Landmålingens fejlteori
Repetition - Kontinuerte stokastiske variable
Lektion 3

Torben Tvedebrink - tvede@math.aau.dk
http://www.math.aau.dk/∼tvede/teaching/L4

Institut for Matematiske Fag
Aalborg Universitet

1. maj 2009

Definitioner

Definition 1

Fordelingsfunktion for en stokastisk variabel X er

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

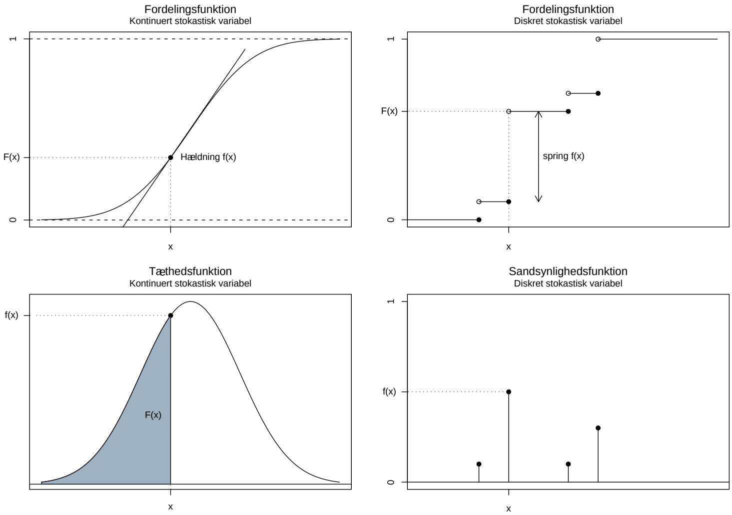
Definition 2

Lad X være en stokastisk variabel med fordelingsfunktion F. Da siges X at være **kontinuert**, hvis F er kontinuert, og hvis der findes en ikke-negativ reel funktion f, således at

$$F(x) = \int_{-\infty}^{x} f(y)dy.$$

Funktionen f kaldes **tæthedsfunktion** for X

Tæthed-, sandsynlighed- og fordelingsfunktion



Middelværdi og varians

Definition 3

Lad X være en kontinuert stokastisk variabel med tæthedsfunktion f.
Middelværdien af X defineres som

$$\mu = \mathbb{E}(X) = \int_{-\infty}^{\infty} x f(x) dx < \infty$$

- også kaldet forventet værdi.

Definition 3

Lad X være en kontinuert stokastisk variabel med tæthedsfunktion f.
Variansen af X defineres som

$$\sigma^2 = \mathbb{V}(X) = \mathbb{E}((X-\mu)^2) = \int_{-\infty}^{\infty} (x-\mu)^2 f(x) dx < \infty,$$

hvor σ kaldes spredningen (eller standard afvigelse) - mål for forventet variabilitet i X.

Der gælder (som vi så ved diskrete stokastiske variable)

$$\sigma^2 = \mathbb{V}(X) = \int_{-\infty}^{\infty} x^2 f(x) dx - \mu^2 = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = \mathbb{E}(X^2) - \mu^2.$$

Mest benyttede kontinuerte fordeling.

Definition 4

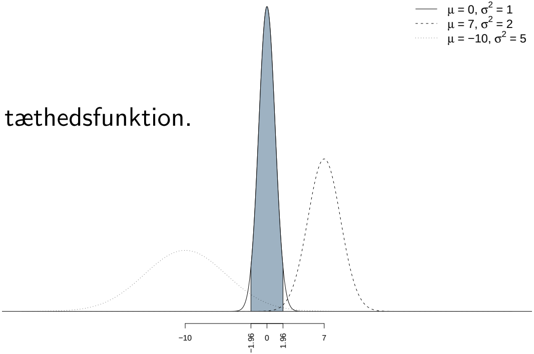
En stokastisk variabel X med **tæthedsfunktion**

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(- \frac{(x-\mu)^2}{2\sigma^2} \right), \quad x \in \mathbb{R},$$

