Flere stokastiske variable

Læsning:

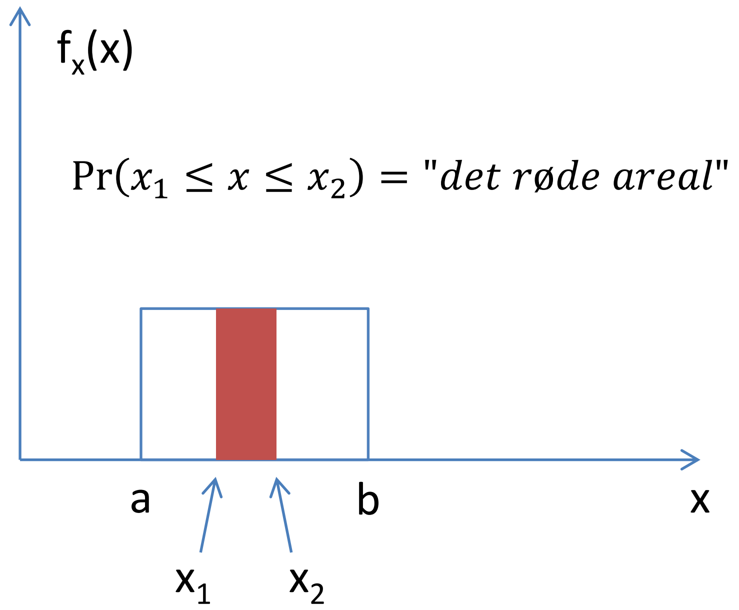
Cooper/McGillem kap. 3.1-3.5

Opsummering

- Vi indførte begrebet *stokastisk variabel* samt tæthedsfunktionen, $f_x(x)$, til at beskrive sandsynligheder for s.v.
- Beregning af sandsynligheder

$$
\begin{aligned}
\Pr(X = x) &= 0 \\
\Pr(X \leq x) &= \int_{-\infty}^{x} f_X(x)dx \\
&= F_X(x) \\
\Pr(x_1 \leq X \leq x_2) &= \int_{x_1}^{x_2} f_X(x)dx \\
&= F_X(x_2) - F_X(x_1) \\
\Pr(X > x) &= 1 - \Pr(X \leq x) \\
&= 1 - F_X(x)
\end{aligned}
$$

- hvor $F_X(x)$ kaldes fordelingsfunktionen.
- Middelværdi og varians.
- Transformationssætningen.



Middelværdi og varians for forskellige fordelinger

- Uniform fordeling i intervallet [a,b]
  - $\mathrm{E}[X] = \frac{1}{2}(b-a)$
  - $\mathrm{Var}(X) = \frac{1}{12}(b-a)^2$
- Binomial fordeling med n trials og sandsynlighedsparameter p
  - $\mathrm{E}[X] = n \cdot p$
  - $\mathrm{Var}(X) = n \cdot p \cdot (1-p)$
- Gauss fordeling (normalfordeling)
  - $f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-(x-\mu)^2/2\sigma^2}$
  - $\mathrm{E}[X] = \bar{X} = \mu$
  - $\mathrm{Var}(X) = \sigma^2$

Exponentialfordelingen

- Beskriver hændelser, som sker tilfældigt i tid.
- Parameteren $\bar{\tau}$ er den gennemsnitlige tid mellem to på hinanden følgende hændelser.
- Sandsynligheden for, at hændelsen indtræffer inden for tidsintervallet Δt, er $\frac{\Delta t}{\bar{\tau}}$.
- Tæthedsfunktion: $f(\tau) = \frac{1}{\bar{\tau}} e^{-\tau/\bar{\tau}}$
- Fordelingsfunktion: $F(\tau) = 1 - e^{-\tau/\bar{\tau}}$

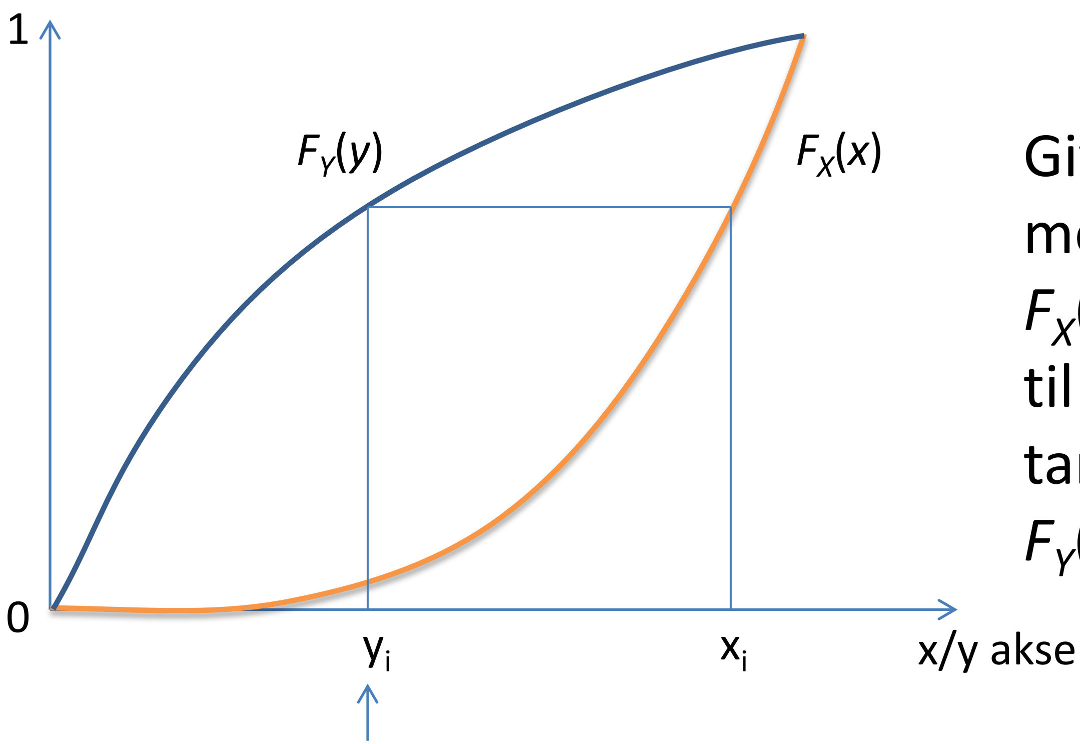
Exponentialfordelingen
Eksempel fra bogen s. 93

- Den gennemsnitlige levetid for komponenter i et rumfartøj er 100 dage.
- Et rumfartøj skal på en 200 dage lang mission.
- Hvad er sandsynligheden for, at missionen kan gennemføres uden komponentfejl?
- Dette er det samme som at spørge: Hvad er sandsynligheden for, at tiden til den første fejl er større end 200 dage?
- Beregn ud fra exponentialfordeling med $\bar{\tau} = 100\ dage$:

$$\Pr(\tau > 200\ dage) = 1 - \Pr(\tau \leq 200\ dage) = 1 - F(\tau)$$
$$= 1 - \left(1 - e^{-\frac{200}{100}}\right) = 0.1352$$

Histogram matching

(dette relaterer til s. 89-91 i bogen)

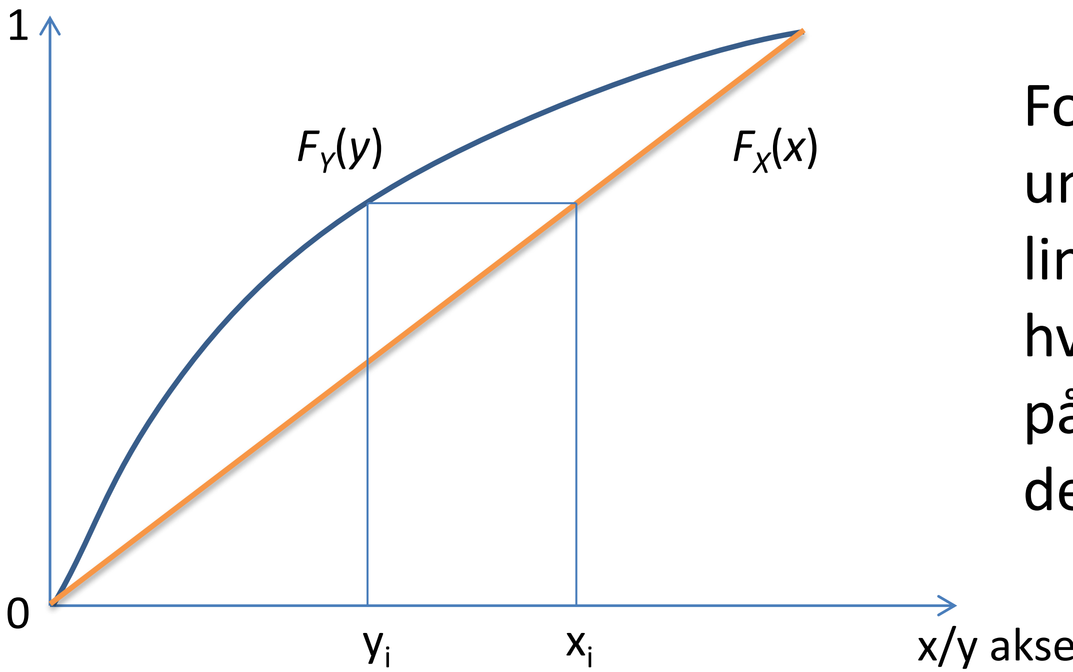


Givet input data $(x_1,x_2,\ldots,x_n)$ med fordelingsfunktion $F_X(x)$, transformer disse data til $(y_1,y_2,\ldots,y_n)$, så disse får target fordelingsfunktion $F_Y(y)$.

$$y = F_Y^{-1}[F_X(x)]$$

(Formel 2-41 side 89)

Histogram matching
hvor input fordelingen er uniform



$$y = F_Y^{-1}(x)$$

(Formel 2-41 side 89)

Fordelingsfunktionen for en uniform fordeling er en ret linje. I det specielle tilfælde, hvor fordelingen er uniform på intervallet [0,1], gælder der, at $F_X(x) = x$

Brug af den uniforme fordeling til at generere data med arbitrær fordeling

1. Specificer en target fordelingsfunktion (F_Y).

```
yrange = -5:0.1:5;
Fy = normcdf(yrange,0,1);
```

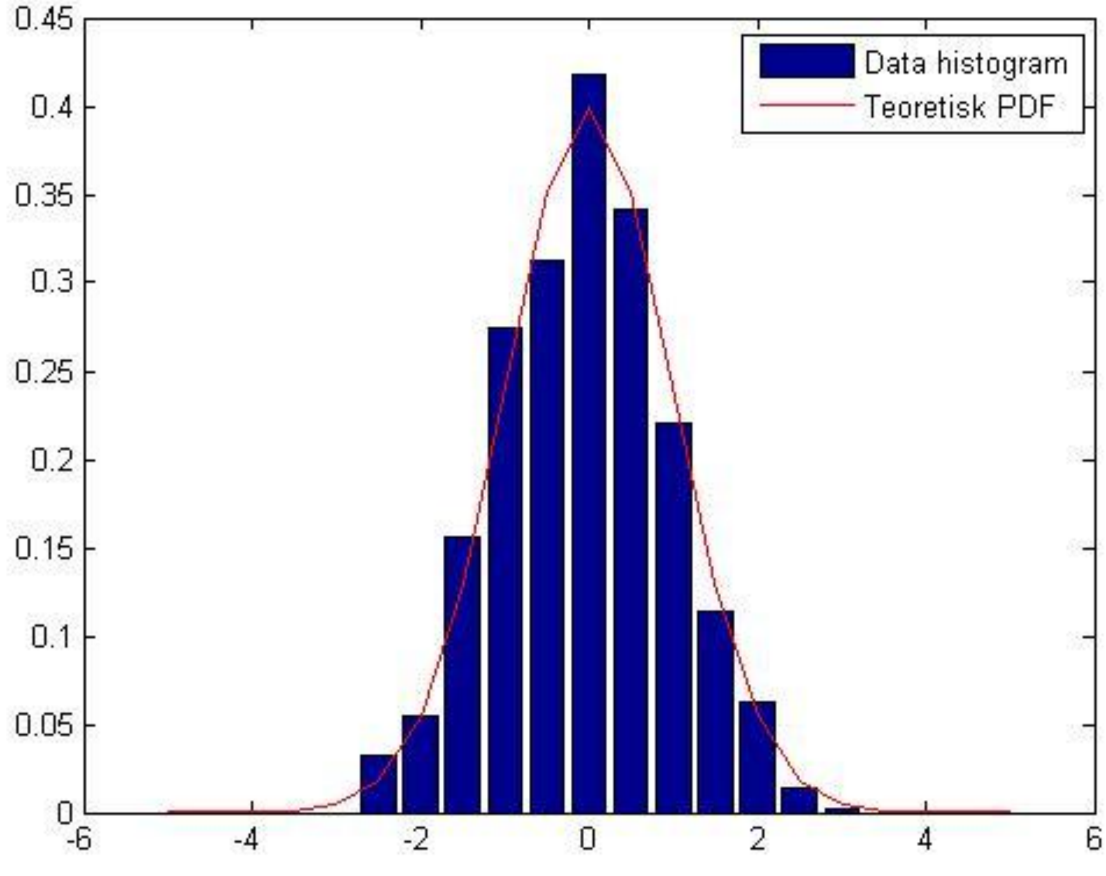
2. Generer tilfældige data, som er uniformt fordelte på intervallet [0,1].

```
N = 1000;
x = rand(1,N);
Fx = unifcdf(x,0,1);
```

3. Histogram matching

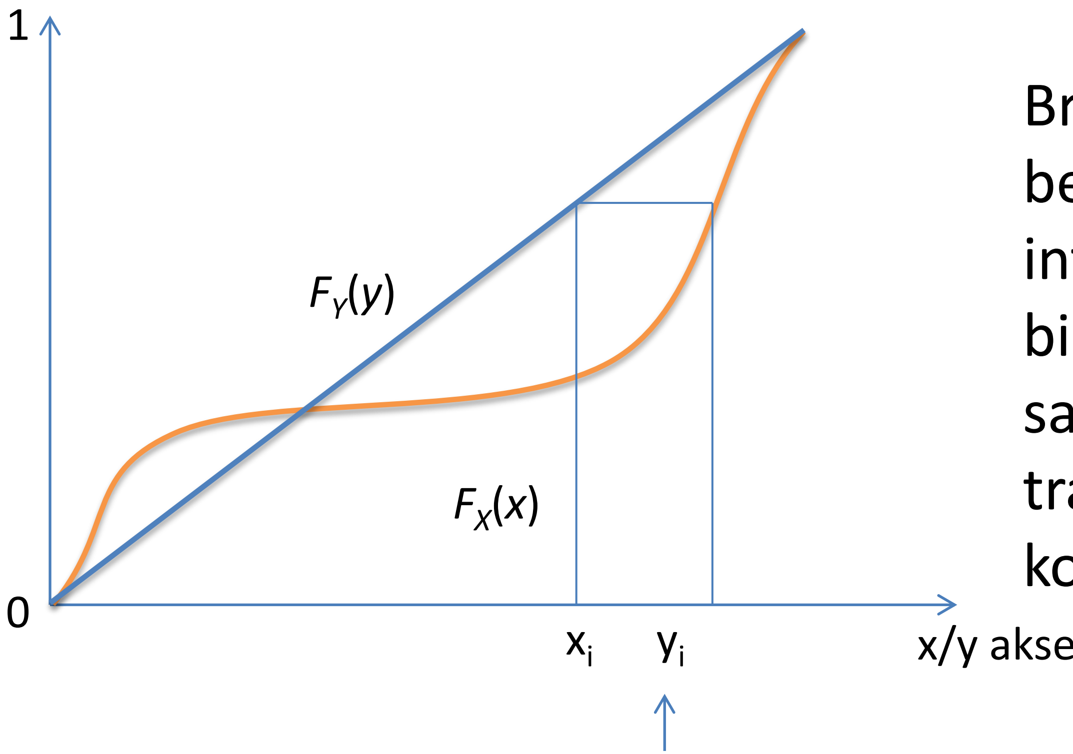
```
for i = 1:N
    distance = abs(Fx(i)-Fy);
    [min_val,min_ix] = min(distance);
    y(i) = yrange(min_ix);
end
```

