Nøkkelbegreper 7.1 – 7.2

$X_1 \dots X_n$ u.i.f. $\sim f(\cdot\, ; \theta)$ θ ukjent parameter (vektor)

Estimator $\hat{\theta}(X_1 \dots X_n)$

(Estimat $\hat{\theta}(x_1 \dots x_n)$ $x_1 \dots x_n$ observerte verdier)

MSE – bruttovarians

$$MSE(\hat{\theta}) = E\left[(\hat{\theta} - \theta)^2\right]$$

$$= V(\hat{\theta}) + (\text{skjevhet})^2$$

der skjevhet $= E(\hat{\theta}) - \theta$
$=$ "bias"

Forventningsrett estimator har
$E(\hat{\theta}) = \theta$ ("unbiased")

$X_1 \dots X_n$ u.i.f. $E(X_i) = \mu$
$V(X_i) = \sigma^2$

Da er
$\hat{\mu} = \bar{X}$ forventningsrett estimator for μ

Geir kom hit $\longrightarrow$

$S^2 = \frac{1}{n-1} \Sigma (X_i - \bar{X})^2$ —— " —— σ^2

I dag: Konsistens – ønsket egenskap for estimatoren
Standard feil + bootstrap
7.2: Metoder for å finne estimatorer
Momentmetoden
Maximum Likelihood - metoden

Forklaring av hvorfor vi bruker

$S^2 = \frac{1}{n-1} \sum (X_i - \bar{X})^2$ som estimator for σ^2

Vet: $V(X_i) \underset{\text{def.}}{=} E\left[(X_i - E(X_i))^2\right]$

Forslag: $\hat{\sigma}^2 = \frac{1}{n} \sum (X_i - \bar{X})^2$

Vil sjekke om $E(\hat{\sigma}^2) = \sigma^2$

$$E\left[\sum (X_i - \bar{X})^2\right] = E\left[\sum (X_i^2 - 2X_i\bar{X} + \bar{X}^2)\right]$$

$$= E\left[\sum X_i^2 - 2n\bar{X}^2 + n\bar{X}^2\right]$$

$$= E\left[\sum X_i^2 - n\bar{X}^2\right]$$

$$= \sum \underline{E(X_i^2)} - n\,\underline{E(\bar{X}^2)}$$

$$= \sum \left(V(X_i) + E(X_i)^2\right) - n\left(V(\bar{X}) + E(\bar{X})^2\right)$$

$$= \sum (\sigma^2 + \mu^2) - n\left(\frac{\sigma^2}{n} + \mu^2\right)$$

$$= n\sigma^2 + \cancel{n\mu^2} - \sigma^2 - \cancel{n\mu^2} = (n-1)\sigma^2$$

$V(Y) = E(Y^2) - E(Y)^2$

$\rightarrow E(Y^2) = V(Y) + E(Y)^2$

$\bar{X} = \frac{1}{n}\sum X_i$

$E(\bar{X}) = \frac{1}{n}\sum E(X_i) = \frac{1}{n} n\mu = \mu$

$V(\bar{X}) = \frac{1}{n^2}\sum V(X_i) = \frac{1}{n^2} n\sigma^2 = \frac{\sigma^2}{n}$

Derfor $E(\hat{\sigma}^2) = \frac{1}{n} E\left(\sum (X_i - \bar{X})^2\right)$

$$= \frac{1}{n}(n-1)\sigma^2 = \frac{n-1}{n}\sigma^2$$

$\hat{\sigma}^2$ er IKKE forventningsrett!

Bruker i stedet estimatoren

$$S^2 = \frac{1}{n-1}\sum (X_i - \bar{X})^2 = \frac{n}{n-1}\hat{\sigma}^2$$

Har da at

$$\underline{E(S^2)} = E\left(\frac{n}{n-1}\hat{\sigma}^2\right) = \frac{n}{n-1}\boxed{\frac{n-1}{n}\sigma^2}$$

$$\underline{\underline{= \sigma^2}}$$

S^2 er forventningsrett!!

