Rimelighetsfunksjonen og sannsynlighetsmaksimeringsestimator

TMA4240/TMA4245 Statistikk

Håkon Tjelmeland
Institutt for matematiske fag
Norges teknisk-naturvitenskapelige universitet

Sannsynlighetsmaksimeringsprinsippet og rimelighetsfunksjonen

- Hvordan finne/bestemme en fornuftig estimator for en parameter?
  - ulike metoder/prinsipper finnes
  - tolkning av parameter kan benyttes, f.eks. $\widehat{\mu} = \overline{X}$
  - vi skal nå se på et annet prinsipp for konstruksjon av estimator

Definisjon (Sannsynlighetsmaksimeringsprinsippet)

Dersom man skal estimere verdien til en parameter θ ut fra observerte verdier $x_1, x_2, \ldots, x_n$, skal man velge som estimat den verdi av θ som gjør det mest sannsynlig å observere de verdiene man faktisk har observert.

Definisjon (Rimelighetsfunksjonen)

La $X_1, X_2, \ldots, X_n$ være stokastiske variabler med simultan sannsynlighetsfordeling $f(x_1, x_2, \ldots, x_n; \theta)$, der θ er en skalar parameter eller en vektor av parametre. Anta at en formel for $f(x_1, x_2, \ldots, x_n; \theta)$ er kjent, men at verdien til θ er ukjent. Hvis man har observerte verdier $x_1, x_2, \ldots, x_n$ for de stokastiske variablene $X_1, X_2, \ldots, X_n$ er rimelighetsfunksjonen gitt som

$$L(\theta; x_1, x_2, \ldots, x_n) = f(x_1, x_2, \ldots, x_n; \theta).$$

- Merk: Velger som estimat for θ den verdien av θ som maksimerer $L(\theta; x_1, x_2, \ldots, x_n)$
  - kan skrive estimatet som $\widehat{\theta} = u(x_1, x_2, \ldots, x_n)$

- Neste side: $X_1, X_2, \ldots, X_n$ tilfeldig utvalg fra eksponensialfordeling med parameter λ

Eksempel: Rimelighetsfunksjon og sannsynlighetsmaksimeringsestimat

- La $X_1, X_2, \ldots, X_n$ være et tilfeldig utvalg fra en eksponensialfordeling med parameter λ
- Simultan sannsynlighetstetthet blir

$$f(x_1, x_2, \ldots, x_n; \lambda) = \prod_{i=1}^{n} f(x_i; \lambda) = \prod_{i=1}^{n} \left[\lambda e^{-\lambda x_i}\right] = \lambda^n \exp\left\{-\lambda \sum_{i=1}^{n} x_i\right\}; x_1, x_2, \ldots, x_n \geq 0$$

- La $x_1, x_2, \ldots, x_n$ være observerte verdier
- Rimelighetsfunksjonen blir

$$L(\lambda; x_1, x_2, \ldots, x_n) = \lambda^n \exp\left\{-\lambda \sum_{i=1}^{n} x_i\right\}, \lambda > 0$$

- Finner maksimumspunkt ved å derivere og sette den deriverte lik null

$$\begin{aligned} L'(\lambda; x_1, x_2, \ldots, x_n) &= n\lambda^{n-1} \exp\left\{-\lambda \sum_{i=1}^{n} x_i\right\} + \lambda^n \exp\left\{-\lambda \sum_{i=1}^{n} x_i\right\} \cdot \left(-\sum_{i=1}^{n} x_i\right) \\ &= \lambda^{n-1} \exp\left\{-\lambda \sum_{i=1}^{n} x_i\right\} \cdot \left(n - \lambda \sum_{i=1}^{n} x_i\right) \end{aligned}$$

$$L'(\lambda; x_1, x_2, \ldots, x_n) = 0 \Leftrightarrow n - \lambda \sum_{i=1}^{n} x_i = 0 \Leftrightarrow \lambda = \frac{n}{\sum_{i=1}^{n} x_i}$$

- Sannsynlighetsmaksimeringsestimatet er altså

$$\widehat{\lambda} = u(x_1, x_2, \ldots, x_n) = \frac{n}{\sum_{i=1}^{n} x_i}$$

Sannsynlighetsmaksimeringsestimator

Prinsipp: Velg som estimat den verdi som gjør det mest sannsynlig å observere det man har observert

Rimelighetsfunksjonen: $L(\theta; x_1, x_2, \dots, x_n)$

Estimat: $\widehat{\theta} = u(x_1, x_2, \dots, x_n)$

Definisjon (Sannsynlighetsmaksimeringsestimator)

La $u(x_1, x_2, \dots, x_n)$ betegne estimatet man får for θ ved å benytte sannsynlighetsmaksimeringsprinsippet. Sannsynlighetsmaksimeringsestimatoren (SME) for θ er da

$$\widehat{\theta} = u(X_1, X_2, \dots, X_n).$$

- ⋆ I eksempel på forrige side: $\widehat{\lambda} = u(x_1, x_2, \dots, x_n) = \frac{n}{\sum_{i=1}^{n} x_i}$
  - SME blir dermed: $\widehat{\lambda} = u(X_1, X_2, \dots, X_n) = \frac{n}{\sum_{i=1}^{n} X_i}$

- ⋆ Merk:
  - for å finne maksimumspunkt for $L(\theta; x_1, x_2, \dots, x_n)$ må vi derivere denne
  - $L(\theta; x_1, x_2, \dots, x_n)$ er vanligvis et produkt av flere faktorer
  - å derivere et produkt av flere faktorer er «tungvint»
  - mye enklere å derivere en sum!

