Lecture 1: Point Estimators

Dr. Mutua Kilai

Department of Pure and Applied Sciences

Jan-April 2024

Kirinyaga University

Basic Concepts

- A point estimator of a parameter θ denoted by $\hat{\theta}$ is a single number that can be considered as a possible value for θ
- An **estimator** is a rule or a formula that tells us how to calculate an estimate based on measurements contained in the sample.
- A point estimate is a single number calculated from available sample data, that is used to estimate the value of an unknown population parameter.
- Some simple examples are:
  i. If $X_1, ..., X_n$ is from $B(1, p)$ then $\hat{p} = \frac{1}{n}\sum_{i=1}^{n} X_i$ the sample proportion of success
  ii. If $X_1, ..., X_n$ is from normal population then the estimates for $\hat{\mu} = \bar{X}, \hat{\sigma^2} = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2 = S^2$

Properties of Point Estimators

1. Unbiasedness
2. Consistency
3. Efficiency
4. Sufficiency
5. Completeness

Unbiasedness

- An estimator $\hat{\theta}$ is said to be unbiased for parameter θ if

$$E(\hat{\theta}) = \theta$$

- If this does not hold $\hat{\theta}$ is said to be biased estimator of θ with bias given by

$$Bias(\hat{\theta}) = E(\hat{\theta} - \theta) = E(\hat{\theta}) - \theta$$

Example 1

An engineer wishes to estimate the mean yield of a chemical process based on the yield measurement x_1, x_2, x_3 from three independent runs of an experiment. Consider the following two estimators of the mean yield θ such that $E(X_i) = \theta$

$$\hat{\theta}_1 = \frac{x_1 + x_2 + x_3}{3} = \frac{1}{3}\sum_{i=1}^{n} x_i$$

$$\hat{\theta}_2 = \frac{x_1 + 2x_2 + x_3}{4}$$

Which one should we prefer?

Solution

$$E(\hat{\theta}_1) = E\left(\frac{x_1 + x_2 + x_3}{3}\right)$$
$$= \frac{\theta + \theta + \theta}{3} \qquad (1)$$
$$= \theta$$

$$E(\hat{\theta}_2) = \frac{E(x_1) + 2E(x_2) + E(x_3)}{3} \qquad (2)$$
$$= \theta$$

Both $\hat{\theta}_1$ and $\hat{\theta}_2$ are unbiased for θ

Exercise

Let $X_1, X_2, ..., X_n$ be a random sample from a population with mean μ. Suppose

$$T_1 = \frac{X_1 + X_2 + X_3 + X_4 + X_5}{5}$$

$$T_2 = \frac{X_1 + X_2}{2} + X_3$$

$$T_3 = \frac{2X_1 + X_2 + \lambda X_3}{3}$$

Where λ is such that T_3 is an unbiased estimator for μ

i Find λ

ii. Are T_1 and T_2 unbiased

Relative Efficiency (RE)

- If we have two unbiased estimators of a parameter $\hat{\theta}_1$ and $\hat{\theta}_2$, we say that $\hat{\theta}_1$ is relatively more efficient than $\hat{\theta}_2$ if $var(\hat{\theta}_2) > var(\hat{\theta}_1)$ and vice versa.
- So clearly there is a necessity for a criterion that enables us to choose between estimators with common property of unbiasedness. Such a criterion based on variances of the sampling distribution of estimators is known as **efficiency**
- Efficiency of $\hat{\theta}_1$ relative to $\hat{\theta}_2$ is given by

$$eff(\hat{\theta}_1, \hat{\theta}_2) = \frac{var(\theta_2)}{var(\theta_1)}$$

- If $eff(\hat{\theta}_1, \hat{\theta}_2) > 1$ choose $\hat{\theta}_1$
- If $eff(\hat{\theta}_1, \hat{\theta}_2) < 1$ choose $\hat{\theta}_2$

