Course 02402/02323 Introducerende Statistik

Forelæsning 3: Kontinuerte fordelinger

Per Bruun Brockhoff

DTU Compute, Statistik og Dataanalyse
Bygning 324, Rum 220
Danmarks Tekniske Universitet
2800 Lyngby – Danmark

e-mail: perbb@dtu.dk

DTU Compute
Institut for Matematik og Computer Science

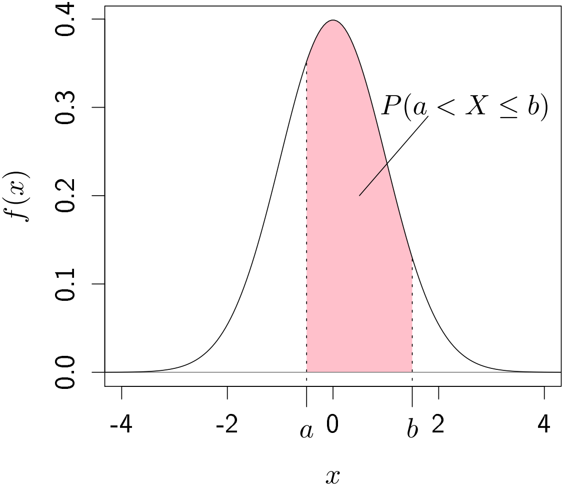
Oversigt

Oversigt

Tæthedsfunktion (probability density function (pdf))

- Tæthedsfunktionen for en stokastisk variabel betegnes ved $f(x)$
- $f(x)$ siger noget om hyppigheden af udfaldet x for den stokastiske variabel X
- For kontinuerte variable svarer tætheden ikke til sandsynligheden, dvs. $f(x) \neq P(X = x)$
- Et godt plot af $f(x)$ er et histogram (kontinuert)

Tæthedsfunktion for en kontinuert variabel



DTU Compute
Institut for Matematik og Computer Science

Tæthedsfunktion for en kontinuert variabel

For en kontinuert stokastisk variabel skrives tæthedsfunktionen som:

$$f(x)$$

Der gælder:

$$f(x) \geq 0 \quad \text{for alle mulige } x$$

$$\int_{-\infty}^{\infty} f(x)dx = 1$$

DTU Compute
Institut for Matematik og Computer Science

Fordelingsfunktion (distribution function eller cumulative density function (cdf))

- Fordelingsfunktion for en kontinuert stokastisk variabel betegnes ved $F(x)$.
  Fordelingsfunktionen svarer til den kumulerede tæthedsfunktion:

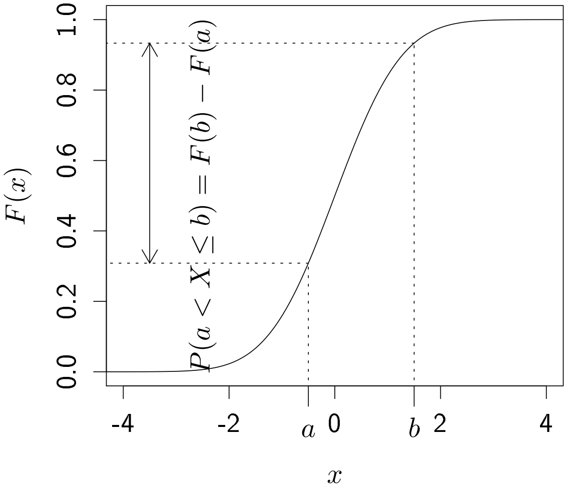
$$F(x) = P(X \leq x)$$

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

$$f(x) = F^{'}(x)$$

- Et godt plot for fordelingsfunktionen er den kumulative fordeling

Fordelingsfunktion (distribution function eller cumulative density function (cdf))

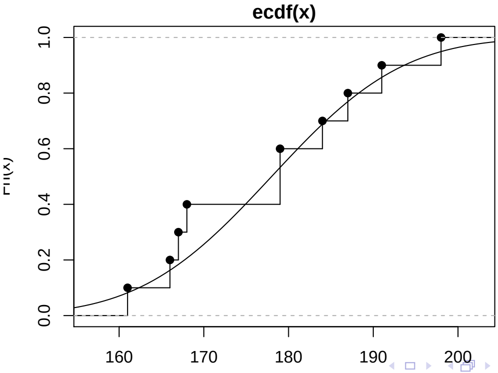


DTU Compute
Institut for Matematik og Computer Science

Den empiriske cumulative distribution function - ecdf

Student height example from Chapter 1:

```
x <- c(168,161,167,179,184,166,198,187,191,179)
plot(ecdf(x), verticals = TRUE)
xp <- seq(0.9*min(x), 1.1*max(x), length.out = 100)
lines(xp, pnorm(xp, mean(x), sd(x)))
```



Middelværdi (mean) af en kontinuert stokastisk variabel

Middelværdien af en kontinuert stokastisk variabel

$$\mu = \int_{-\infty}^{\infty} x \cdot f(x) dx$$

DTU Compute
Institut for Matematik og Computer Science

Middelværdi (mean) af en kontinuert stokastisk variabel

Middelværdien af en kontinuert stokastisk variabel

$$\mu = \int_{-\infty}^{\infty} x \cdot f(x) dx$$

Sammenlign med den diskrete definition:

$$\mu = \sum_{i=1}^{\infty} x_i f(x_i)$$

Varians af en kontinuert stokastisk variabel

Variansen af en kontinuert stokastisk variabel:

$$\sigma^2 = \int_{-\infty}^{\infty} (x-\mu)^2 \cdot f(x)dx$$

DTU Compute
Institut for Matematik og Computer Science

Varians af en kontinuert stokastisk variabel

Variansen af en kontinuert stokastisk variabel:

$$\sigma^2 = \int_{-\infty}^{\infty} (x-\mu)^2 \cdot f(x) dx$$

Sammenlign med den diskrete definition:

$$\sigma^2 = \sum_{i=1}^{\infty} (x_i - \mu)^2 f(x_i)$$

Kovariansen af to stokastiske variable

Kovariansen af to stokastiske variable:

Let X and Y be two random variables, then the covariance between X and Y, is

$$\text{Cov}(X,Y) = E[(X - E[X])(Y - E[Y])]$$

DTU Compute
Institut for Matematik og Computer Science

Oversigt

Konkrete statistiske fordelinger

- Der findes en række statistiske fordelinger, som kan bruges til at beskrive og analysere forskellige problemstillinger med

Vi betragter nu kontinuerte fordelinger

- Uniform fordelingen
- Normal fordelingen
- Log-Normal fordelingen
- Eksponential fordelingen

Uniform fordelingen

Skrivemåde:

$$X \sim U(\alpha, \beta)$$

Uniform fordelingen

Skrivemåde:

$X \sim U(\alpha, \beta)$

Tæthedsfunktion:

$f(x) = \frac{1}{\beta - \alpha}$

DTU Compute
Institut for Matematik og Computer Science

Uniform fordelingen

Skrivemåde:

$X \sim U(\alpha, \beta)$

Tæthedsfunktion:

$f(x) = \frac{1}{\beta - \alpha}$

Middelværdi:

$\mu = \frac{\alpha + \beta}{2}$

Uniform fordelingen

Skrivemåde:

$X \sim U(\alpha, \beta)$

Tæthedsfunktion:

$f(x) = \frac{1}{\beta-\alpha}$

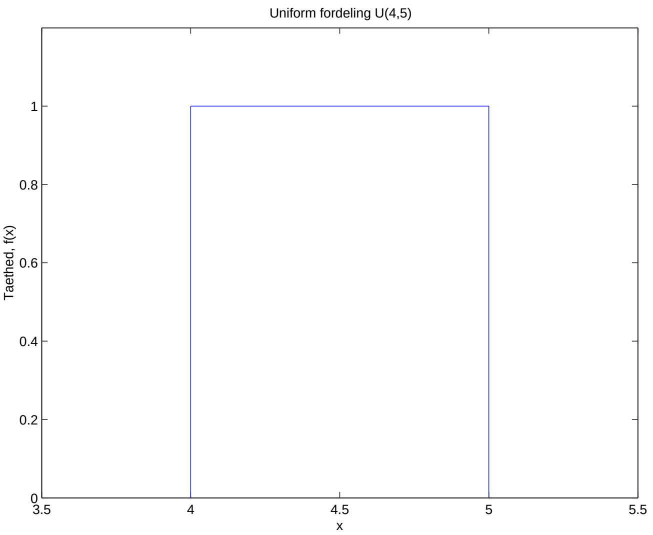
Middelværdi:

$\mu = \frac{\alpha+\beta}{2}$

Varians:

$\sigma^2 = \frac{1}{12}(\beta - \alpha)^2$

Uniform fordelingen



Eksempel 1

Medarbejdere på en arbejdsplads ankommer mellem klokken 8.00 og 8.30. Det antages, at ankomsttiden kan beskrives ved en uniform fordeling.

Spørgsmål:

Hvad er sandsynligheden for at en tilfældig udvalgt medarbejder (Hans) ankommer mellem 8.20 og 8.30?

Eksempel 1

Medarbejdere på en arbejdsplads ankommer mellem klokken 8.00 og 8.30. Det antages, at ankomsttiden kan beskrives ved en uniform fordeling.

Spørgsmål:

Hvad er sandsynligheden for at en tilfældig udvalgt medarbejder (Hans) ankommer mellem 8.20 og 8.30?

Svar:

10/30=1/3

```
punif(30,0,30)-punif(20,0,30)
```

```
[1] 0.33333
```

Eksempel 1 - forts.

Spørgsmål:

Hvad er sandsynligheden for at en tilfældig udvalgt medarbejder (Martin) ankommer efter 8.30?

Eksempel 1 - forts.

Spørgsmål:

Hvad er sandsynligheden for at en tilfældig udvalgt medarbejder (Martin) ankommer efter 8.30?

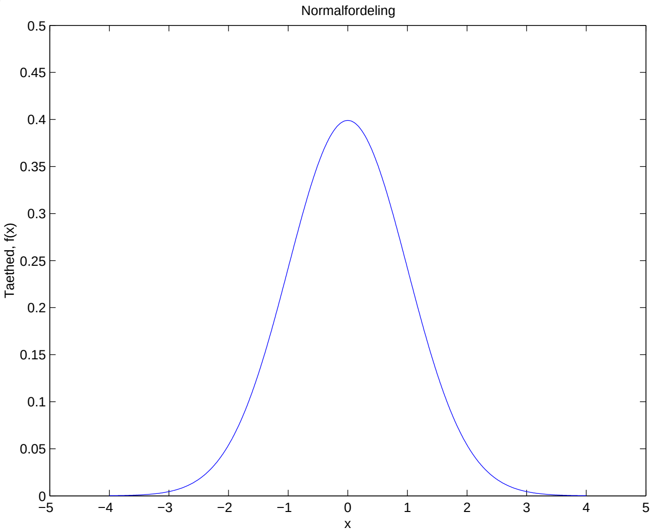
Svar:

0

```
1-punif(30,0,30)
```

```
[1] 0
```

Normalfordelingen



Normal fordelingen

Skrivemåde:

$$X \sim N(\mu, \sigma^2)$$

Normal fordelingen

Skrivemåde:

$$X \sim N(\mu, \sigma^2)$$

Tæthedsfunktion:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Normal fordelingen

Skrivemåde:

$$X \sim N(\mu, \sigma^2)$$

Tæthedsfunktion:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Middelværdi:

$$\mu = \mu$$

Normal fordelingen

Skrivemåde:

$X \sim N(\mu, \sigma^2)$

Tæthedsfunktion:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

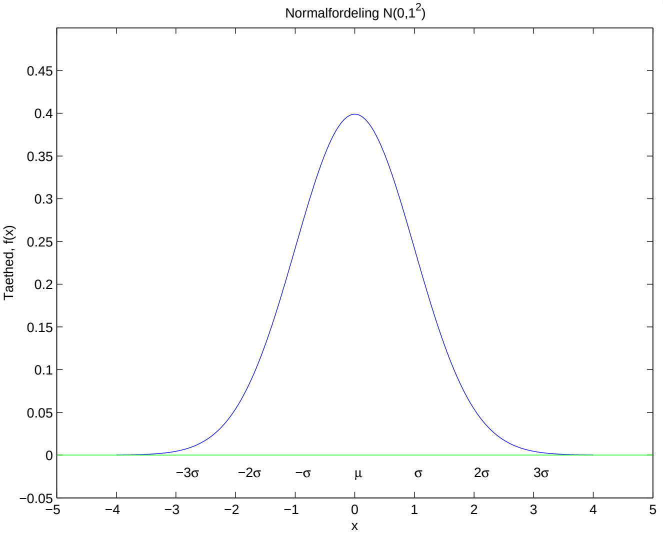
Middelværdi:

$\mu = \mu$

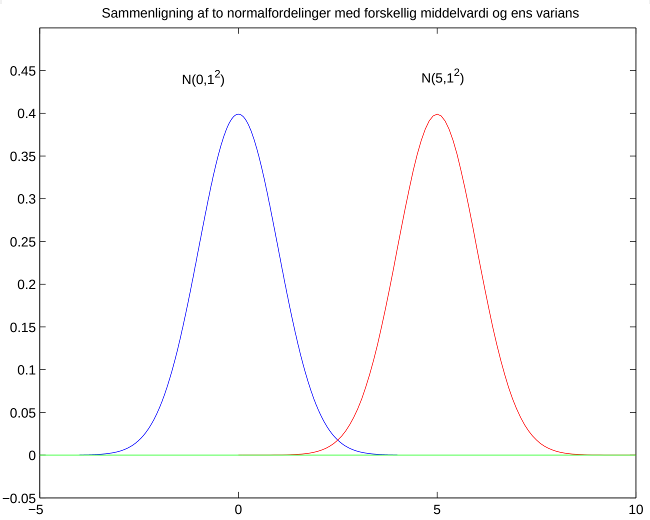
Varians:

$\sigma^2 = \sigma^2$

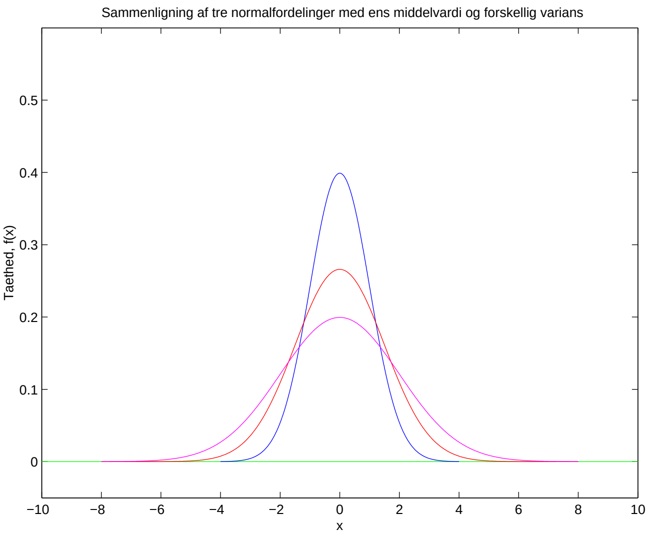
Normalfordelingen



Normalfordelingen



Normalfordelingen



Normal fordelingen

En standard normal fordeling:

$$Z \sim N(0, 1^2)$$

En normalfordeling med middelværdi 0 og varians 1.

Normal fordelingen

En standard normal fordeling:

$$Z \sim N(0, 1^2)$$

En normalfordeling med middelværdi 0 og varians 1.

Standardisering:

En vilkårlig normal fordelt variabel $X \sim N(\mu, \sigma^2)$ kan standardiseres ved at beregne

$$Z = \frac{X - \mu}{\sigma}$$

Eksempel 2

Målefejl:

En vægt har en målefejl, Z, der kan beskrives ved en standard normalfordeling, dvs

$$Z \sim N(0, 1^2)$$

dvs. middelværdi $\mu = 0$ og spredning $\sigma = 1$ gram.
Vi måler nu vægten af ét emne

Eksempel 2

Målefejl:

En vægt har en målefejl, Z, der kan beskrives ved en standard normalfordeling, dvs

$$Z \sim N(0, 1^2)$$

dvs. middelværdi $\mu = 0$ og spredning $\sigma = 1$ gram.
Vi måler nu vægten af ét emne

Spørgsmål a):

Hvad er sandsynligheden for at vægten måler mindst 2 gram for lidt?

