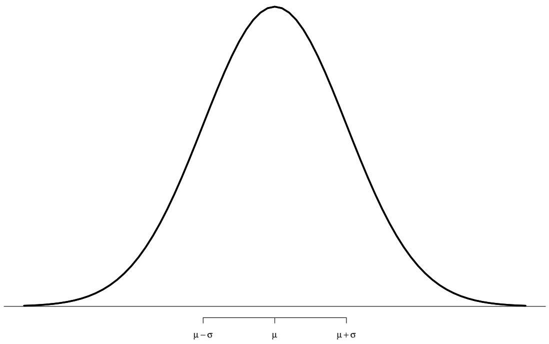
Kapittel 6: Kontinuerlige sannsynlighetsfordelinger

TMA4240 Statistikk (F2 og E7)

Foreleses 15. september, 2004.



6.1 Kontinuerlig uniform fordeling

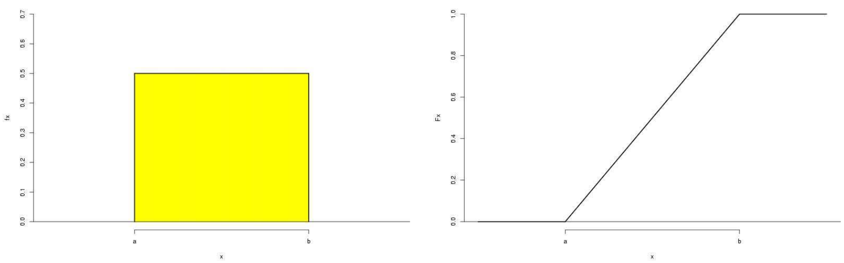
Kontinuerlig uniform fordeling: Sannsynlighetstettheten til den kontinuerlige uniforme stokastiske variabelen X på intervallet $[A, B]$ er

$$
\begin{aligned}
f(x; A, B) &= \frac{1}{B - A}, \qquad A \le x \le B \\
&= 0 \qquad \text{ellers.}
\end{aligned}
$$

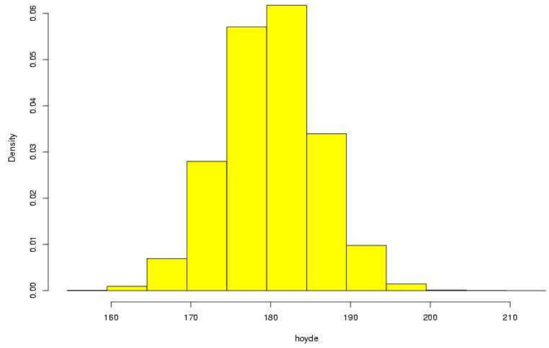
$$\mu = E(X) = \frac{A + B}{2}$$

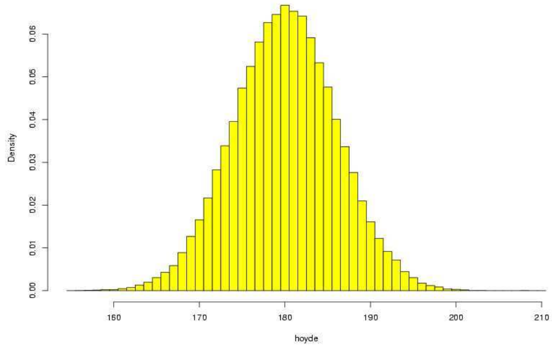
og varians

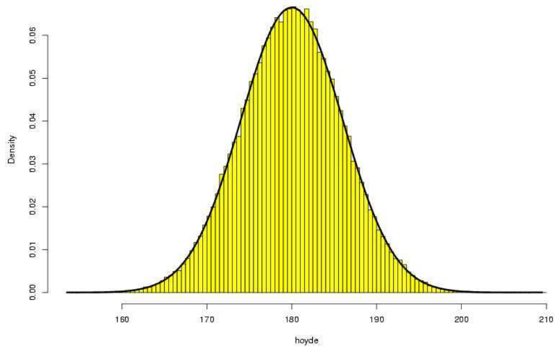
$$\sigma^2 = \mathrm{Var}(X) = \frac{(B - A)^2}{12}$$



