18.10.2016
TMA 4240

Egenskaper til estimatorer [9.3]

EKS: MapleTA øving 8, oppgave 5

$X_1, X_2, \dots, X_8$ u.i.f $\text{bin}(n=9, p)$

ukjent

$\hat{p}_1 = \frac{X_1}{9}$ og $\hat{p}_2 = \frac{\sum_{i=1}^{8} X_i}{8 \cdot 9}$

Hvilken estimator skal jeg velge å bruke?

1) Forventningsrett

$$E(\hat{p}_1) = E\left(\frac{X_1}{9}\right) = \frac{1}{9}\overbrace{E(X_1)}^{n \cdot p = 9}$$

$$= \frac{1}{9} 9 \cdot p = \underline{p}$$ forventningsrett

$$E(\hat{p}_2) = E\left(\frac{\sum_{i=1}^{8} X_i}{8 \cdot 9}\right) = \frac{1}{8 \cdot 9} E\left(\sum_{i=1}^{8} X_i\right)$$

$$= \frac{1}{8\cdot 9} \sum_{i=1}^{8} \underbrace{E(X_i)}_{\substack{n \cdot p \\ \| \\ 9}} = \frac{1}{8\cdot 9} \sum_{i=1}^{8} 9 \cdot p$$

$$= \frac{1}{8\cdot 9} \cdot 8 \cdot 9 \cdot p = \underline{\underline{p}}$$ forventningsrett

Hvilken estimator skal jeg velge videre?

→ Vi vil ha den med minst varians!

$$Var(\hat{p}_1) = Var\left(\frac{X_1}{9}\right) = \left(\frac{1}{9}\right)^2 \underbrace{Var(X_1)}_{\substack{n \cdot p \cdot (1-p) \\ \| \\ 9}}$$

$$= \left(\frac{1}{9}\right)^2 9 \cdot p(1-p) = \underline{\underline{\frac{p(1-p)}{9}}}$$

$$Var(\hat{p}_2) = Var\left(\frac{\sum^{8} X_i}{8 \cdot 9}\right) = \left(\frac{1}{8\cdot 9}\right)^2 Var\left(\sum_{i=1}^{8} X_i\right)$$

$$= \left(\frac{1}{8\cdot 9}\right)^2 \sum_{i=1}^{8} \underbrace{Var(X_i)}_{\substack{n \cdot p(1-p) \\ \| \\ 9}} = \left(\frac{1}{8\cdot 9}\right)^2 8 \cdot 9 \cdot p(1-p)$$

uavh

$$= \left(\frac{1}{8 \cdot 9}\right) \cdot p(1-p)$$

$\Rightarrow$ Velger $\hat{p}_2$ fordi den er forventningsrett og har minst varians av $\hat{p}_1$ og $\hat{p}_2$.

[9.14] Sannsynlighetsmaksimerings-estimator

= matematisk metode for å finne en estimator

Studer populasjon som beskrives av fordeling

$$f(x;\theta)$$

kjent

ukjent

normal
eksponential
Poisson

μ, σ^2
β
μ

og vil finne en estimator for θ basert på et tilfeldig utvalg

$$X_1, X_2, \dots, X_n \quad \text{u.i.f} \quad f(x;\theta)$$

uavh identisk fordelt

og vi observerer

$$X_1 = x_1, \; X_2 = x_2, \dots, \; X_n = x_n$$

Den simultane sanns. ford til det vi har observert

$$f(x_1, x_2, \dots, x_n; \theta) \overset{\text{uavhengig}}{=} f(x_1;\theta) \cdot f(x_2;\theta) \cdots f(x_n;\theta)$$

som vi kaller rimelighetsfunksjonen (likelihood)

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

Sannsynlighetsmaksimeringsestimatoren (SME) for θ, $\hat{\theta}$, finnes ved å maksimere rimelighetsfunksjonen.

SME i Poisson

Skader i norske skianlegg:
Poissonprosess med underliggende intensitet μ. $f(x; \mu) = \frac{\mu^x}{x!} e^{-\mu}$

$$x > 0$$

μ = forventet antall skader pr 1000 skidager

en person en dag i anlegget

Observerer $X_1, X_2, \ldots, X_n$ antall skader pr 1000 skidager med n tilfeller.

1) Ønsker estimator for μ.

