Lecture (18) — Properties of Estimators

Recall :-

To select the sample size (n)

we have $\mathcal{E}$ = error.

Margin of error $\Rightarrow$ P, $P_1 - P_2$, μ, $\mu_1 - \mu_2$

$\Downarrow$

It is The 2nd Part of the CI

$\downarrow$

after $\pm$

$\mathcal{E}$ = 2nd part of the CI.

Proof $\Rightarrow$ of why we assume $P = 0.5$

$$Var(\hat{P}) = \frac{pq}{n}$$ $\longrightarrow$ inverse relationship between n & $var(\hat{P})$

So to minimize (n) we need to maximize

the variance. $\hookrightarrow$ has a corresponding P.

* to get the P that maximize $Var(\hat{P})$

① $var'(\hat{P}) = 0$ $\Rightarrow$ (derivative, equate = 0)

$$var'(\hat{P}) = \frac{1-2P}{n} = 0 \quad \therefore P = \frac{1}{2}$$

Properties of Estimators :-

Recall: we said that we seek good properties of estimators which are ① unbiasness

② low variability

↳ now we call this with Efficiency

* Relative Efficiency : we use It to compare efficiency of estimators

two cases

All estimators are unbiased

* find the efficiency of $\hat{\theta}_1$ relative to $\hat{\theta}_2$

$$R\text{-}eff = \frac{var(\hat{\theta}_2)}{var(\hat{\theta}_1)} \begin{matrix} \rightarrow \text{what we end with} \\ \rightarrow \text{what we start with} \end{matrix}$$

at least one is biased

* find the efficiency of $\hat{\theta}_1$ relative to $\hat{\theta}_2$

$$R.eff = \frac{MSE(\hat{\theta}_2)}{MSE(\hat{\theta}_1)}$$

If R-eff > 1 ⟹ $\hat{\theta}_1$ is more eff. $(var(\hat{\theta}_2) > Var(\hat{\theta}_1))$

Reff < 1 ⟹ $\hat{\theta}_2$ is more eff. $(Var(\hat{\theta}_1) > Var(\hat{\theta}_2))$

Reff = ⟹ $\hat{\theta}_1$ and $\hat{\theta}_2$ are equally eff.

⇓

Similarly for MSE

Example ① $\Rightarrow$ in the slides.

$Y \sim uni(0, \theta)$

$$Reff = \frac{var(\hat{\theta}_2)}{var(\hat{\theta}_1)}$$

$$var(\hat{\theta}_1) = V(2\bar{Y}) = 4\ V(\bar{Y}) = 4\ \frac{var(Y)}{n}$$

$$= 4 \cdot \frac{(b-a)^2}{12} \cdot \frac{1}{n} = 4 \cdot \frac{\theta^2}{12} \cdot \frac{1}{n} = \frac{\theta^2}{3n}$$

$$var(\hat{\theta}_2) = var\left(\frac{n+1}{n} Y_{(n)}\right) = \left(\frac{n+1}{n}\right)^2 var(Y_{(n)})$$

$$Var(Y_{(n)}) = E(Y_{(n)}^2) - [E(Y_{(n)})]^2$$

$$\downarrow$$

$$\int_0^\theta y \cdot f_{(n)}(Y)\, dy$$

$$f_{(n)}(Y) = n\,(F(Y))^{n-1} f(y)$$

$$f(y) = \frac{1}{\theta}, \quad F(Y) = \frac{Y}{\theta}$$

$$f_{(n)}(Y) = n\left(\frac{Y}{\theta}\right)^{n-1} \cdot \frac{1}{\theta}$$

$$\int_0^\theta n\, \frac{Y^n}{\theta^{n-1}}\, \frac{1}{\theta}\, dy$$

$$E(Y_{(n)}^2) = \int_0^\theta n\, \frac{Y^{n+1}}{\theta^n}\, dy$$

we get $V(\hat{\theta}_2) = \frac{\theta^2}{n(n+2)}$ $\Longleftarrow$

$$R.eff = \frac{\frac{\theta^2}{n(n+2)}}{\frac{\theta^2}{3n}}$$

$$= \frac{3}{n+2}$$

If $n = 1 \Rightarrow$ both are equally eff.

$n > 1 \Rightarrow R.eff < 1$

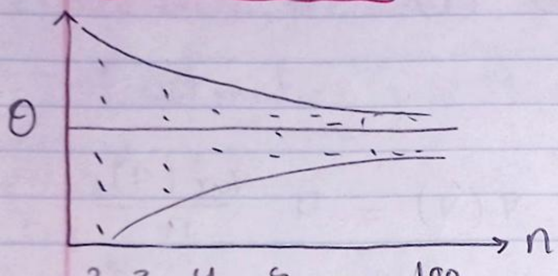
$\Rightarrow var(\hat{\theta}_2) < var(\hat{\theta}_1)$

$\rightarrow \hat{\theta}_2$ is more eff.

Lecture (19)

follow: Properties of Estimators - Consistency

* Consistency :-



when the sample size increases, n approaches the population size thus the estimator becomes closer to the true parameter

* If this happens we call $\hat{\theta}$ a consistent estimator for θ.