6.2 Normalfordelingen







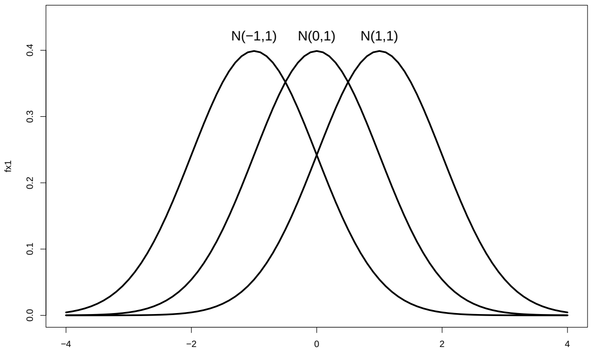
Normalfordelingen DEF

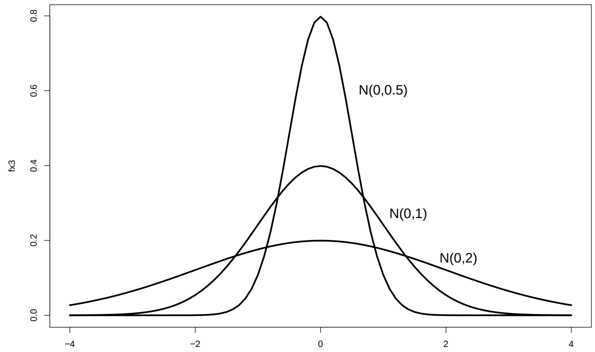
Normalfordeling: Sannsynlighetstettheten til en normalfordelt stokastisk variabel, X, med forventning $E(X) = \mu$ og varians $\mathrm{Var}(X) = \sigma^2$, er gitt ved

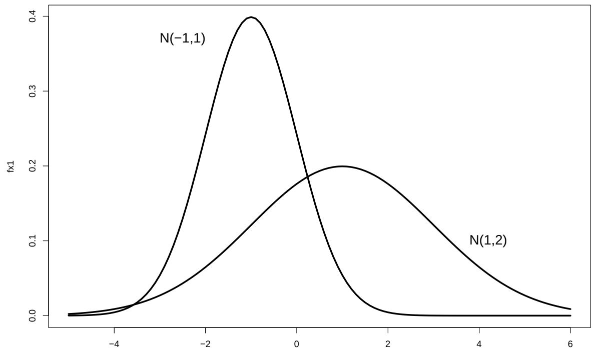
$$n(x;\mu,\sigma) = \frac{1}{\sqrt{2\cdot\pi}\sigma} e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}},$$

for $-\infty < x < \infty$, der π=3.14159... og $\exp = 2.71828$.

Lokasjon og spredning

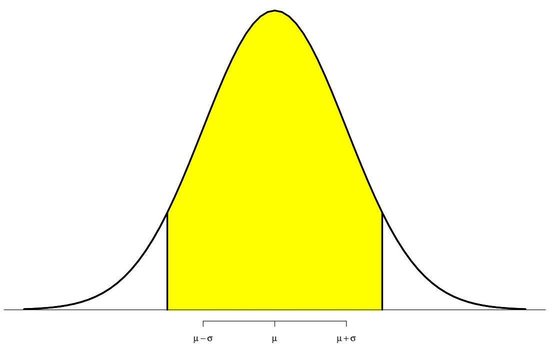






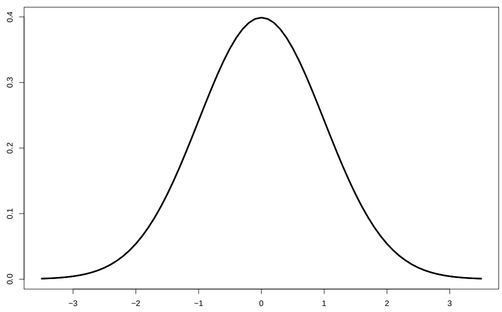
6.3 Sannsynligheter i normalfordelingen

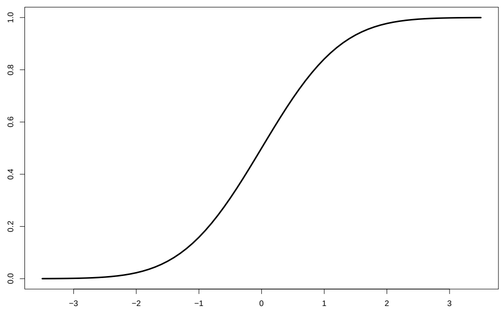
$$
\begin{aligned}
P(x_1 < X < x_2) &= \int_{x_1}^{x_2} n(x;\mu,\sigma)dx \\
&= \frac{1}{\sqrt{2\cdot\pi}\sigma}\int_{x_1}^{x_2} e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}}dx
\end{aligned}
$$



Standard normalfordeling

DFF 6.1: Fordelingen til en normalfordelt stokastisk variabel, X, med forventning $E(X) = 0$ og varians $\mathrm{Var}(X) = 1$ kalles en *standard normalfordeling*.