Brug af den uniforme fordeling til at generere data med arbitrær fordeling



Histogram over de genererede y-værdier

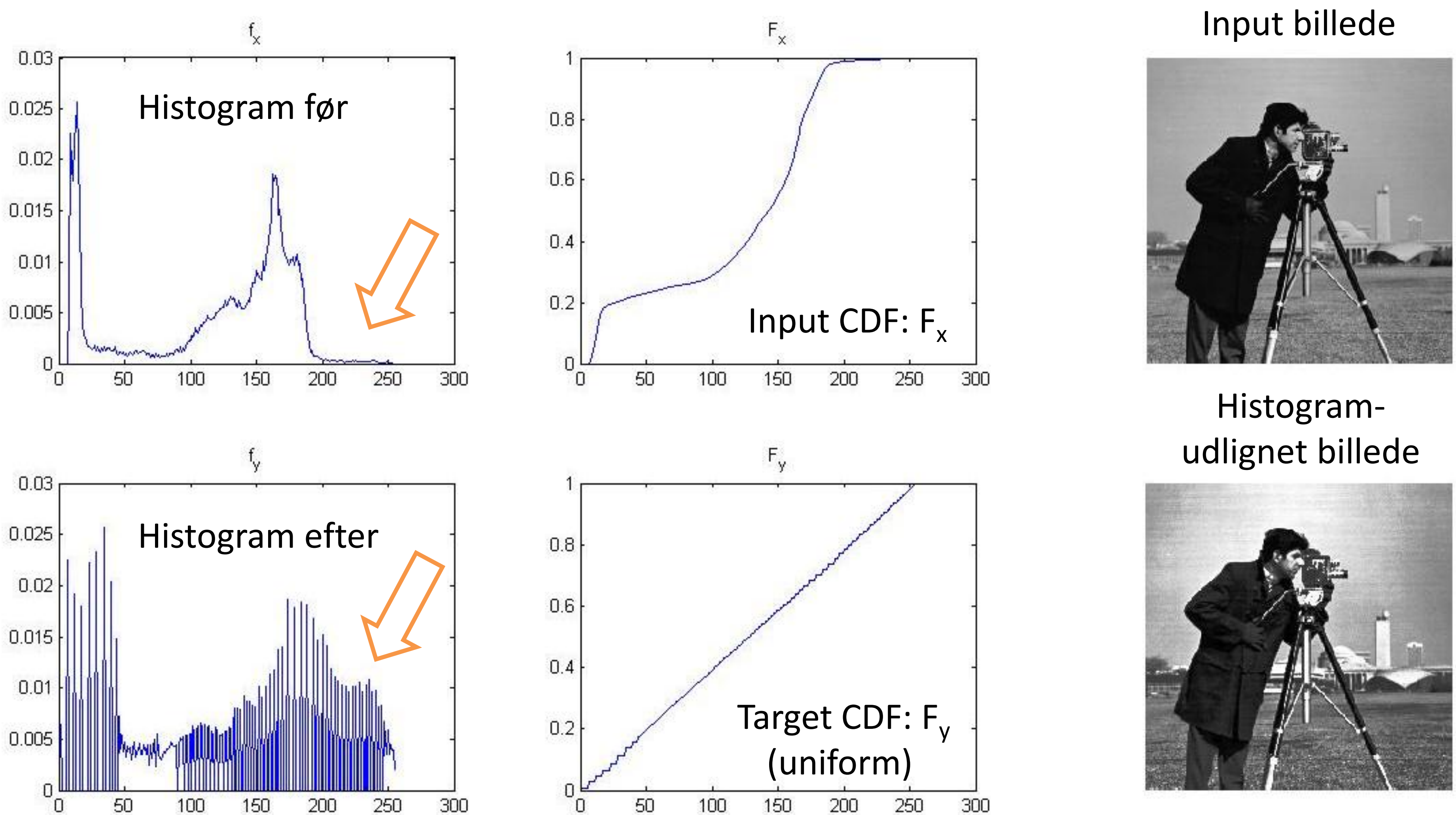
Histogram matching
hvor target fordelingen er uniform



$$y = F_Y^{-1}[F_X(x)]$$

Bruges ofte i billed-behandling. Hvis intensiteterne (x_i) i input billedet er klumpet sammen, vil denne tranformation øge kontrasten i output billedet.

Histogram udligning i billeder



Simultan tæthedsfunktion *f*(*x*,*y*)

	1Ω	10Ω	100Ω	1000Ω	Total
1W	50	300	90	0	440
2W	50	50	0	100	200
5W	0	150	60	150	360
Total	100	500	150	250	1000

Omregn til sandsynligheder (relativ frekvens)

	1Ω	10Ω	100Ω	1000Ω	Total
1W	0.05	0.30	0.09	0	0.44
2W	0.05	0.05	0	0.10	0.20
5W	0	0.15	0.06	0.15	0.36
Total	0.10	0.50	0.15	0.25	1.0

Simultan tæthedsfunktion *f*(*x*,*y*)

	$\mathbf{X = 1}$	$\mathbf{X = 2}$	$\mathbf{X = 3}$	$\mathbf{X = 4}$	$f_Y(y)$
$\mathbf{Y = 1}$	$f(1,1)$	$f(2,1)$	$f(3,1)$	$f(4,1)$	$f_Y(1)$
$\mathbf{Y = 2}$	$f(1,2)$	$f(2,2)$	$f(3,2)$	$f(4,2)$	$f_Y(2)$
$\mathbf{Y = 3}$	$f(1,3)$	$f(2,3)$	$f(3,3)$	$f(4,3)$	$f_Y(3)$
$f_X(x)$	$f_X(1)$	$f_X(2)$	$f_X(3)$	$f_X(4)$	1

Indfør stokastiske variable samt *f*(*x*,*y*)

	1Ω	**10Ω**	**100Ω**	**1000Ω**	**Total**
1W	0.05	0.30	0.09	0	0.44
2W	0.05	0.05	0	0.10	0.20
5W	0	0.15	0.06	0.15	0.36
Total	0.10	0.50	0.15	0.25	1.0

Ex.: $f(x = 1, y = 1) = 0.05$

Marginale tæthedsfunktioner

	$\mathbf{X = 1}$	$\mathbf{X = 2}$	$\mathbf{X = 3}$	$\mathbf{X = 4}$	$f_Y(y)$
$\mathbf{Y = 1}$	$f(1,1)$	$f(2,1)$	$f(3,1)$	$f(4,1)$	$f_Y(1)$
$\mathbf{Y = 2}$	$f(1,2)$	$f(2,2)$	$f(3,2)$	$f(4,2)$	$f_Y(2)$
$\mathbf{Y = 3}$	$f(1,3)$	$f(2,3)$	$f(3,3)$	$f(4,3)$	$f_Y(3)$
$f_X(x)$	$f_X(1)$	$f_X(2)$	$f_X(3)$	$f_X(4)$	1

- $f_X(x)$ og $f_Y(y)$ kaldes for de *marginale* tæthedsfunktioner.

- Hvordan udregner man $f_X(2)$?

$$f_X(2) = \sum_{y=1}^{3} f(2, y) = 0.3 + 0.05 + 0.15 = 0.5$$

Betingede tæthedsfunktioner

	$\mathbf{X = 1}$	$\mathbf{X = 2}$	$\mathbf{X = 3}$	$\mathbf{X = 4}$	$f_Y(y)$
$\mathbf{Y = 1}$	$f(1,1)$	$f(2,1)$	$f(3,1)$	$f(4,1)$	$f_Y(1)$
$\mathbf{Y = 2}$	$f(1,2)$	$f(2,2)$	$f(3,2)$	$f(4,2)$	$f_Y(2)$
$\mathbf{Y = 3}$	$f(1,3)$	$f(2,3)$	$f(3,3)$	$f(4,3)$	$f_Y(3)$
$f_X(x)$	$f_X(1)$	$f_X(2)$	$f_X(3)$	$f_X(4)$	1

- Hvordan udregner man de betingede tæthedsfunktioner $f(x|y)$ og $f(y|x)$?

- Eksempel: $f(x = 2|y = 3)$

$$f(x = 2|y = 3) = \frac{f(x = 2, y = 3)}{f_Y(y = 3)} = \frac{0.15}{0.36}$$

Betingede tæthedsfunktioner

	$\mathbf{X=1}$	$\mathbf{X=2}$	$\mathbf{X=3}$	$\mathbf{X=4}$	$f_Y(y)$
$\mathbf{Y=1}$	$f(1\|1)$	$f(2\|1)$	$f(3\|1)$	$f(4\|1)$	$f_Y(1)$
$\mathbf{Y=2}$	$f(1\|2)$	$f(2\|2)$	$f(3\|2)$	$f(4\|2)$	$f_Y(2)$
$\mathbf{Y=3}$	$f(1\|3)$	$f(2\|3)$	$f(3\|3)$	$f(4\|3)$	$f_Y(3)$
$f_X(x)$	$f_X(1)$	$f_X(2)$	$f_X(3)$	$f_X(4)$	1

- Givet de betingede og de marginale tætheder, hvordan beregner man de simultane tætheder?

- Eksempel: $f(x=2, y=3)$

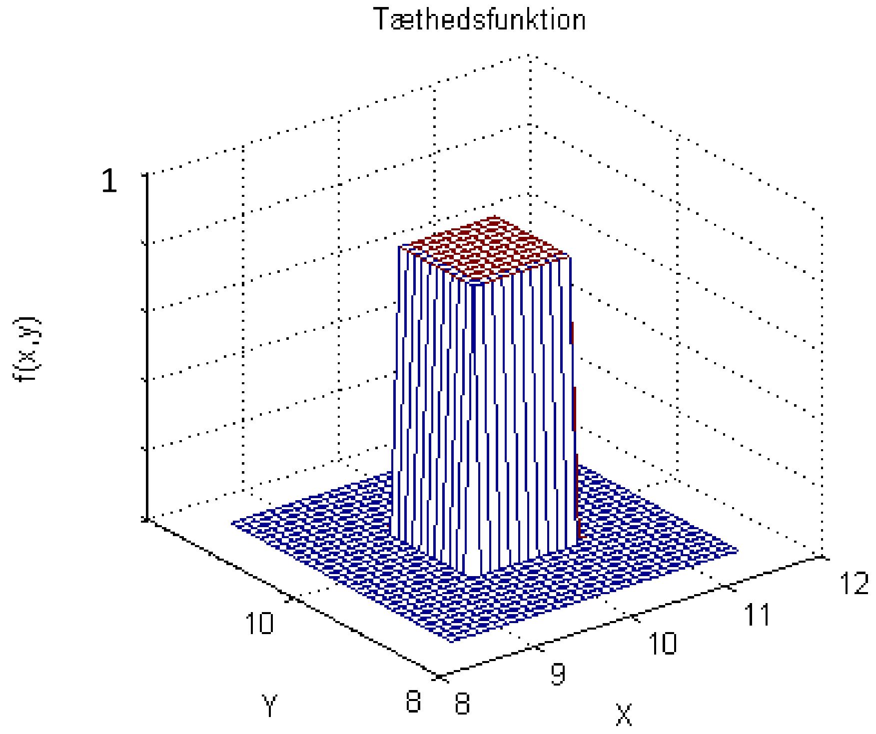
$$\begin{aligned} f(x=2, y=3) &= f(x=2|y=3) \cdot f_Y(y=3) \\ &= f(y=3|x=2) \cdot f_X(x=2) \end{aligned}$$

Lad os nu kigge på kontinuerte simultane tæthedsfunktioner

Eksempel 1

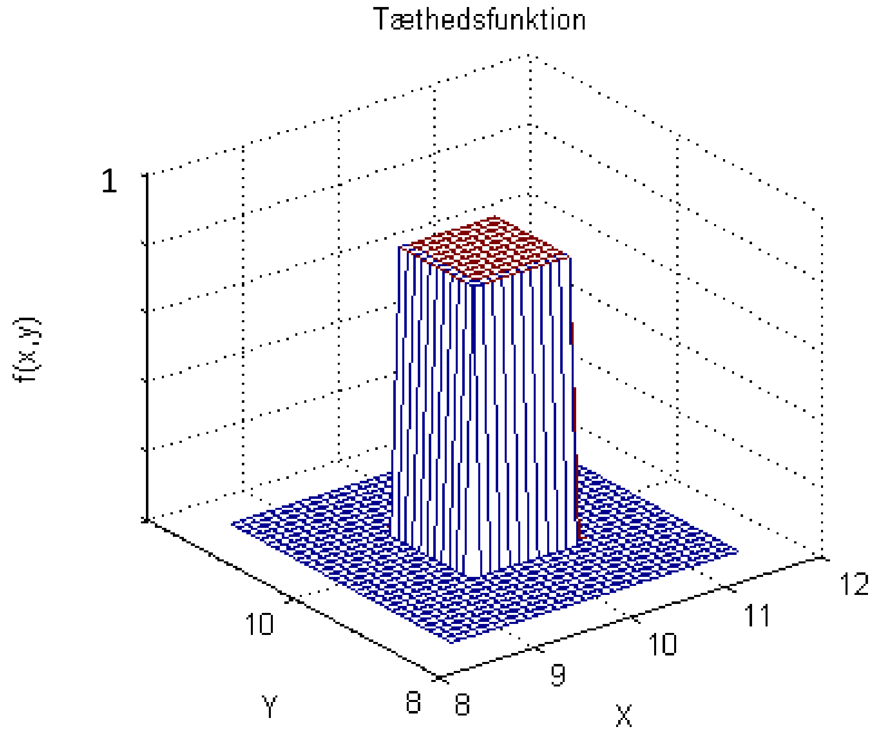
- En maskine skal udskære træplader på 10×10 cm.
- Maskinen er ikke perfekt, så den faktiske størrelse af pladen er stokastisk.
- Vi indfører derfor to stokastiske variable X og Y, således at det faktiske areal er $X \times Y$.
- De marginale tæthedsfunktioner er begge uniforme:
  - $f_X(x) = \begin{cases} 1 & for\ 9.5\ cm \le x \le 10.5\ cm \\ 0 & ellers \end{cases}$
  - $f_Y(y) = \begin{cases} 1 & for\ 9.5\ cm \le y \le 10.5\ cm \\ 0 & ellers \end{cases}$
- Hvordan ser den simultane tæthedsfunktion for (X, Y) ud?

Eksempel 1
Simultan tæthedsfunktion



$$f(x,y) = f_X(x) \cdot f_Y(y) = \begin{cases} 1 & \text{for } 9.5 \leq x, y \leq 10.5 \\ & 0 \text{ ellers} \end{cases}$$

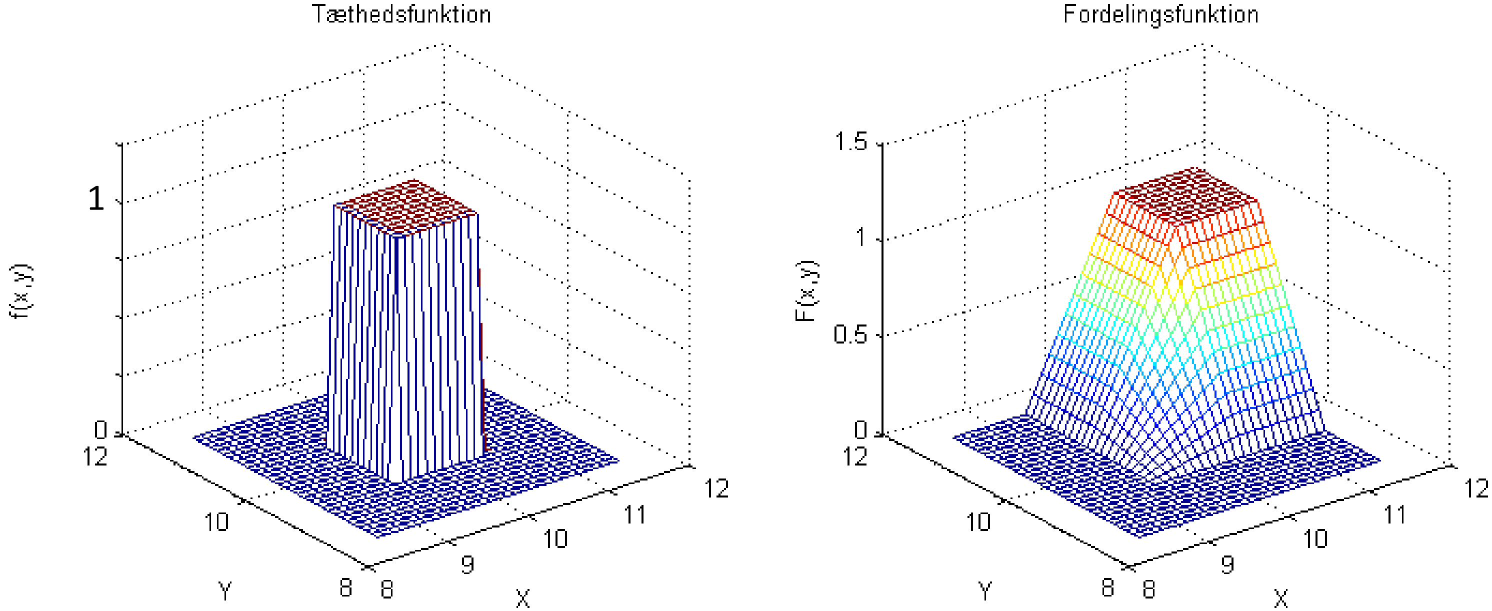
Eksempel 1
Simultan tæthedsfunktion



Hvad er $\Pr(X \leq 10, Y \leq 10)$?

$$\Pr(X \leq 10, Y \leq 10) = \int_{x=9.5}^{10} \int_{y=9.5}^{10} f(x,y)dxdy = \frac{1}{4}$$

