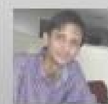
Estimator:

To estimate the value of a population parameter, you can use information from the sample in the form of an estimator.

Estimators are used in two forms:

Point estimation:

Interval estimation:



Point estimation: Based on sample data, a single number is calculated to estimate the population parameter. The rule or formula that describes this calculation is called the point estimator, and the resulting number is called a point estimate.

Interval estimation: Based on sample data, two numbers are calculated to form an interval within with the population parameter is expected to lie. The rule or formula that describes this calculation is called the interval estimator, and the resulting number is called an interval estimate or confidence interval.

Point estimator:

In practical situation, there may be several statistics that can be used as point estimators for a population parameter.

Example. The mean of sample x_1, x_2, x_3, x_4 is either

$$t_1 = \frac{x_1+x_2+x_3+x_4}{4}$$

$$t_2 = \frac{a_1x_1+a_2x_2+a_3x_3+a_4x_4}{a_1+a_2+a_3+a_4}$$

$$t_3 = \frac{x_1+x_2}{2} + \frac{2x_3+x_4}{3}$$

2:17 / 16:01

To decide which of several choices is best, you need to know "how the estimator behaves in repeated sampling situation by its sampling distribution".

Sampling distribution provide information that can be used to select the best estimator.

What characteristics would be available?

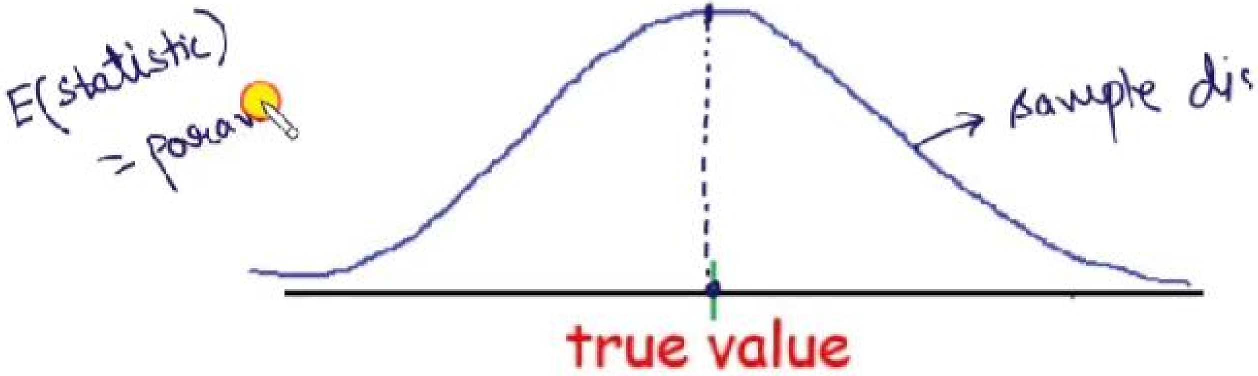
- Sampling distribution of the point estimator should be centered over the true value of the parameter to be estimated.

  Estimator should not consistently underestimate or overestimate the parameter of the interest. Such an estimator is said to be unbiased.

Sampling distribution provide information that can be used to select the best estimator.

What characteristics would be available?