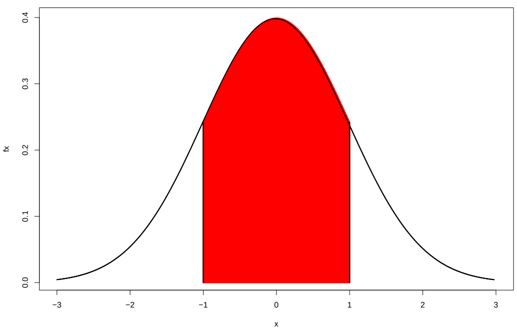


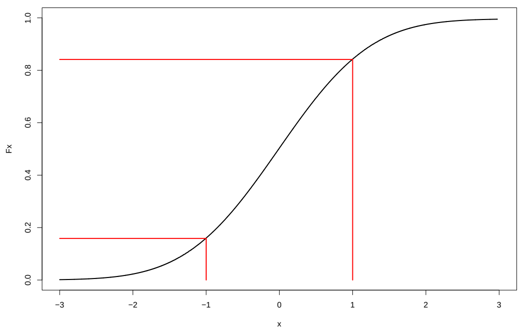


$N(\mu, \sigma)$ og $N(0, 1)$

- X har fordeling $n(x; \mu, \sigma)$
- $Z = \frac{X-\mu}{\sigma}$ har fordeling $n(z; 0, 1)$

$$\begin{aligned} P(x_1 \leq X \leq x_2) &= P(\frac{x_1 - \mu}{\sigma} \leq Z \leq \frac{x_2 - \mu}{\sigma}) \\ &= F(\frac{x_2 - \mu}{\sigma}) - F(\frac{x_1 - \mu}{\sigma}) \\ &= \Phi(\frac{x_2 - \mu}{\sigma}) - \Phi(\frac{x_1 - \mu}{\sigma}) \end{aligned}$$





IQ

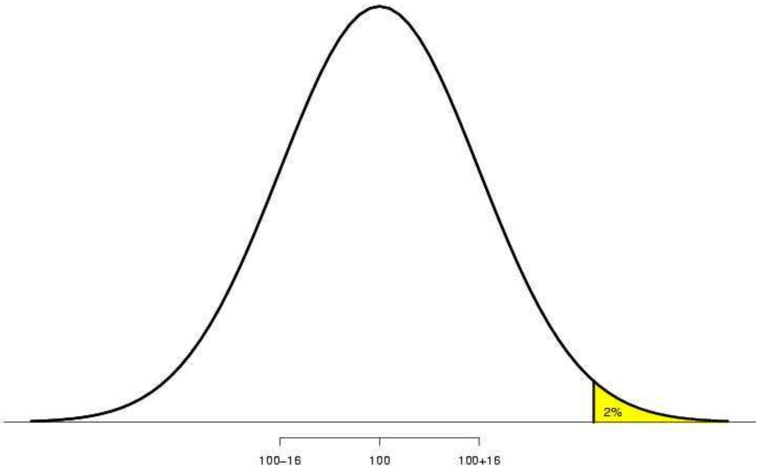
- Det finnes mange IQ-tester. Man antar at poengsummen fra en IQ-test er normalfordelt.
- Flere av IQ-testene antas å ha en forventningsverdi på 100 og et standardavvik på 16.
- Skalaen kan deles inn i kategorier:

140 and over	Genius or near genius
120-140	Very superior intelligence
110-120	Superior intelligence
90-110	Normal or average intelligence
80-90	Dullness
70-80	Borderline deficiency
Below 70	Definite feeble-mindedness

IQ (forts.)

- Hva er sannsynligheten for å ha en IQ lavere enn 100?
- Hva er sannsynligheten for å ha en IQ høyere enn 120?
- Hva er sannsynligheten for å ha en IQ mellom 80 og 120?
- For å bli med i Mensa må man oppnå en poengsum høyere enn 98 percentilen i fordelingen for testen. Hvor høy poengsum må man ha for å blir medlem av Mensa?

Fra halesannsynlighet til verdi



$$z = \frac{x-\mu}{\sigma} \text{ kan gi } x = \sigma z + \mu$$

Halesannsynligheter gir z. Vi kan regne ut x.

