

1. Introduction

Welcome to the series of E-learning modules on Practicals on Estimation by the Method of Maximum Likelihood. In this module we are going to cover the basic principle of method of maximum likelihood-Problems to demonstrate estimation of certain population parameters by the method of maximum likelihood.

By the end of this session, you will be able to:

- Understand the principle of method of maximum likelihood
- Describe the procedure to estimate certain parameters of the population by the method of maximum likelihood

In statistics, maximum-likelihood estimation (MLE) is a method of estimating the [parameters](#) of a statistical model which is generally used in estimation.

When applied to a data set and given a statistical model, maximum-likelihood estimation provides estimates for the model's parameters.

In case the sample observations are independent, the likelihood function happens to be the product of the density functions of the random observations. If x_1, x_2, \dots, x_n are n independent and identically distributed observations, from a population with an unknown parameter θ , then the likelihood function of the random observations denoted by L is equal to $L(x_1, x_2, \dots, x_n, \theta)$ is given by

$L(x_1, x_2, \dots, x_n, \theta)$ is equal to product of $f(x_i, \theta)$

where $f(x_i, \theta)$ is the probability density function of the population.

The method of maximum likelihood is the one in which for a given set of values x_1, x_2, \dots, x_n , an estimator of θ is found that maximizes L . Thus if there exists $\hat{\theta}$, a function of x_1, x_2, \dots, x_n for which L is maximum for variations in θ . Then $\hat{\theta}$ is the maximum likelihood expectation of θ if

$\frac{\partial L}{\partial \theta} = 0$ and $\frac{\partial^2 L}{\partial \theta^2} < 0$

In practice it is often more convenient to work with the logarithm of the likelihood function, called the log-likelihood.

Since $\log L$ attains maximum when L attains maximum the solution of the equation,

$\log L(x_1, x_2, \dots, x_n, \theta)$ is equal to summation $\log f(x_i, \theta)$

$\frac{\partial \log L}{\partial \theta} = 0$ and $\frac{\partial^2 \log L}{\partial \theta^2} < 0$ also gives the maximum likelihood expectation of θ .

2. Illustrative Examples

Illustrative Examples:

Problem 1:

A random variable X takes values zero and one. Obtain on the basis of a random sample of size n , the maximum likelihood expectation for the probability of zero. If the following were the observations, what is the estimate of the probability of zero?

The Observations are: zero, zero, one, zero, one, zero, one, one, one, zero, one, zero, zero, one, one, zero, one, zero, zero, and zero.

Solution:

Let x_1, x_2, \dots, x_n be a random sample of size n drawn from a Bernoulli population. Let p denote the probability of zero. Then $(1 - p)$ is the probability of one. Out of n observations let there be m number of zeroes. Thus there will be $n - m$ number of one's.

Therefore the likelihood function is

L is equal to $L(x_1, x_2, \dots, x_n, p)$ which is equal to product of $f(x_i, p)$, i runs from one to n which is equal to $p^m (1 - p)^{n - m}$
 $\ln L$ is equal to $m \ln p + (n - m) \ln(1 - p)$

Differentiating with respect to ' p ' and equating it to zero

$\frac{\Delta \ln L}{\Delta p}$ is equal to zero implies $m \ln p + (n - m) \ln(1 - p)$ is equal to zero implies $m \ln p + (n - m) \ln(1 - p) = 0$ which implies $p^m (1 - p)^{n - m} = 0$

Therefore estimate of p is number of zeroes divided by total number of observations.

Here n is equal to twenty and m is equal to eleven

Therefore \hat{p} is equal to m/n which is equal to eleven by twenty which is equal to zero point five, five.

Hence maximum likelihood estimate of the probability of zero is zero point five, five.

Problem 2:

Obtain maximum likelihood expectation of θ based on the sample observations for the following function $f(x, \theta)$ is equal to $\frac{1}{\theta} e^{-x/\theta}$, $x > 0$.

The observations are as follows:

ten, three point seven two, twelve point one, zero point seven eight, five point three one, two point nine eight, eight point eight, eight, nine point three.

Solution:

Let x_1, x_2, \dots, x_n be a random sample of size n drawn from the given population with probability density function. Given $f(x, \theta) = \frac{1}{\theta} e^{-x/\theta}$

L is equal to $L(x_1, x_2, \dots, x_n, \theta)$ which is equal to product of $f(x_i, \theta)$, i runs from one to n which is equal to $\frac{1}{\theta^n} e^{-\sum x_i / \theta}$

summation x_i by θ

$\ln L$ is equal to $-\ln \theta - \sum x_i \ln \theta$.

Differentiating with respect to ' θ ' and equating it to zero

$\frac{\partial \ln L}{\partial \theta} = 0$ implies $-\frac{1}{\theta} + \sum \frac{x_i}{\theta^2} = 0$
implies $-\frac{1}{\theta} + \frac{\sum x_i}{\theta^2} = 0$
implies $\theta = \frac{\sum x_i}{n}$ which is equal to \bar{x} .

