	0,9988	0,9989	0,9989	0,9989	0,9990	0,9990
3,1	0,9990	0,9991	0,9991	0,9991	0,9992	0,9992	0,9992	0,9992	0,9993	0,9993
3,2	0,9993	0,9993	0,9994	0,9994	0,9994	0,9994	0,9994	0,9995	0,9995	0,9995
3,3	0,9995	0,9995	0,9995	0,9996	0,9996	0,9996	0,9996	0,9996	0,9996	0,9997
3,4	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9998
3,5	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998
3,6	0,9998	0,9998	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999
3,7	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999
3,8	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999
3,9	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000

Denne tabellen får vi i Maple med kommandoen NormalFordelingsTabell.

> *NormalFordelingsTabell*$(0, 1, 0, 3.9, 6)$ #μ, σ, $z = 0\,..3.9$, *antall siffer*



μ = 0 ,σ = 1
P(Z ≤ 2.0) = 0.977250
0.3
0.2
0.1
−4
−3
−2
−1
0
1
2
3
4
z

z	0.	0.01	0.02	0.03	0.04	$\cdots$
0.	0.5000	0.5040	0.5080	0.5120	0.5160	$\cdots$
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	$\cdots$
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	$\cdots$
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	$\cdots$
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	$\cdots$
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	$\cdots$
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	$\cdots$
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	$\cdots$
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	$\cdots$
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	$\cdots$
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	$\cdots$
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	$\cdots$
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	$\cdots$
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	$\cdots$
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	$\cdots$
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	$\cdots$
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	$\cdots$

De enkelte tabellverdiene kan vi finne med

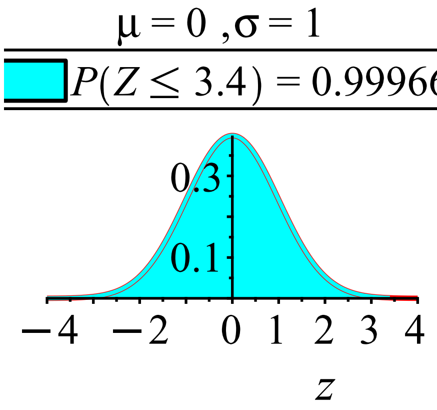
> *ProbTable*([*Normal*, 0, 1], $x = 3.4$)

0.999663070734323

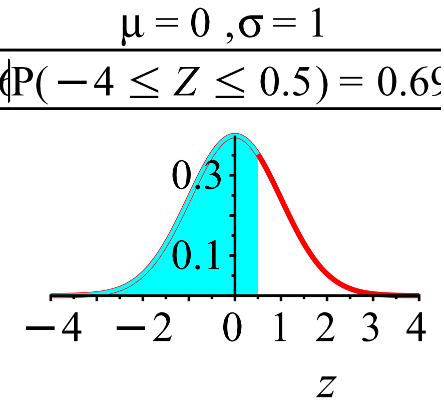
> *ProbTable*([*Normal*, 0, 1], $x = 0.5$)

0.691462461274013

> *NormalFordeling*(0, 1, 3.4, *venstre*)



> *NormalFordeling*(0, 1, −4, 0.5)



Kvantiltabellen fra www.nkhansen.com/normalfordelingstabell_kvantil
som gir **z** slik at arealet til høyre for **z** under en standard normalfordelingskurve er lik **a**.

a	z
0,50	0,0000
0,49	0,0251
0,48	0,0502
0,47	0,0753
0,46	0,1004
0,45	0,1257
0,44	0,1510
0,43	0,1764
0,42	0,2019
0,41	0,2275
0,40	0,2533
0,39	0,2793
0,38	0,3055
0,37	0,3319
0,36	0,3585
0,35	0,3853
0,34	0,4125
0,33	0,4399
0,32	0,4677
0,31	0,4959
0,30	0,5244
0,29	0,5534
0,28	0,5828
0,27	0,6128
0,26	0,6433
0,25	0,6745
0,24	0,7063
0,23	0,7388
0,22	0,7722
0,21	0,8064
0,20	0,8416
0,19	0,8779
0,18	0,9154
0,17	0,9542
0,16	0,9945
0,15	1,0364
0,14	1,0803
0,13	1,1264
0,12	1,1750
0,11	1,2265
0,10	1,2816
0,09	1,3408
0,08	1,4051
0,07	1,4758
0,06	1,5548
0,05	1,6449

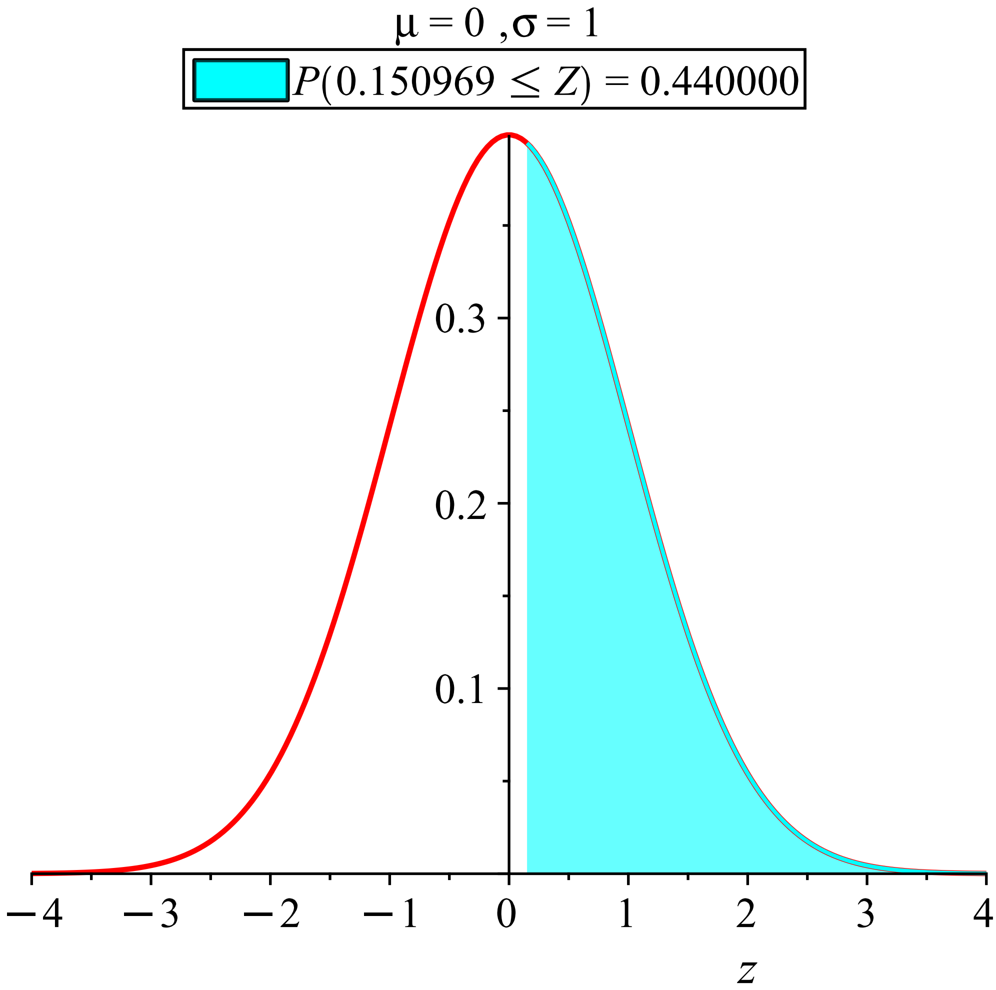
a	z
0,050	1,6449
0,049	1,6546
0,048	1,6646
0,047	1,6747
0,046	1,6849
0,045	1,6954
0,044	1,7060
0,043	1,7169
0,042	1,7279
0,041	1,7392
0,040	1,7507
0,039	1,7624
0,038	1,7744
0,037	1,7866
0,036	1,7991
0,035	1,8119
0,034	1,8250
0,033	1,8384
0,032	1,8522
0,031	1,8663
0,030	1,8808
0,029	1,8957
0,028	1,9110
0,027	1,9268
0,026	1,9431
0,025	1,9600
0,024	1,9774
0,023	1,9954
0,022	2,0141
0,021	2,0335
0,020	2,0537
0,019	2,0749
0,018	2,0969
0,017	2,1201
0,016	2,1444
0,015	2,1701
0,014	2,1973
0,013	2,2262
0,012	2,2571
0,011	2,2904
0,010	2,3263
0,009	2,3656
0,008	2,4089
0,007	2,4573
0,006	2,5121
0,005	2,5758

a	z
0,0050	2,5758
0,0049	2,5828
0,0048	2,5899
0,0047	2,5972
0,0046	2,6045
0,0045	2,6121
0,0044	2,6197
0,0043	2,6276
0,0042	2,6356
0,0041	2,6437
0,0040	2,6521
0,0039	2,6606
0,0038	2,6693
0,0037	2,6783
0,0036	2,6874
0,0035	2,6968
0,0034	2,7065
0,0033	2,7164
0,0032	2,7266
0,0031	2,7370
0,0030	2,7478
0,0029	2,7589
0,0028	2,7703
0,0027	2,7822
0,0026	2,7944
0,0025	2,8070
0,0024	2,8202
0,0023	2,8338
0,0022	2,8480
0,0021	2,8627
0,0020	2,8782
0,0019	2,8943
0,0018	2,9112
0,0017	2,9290
0,0016	2,9478
0,0015	2,9677
0,0014	2,9889
0,0013	3,0115
0,0012	3,0357
0,0011	3,0618
0,0010	3,0902
0,0009	3,1214
0,0008	3,1559
0,0007	3,1947
0,0006	3,2389
0,0005	3,2905

a	z
0,00050	3,2905
0,00049	3,2962
0,00048	3,3020
0,00047	3,3079
0,00046	3,3139
0,00045	3,3201
0,00044	3,3263
0,00043	3,3327
0,00042	3,3393
0,00041	3,3460
0,00040	3,3528
0,00039	3,3598
0,00038	3,3670
0,00037	3,3743
0,00036	3,3818
0,00035	3,3896
0,00034	3,3975
0,00033	3,4057
0,00032	3,4141
0,00031	3,4227
0,00030	3,4316
0,00029	3,4408
0,00028	3,4503
0,00027	3,4601
0,00026	3,4702
0,00025	3,4808
0,00024	3,4917
0,00023	3,5030
0,00022	3,5149
0,00021	3,5272
0,00020	3,5401
0,00019	3,5536
0,00018	3,5678
0,00017	3,5827
0,00016	3,5985
0,00015	3,6153
0,00014	3,6331
0,00013	3,6522
0,00012	3,6727
0,00011	3,6949
0,00010	3,7190
0,00009	3,7455
0,00008	3,7750
0,00007	3,8082
0,00006	3,8461
0,00005	3,8906

Tabellen i Maple får vi med kommandoen TabellKvantiler

> *TabellKvantiler*(*Normal*, 5)



α	z_{α}	α	z_{α}	α	z_{α}	$\cdots$
0.5	0.	0.05	1.6449	0.005	2.5758	$\cdots$
0.49	0.025069	0.049	1.6546	0.0049	2.5828	$\cdots$
0.48	0.050154	0.048	1.6646	0.0048	2.5899	$\cdots$
0.47	0.075270	0.047	1.6747	0.0047	2.5972	$\cdots$
0.46	0.10043	0.046	1.6849	0.0046	2.6045	$\cdots$
0.45	0.12566	0.045	1.6954	0.0045	2.6121	$\cdots$
0.44	0.15097	0.044	1.7060	0.0044	2.6197	$\cdots$
0.43	0.17637	0.043	1.7169	0.0043	2.6276	$\cdots$
0.42	0.20189	0.042	1.7279	0.0042	2.6356	$\cdots$
0.41	0.22754	0.041	1.7392	0.0041	2.6437	$\cdots$
0.40	0.25335	0.040	1.7507	0.0040	2.6521	$\cdots$
0.39	0.27932	0.039	1.7624	0.0039	2.6606	$\cdots$
0.38	0.30548	0.038	1.7744	0.0038	2.6693	$\cdots$
0.37	0.33185	0.037	1.7866	0.0037	2.6783	$\cdots$
0.36	0.35846	0.036	1.7991	0.0036	2.6874	$\cdots$
0.35	0.38532	0.035	1.8119	0.0035	2.6968	$\cdots$
0.34	0.41246	0.034	1.8250	0.0034	2.7065	$\cdots$

>

Student-t-fordeling

> $restart:$

> $X := \mathrm{nu} \rightarrow RandomVariable(StudentT(\mathrm{nu})):$

> $S := (\mathrm{nu}, t) \rightarrow PDF(X(\mathrm{nu}), t):$

> $T := (\mathrm{nu}, t) \rightarrow CDF(X(\mathrm{nu}), t):$

> $'S(\nu, t)' = S(\nu, t)$

$$S(\nu, t) = \frac{\Gamma\left(\frac{\nu}{2} + \frac{1}{2}\right)}{\sqrt{\pi \nu}\, \Gamma\left(\frac{\nu}{2}\right) \left(1 + \frac{t^2}{\nu}\right)^{\frac{\nu}{2} + \frac{1}{2}}}$$

