χ^2-fordelingen

Generelt om χ^2-fordelingen

χ^2-fordelingen er en **kontinuert** fordeling, modsat binomialfordelingen som er en diskret fordeling. Fordelingen er særdeles kompleks at beskrive med matematiske formler. Formlerne opdaget af Pearson omkring år 1900.

http://en.wikipedia.org/wiki/Chi-squared_distribution

http://en.wikipedia.org/wiki/Pearson%27s_chi-squared_test

Græske bogstaver:

χ: chi [udtales "ki"]
ν: ny (**antal frihedsgrader** i χ^2-fordelingen)
μ: my (**middelværdi** i en fordeling)
σ: sigma (**spredning** i en fordeling)

Vi definerer en **stokastisk variabel** ***X***, som er χ^2-fordelt med ν frihedsgrader:

```
> restart
> with(Statistics) :
> X := RandomVariable(ChiSquare(ν)) :
```

Vi beregner middelværdi og spredning (generelt):

```
> μ := Mean(X)
```

$$\mu := \nu \tag{1.1}$$

```
> σ := StandardDeviation(X)
```

$$\sigma := \sqrt{2}\,\sqrt{\nu} \tag{1.2}$$

χ^2-fordelingen har $\underline{\underline{middelværdi = \nu}}$ og $\underline{\underline{spredning = \sqrt{2\cdot\nu}}}$

Test af sandsynlighederne med integralregning:

NB:

$PDF(X, x)$ angiver "ProbabilityDensityFunction" (**tæthedsfunktionen** for χ^2-fordelingen)
$CDF(X, x)$ angiver "CumulativeDistributionFunction" (den kumulerede sandsynlighedsfordeling = **fordelingsfunktionen** for χ^2-fordelingen)

Den samlede sandsynlighed i χ^2-fordelingen skal være 100%, dvs. 1:

$$> \int_0^{\infty} PDF(X, x)\, dx$$

(1.3)

1 **(1.3)**

Grafer over χ^2-fordelingen (grafen afhænger af ν = antal frihedsgrader)

Hvordan ser grafen for χ^2-fordelingen ud?
Lad os variere antal frihedsgrader ν fra 1 til 10.

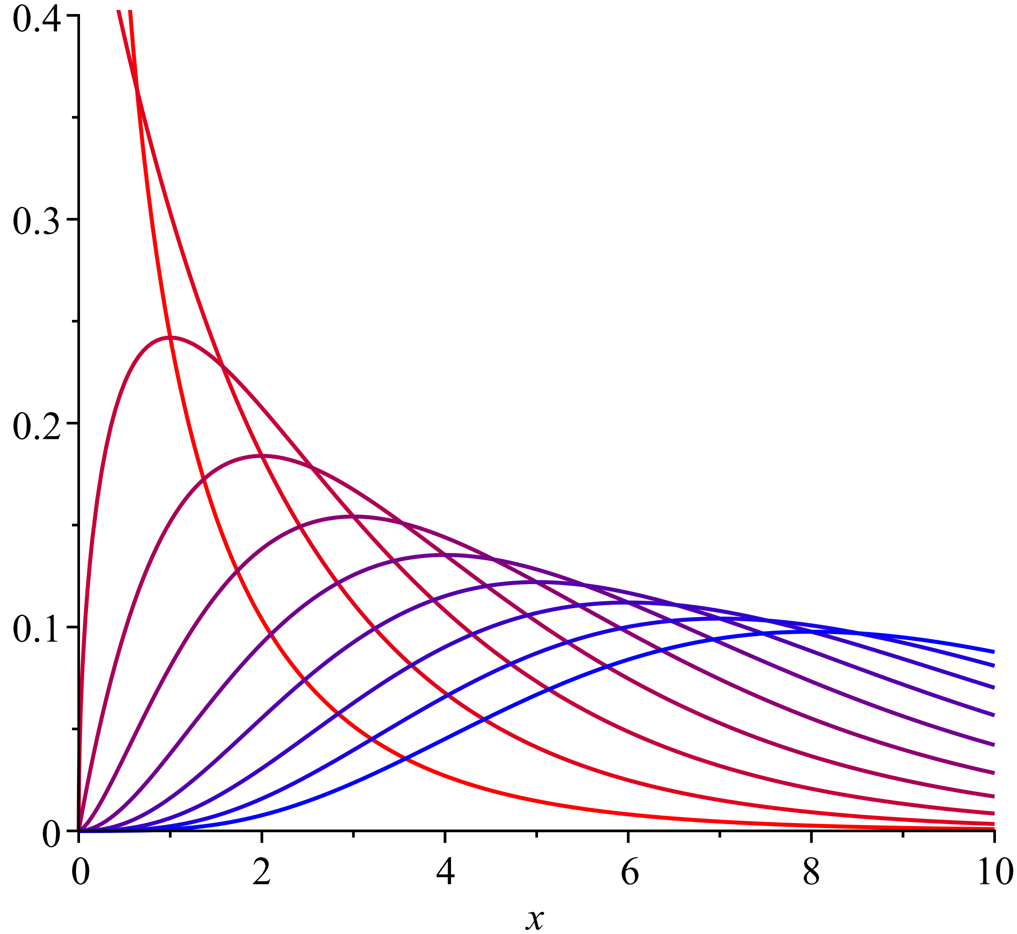
Vi vil gerne tegne graferne i samme koordinatsystem.
Først beregnes alle graferne, og gemmes i variablen $PlotPDFChi2_\nu$ hhv. $PlotCDFChi2_\nu$.
Og graferne skal have forskellig farvetone.
Derefter tegnes alle graferne i samme koordinatsystem med kommandoen *display*:

> *with*(*plots*) :

Tæthedsfunktionen PDF for χ^2-fordelingen:

> **for** ν **from** 1 **by** 1 **to** 10 **do** $PlotPDFChi2_\nu := plot\left(PDF(X, x), x = 0..10, view = [0..10, 0..0.4], color = COLOR\left(RGB, \frac{(10-\nu)}{9}, 0, \frac{(\nu-1)}{9}\right)\right)$ **end do**:

> $display\left(\left[seq\left(PlotPDFChi2_\nu, \nu = 1..10\right)\right], caption = typeset(\text{"PDF for Chi2-fordelingen med antal frihedsgrader } \nu \text{ fra 1 (rød) til 10 (blå)"})\right)$

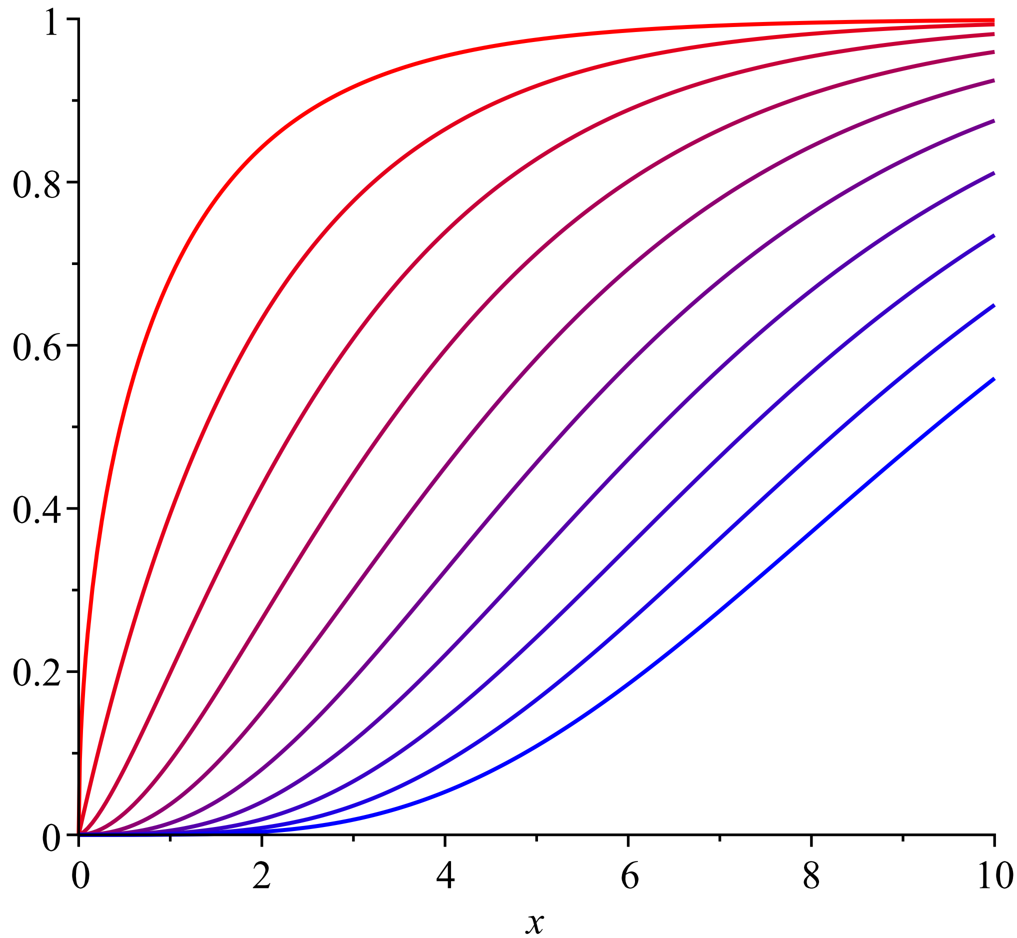


PDF for Chi2-fordelingen med antal frihedsgrader ν fra 1 (rød) til 10 (blå)