Example

Let $X_i, i = 1, 2, ..., n$ be iid random sample obtained from a population with mean μ and variance σ^2. Suppose we are given the following two estimators for the parameter μ

$$\hat{\mu_1} = \frac{1}{2}(X_1 + X_2)$$

$$\hat{\mu_2} = \bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

i. Check the unbiasedness of the estimators $\hat{\mu_1}$ and $\hat{\mu_2}$

ii. Find the efficiency of $\hat{\mu_1}$ relative to $\hat{\mu_2}$

Solution

We are given that

$$E(X_i) = \mu$$

$$E(\hat{\mu_1}) = \frac{E(X_1) + E(X_2)}{2} = \frac{\mu + \mu}{2} = \mu$$

$$\hat{\mu_2} = \bar{X}$$

$$E(\bar{X}) = E\left(\frac{X_1 + X_2 + ... + X_n}{n}\right) = \frac{n\mu}{\mu} = \mu$$

Cont'd

Efficiency of $\hat{\mu}_1$ relative to $\hat{\mu}_2$

$$
\begin{aligned}
Var(\hat{\mu}_1) &= Var\left(\frac{X_2 + X_2}{2}\right) \\
&= \frac{1}{4}[Var(X_1) + Var(X_2)] \qquad (3) \\
&= \frac{\sigma^2}{2}
\end{aligned}
$$

$$
\begin{aligned}
Var(\hat{\mu}_2) &= Var\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) \\
&= \frac{1}{n^2}\sum_{i=1}^{n} Var(X_i) \qquad (4) \\
&= \frac{\sigma^2}{n}
\end{aligned}
$$

Cont'd

Efficiency of $\hat{\mu_1}$ relative to $\hat{\mu_2}$ is thus obtained as

$$\begin{aligned} eff(\hat{\mu_1}, \hat{\mu_2}) &= \frac{var(\hat{\mu_2})}{var(\hat{\mu_2})} \\ &= \frac{\sigma^2}{n} \Big/ \frac{\sigma^2}{2} \\ &= \frac{2}{n} \end{aligned} \tag{5}$$

Exercise 1

Let

$$\hat{\sigma_1^2} = \frac{1}{n-1}\sum_{i=1}^{n}(\mu_i - \mu)^2$$

and

$$\hat{\sigma_2^2} = \frac{1}{2}\sum_{i=1}^{n}(\mu_i - \mu)^2$$

Find the efficiency of $\hat{\sigma_1^2}$ relative to $\hat{\sigma_2^2}$

Exercise 2

Let

$$X_1, X_2, ..., X_n$$

be a random sample with $E(X_i) = \mu_i$ and $Var(X_i) = \sigma_i^2$. Show that

i. $S_1^2 = \frac{1}{n}\sum(X_i - \bar{X})^2$ is a biased estimator for σ^2

ii. $S_2^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$ is unbiased estimator for σ^2

Note

Population and sample variances are given respectively as:

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

$$S_i^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

Consistency

- An estimator $\hat{\theta}_n$ is a consistent estimator of the parameter θ if $\hat{\theta}_n \to \theta$ i.e if $\hat{\theta}_n$ converges in probability to θ

$$\lim_{n\to\infty} p(|\hat{\theta}_n - \theta| > \epsilon) = 0$$

for any $\epsilon > 0$

Theorem

An unbiased estimator $\hat{\theta}_n$ for θ is consistent if

$$\lim_{n\to\infty} Var(\hat{\theta}_n) = 0$$

Note: Proof is omitted.