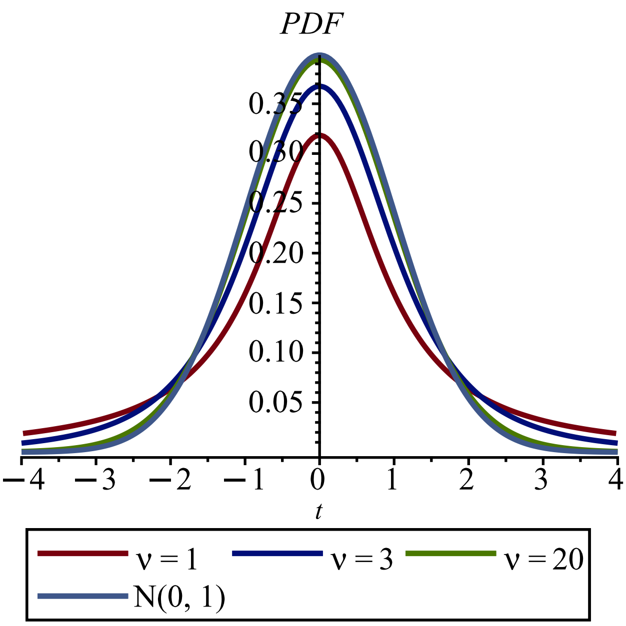
> $\mu = Mean(X(\nu)), \sigma = StandardDeviation(X(\nu))$

$$\mu = \begin{cases} undefined & \nu \leq 1 \\ 0 & otherwise \end{cases}, \sigma = \sqrt{\begin{cases} undefined & \nu \leq 2 \\ \frac{\nu}{\nu - 2} & otherwise \end{cases}}$$

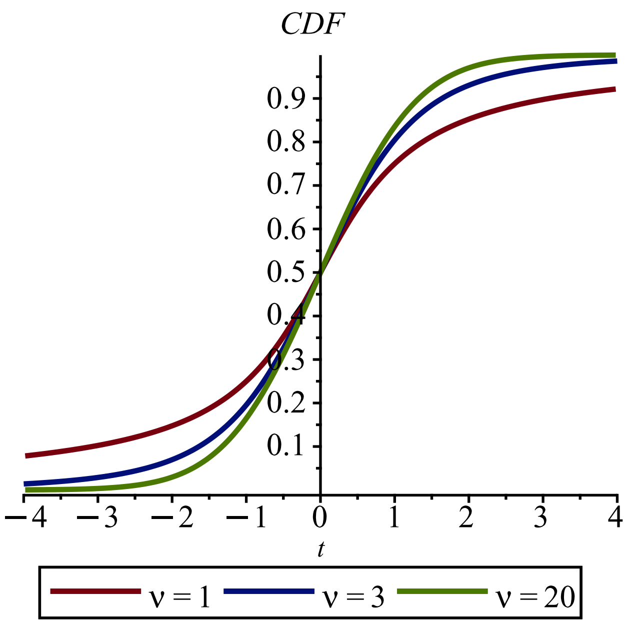
> $Xn := RandomVariable(Normal(0, 1)):$

> $N := t \rightarrow PDF(Xn, t):$

> $plot([S(1, t), S(3, t), S(20, t), N(t)], t = -4..4, thickness = 2, legend = [typeset(\nu = 1), typeset(\nu = 3), typeset(\nu = 20), \text{"N(0, 1)"}], title = PDF)$

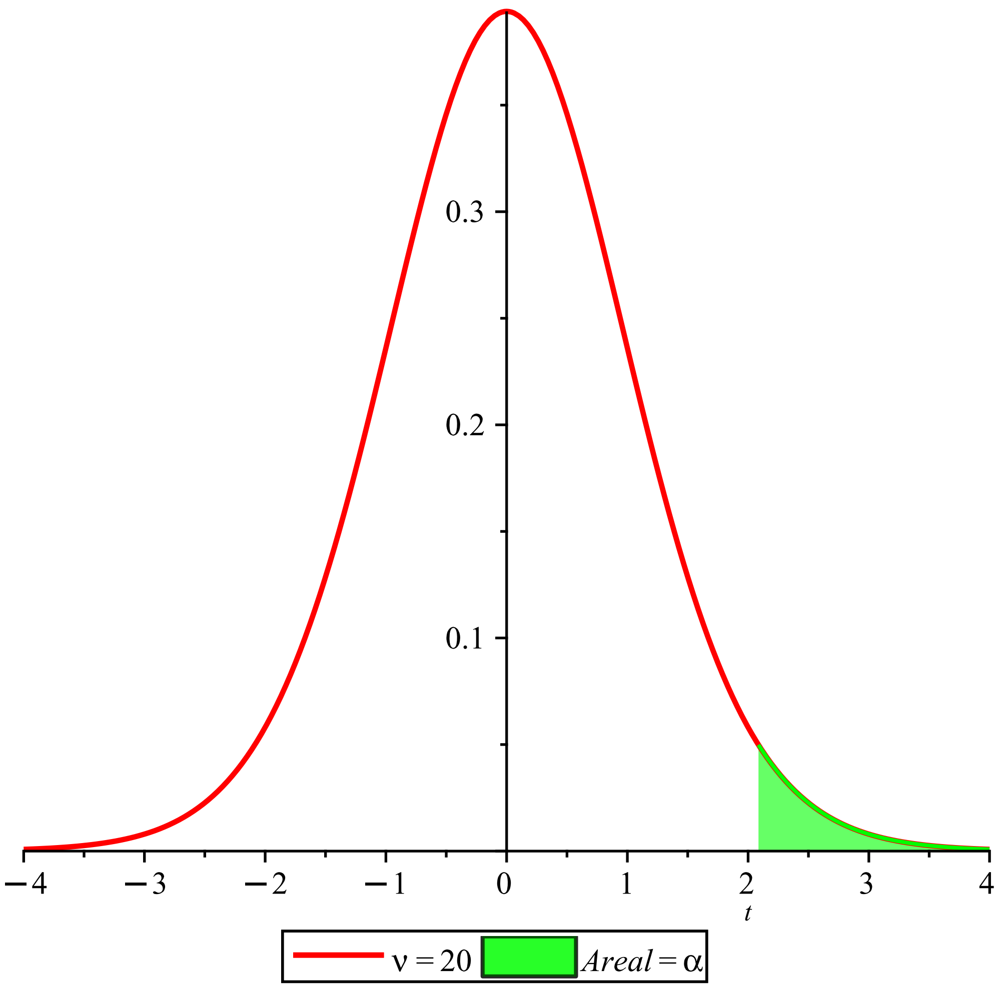


> $plot([T(1, t), T(3, t), T(20, t)], t = -4..4, thickness = 2, legend = [typeset(\nu = 1), typeset(\nu = 3), typeset(\nu = 20)], title = CDF)$



> $p1 := plot(S(20, t), t = -4..4, color = red, thickness = 2, legend = [\nu = 20]):$

> $p2 := plot(S(20, t), t = 2.086..4, color = green, filled = true, legend = typeset(Areal = \alpha)):$
$display(p1, p2)$



$> K := (\text{nu}, \text{alpha}) \rightarrow evalf(-Quantile(X(\text{nu}), \text{alpha}), 5) :$

$> t_{0.025} = K(20, 0.025)$

$$t_{0.025} = 2.0860$$

$t_{0.025}$ kommer frem ved å beregne t av

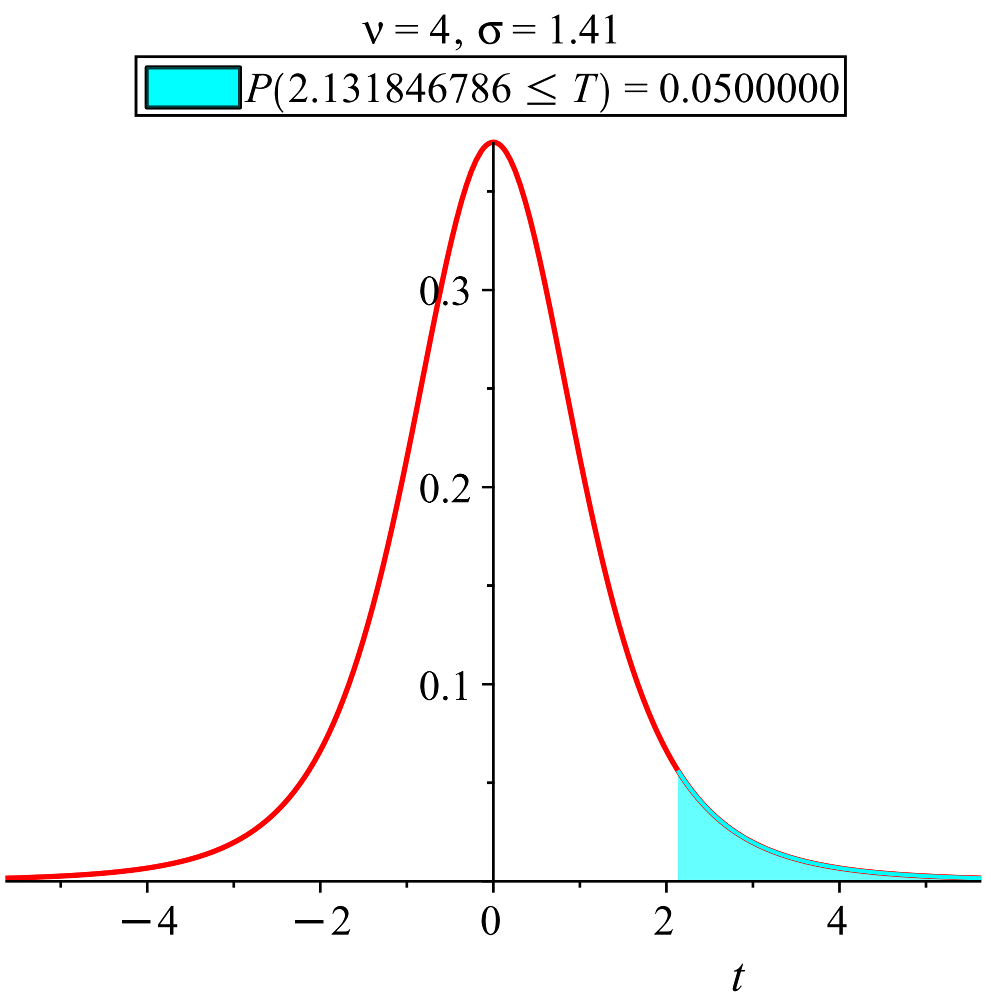
$> 1 - T(6, t) = 0.025$

$$\frac{1}{2} - \frac{5\, t\, \text{hypergeom}\left(\left[\frac{1}{2}, \frac{7}{2}\right], \left[\frac{3}{2}\right], -\frac{t^2}{6}\right)\sqrt{6}}{32} = 0.025$$

$> fsolve(\%, t)$

$$2.446911851$$

$> TabellKvantiler(StudentT, 1, 6, 5)$



	$t_{0.150}$	$t_{0.100}$	$t_{0.075}$	$t_{0.050}$	$t_{0.025}$	⋯
1.	1.9626	3.0777	4.1653	6.3138	12.706	⋯
2.	1.3862	1.8856	2.2819	2.9200	4.3027	⋯
3.	1.2498	1.6377	1.9243	2.3534	3.1824	⋯
4.	1.1896	1.5332	1.7782	2.1318	2.7763	⋯
5.	1.1558	1.4759	1.6994	2.0150	2.5706	⋯
6.	1.1342	1.4398	1.6502	1.9432	2.4469	⋯

Chi-kvadrat-fordeling (χ^2)

$> X := \mathrm{nu} \rightarrow RandomVariable(ChiSquare(\mathrm{nu})) :$

$> P := (\mathrm{nu}, t) \rightarrow PDF(X(\mathrm{nu}), t) :$

$> C := (\mathrm{nu}, t) \rightarrow CDF(X(\mathrm{nu}), t) :$

$> \chi_{\nu}^{2} = P(\nu, t)$

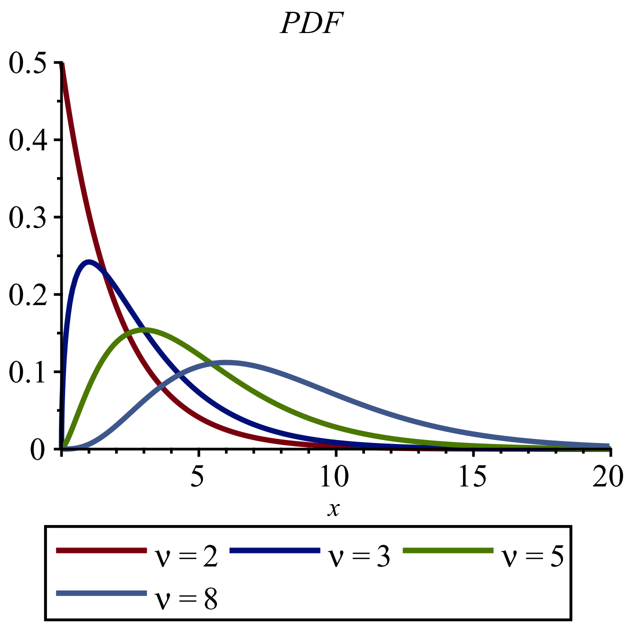
$$\chi_{\nu}^{2} = \begin{cases} 0 & t < 0 \\ \dfrac{t^{\frac{\nu}{2}-1}\, \mathrm{e}^{-\frac{t}{2}}}{2^{\frac{\nu}{2}}\, \Gamma\left(\frac{\nu}{2}\right)} & otherwise \end{cases}$$