Eksempel 2

Målefejl:

En vægt har en målefejl, Z, der kan beskrives ved en standard normalfordeling, dvs

$$Z \sim N(0, 1^2)$$

dvs. middelværdi $\mu = 0$ og spredning $\sigma = 1$ gram.
Vi måler nu vægten af ét emne

Spørgsmål a):

Hvad er sandsynligheden for at vægten måler mindst 2 gram for lidt?

Svar:

$P(Z \leq -2) = 0.02275$

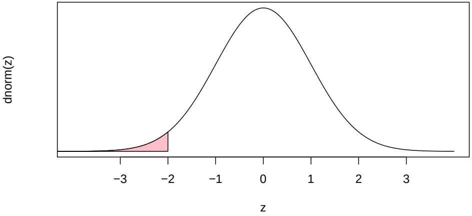
```
pnorm(-2)
```

Eksempel 2

Svar:

```
pnorm(-2)
```

```
[1] 0.02275
```



DTU Compute
Institut for Matematik og Computer Science

Eksempel 2

Spørgsmål b):

Hvad er sandsynligheden for at vægten måler mindst 2 gram for meget?

Eksempel 2

Spørgsmål b):

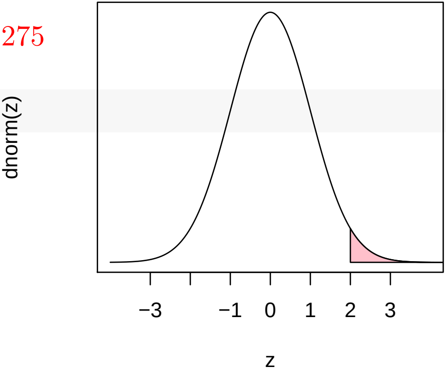
Hvad er sandsynligheden for at vægten måler mindst 2 gram for meget?

Svar:

$P(Z \geq 2) = 0.02275$

```
1-pnorm(2)
```

```
[1] 0.02275
```



Eksempel 2

Spørgsmål c):

Hvad er sandsynligheden for at vægten måler højst ± 1 *gram forkert?*

DTU Compute
Institut for Matematik og Computer Science

Eksempel 2

Spørgsmål c):

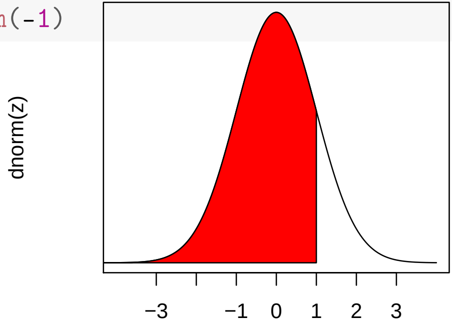
Hvad er sandsynligheden for at vægten måler højst ± 1 *gram forkert?*

Svar:

$$P(|Z| \leq 1) = P(-1 \leq Z \leq 1) = P(Z \leq 1) - P(Z \leq -1) = 0.683$$

```
pnorm(1)-pnorm(-1)
```

[1] 0.68269



Eksempel 2

Spørgsmål c):

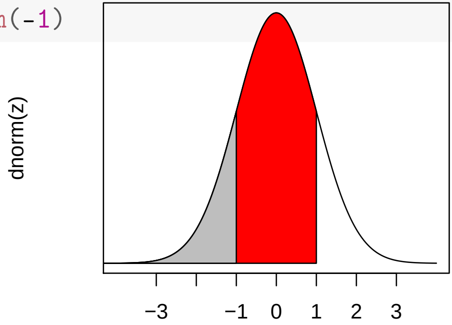
Hvad er sandsynligheden for at vægten måler højst ± 1 gram forkert?

Svar:

$P(|Z| \leq 1) = P(-1 \leq Z \leq 1) = P(Z \leq 1) - P(Z \leq -1) = 0.683$

```
pnorm(1)-pnorm(-1)
```

```
[1] 0.68269
```



Eksempel 3

Indkomstfordeling:

Det antages, at blandt en gruppe lærere i folkeskolen, at lønnen kan beskrives ved en normalfordeling med middelværdi $\mu = 280.000$ og spredning $\sigma = 10.000$.

DTU Compute
Institut for Matematik og Computer Science

Eksempel 3

Indkomstfordeling:

Det antages, at blandt en gruppe lærere i folkeskolen, at lønnen kan beskrives ved en normalfordeling med middelværdi $\mu = 280.000$ og spredning $\sigma = 10.000$.

Spørgsmål a):

Hvad er sandsynligheden for at en tilfældig udvalgt lærer tjener mere end 300.000?

Eksempel 3

Indkomstfordeling:

Det antages, at blandt en gruppe lærere i folkeskolen, at lønnen kan beskrives ved en normalfordeling med middelværdi $\mu = 280.000$ og spredning $\sigma = 10.000$.

Spørgsmål a):

Hvad er sandsynligheden for at en tilfældig udvalgt lærer tjener mere end 300.000?