Kumulerede sandsynlighed CDF for χ^2-fordelingen:

> **for** ν **from** 1 **by** 1 **to** 10 **do** $PlotCDFChi2_{\nu} := plot\left(CDF(X, x), x = 0 ..10, view = [0 ..10, 0 ..1], color = COLOR\left(RGB, \frac{(10-\nu)}{9}, 0, \frac{(\nu-1)}{9}\right)\right)$ **end do**:

> $display\left(\left[seq\left(PlotCDFChi2_{\nu}, \nu = 1 ..10\right)\right], caption = typeset(\text{"CDF for Chi2-fordelingen med antal frihedsgrader ν fra 1 (rød) til 10 (blå)"})\right)$



CDF for Chi2-fordelingen med antal frihedsgrader ν fra 1 (rød) til 10 (blå)

Brug af Gym-pakken

statistiske funktioner

chicdf chipdf invchi

statistiske test og værktøjer

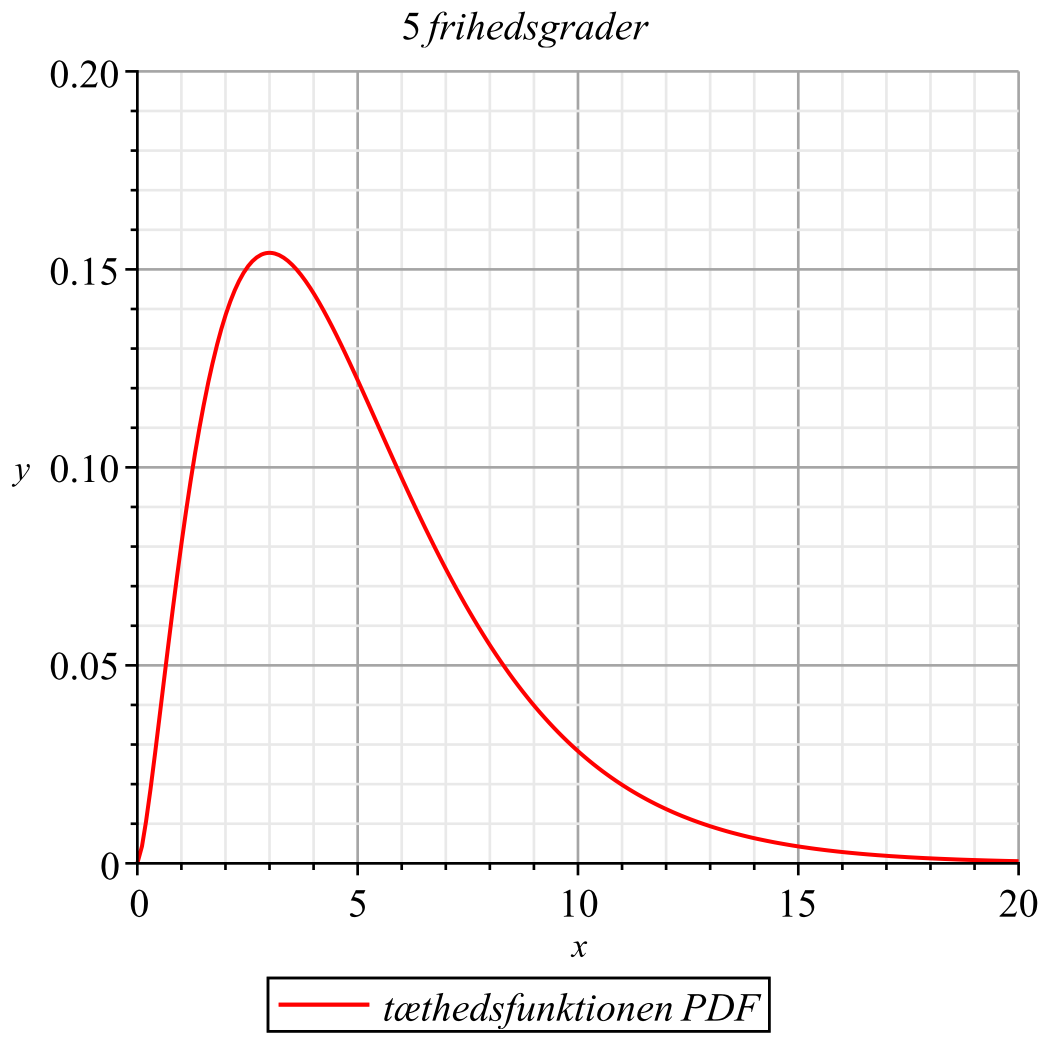
chiKvadratGOFtest chiKvadratUtest

```
> restart
> with(Gym) :
```

Gym-pakken indeholder bl.a. "**chipdf**" og "**chicdf**".
Hermed kan χ^2-fordelingen lettere tegnes.