2) $X_1, \ldots, X_n$ u.i.f Poisson

$$f(x;\mu) = \frac{\mu^x}{x!} e^{-\mu} \qquad x > 0$$

3) Rimelighetsfunksjonen.

$$L(x_1, x_2, \ldots, x_n; \mu) = f(x_1;\mu) \cdot f(x_2;\mu) \cdots f(x_n,\mu)$$

$$= \frac{\mu^{x_1}}{x_1!} e^{-\mu} \cdot \frac{\mu^{x_2}}{x_2!} e^{-\mu} \cdots \frac{\mu^{x_n}}{x_n!} e^{-\mu}$$

$$= \frac{\mu^{\sum_{i=1}^{n} x_i} \cdot e^{-n\mu}}{\prod_{i=1}^{n} (x_i!)}$$

4) Tar ln (naturlig logaritme) til L for å lette regningen. Ok?

Ja, fordi L og ln L har samme maksimum.

$$\ln L = \left(\sum_{i=1}^{n} x_i\right) \cdot \ln \mu + (-n\mu) - \ln\left(\prod_{i=1}^{n} (x_i!)\right)$$

$$\ln\left(\frac{a}{b}\right) = \ln a - \ln b \qquad \ln x^a = a \ln x$$

$$\ln(a \cdot b) = \ln a + \ln b$$

5) Deriver mhp μ

$$\frac{\partial \ln \mu}{\mu} = \frac{1}{\mu}$$

$$\frac{\partial \ln L}{\partial \mu} = \left(\sum_{i=1}^{n} x_i\right) \cdot \frac{1}{\mu} - n - 0$$

6) Setter $\frac{\partial \ln L}{\partial \mu} = 0$ og løser mhp μ.

$$\frac{\partial \ln L}{\partial \mu} = 0$$

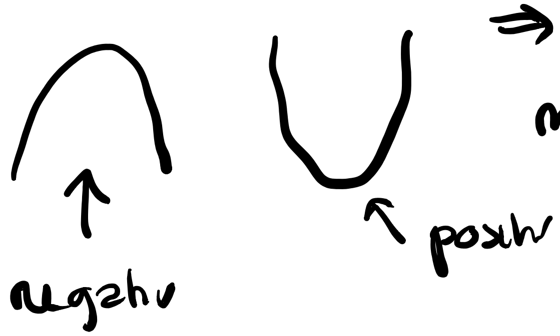
$$\frac{1}{\mu}\sum_{i=1}^{n} x_i - n = 0$$

$$\frac{1}{\mu}\sum x_i = n$$

$$\mu = \frac{\sum_{i=1}^{n} x_i}{n}$$

Dermed SME blir

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} X_i = \bar{X}$$



⇒ men er dette et maksimum?

Sjekk om andrederiverte er negativ.

(7) Sjekke at vi har funnet et maksimum og ikke et minimum

$$\left.\frac{\partial^2 \ln L}{\partial \mu^2}\right|_{\hat{\mu}} < 0 \quad ?$$

$$\frac{\partial^2 \ln L}{\partial \mu^2} = \frac{\partial}{\partial \mu}\left(-n + \frac{1}{\mu}\sum x_i\right)$$

$$= -\frac{1}{\mu^2} \cdot \sum x_i \qquad \left(\frac{\partial}{\partial \mu}\frac{1}{\mu} = -\frac{1}{\mu^2}\right)$$

Setter inn $\mu = \frac{1}{n}\sum x_i$

$$\frac{\partial^2 \ln L}{\partial \mu^2} = -\frac{1}{\left(\frac{1}{n}\sum x_i\right)^2}\sum x_i$$

$$= \frac{-n^2}{\sum x_i} < 0 \Rightarrow \text{OK}$$

maksimum

alle $x_i > 0$

Derved SME $\hat{\mu} = \overline{X}$

Er $\hat{\mu}$ en god estimator for μ?

$$E(\hat{\mu}) = E(\bar{X}) = \frac{1}{n}\sum_{i=1}^{n} \overbrace{E(X_i)}^{\mu} = \underline{\underline{\mu}}$$

$$Var(\hat{\mu}) = Var(\bar{X}) = \frac{1}{n^2}\sum_{i=1}^{n} \underbrace{Var(X_i)}_{\mu} = \underline{\underline{\frac{\mu}{n}}}$$

$X_i \sim Poisson(\mu)$

$E(X_i) = \mu, \quad Var(X_i) = \mu$

SME vil alltid være slik at

$$\lim_{n \to \infty} [E(\hat{\theta}) - \theta] = 0$$

$$\lim_{n \to \infty} [Var(\hat{\theta})] = 0$$

(konsistent).

Asymptotisk forventningsrett