Svar:

$P(X > 300) = P(Z > \frac{300-280}{10}) = P(Z > 2) = 0.023$

$$X \sim N(300, 10^2) \Rightarrow Z = \frac{X - 280}{10} \sim N(0, 1^2)$$

Eksempel 3

Spørgsmål a):

Hvad er sandsynligheden for at en tilfældig udvalgt lærer tjener mere end 300.000?

Eksempel 3

Spørgsmål a):

Hvad er sandsynligheden for at en tilfældig udvalgt lærer tjener mere end 300.000?

Eksempel 3

Spørgsmål a):

Hvad er sandsynligheden for at en tilfældig udvalgt lærer tjener mere end 300.000?

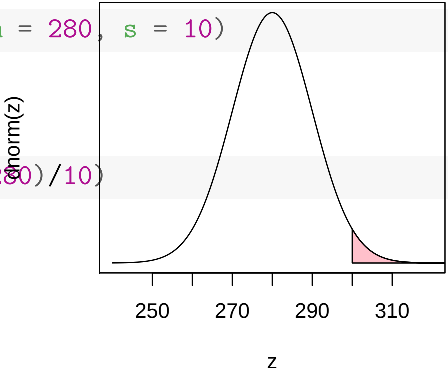
Svar:

```
1-pnorm(300, m = 280, s = 10)
```

[1] 0.02275

```
1-pnorm((300-280)/10)
```

[1] 0.02275



Eksempel 4

En mere smal fordeling:

Det antages, at blandt en gruppe lærere i folkeskolen, at lønnen kan beskrives ved en normalfordeling med middelværdi $\mu = 290.000$ og spredning $\sigma = 4.000$.

Eksempel 4

En mere smal fordeling:

Det antages, at blandt en gruppe lærere i folkeskolen, at lønnen kan beskrives ved en normalfordeling med middelværdi $\mu = 290.000$ og spredning $\sigma = 4.000$.

Spørgsmål a):

Hvad er sandsynligheden for at en tilfældig udvalgt lærer tjener mere end 300.000?

Eksempel 4

Spørgsmål a):

Hvad er sandsynligheden for at en tilfældig udvalgt lærer tjener mere end 300.000?

Eksempel 4

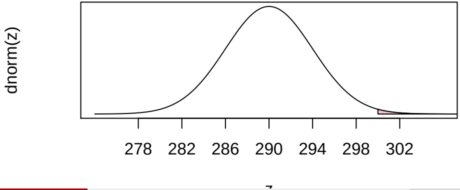
Spørgsmål a):

Hvad er sandsynligheden for at en tilfældig udvalgt lærer tjener mere end 300.000?

Svar:

```
1-pnorm(300, m = 290, s = 4)
```

[1] 0.0062097



Eksempel 5

Samme indkomstfordeling

Det antages, at blandt en gruppe lærere i folkeskolen, at lønnen kan beskrives ved en normalfordeling med middelværdi $\mu = 290.000$ og spredning $\sigma = 4.000$

DTU Compute
Institut for Matematik og Computer Science

Eksempel 5

Samme indkomstfordeling

Det antages, at blandt en gruppe lærere i folkeskolen, at lønnen kan beskrives ved en normalfordeling med middelværdi $\mu = 290.000$ og spredning $\sigma = 4.000$

"Omvendt spørgsmål"

Angiv det interval, der dækker over 95% af læreres løn

Eksempel 5

Samme indkomstfordeling

Det antages, at blandt en gruppe lærere i folkeskolen, at lønnen kan beskrives ved en normalfordeling med middelværdi $\mu = 290.000$ og spredning $\sigma = 4.000$

"Omvendt spørgsmål"

Angiv det interval, der dækker over 95% af læreres løn

Svar:

```
qnorm(c(0.025, 0.975), m = 290, s = 4)
```

[1] 282.16 297.84

Log-Normal fordelingen

Skrivemåde:

$X \sim LN(\alpha, \beta)$

Log-Normal fordelingen

Skrivemåde:

$X \sim LN(\alpha, \beta)$

Tæthedsfunktion:

$$f(x) = \begin{cases} \frac{1}{\beta\sqrt{2\pi}} x^{-1} e^{-(\ln(x)-\alpha)^2/2\beta^2} & x > 0,\ \beta > 0 \\ 0 & \text{ellers} \end{cases}$$

DTU Compute
Institut for Matematik og Computer Science

Log-Normal fordelingen

Skrivemåde:

$X \sim LN(\alpha, \beta)$

Tæthedsfunktion:

$$f(x) = \begin{cases} \frac{1}{\beta\sqrt{2\pi}} x^{-1} e^{-(\ln(x)-\alpha)^2/2\beta^2} & x > 0,\ \beta > 0 \\ 0 & \text{ellers} \end{cases}$$

Middelværdi:

$\mu = e^{\alpha + \beta^2/2}$

Log-Normal fordelingen

Skrivemåde:

$X \sim LN(\alpha, \beta)$

Tæthedsfunktion:

$$f(x) = \begin{cases} \frac{1}{\beta\sqrt{2\pi}} x^{-1} e^{-(\ln(x)-\alpha)^2/2\beta^2} & x > 0,\ \beta > 0 \\ 0 & \text{ellers} \end{cases}$$

