Course: Mathematical statistics

Week 4: Methods Of Point Estimators

Lecturer: Nagulama Moses

Kumi University

April 11, 2025

Outline

1. Consistent Estimation
2. methods of point estimation
   - Method of moments
   - Method of maximum likelihood estimator (MLE)
   - limitation of the MLE

Intended learning outcomes

- Explain the conditions under which an estimator is considered consistent.
- Apply the method of moments to find estimators for parameters of common distributions.
- Solve for parameter estimates using MLE for standard distributions.

Consistent Estimator

if $\hat{\theta}$ is consistent when $\lim_{n\to\infty} P(|\theta - \hat{\theta}| \geq c) = 0$

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

consistency

Example: consider $x_1, x_2, ..., x_n$ being random sample taken from a population with mean μ and varianceσ^2. show that sample mean $\overline{x} = \frac{\Sigma_{i=1}^{n} x_i}{n}$ is a consistent estimator of μ.

solution

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

$$\lim_{n\to\infty} var(\overline{x}) = \lim_{n\to\infty} \frac{\sigma^2}{n}$$

$$= \sigma^2 \lim_{n\to\infty} \frac{1}{n} = 0$$

Exercise

A random sample of size $2n + 1$ is taken from a normal population with μ and σ^2. Determine whether or not the sample mean is a consistent estimator of a population mean.

We are given a sample of size $2n + 1$ from a normal population with mean μ and variance σ^2. The sample mean is

$$\bar{x} = \frac{1}{2n+1} \Sigma_{i=1}^{2n+1} x_i$$

To check if $\bar{x}$ is a consistent estimator of μ we check

Unbiasedness

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

variance goes to zero as sample increases

$$var(\bar{x}) = \frac{\sigma^2}{2n+1} \to 0 as n \to \infty$$

Since the sample mean is unbiased and its variance tends to zero, it

converges to μ as the sample size increases.

Therefore the sample mean $\bar{x}$ is a consistent estimator of μ.

Methods of point estimation

There are two methods

- Method of moments
- Method of maximum likelihood estimate (MLE)

Method of moments

Depending on the number of parameters to be estimated equate the sample moments to population moments and solve the resulting equations for the point estimate

Definition: The k^{th} sample moment $m_k = \frac{\sum_{i=1}^{n} x_i^k}{n}$
the k^{th} population moment $\mu_k = E(x^k)$

1^{st} sample moment is a sample mean

2^{nd} sample moment is $M_2 = \sum_{i=1}^{n} \frac{x_i^2}{n}$

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

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

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

$$= \Sigma_{i=1}^{n}x_i^2 - 2\overline{x}\Sigma_{i=1}^{n}x_i + \Sigma_{i=1}^{n}\overline{x}^2 = \Sigma_{i=1}^{n}x_i^2 - 2n\overline{x}^2 + n\overline{x}^2$$

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

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

$$\mu_1 = E(x) = \mu, \mu_2 = E(x^2)$$

from

$$\sigma^2 = E(x^2) - (E(x))^2$$
$$E(x^2) = \sigma^2 + \mu^2$$

Equate $M_k = \mu_k, k = 1, 2, ...r$

Example: A random variable has a distribution with a pdf

$$f(x) = \begin{cases} (1+k)x^k, 0 \leq x < 1, k > 0 \\ 0, \textit{elsewhere} \end{cases}$$

where k is a parameter for the distribution. The following sample values of x were observed $0.2, 0.4, 0.7, 0.8, 0.8, 0.9, 0.9$

solution

The first population moment $\mathbb{E}[X]$ is derived from the PDF:

$$\mathbb{E}[X] = \int_0^1 x f(x)\, dx = \int_0^1 x(1+k)x^k\, dx = (1+k)\int_0^1 x^{k+1}\, dx.$$

