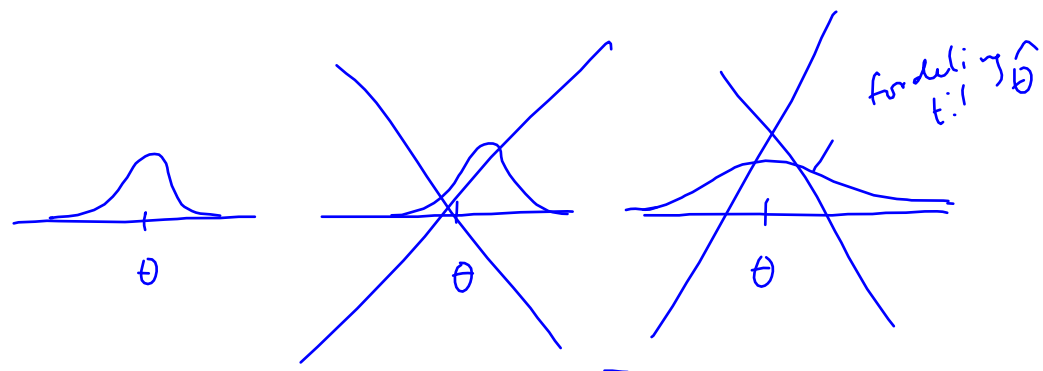
$\hat{\theta} = \hat{\theta}(X_1, \ldots, X_n)$ stok. var. med en fordeling

Bruttovarians

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




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

$E(aX+b) = aE(X)+b$

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

$$= \underbrace{V(\hat{\theta})} + \underbrace{(E(\hat{\theta}) - \theta)^2}$$

$$= \text{variansen} + (\text{skjevhet})^2$$

Skjevhet kalles bias

## Estimering : binomisk fordeling

$Y \sim bin(n, p)$ Har at $E(Y) = np$ $V(Y) = np(1-p)$

To estimatorer

$$\hat{p} = \frac{Y}{n} \qquad p^* = \frac{Y+2}{n+4}$$

Hvilken av disse er "best" ?

Vil se på $MSE(\hat{p}) = E\left[(\hat{p} - p)^2\right]$

og $MSE(p^*) = E\left[(p^* - p)^2\right]$

Vet at $MSE = \text{varians} + (\text{skjevhet})^2$

Ser først på $\hat{p} = \frac{Y}{n}$. Har at

$$V(\hat{p}) = V\left(\frac{Y}{n}\right) = \frac{1}{n^2} V(Y) = \frac{1}{n^2} np(1-p) = \frac{p(1-p)}{n}$$

$$\text{Skjevhet} = E(\hat{p}) - p = E\left(\frac{Y}{n}\right) - p = \frac{1}{n} E(Y) - p = \frac{1}{n} np - p = 0$$

$$\boxed{MSE(\hat{p}) = V(\hat{p}) = \frac{p(1-p)}{n}}$$

ingen skjevhet!

Ser så på $p^* = \frac{Y+2}{n+4} = \frac{1}{n+4} Y + \frac{2}{n+4}$

Har at

$$V(p^*) = \left(\frac{1}{n+4}\right)^2 V(Y) = \frac{np(1-p)}{(n+4)^2} = \frac{p(1-p)}{n+8+16/n}$$

$$\text{Skjevhet} : E(p^*) - p = \frac{1}{n+4} E(Y) + \frac{2}{n+4} - p$$

$$= \frac{np}{n+4} + \frac{2}{n+4} - p = \frac{2/n + 4p/n}{1 + 4/n}$$

$$\boxed{MSE(p^*) = \frac{p(1-p)}{n+8+16/n} + \left(\frac{2/n + 4p/n}{1 + 4/n}\right)^2}$$

Sammenlignes i plot

# Estimering av varians

$X_1, \ldots, X_n$ uavh., alle samme fordeling

$$E(X_i) = \mu \qquad V(X_i) = \sigma^2$$

Har at $\hat{\mu} = \overline{X}$ er en forventningsrett estimator for $\mu$

Hvordan estimere $\sigma^2$?

Hvis vi hadde kjent $\mu$, kunne vi ha estimert $\sigma^2$ med

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} (X_i - \mu)^2$$

Har da at

$$E(\hat{\sigma}^2) = E\left[ \frac{1}{n} \sum (X_i - \mu)^2 \right] = \frac{1}{n} \sum_{i=1}^{n} \underbrace{E\left[(X_i - \mu)^2\right]}_{V(X_i) = \sigma^2}$$

$$= \frac{1}{n} n \sigma^2 = \sigma^2$$

Denne er forv. rett!

Men vi kjenner jo ikke $\mu$! Også $\mu$ må estimeres!

Kan prøve med estimatoren

$$\hat{\hat{\sigma}}^2 = \frac{1}{n} \sum_{i=1}^{n} (X_i - \hat{\mu})^2 = \frac{1}{n} \sum (X_i - \overline{X})^2$$

Vil finne $E(\hat{\hat{\sigma}}^2)$ – er den fremdeles forv. rett?

$$E\left[\sum_{i=1}^{n}(X_i-\bar{X})^2\right] = E\left[\sum(X_i^2 - 2X_i\bar{X} + \bar{X}^2)\right]$$

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

$$= E\left[\sum X_i^2 - n\bar{X}^2\right] = \sum_{i=1}^{n} E(X_i^2) - nE(\bar{X}^2)$$

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

$$= \sum_{i=1}^{n}\left(\sigma^2 + \mu^2\right) - n\left(\frac{\sigma^2}{n} + \mu^2\right)$$

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

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

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

Dermed

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

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

ikke forv. rett!

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

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

Så $S^2$ er forventningsrett!