Example 1

Let $Y_1, Y_2, ..., Y_n$ be a random sample of size n from a population with mean μ and variance $\sigma^2 < \infty$. Show that

$$\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i$$

is a consistent estimator of μ

Solution

$$\begin{aligned} E(\bar{Y}) &= E\left(\frac{1}{n}\sum_{i=1}^{n} Y_i\right) \\ &= \frac{1}{n}\sum_{i=1}^{n} E(Y_i) = \frac{n\mu}{\mu} = \mu \end{aligned} \tag{6}$$

This implies that $\bar{Y}$ is unbiased for μ

Cont'd

$$
\begin{aligned}
Var(\bar{Y}) &= \frac{1}{n^2}\sum_{i=1}^{n} Var(Y_i) \\
&= \frac{n\sigma^2}{n^2} \\
&= \frac{\sigma^2}{n} \to 0 \;\; as \;\; n \to \infty
\end{aligned}
\tag{7}
$$

Hence $\bar{Y}$ is consistent.

How to Find Estimators

There are two main methods namely

1. Method of Moments (MoM)
2. Method of Maximum Likelihood (MLE)

Method of Moments (MoM)

- Is a very simple procedure for
nding an estimator for one or more parameters of a statistical model. It's one of the oldest methods for deriving point estimators.
- **Recall** k^{th} moment of a random variable is $\mu_k = E(Y^k)$. The corresponding k^{th} sample moment is

$$m_k = \frac{1}{n}\sum_{i=1}^{n} Y_i^k$$

  m_k is an estimator for μ_k
- The estimator based on the method of moments will be the solution to the equation

$$\mu_k = m_k$$

Example 1

Let $Y_1, Y_2, ..., Y_n \sim N(\mu, \sigma^2)$. Find the MoM estimators of μ and σ^2

Solution

$$\mu_1 = E(Y_i) = \mu$$

$$\mu_2 = E(Y_i^2) = \sigma^2 + \mu^2 \Rightarrow \sigma^2 = E(Y_i^2) - (E(Y_i))^2$$

$$m_1 = \frac{1}{n} \sum Y_i = \bar{Y}$$

$$m_2 = \frac{1}{n} \sum Y_i^2$$

Cont'd

$$\mu_1 = m_1 \Rightarrow \hat{\mu} = \bar{Y}$$

$$\mu_2 = m_2 \Rightarrow \sigma^2 + \mu^2 = \frac{1}{n}\sum Y_i^2$$

$$\begin{aligned}\hat{\sigma^2} &= \frac{1}{n}\sum Y_i^2 - \bar{Y}^2 \\ &= \frac{1}{n}\sum(Y_i - \bar{Y})^2\end{aligned} \tag{8}$$

Hence $\hat{\sigma^2}$ is biased for σ

Exercise

Let $Y_1, Y_2, ..., Y_n$ be a random sample obtained from $Unif(0, \theta)$. Use MoM to estimate θ.

Maximum Likelihood Estimators (MLE)

- Suppose the likelihood function depends on k parameters $\theta_1, \theta_2, ..., \theta_k$. Choose as estimates those values of the parameters that maximize the likelihood

$$L(\theta_1, \theta_2, ..., \theta_k) = L(y_1, y_2, ..., y_n|\theta_1, ..., \theta_k)$$

- $l(\theta) = \ln(L(\theta))$ is the log-likelihood function
- Both the likelihood function and the log-likelihood function have their maxima at the same value of θ but its often easier to maximize log-likelihood function $l(\theta)$

Example 1

Let $Y_1, Y_2, ..., Y_n \sim i.i.dN(\mu, \sigma^2)$. Find the MLEs of μ and σ^2

Solution

$$
\begin{aligned}
L(\mu, \sigma^2) &= f(y_1|\mu, \sigma^2) \times ... \times f(y_n|\mu, \sigma^2) \\
&= \left[\frac{1}{\sqrt{2\pi\sigma^2}}\right]^n e^{-\frac{\sum (y_i-\mu)^2}{2\sigma^2}} \\
&= (2\pi\sigma^2)^{-\frac{n}{2}} e^{-\frac{\sum (y_i-\mu)^2}{2\sigma^2}}
\end{aligned}
\tag{9}
$$

$$
l(\mu, \sigma^2) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln\sigma^2 - \frac{1}{2\sigma^2}\sum (y_i - \mu)^2
$$

