Frequently Asked Questions

1. What do you mean by estimation of parameters?

Answer:

Any statistical investigation aims at generalizing from sample to population. Moreover selecting a random sample is essential for drawing valid conclusions about the population. Methods of estimation develops theoretical basis of connection between sample information and population model. This in turn permits inference about the population.

True value of a parameter is an unknown constant that can be correctly ascertained only by an exhaustive study of the population, if indeed that were possible. . Our objective may be to obtain a guess or an estimate of the unknown true value along with the determination of its accuracy. This type of inference is called estimation of parameters.

Estimation is the process by which sample data are used to indicate the value of an unknown quantity in a population.

2. Explain two different ways of estimation in statistical Inference.

Answer:
The estimation is done in two different ways:

- Point Estimation
- Interval Estimation

In Point Estimation the estimated value is given by a single quantity which is the function of a sample observations (that is a statistic). This function is called the estimator of the parameter and the value of the estimator is called an estimate.

In Interval Estimation, an interval within which the parameter is expected to lie is given by using two quantities based on the sample values. This is known as a confidence Interval.

3. Explain briefly point estimation and the role of error in the estimation

Answer:
An objective of point estimation is to compute a single value from the sample data that is likely to be close to the unknown value of the parameter. A statistic intended for estimating a parameter is called a point estimator.

That means when we provide a single numerical estimate of the population parameter (θ) on the basis of the sample information it is known as point estimation.

A point estimate is often insufficient, because it is either right or wrong. If we are told only that a point estimate of average number of students for a particular course is wrong, you do not know how wrong it is, and you cannot be certain of the estimates reliability. If you learn that it is off by only ten students you can accept the closer figure of enrolment of students as a good estimate of future enrolment. Without an assessment in accuracy a single number quoted as an estimate may not serve a very useful purpose We must indicate an estimate of variability in the distribution of an estimator. The standard error provides information about its variability.

Therefore, point estimate is much more useful if it is accompanied by an estimate of the error that might be involved.

4. What are the different Methods of Point Estimation?

Answer:
Point estimation refers to the process of estimating a parameter from a probability distribution, based on observed data from the distribution. There are many methods available for the estimation of the population parameters using Point Estimation method. Some of them are

- Method of maximum Likelihood
- Method of Moments
- Method of minimum variance
- Method of Chi-square
- Bayesian's estimators etc

5. Explain briefly the method of moments.

Answer:
The method of moments is the oldest method of deriving point estimators. It usually produces some asymptotically unbiased estimators, although they may not be the best estimators.

The method was moments were discovered by Karl Pearson. As the name itself suggests in this technique moments are utilized for the estimation of the unknown parameters. This method is based on the principle that the unknown population parameters can be estimated by making use of sample moments. By equating the sample moments to the population raw moments the unknown parameters of the population may be estimated.

6. Write a note on the principle behind the method of moments.

Answer:

The method of moments is based on the moments of population as well as that of sample.

Let $f(x, \theta 1, \theta_2, \ldots, \theta_k)$ is the density function of the population under consideration. To estimate the unknown parameters $\theta_1, \theta_2, \ldots, \theta_k$ if μ_r' denotes the r the moment (about zero) then by definition

$\mu_r' = \int x^r f(x, \theta_1, \theta_2, \ldots, \theta_k)dx$, r=12,…..

That is

$\mu_1' = \int x f(x, \theta 1, \theta_2, \ldots, \theta_k)dx$

$\mu_2' = \int x^2 f(x, \theta 1, \theta_2, \ldots, \theta_k)dx$

$.\mu_{k'} = \int x^k f(x, \theta 1, \theta_2, \ldots, \theta_k)dx$

$\mu_1', \mu_2', \mu_3', \ldots, \mu_k'$ are in general functions of the parameters $\theta_1, \theta_2, \ldots, \theta_k$. Thus the above is a set of k equations involving k unknown parameters $\theta_1, \theta_2, \ldots, \theta_k$.

Now solving the equations $\theta_1, \theta_2, \ldots, \theta_k$ can be written as the functions of $\mu_1', \mu_2', \mu_3', \ldots, \mu_k'$. But in general $\mu_1', \mu_2', \mu_3', \ldots, \mu_k'$ are unknown and hence their estimators by sample moments $m_1', m_2', m_3', \ldots, m_k'$ respectively where $m_{r'} = \sum x_i^r / n$, x1,x2,…xn being sample observations are determined.

In case of frequency distribution of the sample observations r^{th} sample moment $m_r 1$ is given by $m_{r'} = \sum_{fixi}{}^r /N$. Thus the method of moments consistes in equating rth raw moments about the origin in the population to the r th raw moments about the origin in the sample by giving values r=1,2,…. In addition, obtaining various equations containing parameters and solving these equations to obtain the estimate of the parameters.