Estimering i uniform fordeling (kont.)

$X_1 \dots X_n$ u.i.f. Unif $[0, \theta]$ θ ukjent

$E(X_i) = \frac{\theta}{2}$
$V(X_i) = \frac{\theta^2}{12}$

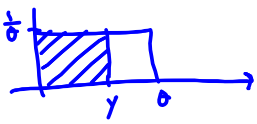
En mulig estimator for θ er

$$\theta^* = 2\overline{X}$$

Har da

$$\boxed{\begin{aligned} E(\theta^*) &= 2E(\overline{X}) = 2 \cdot \frac{\theta}{2} = \theta \quad \text{forv. rett!} \\ V(\theta^*) &= 4V(\overline{X}) = 4 \cdot \frac{\theta^2}{12n} = \frac{\theta^2}{3n} \end{aligned}}$$

Alternativ estimator tar utgangspunkt i $Y_n = \max_n X_i$



Kumulativ fordeling for Y_n:

$$F_{Y_n}(y) = P(Y_n \leq y) = P(\max X_i \leq y) = P(\text{alle } X_i \leq y)$$
$$= P(X_1 \leq y) \cdot P(X_2 \leq y) \cdots P(X_n \leq y)$$
$$= \frac{y}{\theta} \cdot \frac{y}{\theta} \cdots \frac{y}{\theta} = \left(\frac{y}{\theta}\right)^n$$

Tettheten blir

$$f_{Y_n}(y) = F_{Y_n}'(y) = \frac{n}{\theta}\left(\frac{y}{\theta}\right)^{n-1} \qquad 0 \leq y \leq \theta$$

Finner

$$E(Y_n) = \int_0^\theta y \, \frac{n}{\theta}\left(\frac{y}{\theta}\right)^{n-1} dy = \frac{n}{n+1}\theta \longleftarrow$$

$$E(Y_n^2) = \int_0^\theta y^2 \, \frac{n}{\theta}\left(\frac{y}{\theta}\right)^{n-1} dy = \frac{n}{n+2}\theta^2$$

$$V(Y_n) = E(Y_n^2) - E(Y_n)^2 = \frac{n\theta^2}{(n+1)^2(n+2)} \longleftarrow$$

$\hat{\theta} = \frac{n+1}{n} Y_n$ forslag, fordi da får vi

$$\boxed{\begin{aligned} E(\hat{\theta}) &= \frac{n+1}{n} E(Y_n) = \frac{n+1}{n} \frac{n}{n+1} \theta = \theta \quad \text{forv. rett!} \\ V(\hat{\theta}) &= \frac{(n+1)^2}{n^2} V(Y_n) = \frac{\theta^2}{n(n+2)} \end{aligned}}$$

Bootstrap – parametrisk

for å finne standardfeilen til en estimator $\hat{\theta}$

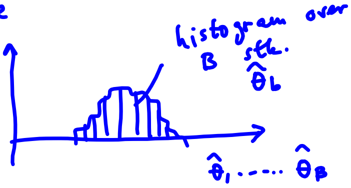
$X_1, \dots, X_n$ u.i.f. $\sim f(\cdot\,;\theta)$

Estimerer θ med estimator $\hat{\theta}$

Vil deretter estimere standardfeilen til $\hat{\theta}$ $(\sigma_{\hat{\theta}})$

1) Trekker B utvalg på størrelse n fra $f(\cdot\,;\hat{\theta})$

2) Estimerer B ulike $\hat{\theta}_b$, fra hvert av de simulerte utvalgene



3) Regner ut

$$s_{\hat{\theta}} = \sqrt{\frac{1}{B-1}\sum_{b=1}^{B}\left(\hat{\theta}_b - \bar{\theta}_b\right)^2}$$

der $\bar{\theta}_b = \frac{1}{B}\sum_{b=1}^{B}\hat{\theta}_b$

som estimert standardfeil for $\hat{\theta}$