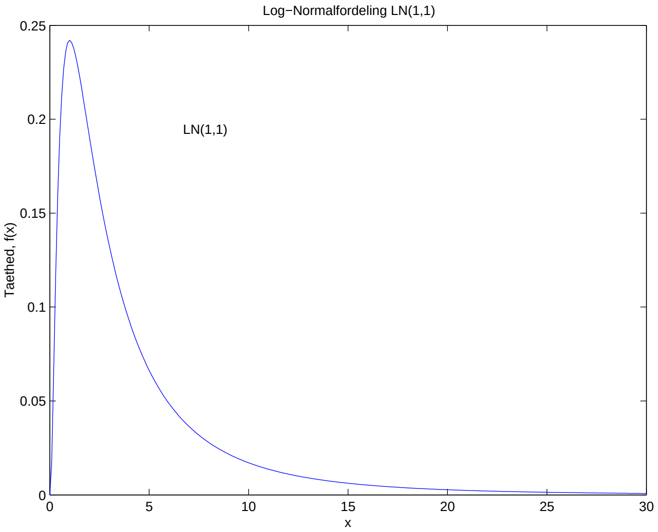
Middelværdi:

$\mu = e^{\alpha+\beta^2/2}$

Varians:

$\sigma^2 = e^{2\alpha+\beta^2}(e^{\beta^2} - 1)$

Log-Normal fordelingen



DTU Compute
Institut for Matematik og Computer Science

Log-Normal fordelingen

Lognormal og Normalfordelingen:

En log-normal fordelt variabel $Y \sim LN(\alpha, \beta)$, kan transformeres til en standard normal fordelt variabel X ved

$$X = \frac{\ln(Y) - \alpha}{\beta}$$

dvs.

$$X \sim N(0, 1^2)$$

DTU Compute
Institut for Matematik og Computer Science

Kontinuerte fordelinger i R

R	Betegnelse
`norm`	Normalfordelingen
`unif`	Den uniforme fordeling
`lnorm`	Log-normalfordelingen
`exp`	Exponentialfordelingen

- `d` Tæthedsfunktion $f(x)$ (probability density function).
- `p` Fordelingsfunktion $F(x)$ (cumulative distribution function).
- `q` Fraktil (quantile) i fordeling.
- `r` Tilfældige tal fra fordelingen.

Eksponential fordelingen

Tæthedsfunktionen

$$f(x) = \begin{cases} \frac{1}{\beta} e^{-x/\beta} & x > 0, \ \ \beta > 0 \\ 0 & \text{ellers} \end{cases}$$

DTU Compute
Institut for Matematik og Computer Science

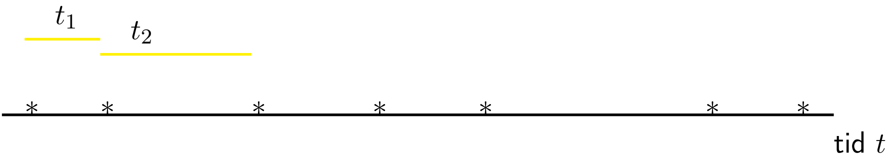
Eksponentialfordelingen

- Eksponential fordelingen er et special tilfælde af Gamma fordelingen
- Eksponential fordelingen anvendes f.eks. til at beskrive levetider og ventetider
- Eksponential fordelingen kan bruges til at beskrive (vente)tiden mellem hændelser i poisson fordelingen
- Middelværdi $\mu = \beta$
- Varians $\sigma^2 = \beta^2$

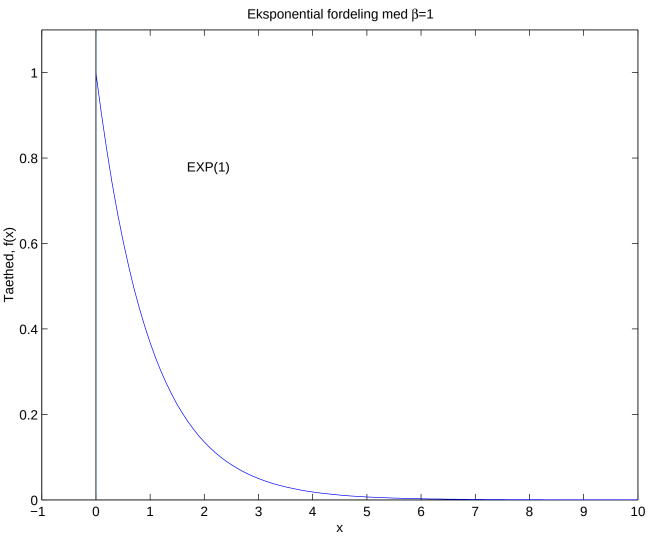
Sammenhæng mellem Eksponential og Poisson fordelingen

Poisson: Diskrete hændelser pr. enhed

Eksponential: Kontinuert afstand mellem hændelser



Eksponential fordelingen



Eksempel 6

Kø-model - poisson proces

Tiden mellem kundeankomster på et posthus er eksponential fordelt med middelværdi $\mu = 2$ minutter.

Eksempel 6

Kø-model - poisson proces

Tiden mellem kundeankomster på et posthus er eksponential fordelt med middelværdi $\mu = 2$ minutter.

Spørgsmål:

En kunde er netop ankommet. Hvad er sandsynligheden for at der ikke kommer flere kunder indefor en periode på 2 minutter?

Eksempel 6

Kø-model - poisson proces

Tiden mellem kundeankomster på et posthus er eksponential fordelt med middelværdi $\mu = 2$ minutter.

