Typiske spørsmål kap. 9 og kap 10.

- Kva er gjennomsnittshøgda for kvinnelege NTNU-studentar. Estimering
- Er eg høgare enn gjennomsnittleg kvinnelege NTNU-studentar. Hypotesetest
- Kva er forskjellen i kvalitet på laks for lagringsmetode 1 og lagringsmetode 2. Estimering
- Er kvaliteten på laks ved lagringsmetode 2 betre enn ved lagringsmetode 1. Hypotesetest

- Antar fordeling frå kjennskap til fenomenet; $X_i \sim f(x;\theta)$.
- Har data $x_1, x_2, \ldots, x_n$
- *Estimerer/ berekner* θ frå data v.h.a. ein *estimator* $\hat{\theta}$.

- Kva er ein god estimator (forventningsrett og minst mogeleg varians).
- Korleis finne ein estimalor (SME)
- Korleis kvantifisere usikkerheita i estimat (i dag: Konfidensintervall (i dag og neste veke))

Eksempel NTNU kvinner, case 1 og 2

Høgde kvinnelege NTNU-studentar

DATA

- CASE 1
  - $n = 149$
  - $\bar{x} = 169.5$, kjent varians $\sigma^2 = 6.5^2$
- CASE 2
  - n=5
  - $\bar{x} = 169.5$, kjent varians $\sigma^2 = 6.5^2$

ANTAR

- $X_i \overset{u.i.f.}{\sim} N(\mu, \sigma^2)$ for $i = 1, 2, \ldots, n$

ESTIMATOR

- $\hat{\mu} = \bar{X} \sim N(\mu, \sigma^2/n)$
- CASE 1: $Var(\hat{\mu}) = 6.0^2/149 = 0.5^2$
- CASE 2: $Var(\hat{\mu}) = 6.0^2/5 = 2.9^2$

Viktige observatorar

Dersom X_i, $i = 1, 2, \ldots, n$ er uavhengig identisk fordelt med $E(X_i) = \mu$ og $Var(X_i) = \sigma^2$, så

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \overset{n\to\infty}{\to} N(0,1)$$

Dersom $X_i \overset{u.i.f.}{\sim} N(\mu, \sigma^2)$ og $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$

$$T = \frac{\bar{X} - \mu}{S/\sqrt{n}} \sim t_{n-1}$$

Student-t fordelt med $\nu = n - 1$ fridomsgrader.

Dersom $X_i \overset{u.i.f.}{\sim} N(\mu, \sigma^2)$ og $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$

$$X^2 = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$$

Kji-kvadrat fordelt med $\nu = n - 1$ fridomsgrader.

Eks, NTNU kvinner prediksjonsintervall

- X_{ny}; en ny observasjon
- $X_{ny} \sim N(\mu, \sigma^2)$
- μ ukjend, men har estimat frå $n = 149$ data; $\mu^* = 169.5$. Varians antatt kjent $\sigma^2 = 6.52^2$.
- X_{ny} uavhengig av X_i-ane i $\bar{X}$.
- $\bar{X} \sim N(\mu, \sigma^2/n)$
- Ser på $X_{ny} - \bar{X} \sim N(0, \sigma^2 + \sigma^2/n) = N(0, \sigma^2_{diff}) = N(0, 6.04^2)$
- $Z = \dfrac{X_{ny} - \bar{X}}{\sigma_{diff}} \sim N(0, 1)$
- Prediksjonsintervall: $\alpha = 0.05$
- $[\bar{x} - z_{\alpha/2}\sigma_{diff}; \bar{x} - z_{\alpha/2}\sigma_{diff}] = [156.7, 182.3]$

Konfidensintervall: Dersom forsøket blir repetert mange gongar, vil sann parameter vere i KI i andel $1 - \alpha$ av forsøk.