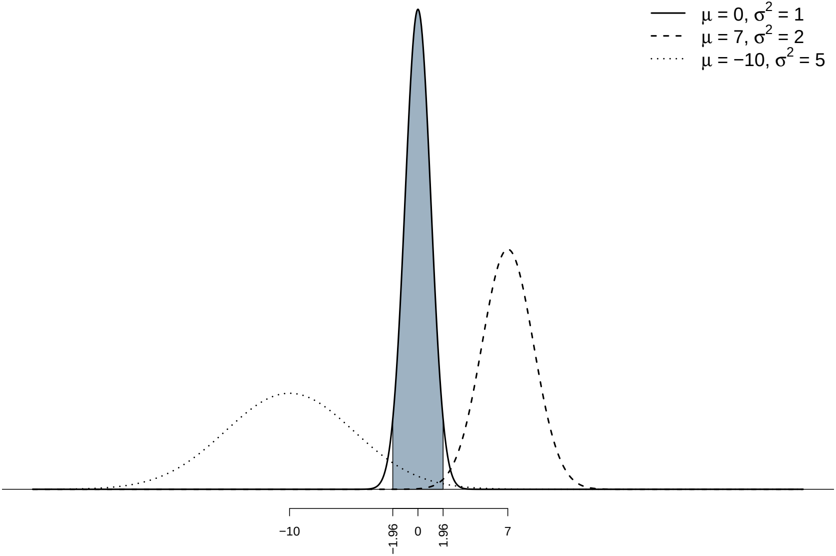
siges at være **normalfordelt** med parameterne μ og σ^2, hvor μ og σ er reelle tal og $\sigma > 0$.

Notation: $X \sim \mathcal{N}(\mu, \sigma^2)$.

Klokkeformet symmetrisk tæthedsfunktion.



Normalfordelingen



Sætning 2

Hvis $X \sim \mathcal{N}(\mu, \sigma^2)$, så er **middelværdi** og **varians** for X

$$
\begin{aligned}
\mathbb{E}(X) &= \mu, \\
\mathbb{V}(X) &= \sigma^2.
\end{aligned}
$$

Fordelingsfunktion for $X \sim \mathcal{N}(\mu, \sigma^2)$ er givet som

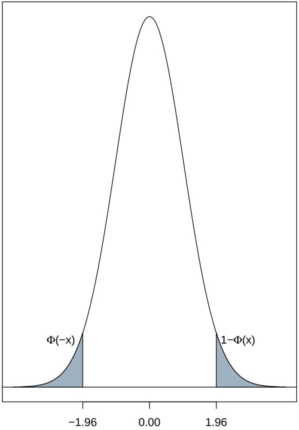
$$
F(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(- \frac{(y-\mu)^2}{2\sigma^2} \right) dy.
$$

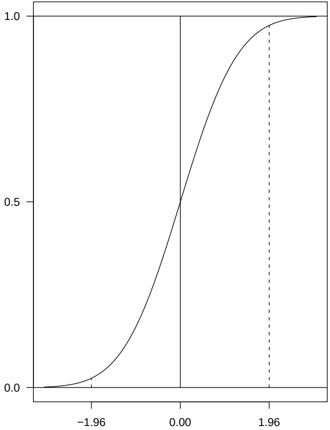
Der findes ikke et pænt lukket udtryk.

Definition

Fordelingen $\mathcal{N}(0,1)$ kaldes **standard normalfordeling**.
Den tilhørende fordelingsfunktion betegnes Φ - tabellagt.

Tæthed- og fordelingsfunktion for $\mathcal{N}(0,1)$.





Normalfordelingen

Definition

Fordelingen $\mathcal{N}(0,1)$ kaldes **standard normalfordeling**.
Den tilhørende fordelingsfunktion betegnes Φ - tabellagt.
Der gælder

$$\Phi(-x) \quad = \quad 1 - \Phi(x).$$

$\Rightarrow$ kun nødvendigt med tabeller for $x \geq 0$.

Sætninger

Sætning **Lineær transformation**

Hvis $X \sim \mathcal{N}(\mu, \sigma^2)$, og $a, b \in \mathbb{R}$ og $a \neq 0$, så er

$$Y = aX + b \sim \mathcal{N}(a\mu + b, a^2\sigma^2).$$

Sætning **Standardisere**

Hvis $X \sim \mathcal{N}(\mu, \sigma^2)$, så gælder

$$Z = \frac{X - \mu}{\sigma} \sim \mathcal{N}(0, 1).$$

Dvs.

$$P(X \leq x) = P\left(\frac{X - \mu}{\sigma} \leq \frac{x - \mu}{\sigma}\right) = P\left(Z \leq \frac{x - \mu}{\sigma}\right) = \Phi\left(\frac{x - \mu}{\sigma}\right).$$

Sandsynligheder vha. tabel for standard normalfordelingen.

