I dag

- Konfidenintervall for andel
- Eksamen juni 2007, oppgåve 1 (Sannsynsmaksimeringsestimator (SME) og KI for å samanlikne to andelar)
- Om T og χ^2 fordeling
- Kahoot

Har eit nivå av trygghet for at sann parameter ligg i intervallet.

1. Finn estimator for parameteren vi ønsker å finne KI for.
2. Finn observator der parameter av interesse og estimator inngår:
   - $Z = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$,
     når SGT eller $X_i \overset{u.i.f.}{\sim} N(\mu, \sigma^2)$
   - $T = \frac{\bar{X}-\mu}{S/\sqrt{n}} \sim t_{n-1}$,
     når $X_i \overset{u.i.f.}{\sim} N(\mu, \sigma^2)$ og $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$
   - $X^2 = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$,
     når $X_i \overset{u.i.f.}{\sim} N(\mu, \sigma^2)$ og $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$
3. Har at (f.eks) $P(-z_{\alpha/2} < Z < z_{\alpha/2}) = 1 - \alpha$.
4. Løyser ut for parameter, (f.eks. μ); $P(\hat{\mu}_L < \mu < \hat{\mu}_U) = 1 - \alpha$.
5. Konfidensintervall; sett inn for data: $[\mu_L^*, \mu_U^*]$

Eksempel Defektsannsyn

Du har fått sommarjobb på vindusfabrikken. Dei tar i bruk ny produksjons teknikk, og ønsker å finne sannsynet p for defekt vindu.

- $m_1 = 40$ vindu produsert i første skiftet.
- $m_2 = 60$ vindu i andre skift.
- $u = 5$ defekte i første skiftet.
- $v = 15$ defekte i andre skiftet.

Estimator $\hat{p} = \frac{U+V}{m_1+m_2} = \frac{X}{n}$
$X \sim Bin(n, p)$

Sentralgrenseteoremet, teorem 8.2

La X_i, $i = 1, 2, \ldots, n$ vere uavhengige identisk fordelte (u.i.f.) stokastiske vaiable 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

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

når $n \to \infty$.
Dette tilsvarer

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

når $n \to \infty$.

PS: Gjeld uansett fordeling for X_i

Modell:

- Spm.lagar A: Tilfeldig deltakar klarar < 5 med sanns. q_1.
- Spm.lagar B: Tilfeldig deltakar klarar < 5 med sanns. q_2.

Data:

- Spm.lagar A: n_1 deltakarar, Z_1 klarte < 5
- Spm.lagar B: n_2 deltakarar, Z_2 klarte < 5

Estimatorar:

- $\hat{q}_1 = \frac{Z_1}{n_1}$
- $\hat{q}_2 = \frac{Z_2}{n_2}$

Oppgåve 1c):

- Utlei eit tilnærma 95% konfidensintervall for $q_1 - q_2$ basert på normaltilnærming til binomisk fordeling.
- Beregn intervallet når $n_1 = n_2 = 64$, og $z_1 = 34$ og $z_2 = 18$.
- Gjer KI grunnlag for å seie at oppgåvene frå A og B har ulik vanskelighetsgrad? Begrunn svaret.

REP: Geometrisk fordeling

Fordelinga til antall terningkast inntil første sekesar.

- Uavhengige forsøk med suksess / ikkje-suksess.
- Likt suksess-sannsyn i alle forsøka.
- X: antall forsøk t.o.m. første suksess.
- X er då geometrisk fordelt

REP: Geometrisk fordeling, pengespelet

- Uavhengige forsøk med suksess / ikkje-suksess.
  Svarar feil / svarar rett
- Likt suksess-sannsyn i alle forsøka.
  Likt sannsyn p for å svare feil.
- X: antall forsøk t.o.m. første suksess.
  t.o.m. første feile svar.
- X er då geometrisk fordelt

Finn den verdien for parameteren θ (pengespelet $\theta = p$) som gjev høgast sannsyn for å observere dei dataene vi har observert.

OPPSKRIFT

- Finn likelihoodfunksjonen
  $L(\theta; x_1, x_2, \ldots, x_n) = f(x_1, x_2, \ldots, x_n; \theta) \overset{uavh}{=} \prod_{i=1}^{n} f(x_i; \theta)$
- Finn toppunkt av likelihoodfunksjonen:
  - Tar ln av L; $l(\theta; x_1, x_2, \ldots, x_n) = \ln(L(\theta; x_1, x_2, \ldots, x_n))$. *Reknetriks som nesten alltid blir brukt. L og l har same toppunkt.*
  - Deriverer og set lik 0; $\dfrac{\partial}{\partial\theta} l(\theta; x_1, x_2, \ldots, x_n) = 0$
  - Løyser ut for θ.

REP, Bernoulli-prosess, binomisk fordeling

Bernoulli-prosess

- n uavhengige forsøk.
- Kvart forsøk resulterer i suksess eller ikkje-suksess.
- Sannsynet p for suksess er likt i alle forsøk.

La X vere antall suksess. Då er $X \sim bin(n, p)$

Bernoulli-prosess

- n uavhengige forsøk.
  n_1 uavhengige deltakarar
- Kvart forsøk resulterer i suksess eller ikkje-suksess.
  Kvar deltakar klarer enten ferre enn 5 oppgåver (C) eller 5 eller fleire oppgåver (C').
- Sannsynet p for suksess er likt i alle forsøk.
  $P(C) = q_1$ i alle forsøka

La X vere antall suksess. Då er $X \sim bin(n, p)$
Lar Z_1 vere antall deltakarar med ferre enn 5 rette. Då er $Z_1 \sim bin(n_1, q_1)$