Course: Mathematical statistics

Week 3: Point Estimation, Properties of estimators

Lecturer: Nagulama Moses

Kumi University

March 31, 2025

Outline

Intended learning outcomes

- Explain the meaning of an estimator and an estimate in statistical inference
- Define unbiasedness and state the condition for an estimator to be unbiased
- Explain how the Cramér-Rao lower bound provides a benchmark for efficiency
- Apply the properties of unbiasedness, efficiency, and asymptotics to compare common estimators

Estimation

- Statistics are used to estimate the corresponding parameters. Since parameters are generally unknown, sample statistics serve as point estimates or interval estimates.
- A single value used to estimate an unknown parameter. For example: Sample Mean $(\bar{x})$ is a point estimate of μ, Sample Proportion $(\hat{p})$ is a point estimate of p.
- A statistic is unbiased if its expected value equals the parameter it estimates. For instance:

$$E[\bar{x}] = \mu, \quad E[s^2] = \sigma^2.$$

- Inferential statistics consist of methods used to make/ draw conclusion about a population using sample values and it has two branches i.e. estimation and hypothesis testing.
- **Definition:** Estimation as used in statistic is any statistical method used to approximate a population parameter using known sample statistic.
- There are two branches of estimation i.e. point estimation and interval estimation.

Point estimation

- let θ be a population parameter (any) and let $\hat{\theta}$ be its estimate.
- A single numerical value statistic $\hat{\theta}$ which approximate the value of θ of a population parameter is called a point estimator.
- when a single value of a statistic is used to estimate the population parameter the process is called point estimation.

Properties of point estimators

Unbiased estimator

- A statistic $\hat{\theta}$ used to estimate a parameter θ is said to be unbiased if and only if $E(\hat{\theta}) = \theta$. Otherwise it is said to be biased.
- If an estimator $\hat{\theta}$ of θ is not unbiased, then the difference $b(\hat{\theta}) = E(\hat{\theta}) - \theta$ is called the biased estimator.
- For unbiased estimator $b(\hat{\theta}) = 0$

Example

Given that $\hat{\theta_1}$ and $\hat{\theta_2}$ are both estimators of a parameter θ. Determine the condition that must be imposed on a constant k_1 and k_2 so that $k_1\hat{\theta} + k_2\hat{\theta}$ is unbiased estimator of θ.

solution

- θ_1 is unbiased estimator *iff* $E(\hat{\theta}) = \theta$
- $E(k_1\hat{\theta_1} + k_2\hat{\theta_2}) = E(k_1\hat{\theta_1}) + E(k_2\hat{\theta_2})$
  $= k_1 E(\hat{\theta_1}) + k_2 E(\hat{\theta_2})$
- $= k_1\theta + k_2\theta$, where $\hat{\theta_1}$ and $\hat{\theta_2}$ are unbiased estimators of θ
- $= (k_1 + k_2)\theta$
- $= \theta$ *iff* $k_1 + k_2 = 1$

Example let $x_1, x_2, x_3, ..., x_n$ constitute a random sample taken from a population with a probability density function given by

$$f(x) = \begin{cases} e^{-(x-\delta)}, x > 0 \\ 0, otherwise \end{cases}$$

where δ is the population parameter so that the sample mean is a biased estimator of a parameter δ and find the bias of the estimator.

solution

let $\overline{x}$ be sample mean

if $\overline{x}$ is biased estimator of δ, then $E(\overline{x} \neq \delta$

sampling distribution of $\overline{x}$

$$E(\overline{x}) = \mu, \mu = E(x)$$

$$= \int_0^{\infty} xe^{-(x-\delta)} dx$$

$$= \lim_{b \to \infty} \int_0^{b} xe^{-(x-\delta)} dx$$

$$= \lim_{b \to \infty} \int_0^{b} e^{\delta} xe^{-x} dx$$

$$= e^{\delta} \lim_{b\to\infty} \int_0^b xe^{-x} dx$$

$$= e^{\delta} \lim_{b\to\infty} \left[-xe^{-x} - e^{-x}\right]_0^b$$

$$= e^{\delta} \lim_{b\to\infty} \left[-be^{-b} - e^{-b} + 1\right]_0^b$$

$$\mu = e^{\delta}$$

$$E(\overline{x}) = e^{\delta} \neq \sigma$$

therefore $\overline{x}$ is a biased estimator of σ
bias $= b(\overline{x}) = E(\overline{x}) - \sigma$

$$= e^{\sigma} - \sigma$$

Theorem:

- Suppose x is a random variable with mean μ variance σ^2.
- let $x_1, x_2, \ldots\ldots, x_n$ be random samples of size n taken from a population represented by x, then the sample mean $\overline{x}$ and sample variance s^2 are unbiased estimators of the population mean μ and a population variance σ^2 respectively.

Proof

$$E(\overline{x}) = E\left(\frac{x_1 + x_2 + ... + x_n}{n}\right)$$

$$\frac{1}{n}E(x_1 + x_2 + ... + x_n)$$

$$\frac{1}{n}[E(x_1) + E(x_2) + ... + E(x_n)]$$

$$= \frac{1}{n}n\mu = \mu$$

$$E(s^2) = \frac{\Sigma_{i=1}^{n}(x_i - \overline{x})^2}{n-1}$$

$$\frac{1}{n-1}E\left[\Sigma_{i=1}^{n}(x_i - \overline{x})^2\right]$$

$$\frac{1}{n-1}E[\Sigma_{i=1}^{n}(x_i^2 - 2\overline{x}x_i + \overline{x}^2)]$$

$$= \frac{1}{n-1}E\left[\Sigma_{i=1}^{n}x_i^2 - 2\overline{x}\Sigma_{i=1}^{n}x_i + \Sigma_{i=1}^{n}\overline{x}^2\right]$$

$$= \frac{1}{n-1}E\left[\Sigma_{i=1}^{n}x_i^2 - 2\overline{x}n\overline{x} + n\overline{x}^2\right]$$

$$= \frac{1}{n-1} E\left[\Sigma_{i=1}^{n} x_i^2 - n\overline{x}^2\right]$$

$$= \frac{1}{n-1} E\left[\Sigma_{i=1}^{n} x_i^2 - E(n\overline{x}^2)\right]$$

$$= \frac{1}{n-1} \left[\Sigma_{i=1}^{n} E(x_i^2) - nE(\overline{x}^2)\right]$$

from

$$\sigma^2 = E(x_i^2) - \mu^2$$

$$E(x_i^2) = \sigma^2 + \mu^2$$

$$var(\overline{x}) = E(\overline{x}^2) - [E(\overline{x})]^2$$

$$E(\bar{x}^2) = \frac{\sigma^2}{n} + \mu^2$$

$$E(s^2) = \frac{1}{n-1}\left[n(\sigma^2 + \mu^2) - n(\frac{\sigma^2}{n} + \mu)\right]$$

$$E(s^2) = \frac{1}{n-1}\left[n(\sigma^2 + \mu^2) - \sigma^2 - n\mu^2)\right]$$

$$E(s^2) = \frac{1}{n-1}\left[n\sigma^2 + n\mu^2 - \sigma^2 - n\mu^2)\right]$$

$$E(s^2) = \frac{1}{n-1}\left[(n-1)\sigma^2)\right]$$

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

The sample standard deviation is not unbiased estimator of the sample standard deviation σ

Asymptotically unbiased estimator

let $\hat{\theta}$ be an estimator of θ then $\hat{\theta}$ is said to asymptotically unbiased estimator of θ if the

$$\lim_{n\to\infty} b(\theta) = \lim_{n\to\infty} E(\hat{\theta}) - \theta = 0$$

Example

show that for a random sample taken from a population with pdf

$$f(x) = \begin{cases} e^{-(x-\sigma)}, x > 0 \\ 0, elsewhere \end{cases}$$

$\overline{x}$ is not asymptotically unbiased estimator of σ.