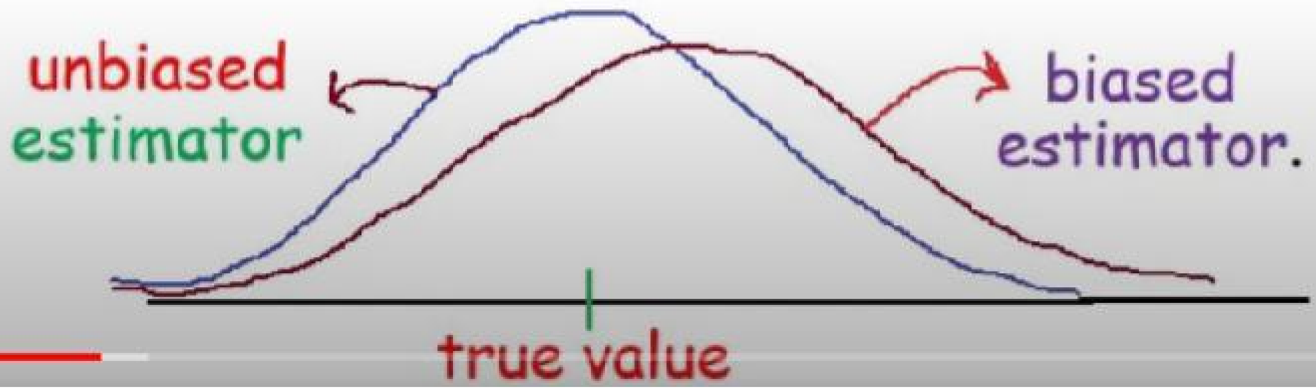
- Sampling distribution of the point estimator should be centered over the true value of the parameter to be estimated.



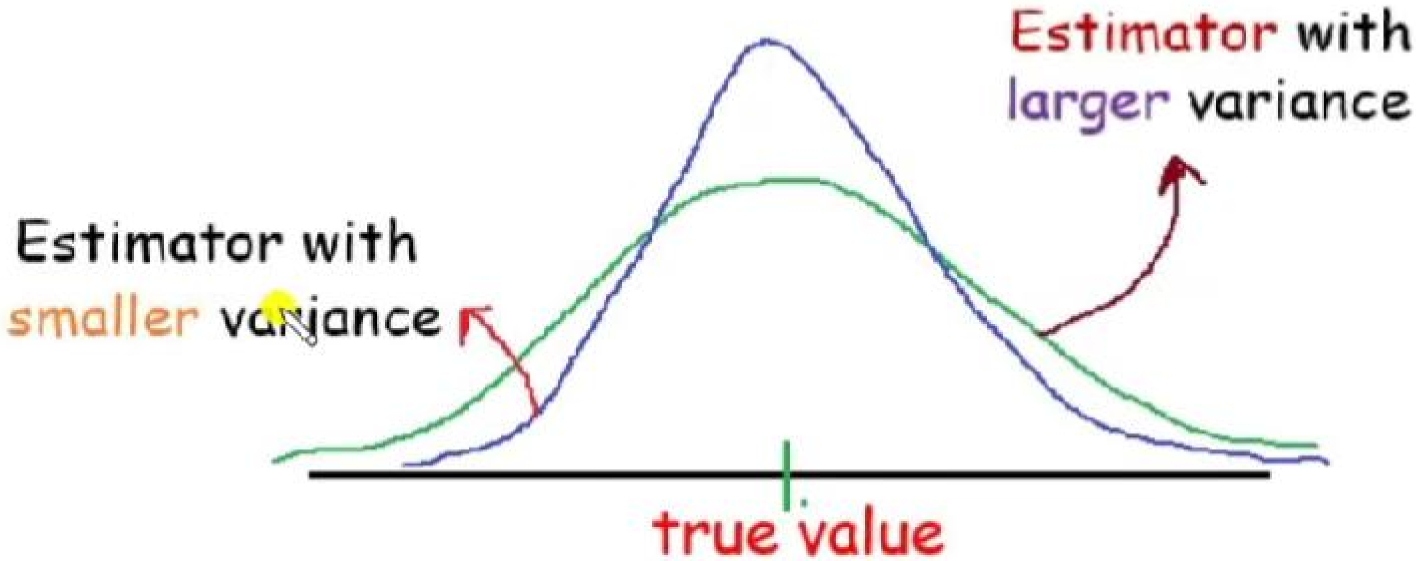
Sampling distribution provide information that can be used to select the best estimator.

What characteristics would be available?

- Sampling distribution of the point estimator should be centered over the true value of the parameter to be estimated.



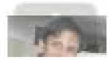
The second desirable characteristics of an estimator is that spread (as measured by variance) of the sampling distribution should be small as possible.