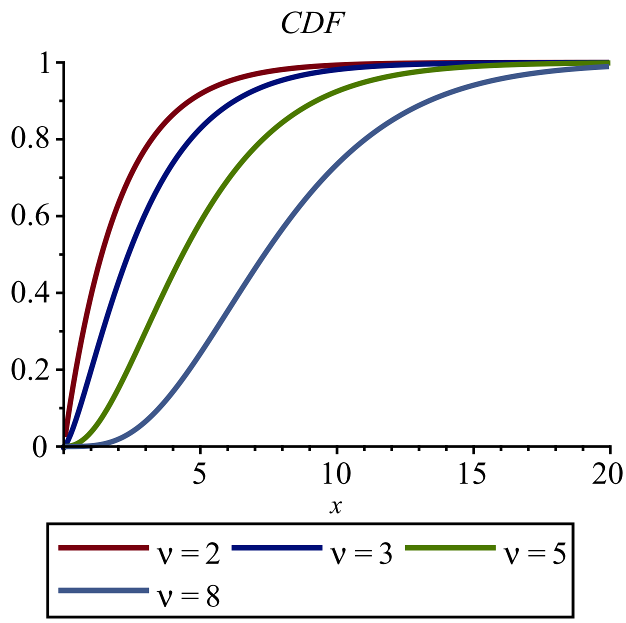
$> \mu = Mean(X(\nu)), \sigma = StandardDeviation(X(\nu))$

$$\mu = \nu, \sigma = \sqrt{2}\, \sqrt{\nu}$$

$> plot([P(2, x), P(3, x), P(5, x), P(8, x)], x = 0..20, thickness = 2, legend = [\nu = 2, \nu = 3, \nu = 5, \nu = 8], title = PDF)$

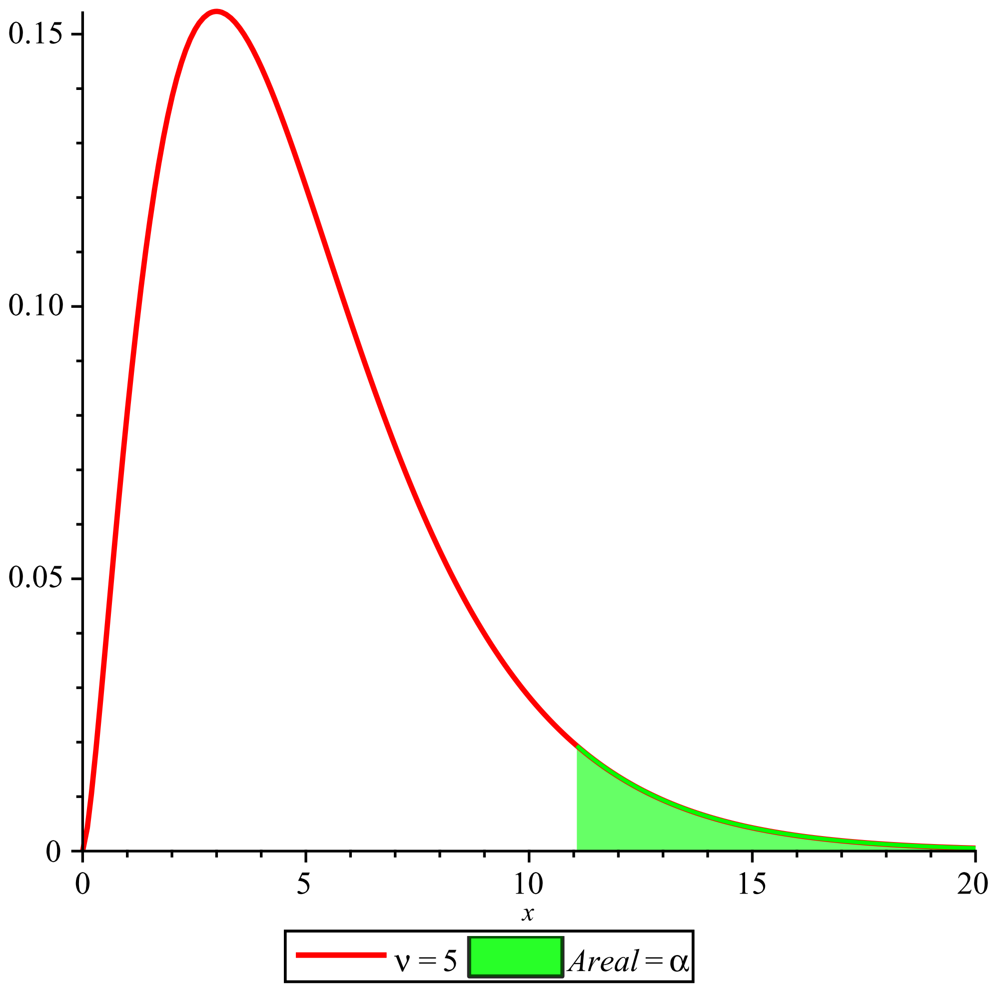
$> plot([C(2, x), C(3, x), C(5, x), C(8, x)], x = 0..20, thickness = 2, legend = [\nu = 2, \nu = 3, \nu = 5, \nu = 8], title = CDF)$





$> p1 := plot(P(5, x), x = 0..20, color = red, thickness = 2, legend = [\nu = 5]) :$

$> p2 := plot(P(5, x), x = 11.07..20, color = green, filled = true, legend = typeset(Areal = \alpha)) :$
$display(p1, p2)$



> $1 - C(5, t) = 0.05$:

> $\chi_{0.05}^{2} = fsolve(\%, t)$

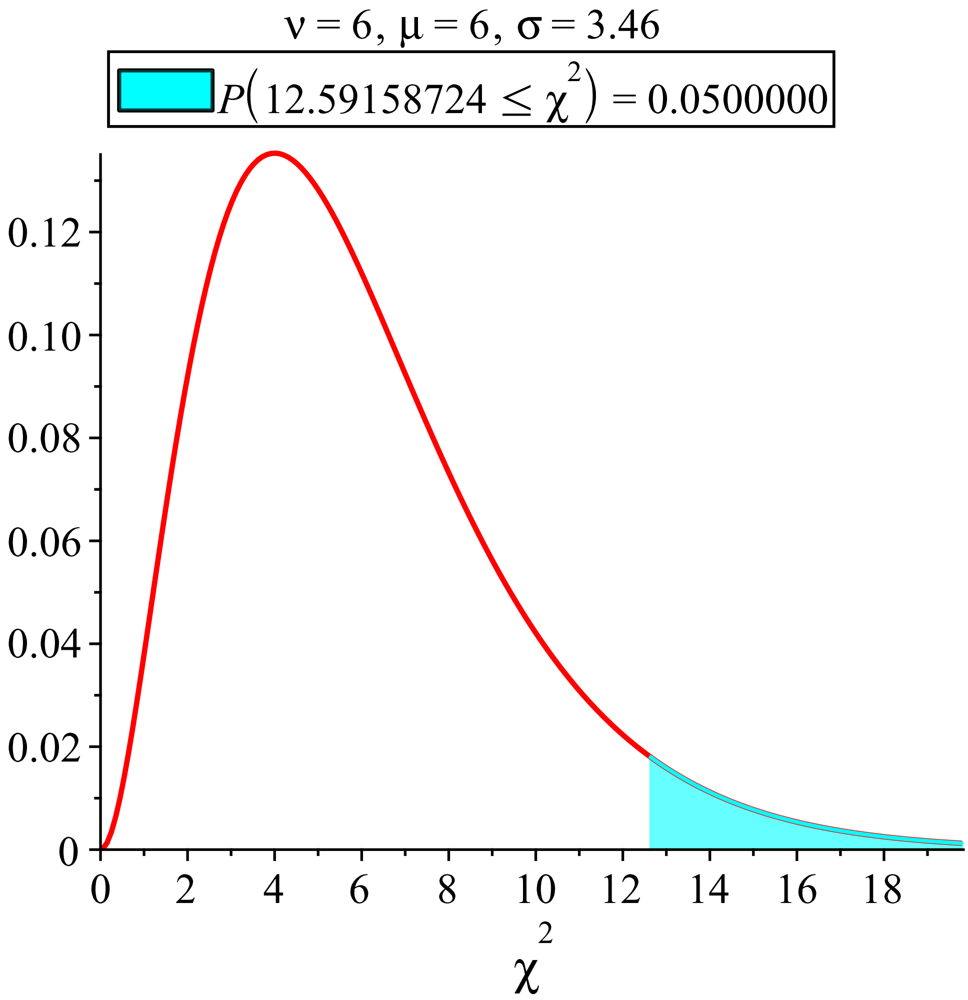
$$\chi_{0.05}^{2} = 11.07049769$$

> $K := (\text{nu}, \text{alpha}) \rightarrow evalf(Quantile(X(\text{nu}), 1 - \text{alpha}), 4)$:

> $\chi_{0.05}^{2} = K(5, 0.05)$

$$\chi_{0.05}^{2} = 11.07$$

> $TabellKvantiler(ChiSquare, 4, 8, 5)$



	$\chi^2_{0.995}$	$\chi^2_{0.990}$	$\chi^2_{0.975}$	$\chi^2_{0.950}$	$\chi^2_{0.05}$ ⋯
4.	0.20699	0.29711	0.48442	0.71072	9.487 ⋯
5.	0.41174	0.55430	0.83121	1.1455	11.07 ⋯
6.	0.67573	0.87209	1.2373	1.6354	12.59 ⋯
7.	0.98926	1.2390	1.6899	2.1674	14.06 ⋯
8.	1.3444	1.6465	2.1797	2.7326	15.50 ⋯

F-fordeling

$> X := (nu1, nu2) \rightarrow RandomVariable(FRatio(nu1, nu2)) :$

$> P := (nu1, nu2) \rightarrow PDF(X(nu1, nu2), f) :$

$> C := (nu1, nu2) \rightarrow CDF(X(nu1, nu2), f) :$

> $F(\nu_1, \nu_2) = P(\nu_1, \nu_2)$

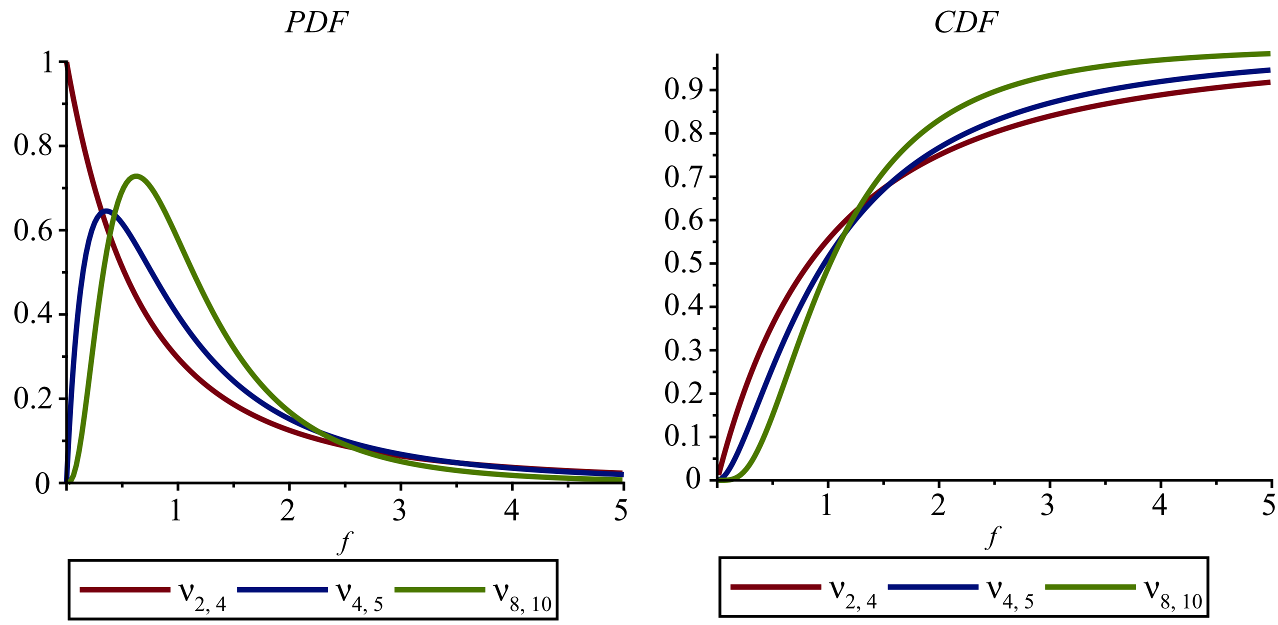
$$F(\nu_1, \nu_2) = \begin{cases} 0 & f < 0 \\ \dfrac{\Gamma\left(\frac{\nu_1}{2} + \frac{\nu_2}{2}\right) \left(\frac{\nu_1}{\nu_2}\right)^{\frac{\nu_1}{2}} f^{\frac{\nu_1}{2} - 1}}{\Gamma\left(\frac{\nu_1}{2}\right) \Gamma\left(\frac{\nu_2}{2}\right) \left(1 + \frac{\nu_1 f}{\nu_2}\right)^{\frac{\nu_1}{2} + \frac{\nu_2}{2}}} & \textit{otherwise} \end{cases}$$

