STK1110 Høsten 2021

Likelihood-ratio-tester

Tilsvarer Avsnitt 9.5 (s. 475-477)

Ingrid Hobæk Haff
Matematisk institutt
Universitetet i Oslo

Likelihood-ratio-tester

- Anta at $X_1, \ldots, X_n$ er stokastiske variabler med simultantetthet/-punktsannsynlighet $f(x_1, \ldots, x_n; \theta)$, der θ kan være en vektor, dvs. $\theta = (\theta_1, \ldots, \theta_k)$.
- Parameterrommet Ω er mengden av mulige verdier for θ, dvs. $\theta \in \Omega$.
- La $\Omega_0 \subset \Omega$ og $\Omega_a = \Omega \setminus \Omega_0$.
- Vi vil teste hypotesen $H_0 : \theta \in \Omega_0$ mot alternativet $H_a : \theta \in \Omega_a$.

Likelihood-ratio-tester (forts.)

- La $x_1, \ldots, x_n$ være observerte verdier av $X_1, \ldots, X_n$.
- Da er likelihood-funksjonen gitt ved $L(\theta) = f(x_1, \ldots, x_n; \theta)$.
- La nå $\hat{\theta}$ være MLE for θ.
- Videre, la $\hat{\theta}_0$ være MLE for θ under H_0, altså den verdien av θ som maksimerer $L(\theta)$ under begrensningen $\theta \in \Omega_0$.
- Da er **likelihood-ratioen** (sannsynlighetskvoten) gitt ved:

$$LR = \frac{\max_{\theta \in \Omega_0} L(\theta)}{\max_{\theta \in \Omega} L(\theta)} = \frac{L(\hat{\theta}_0)}{L(\hat{\theta})}.$$

Likelihood-ratio-tester (forts.)

- Da $\Omega_0 \subset \Omega$, må nødvendigvis $L(\hat{\theta}) \geq L(\hat{\theta}_0)$, slik at $LR \leq 1$.
- Dersom LR er liten, betyr det at dataene samsvarer mye mer med H_a enn med H_0.
- Dermed forkaster vi H_0 dersom $LR \leq k$, der k velges slik at vi får ønsket signifikansnivå α.

Eksempel

$X_1, \ldots, X_n \overset{uif}{\sim} N(\mu, \sigma^2)$, der σ er kjent. Vi vil teste $H_0 : \mu = \mu_0$ mot $H_a : \mu \neq \mu_0$.

Eksempel

$X_1, \ldots, X_n \overset{uif}{\sim} N(\mu, \sigma^2)$, der σ er ukjent. Vi vil teste $H_0 : \mu = \mu_0$ mot $H_a : \mu \neq \mu_0$.

Likelihood-ratio-tester for store utvalg

- Anta at $X_1, \ldots, X_n$ er uif med tetthet/punktsannsynlighet $f(x; \theta)$, som avhenger av én parameter θ.
- Vi vil teste $H_0 : \theta = \theta_0$ mot $H_a : \theta \neq \theta_0$.
- Hvis n er stor, er $-2\log(LR) \overset{tiln.}{\sim} \chi^2_1$ når H_0 er sann.
- Vi får dermed en test med tilnærmet signifikansnivå α hvis vi forkaster H_0 så sant $-2\log(LR) \geq \chi^2_{\alpha,1}$.

Eksempel

$X \sim Bin(n, p)$. Vi vil teste $H_0 : p = p_0$ mot $H_a : p \neq p_0$.

Likelihood-ratio-tester for store utvalg (forts.)

- Anta at $X_1, \ldots, X_n$ er uif med tetthet/punktsannsynlighet $f(x; \theta_1, \ldots, \theta_k)$, som avhenger av k parametere $\theta_1, \ldots, \theta_k$.
- Vi vil teste $H_0 : (\theta_1, \ldots, \theta_k) \subset \Omega_0$ mot $H_a : (\theta_1, \ldots, \theta_k) \subset \Omega_a$.
- Videre antar vi at hvis H_0 er sann, har vi $m < k$ parametere.
- Hvis n er stor, er da $-2\log(LR) \overset{tiln.}{\sim} \chi^2_{k-m}$ når H_0 er sann.

Eksempel

Nedbørsdataene.