

- ① Kap 2-7: Sannsynsteori
- ② Kap 8-11: Statistikk

Statistikk er innsamling, bearbeiding, analyse og tolkning av *data*.
Tar beslutningar under usikkerheit.

Del 1: Antar at vi kjenner $f(x)$ (inkludert parametre).

Del 2: Ønsker å finne parametre / teste hypoteser om parametre frå **data**.

Kap 8: Utvalfordelingar og databeskrivelse

Kap 9: Estimering

Kap 10: Hypotesetesting

Kap 11: Enkel lineær regresjon

Idag

- Utvalsfordelingar motivert frå hypotesetesting og estimering.

Øvingar

- Må ha godkjent 4 av 6 øvingar i kvar bolc.
- Øving 7 (b2-1) vil er ei nesten rein Matlab-øving.
- Må levere enten denne eller øving 1.

To typer ventilar til undervassinstallasjonen

Ønsker å finne ut om det er kvalitetsforskjell på ventilane A og B .

- Har testa $n = 30$ av kvar, og fått $\bar{x}_A = 2100$ og $\bar{x}_B = 2157$.
- Antar kjente variansar $\sigma_A^2 = 100^2$ og $\sigma_B^2 = 200^2$
- Forskjell mellom dei, eller tilfeldig variasjon?

Lineærkombinasjon av normalfordelte variable er normalfordelt.

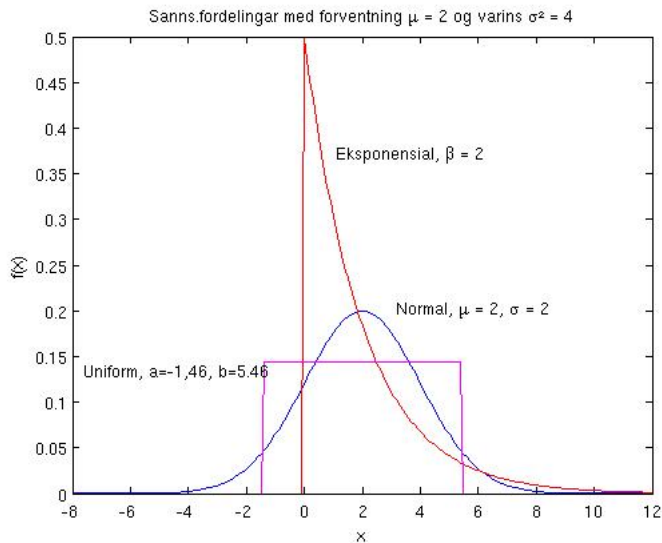
Teorem 7.11

La $X_1, X_2, \dots, X_n$ vere *uavhengige* normalfordelte variable med $E(X_i) = \mu_i$ og $Var(X_i) = \sigma_i^2$.

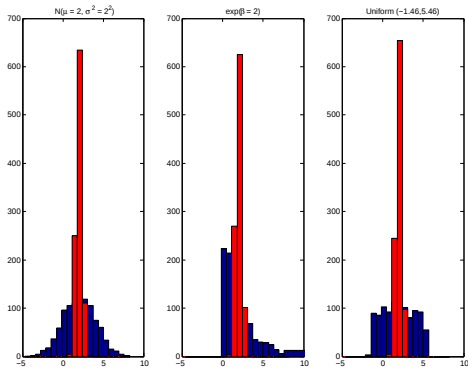
La $Y = a_0 + a_1X_1 + \dots + a_nX_n$. Då er

$$y \sim N(a_0 + a_1\mu_1 + \dots + a_n\mu_n, a_1^2\sigma_1^2 + \dots + a_n^2\sigma_n^2)$$

Felles forventning og varians



Histogram populasjon og gj.snitt av 30



Vi har sett på fordelinga til gjennomsnittet, $\bar{X}$.

Sentralgrenseteoremet, teorem 8.2

La X_i , $i = 1, 2, \dots, n$ vere uavhengige identisk fordelte (u.i.f.) stokastiske variable med $E(X_i) = \mu$ og $Var(X_i) = \sigma^2 < \infty$ (endeleg varians). La $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$ og $Z = \frac{\bar{X} - \mu}{\sqrt{\sigma^2/n}}$. Då

$$Z \rightarrow N(0, 1)$$

når $n \rightarrow \infty$.

Dette tilsvarer

$$\bar{X} \rightarrow N\left(\mu, \frac{\sigma^2}{n}\right)$$

når $n \rightarrow \infty$.

PS: Gjeld uansett fordeling for X_i