> $\mu = Mean(X(\nu_1, \nu_2)), \sigma = StandardDeviation(X(\nu_1, \nu_2))$

$$\mu = \begin{cases} \textit{undefined} & \nu_2 \le 2 \\ \dfrac{\nu_2}{\nu_2 - 2} & \textit{otherwise} \end{cases}, \sigma = \sqrt{\begin{cases} \textit{undefined} & \nu_2 \le 4 \\ \dfrac{2\,\nu_2^2\,(\nu_1 + \nu_2 - 2)}{\nu_1\,(\nu_2 - 2)^2\,(\nu_2 - 4)} & \textit{otherwise} \end{cases}}$$

> $plot([P(2, 4), P(4, 5), P(8, 10)], f = 0..5, thickness = 2, legend = [typeset(\nu_{2,4}), typeset(\nu_{4,5}), typeset(\nu_{8,10})], title = PDF)$

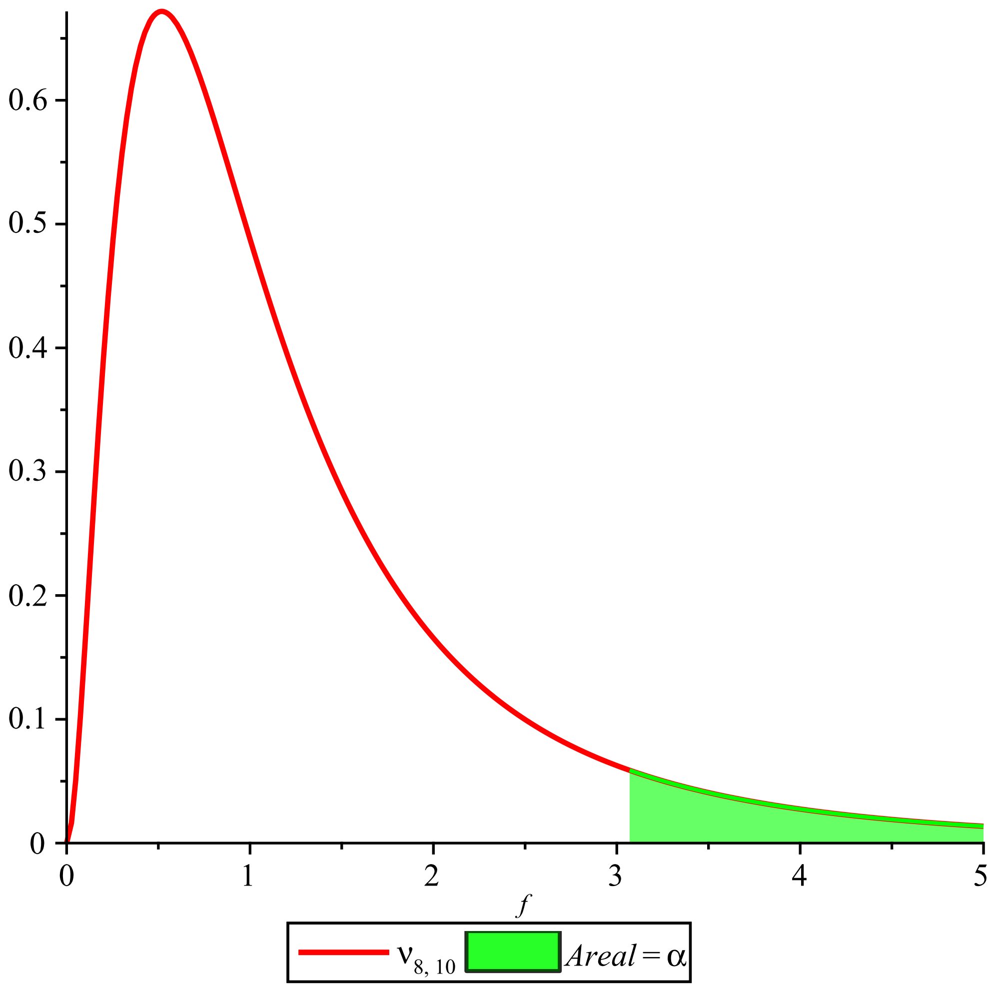
> $plot([C(2, 4), C(4, 5), C(10, 8)], f = 0..5, thickness = 2, legend = [typeset(\nu_{2,4}), typeset(\nu_{4,5}), typeset(\nu_{8,10})], title = CDF)$



> $K := (nu1, nu2, \alpha) \rightarrow evalf(Quantile(X(nu1, nu2), 1 - \alpha), 4):$

> $p1 := plot(P(6, 7), f = 0..5, color = red, thickness = 2, legend = \nu_{8,10}):$

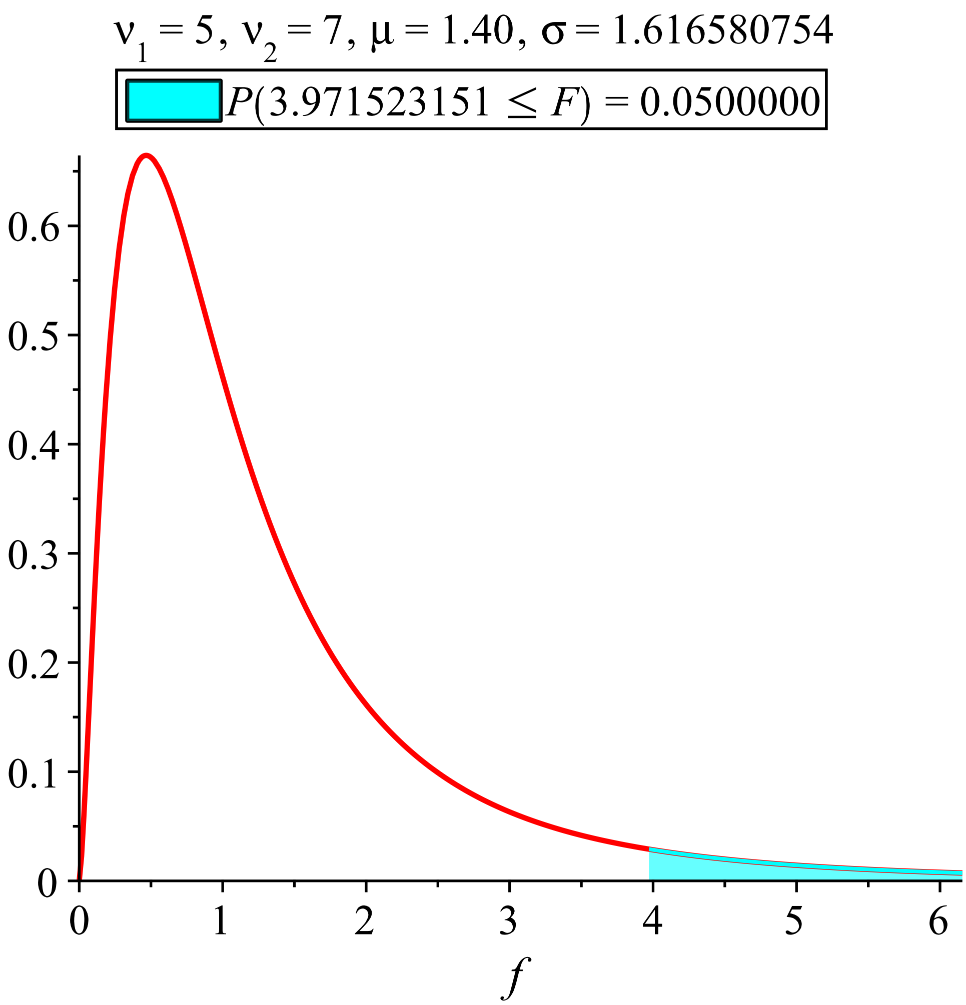
> $p2 := plot(P(6, 7), f = 3.07..5, color = green, filled = true, legend = typeset(Areal = \alpha)):$
$display(p1, p2)$



> $K(6, 7, 0.05)$

$$3.866$$

> $TabellKvantiler([F, [1, 8], [4, 9], 0.05], 5)$



ν1 = 5, ν2 = 7, μ = 1.40, σ = 1.616580754
P(3.971523151 ≤ F) = 0.0500000
0.6
0.5
0.4
0.3
0.2
0.1
0
0
1
2
3
4
5
6
f

$\alpha = 0.05$					$\cdots$
$\dfrac{\nu_1}{\nu_2}$	1.	2.	3.	4.	$\cdots$
4.	7.7086	6.9443	6.5914	6.3882	6.$\cdots$
5.	6.6079	5.7861	5.4095	5.1922	5.$\cdots$
6.	5.9874	5.1433	4.7571	4.5337	4.$\cdots$
7.	5.5914	4.7374	4.3468	4.1203	3.$\cdots$
8.	5.3177	4.4590	4.0662	3.8379	3.$\cdots$
9.	5.1174	4.2565	3.8625	3.6331	3.$\cdots$

Binomial-fordeling

```
> X := (n, p) → RandomVariable(Binomial(n, p)) :
> B := (n, p, x) → ProbabilityFunction(X(n, p), x) :
> C := (n, p, x) → CDF(X(n, p), x) :
> f(x) = B(n, p, x)
```

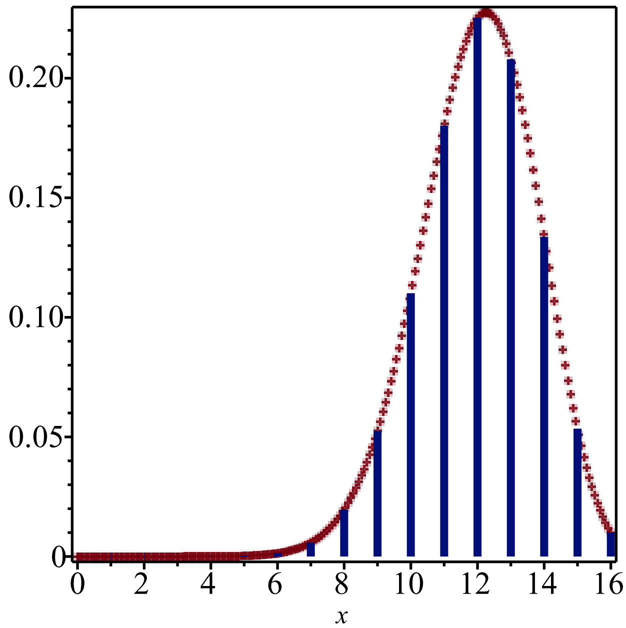
$$f(x) = \begin{cases} 0 & x < 0 \\ \binom{n}{x} p^x (1-p)^{n-x} & \mathit{otherwise} \end{cases}$$

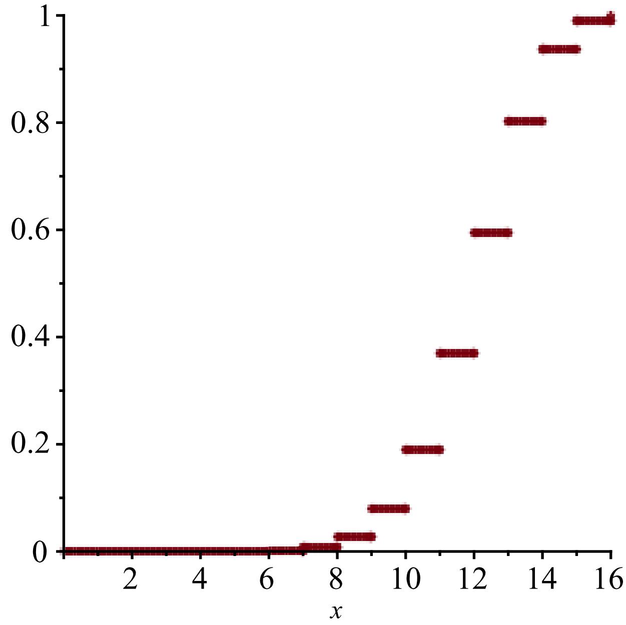
```
> μ = Mean(X(n, p)), σ = StandardDeviation(X(n, p))
```

$$\mu = p\,n, \sigma = \sqrt{n\,p\,(1-p)}$$

```
> p1 := plot(B(16, 0.75, x), x = 0 ..16, style
      = point) :
  p2 := DensityPlot(X(16, 0.75), range = 0
      ..16) :
  display(p1, p2)
```

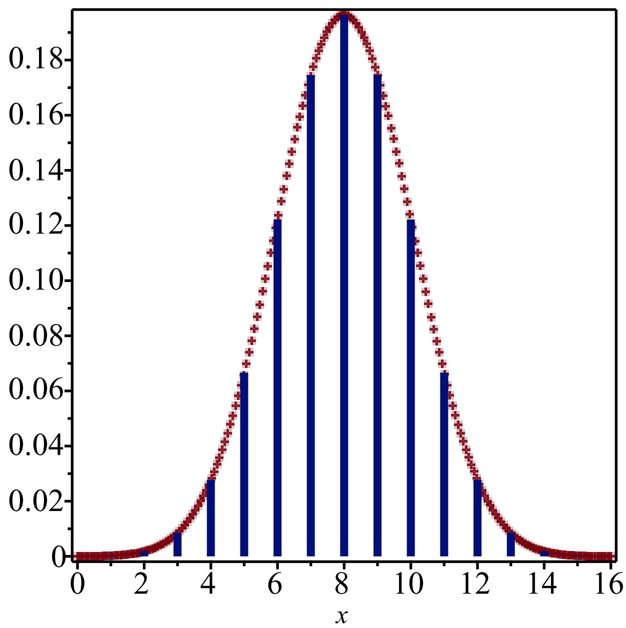
```
> plot(C(16, 0.75, x), x = 0 ..16, style = point)
```