$\frac{\partial^2 \ln L}{\partial \theta^2} = -\frac{1}{\theta^2} - \frac{2 \sum x_i}{\theta^3}$ which is equal to $-\frac{1}{\theta^2} - \frac{2 \sum x_i}{\theta^3}$ which is less than zero.

Therefore Maximum likelihood estimation of θ is $\hat{\theta}$ which is equal to $\frac{\sum x_i}{n}$ which is equal to $\frac{53.7}{8} = 6.7125$.

3. Examples (Part – 1)

Problem 3:

Find the M. L. E of θ in the following Uniform distribution $U(0, \theta)$

$f(x, \theta)$ is equal to $\frac{1}{\theta}$ for $0 < x < \theta$ using the sample observations

two point five, one point six three, five point two one, zero point eight two, six point three one, four point three four, one point five two, three point six two, four point seven three, three point nine five, five point six three, six point five.

Solution:

Let x_1, x_2, \dots, x_n be a random sample of size n drawn from an Uniform population with probability density function

$f(x, \theta)$ is equal to $\frac{1}{\theta}$ for x varies between zero and θ .

L is equal to $L(x_1, x_2, \dots, x_n, \theta)$ is equal to product of $f(x_i, \theta)$. This is equal to $(\frac{1}{\theta})^n$.

$\log L$ is equal to $-n \log \theta$.

$\frac{\Delta \log L}{\Delta \theta} = 0$ implies $-\frac{n}{\theta} = 0$ which implies $\theta \rightarrow \infty$ or $n = 0$ which is meaningless.

Hence we use the basic principle of the maximum likelihood estimation

maximum likelihood expectation is that value of θ for which L is maximum

L is maximum when $(\frac{1}{\theta})^n$ is maximum. That is when θ is minimum.

That is when θ is minimum.

But x varies between zero and θ . This implies,

$0 < x_1 < \theta, 0 < x_2 < \theta, \dots, 0 < x_n < \theta$

Minimum value of θ is x_n , the maximum of the observations.

Therefore maximum likelihood expectation of θ is $\hat{\theta} = x_n$, which is equal to six point five, the maximum of the observations in the given set of data.

Problem 4:

The following observations were obtained from Uniform (α, β) distribution. Find the maximum likelihood estimates of α and β . The observations are:

minus two, zero point one, minus one point two, four point six, minus two point five, zero, zero point five, two point eight, minus one point six, one point five.

Solution:

The given random sample is drawn from Uniform population with parameters α and β .

Hence when x_i follows Uniform (α, β), the probability density function is given by

$f(x)$ is equal to $\frac{1}{\beta - \alpha}$ for $\alpha \leq x \leq \beta$

beta.

L is equal to L of $(x_1, x_2, \text{etc}, x_n, \alpha, \beta)$ is equal to product of f of (x_i, α, β)
Equal to $(1 - \beta + \alpha)^n$.

$\log L$ is equal to $n \log (\beta - \alpha)$.

$\Delta \log L / \Delta \beta = 0$ implies $n / (\beta - \alpha) = 0$ which implies $\beta - \alpha$ is equal to infinity Which is meaningless.

Hence the usual method of finding maximum likelihood expectation fails. Now we use the basic principle of the maximum likelihood estimation.

Maximum likelihood expectations of α and β are that values for which L is maximum.

L is maximum when $(1 - \beta + \alpha)^n$ is maximum. That is when $(\beta - \alpha)$ is minimum.

That is when $(\beta - \alpha)$ is minimum.

$(\beta - \alpha)$ is minimum when α is maximum and β is minimum.

But $\alpha \leq x_1 \leq x_2 \leq \dots \leq x_n \leq \beta$ implies $\alpha \leq x_1$ and $x_n \leq \beta$.
less than or equal to x_1 less than or equal to x_2 less than or equal to etc till less than or equal to x_n less than or equal to β .

Minimum value of β is x_n , the maximum of the observations and the maximum value of α is x_1 , the minimum of the observations.

Therefore maximum likelihood expectation of α is x_1 β is x_n . Hence in the given data minimum of the observations is minus two point five and maximum of the observations is four point six.

Therefore maximum likelihood estimate of α is x_1 which is minimum of the observations which is equal to minus two point five and maximum likelihood estimate of β is x_n , which is maximum of the observations which is equal to four point six.

4. Examples (Part - 2)

Problem 5:

The following sample is taken from a Normal population $N(\theta, \sigma^2)$. Find the maximum likelihood estimators for θ and σ^2 and hence their respective estimates. The observations are : fifty two point seven, fifty nine, sixty two, sixty two point seven, forty seven, fifty two, fifty five, fifty six, sixty five, and forty two.

Solution:

Let x_1, x_2, \dots, x_n be a random sample of size n drawn from a given population.