7. What are the properties of moment estimators?

Answer:
The basic properties of moment estimators are

- Moment estimators are consistent provided the population moments exists
- Moment estimators need not be unbiased
- Under certain conditions moment estimators have an asymptotic normal distribution
- Moment estimators are less efficient than maximum likelihood estimators

8. What do you mean by method of maximum likelihood?

Answer:

In statistics, maximum-likelihood estimation (MLE) is a method of estimating the parameters of a statistical model. When applied to a data set and given a statistical model, maximum-likelihood estimation provides estimates for the model's parameters.

The method of maximum likelihood corresponds to many well-known estimation methods in statistics. For example, one may be interested in the heights of adult female giraffes, but be unable to measure the height of every single giraffe in a population due to cost or time constraints.

Assuming that the heights are normally (Gaussian) distributed with some unknown mean and variance, the mean and variance can be estimated with MLE while only knowing the heights of some sample of the overall population. MLE would accomplish this by taking the mean and variance as parameters and finding particular parametric values that make the observed results the most probable. The method aims at obtaining the values of unknown parameters of the population by maximizing the likelihood function

The principle of maximum likelihood estimation (MLE) originally developed by R.A. Fisher in the 1920s, states that the desired probability distribution is the one that makes the observed data ‘‘most likely, ‘which means that one must seek the value of the parameter vector that maximizes the likelihood function.

9. The method of maximum likelihood provides a basis for many of the techniques. Give reasons.

Answer:

The method of maximum likelihood provides a basis for many of the techniques. The reasons are:

- The method has a good intuitive foundation. The underlying concept is that the best estimate of a parameter is that giving the highest probability that the observed set of measurements will be obtained
- The least-squares method and various approaches to combining errors or calculating weighted averages, etc., can be derived or justified in terms of the maximum likelihood approach
- The method is of sufficient generality that most problems are amenable to a straightforward application of this method, even in cases where other techniques become difficult. Inelegant but conceptually simple approaches often provide useful results where there is no easy alternative

10. What are the applications of the maximum likelihood method?

Answer

Applications of the method

Maximum likelihood estimation is used for a wide range of statistical models, including:

- linear models and generalized linear models
- exploratory and confirmatory factor analysis
- structural equation modelling
- many situations in the context of hypothesis testing and confidence interval formation
- Discrete choice models

These uses arise across applications in widespread set of fields, including:

- communication systems
- psychometrics
- econometrics
- time-delay of arrival (TDOA) in acoustic or electromagnetic detection
- data modelling in nuclear and particle physics
- magnetic resonance imaging
- computational Phylogenetics
- origin/destination and path-choice modelling in transport networks
- geographical satellite-image classification

11. Describe the principle behind the method of maximum likelihood.

Answer:

Suppose there is a sample $x_1, x_2, \ldots, x_n$ of n independent and identically distributed observations, coming from a distribution with an unknown pdf $f(\cdot)$. It is however surmised that the function f belongs to a certain family of distributions $\{ f(\cdot,\theta), \theta \in \Theta \}$, called the parametric model, so that $f = f(\cdot,\theta)$.

The value θ is unknown and is referred to as the "*true value*" of the parameter. It is desirable to find an estimator $\hat{\theta}$ which would be as close to the true value θ as possible. Both the observed variables x_i and the parameter θ can be vectors.

To use the method of maximum likelihood, one first specifies the joint density function for all observations. For an i.i.d. sample, this joint density function is

f(x1,x2,…,xn, θ)= f(x1, θ)f(x2, θ)……f(xn, θ)

Now we look at this function from a different perspective by considering the observed values $x_1, x_2, \ldots, x_n$ to be fixed "parameters" of this function, whereas θ will be the function's variable and allowed to vary freely; this function will be called the likelihood:

L(x1,x2,…,xn, θ)= f(x1,,x2,…,xn, θ)= $\prod_{i=1}^{n} f(xi,\theta)$

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

ln L(x1,x2,…,xn, θ)= $\sum_{i=1}^{n} \ln f(xi,\theta)$

12. Briefly explain the properties of maximum likelihood estimators.

Answer:

A maximum-likelihood estimator is an extremum estimator obtained by maximizing, as a function of θ, the *objective function*

Maximum-likelihood estimators have no optimum properties for finite samples, in the sense that (when evaluated on finite samples) other estimators have greater concentration around the true parameter-value.