The integral of x^{k+1} over $[0,1]$ is:

$$\int_0^1 x^{k+1}\, dx = \frac{x^{k+2}}{k+2}\Big|_0^1 = \frac{1}{k+2}.$$

Thus:

$$\mathbb{E}[X] = (1+k)\cdot\frac{1}{k+2} = \frac{1+k}{k+2}.$$

The first sample moment is:

$$m_1 = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

Substitute the sample values $0.2, 0.4, 0.7, 0.8, 0.8, 0.9, 0.9$:

$$m_1 = \frac{1}{7}(0.2 + 0.4 + 0.7 + 0.8 + 0.8 + 0.9 + 0.9).$$

Compute the sum:

$$\text{Sum} = 0.2 + 0.4 + 0.7 + 0.8 + 0.8 + 0.9 + 0.9 = 4.7.$$

So:

$$m_1 = \frac{4.7}{7} \approx 0.6714.$$

Equate the population moment to the sample moment:

$$\frac{1+k}{k+2} = m_1.$$

Substitute $m_1 = 0.6714$:

$$\frac{1+k}{k+2} = 0.6714.$$

Multiply through by $k+2$ to eliminate the denominator:

$$1 + k = 0.6714(k+2).$$

Simplify:

$$1 + k = 0.6714k + 1.3428.$$

Rearrange terms to isolate k:

$$1 - 1.3428 = 0.6714k - k.$$

$$-0.3428 = -0.3286k.$$

Solve for k:

$$k = \frac{-0.3428}{-0.3286} \approx 1.043.$$

The method of moments estimate for k is approximately:

$$\hat{k} \approx 1.043.$$

Let $X_1, X_2, \ldots, X_n$ be a random sample from a distribution with the following probability density function:

$$f(x;\theta) = \begin{cases} 2\theta x, & 0 < x < 1 \\ 0, & \text{otherwise} \end{cases}$$

Use the method of moments to find an estimator for θ.

Solution

Find the population moment (theoretical mean)

We calculate the expected value $E[X]$:

$$E[X] = \int_0^1 x \cdot 2\theta x \, dx = 2\theta \int_0^1 x^2 \, dx = 2\theta \cdot \frac{1}{3} = \frac{2\theta}{3}$$

Set the sample moment equal to the population moment
The first sample moment (sample mean) is:

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

Set $\bar{X} = E[X]$:

$$\bar{X} = \frac{2\theta}{3}$$

Solve for θ

$$\theta = \frac{3}{2}\bar{X}$$

So, the **method of moments estimator** for θ is:

$$\boxed{\hat{\theta} = \frac{3}{2}\bar{X}}$$

Let $X_1, X_2, \ldots, X_n$ be a random sample from a Normal distribution with unknown mean μ and unknown variance σ^2, i.e.,

$$X_i \sim \mathcal{N}(\mu, \sigma^2)$$

Use the **Method of Moments** to estimate both μ and σ^2.

Solution

Find the first and second theoretical (population) moments

$$E[X] = \mu$$
$$E[X^2] = \mu^2 + \sigma^2$$

Find the corresponding sample moments
First sample moment (sample mean)

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

Second sample moment

$$M_2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2$$

Equate sample moments to population moments

$$\bar{X} = \mu \quad (1)$$

$$M_2 = \mu^2 + \sigma^2 \quad (2)$$

Substitute equation (1) into equation (2):

$$M_2 = \bar{X}^2 + \sigma^2 \quad \Rightarrow \quad \sigma^2 = M_2 - \bar{X}^2$$

Final Method of Moments Estimators:

$$\hat{\mu} = \bar{X}, \quad \hat{\sigma}^2 = M_2 - \bar{X}^2$$

Method of maximum likelihood estimator (MLE)

Need to get estimates of the parameters that will maximize the likelihood function.

let $x_1, x_2, ..., x_n$ be a random sample of size n from a population having a distribution $f(x, \theta)$ where θ is a single unknown parameter. The likely hood function for the sample values of $x_1, x_2, ..., x_n$ is going to be given by

$$L(\theta) = f(x_1, \theta), f(x_2, \theta), ..., f(x_n, \theta)$$

$$= \pi_{i=1}^{n} f(x_i, \theta)$$

for continuous distribution

$\hat{\theta}$ of θ that maximises $L(\theta)$
$P(X = x, \theta)$ is discrete function

$$L(\theta) = P(X = x_1, \theta).P(X = x_2, \theta)......P(X = x_n, \theta)$$

$$= \pi_{i=1}^{n} P(X = x_i, \theta)$$

Example

A random sample $x_1, x_2, ..., x_n$ is taken from a population having a poison distribution with a parameter λ. find the maximum likelihood estimate $\hat{\lambda}$ for the parameter λ

solution

To find the maximum likelihood estimate $\hat{\lambda}$ for the parameter λ of a Poisson distribution, follow these steps: The probability mass function (PMF) of a Poisson random variable is given by:

$$f(x;\lambda) = \frac{\lambda^x e^{-\lambda}}{x!}, \quad x = 0, 1, 2, \ldots, \lambda > 0.$$