$$E(\overline{x}) = e^{\sigma} \neq \sigma$$

$\overline{x}$ is a biased estimator of σ

$$b(\sigma) = E(\overline{x}) - \sigma$$

$$e^{\sigma} - \sigma$$

$$\lim_{n\to\infty} b(\sigma) = \lim_{n\to\infty} (e^{\sigma} - \sigma) \neq 0$$

The sample mean is not asymptotically unbiased estimator of the parameter σ

Example:
let x be the number of successes in n independent trials from a binomial population with a parameter P. show that $\frac{x+1}{n+2}$ is a biased estimator of the parameter P. Is this estimator asymptotically unbiased.

Solution
for unbiasedness $E\left(\frac{x+1}{n+2}\right) = P$

$$E\left(\frac{x+1}{n+2}\right) = \frac{1}{n+2}E(x+1) = \frac{1}{n+2}\left[E(x)+E(1)\right]$$

$$= \frac{1}{n+2}\left[np+1\right] = \frac{np+1}{n+2} \neq P$$

Therefore $\frac{x+1}{n+2}$ is a biased estimator of P.

$$b(P) = E\left(\frac{x+1}{n+2}\right) - P = \frac{np+1}{n+2} - P$$

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

$$\lim_{n\to\infty} b(p) = \lim_{n\to\infty} \frac{1-2p}{n+2}$$

$$(1-2p)\lim_{n\to\infty}\frac{1}{n+2} = (1-2p).0 = 0$$

Therefore $\frac{x+1}{n+2}$ is asymptotically unbiased estimators of a parameter p

Efficient Estimator

- Suppose $\hat{\theta}_1$ and $\hat{\theta}_2$ are both unbiased estimators of a parameter θ.
- Distribution of each estimator is centred at θ. Both distribution have different variance $\hat{\theta}_2$ is more spread out than $\hat{\theta}_1$.
- Implying $\hat{\theta}_1$ has a smaller variance. logical choice for a better estimator is one with a smaller variance.

Definition: if we consider all unbiased estimators of the parameter θ, The one with the smallest variance is called the minimum variance unbiased estimator (MVUE).

- Thus if $\hat{\theta_1}$ and $\hat{\theta_2}$ are unbiased estimator of θ then $\hat{\theta_1}$ is said to be MVUE if the variance $var(\hat{\theta_1}) \leq var(\hat{\theta_2})$
- $\hat{\theta_1}$ is then said to be a more efficient estimator than finding MVUE is generally difficult.
- However checking whether or not a particular unbiased estimator possess a minimum variance can be done using the cramer-Rao's inequality.

Cramer-Rao's inequality states that of $\hat{\theta}$ is unbiased estimator of θ, then the

$$var(\hat{\theta}) = \frac{1}{n\left[E(\frac{\partial lnf(x)}{\partial\theta})^2\right]}$$

where $f(x)$ is the value of the probability density at x and n s the size of the random sample.

Example

show that the sample mean $\overline{x}$ is a minimum variance unbiased estimator of he population mean μ of a normal population/distribution

solution

Normal population $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$

$$lnf(x) = -ln\sigma - \frac{1}{2}ln2\pi - \frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2$$

$$\frac{\partial}{\partial\mu}[inf(x)] = +\left(\frac{x-\mu}{\sigma}\right).\frac{1}{\sigma} = \frac{x-\mu}{\sigma^2}$$

$$\left[\frac{\partial}{\partial\mu}[lnf(x)]\right]^2 = \frac{1}{\sigma^4}(x-\mu)^2 = \frac{1}{\sigma^2}\left(\frac{x-\mu}{\sigma}\right)^2$$

$$(\frac{\partial}{\partial\mu} lnf(x))^2 = E\left(\frac{1}{\sigma^2}\left[\frac{x-\mu}{\sigma}\right]^2\right)$$

$$\frac{1}{\sigma^2}E(\frac{x-\mu}{\sigma})^2 = \frac{1}{\sigma^4}E(x-\mu)^2$$

$$\frac{1}{\sigma^4}[E(x^2-2\mu x+\mu^2)]$$

$$\frac{1}{\sigma^4}[E(x^2)-2\mu E(x)+\mu^2)]$$

$$\frac{1}{\sigma^4}[E(x^2-\mu^2)] = \frac{\sigma^2}{\sigma^4} = \frac{1}{\sigma^2} = \frac{1}{n.\frac{1}{\sigma^2}} = \frac{\sigma^2}{n}$$