Cont'd

$$l'(\mu) = \frac{2}{2\sigma^2} \sum (y_i - \mu) = 0 \Rightarrow \hat{\mu} = \frac{1}{n} \sum y_i = \bar{y}$$

$$l''(\mu) = -\frac{1}{\sigma^2} \sum_{i=1}^{n} (y_i - \mu)^{1-1} = -\frac{1}{\sigma^2} \sum 1 = -\frac{n}{\sigma^2} < 0$$

$$\hat{\mu} = \bar{y}$$

is maximum.

$$l'(\sigma^2) = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum(y_i - \mu)^2 = 0$$

$$n = \frac{1}{\sigma^2}\sum(y_i - \mu)^2 \Rightarrow \hat{\sigma^2} = \frac{1}{n}\sum(y_i - \bar{y})^2$$

Exercise

Show that $\hat{\sigma^2}$ gives the maximum.

Exercise 1

The pareto distribution has probability density function

$$f(x) = \theta\alpha^{\theta}x^{-\theta-1}, \;\; \textit{for} \;\; x \geq \alpha, \theta > 1$$

where θ and α are positive parameters of the distribution. Assume that α is known and that $X_1, X_2, ..., X_n$ is a random sample of size n

a. Find the MoM estimator for θ

b. Find the MLE for θ. Does this estimator differ from that found in part (a)?

1. Estimate θ based on these data $3, 5, 2, 3, 4, 1, 4, 3, 3, 3$

Exercise 2

Let $X_i, i = 1, 2, ..., n$ be an *i.i.d* collection of Poisson random variables with parameter λ where $\lambda > 0$. Find the MLE of λ

Exercise 3

Let $X_i, i = 1, ..., n$ be a random sample from a geometric distribution with

$$f(x, p) = (1 - p)^{x-1}, x = 1, 2, 3, ...$$

Find the estimator for p

Solution: Exercise 1

solution (a)

$$\mu_1 = m_1$$

$$\begin{aligned}
\mu_1 = E(X) &= \int_{\alpha}^{\infty} \theta\alpha^{\theta} X^{-\theta} dx \\
&= \theta\alpha^{\theta} \int_{\alpha}^{\infty} X^{-\theta} dx \\
&= \theta\alpha^{\theta} \left[\frac{X^{-\theta+1}}{-\theta+1} \right]_{x=\alpha}^{x=\infty} \\
&= -\theta\alpha^{\theta} \frac{\alpha^{-\theta+1}}{-\theta+1} \\
&= \frac{\theta\alpha}{\theta-1}
\end{aligned}$$

$$m_1 = \bar{X}$$

$$\bar{X} = \frac{\theta\alpha}{\theta - 1} \implies \hat{\theta} = \frac{\bar{X}}{\bar{X} - \alpha}$$

solution (b)

$$\begin{aligned} L(\theta) &= \prod_{i=1}^{n} \theta\alpha^{\theta} X_i^{-\theta-1} \\ &= \theta^n \alpha^{n\theta} \prod X_i^{-\theta-1} \end{aligned}$$

$$\begin{aligned} l(\theta) &= lnL(\theta) \\ &= nln(\theta) + n\theta ln\alpha - (\theta + 1)\sum lnX_i \end{aligned}$$

solution (c) We are given

$X_i = 3, 5, 2, 3, 4, 1, 4, 3, 3, 3$

$n = 10,\ \bar{X} = 3.1\ ,\ \sum lnX_i = 10.57$

$$
\begin{aligned}
MoM:\ \hat{\theta} &= \frac{\bar{X}}{\bar{X} - \alpha} \\
&= \frac{3.1}{3.1 - \alpha}
\end{aligned}
$$

$$
\begin{aligned}
MLE:\ \hat{\theta} &= \frac{n}{\sum lnX_i - nln\alpha} \\
&= \frac{10}{10.57 - 10ln\alpha}
\end{aligned}
$$

Sufficiency Principle

- A **sufficient statistic** for a parameter θ is a statistic that in a certain sense captures all the information about θ contained in the sample.
- **Sufficiency Principle** If $T(\mathbf{X})$ is a sufficient statistic for θ then any inference about θ should depend on the sample $\mathbf{X}$ only through the value $T(\mathbf{X})$

Sufficient Statistic

A statistic $T(X)$ is a sufficient statistic for θ if the conditional distribution of the sample $\mathbf{X}$ given the value of $T(X)$ does not depend on θ.

