Lezione 10 – Point Estimation

Distribution of a characteristic in a population.

Typically f and F are unknown or f is a special case for which we need specific parameters.

Θ can be the characteristic of a distribution like λ in Exp or μ in N, which is unknown, and that we are therefore trying to estimate.

If an estimator T of θ is unbiased it means:

$$E(T) = \Theta \quad \forall \theta$$

So on average the error is 0.

Definition of Bias (Distorsione)

↳ The bias of an estimator T of Θ is the difference

$$b_\theta(T) = b(T,\theta) = E(T) - \theta \lesseqgtr 0$$

If an estimator is unbiased for θ, then $b(T,\theta) = 0$

Definition: Mean Square Error (MSE)

The mean square error of an estimator T for θ is:

$$MSE_\theta(T) = E\left((T-\theta)^2\right)$$

The smaller the MSE, the better the estimator.

If T is an estimator for θ:

$$MSE_\theta(T) = Var_\theta(T) + (b_\theta(T))^2$$

↳ The MSE is what we use to measure the quality of an estimator.

Property 1:

Suppose we have, T_1 and T_2.

T_1 is better than T_2 if $MSE(T_1) \leq MSE(T_2)$ $\forall \theta$

Property 2:

If T is an estimator for θ:

$$MSE_\theta(T) = Var_\theta(T) + (b_\theta(T))^2$$

↳ Proof:

$$MSE_\theta(T) = E((T-\theta)^2) = E[(T - E_\theta(T) + E_\theta(T) - \theta)^2]$$

$$= E((T - E_\theta(T))^2 + (\underbrace{E_\theta(T) - \theta}_{\text{constant}})^2 - 2(T - E_\theta(T))(E_\theta(T) - \theta)]$$

$$= \underbrace{E[(T - E_\theta(T))^2]}_{\text{Variance of } T} + \underbrace{(E_\theta(T) - \theta)^2}_{\text{Bias}(T)^2} + 2(E_\theta(T) - \theta)\cancel{E(T - E_\theta(T))}^{0}$$

$$= Var_\theta(T) + b_\theta(T, \theta)^2$$

If T is an unbiased estimator of θ:

$$MSE_\theta(T) = Var_\theta(T)$$

Exercise

We measure thermal conductivity of ARMCO pure iron under certain conditions.

$X \rightarrow$ thermal conductivity $\sim f$

$n = 10$ measurements

$\underbrace{X_1, \dots, X_{10}}_{\text{variables, they are not the measure}} \overset{iid}{\sim} f \rightarrow$ unknown $\Rightarrow \mu$ unknown

Estimate of $\mu = E(X) =$ mean of the density f

$x_1 = 41{,}60 \quad x_3 = 42{,}34 \quad x_5 = 41{,}86 \quad x_7 = 41{,}72 \quad x_9 = 41{,}81$

$x_2 = 41{,}48 \quad x_4 = 41{,}95 \quad x_6 = 42{,}18 \quad x_8 = 42{,}26 \quad x_{10} = 42{,}04$

$\hookrightarrow$ Realization of the random variables we are measuring.

$x_1, x_2, \dots, x_{10} \rightarrow$ observed sample

We want an estimator for μ: $\hat{\mu}$

sample mean: $\bar{X}_{10} = \dfrac{X_1 + \dots + X_{10}}{10}$

It's a good estimator since it's unbiased for μ.

$$E(\bar{X}_{10}) = \mu$$

$$MSE(\bar{X}_n) = Var(\bar{X}_n) = \frac{Var(X_i)}{n} = \frac{\text{var of } f}{n}$$

$\hookrightarrow$ also computable

$\hat{\mu}$: estimate for $\mu \rightarrow$ using estimator using the observed sample.

$\hookrightarrow$ same letter $\quad \hookrightarrow$ realisation of an estimator

$$\bar{x}_{10} = \frac{x_1 + \dots + x_{10}}{10} = 41{,}924$$

We can estimate $Var(\bar{X}_n)$

$\sigma^2 = Var(X_i)$ = variance of f we saw this last time

$\hat{\sigma}^2$: estimator for σ^2 = $S_n^2 = \frac{1}{n-1}\sum_{j=1}^{n}(x_j - \bar{X}_n)^2$

sample variance

if it were to be $\frac{1}{n}$, it would not be unbiased.

unbiased for σ^2

$E[S_n^2] = \sigma^2 \quad \forall \sigma^2$

estimate for σ^2: $\hat{\sigma}^2 \to$ realisation of the sample variance

$$\hat{\sigma}^2 = \frac{1}{9}\sum_{j=1}^{10}(x_j - \bar{x}_{10})^2 \simeq 0{,}0807$$

Most sensible estimator of σ.