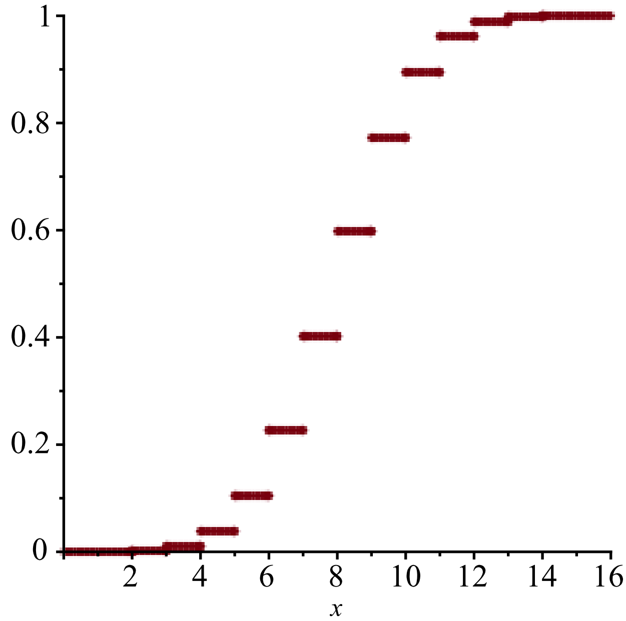




> $p1 := plot(B(16, 0.5, x), x = 0..16, style = point):$
$p2 := DensityPlot(X(16, 0.5), range = 0..16):$
$display(p1, p2)$

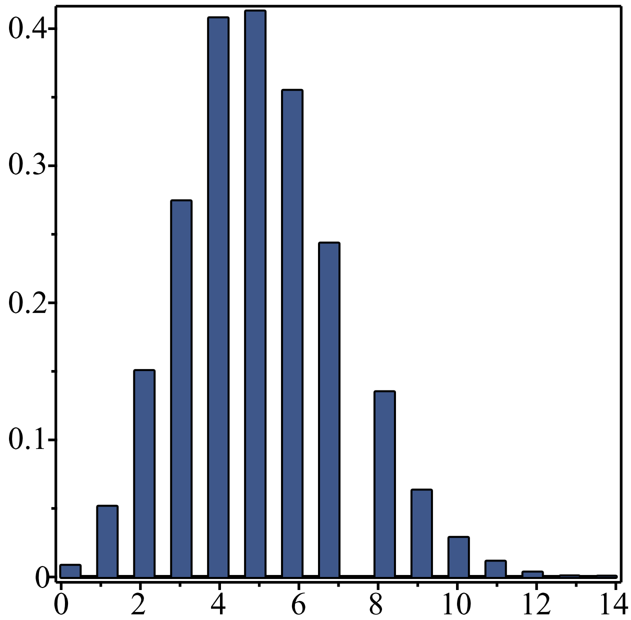
> $plot(C(16, 0.5, x), x = 0..16, style = point)$

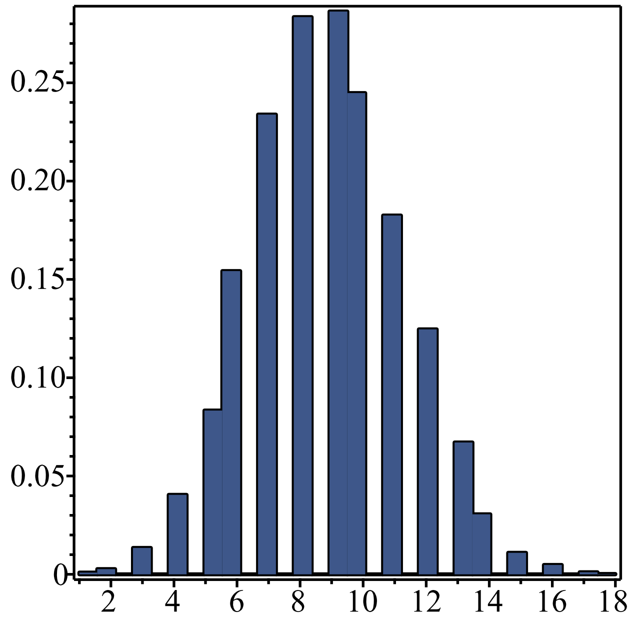




> $S := Sample(X(25, 0.2), 10000):$
$Histogram(S)$

> $S := Sample(X(25, 0.35), 10000):$
$Histogram(S)$





Poisson-fordeling

> $X := \text{lambda} \rightarrow RandomVariable(Poisson(\text{lambda}))$:

> $P := (\text{lambda}, x) \rightarrow ProbabilityFunction(X(\text{lambda}), x)$:

> $C := (\text{lambda}, x) \rightarrow CumulativeDistributionFunction(X(\text{lambda}), x)$:

> $f(x) = P(\lambda, x)$

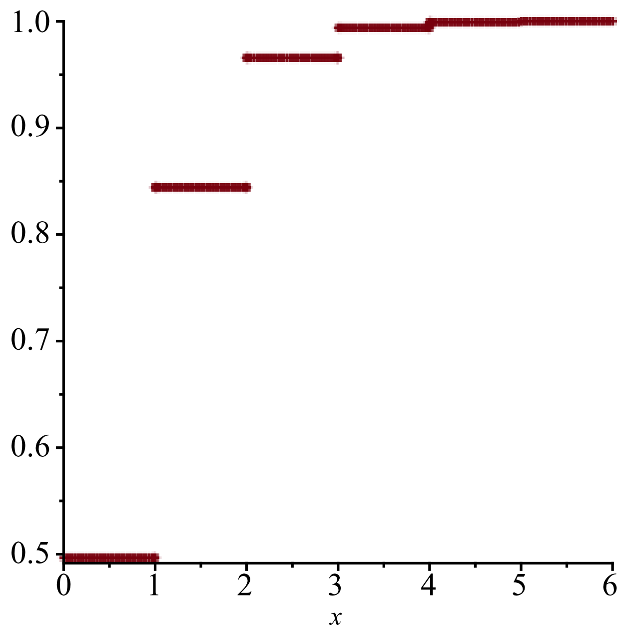
$$f(x) = \begin{cases} 0 & x < 0 \\ \dfrac{\lambda^x e^{-\lambda}}{x!} & otherwise \end{cases}$$

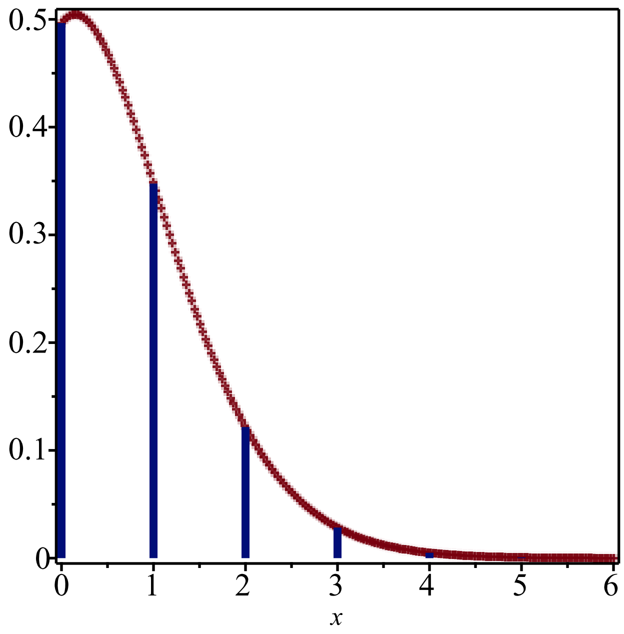
> $\mu = Mean(X(\lambda)), \sigma = StandardDeviation(X(p))$

$$\mu = \lambda, \sigma = \sqrt{p}$$

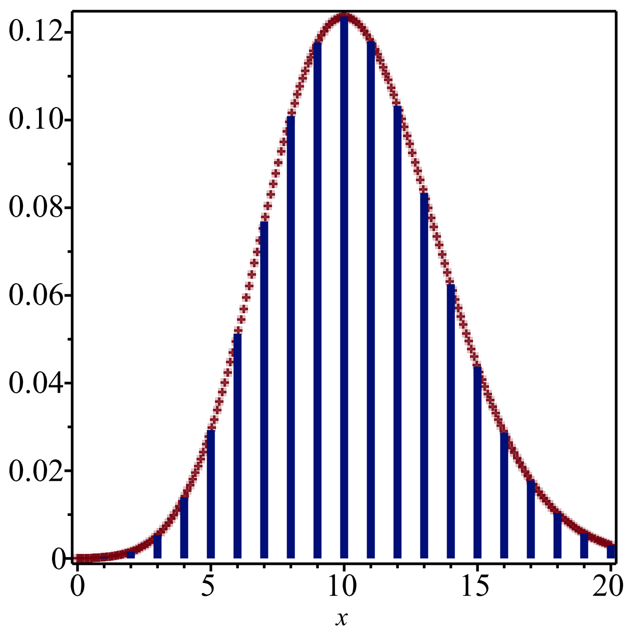
> $p1 := plot(P(0.7, x), x = 0..6, style = point)$:
$p2 := DensityPlot(X(0.7), range = 0..6)$:
$display(p1, p2)$

> $plot(C(0.7, x), x = 0..6, style = point)$



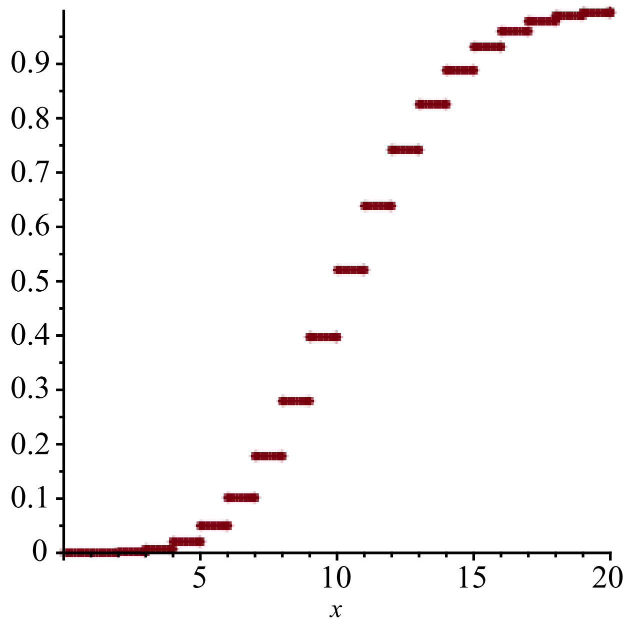


> $p1 := plot(P(10.5, x), x = 0 ..20, style = point) :$
$p2 := DensityPlot(X(10.5), range = 0 ..20) :$
$display(p1, p2)$

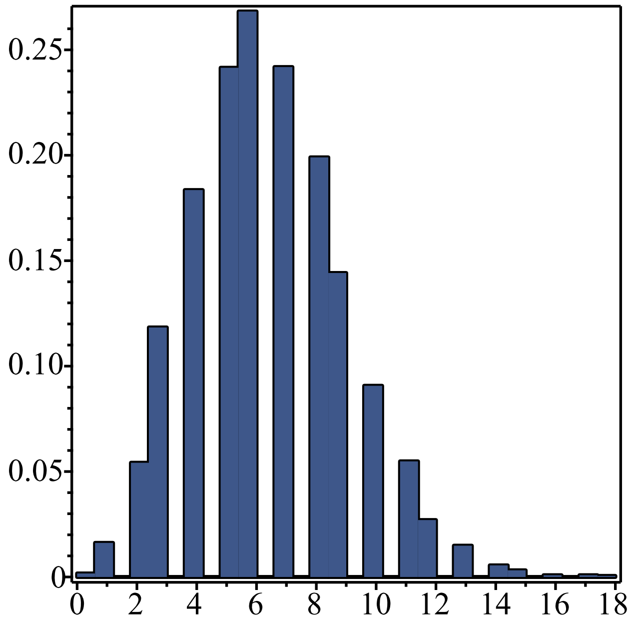
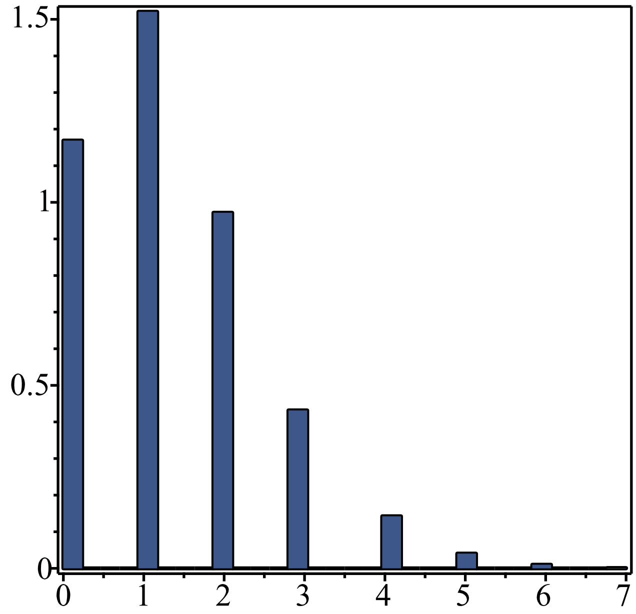


> $S := Sample(X(6.5), 10000) :$
$Histogram(S)$

> $plot(C(10.5, x), x = 0 ..20, style = point)$



> $S := Sample(X(1.3), 10000) :$
$Histogram(S)$

9.3 Befolkningsmodell

Parabelmodell

Folkemengden i Norge per 1. januar er gitt som følger

År	**1769**	**1801**	**1815**	**1825**	**1835**	**1845**	**1855**	**1865**
Antall	723618	883603	885431	1051318	1194827	1328471	1490047	1701756
År	1875	1890	1900	1910	1920	1930	1940	1950
Antall	1813424	2000917	2217971	2376952	2616274	2799713	2963909	3249954
År	1960	1970	1980	1990	2000	2005	2007	
Antall	3567707	3863221	4078900	4233116	4478497	4606363	4681134	

a) Tilpass et andregradspolynom til de gitte dataene for folkemengden.
b) For hvilke år gir modellen en brukbar tilnærming?
Statistisk sentralbyrå antar at med middels nasjonal vekst vil folketallet i årene 2010 til 2060 basert på folkemengden i 2005 være

År	**2010**	**2020**	**2030**	**2040**	**2050**	**2060**
Antall	4748000	5045000	5367000	5623000	5843000	6061000

c) Hvordan passer modellen med tallene fra Statistisk sentralbyrå når modellen anvendes fremover i tid?