Tæthedsfunktionen med 5 frihedsgrader:

> $plot(chipdf(5, x), x = 0..20, y = 0..0.2, gridlines, color = red, legend = tæthedsfunktionen\ PDF, title = 5\ frihedsgrader)$



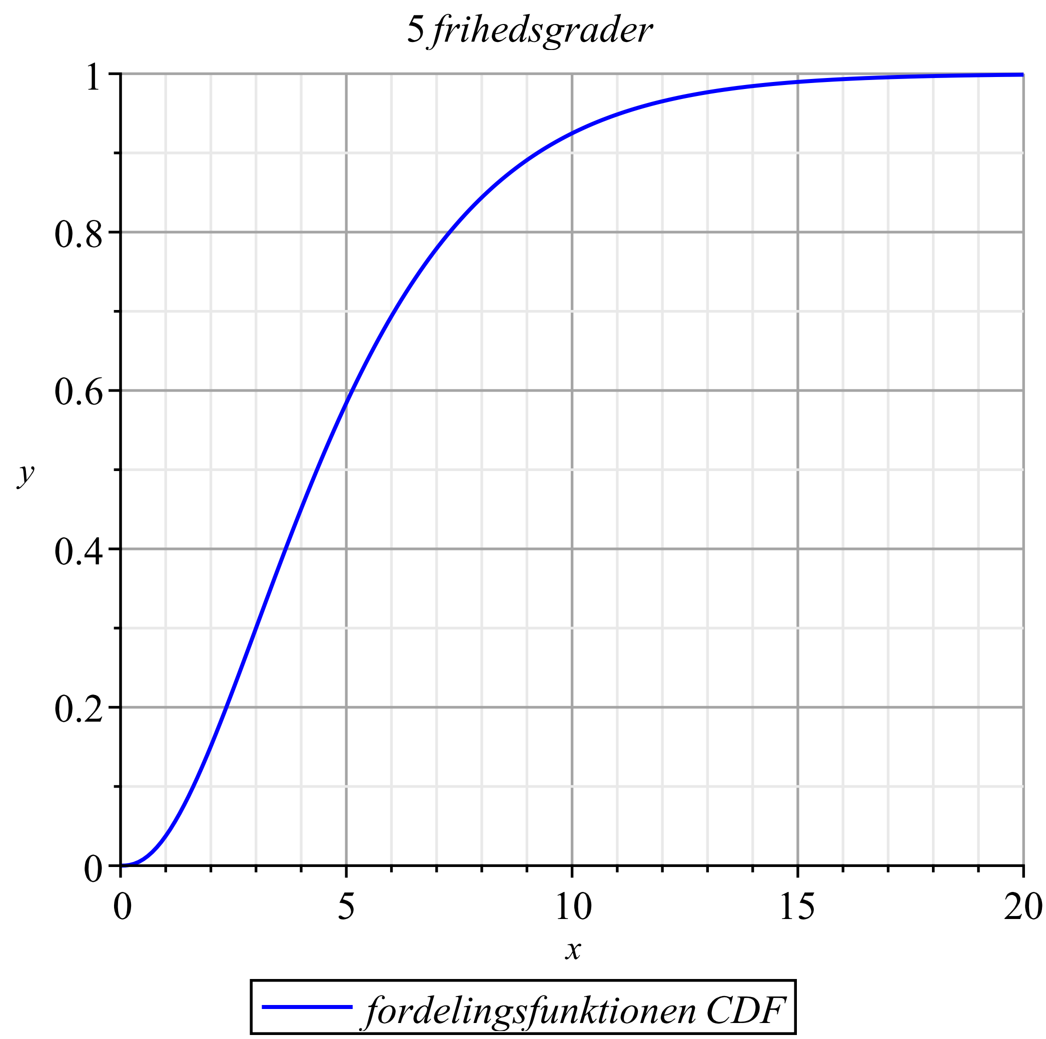
Arealet under grafen bør give 1 (dvs. 100% i sandsynlighed):

> $\int_0^\infty chipdf(5, x)\, dx$

$$0.9999999997 \tag{3.1}$$

Fordelingsfunktionen med 5 frihedsgrader:

> $plot(chicdf(5, x), x = 0..20, y = 0..1, gridlines, color = blue, legend = fordelingsfunktionen\ CDF, title = 5\ frihedsgrader)$



Grafen er voksende, og skal nærme sig 1 (dvs. 100%), når $x \rightarrow \infty$:

> $\lim_{x \rightarrow \infty} chicdf(5, x)$

$$1. \tag{3.2}$$