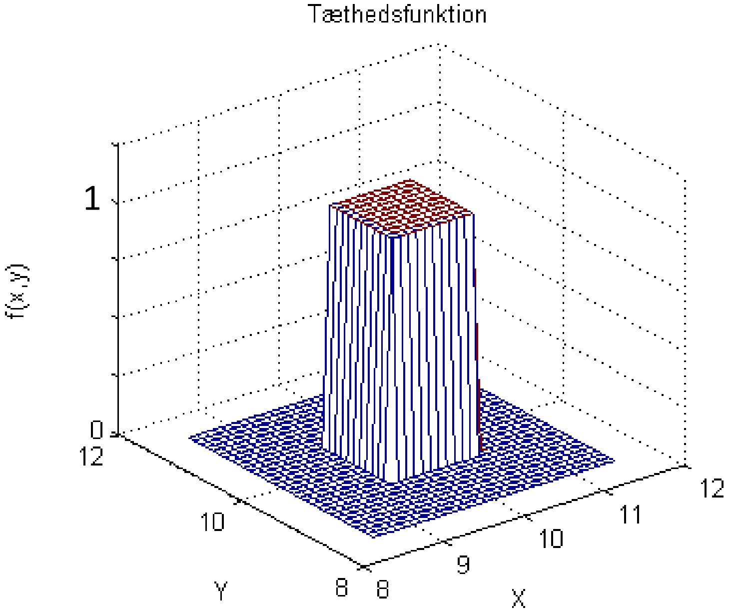
Eksempel 1
Simultan fordelingsfunktion

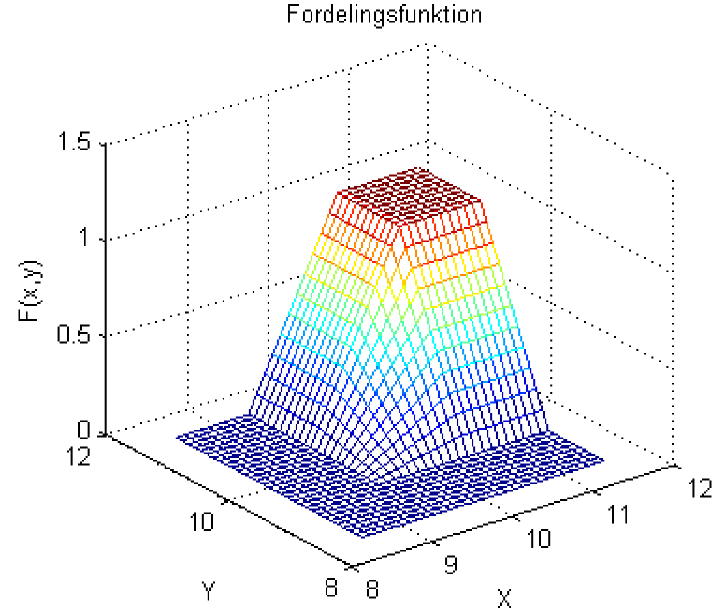


$$f(x,y) = \begin{cases} 1 & \text{for } 9.5 \le x, y \le 10.5 \\ & 0 \text{ ellers} \end{cases}$$

$$F(x,y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(u,v)\,du\,dv$$

Eksempel 1
Simultan fordelingsfunktion





Hvad er $\Pr(X \leq 10, Y \leq 10)$?

$$\Pr(X \leq 10, Y \leq 10) = F(10,10) = \int_{-\infty}^{10} \int_{-\infty}^{10} f(u,v)dudv = \frac{1}{4}$$

Vigtige relationer

- Givet $f(x,y)$, hvordan beregner man $F(x,y)$?
  - $F(x, y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(u, v) du dv$
- Givet $F(x,y)$, hvordan beregner man $f(x,y)$?
  - $f(x, y) = \frac{\partial F(x,y)}{\partial x \partial y}$
- Givet $f(x,y)$, hvordan beregner man de marginale sandsynligheder $f_X(x)$ og $f_Y(y)$?
  - $f_X(x) = \int_{-\infty}^{\infty} f(x, y) dy$
  - $f_Y(y) = \int_{-\infty}^{\infty} f(x, y) dx$

Vigtige relationer (fortsat)

- Hvordan beregnes de betingede tætheder $f(x|y)$ og $f(y|x)$?
  - $f(x|y) = \frac{f(x,y)}{f_Y(y)}$
  - $f(y|x) = \frac{f(x,y)}{f_X(x)}$
- Bayes regel:
  - $f(y|x) = \frac{f(x|y) \cdot f_Y(y)}{f_X(x)}$ og tilsvarende for $f(x|y)$.
- Hvornår er X og Y statistisk uafhængige?
  - Når $f(x|y) = f_X(x)$ og $f(y|x) = f_Y(y)$
  - Dvs. når

$$f(x,y) = f(x|y) \cdot f_Y(y) = f_X(x) \cdot f_Y(y)$$

  - samme som at

$$f(x,y) = f(y|x) \cdot f_X(x) = f_Y(y) \cdot f_X(x)$$

Integration og differentiation
(se youtube videoer – links i kalenderen)

- Fra tæthedsfunktion (f) til fordelingsfunktion (F):
  - Husk, at $F(x,y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(u,v)dudv$
  - Løs først det indre integral (mht. u).
    - Husk, at grænserne (generelt fra $-\infty$ til y) skal sættes ind på u's plads.
    - Den anden integrationsvariabel (v) behandles her som en konstant.
  - Løs derefter det ydre integral (mht. v).
    - Den første integrationsvariabel (u) vil være erstattet med x'er, som nu skal behandles som en konstant.
    - Husk, at grænserne (generelt fra $-\infty$ til x) skal sættes ind på v's plads.
- Fra fordelingsfunktion (F) til tæthedsfunktion (f):
  - $f(x,y) = \frac{\partial F(x,y)}{\partial x \partial y}$
  - Differentier først mht. x og behandl y som en konstant
  - Differentier det resulterende udtryk mht. y og behandl x som en konstant.

Regneeksempel (s. 126-127)

Korrelation

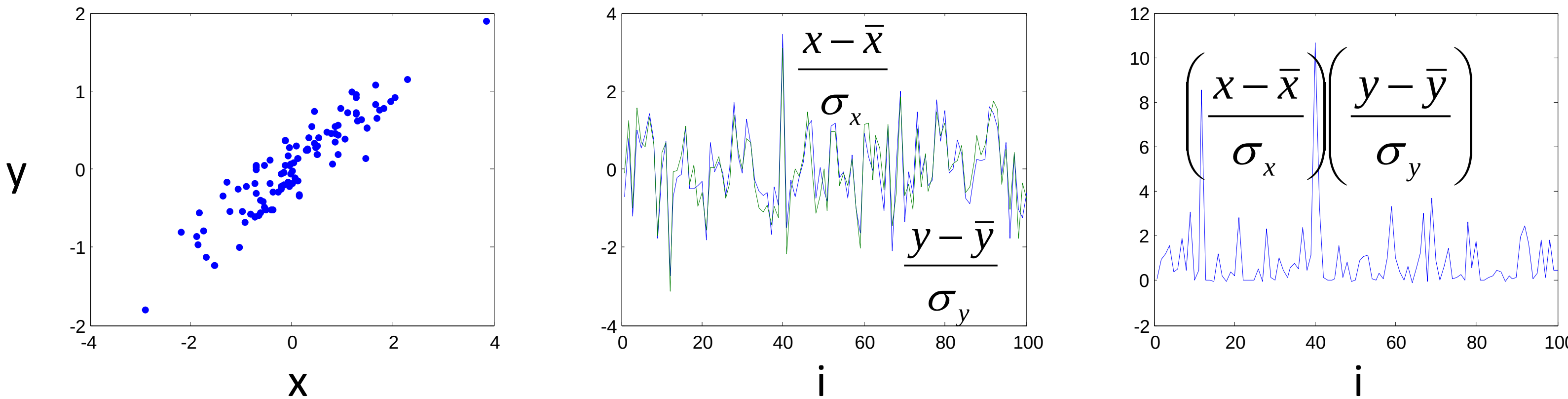
- Afhænger X og Y af hinanden?
- Korrelation
  - $E[X \cdot Y] = \int_{-\infty}^{\infty} x \cdot y \cdot f(x, y) dxdy$
  - Hvis X og Y er uafhængige er $E[X \cdot Y] = E[X] \cdot E[Y]$

- Kovarians
  - $E[(X - \bar{X}) \cdot (Y - \bar{Y})] = \int_{-\infty}^{\infty} (x - \bar{X}) \cdot (y - \bar{Y}) \cdot f(x, y) dxdy$

- Korrelationskoefficient
  - $\rho = E\left[(\frac{X - \bar{X}}{\sigma_X}) \cdot (\frac{Y - \bar{Y}}{\sigma_Y})\right] = \frac{E[X \cdot Y] - E[X] \cdot E[Y]}{\sigma_X \cdot \sigma_Y}$
  - $-1 \leq \rho \leq 1$
  - Hvis X og Y er uafhængige er $\rho = 0$

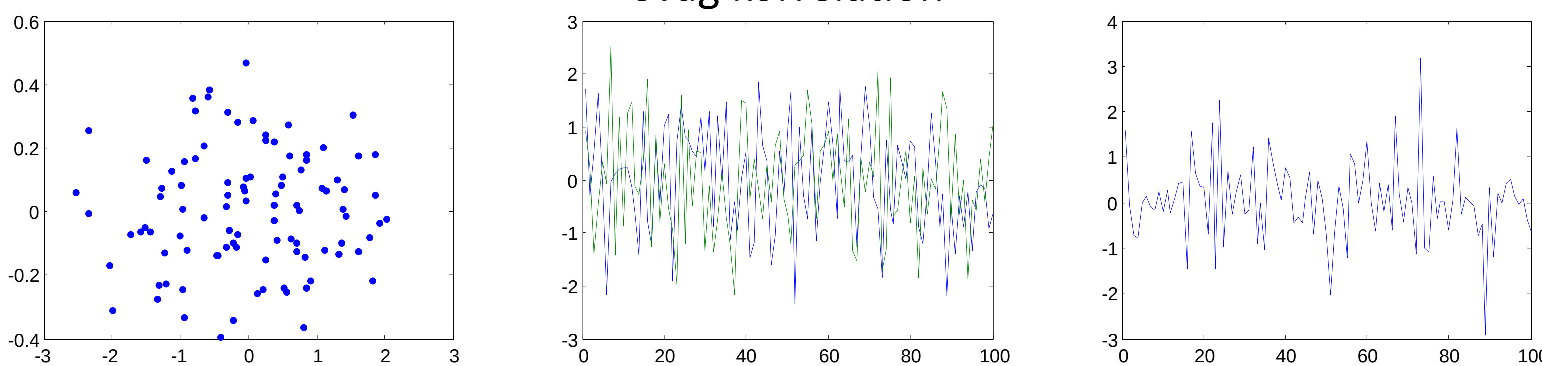
Korrelationskoefficienten for diskrete data

Stærk korrelation

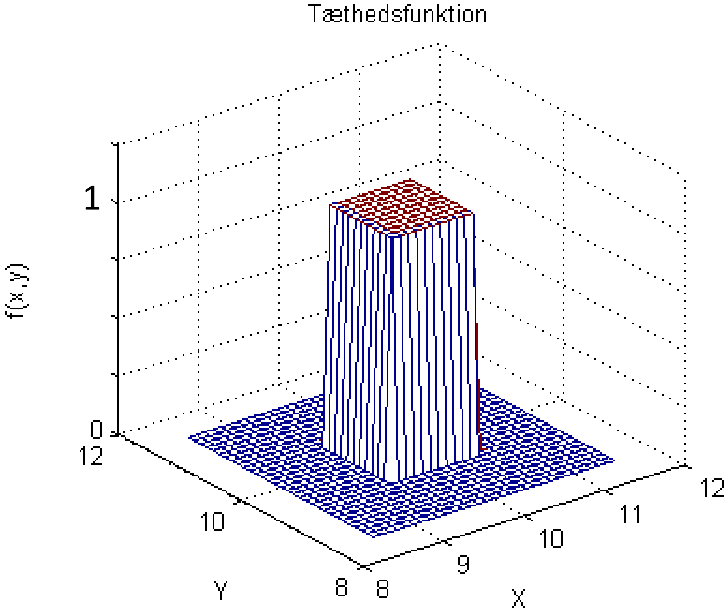


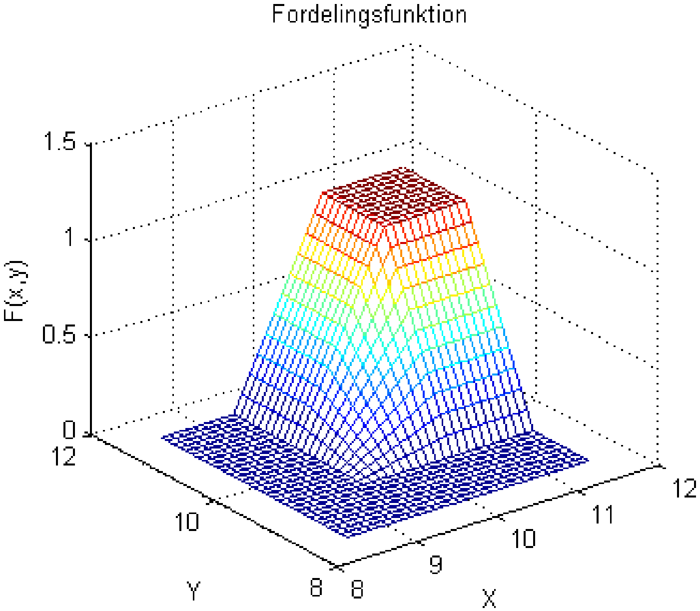
$$\rho = \frac{1}{n-1}\sum_{i=1}^{n}\left(\frac{x-\bar{x}}{\sigma_x}\right)\left(\frac{y-\bar{y}}{\sigma_y}\right)$$

Svag korrelation



Uniform fordeling





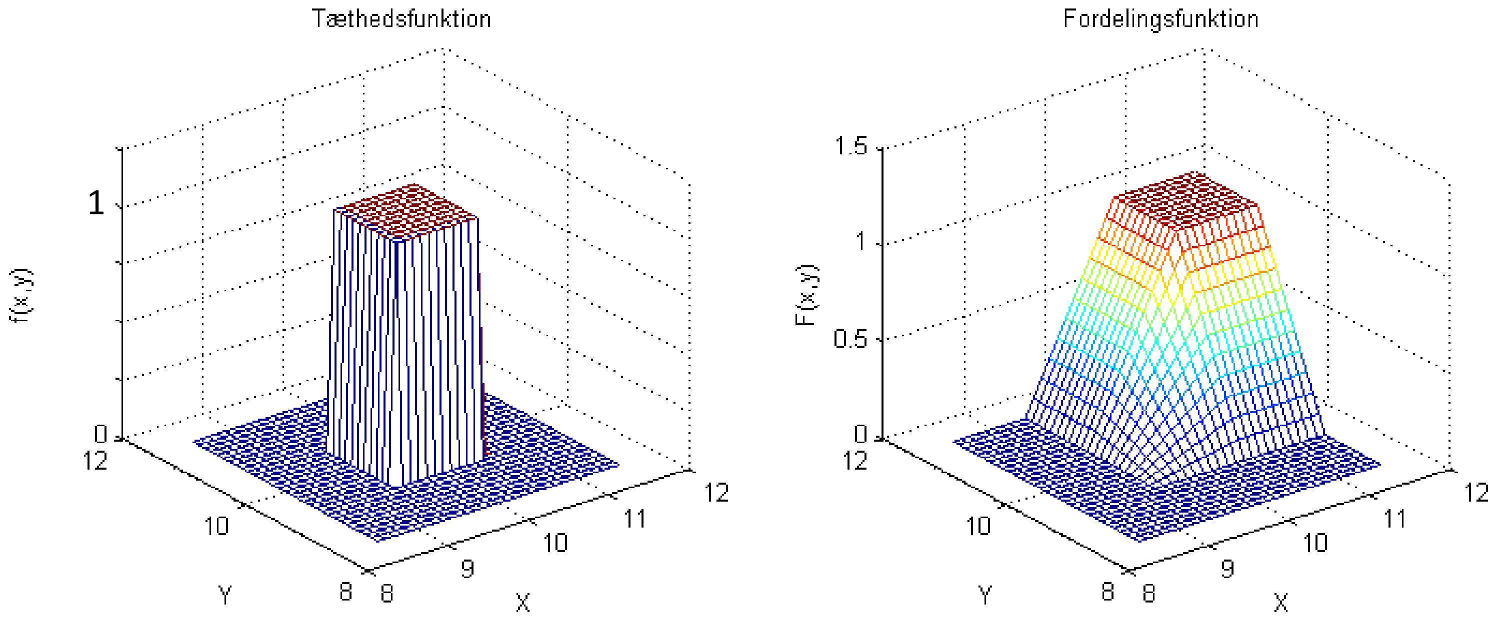
$$f(x,y) = \begin{cases} 1 & \text{for } 9.5 \le x, y \le 10.5 \\ & 0 \text{ ellers} \end{cases}$$

Er $f_X(x)$ og $f_Y(y)$ uafhængige?

Svar: Ja, fordi

$$f_X(x) \cdot f_Y(y) = \begin{cases} 1 & \text{for } 9.5 \le x, y \le 10.5 \\ & 0 \text{ ellers} \end{cases} = f(x,y)$$

Uniform fordeling



Hvad er korrelationen mellem X og Y, når X og Y er statistisk uafhængige?