Characteristics of Estimators:

For a good estimator, it must have the following characteristics

- Unbiasedness $\rightarrow$
- Consistency $\rightarrow$ variance $\rightarrow 0$
- Efficiency $\rightarrow$
- Sufficiency –



Unbiasedness

Mean

An estimator $T = t(x_1, x_2, \dots, x_n)$ is said to be unbiased estimator of γ,

if $E(T) = \gamma$

$t_1 =$

$t_2 =$ —

$t_3 =$

E(statistic) = parameter

Sample Inform

Population

$E(\bar{x}) = \mu$

$E(s^2) = \sigma^2$

Unbiasedness

An estimator $T = t(x_1, x_2, \ldots, x_n)$ is said to be unbiased estimator of γ,

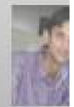
if $\qquad E(T) = \gamma \qquad\qquad E(T) \neq \gamma$

Biased.

If $E(T) > \gamma$, T is said to be positively biased.

If $E(T) < \gamma$, T is said to be negatively biased.

Then amount of bias is $E(T) - \gamma$



Example: If $x_1, x_2, \ldots, x_n$ is a random sample from a normal population $N(\mu, 1)$, show that $T = \frac{1}{n}\sum_{i=1}^{n} x_i^2$ is an unbiased estimator of $\mu^2 + 1$.

Solution: As $x_1, x_2, \ldots, x_n$ is a random sample from a normal population $N(\mu, 1)$

$$\Rightarrow E(x_i) = \mu; \quad Var(x_i) = 1 \text{ for all } i = 1,2,\ldots,n$$

For unbiased estimator, we have to show that

$$E(T) = \mu^2 + 1 \quad ; \quad T = \frac{x_1^2 + x_2^2 + \cdots + x_n^2}{n}$$

$$E(T) = \frac{1}{n}[E(x_1^2) + E(x_2^2) + \cdots + E(x_n^2)]$$

$$= \frac{1}{\cancel{n}}[\cancel{n}(\mu^2 + 1)]$$

$$= \mu^2 + 1$$

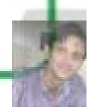
i.e., T is unbiased estimator of $\mu^2 + 1$.

From

$$Var(X) = E(X^2) - (E(X))^2$$

$$\Rightarrow \quad 1 = E(x_i^2) - \mu^2$$

$$\Rightarrow E(x_i^2) = \mu^2 + 1$$



Example: If T is unbiased estimator for β, show that T^2 is a biased estimator for β^2.

Solution: Given that T is unbiased estimator for β

$$\Rightarrow E(T) = \beta$$

To show $E(T^2) \neq \beta^2$

We know that,

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

$$\Rightarrow Var(T) = E(T^2) - \beta^2$$

$$\Rightarrow E(T^2) = \beta^2 + Var(T)$$

$$\neq \beta^2 \text{ as } Var(T) \neq 0$$

Example: For the sample $x_1, x_2, \dots, x_n$ drawn on X which takes the values 1 and 0 with respective probabilities p and $1-p$, show that $\frac{\sum x_i(\sum x_i - 1)}{n(n-1)}$ is an unbiased estimator of p^2.

Solution: As $x_1, x_2, \dots, x_n$ takes values 1 and 0 only,

$\Rightarrow x_1, x_2, \dots, x_n$ follows Bernoulli distribution, with mean p and variance $p(1-p)$

$\boxed{\sum x_i \text{ follows Binomial distribution with mean } np \text{ and variance } np(1-p).}$

Let $H = \sum x_i$ such that $E(H) = np$; $Var(H) = np(1-p)$

Our aim is to show $\frac{H(H-1)}{n(n-1)}$ is an unbiased estimator of p^2, i.e.,

$$E\left(\frac{H(H-1)}{n(n-1)}\right) = p^2$$

Now,

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

unbias

$$= \frac{1}{n(n-1)}[E(H^2) - E(H)]$$

$$= \frac{1}{n(n-1)}[np(1-p) + n^2p^2 - np]$$

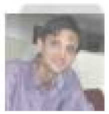
$$= \frac{1}{n(n-1)}p^2(n^2 - n)$$

$$= p^2$$

Hence, the result.