Eksempel

Opgave 3.20 a)

Hvad er det forventede antal værneplitige over 170 cm når højden af værneplitige X antages at være normal fordelt med middelværdi 176.3 cm og spredning 6.4 cm, $X \sim \mathcal{N}(176.3, 6.4^2)$, til en session med 1000 fremmødte?

$P(X \geq 170) =$

Lad Y være antallet af værnepligtige over 170 cm ved en session med 1000 værnepligtige. Dvs. $Y \sim b(1000, P(X \geq 170))$ hvor vi ved $\mathbb{E}(Y) = np$.

$\mathbb{E}(Y) = 1000 \cdot P(X \geq 170) =$

Opslag i normalfordelingstabel

	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365

Tæthedsfunktion - fraktiler

Lad $X \sim \mathcal{N}(\mu, \sigma^2)$, dvs. $Z = \frac{X-\mu}{\sigma} \sim \mathcal{N}(0,1)$.
Betragt sandsynligheden

$$
\begin{aligned}
p = P(\mu - z\sigma < X < \mu + z\sigma) &= P(-z < \frac{X-\mu}{\sigma} < z) \\
&= P(-z < Z < z) \\
&= P(Z < z) - P(Z \leq -z) \\
&= \Phi(z) - \Phi(-z) \\
&= \Phi(z) - (1 - \Phi(z)) \\
&= 2\Phi(z) - 1.
\end{aligned}
$$

Tæthedsfunktion - fraktiler

Lad $X \sim \mathcal{N}(\mu, \sigma^2)$, dvs. $Z = \frac{X-\mu}{\sigma} \sim \mathcal{N}(0,1)$.

Vi har $p = 2\Phi(z) - 1$ for et valgt p hvor z findes ved tabelopslag.

For $p = 0.95$ gælder

$$2\Phi(z) - 1 = 0.95 \;\Leftrightarrow\; \Phi(z) = 0.975 \;\Leftrightarrow\; z = 1.96.$$

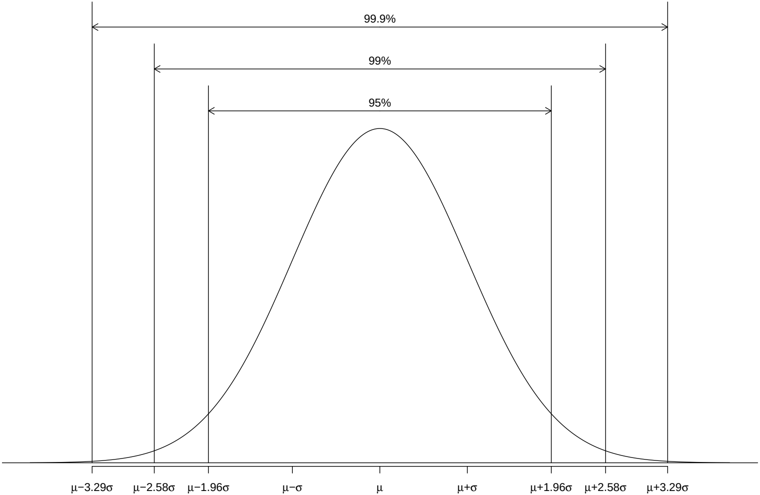
Dvs.

$$P(\mu - 1.96\sigma < X < \mu + 1.96\sigma) = 0.95.$$

Tilsvarende for $p = 0.99$ og $p = 0.999$

$$\begin{aligned} P(\mu - 2.58\sigma < X < \mu + 2.58\sigma) &= 0.99, \\ P(\mu - 3.29\sigma < X < \mu + 3.29\sigma) &= 0.999. \end{aligned}$$

Tæthedsfunktion



Dvs. middelværdi $\pm 2\times$standard afvigelse dækker ca. 95% af sandsynlighedsmassen.

Normalfordelingen i landmåling

Når vinkler og længder måles i landmåling antages disse størrelser at være normalfordelte (Kapitel 1 i Karsten Jensens noter).

Dvs. når en vinkel/længde måles antages det at målingen x er en realisation af en stokastisk variabel som følger en normalfordeling.

Eksempelvis er en vinkelmåling normalfordelt $\mathcal{N}(\beta, \sigma_\beta^2)$ hvor udtryk for σ_β udledes i dagens forlæsning.