Estimator for $\sigma := \sqrt{Var X_i} = \sqrt{\sigma^2}$: $\sqrt{S_n^2}$

theoretical standard deviation of f

Generally not unbiased

Estimate for σ:

$$\hat{\sigma} = \sqrt{S_{10}^2} = \sqrt{0{,}0807} = 0{,}284$$

Example

Estimation of smokers of a proportion $p \in (0,1)$

Smokers in [18-25] year olds

$n = 40$

$X_1, \dots, X_n \overset{iid}{\sim} Be(p) = f \to$ since $X_i = \begin{cases} 1, & \text{if you smoke} \\ 0, & \text{otherwise} \end{cases}$

Mean of f is $p = E(X_i)$

Estimator for $p \to$ sample mean

$$\hat{p} = \frac{X_1 + \ldots + X_{40}}{n} = \frac{\text{\# of individual who have property}}{n}$$

Estimate for $p = \frac{7}{40} = 0{,}175 = 17{,}5\%$

Plug-in Estimator

Suppose we have $X_1, \ldots, X_n \overset{iid}{\sim} f$ unknown

Suppose we want to estimate θ which is a characteristic of f.

$\bar{X}_n \to$ estimator mean of $f = g(\theta) = E(X_i)$

How can we find an estimator for θ:

$$\widehat{g(\theta)} = \bar{X}_n \Rightarrow \text{solve } g(\hat{\theta}) = \bar{X}_n \text{ with respect to } \hat{\theta}$$

Example

Suppose $X_1, \ldots, X_n \overset{iid}{\sim} Exp(\theta)$ $\quad \theta > 0$

1) estimator for θ

$$\underbrace{E(X_i) = \frac{1}{\theta}}_{g(\theta)}$$

plug-in

$$\widehat{\left(\frac{1}{\theta}\right)} = \bar{X}_n \Rightarrow \left(\frac{1}{\hat{\theta}}\right) = g(\hat{\theta}) = \bar{X}_n \Rightarrow \frac{1}{\hat{\theta}} = \bar{X}_n$$

$$\Rightarrow \hat{\theta} = \frac{1}{\bar{X}_n}$$

This method allows us to find estimators beyond mean and variance.

2) Suppose we want to find an estimator for $P(X_i > 1)$

$$P(X_i \geq 1) = 1 - F_{X_i}(1) = e^{-\theta \cdot 1} = e^{-\theta} = g(\theta)$$

$$\Rightarrow \widehat{e^{-\theta}} = e^{-\hat{\theta}} = g(\hat{\theta}) = e^{-1/\bar{X}_n}$$

Plug in estimator → estimator of whole function just plug-in of $\hat{\theta}$ of the estimator.

we plug in $\hat{\theta}$ to g

Definition

If T is an estimator for $\theta > 0$ starting from a sample. $X_1, \ldots, X_n \overset{iid}{\sim} f(-, \theta)$, then the standard error of T is the standard deviation of the sampling distribution of a point estimator, i.e.

$$\text{standard error} = \sqrt{Var_\theta(T)}$$

Example

$X_1, \ldots, X_{49} \overset{iid}{\sim} Poi(\lambda)$ $\lambda > 0$ λ is unknown

$n = 49$

a) estimator for λ and compute it's standard error

$$\lambda = E(X_i) \Rightarrow \hat{\lambda} = \bar{X}_n$$

$$\text{standard error} = \sqrt{Var(\bar{X}_n)} = \sqrt{\frac{Var(X_i)}{n}} = \sqrt{\frac{\lambda}{49}}$$

b) Point estimator for $P(X_i = 0) = e^{-\lambda} = g(\lambda)$

Plug-in estimator $\widehat{g(\lambda)} = g(\hat{\lambda}) = e^{-\hat{\lambda}} = e^{-\bar{X}_n}$

c) Data: $\sum_{i=1}^{49} x_i = 72$

Point Estimate of λ and $P(X_i = 0)$

$$\hat{\lambda} = \frac{\sum_{i=1}^{49} x_i}{49} = \frac{72}{49}$$

$$P(X_i = 0) = e^{-\hat{\lambda}} = e^{-\bar{x}_{49}} = e^{\frac{-72}{49}}$$