$E(H) = np$;

$Var(H) = np(1-p)$



Example: Let X be distributed in the Poisson form with parameter λ. Show that only unbiased estimator of $e^{-(k+1)\lambda}, k > 0$ is $T(X) = (-k)^X$.

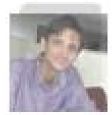
Solution: Since X follows Poisson distribution with parameter λ.

Thus, $E(X) = \lambda$, $Var(X) = \lambda$

$$
\begin{aligned}
E(T) &= E((-k)^x) \\
&= \sum_{x=0}^{\infty} (-k)^x \frac{e^{-\lambda}\lambda^x}{x!} \\
&= e^{-\lambda} \sum_{x=0}^{\infty} \frac{(-\lambda k)^x}{x!} \\
&= e^{-\lambda} e^{-\lambda k} \\
&= e^{-(1+k)\lambda}
\end{aligned}
$$

To show,

$E(T) = e^{(-k+1)\lambda}$



Example: Let t_1 and t_2 be two unbiased estimates of θ with variances σ_1^2 and σ_2^2 (both known) and correlation ρ. Consider the estimate $\hat{\theta} = \alpha t_1 + (1-\alpha)t_2$, check whether $\hat{\theta}$ is unbiased or not?

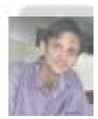
Solution. Given that t_1 and t_2 are the two unbiased estimates of θ

$$\Rightarrow E(t_1) = \theta \quad ; \quad E(t_2) = \theta$$

Now

$$\boxed{\begin{aligned} E(\hat{\theta}) &= E(\alpha t_1 + (1-\alpha)t_2) \\ &= \alpha E(t_1) + (1-\alpha)E(t_2) \\ &= \alpha\theta + (1-\alpha)\theta \\ &= \theta \end{aligned}}$$

Hence $\hat{\theta}$ is unbiased estimate of θ.



Example: A random sample x_1, x_2, x_3, x_4, x_5 of size 5 is drawn from a normal distribution with unknown mean μ. Consider the following estimators to estimate μ:

$$t_1 = \frac{x_1+x_2+x_3+x_4+x_5}{5}\ ; \quad t_2 = \frac{x_1+x_2}{2} + x_3\ ; \quad t_3 = \frac{2x_1+x_2+\lambda x_3}{4}$$

where λ is such that t_3 is an unbiased estimator of μ.

Find λ. Check whether t_1, t_2 are unbiased or not?

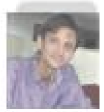
Solution: Since t_3 is an unbiased estimator of μ.

$$\Rightarrow E(t_3) = \mu$$

$$\Rightarrow E\left(\frac{2x_1 + x_2 + \lambda x_3}{4}\right) = \mu$$

$$\Rightarrow \frac{1}{4}[2E(x_1) + E(x_2) + \lambda E(x_3)] = \mu$$

$$\Rightarrow \frac{1}{4}[2\mu + \mu + \lambda\mu] = \mu \quad \Rightarrow \boxed{\lambda = 1}$$



To check whether t_1, t_2 are unbiased or not,

we have to show that $E(t_1) = \mu$, $E(t_2) = \mu$

$$E(t_1) = E\left(\frac{x_1 + x_2 + x_3 + x_4 + x_5}{5}\right)$$

$$= \frac{1}{5}[E(x_1) + E(x_2) + \cdots + E(x_5)]$$

$$= \frac{1}{5}[\mu + \mu + \mu + \mu + \mu]$$

$$= \mu$$

Hence, t_1 is an unbiased estimator of μ.

$$E(t_2) = E\left(\frac{x_1 + x_2}{2} + x_3\right)$$

$$= \frac{1}{2}[E(x_1) + E(x_2)] + E(x_3)$$

$$= \frac{1}{2}(\mu + \mu) + \mu$$

$$= 2\mu \neq \mu$$

Hence, t_2 is not an unbiased estimator of μ.