Sannsynsmaksimeringsestimator

Torstein Fjeldstad
Institutt for matematiske fag, NTNU
15.10.2018

I dag

- Repetisjon
- Sannsynsmaksimeringsestimatoren

Repetisjon

Definisjon

Anta at me har eit tilfeldig utval $X_1, X_2, \ldots, X_n$ frå $f(x; \theta)$-populasjonen, der verdien til parameteren θ er ukjend. Ein estimator er ein observator som nyttast til å anslå verdien til θ.

Forventningsrett estimator

Ein observator $\hat{\theta}$ seies å verre ein forventningsrett (eng: unbiased) estimator for parameteren θ dersom

$$E(\hat{\theta}) = \theta.$$

Effisient estimator

Av fleire forventningsrette estimatorar for θ seier me at den med minst varians er den mest effisiente.
Me føretrekk den mest effisiente estimatoren.

Kva om $\bar{X}_n$ ikkje er ein estimator for θ?

Sannsynsmaksimeringsestimator (SME)

Problemstilling

- **Situasjon:** Anta $X_1, X_2, \ldots, X_n$ tilfelding utval frå $f(x;\theta)$-populasjonen.

Problemstilling

- **Situasjon:** Anta $X_1, X_2, \ldots, X_n$ tilfelding utval frå $f(x;\theta)$-populasjonen.
- **Mål:** Definere ei oppskrift på korleis me kan nytte observerte verdiar $x_1, x_2, \ldots, x_n$ til å finne ein estimator for θ, $\hat{\theta}$.

Problemstilling

- **Situasjon:** Anta $X_1, X_2, \ldots, X_n$ tilfelding utval frå $f(x;\theta)$-populasjonen.
- **Mål:** Definere ei oppskrift på korleis me kan nytte observerte verdiar $x_1, x_2, \ldots, x_n$ til å finne ein estimator for θ, $\hat{\theta}$.

Forrige veke snakka me om (punkt)estimatorar $\hat{\theta}$, men desse blei ofte foreslått "ut av det blå".

Problemstilling

- **Situasjon:** Anta $X_1, X_2, \ldots, X_n$ tilfelding utval frå $f(x;\theta)$-populasjonen.
- **Mål:** Definere ei oppskrift på korleis me kan nytte observerte verdiar $x_1, x_2, \ldots, x_n$ til å finne ein estimator for θ, $\hat{\theta}$.

Forrige veke snakka me om (punkt)estimatorar $\hat{\theta}$, men desse blei ofte foreslått "ut av det blå".

- I dag skal me sjå på ei oppskrift på korleis me kan foreslå ein estimator.

Eksempel

- Undersøk levetida X til elektroniske komponentar
- Anta at levetida er eksponentialfordelt ($E(X) = 1/\lambda$, λ ukjend)
- Test n komponenter

$$\begin{aligned} X_1 &\sim f(x_1;\lambda) \\ X_2 &\sim f(x_2;\lambda) \\ &\vdots \\ X_n &\sim f(x_n;\lambda) \end{aligned}$$

Sannsynsmaksimeringsestimator (SME)

Oppskrift for å finne SME

1. Finn simultanfordelinga til $X_1, X_2, \ldots, X_n$:

$$f(x_1, x_2, \ldots, x_n; \theta) = f(x_1; \theta) f(x_2; \theta) \ldots f(x_n; \theta) = \prod_{i=1}^{n} f(x_i; \theta)$$

Sannsynsmaksimeringsestimator (SME)

Oppskrift for å finne SME

1. Finn simultanfordelinga til $X_1, X_2, \ldots, X_n$:

$$f(x_1, x_2, \ldots, x_n; \theta) = f(x_1; \theta) f(x_2; \theta) \ldots f(x_n; \theta) = \prod_{i=1}^{n} f(x_i; \theta)$$

2. Definer rimelighetsfunksjonen (eng: likelihood function)

$$L(\theta) = \prod_{i=1}^{n} f(x_i; \theta)$$

Sannsynsmaksimeringsestimator (SME)

Oppskrift for å finne SME

1. Finn simultanfordelinga til $X_1, X_2, \ldots, X_n$:

$$f(x_1, x_2, \ldots, x_n; \theta) = f(x_1; \theta) f(x_2; \theta) \ldots f(x_n; \theta) = \prod_{i=1}^{n} f(x_i; \theta)$$

2. Definer rimelighetsfunksjonen (eng: likelihood function)

$$L(\theta) = \prod_{i=1}^{n} f(x_i; \theta)$$

3. For å forenkle rekninga ser me ofte på log-rimelighetsfunksjonen

$$l(\theta) = \ln L(\theta) = \sum_{i=1}^{n} \ln f(x_i; \theta)$$

Sannsynsmaksimeringsestimator (SME)

Oppskrift for å finne SME

1. Finn simultanfordelinga til $X_1, X_2, \ldots, X_n$:

$$f(x_1, x_2, \ldots, x_n; \theta) = f(x_1; \theta) f(x_2; \theta) \ldots f(x_n; \theta) = \prod_{i=1}^{n} f(x_i; \theta)$$

2. Definer rimelighetsfunksjonen (eng: likelihood function)

$$L(\theta) = \prod_{i=1}^{n} f(x_i; \theta)$$

3. For å forenkle rekninga ser me ofte på log-rimelighetsfunksjonen

$$l(\theta) = \ln L(\theta) = \sum_{i=1}^{n} \ln f(x_i; \theta)$$

4. Maksimer (log-)likelihood mhp θ:

$$\hat{\theta}_{SME} = \text{arg max}_\theta L(\theta) = \text{arg max}_\theta\ l(\theta)$$

Fredag

- Konfidensintervall