However, like other estimation methods, maximum-likelihood estimation possesses a number of attractive limiting properties: As the sample-size increases to infinity, sequences of maximum-likelihood estimators have these properties:

- Consistency: a subsequence of the sequence of MLEs converges in probability to the value being estimated.
- Asymptotic Normality: as the sample size increases, the distribution of the MLE tends to the Gaussian distribution with mean θ and covariance matrix equal to the inverse of the Fisher information matrix.
- Efficiency: i.e., it achieves the Cramér–Rao lower bound when the sample size tends to infinity. This means that no asymptotically unbiased estimator has lower asymptotic mean squared error than the MLE (or other estimators attaining this bound).

13. Explain the applications of the methods of estimation.

Answer:

Either methods of estimation give us the point estimates or interval estimates which are used as parts of other statistical calculations. For example, a point estimates of the standard deviation is used in the calculation of a confidence interval for μ. Estimates of parameters are often used in the formulas for significance testing. Even though point estimates are not as informative as interval estimates, the importance of point estimates lies in the fact that many statistical formulas are based on them.

Nevertheless, there are some instances in which the point estimates obtained using method of maximum likelihood or method of moments are very much useful. Two examples are:

i. Interval estimates of certain fundamental physical constants would be very difficult or inconvenient to work with in calculations. Thus, although quantities such as the gravitational acceleration, g; Avogadro's number, N; and so forth are numbers, which are experimentally determined and thus subject to sampling errors of one sort or another, we normally use point estimates of them rather than interval estimates. Of course, many of these fundamental constants have been estimated with high precision so that errors in their estimates are not significant for many applications.
ii. As you'll see shortly, when we carry out various procedures of statistical inference focusing on one population parameter of greatest interest, the formulas that result may involve the values of other population parameters. In such situations, we can usually obtain adequately accurate results by using point estimates for the parameters of secondary interest in order to derive formulas for an interval estimate of the parameter of greatest interest.
iii. For example, in deriving formulas for interval estimates of the population mean, μ , we require the value of the population standard deviation, σ . Since μ is unknown, it is very unlikely that we'll know the value of σ (though in some instances we might). Rather than backing up one more step and determining an interval estimate for σ , it is more usual to use the available value of s as a point estimate of σ in the formula for the interval estimate of μ .

14. How do you estimate the parameters using interval estimation technique?

Answer:

Assume that we have a sample (*x1,x2,,,xn*) from a given population. All parameters of the population are known except some parameterθ.

We want to determine from the given observations unknown parameter - θ.

In other words we want to determine a number or range of numbers from the observations that can be taken as a value of an unknown parameter θ.

An interval estimation is a process of defining two numbers, between which a population parameter is said to lie. For example, *a* less than x less than *b* is an interval estimate of the population mean μ. It indicates that the population mean is greater than *a* but less than b.

An interval estimate is a range of value used to estimate a population parameter

However, because of the presence of sampling error, sometimes it is more useful to start with the point estimate, and then establish a range of values both above and below the point estimate.

In practice, confidence interval estimates are used more commonly by far than point estimates. Nevertheless, since point estimates are used in certain important ways in statistics, and carry with them some important concepts and terms, we need to look at them briefly in this course.

15. When the Maximum-likelihood estimators can lack asymptotic normality and can be inconsistent?

Answer:

Maximum-likelihood estimators can lack asymptotic normality and can be inconsistent if there is a failure of one (or more) of the below regularity conditions:

Estimate on boundary

Sometimes the maximum likelihood estimate lies on the boundary of the set of possible parameters, or (if the boundary is not, strictly speaking, allowed) the likelihood gets larger and larger as the parameter approaches the boundary. Standard asymptotic theory needs the assumption that the true parameter value lies away from the boundary. If we have enough data, the maximum likelihood estimate will keep away from the boundary too.

Data boundary parameter-dependent

For the theory to apply in a simple way, the set of data values which has positive probability (or positive probability density) should not depend on the unknown parameter. A simple example where such parameter-dependence does hold is the case of estimating θ from a set of independent identically distributed when the common distribution is uniform on the range $(0,\theta)$. For estimation purposes the relevant range of θ is such that θ cannot be less than the largest observation. Because the interval $(0,\theta)$ is not compact, there exists no maximum for the likelihood function:

Nuisance parameters

For maximum likelihood estimations, a model may have a number of nuisance parameters. For the asymptotic behaviour outlined to hold, the number of nuisance parameters should not increase with the number of observations (the sample size).

Increasing information

For the asymptotic to hold in cases where the assumption of independent identically distributed observations does not hold, a basic requirement is that the amount of information in the data increases indefinitely as the sample size increases. Such a requirement may not be met if either there is too much dependence in the data (for example, if new observations are essentially identical to existing observations), or if new independent observations are subject to an increasing observation error.