Spørgsmål:

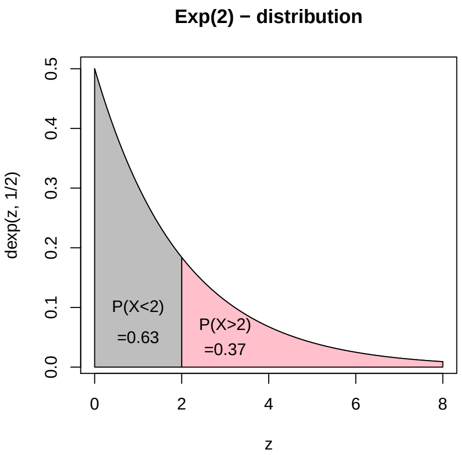
En kunde er netop ankommet. Hvad er sandsynligheden for at der ikke kommer flere kunder indefor en periode på 2 minutter?

Svar:

```
1-pexp(2, rate = 1/2)
```

[1] 0.36788

Eksempel 6



DTU Compute
Institut for Matematik og Computer Science

Eksempel 6

```
z=seq(0,8,by=0.01)

plot(z,dexp(z, 1/2),type = "l", main = "Exp(2) - distribution")

polygon(c(2, seq(2, 8, by = 0.01), 8, 2),
c(0, dexp(seq(2, 8, by = 0.01), 1/2), 0, 0),
col = "pink")

text(3,0.07,"P(X>2)")
text(3,0.03,"=0.37")

polygon(c(2, seq(2, 0, by = -0.01), 0, 2),
c(0, dexp(seq(2, 0, by =- 0.01), 1/2), 0, 0),
col = "grey")

text(1,0.1,"P(X<2)")
text(1,0.05,"=0.63")
```

Eksempel 7

Spørgsmål:

En kunde er netop ankommet. Beregn sandsynligheden for at der ikke kommer flere kunder indefor en periode på 2 minutter vha. Poissonfordelingen

Eksempel 7

Spørgsmål:

En kunde er netop ankommet. Beregn sandsynligheden for at der ikke kommer flere kunder indefor en periode på 2 minutter vha. Poissonfordelingen

Eksempel 7

Spørgsmål:

En kunde er netop ankommet. Beregn sandsynligheden for at der ikke kommer flere kunder indefor en periode på 2 minutter vha. Poissonfordelingen

Svar:

$\lambda_{2min} = 1, \; P(X = 0) = \frac{e^{-1}}{1!} 1^0 = e^{-1}$

```
dpois(0,1)
```

```
[1] 0.36788
```

```
exp(-1)
```

```
[1] 0.36788
```

Eksempel 8

Andre tidsperioder:

Tiden mellem kundeankomster på et posthus er eksponential fordelt med middelværdi $\mu = 2$ minutter. Vi betragter nu en periode på 10 minutter

DTU Compute
Institut for Matematik og Computer Science

Eksempel 8

Andre tidsperioder:

Tiden mellem kundeankomster på et posthus er eksponential fordelt med middelværdi $\mu = 2$ minutter. Vi betragter nu en periode på 10 minutter

Spørgsmål:

Beregn sandsynligheden for at der ikke kommer nogen kunder i perioden vha. Poissonfordelingen

DTU Compute
Institut for Matematik og Computer Science

Eksempel 8

Andre tidsperioder:

Tiden mellem kundeankomster på et posthus er eksponential fordelt med middelværdi $\mu = 2$ minutter. Vi betragter nu en periode på 10 minutter

Spørgsmål:

Beregn sandsynligheden for at der ikke kommer nogen kunder i perioden vha. Poissonfordelingen

Svar:

$\lambda_{10min} = 5,\ P(X = 0) = \frac{e^{-5}}{1!} 5^0 = e^{-5}$

```
dpois(0,5)
```

```
[1] 0.0067379
```

Oversigt

Regneregler for stokastiske variable

(Gælder BÅDE kontinuert og diskret)

X er en stokastisk variabel

. Vi antager at a og b er konstanter Da gælder:

Regneregler for stokastiske variable

(Gælder BÅDE kontinuert og diskret)

X er en stokastisk variabel

. Vi antager at a og b er konstanter Da gælder:

Middelværdi-regel:

$$E(aX + b) = aE(X) + b$$

Regneregler for stokastiske variable

(Gælder BÅDE kontinuert og diskret)

X er en stokastisk variabel

. Vi antager at a og b er konstanter Da gælder:

Middelværdi-regel:

$$E(aX + b) = aE(X) + b$$

Varians-regel:

$$Var(aX + b) = a^2 Var(X)$$

Eksempel 9

X er en stokastisk variabel

. En stokastisk variabel X har middelværdi 4 og varians 6.

Eksempel 9

X er en stokastisk variabel

. En stokastisk variabel X har middelværdi 4 og varians 6.

Spørgsmål:

Beregn middelværdi og varians for $Y = -3X + 2$

DTU Compute
Institut for Matematik og Computer Science

Eksempel 9

X er en stokastisk variabel

. En stokastisk variabel X har middelværdi 4 og varians 6.