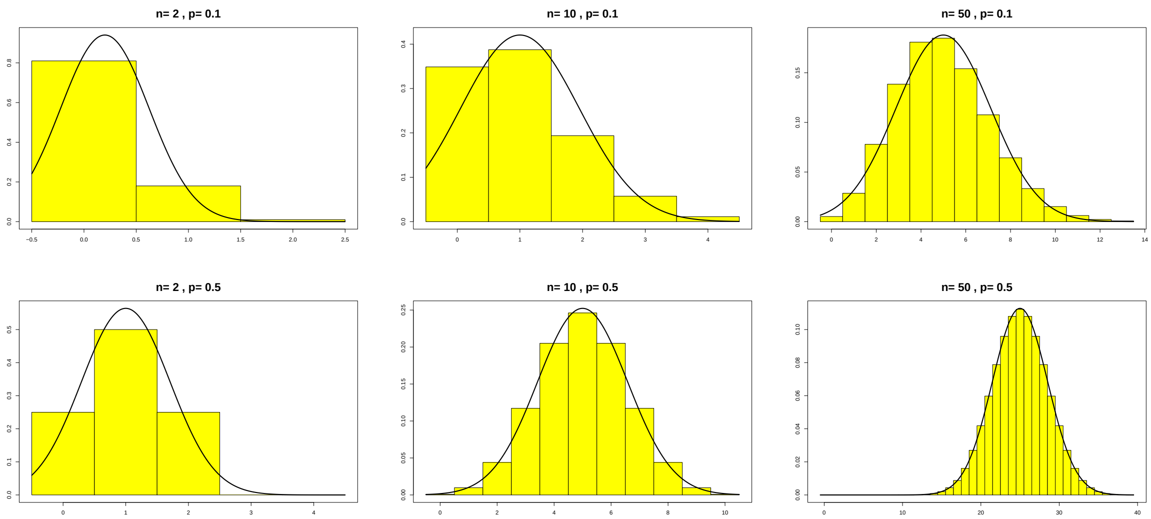
6.5 Normalapproksimasjon til binomisk fordeling

TEO 6.2 Hvis X er en binomisk stokastisk variabel med forventning $\mu = np$ og varians $\sigma^2 = np(1-p)$, så vil den stokastiske variablen

$$Z = \frac{X - np}{\sqrt{np(1-p)}}$$

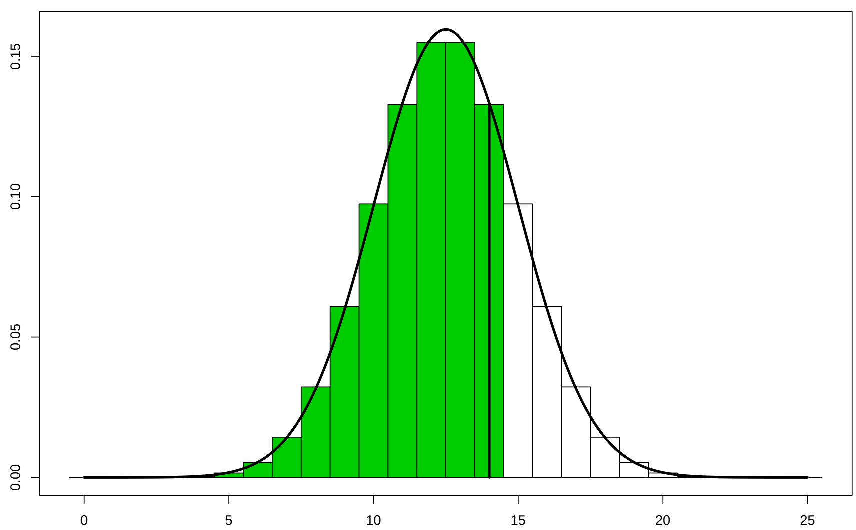
når $n \to \infty$ være tilnærmet standard normalfordelt, $n(z; 0, 1)$.

Binomisk til normal



Regneeksempel fra Cartoon Guide

- Binomisk med $n = 25$, $p = 0.5$.
- Hva er $P(x \leq 14)$?



Normalapproksimasjon til binomisk fordeling med korreksjon

- X er en binomisk stokastisk variabel med forventning $\mu = np$ og varians $\sigma^2 = np(1-p)$.
- Ønsker å beregne: $P(x_1 \leq X \leq x_2)$.
- Har at $Z = \frac{X-np}{\sqrt{np(1-p)}}$ er tilnærmet standard normalfordelt for stor n .
- På grunn av "kantene" på binomisk fordeling kan vi gjøre approksimasjonen bedre med en "kontinuitets-korreksjon":

$$P(x_1 \leq X \leq x_2) \approx P(\frac{x_1 - 0.5 - np}{\sqrt{np(1-p)}} \leq Z \leq \frac{x_2 + 0.5 - np}{\sqrt{np(1-p)}})$$

- Approksimasjonen er er akseptabel hvis både $np > 5$ og $n(1-p) > 5$.

Eksamen Des2003 siste tog...

- P(mer enn 2 minutter forsinket)=0.09 fra tidligere.
- La V være antall ganger i løpet av en måned (= 22 hverdager) at toget er mer enn 2 minutter forsinket. Foreslå en sannsynlighetsfordeling for V og sett opp de forutsetninger som ligger til grunn for denne.
- Hva er sannsynligheten for at toget er mer enn 2 minutter forsinket minst 2 ganger i løpet av en måned (= 22 hverdager)?
- Hva er (tilnærmet) sannsynligheten for at toget er mer enn 2 minutter forsinket mer enn 30 ganger i løpet av 220 hverdager?