- To verify $T(X)$ is a sufficient statistic for θ we must verify that for any fixed values of x and t the conditional probability $P_\theta(X = x | T(X) = t)$ is the same for all values of θ

Theorem 1

If $p(x|\theta)$ is the joint pdf or pmf of $\mathbf{X}$ and $q(t|\theta)$ is the pdf or pmf of $T(\mathbf{X})$ then $T(\mathbf{X})$ is a sufficient statistic for θ if for every x in the sample space the ratio $p(x|\theta)/q(T(x)|\theta)$ is constant as a function of θ.

Example 1:Binomial Sufficient Statistic

Let $X_1, X_2, ..., X_n$ be iid Bernoulli random variables with parameter $\theta, 0 < \theta < 1$. Show that the statistic $T(X) = X_1 + X_2 + ... + X_n$ is a sufficient statistic for θ

Solution

$$p(x|theta) = \prod \theta^{x_i}(1-\theta)^{1-x_i}$$

$$q(T(X)|\theta) = \binom{n}{t}\theta^t(1-\theta)^{n-t}$$

Cont'd

Therefore the ratio is

$$\begin{aligned}
\frac{p(x|\theta)}{q(T(x)|\theta)} &= \frac{\prod \theta^{x_i}(1-\theta)^{1-x_i}}{\binom{n}{t}\theta^t(1-\theta)^{n-t}} \\
&= \frac{\theta^{\sum x_i}(1-\theta)^{\sum(1-x_i)}}{\binom{n}{t}\theta^t(1-\theta)^{n-t}} \\
&= \frac{\theta^t(1-\theta)^{n-t}}{\binom{n}{t}\theta^t(1-\theta)^{n-t}} = \frac{1}{\binom{n}{t}} = \frac{1}{\binom{n}{\sum x_i}}
\end{aligned} \tag{10}$$

Factorization Theorem

Let $X_1, X_2, ..., X_n$ denote a random sample from a distribution $f(x, \theta), \theta \in \Omega$. The statistic $T(x)$ is said to be a sufficient statistic of θ if and only if we can find two nonnegative functions K_1 and K_2 such that

$$\prod_{i=1}^{n} f(x_i; \theta) = K_1(t, \theta) \times K_2(x_1, x_2, ..., x_n)$$

where $K_2(x_1, x_2, ..., x_n)$ does not depend on θ.

Theorem

Let $X_1, X_2, ..., X_n$ denote a random sample from a distribution that has probability distribution $f(x, \theta), \theta \in \Omega$. If a sufficient statistic $T(x)$ of θ exist and if a maximum likelihood estimator $\hat{\theta}$ of θ also exists uniquely then $\hat{\theta}$ is a function of $T(x)$

Example 2 Poisson Distribution

Let $X_1, x_2, ..., X_n$ be a random sample with Poisson pmf and parameter μ

$$f(x,\mu) = \frac{e^{-\mu}\mu^x}{x!}, x = 0, 1, 2, ...$$

Show that the MLE of μ is unbiased, consistent and sufficient statistic estimator.