Løsning
a)

> $restart:$

> $N := [885431, 1051318, 1194827, 1328471, 1490047, 1701756, 1813424, 2000917, 2217971, 2376952, 2616274, 2799713, 2963909, 3249954, 3567707, 3863221, 4078900, 4233116,$

4478497, 4606363, 4681134] :

> $T := [1815, 1825, 1835, 1845, 1855, 1865, 1875, 1890, 1900, 1910, 1920, 1930, 1940, 1950, 1960, 1970, 1980, 1990, 2000, 2005, 2007]$:

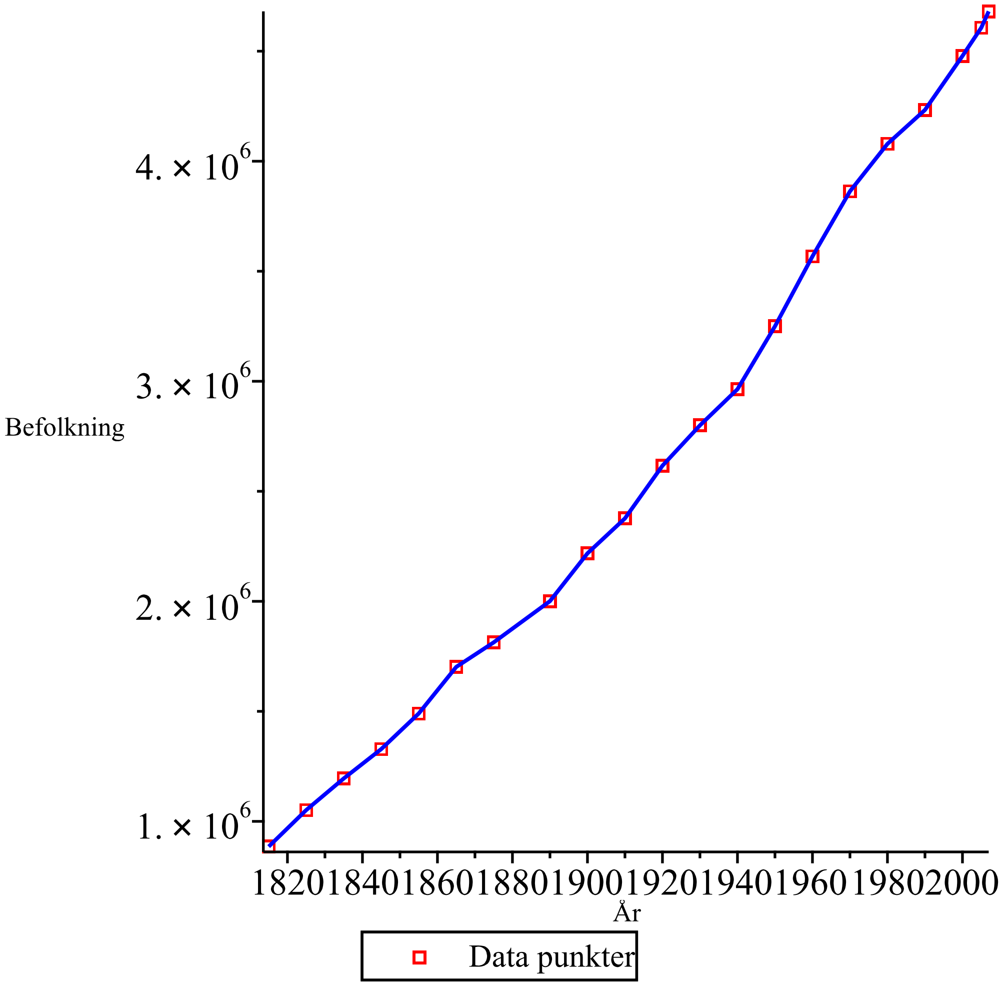
> $g := (x, y) \rightarrow [x, y]$:

> $P := zip(g, T, N)$: % #` tar x fra listen T og y fra listen N og lager samtlige punkter

[[1815, 885431], [1825, 1051318], [1835, 1194827], [1845, 1328471], [1855, 1490047], [1865, 1701756], [1875, 1813424], [1890, 2000917], [1900, 2217971], [1910, 2376952], [1920, 2616274], [1930, 2799713], [1940, 2963909], [1950, 3249954], [1960, 3567707], [1970, 3863221], [1980, 4078900], [1990, 4233116], [2000, 4478497], [2005, 4606363], [2007, 4681134]]

> $plt1 := plot(P, style = point, color = red, symbol = box, legend =$ "Data punkter") :
$plt2 := plot(P, color = blue)$: #plotter linjer mellom punktene

> $p1 := display(plt1, plt2, labels =$ ["År", "Befolkning"], $font = [times, roman, 14])$: %



>

b)

Her vil det være naturlig å prøve å finne et andregradspolynom $f(x) = a\,t^2 + b\,t + c$ som tilnærmet faller sammen med den gitte kurven. Vi må finne mulige verdier for a, b og c. Til det trenger vi å bruke 3 utvalgte punkter på kurven. Vi kan f.eks. bruke det første punktet $(1769, 723618)$ det tiende punktet $(1890, 2000917)$ og det siste punktet $(2007, 4681134)$.
Vi krever altså at kurven skal gå gjennom disse tre punktene. Det gir følgende tre ligninger med tre ukjente størrelser a, b og c.

> $f := x \mapsto a \cdot x^2 + b \cdot x + c$

$$f := x \mapsto a \cdot x^2 + b \cdot x + c$$

> $lign1 := N_1 = f(T_1) : \%; \; lign2 := N_{10} = f(T_{10}) : \%; \; lign3 := N_{21} = f(T_{21}) : \%$

$$885431 = 3294225\,a + 1815\,b + c$$

$$2376952 = 3648100\,a + 1910\,b + c$$

$$4681134 = 4028049\,a + 2007\,b + c$$

> $solve(\{lign1, lign2, lign3\})$

$$\left\{a = \frac{74219753}{1769280},\; b = -\frac{248690492821}{1769280},\; c = \frac{6948108466779}{58976}\right\}$$

> $evalf(\%)$

$$\{a = 41.94912789,\; b = -140560.2804,\; c = 1.178124740 \times 10^8\}$$

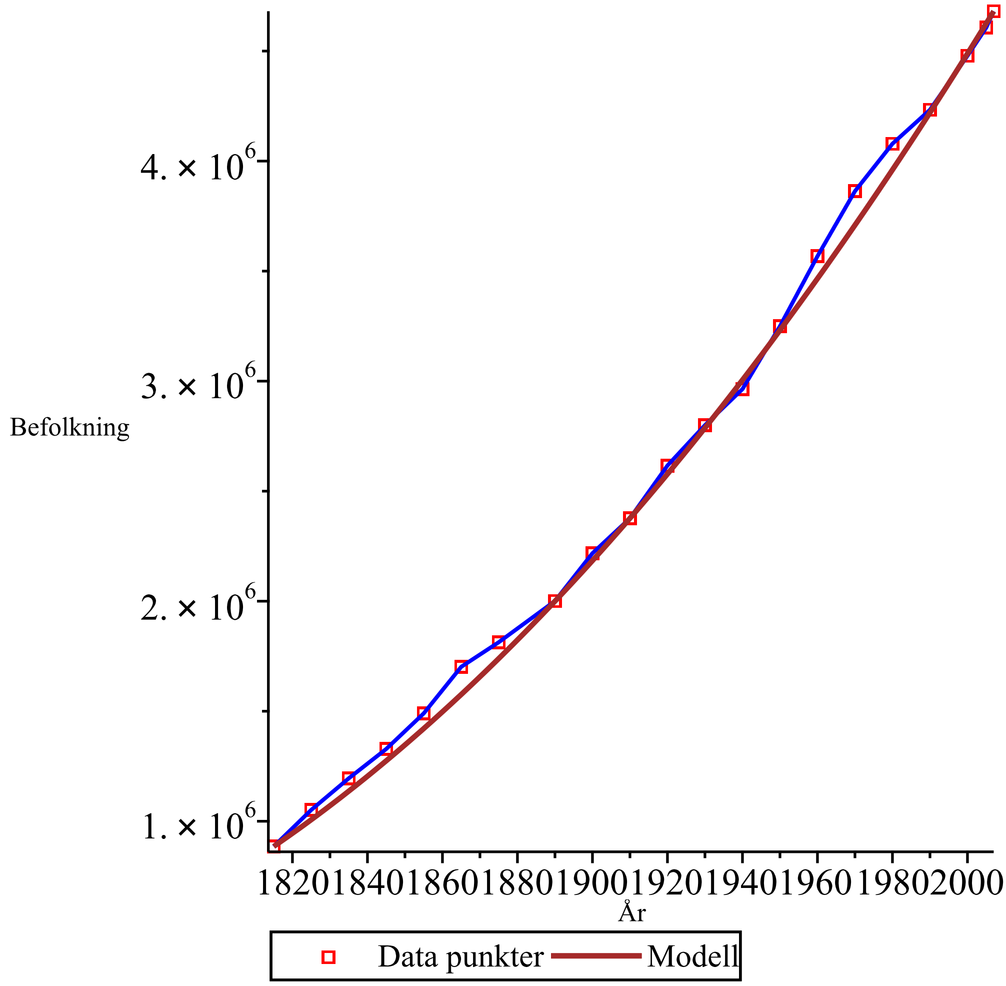
> assign(%) :#` `gjør at a, b og c blir tilordnet tallverdiene (navn på tallverdiene)

> $'f(x)' = f(x)$

$$f(x) = 41.94912789\,x^2 - 140560.2804\,x + 1.178124740 \times 10^8$$

> $p := plot(f(x), x = 1815..2007, color = brown, thickness = 2, legend = \text{"Modell"}) :$

> $plt := display(p1, p) : \%$



>

Vi ser at tilnærmingen stort sett er bra. Det er størst avvik rundt 1815 og i årene 1960-1980.

c)

Så plotter vi de nye punktene og de tilsvarende punktene basert på modellen.

