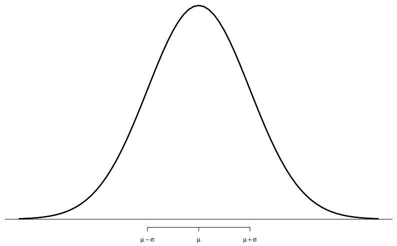
Kapittel 9: Estimering

TMA4240 Statistikk (F2 og E7)

9.7 og 9.8: To-utvalg og par-observasjoner.

Foreleses 20. oktober, 2004.



8.7 t-fordeling

TEO 8.5: La Z være en standard normalfordelt stokastisk variabel og V være en kjikvadrat-fordelt stokastisk variabel med ν frihetsgrader. Hvis Z og V er uavhengige er fordelingen til den stokastiske variablen T

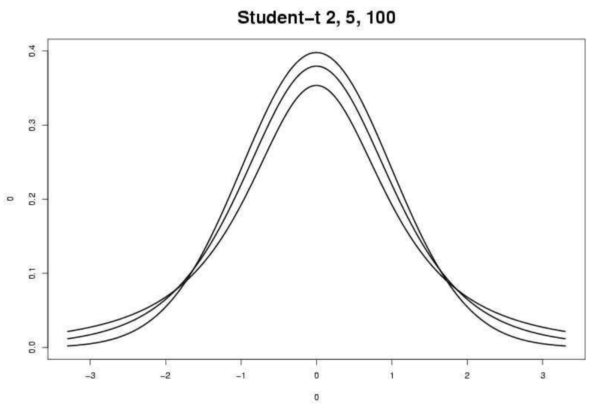
$$T = \frac{Z}{\sqrt{V/\nu}}$$

gitt ved sannsynlighetstettheten

$$h(t) = \frac{\Gamma[(\nu+1)/2]}{\Gamma(\nu/2)\sqrt{\pi\nu}}(1 + \frac{t^2}{\nu})^{-(\nu+1)/2}$$

for $-\infty < t < \infty$. Denne fordelingen har navnet (Student) $t-$fordelingen med ν frihetsgrader.

t-fordelingen



t-fordeling (forts.)

COR: La $X_1, X_2, ..., X_n$ være uavhengige stokastiske variabler som alle er normalfordelte med samme forventning μ og samme standardavvik σ. La

$$\bar{X} = \sum_{i=1}^{n} \frac{X_i}{n} \qquad \text{og} \qquad S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

Da er den stokastiske variablen

$$T = \frac{\bar{X} - \mu}{S/\sqrt{n}}$$

t-fordelt med $\nu = n - 1$ frihetsgrader.

Konfidensintervaller, litt oppsummering

- Hva vil vi estimere?
  - μ
  - $\mu_1 - \mu_2$
- Hva vet vi om fordeling(ene) til det/de tilfeldige utvalget/utvalgene?
  - σ kjent eller ukjent
  - σ_1 og σ_2 kjente
  - σ_1 og σ_2 ukjente, $\sigma_1 = \sigma_2$ eller $\sigma_1 \neq \sigma_2$
- Sett opp observatoren basert på kunnskapen om fordelingen(e)
  - $Z = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}}$ (estimere μ, σ kjent)
  - $T_{n-1} = \frac{\bar{X}-\mu}{S/\sqrt{n}}$ (estimere μ, σ ukjent)
  - $Z = \frac{\Delta\bar{X}-(\mu_1-\mu_2)}{\sigma\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$ (estimere $\mu_1 - \mu_2$, σ kjent)
  - $T_{n_1+n_2-2} = \frac{\Delta\bar{X}-(\mu_1-\mu_2)}{S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$ (estimere $\mu_1 - \mu_2$, σ_1 og σ_2 ukjent, men $\sigma_1 = \sigma_2$)
  - $T'_\nu = \frac{\Delta\bar{X}-(\mu_1-\mu_2)}{\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}}, \nu = g(n_1, s_1, n_2, s_2)$ (som over, men $\sigma_1 \neq \sigma_2$)

Konfidensintervaller, litt oppsummering til

- Konstruer konfidensintervallet ved hjelp av sannsynlighetsfordelingen til observatoren
  - $P(-z_{\alpha/2} < Z < z_{\alpha/2}) = 1 - \alpha$
  - $P(-t_{\alpha/2,\nu} < T_\nu < t_{\alpha/2,\nu}) = 1 - \alpha$
  - $P(-t_{\alpha/2,\nu} < T'_\nu < t_{\alpha/2,\nu}) \approx 1 - \alpha$
- Finn øvre og nedre grense ved å sette inn utrykket for observatoren, og omorganisere slik at det du er interessert i å finne et intervall for står alene i midten
  - $P(\text{nedre grense} < \mu \text{ eller } \mu_1 - \mu_2 < \text{ovre grense}) = 1 - \alpha$