Solution

- The MLE of μ is $\bar{X}$ therefore

$$E(\bar{X}) = E\left(\frac{\sum X_i}{n}\right) = \frac{1}{n}\sum E(X_i) = \frac{1}{n}n\mu = \mu$$

-

$$Var(\bar{X}) = Var\left(\frac{\sum X_i}{n}\right) = \frac{1}{n^2}\sum Var(X_i) = \frac{1}{n^2}(n\mu) = \frac{\mu}{n}$$

$$\lim_{n\to\infty} Var(\bar{X}) = \lim_{n\to\infty}\frac{\mu}{n} = 0$$

Cont'd

Now

$$\prod_{i=1}^{n} f(x_i, \mu) = \prod_{i=1}^{n} \frac{e^{-\mu}\mu^{x_i}}{x_i!} = \frac{e^{-n\mu}\mu^{\sum x_i}}{\prod x_i!} = \frac{e^{n\mu}\mu^{n\bar{x}}}{\prod x_i!}$$

- Thus $\prod_{i=1}^{n} f(x_i, \mu)$ can be written by a product of two functions $K_1(t, \theta) = e^{-n\mu}\mu^{n\bar{x}}$ which depends on the parameter μ and the MLE $T = \bar{x}$ and $K_2(x_1, x_2, ..., x_n) = \frac{1}{\prod x_i!}$ which depends only on the random sample
- Therefore we conclude that $T = \bar{X}$ is a sufficient statistic estimator.
- Thus the MLE $\bar{X}$ is unbiased, consistent and sufficient statistic estimator of μ

Theorem

Let $X_1, X_2, ..., X_n$ denote a random sample from a distribution $f(x, \theta'), \theta' = (\theta_1, \theta_2, ..., \theta_k)$ Then the statistic $T' = (T_1, T_2, ..., T_k)$ are **joint sufficient statistic** of $\theta' = (\theta_1, \theta_2, ..., \theta_k)$ if and only if

$$L(x, \theta') = \prod_{i=1}^{n} f(x_i, \theta') = K_1(t', \theta') \times K_2(x_1, x_2, ..., x_n)$$

where $K_2(x_1, x_2, ..., x_n)$ does not depend on θ

Example

Let $X_1, X_2, ..., X_n$ denote a random sample from a distribution that is $N(\mu, \sigma^2), -\infty < \mu < \infty, \sigma^2 > 0$. Find the sufficient statistic for μ and σ^2

Solution

The likelihood function of $N(\mu, \sigma^2)$ is obtained as

$$\prod_{i=1}^{n} f(x_i;\ \mu, \sigma^2) = \left(\sqrt{2\pi}\sigma\right)^{-n} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(X_i-\mu)^2}$$

$$= \left(\sqrt{2\pi}\sigma\right)^{-n} e^{-\frac{1}{2\sigma^2}\left[\sum_{i=1}^{n}(X_i-\bar{X})^2+n(\bar{X}-\mu)^2\right]}$$

$$= \left(\sqrt{2\pi}\sigma\right)^{-n} e^{-\frac{1}{2\sigma^2}\left[nS_1^2+n(\bar{X}-\mu)^2\right]}$$

Let $T_1 = \bar{X}, T_2 = S_1^2$. Then, we can write

$$\prod_{i=1}^{n} f(x_i;\ \mu, \sigma^2) = K_1(T_1, T_2;\ \mu, \sigma^2).K_2(X)$$

where $K_1(T_1, T_2, \mu, \sigma^2) = \left(\sqrt{2\pi}\sigma\right)^{-n} e^{-\frac{1}{2\sigma^2}\left[nT_2+n(T_1-\mu)^2\right]}$ and $K_2(X) = 1$.

Therefore, (T_1, T_2) are jointly sufficient statistic of (μ, σ^2).

Exercise

Let $X_1, X_2, ..., X_n$ be a random sample drawn from continuous uniform distribution where $x \in (0, \theta)$. Find the following

- The MLE of θ
- Prove that $Y_n = Maximum(X_1, .., X_n)$ is a sufficient statistic, asymptotically unbiased and consistent estimator of θ
- An unbiased estimator of θ

Note: Solution will be provided after one week.
Make sure you attempt before then

Thank You!