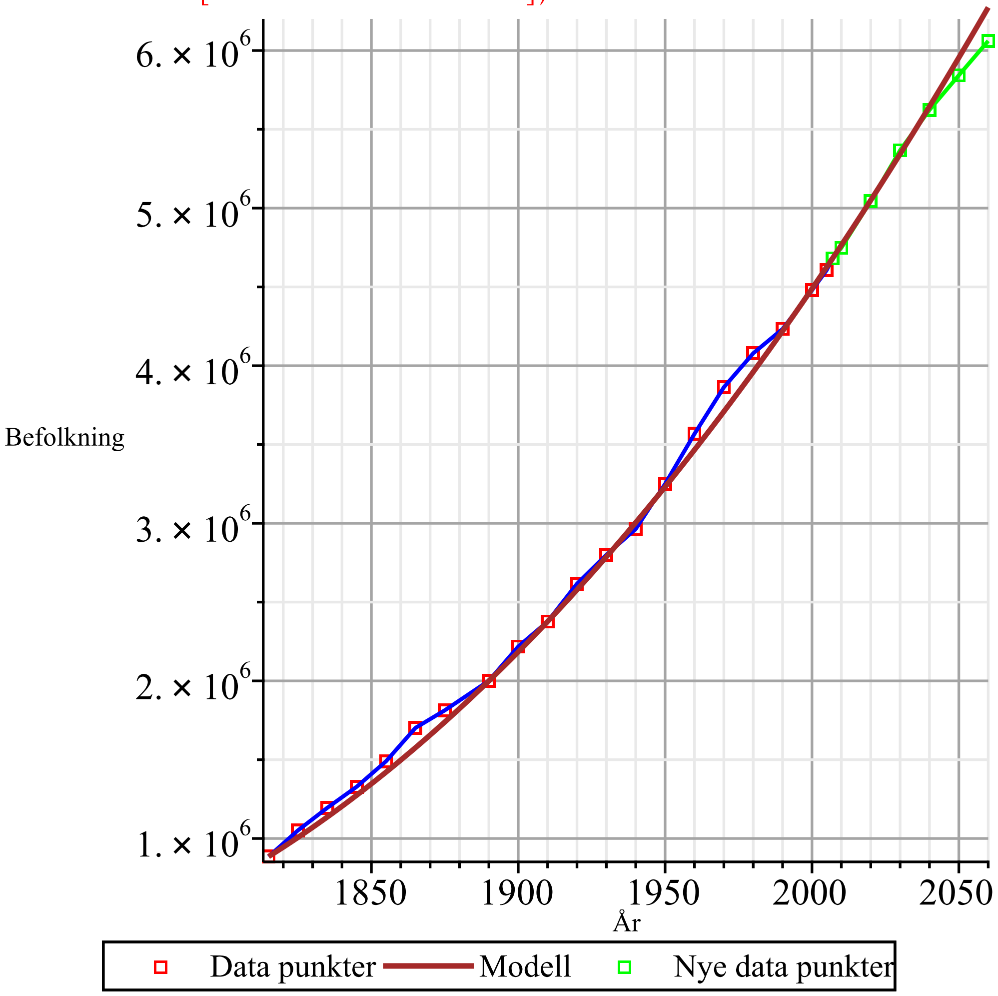
>

$P2 := [[2007, 4681134], [2010, 4748000], [2020, 5045000], [2030, 5367000], [2040, 5623000], [2050, 5843000], [2060, 6061000]]:$

> $plt3 := plot(P2, style = point, color = green, symbol = box, legend = \text{"Nye data punkter"}):$
$plt4 := plot(P2, color = green) : plt5 := plot(f(x), x = 2005..2060, color = brown, thickness = 2, legend = \text{"Modell"}) : plt6 := display(plt3, plt4, plt5, labels = [\text{"År"}, \text{"Befolkning"}], gridlines):$

Hekter vi denne grafen på den forrige, ser vi bedre sammenhengen.

> $display\left(plt, plt6, view = \left[1815..2060, N_1..6.2\ 10^6\right]\right)$



Parabelmodellen er brukbar fram til 2020?

>

Logistisk modell

a) Fremstill grafisk befolkningen som en funksjon av tiden (år). De enkelte punktene skal markeres og forbindes med rette linjer.

På grunnlag av befolkningsdataene skal du lage en logistisk modell av typen

$$\frac{dP}{dt} = k\,P\,(M - P), \text{ der } P(0) = P_0 \qquad (1)$$

der $P = P(t)$ er folkemengden ved tiden t og $P(0) = P_0$ er folkemengden ved tiden $t = 0$, som settes til år 1900. k er en konstant, og M er den asymptotiske verdien som folkemengden nærmer seg mot når $t \rightarrow \infty$. Ligning (1) kan skrives

$$\frac{dP}{dt} = a\,P + b\,P^2 \qquad (2)$$

der $a = k\,M$ og $b = -k$. Ligning (2) kan omformes til

$$y = \frac{1}{P}\ \frac{dP}{dt} = a + b\,P \qquad (3)$$

som er ligningen for en rett linje.

b) Fremstill punktmengden y grafisk som funksjon av punktmengden P.

c) Finn de numeriske verdiene på a og b ved hjelp av minste kvadraters metode.
Fremstill den rette linjen $y = a + b\,P$ i samme aksesystem som grafen i **b)**.

d) Løs differensialligningen $\frac{d}{dt}\,P(t) = k\,P(t)\,(M - P(t))$, $P(0) = 2.217971 \cdot 10^6$

der k og M er gitt over ved hjelp av de beregnete verdiene for a og b.
Hvilket år er endringen i folkemengden per tidsenhet størst?
Hvor stor er da folkemengden?

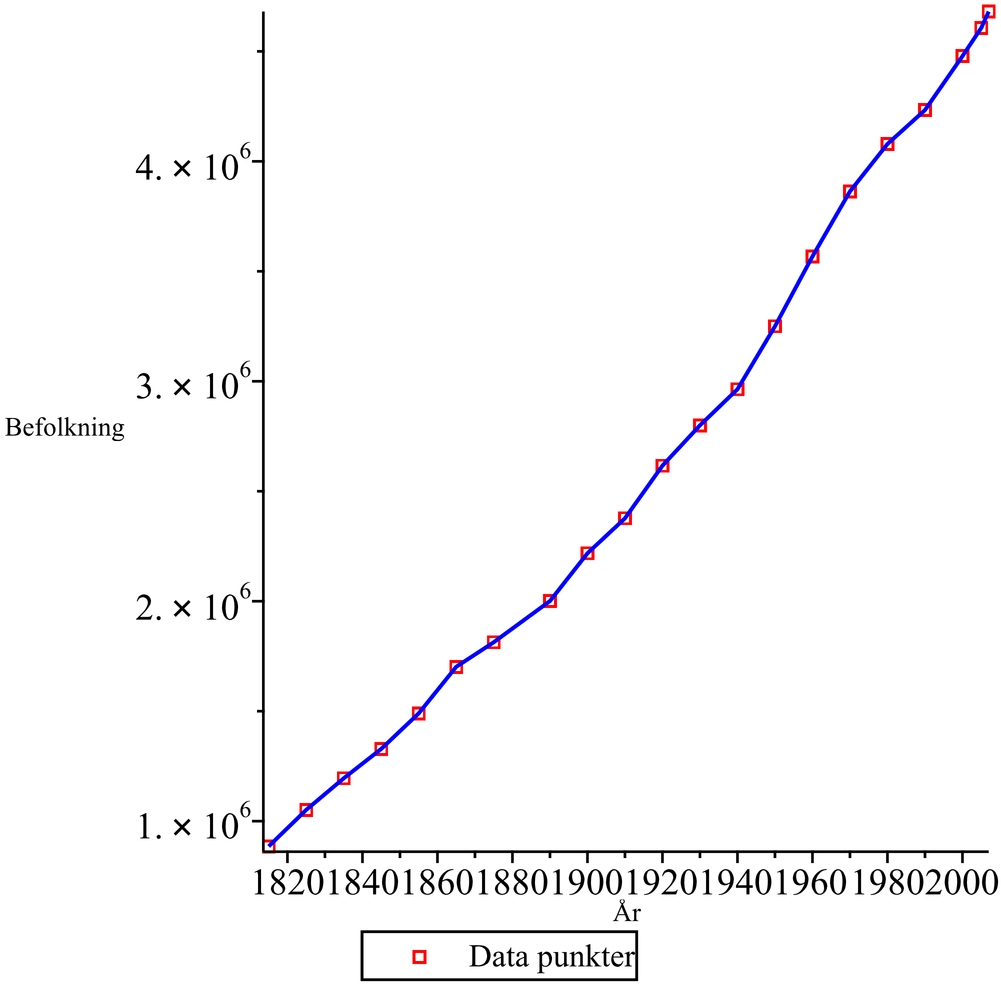
e) Fremstill grafisk i samme aksesystem grafen i **a)**, $P(t)$ og punktmengden L i **d)**.

f) Justér ved prøving og feiling verdien på M slik at den logistiske modellen for folkemengden, $P(t)$, bedre samsvarer med dataene i listen L fra Statistisk sentralbyrå.

Løsning

a)

```
> P := 'P' :
> B := [885431, 1051318, 1194827, 1328471, 1490047, 1701756, 1813424, 2000917, 2217971,
    2376952, 2616274, 2799713, 2963909, 3249954, 3567707, 3863221, 4078900, 4233116,
    4478497, 4606363, 4681134] :
> T := [1815, 1825, 1835, 1845, 1855, 1865, 1875, 1890, 1900, 1910, 1920, 1930, 1940, 1950,
    1960, 1970, 1980, 1990, 2000, 2005, 2007] :
> p1
```



>

b) Fremstill punktmengden y grafisk som funksjon av punktmengden P.

Løsning

> $N := nops(B)$:

Den relative befolkningsraten y = $\frac{1}{P}\ \frac{dP}{dt}$ kan approksimeres ved

> $Pr := i \rightarrow evalf((B[i+1]-B[i-1])/((T[i+1]-T[i-1]) * B[i]))$

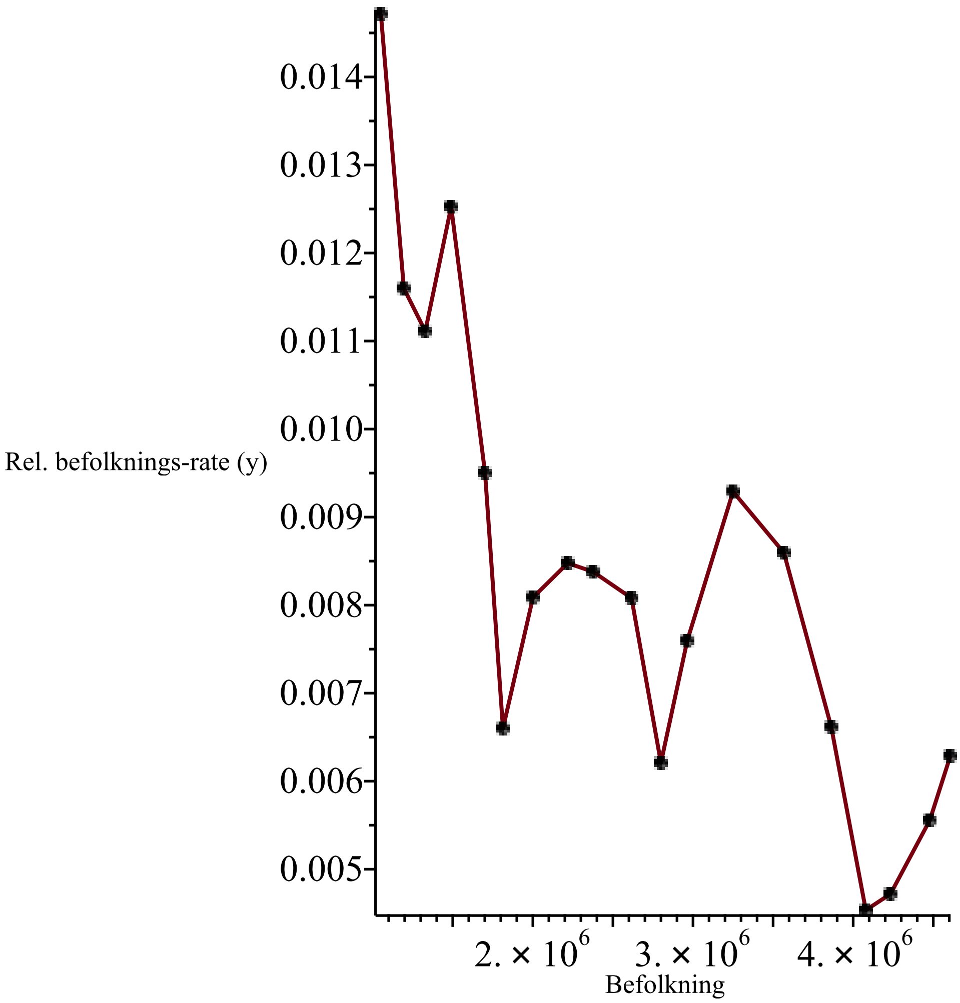
$$Pr := i \mapsto evalf\left(\frac{B_{i+1}-B_{i-1}}{(T_{i+1}-T_{i-1})\cdot B_i}\right)$$

> $X := B_{2..N-1}$:

> $Y := [seq(Pr(i), i=2..N-1)]$:

> $p2 := PlotData(X, Y, labels = [$"Befolkning", "Rel. befolknings-rate (y)"$], font = [times, roman, 14])$:

%



c) Finn de numeriske verdiene på a og b ved hjelp av minste kvadraters metode.
Fremstill den rette linjen $y = a + b\,P$ i samme aksesystem som grafen i **b)**.