L is equal to $L(x_1, x_2, \dots, x_n, \theta)$ is equal to product of $f(x_i, \theta)$

Which is equal to product of $\frac{1}{\sigma \sqrt{2\pi}} \exp\left(-\frac{(x_i - \theta)^2}{2\sigma^2}\right)$

Which is equal to $\frac{1}{\sigma^n (2\pi)^{n/2}} \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \theta)^2\right)$

$\ln L$ is equal to $-\frac{n}{2} \ln \sigma^2 - \frac{n}{2} \ln (2\pi) - \frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \theta)^2$. Maximum likelihood estimates of θ and σ^2 when both are unknown are: $\frac{\partial \ln L}{\partial \theta} = 0$ implies $\sum_{i=1}^n (x_i - \theta) = 0$ implies $\hat{\theta} = \frac{\sum_{i=1}^n x_i}{n}$ which is \bar{x} . Call this as equation one.

$\frac{\partial \ln L}{\partial \sigma^2} = 0$ implies $-\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4} \sum_{i=1}^n (x_i - \theta)^2 = 0$

Implies $-\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4} \sum_{i=1}^n (x_i - \bar{x})^2 = 0$

From equation one, Which implies $\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$

Therefore maximum likelihood expectation of θ is, $\hat{\theta}$ is equal to $\sum_{i=1}^n x_i$ by n , which is equal to \bar{x} .

The maximum likelihood expectation of σ^2 when θ is unknown is $\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$

In the given problem the maximum likelihood estimate of θ is:

For the given problem the maximum likelihood estimate of θ is, $\hat{\theta}$ is equal to $\sum_{i=1}^n x_i$ by n . Which is equal to five hundred fifty three point four by ten which is equal to fifty five point three four and the maximum likelihood estimate of σ^2 when θ is unknown is, $\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$ by n , equal to four hundred and seventy one point four two four by ten which is equal to forty seven point one four two four.

5. Examples (Part - 3)

Problem 6:

From a population with density function, f of x is equal to one by two into 'e' to the power minus modulus of $(x_i - \theta)$, the following ten random observations are drawn.

Ten point seven, two point one three, three point four seven, twelve, five point four, zero point zero seven, minus two, seven point five, minus zero point one, four point four five.

Obtain Maximum likelihood estimate of θ .

Solution:

Let x_1, x_2, \dots, x_n be a random sample of size n drawn from a given population.

L is equal to L of $(x_1, x_2, \dots, x_n, \theta)$ is equal to product of f of (x_i, θ) , i runs from one to n which is equal to product of one by two into 'e' to the power minus modulus of $(x_i - \theta)$

Which is equal to $(1/2)^n$ into 'e' to the power minus summation modulus of $(x_i - \theta)$.

$\ln L$ is equal to \ln of $(1/2)^n$ into two minus summation modulus of $(x_i - \theta)$.

Since the function minus summation modulus of $(x_i - \theta)$

is not differentiable, we use the definition to find the maximum likelihood expectation of θ .

That is, the value of θ for which L is maximum. L is maximum when 'e' to the power minus modulus of $(x_i - \theta)$

is maximum.

L is maximum when summation modulus of $(x_i - \theta)$ is minimum.

L is maximum when $1/n$ into summation modulus of $(x_i - \theta)$ is minimum.

L is maximum when Mean deviation about θ is minimum. But mean deviation about θ is minimum when θ is equal to Median. Therefore the maximum likelihood expectation of θ is $\hat{\theta}$ is equal to Median which can be obtained as follows :

Arranging the given observations in an ascending order we get

Minus two, minus zero point one, zero point zero seven, two point one three, three point four seven, four point four five, five point four, seven point five, ten point seven and twelve.

Median is the value of $(n + 1)/2$ th item is equal to Value of five point fifth item.

Which is equal to value of fifth item plus (zero point five) into (Value of sixth item minus Value of fifth item).

Which is equal to three point four seven plus (zero point five) into (four point four five minus three point four seven) which is equal to three point nine six.

Therefore the maximum likelihood expectation of θ is $\hat{\theta}$ which is

Equal to median, this is equal to three point nine six.

Problem 7:

From a population with density

$f(x, \alpha, \beta)$ is equal to $\alpha x^{\alpha-1} (1+x)^{-(\alpha+1)}$, where x varies between zero and infinity and α is greater than zero.

The following ten random observations are drawn. Obtain Maximum likelihood estimate of α .

The observations are seventy five, sixty eight, fifty five, sixty two, fifty, eighty two, seventy seven, sixty, seventy two and fifty.

Solution:

L is equal to product of $f(x_i, \alpha)$ which is equal to $\alpha^n \prod x_i^{\alpha-1} (1+x_i)^{-(\alpha+1)}$

$\ln L$ equal to $n \ln \alpha + (\alpha-1) \sum \ln x_i - (\alpha+1) \sum \ln(1+x_i)$.

Differentiating with respect to α and then equating to zero we get, $\hat{\alpha}$ is equal to n divided by $\sum \ln(1+x_i) - \sum \ln x_i$.

Therefore the maximum likelihood expectation of α is,

$\hat{\alpha}$ is equal to n by $\sum \ln(1+x_i) - \sum \ln x_i$. This is equal to ten divided by forty one point seven, seven three nine minus forty one point six one seven zero, which is equal to sixty three point seven three four nine.

Here's a summary of our learning in this session where we have :

- Discussed the principle of method of maximum likelihood estimation
- Explained illustrative examples to obtain maximum likelihood expectation of the unknown parameters of the population