Svar:

$$\mathrm{E[XY]} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf(x,y)dxdy = \int_{-\infty}^{\infty} xf_X(x)dx \int_{-\infty}^{\infty} yf_Y(y)dy = E[X]\cdot E[Y]$$

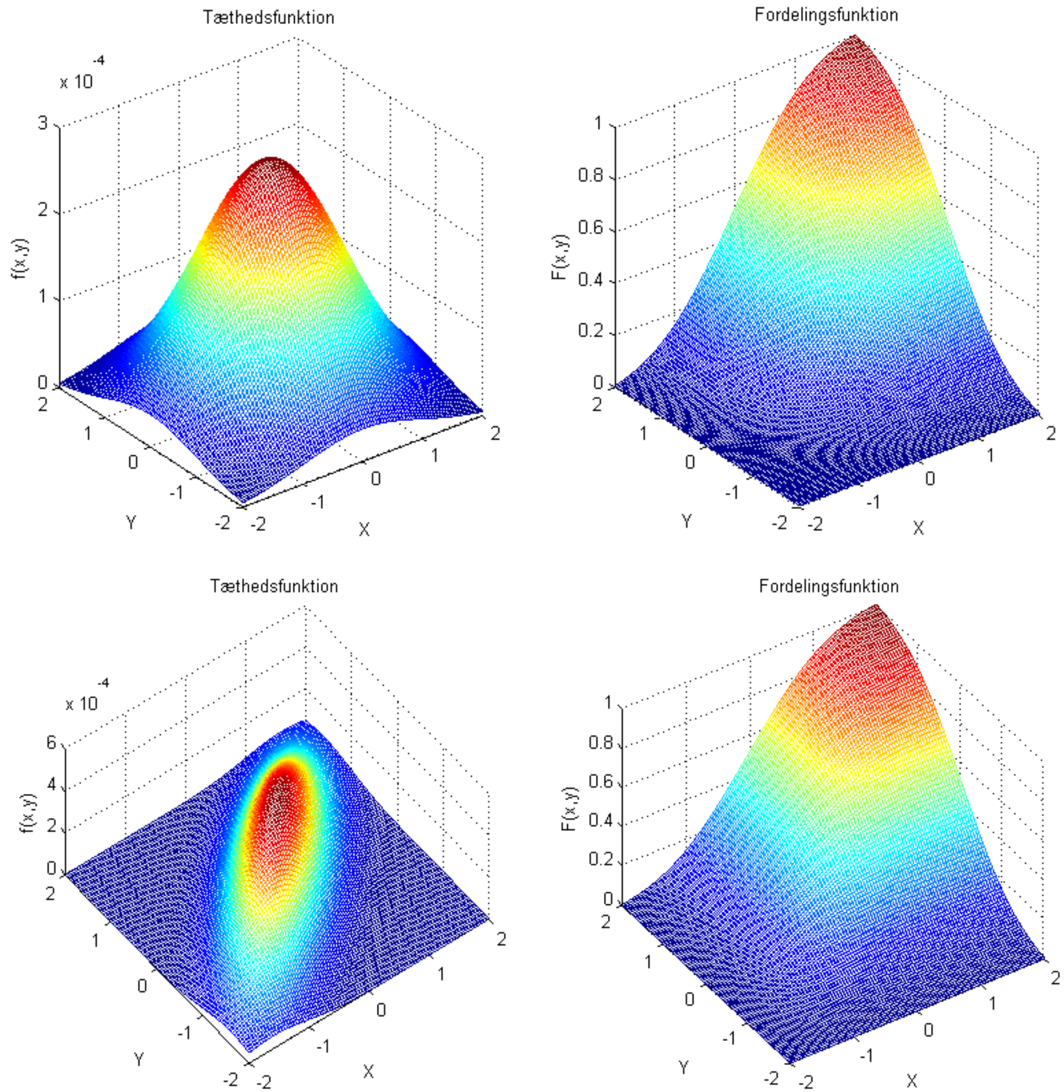
Bivariat Gaussfordeling

- Wikipedia
  - http://en.wikipedia.org/wiki/Multivariate_normal_distribution

$$f(x,y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_x)^2}{\sigma_x^2} + \frac{(y-\mu_y)^2}{\sigma_y^2} - \frac{2\rho(x-\mu_x)(y-\mu_y)}{\sigma_x\sigma_y}\right]\right),$$

  - hvor ρ er korrelationskoefficienten mellem X og Y.

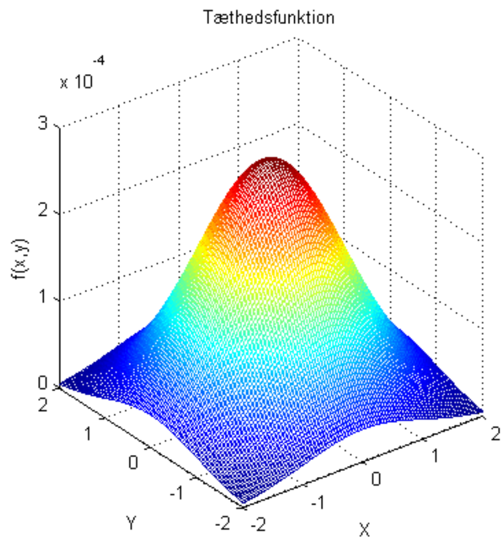
Bivariat Gaussfordeling

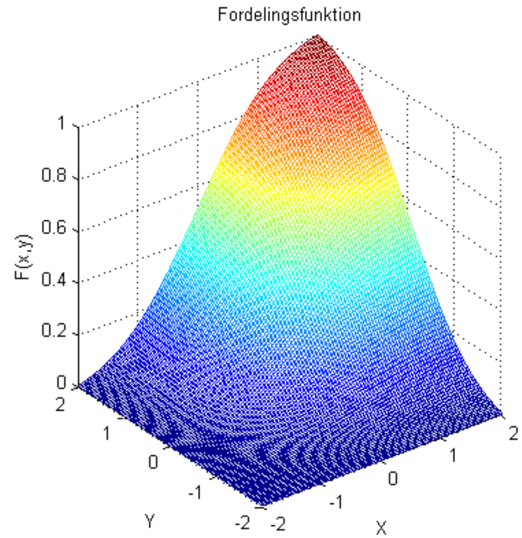


<u>ρ=0:</u>
Symmetrisk
fordeling

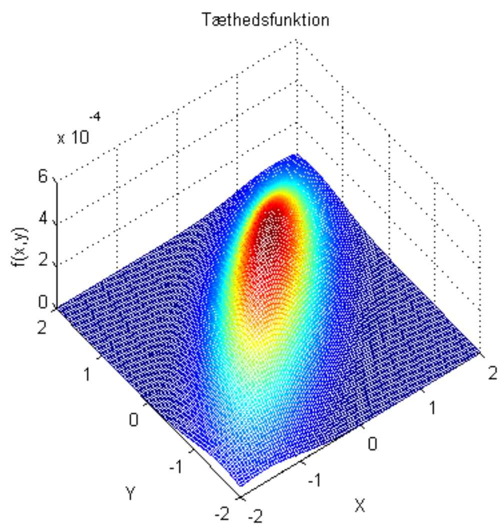
<u>ρ=0.8:</u>
Asymmetrisk
fordeling

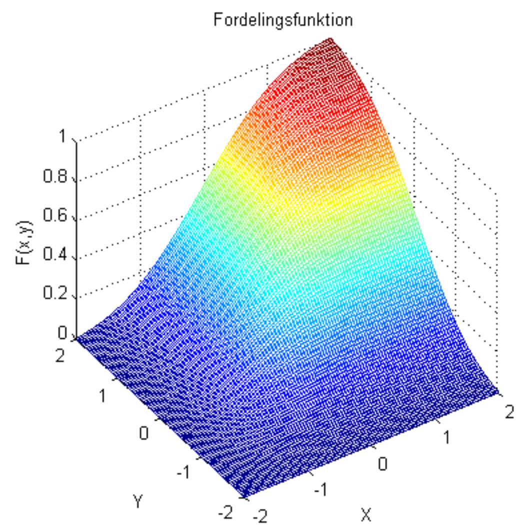
Bivariat Gaussfordeling





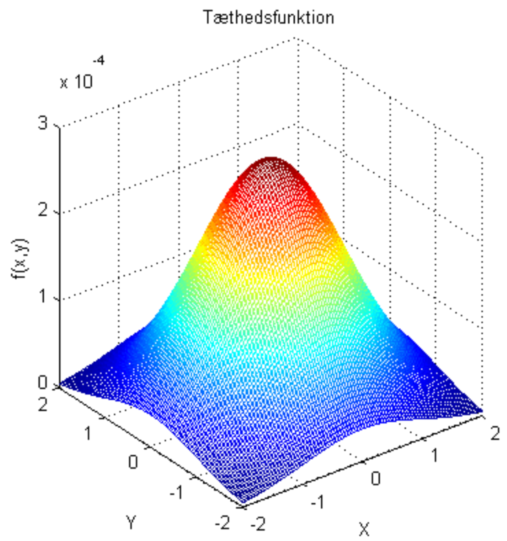
Her er X og Y uafhængige

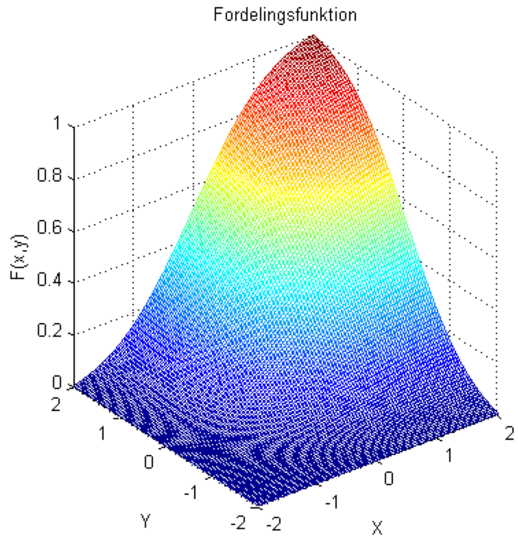




Her er X og Y <u>ikke</u> uafhængige!

Symmetrisk Gaussfordeling



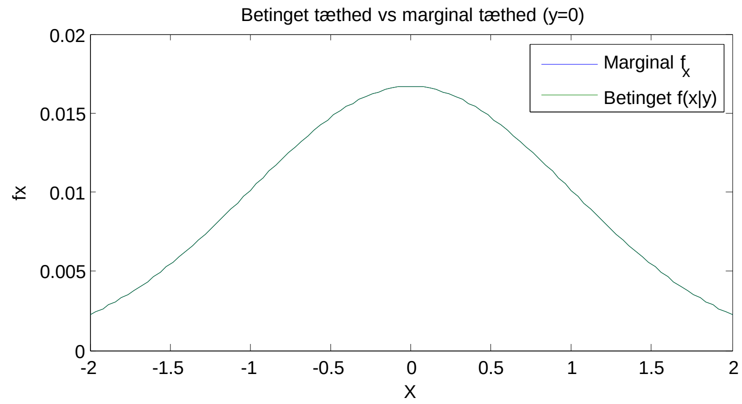


$\underline{\rho=0:}$
Symmetrisk fordeling

X og Y uafhængige

Hvad er $f_X(x)$?

Svar: En 1D gauss

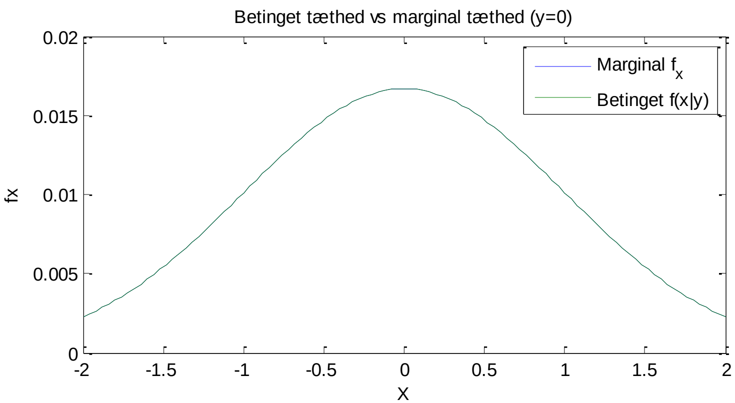



```
% Matlab
fx = sum(f,2);
fy = sum(f,1);
```

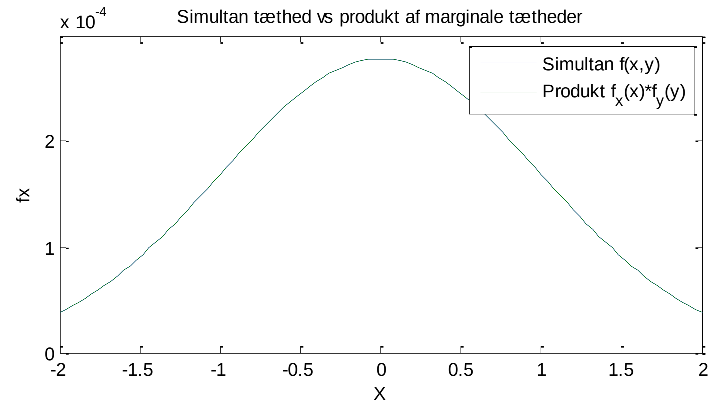
Bemærk, at når X og Y er uafhængige, så $f(x|y) = f_X(x)$ og $f(y|x) = f_Y(y)$.

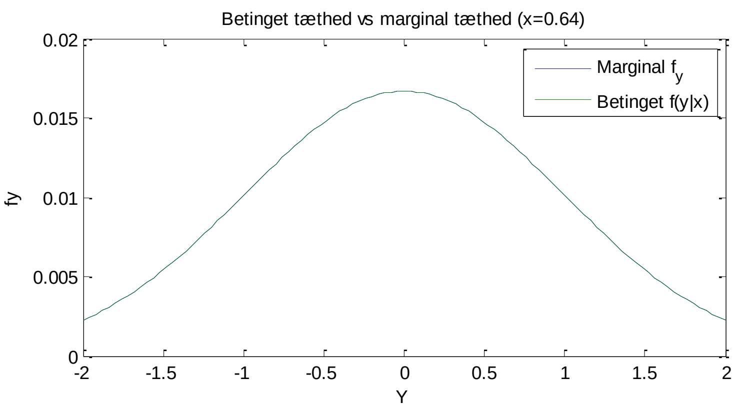
Symmetrisk Gaussfordeling

$f(x|y=0)$ og $f_X(x)$

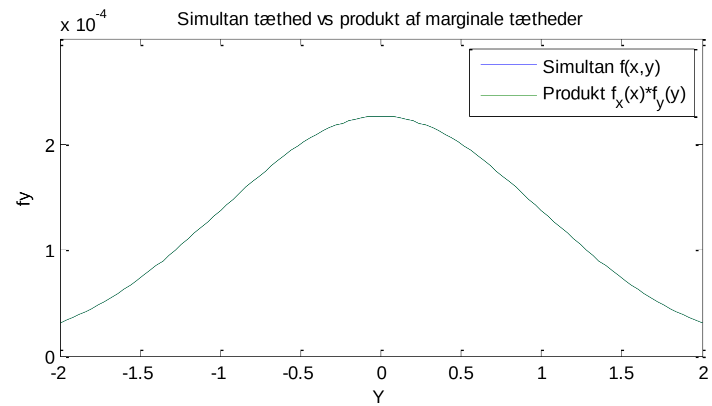


$f(x, y=0)$ og $f_X(x) \cdot f_Y\,(y=0)$



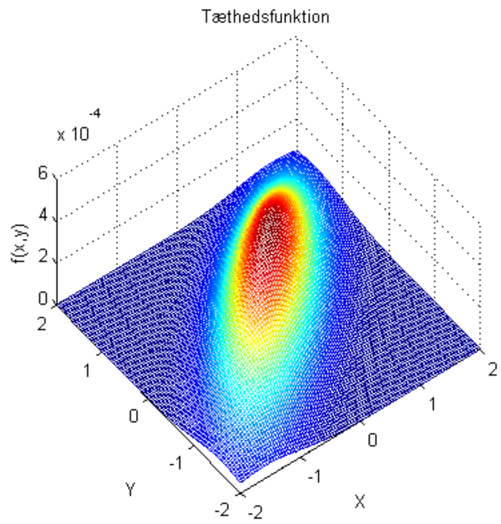


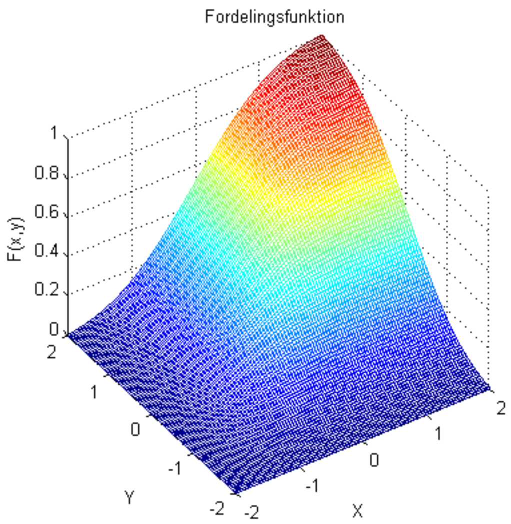
$f(y|x=0.64)$ og $f_Y(y)$



$f(x=0.64, y)$ og $f_X(x=0.64) \cdot f_Y\,(y)$

Asymmetrisk Gaussfordeling



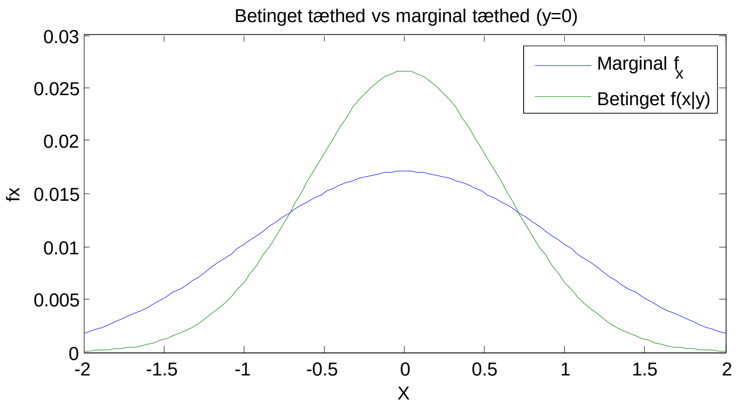


$\underline{\rho=0.8:}$
Asymmetrisk fordeling

X og *Y ikke* uafhængige

Hvad er $f_X(x)$?