Therefore, the $var(\overline{x}) = \frac{\sigma^2}{n}$

- The standard error of an estimator is a standard deviation of its distribution.
- If we are sampling from a normal distribution with μ and σ^2, then sampling distribution with $\mu_{\overline{X}} = \mu$ and $\sigma^2_{\overline{X}} = \frac{\sigma^2}{n}$.
- The standard error $\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}}$ and the standard error of the sample error involves unknown parameter σ.
- In circumstances where σ is not known we can replace σ by the sample standard deviation $\sigma_{\overline{X}} = \frac{s}{\sqrt{n}}$

- Sometimes we need to compare two estimators that are not necessary unbiased. we can make use of what is referred to as the mean square error (MSE)

- let $\hat{\theta}$ be a point estimator of a population parameter θ then

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

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

- MSE is useful in comparing two estimators.
- let $\hat{\theta}_1$ and $\hat{\theta}_2$ be any two estimators of a parameter θ and let $MSE(\hat{\theta}_1)$ and $MSE(\hat{\theta}_2)$ respectively. we refer to that quantity as the relative efficiency of $\hat{\theta}_2$ and $\hat{\theta}_1$
- If $\frac{MSE(\hat{\theta}_1)}{MSE(\hat{\theta}_2)} < 1$, then is a more efficient estimator in the sense of the MSE.

Example

- suppose that $\hat{\theta_1}, \hat{\theta_2}, \hat{\theta_3}$ are estimators of the parameter θ and we are given that $E(\hat{\theta_1}) = E(\hat{\theta_2}) = \theta$ and $E(\hat{\theta_3}) \neq \theta$ and the variance of $\theta_1, var(\hat{\theta_1}) = 12$ and $var(\hat{\theta_2}) = 10$ and $E(\hat{\theta_3} - \theta)^2 = 12.$ compare the 3 estimators which estimator do you prefer and why?.

solution

- $\hat{\theta_1}$ and $\hat{\theta_2}$ are both unbiased estimators of θ, $\hat{\theta_3}$ is a biased estimator of θ

$$var(\hat{\theta_2}) < var(\hat{\theta_1})$$

- $\hat{\theta_2}$ is a more efficient estimator of θ in the sense of MVUE.

- $MSE(\hat{\theta_1}) = 12, MSE(\hat{\theta_2}) = 10, MSE(\hat{\theta_3}) = 12$ i.e $MSE(\hat{\theta_1}) = 12 = var(\hat{\theta_1}) + (bia(\theta_1))^2$

$$MSE(\hat{\theta_2}) < MSE(\hat{\theta_1})$$

- $\hat{\theta_2}$is a better estimator than $\hat{\theta_1}$ in a sense of MSE.

$$MSE(\hat{\theta_2}) < MSE(\hat{\theta_3})$$

- $\hat{\theta_2}$is a better estimator of θ in a sense of MSE.

$$MSE(\hat{\theta_1}) = MSE(\hat{\theta_3})$$

  are equally efficient in the sense of MSE.
- $\hat{\theta_2}$ is the most estimator in the sense of MVUE and MSE.

References

- Hogg,R;Mckean,J;Craig,A(2012).Introduction to mathematical statistics, 7th edition, pearson Prentice Hall, 2012.

- Hastings K.J,(1997) Probability and statistics, Addison Wesley reading,massachusetts.

Thank You!

Next Lecture: Properties of Estimators