Given a random sample $x_1, x_2, \ldots, x_n$, the goal is to find the MLE $\hat{\lambda}$.
The likelihood function is the joint probability of observing the sample:

$$L(\lambda) = \prod_{i=1}^{n} f(X = x_i, \lambda)$$

$$\frac{e^{-\lambda}\lambda^{x_1}}{x_1!} \cdot \frac{e^{-\lambda}\lambda^{x_2}}{x_2!} \cdots\cdots\cdots \frac{e^{-\lambda}\lambda^{x_n}}{x_n!}$$

$$= e^{-n\lambda}\left(\frac{\lambda^{x_1}}{x_1!} \cdot \frac{\lambda^{x_2}}{x_2!} \cdots\cdots \frac{\lambda^{x_n}}{x_n!}\right)$$

Take the natural logarithm of the likelihood function to obtain the log-likelihood function

$$ln(L(\lambda)) = lne^{-n\lambda} + ln\left(\frac{\lambda^{x_1}}{x_1!}\right) + ln\left(\frac{\lambda^{x_2}}{x_2!}\right) + ... + ln\left(\frac{\lambda^{x_n}}{x_n!}\right)$$

$$ln(L(\lambda)) = -n\lambda + \{x_1 ln\lambda - lnx_1! + x_2 ln\lambda - lnx_2! + ... + x_n ln\lambda - lnx_n!\}$$

$$ln(L(\lambda)) = -n\lambda + \{x_1 ln\lambda - lnx_1! + x_2 ln\lambda - lnx_2! + ... + x_n ln\lambda - lnx_n!\}$$

$$\frac{\partial}{\partial\lambda}(ln(L(\lambda))) = -n + \{\frac{x_1}{\lambda} + \frac{x_2}{\lambda} + ... + \frac{x_n}{\lambda}\}$$

$$0 = -n + \frac{1}{\lambda}\Sigma_{i=1}^{n} x_i$$

$$\frac{1}{\lambda}\Sigma_{i=1}^{n} x_i = n$$

let $\hat{\lambda}$ be the estimate for λ

$$\frac{1}{\hat{\lambda}}\Sigma_{i=1}^{n} x_i = n$$

$$\hat{\lambda} = \frac{\Sigma_{i=1}^{n} x_i}{n} = \bar{x}$$

Example

A random sample $x_1, x_2, ..., x_n$ from an exponential distribution

$$f(x, \lambda) = \lambda e^{-\lambda x}, x > 0, \lambda > 0$$

The likelihood function is:

$$L(\lambda) = \prod_{i=1}^{n} f(x_i; \lambda) = \prod_{i=1}^{n} \lambda e^{-\lambda x_i} = \lambda^n e^{-\lambda \sum_{i=1}^{n} x_i}$$

Take the natural logarithm of the likelihood:

$$\ln L(\lambda) = \ln(\lambda^n e^{-\lambda \sum x_i}) = \ln(\lambda^n) + \ln\left(e^{-\lambda \sum x_i}\right)$$

$$\ln L(\lambda) = n \ln \lambda - \lambda \sum_{i=1}^{n} x_i$$

$$\frac{d}{d\lambda} \ln L(\lambda) = \frac{n}{\lambda} - \sum_{i=1}^{n} x_i$$

$$\frac{n}{\lambda} - \sum x_i = 0 \quad \Rightarrow \quad \frac{n}{\hat{\lambda}} = \sum x_i \quad \Rightarrow \quad \hat{\lambda} = \frac{n}{\sum x_i}$$

$$\boxed{\hat{\lambda} = \frac{1}{\bar{x}}}$$

This means the MLE of λ is the reciprocal of the sample mean.

properties of maximum likelihood estimators

This method of estimation is usually referred by mathematical statisticians because it is often easy to produce estimators with good statistical properties.
provided that a sample size is large the estimators $\hat{\theta}$ obtained by MLE estimate will have the following properties.

(i) $\hat{\theta}$ will be approximately unbiased estimators $\hat{\theta} = \theta$

(ii) the $var(\hat{\theta})$ is nearly as small as the variance that could b obtained with any other estimators

(iii) $\hat{\theta}$ has an approximate normal distribution.

- it is not usually easy to maximise a likelihood function because $\frac{d}{d\theta}L(\theta) = 0$ may be difficult to solve
- sometimes the calculus method to obtain MLE may fail.

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: Interval Estimation