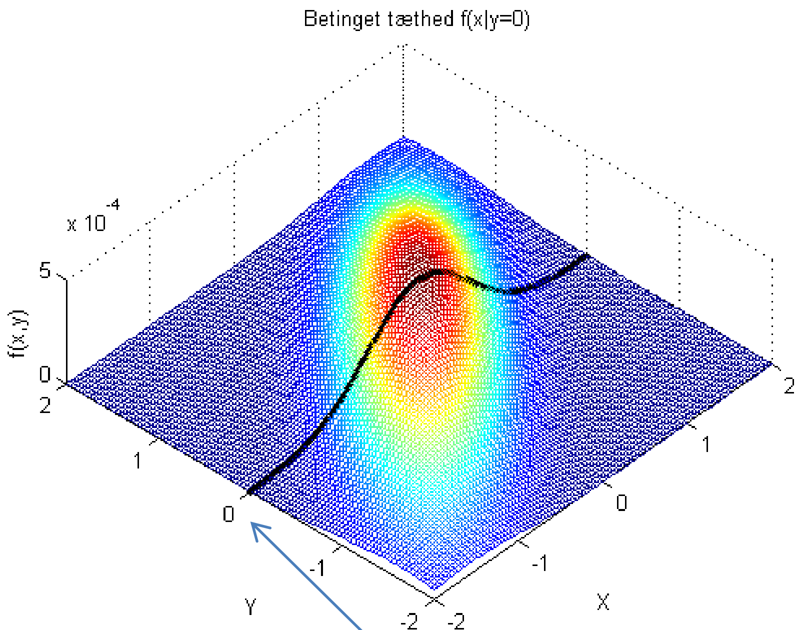
Svar: En 1D gauss

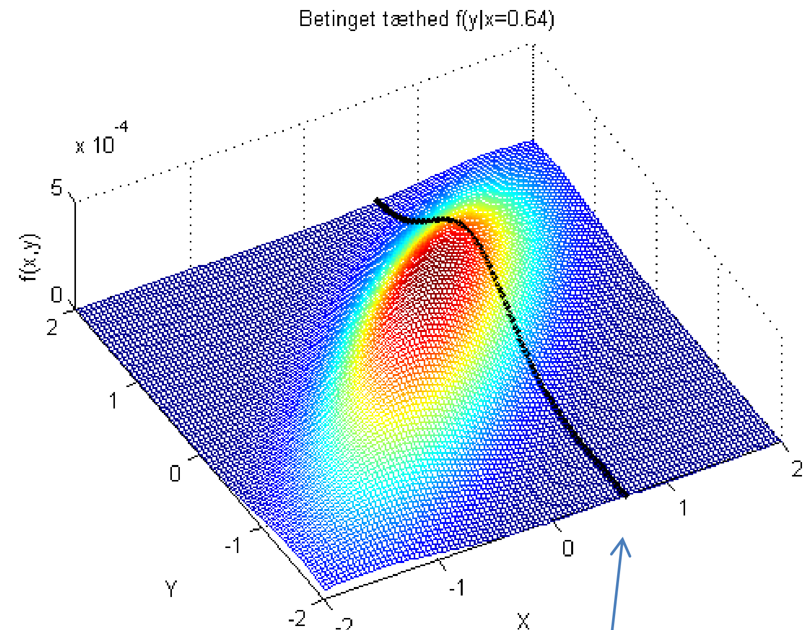



```
% Matlab
fx = sum(f,2);
fy = sum(f,1);
```

Her er $f(x|y) \neq f_X(x)$. Hvordan beregner man den betingede tæthedsfunktion, f(x|y)?

Asymmetrisk Gaussfordeling





Hvad er arealet under den sorte kurve?

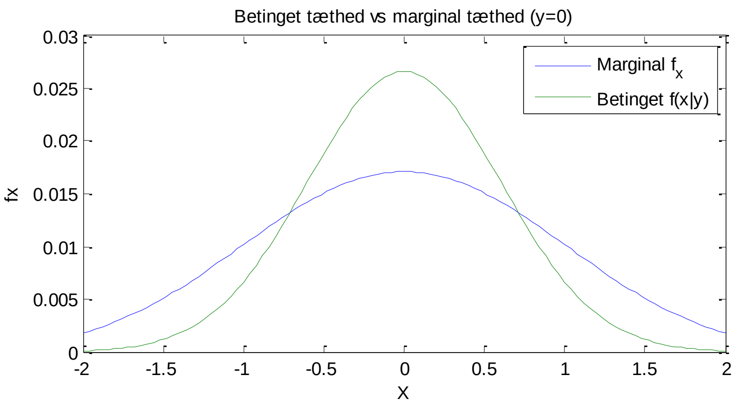
$$f(x|y=0) = \frac{f(x, y=0)}{f_Y(y=0)}$$

$$f(y|x=0.64) = \frac{f(x=0.64, y)}{f_X(x=0.64)}$$

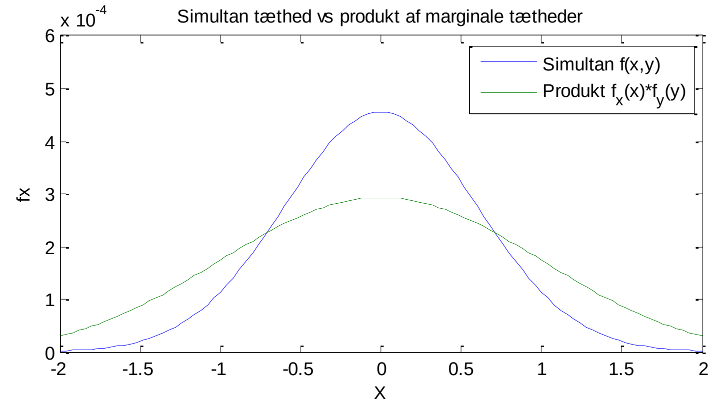
Normaliseringskonstant
(arealet under den sorte kurve)

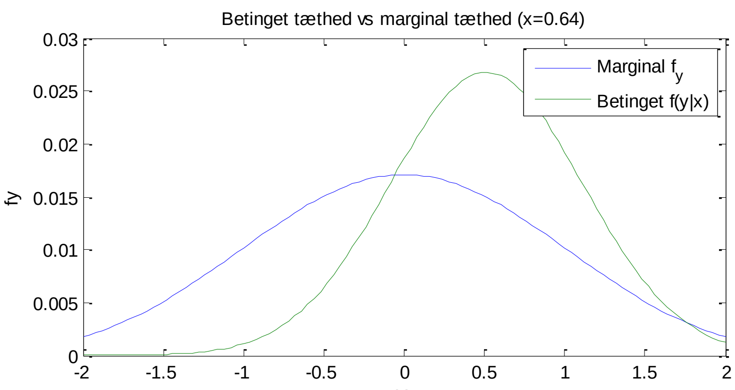
Asymmetrisk Gaussfordeling

$f(x|y=0)$ og $f_X(x)$

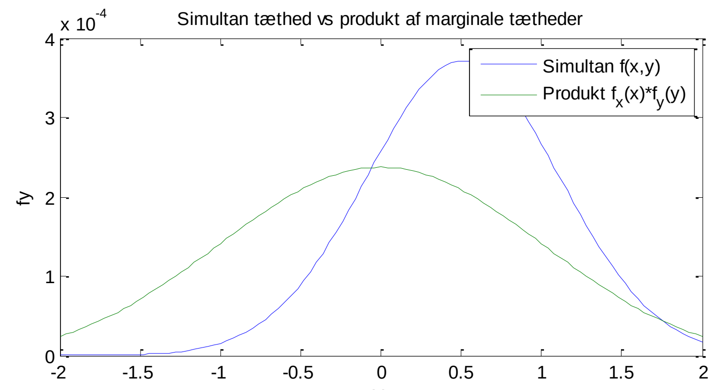


$f(x, y=0)$ og $f_X(x) \cdot f_Y\ (y=0)$





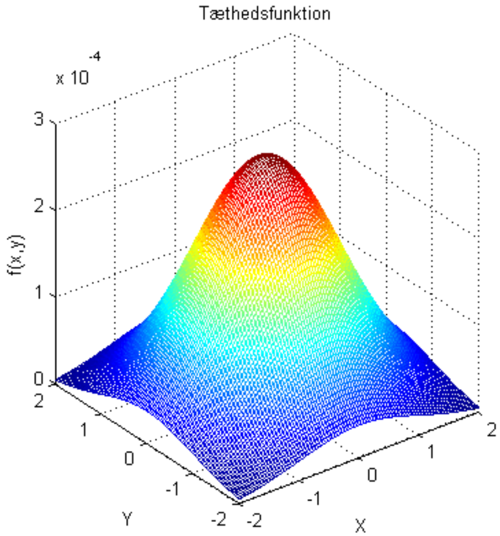
$f(y|x=0.64)$ og $f_Y(y)$

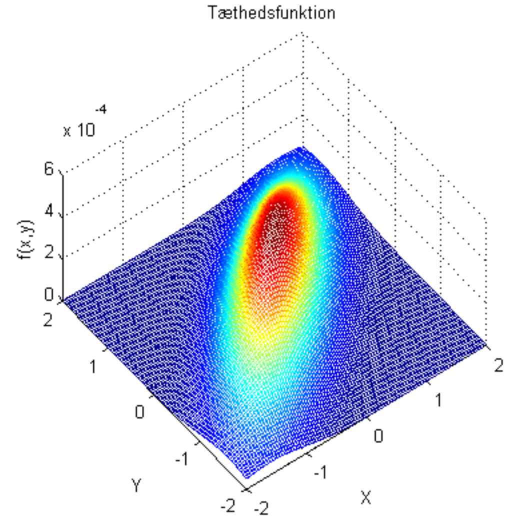


$f(x=0.64, y)$ og $f_X(x=0.64) \cdot f_Y\ (y)$

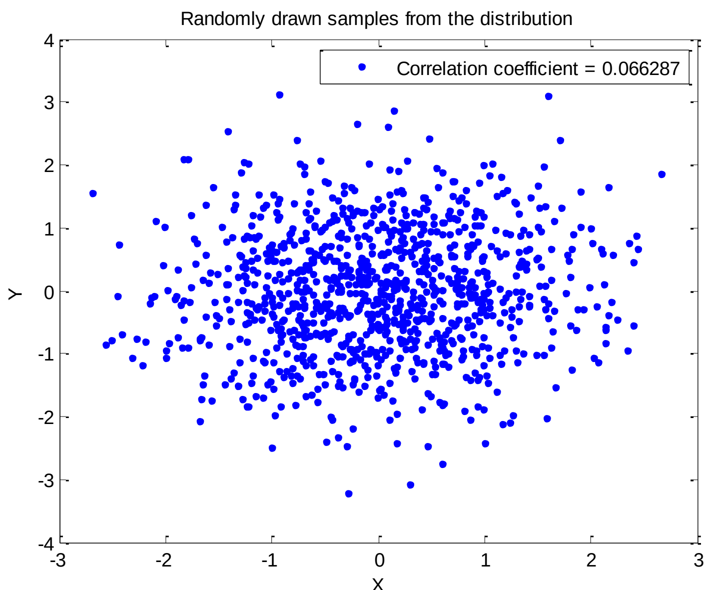
Sampling af data fra fordeling

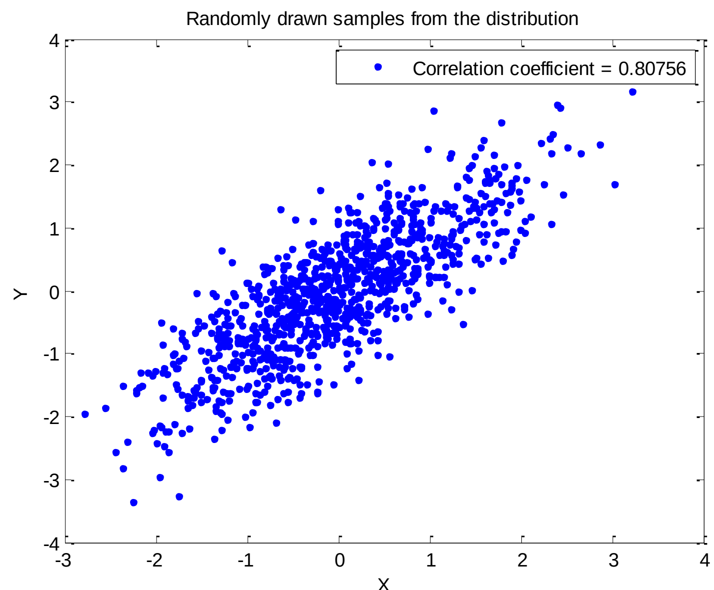
$\underline{\rho=0:}$
Symmetrisk fordeling





$\underline{\rho=0.8:}$
Asymmetrisk fordeling





Hvor er robotten?

- Y = X + N

    Y = målt position

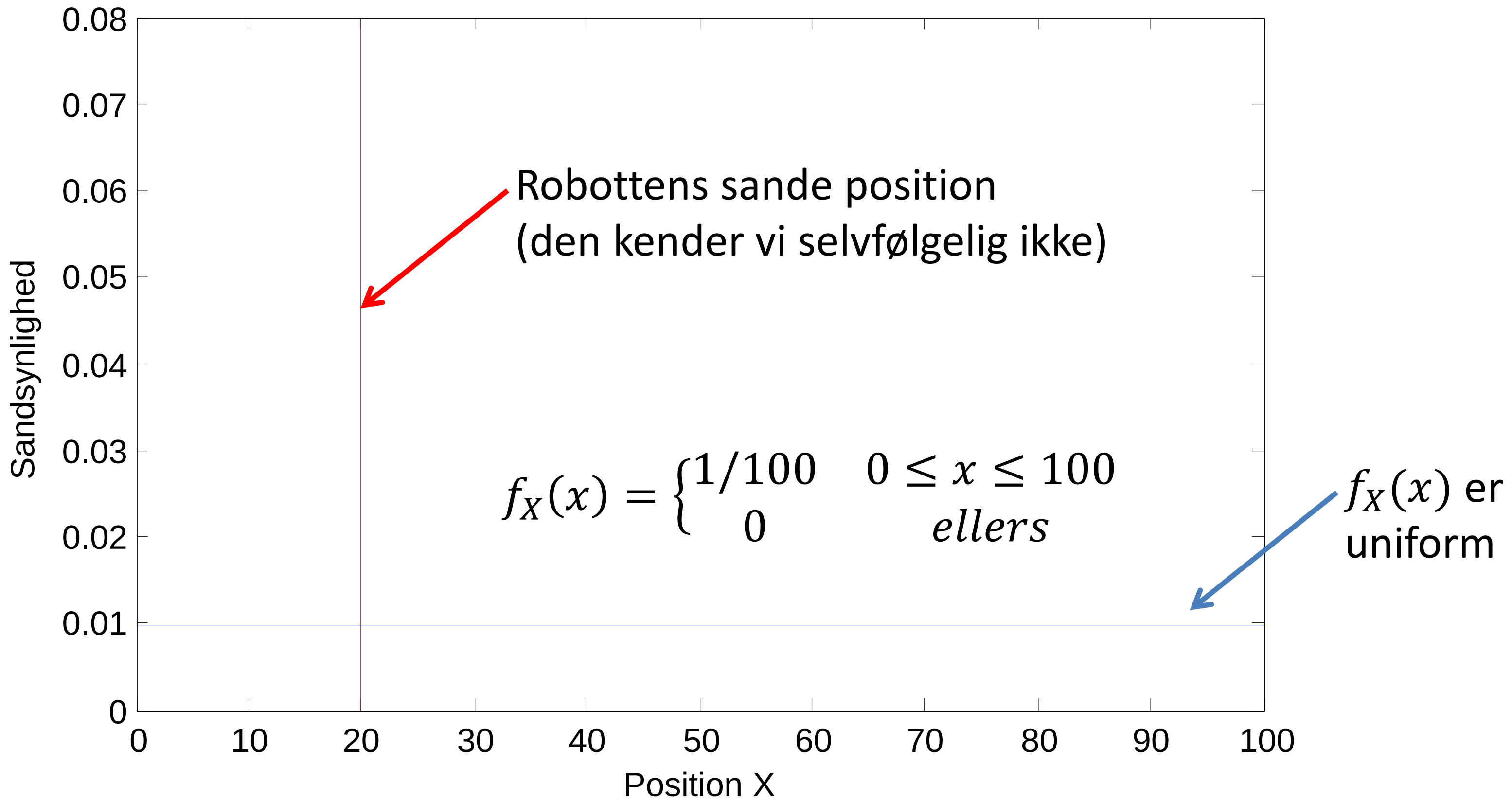
    X = sand position

    N = målestøj

- Givet målingen Y samt viden om fordelingen af hhv. X og N, hvad kan vi sige om robottens position?