Løsning

Ligningen for den rette linjen $y = a + b\,P$ blir

> a := 'a' : b := 'b' :

> with(CurveFitting) :

> lign := y = LeastSquares(X, Y, P, curve = a + b P) : %

$$y = 0.0133497194149881 - 1.84326460613563 \times 10^{-9}\,P$$

> a := op(1, rhs(lign)); b := op([2, 1], rhs(lign))

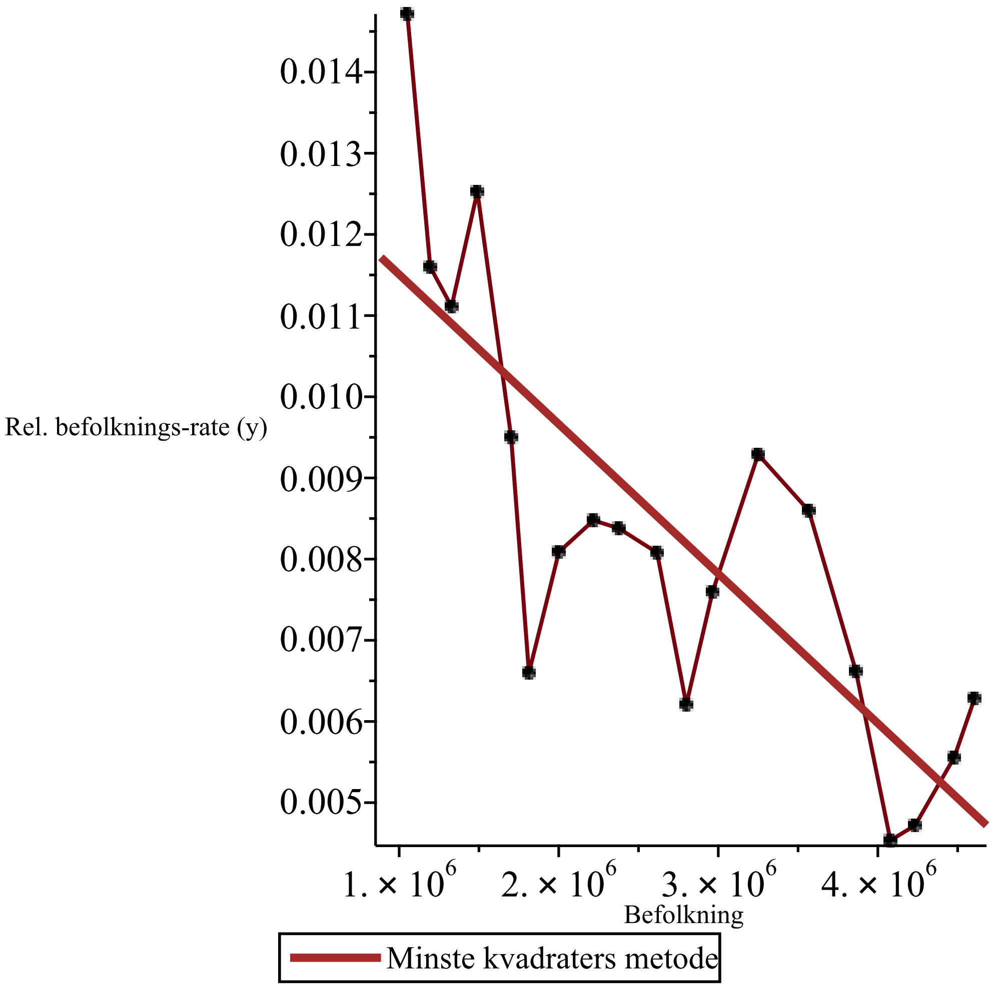
$$a := 0.0133497194149881$$

$$b := -1.84326460613563 \times 10^{-9}$$

> yk := rhs(lign) :

> p3 := plot(yk, P = B[1] ..B[-1], color = brown, thickness = 3, legend = "Minste kvadraters metode") :

> display(p2, p3)



>

d) Løs differensiallignigen $\frac{d}{dt} P(t) = k\, P(t)\, (M - P(t))$, $P(0) = 2.217971 \cdot 10^6$

der k og M er gitt over ved hjelp av de beregnete verdiene for a og b.
Hvilket år er endringen i folkemengden per tidsenhet størst?
Hvor stor er da folkemengden?

Løsning

Av $a = k\, M$ og $b = -k$ fåes

$> k := -b :$

$> M := \frac{a}{k} :$

Differensiallignigen $\frac{d}{dt} P(t) = k\, P(t)\, (M - P(t))$ med startbetingelsen $P(0) = 885431$ kan nå løses.

$> dlign := \frac{\mathrm{d}}{\mathrm{d}t} P(t) = k\, P(t)\, (M - P(t)) :\ \%$

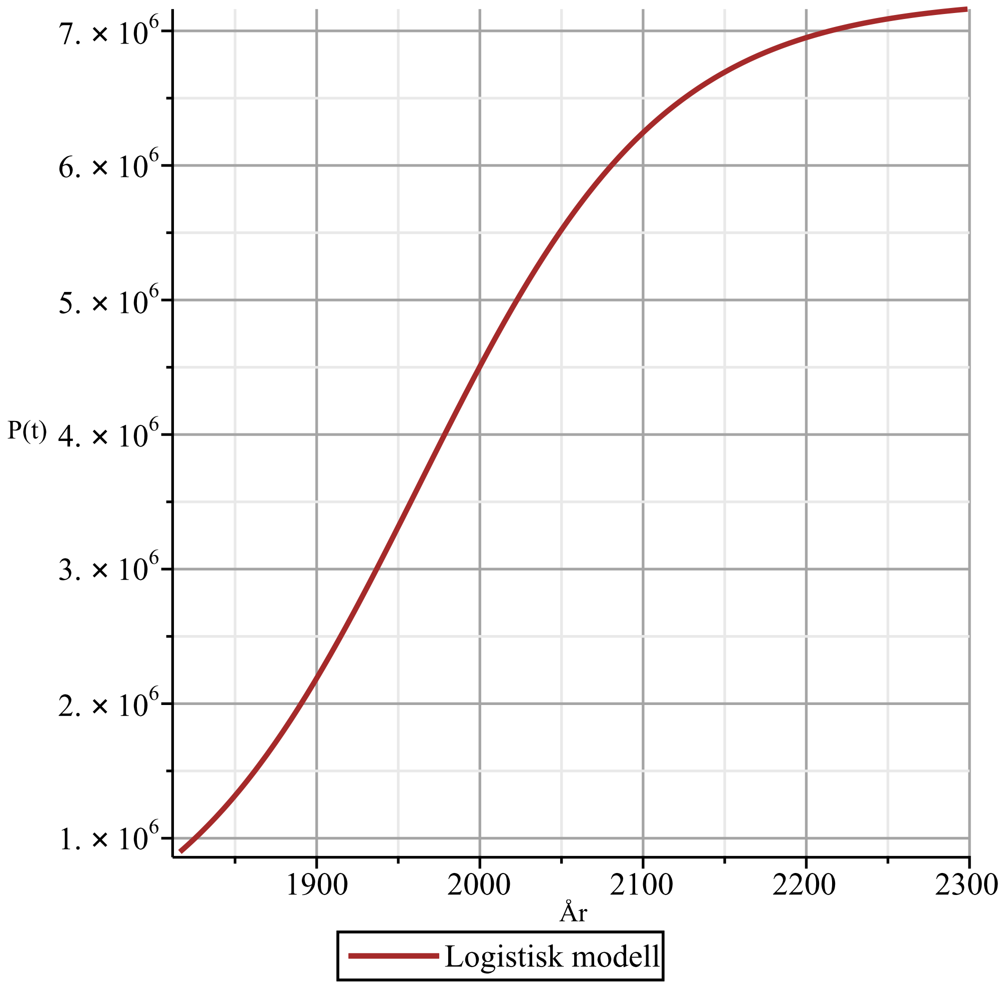
$$\frac{\mathrm{d}}{\mathrm{d}t} P(t) = 1.84326460613563 \times 10^{-9}\, P(t)\, \left(7.24243245953466 \times 10^{6} - P(t)\right)$$

$> sol := dsolve\left(\left\{dlign, P(0) = B_1\right\}, P(t)\right) :\ \%$

$$P(t) = \frac{430347281775991922807}{59420268560384 + 426611146395713\,\mathrm{e}^{-\frac{4332231636297746985122422468237\,t}{324518553658426726783156020576256}}}$$

> $Pf := unapply(rhs(\%), t)$:

> $p4 := plot(Pf(t - 1815), t = 1815..2300, color = brown, thickness = 2, labels = [$"År", "P(t)"$],$
$legend =$ "Logistisk modell", $gridlines)$:
$\%$



>

Hvilket år er endringen i folkemengden per tidsenhet størst.?
Hvor stor er da folkemengden?

Her spørres det etter vendepunktet for kurven, dvs. der $\frac{d^2P(t)}{dt^2} = 0$

> $\frac{\mathrm{d}^2}{\mathrm{d}t^2} P(t) = \mathrm{D}^{(2)}(Pf)(t)$:

som settes lik 0 og løses med hensyn på t.

> $t = fsolve(rhs(\%) = 0, t)$

$$t = 147.6613465$$

Befolkningen er da

> $P(rhs(\%)) = Pf(rhs(\%))$

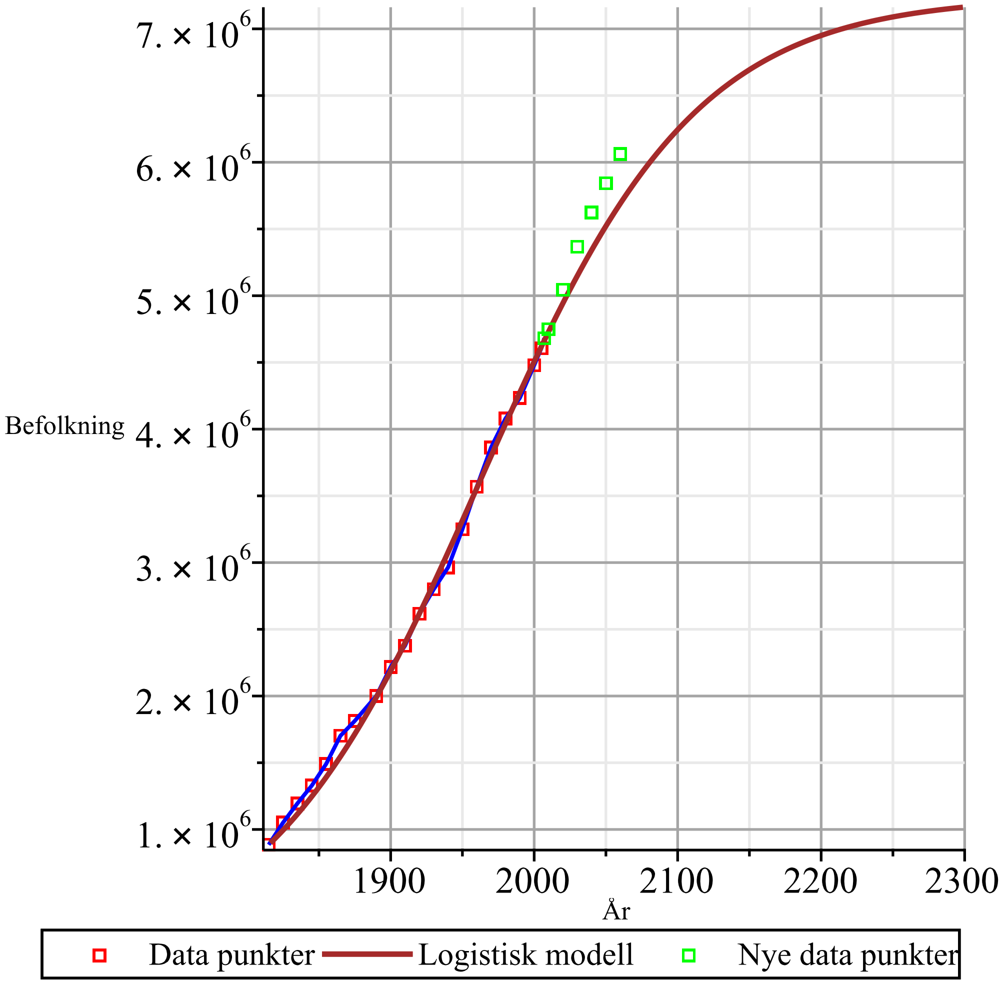
$$P(147.6613465) = 3.621216227 \times 10^6$$

e) Fremstill grafisk i samme aksesystem grafen i **a)**, $P(t)$ og punktmengden L i **d)**.

Løsning

> $p5 := PlotData(X, Y, color = green) :$

> $display(p1, p4, plt3)$



Den logistiske modellen er bra tilpasset den observerte befolkningsutviklingen. De siste dataene (grønne punkter) ligger over den logistiske kurven. Ved å øke den asymptotiske verdien $M = \lim_{t\to\infty} P(t)$.

> $\lim_{t\to\infty} P(t) = evalf\left(\lim_{t\to\infty} Pf(t)\right)$

$$\lim_{t\to\infty} P(t) = 7.242432460 \times 10^{6}$$

oppnår vi en bedre tilpasning.

f) Justér ved prøving og feiling verdien på M slik at den logistiske modellen for folkemengden, $P(t)$, bedre samsvarer med dataene i listen L fra Statistisk sentralbyrå.

Løsning

Vi prøver med

> $M := 7.5\ 10^{6} :$

> $dlign := \frac{d}{dt} P(t) = k\, P(t)\ (M - P(t)) : \%$

$$\frac{d}{dt} P(t) = 1.84326460613563 \times 10^{-9} P(t) \left(7.5000000 \times 10^{6} - P(t)\right)$$

> $dsolve(\{dlign, P(0) = B_1\}, P(t))$

$$P(t) = \frac{6640732500000}{885431 + 6614569 \, e^{-\frac{2089097038817958046875 \, t}{1511157274518286468382 72}}}$$

> $Pf := unapply(rhs(\%), t)$:

> $p4 := plot(Pf(t - 1815), t = 1815 ..2300, color = brown, thickness = 2, legend$ = "Logistisk modell") :

> $display(p1, p4, plt3)$