Log-rimelighetsfunksjonen

Definisjon (Log-rimelighetsfunksjonen)

La $L(\theta; x_1, x_2, \ldots, x_n)$ *være en rimelighetsfunksjon. Log-rimelighetsfunksjonen er da*

$$\ell(\theta; x_1, x_2, \ldots, x_n) = \ln L(\theta; x_1, x_2, \ldots, x_n)$$

for de verdier av θ *hvor* $L(\theta; x_1, x_2, \ldots, x_n) > 0$*.*

- Merk: ln er en strengt voksende funksjon
  - $L(\theta; x_1, x_2, \ldots, x_n)$ og $\ell(\theta; x_1, x_2, \ldots, x_n)$ har maksimum for samme verdi av θ
- Formulert matematisk:

$$\begin{aligned}\ell'(\theta; x_1, x_2, \ldots, x_n) &= \frac{d}{d\theta} \ln L(\theta; x_1, x_2, \ldots, x_n) \\ &= \frac{1}{L(\theta; x_1, x_2, \ldots, x_n)} \cdot L'(\theta; x_1, x_2, \ldots, x_n)\end{aligned}$$

$$\ell'(\theta; x_1, x_2, \ldots, x_n) = 0 \quad \Leftrightarrow \quad L'(\theta; x_1, x_2, \ldots, x_n) = 0$$

Eksempel: Log-rimelighetsfunksjon og SME

- La $X_1, X_2, \ldots, X_n$ være et tilfeldig utvalg fra en eksponensialfordeling med parameter λ
- Simultan sannsynlighetstetthet blir

$$f(x_1, x_2, \ldots, x_n; \lambda) = \prod_{i=1}^{n} f(x_i; \lambda) = \prod_{i=1}^{n} \left[\lambda e^{-\lambda x_i}\right] = \lambda^n \exp\left\{-\lambda \sum_{i=1}^{n} x_i\right\}; x_1, x_2, \ldots, x_n \geq 0$$

- La $x_1, x_2, \ldots, x_n$ være observerte verdier
- Rimelighetsfunksjonen blir

$$L(\lambda; x_1, x_2, \ldots, x_n) = \lambda^n \exp\left\{-\lambda \sum_{i=1}^{n} x_i\right\}, \lambda > 0$$

- Log-rimelighetsfunksjonen blir

$$\ell(\lambda; x_1, x_2, \ldots, x_n) = \ln L(\lambda; x_1, x_2, \ldots, x_n) = n \ln \lambda - \lambda \sum_{i=1}^{n} x_i$$

- Finner maksimumspunkt ved å derivere og sette den deriverte lik null

$$\ell'(\lambda; x_1, x_2, \ldots, x_n) = n \cdot \frac{1}{\lambda} - \sum_{i=1}^{n} x_i$$

$$\ell'(\lambda; x_1, x_2, \ldots, x_n) = 0 \Leftrightarrow \lambda = \frac{n}{\sum_{i=1}^{n} x_i}$$

- Sannsynlighetsmaksimeringsestimatet er altså

$$\widehat{\lambda} = u(x_1, x_2, \ldots, x_n) = \frac{n}{\sum_{i=1}^{n} x_i}$$

- Og SME blir

$$\widehat{\lambda} = u(X_1, X_2, \ldots, X_n) = \frac{n}{\sum_{i=1}^{n} X_i}$$

Oppsummering

- Har definert
  - sannsynlighetsmaksimeringsprinsippet
  - rimelighetsfunksjonen, $L(\theta; x_1, x_2, \ldots, x_n) = f(x_1, x_2, \ldots, x_n; \theta)$
  - sannsynlighetsmaksimeringsestimatoren (SME)
  - log-rimelighetsfunksjonen, $\ell(\theta; x_1, x_2, \ldots, x_n) = \ln L(\theta; x_1, x_2, \ldots, x_n)$

- Regneprosedyre for å utlede SME
  - finne simultanfordeling for aktuelle stokastiske variabler, $f(x_1, x_2, \ldots, x_n; \theta)$
  - bestemme $L(\theta; x_1, x_2, \ldots, x_n)$
  - bestemme $\ell(\theta; x_1, x_2, \ldots, x_n)$
  - finne hvilken verdi av θ som maksimerer $\ell(\theta; x_1, x_2, \ldots, x_n)$
  - formulere SME ved å erstatte observerte verdier $x_1, x_2, \ldots, x_n$ med stokastiske variabler $X_1, X_2, \ldots, X_n$

- Merk: $\ell(\theta; x_1, x_2, \ldots, x_n)$ kan ha sitt maksimum i et
  - kritisk punkt, dvs en verdi av θ der $\ell'(\theta; x_1, x_2, \ldots, x_n) = 0$
  - singulært punkt, dvs en verdi av θ der $\ell'(\theta; x_1, x_2, \ldots, x_n)$ ikke eksisterer
    - et knekkpunkt
    - diskontinuitetspunkt
  - i en verdi på randen av definisjonsområdet