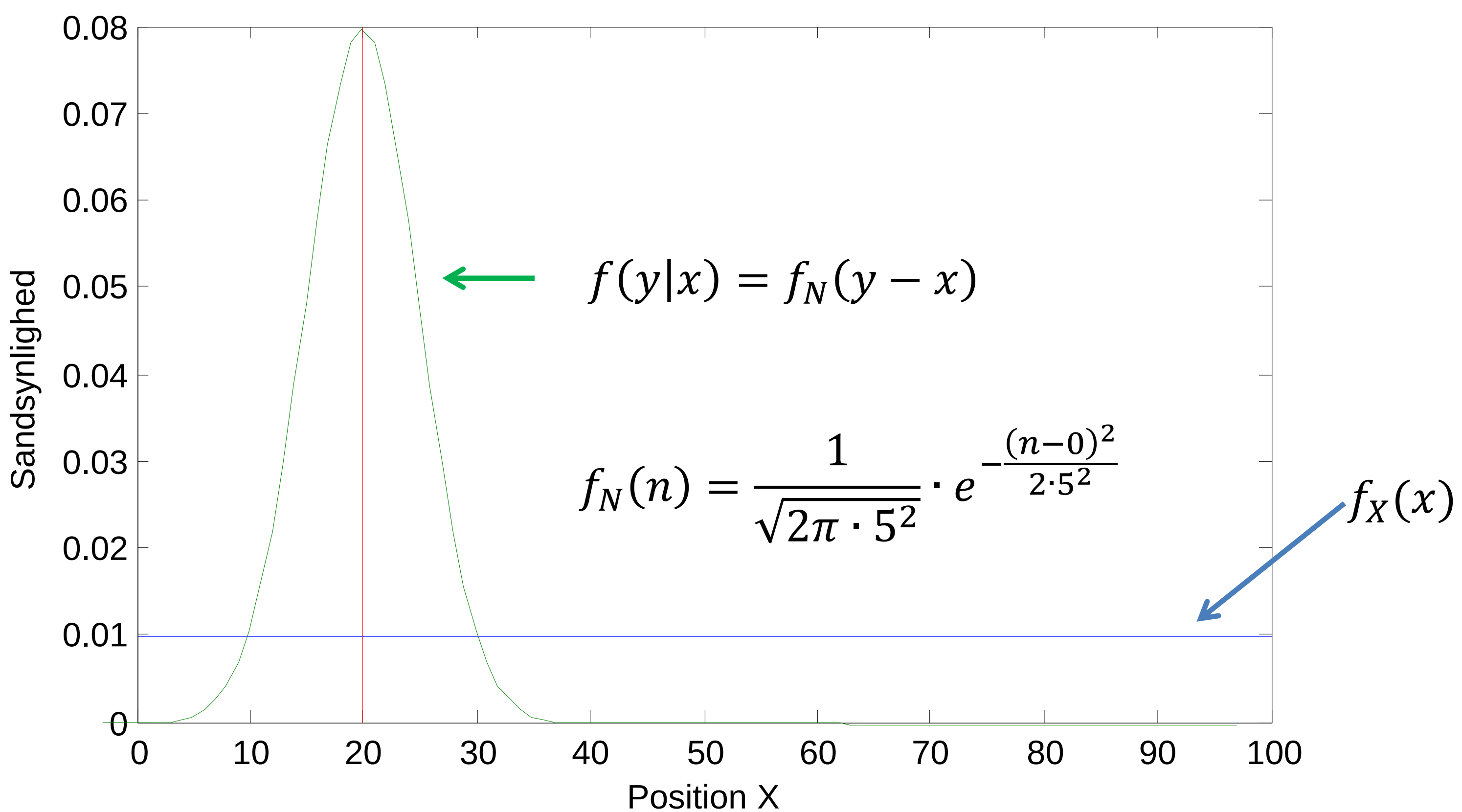
Hvor er robotten?

- Til at begynde med, aner vi ikke, hvor robotten er.
- Dvs., den marginale tæthed $f_X(x)$ for positionen x er uniform.
- Dette er vores a priori ("før måling") viden om positionen.



Hvor er robotten?

- Vi vil nu foretage en måling (y) af robottens position.
- Lad as antage, at målestøjen er normalfordelt: $N \sim \mathcal{N}(0{,}5)$
- Så må $f(y|x) = f(x+n|x) = f_N(n) = f_N(y-x)$



Hvor er robotten?

- Y = X + N

  Y = målt position

  X = sand position

  N = målestøj

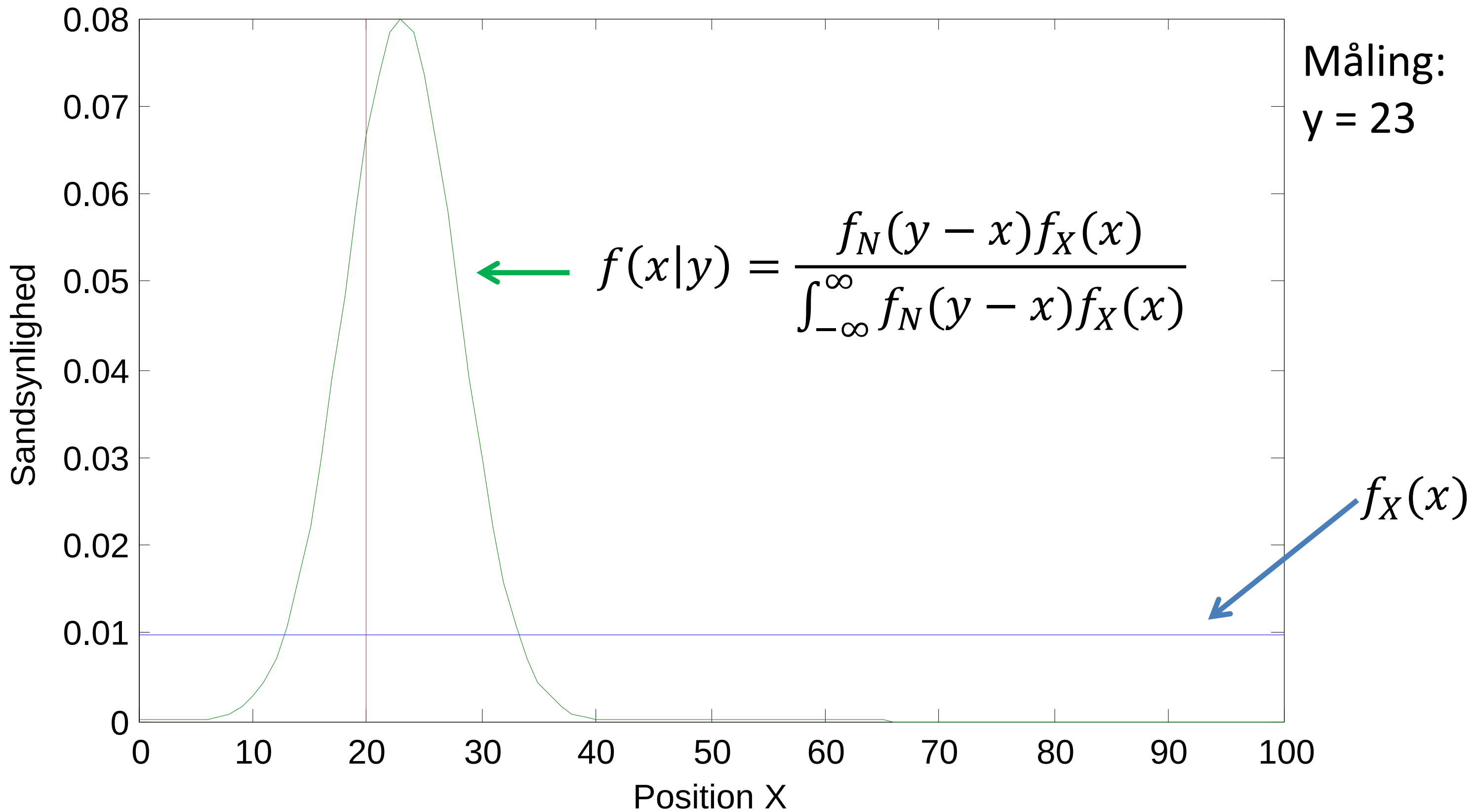
- Givet målingen Y samt viden om fordelingen af hhv. X og N, hvad kan vi sige om robottens position?
- Bayes regel:

$$f(x|y) = \frac{f(x,y)}{f_Y(y)} = \frac{f(y|x)f_X(x)}{\int_{-\infty}^{\infty} f(y|x)f_X(x)dx}$$

$$= \frac{f_N(y-x)f_X(x)}{\int_{-\infty}^{\infty} f_N(y-x)f_X(x)dx}$$

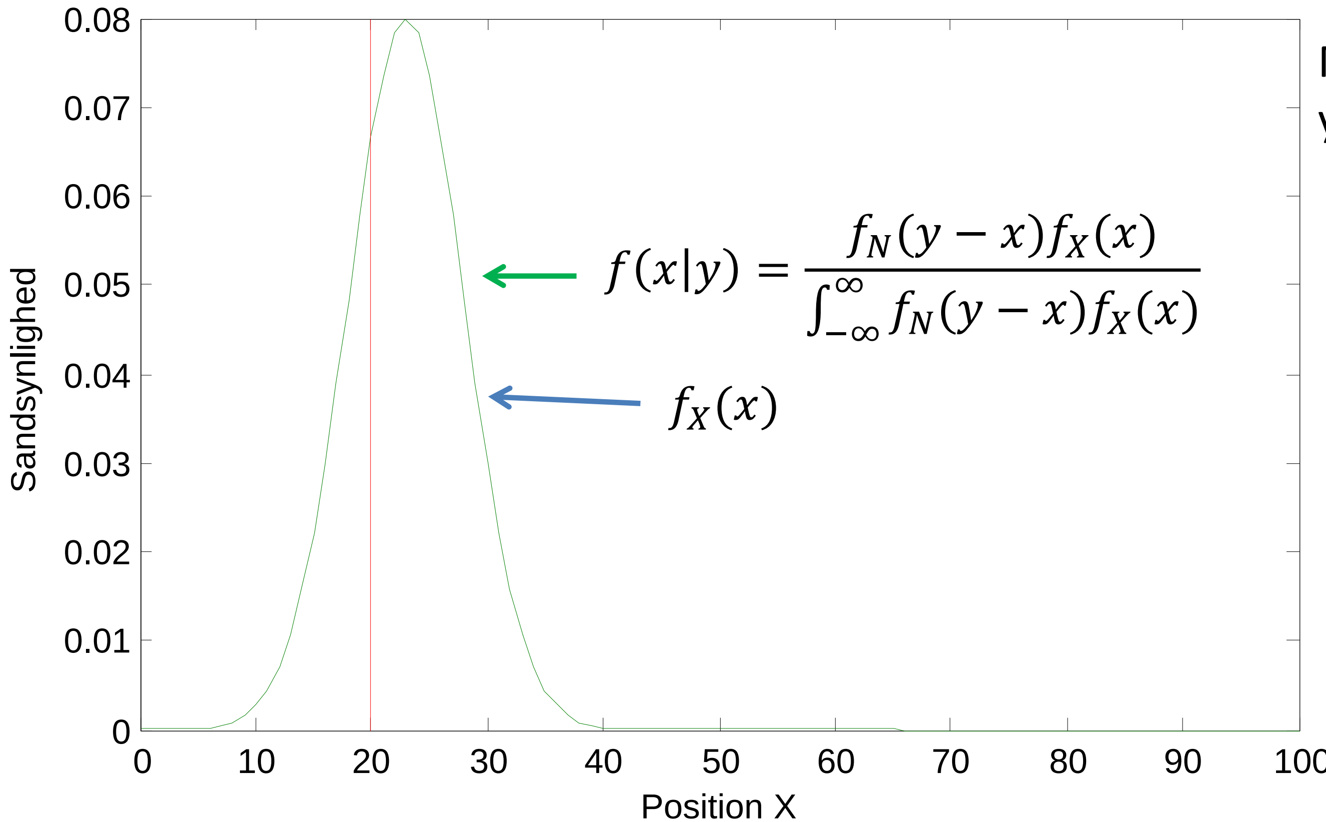
Hvor er robotten?

- Idet vi foretager en måling (y) af positionen, kan vi bruge Bayes formel til at opdatere vores viden om positionen.
- Den opdaterede tæthed $f(x|y)$ repræsenterer vores a posteriori ("efter måling") viden om positionen.



Hvor er robotten?

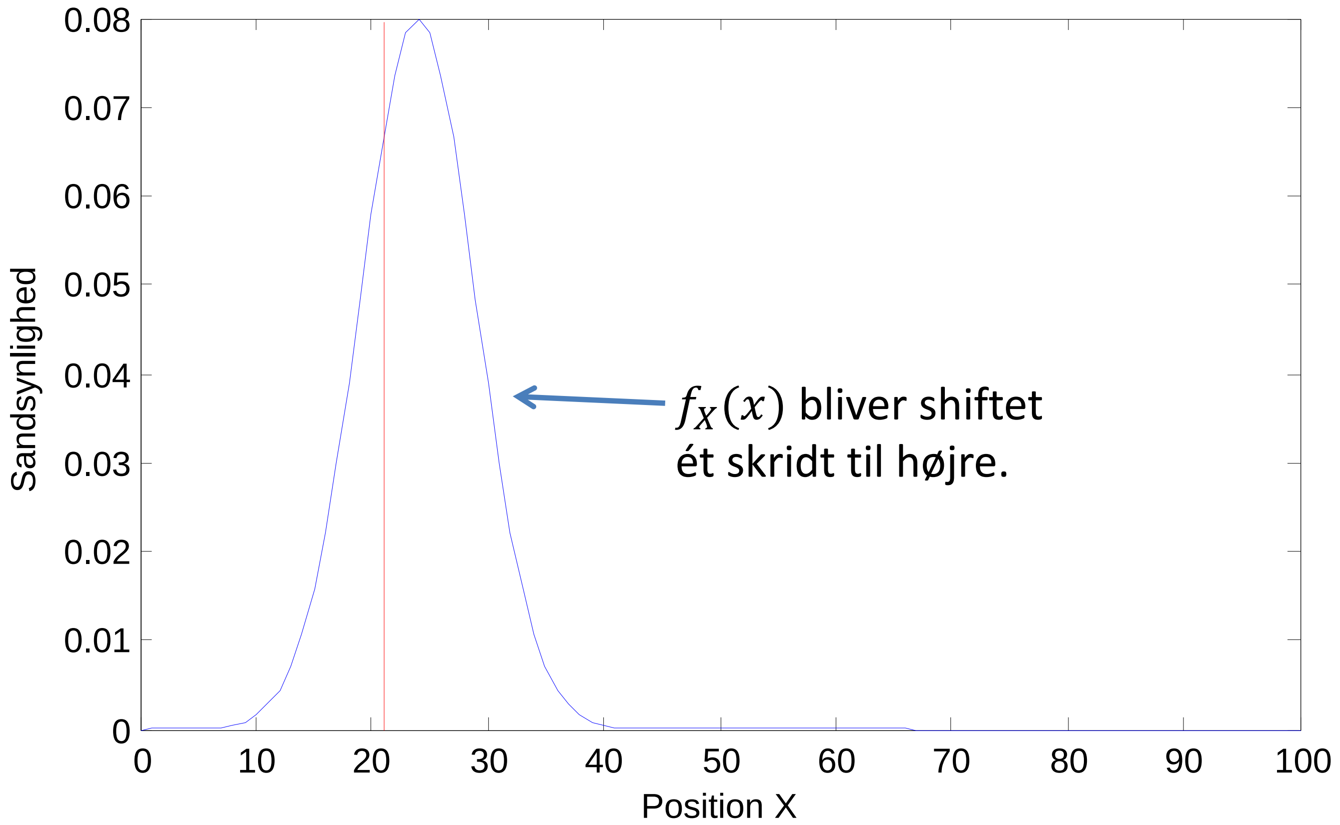
- Vi kan nu opdatere vores a priori viden om positionen, x.
- Konkret lader vi $f(x|y)$ være vores nye marginale tæthed for X, $f_X(x)$.
- Hvad sker der med $f_X(x)$, når vi flytter robotten et skridt til højre?



Måling:
y = 23

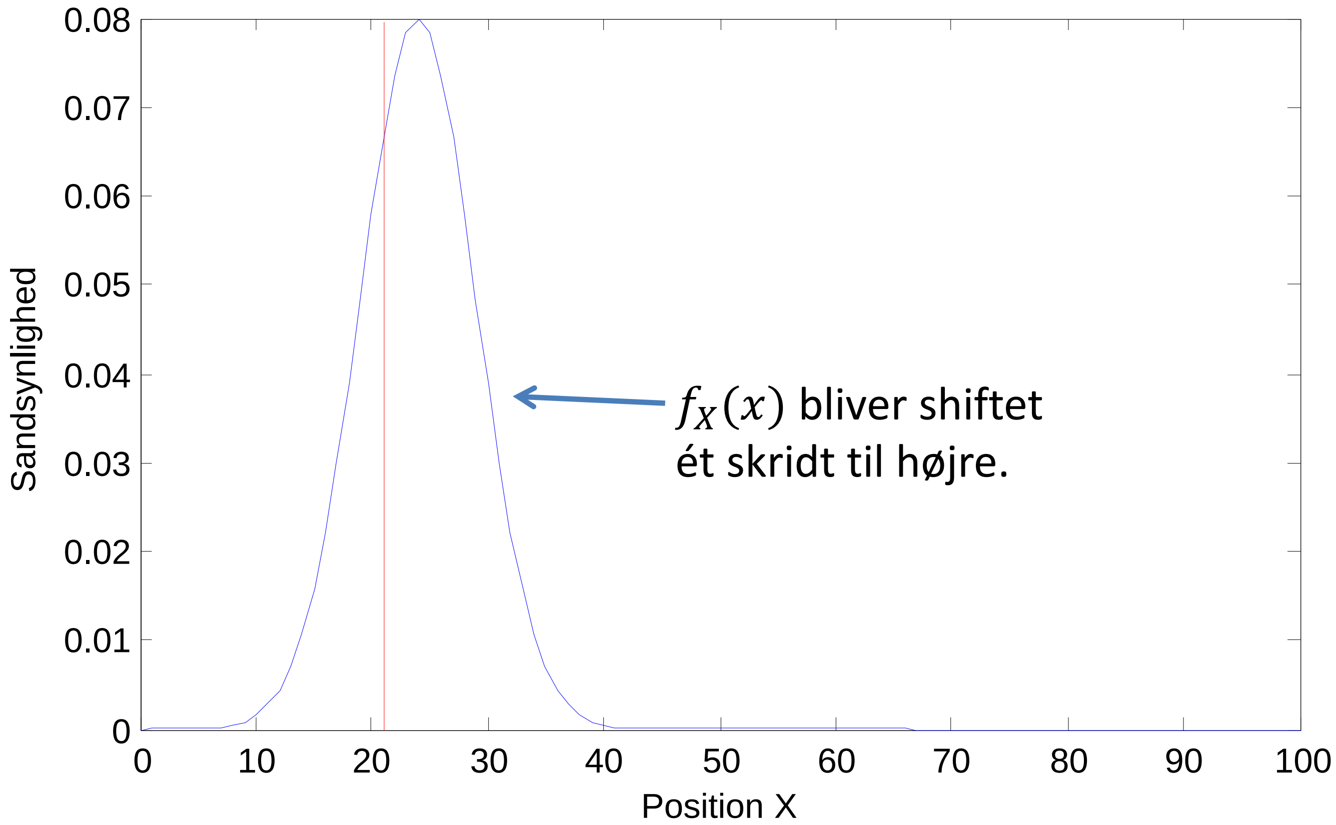
Hvor er robotten?