Def. the estimator $\hat{\theta}_n$ is said to be a consistent of θ Iff for any +ve number ε

$$\lim_{n\to\infty} P(|\hat{\theta}_n - \theta| \leq \varepsilon) = 1$$

or $$\lim_{n\to\infty} P(|\hat{\theta}_n - \theta| \geq \varepsilon) = 0$$

Theorem (1) : If $\lim_{n\to\infty} Var(\hat{\theta}) = 0$

then $\hat{\theta}$ is consistent estimator for θ.

Example ②

* to Check If an estimator is consistent

we first Check If It is biased or unbiased

$\Downarrow$ we Calculate lim for MSE

$\Downarrow$ we Calculate lim for Variance

$E(X) = \mu$, $var(X) = \sigma^2$

$\bar{X}$ is unbiased for μ

$\lim_{n\to\infty} Var(\bar{x}) = ?$

$var(\bar{x}) = \frac{Var(X)}{n} = \frac{\sigma^2}{n}$ $\rightarrow$ $\lim_{n\to\infty} Var(\bar{X}) = \lim_{n\to\infty} \frac{\sigma^2}{n} = \frac{\sigma^2}{\infty} = 0$

$\therefore$ $\bar{X}$ is consistent for μ.

Example ③ Prove :

S^2 is unbiased & consistent for σ^2.

① unbiased

$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{(n-1)}$

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

$\therefore$ S^2 is unbiased for σ^2.

② Consistent :-

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$$

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

$$\frac{(n-1)^2}{\sigma^4} Var(S^2) = 2(n-1)$$

$$Var(S^2) = \frac{2\cancel{(n-1)} \cdot \sigma^4}{(n-1)^{\cancel{2}}} = \frac{2\sigma^4}{n-1}$$

$$\lim_{n\to\infty} Var(S^2) = \lim_{n\to\infty} \frac{2\sigma^4}{n-1} = 0$$

$\therefore S^2$ is Consistent for σ^2.

---

* Sufficiency :-

The estimator doesn't miss any info from the Population.

① By definition method :-

If $f(x_1, x_2, x_3, \ldots, x_n \mid \hat{\theta}) = \dfrac{f(x_1, x_2, \ldots, x_n, \hat{\theta})}{g(\hat{\theta})}$

$= \dfrac{f(x_1, x_2, \ldots, x_n)}{g(\hat{\theta})}$ is free of the

Parameter θ, then $\hat{\theta}$ is sufficient.

Remarks :-

① we get the joint of $X_1, X_2, \dots, X_n$ given $\hat{\theta}$ to know if the result will be free of θ :-

Yes

this means the X's have no contribution

No

this means that the X's contribute some info & $\hat{\theta}$ is not sufficient to estimate θ

② Numerator → joint → $\prod_{i=1}^{n} f(x_i)$ (if they're indep)

denominator → Marginal pdf of $\hat{\theta}$

Example ④

$$\hat{\theta} = \frac{\sum x_i}{n} = \frac{X}{n}$$

$$f(x_1, x_2, x_3, \dots, x_n \mid \hat{\theta}) = \frac{f(x_1, x_2, \dots, x_n)}{g(\hat{\theta})}$$

Numerator → $f(x_1, x_2, \dots, x_n) = \prod_{i=1}^{n} f(x_i)$

$$= \theta^{x_1} (1-\theta)^{1-x_1} \cdot \theta^{x_2} (1-\theta)^{1-x_2} \dots$$

$$= \theta^{\sum x_i} (1-\theta)^{\sum 1 - x_i} = \theta^{\sum x_i} (1-\theta)^{n - \sum x_i}$$

$$= \theta^{X} (1-\theta)^{n-X} = \theta^{n\hat{\theta}} (1-\theta)^{n - n\hat{\theta}}$$

Denominator → $g(\hat{\theta})$ $\quad \hat{\theta} = \frac{\sum x_i}{n} = \frac{X}{n}$

i.e. $X = \hat{\theta} n$

So substitute X in $f(x)$ with $n\hat{\theta}$ to get $g(\hat{\theta})$.

② Factorization Method :-

$f(x_1, x_2, \dots, x_n) = g(\theta, \hat{\theta}) \cdot h(x_1, x_2, \dots, x_n)$

the likelyhood can also be written like This :-

$f(x_1, x_2, \dots, x_n \mid \theta)$ → note that This is θ not $\hat{\theta}$

$= f(x_1, x_2, \dots x_n ; \theta)$

$g(\theta, \hat{\theta})$ — Conditions :-

① function in θ and $\hat{\theta}$

② X's enter this Part only through $\hat{\theta}$

$h(x_1, x_2, \dots, x_n)$ — Conditions :-

① function in Xs only

② free of θ

③ Can be Const.

If this happens, we say that $\hat{\theta}$ is minimal sufficient estimator for θ.

* what If the domain is dependent on Parameter θ

in this Case we write $I_{(x_{(n)} < \theta)}$

or $I_{(x_{(1)} > \theta)}$

↓

the min or the max is the estimator

* Note : any one to one function in the minimal sufficient estimator is a sufficient estimator.

for example If $\hat{\theta} = \sum x_i$ is minimal Suff. est.

Then $\hat{\theta}_1 = \frac{\sum x_i}{n}$ is also a Suff. estimator (but not minimal)