Spørgsmål:

Beregn middelværdi og varians for $Y = -3X + 2$

Svar:

$$\mathsf{E}(Y) = -3\mathsf{E}(X) + 2 = -3 \cdot 4 + 2 = -10$$

$$\mathsf{Var}(Y) = (-3)^2 \mathsf{Var}(X) = 9 \cdot 6 = 54$$

Regneregler for stokastiske variable

$X_1, \ldots, X_n$ er stokastiske variable

Da gælder (når de er uafhængige):

DTU Compute
Institut for Matematik og Computer Science

Regneregler for stokastiske variable

$X_1, \ldots, X_n$ er stokastiske variable

Da gælder (når de er uafhængige):

Middelværdi-regel:

$$E(a_1X_1 + a_2X_2 + .. + a_nX_n)$$
$$= a_1E(X_1) + a_2E(X_2) + .. + a_nE(X_n)$$

DTU Compute
Institut for Matematik og Computer Science

Regneregler for stokastiske variable

$X_1, \ldots, X_n$ er stokastiske variable

Da gælder (når de er uafhængige):

Middelværdi-regel:

$$E(a_1 X_1 + a_2 X_2 + .. + a_n X_n)$$
$$= a_1 E(X_1) + a_2 E(X_2) + .. + a_n E(X_n)$$

Varians-regel:

$$Var(a_1 X_1 + a_2 X_2 + .. + a_n X_n)$$
$$= a_1^2 Var(X_1) + .. + a_n^2 Var(X_n)$$

Eksempel 10

Flypassager-planlægning

Vægten af passagerer på en flystrækning antages normalfordelt $X \sim N(70, 10^2)$.

Et fly, der kan tage 55 passagerer, må max. lastes med 4000 kg (kun passageres vægt betragtes som last).

Eksempel 10

Flypassager-planlægning

Vægten af passagerer på en flystrækning antages normalfordelt $X \sim N(70, 10^2)$.

Et fly, der kan tage 55 passagerer, må max. lastes med 4000 kg (kun passageres vægt betragtes som last).

Spørgsmål:

Beregn sandsynligheden for at flyet bliver overlastet

DTU Compute
Institut for Matematik og Computer Science

Eksempel 10

Flypassager-planlægning

Vægten af passagerer på en flystrækning antages normalfordelt $X \sim N(70, 10^2)$.

Et fly, der kan tage 55 passagerer, må max. lastes med 4000 kg (kun passageres vægt betragtes som last).

Spørgsmål:

Beregn sandsynligheden for at flyet bliver overlastet

Hvad er Y=Total passagervægt?

Eksempel 10

Flypassager-planlægning

Vægten af passagerer på en flystrækning antages normalfordelt $X \sim N(70, 10^2)$.

Et fly, der kan tage 55 passagerer, må max. lastes med 4000 kg (kun passageres vægt betragtes som last).

Spørgsmål:

Beregn sandsynligheden for at flyet bliver overlastet

Hvad er Y=Total passagervægt?

Hvad er Y?

I hvert fald IKKE: $Y = 55 \cdot X$!!!!!!

Eksempel 10

Hvad er Y=Total passagervægt?

$Y = \sum_{i=1}^{55} X_i$, hvor $X_i \sim N(70, 10^2)$

Eksempel 10

Hvad er Y=Total passagervægt?

$Y = \sum_{i=1}^{55} X_i$, hvor $X_i \sim N(70, 10^2)$

Middelværdi og varians for Y:

$$E(Y) = \sum_{i=1}^{55} E(X_i) = \sum_{i=1}^{55} 70 = 55 \cdot 70 = 3850$$

$$\mathrm{Var}(Y) = \sum_{i=1}^{55} \mathrm{Var}(X_i) = \sum_{i=1}^{55} 100 = 55 \cdot 100 = 5500$$

DTU Compute
Institut for Matematik og Computer Science

Eksempel 10

Hvad er Y=Total passagervægt?

$Y = \sum_{i=1}^{55} X_i$, hvor $X_i \sim N(70, 10^2)$

Middelværdi og varians for Y:

$$E(Y) = \sum_{i=1}^{55} E(X_i) = \sum_{i=1}^{55} 70 = 55 \cdot 70 = 3850$$

$$\text{Var}(Y) = \sum_{i=1}^{55} \text{Var}(X_i) = \sum_{i=1}^{55} 100 = 55 \cdot 100 = 5500$$

Bruger normalfordeling for Y:

```
1-pnorm(4000, m = 3850, s = sqrt(5500))
```

```
[1] 0.021557
```

Eksempel 10 - FORKERT ANALYSE

Hvad er Y?

I hvert fald IKKE: $Y = 55 \cdot X$!!!!!!

Eksempel 10 - FORKERT ANALYSE

Hvad er Y?

I hvert fald IKKE: $Y = 55 \cdot X$!!!!!!

Middelværdi og varians for Y:

$$E(Y) = 55 \cdot 70 = 3850$$

$$\text{Var}(Y) = 55^2 \text{Var}(X) = 55^2 \cdot 100 = 550^2$$

Eksempel 10 - FORKERT ANALYSE

Hvad er Y?

I hvert fald IKKE: $Y = 55 \cdot X$!!!!!!

Middelværdi og varians for Y:

$$E(Y) = 55 \cdot 70 = 3850$$

$$\text{Var}(Y) = 55^2 \text{Var}(X) = 55^2 \cdot 100 = 550^2$$

Bruger normalfordeling for Y:

```
1-pnorm(4000, m = 3850, s = 550)
```

[1] 0.39253

Eksempel 10 - FORKERT ANALYSE

Hvad er Y?

I hvert fald IKKE: $Y = 55 \cdot X$!!!!!!

Middelværdi og varians for Y:

$$E(Y) = 55 \cdot 70 = 3850$$

$$\text{Var}(Y) = 55^2 \text{Var}(X) = 55^2 \cdot 100 = 550^2$$

Bruger normalfordeling for Y:

```
1-pnorm(4000, m = 3850, s = 550)
```

[1] 0.39253

Konsekvens af forkert beregning:

MANGE spildte penge for flyselskabet!!!

Oversigt

DTU Compute
Institut for Matematik og Computer Science