- Vi kan nu opdatere vores a priori viden om positionen, x.
- Konkret lader vi $f(x|y)$ være vores nye marginale tæthed for X, $f_X(x)$.
- Hvad sker der med $f_X(x)$, når vi flytter robotten et skridt til højre?



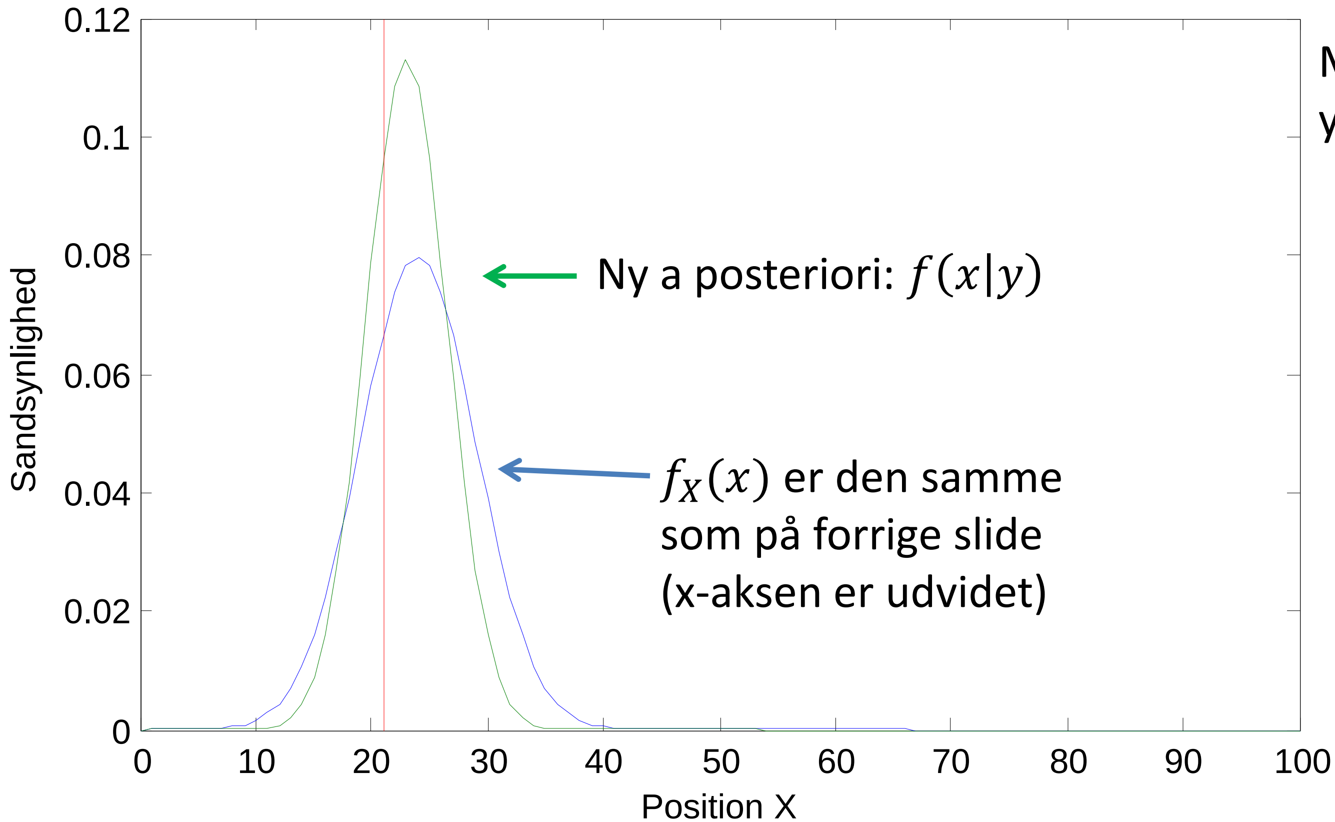
Hvor er robotten?

- Gentag proceduren med den opdaterede a priori tæthed, $f_X(x)$
  - Foretag en ny måling (y)
  - Beregn den nye a posteriori tæthed for positionen ved Bayes regel
  - Opdatér a priori tætheden
  - Flyt robotten et skridt til højre



Hvor er robotten?

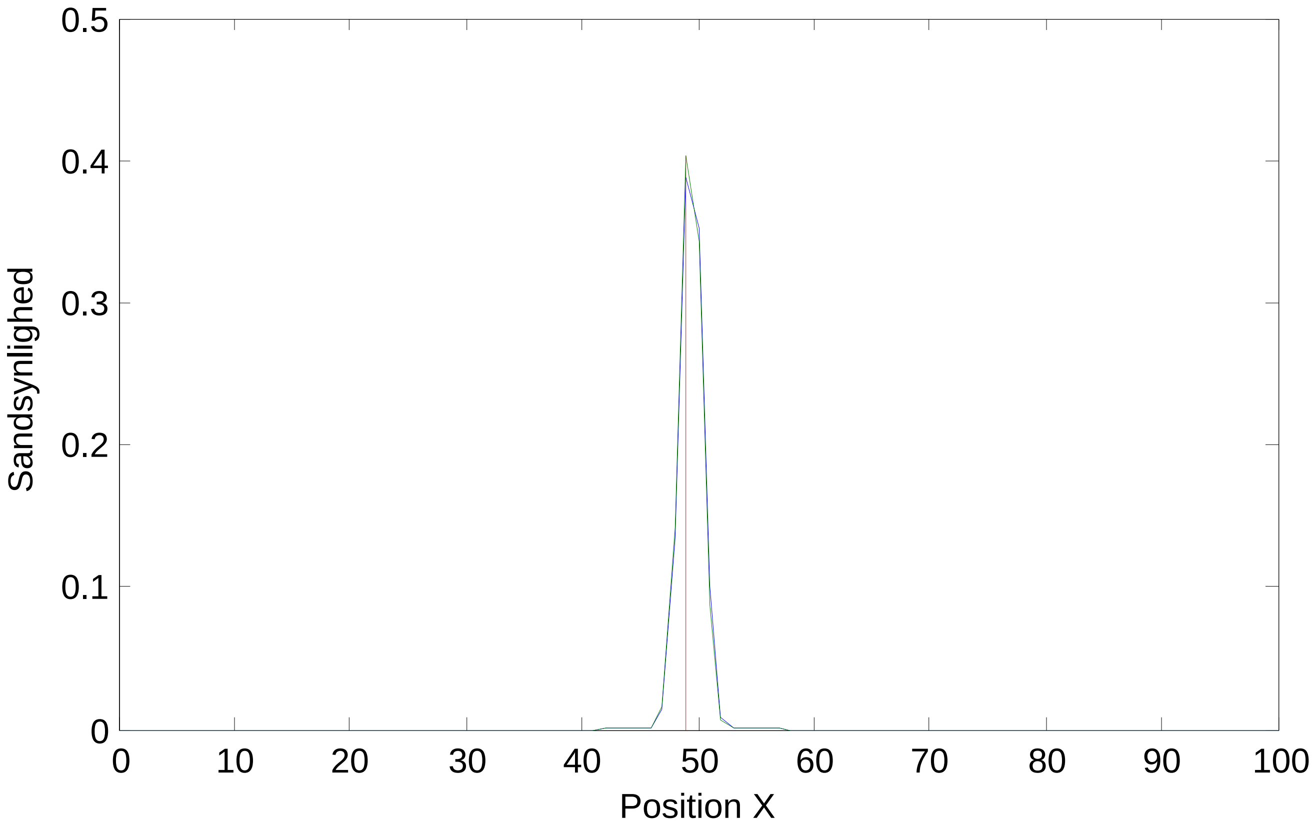
- Gentag proceduren med den opdaterede a priori tæthed, $f_X(x)$
  - Foretag en ny måling (y)
  - Beregn den nye a posteriori tæthed for positionen ved Bayes regel
  - Opdatér a priori tætheden
  - Flyt robotten et skridt til højre



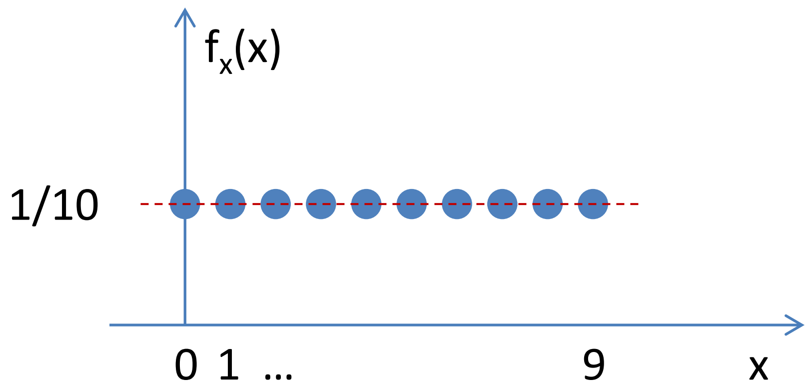
Måling:
y = 22

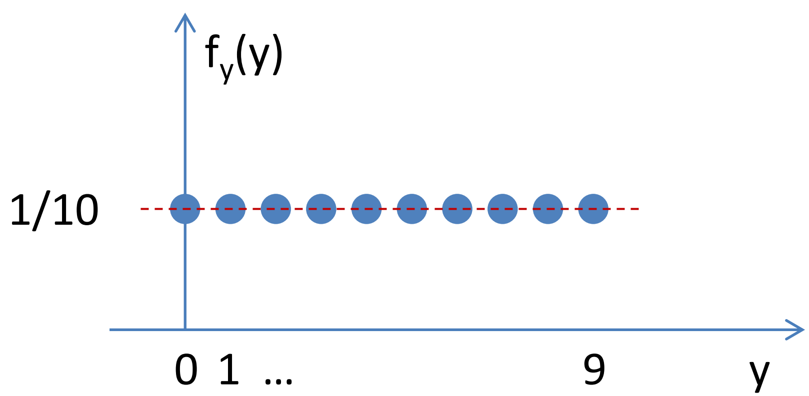
Hvor er robotten?

- Efter blot nogle få iterationer, bliver vores viden om positionen (x) langt mere sikkert.
  - Spredningen på tæthedsfunktionen bliver mindre og tæthedsfunktionen bliver højere (sammenlign tallene på x-aksen med forrige slide)



Sum af to stokastiske variable



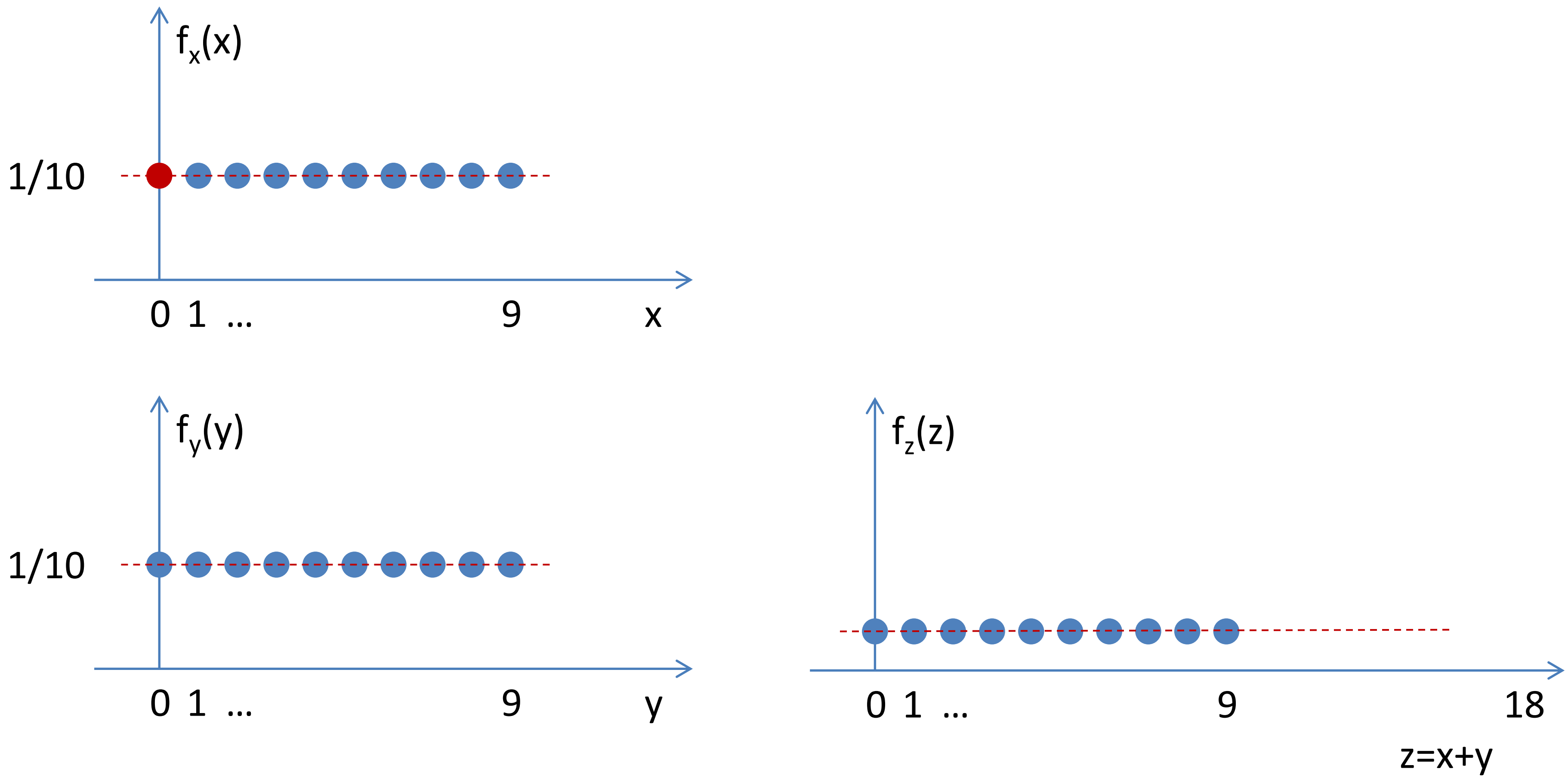


Hvilken fordeling har z = x+y?

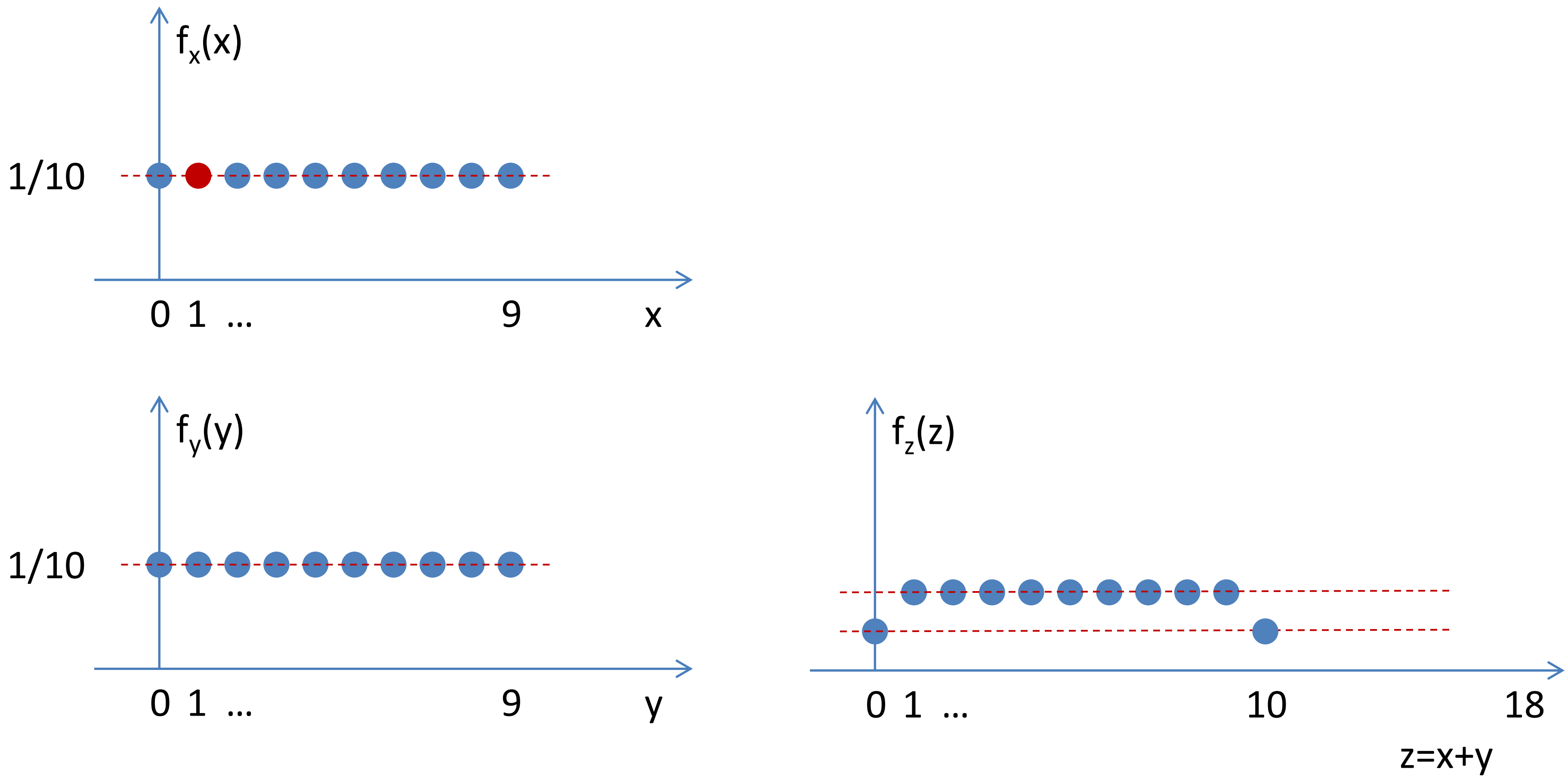
Tabel over mulige hændelser:

x	y	z = x+y
0	0	0
0	1	1
.	.	
.	.	
.	.	
9	9	18

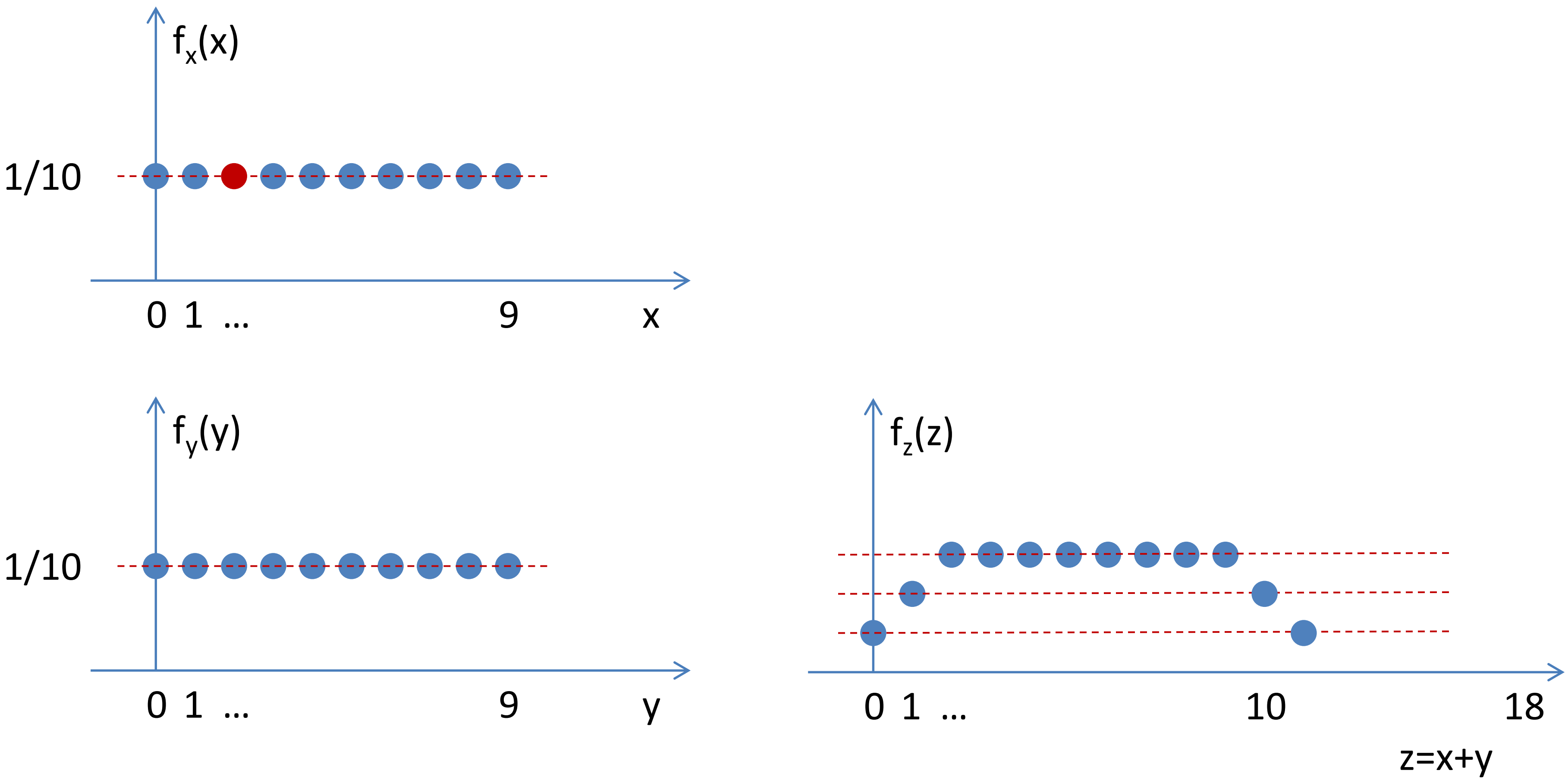
Sum af to stokastiske variable



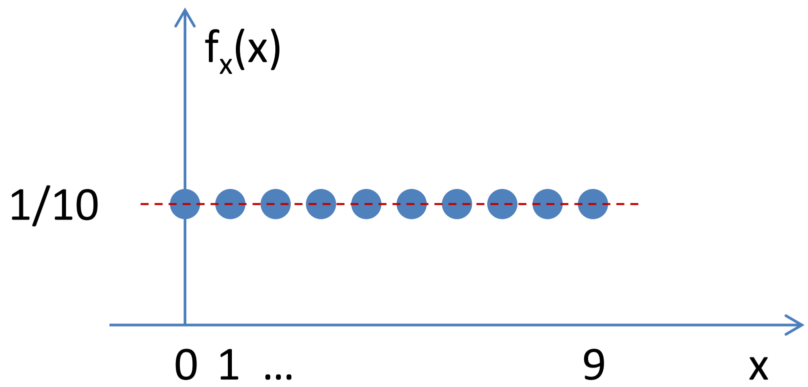
Sum af to stokastiske variable

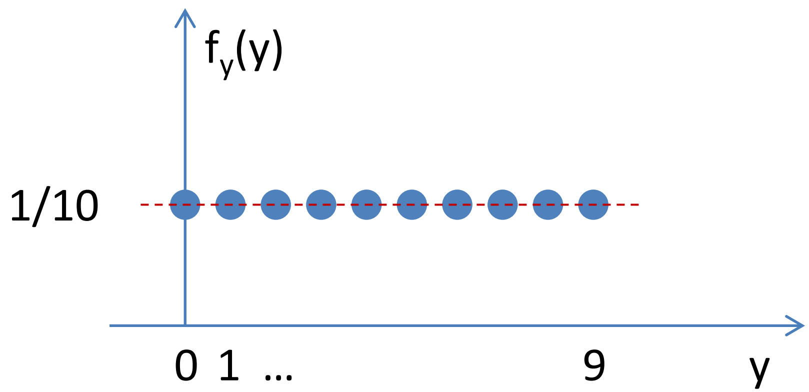


Sum af to stokastiske variable



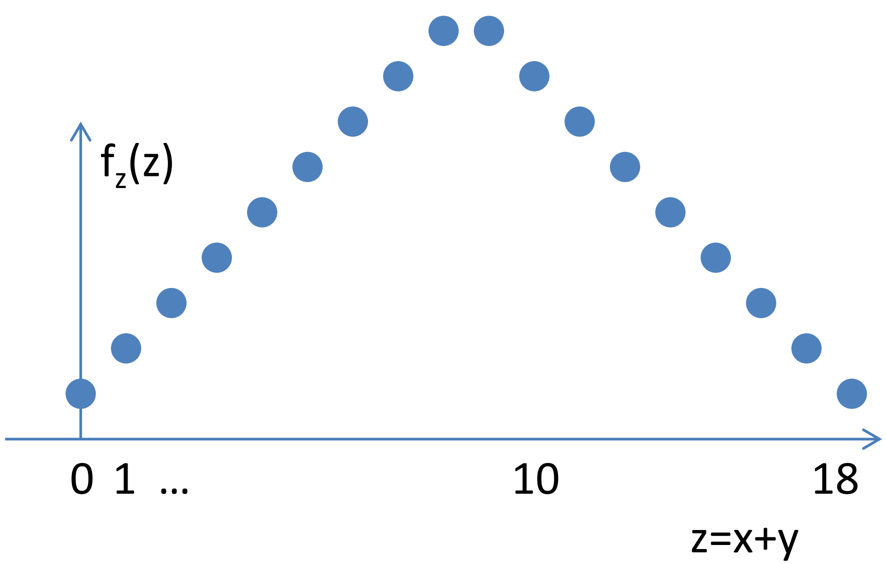
Sum af to stokastiske variable





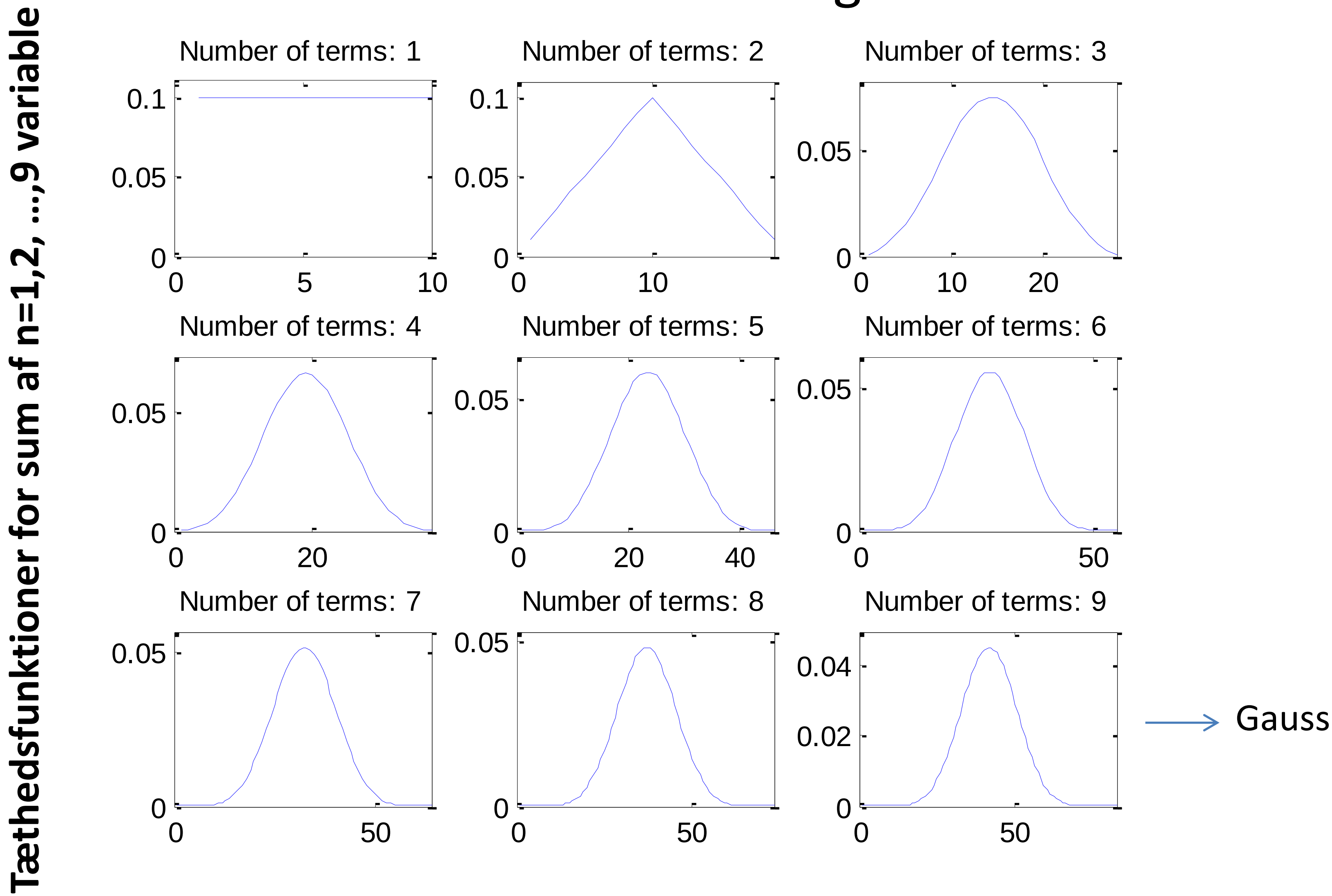
Hvilken fordeling har z = x+y?

f_z er foldningen af f_x og f_y.



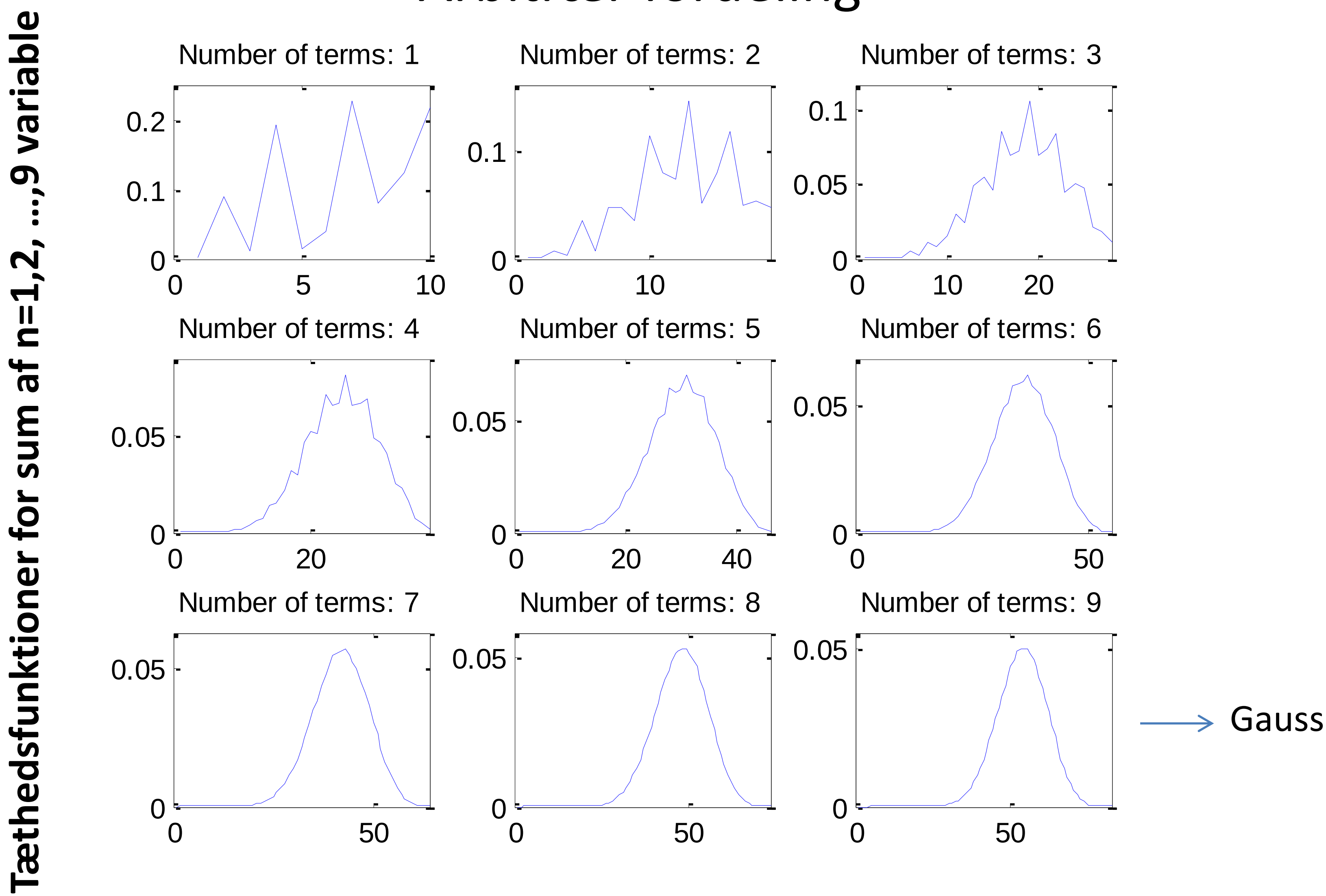
Sum af flere stokastiske variable

Uniform fordeling



Sum af flere stokastiske variable

Arbitrær fordeling



Den centrale grænseværdisætning (Central Limit Theorem)

- Her en én af flere varianter

The Central Limit Theorem

If a random sample of size n is drawn from a population with mean μ and variance σ^2, then the sample mean $\overline{X}$ has approximately a normal distribution with mean μ and variance σ^2/n. That is, the distribution function of

$$\frac{(\overline{X} - \mu)}{\sigma/\sqrt{n}}$$

is approximately a standard normal. The approximation improves as the sample size increases.

- Vi vender tilbage til CLT senere i kurset.

Regneeksempel (s. 137-138)

- Se også youtube video om convolution (link i kalenderen).

Hvor er robotten?

- $X_i = X_{i-1} + dx + N_i$

  $X_0 = 0$ = initiel position.

  $dx = 1$ = kommando til robotten (ryk "dx" skridt til højre).

  N_i = usikkerhed da robotten aldrig rykker sig præcis med dx.

- Givet målingen X_0, dx samt viden om fordelingen af N_i, hvad kan vi sige om robottens position i den i'te iteration?

$$f_i(x) = f_{i-1}(x - dx) * f_N(n)$$

$$f_0(x) = \begin{cases} 1 & for\ x = 0 \\ 0 & ellers \end{cases}$$

$$f_N(n) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{(n-\mu)^2}{2\sigma^2}}$$

Hvor er robotten?

- $X_i = X_{i-1} + dx + N_i$

  $X_0 = 0$ = initiel position.

  $dx = 1$ = kommando til robotten (ryk ”dx” skridt til højre).

  N_i = usikkerhed da robotten aldrig rykker sig præcis med dx.

- Givet målingen X_0, dx samt viden om fordelingen af N_i, hvad kan vi sige om robottens position i den i’te iteration?

$$f_i(x) = f_{i-1}(x - dx) * f_N(n)$$

$$f_0(x) = \begin{cases} 1 & for\ x = 0 \\ 0 & ellers \end{cases}$$

$$f_N(n) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{(n-\mu)^2}{2\sigma^2}}$$