Hypotesetesting

Torstein Fjeldstad
Institutt for matematiske fag, NTNU
06.11.2018

I dag

- Repetisjon
- p-verdi
- Val av tal på observasjonar n

Repetisjon

Hypotesetesting

	H_0 riktig	H_1 riktig
Forkast H_0	Type I-feil	Ok
Ikkje forkast H_0	Ok	Type II-feil

$$P(\text{Type I-feil}) = P(\text{Forkast } H_0 \text{ når } H_0 \text{ er riktig}) \leq \alpha$$

$$\beta = P(\text{Type II-feil}) = P(\text{Ikkje forkast } H_0 \text{ når } H_1 \text{ er riktig})$$

Korleis velge H_0 og H_1?

Ingen symmetri mellom konklusjonane, derfor ingen symmetri mellom H_0 og H_1.
Velg som H_1 det du ynskjer å sjekke/teste.
Velg som H_0 det "motsette" (typisk dagens situasjon).

Teststyrke/styrkefunksjon

Definisjon

Styrken til ein test er sannsynet for å forkaste H_0 gitt at ein spesifikk alternativ hypotese er sann

- teststyrken er $1 - \beta$
- teststyrken er ein funksjon av θ_0

Generell framgangsmåte

Situasjon: $X_1, X_2, \ldots, X_n$ tilfeldig utval med $X_i \sim f(x_i; \theta)$.

1. Ynskjer å teste:
   a) $H_0 : \theta = \theta_0$ mot $H_1 : \theta > \theta_0$
   b) $H_0 : \theta = \theta_0$ mot $H_1 : \theta < \theta_0$
   c) $H_0 : \theta = \theta_0$ mot $H_1 : \theta \neq \theta_0$
2. Finn ein estimator for θ; $\widehat{\theta}$
3. La $Z = h(\widehat{\theta}, \theta_0)$, der $h(\cdot, \cdot)$ er ein funksjon s.a. Z har ei kjend fordeling under H_0
4. Bestem eit forkastningskriterium (antar Z stor når $\widehat{\theta}$ stor)
   a) Forkast H_0 dersom $Z > k$
   b) Forkast H_0 dersom $Z < k$
   c) Forkast H_0 dersom $Z < k_l$ eller $Z > k_u$

   der k bestemmes frå kravet

$$P(\text{Forkast } \mathrm{H}_0 \text{ når } \mathrm{H}_0 \text{ er riktig}) \leq \alpha$$

5. Sett inn tal og konkluder

p-verdi

Eksempel

Situasjon

Ein produsent av fiskesnøre har utvikla eit nytt snøre produsenten påstår at i gjennomsnitt skal tole fisk på 10 kg med standardavvik på 1 kg.

Eksempel

Situasjon

Ein produsent av fiskesnøre har utvikla eit nytt snøre produsenten påstår at i gjennomsnitt skal tole fisk på 10 kg med standardavvik på 1 kg.

Test

$$H_0 : \mu = 10 \qquad \text{mot} \qquad H_1 : \mu \neq 10.$$

Eksempel

Situasjon

Ein produsent av fiskesnøre har utvikla eit nytt snøre produsenten påstår at i gjennomsnitt skal tole fisk på 10 kg med standardavvik på 1 kg.

Test

$$H_0 : \mu = 10 \qquad \text{mot} \qquad H_1 : \mu \neq 10.$$

Produsenten har testa 35 fiskesnører som i gjennomsnitt tolte ein fisk på 10.3 kg.

Eksempel

Situasjon

Ein produsent av fiskesnøre har utvikla eit nytt snøre produsenten påstår at i gjennomsnitt skal tole fisk på 10 kg med standardavvik på 1 kg.

Test

$$H_0 : \mu = 10 \qquad \text{mot} \qquad H_1 : \mu \neq 10.$$

Produsenten har testa 35 fiskesnører som i gjennomsnitt tolte ein fisk på 10.3 kg. Utfør testen ved signifikansnivå $\alpha = 0.05$.

p-verdi

Definisjon

Ein p-verdi er det lågaste signifikansnivået α slik at observert verdi for observatoren gjev at me skal forkaste H_0. Det vil seie, forkast H_0 dersom p-verdien er *lågare* enn α.

Val av tal på observasjonar n

Val av tal på observasjonar *n* I

Vil undersøkje om ein ny medisin *B* er betre enn eksisterande medisin *A*

- Eksisterande medisin *A* verker på 70% av pasientane
- Ny medisin *B* påstås å verke på meir enn 70% av pasientane

Korleis kan me konkludere om *B* er betre enn *A*?
Kor mange pasienter må me gi medisin *B* for å kunne vere "sikre" på konklusjonen vår?

Val av tal på observasjonar n II

Hypotesetest:

$$H_0 : p = p_0 = 0.7 \qquad \text{mot} \qquad H_1 : p > 0.7$$

Hugs:

$$\frac{X - np}{\sqrt{np(1-p)}} \approx n(z; 0, 1) \quad \text{alltid}$$

$$Z = \frac{X - np_0}{\sqrt{np_0(1-p_0)}} \approx n(z; 0, 1) \quad \text{når } H_0 \text{ er riktig}$$

Val av tal på observasjonar n III

Forkast H_0 dersom

$$Z = \frac{X - np_0}{\sqrt{np_0(1 - p_0)}} > z_\alpha$$

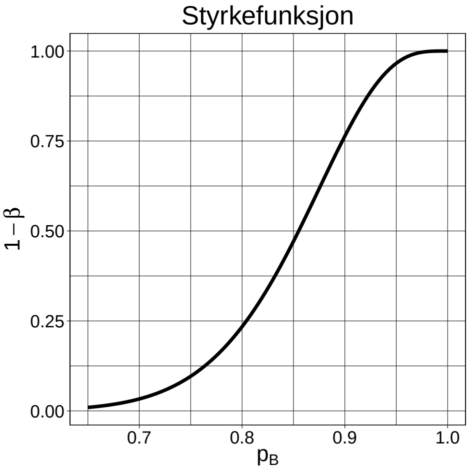
Velg α slik at

$$P(\text{Type I-feil}) = P(\text{forkast } H_0 \text{ når } H_0 \text{ riktig}) \leq \alpha$$

er liten. Itillegg ynskjer me å velge n slik at me kontrollerer

$$P(\text{Type II-feil}) = P(\text{ikkje forkast } H_0 \text{ når } H_1 \text{ riktig}) \leq \beta.$$

Styrkefunksjon $n = 25$



Torsdag

- Lineær regresjon