Frequently Asked Questions

1. Define an estimator or a statistic.

Answer:
An estimator is a rule, usually expressed as a formula that tells us how to calculate an estimate based on information in the sample.

An estimator is a sample characteristic and is a function of the sample observations. Any sample statistic that is used to estimate a population parameter is called an estimator. The sample mean $\overline{x}$ can be an estimator of the population mean μ., the sample proportion can be an estimator p of the population proportion P. We can also use a sample range as an estimator of the population Range.

2. Name two types of estimators.

Answer:

Two types of estimators are Point Estimators and Interval estimators.

Point Estimators yield single-valued results, although this includes the possibility of single vector-valued results and results that can be expressed as a single function.

This is in contrast to an Interval Estimator, where the result would be a range of plausible values (or vectors or functions). Interval Estimator gives an interval which is expected to contain the value of the parameter.

3. What do you mean by an unbiased estimator? Explain.

Answer:
A sample mean is an unbiased estimator of a population mean if the mean of the sampling distribution of sample means taken from the sample population is equal to the population means itself. Any estimator is an unbiased estimator if on average it tends to assume values above the population parameter being estimated as frequently and to the same extent as it tends to assume values that are below the population parameter being estimated.

Suppose t is an estimator of the population mean μ, then it is said to be unbiased for the parameter μ, if Expected value of (t) = μ.

Unbiased estimators determine the tendency, on the average, for the statistics to assume values close to the parameter of interest.

4. Explain briefly consistent estimators.

Answer:
A statistic is a consistent estimator of a population parameter if as the sample size increases it becomes almost certain that the value of statistic comes very close to the value of the population parameter. If an estimator is consistent it becomes more reliable with large

samples. Suppose you are in confusion that whether to increase the sample size to get more information about a population parameter find out first whether your statistic is a consistent estimator. If it is not, you will waste your time and money by taking large sample.

The sample mean and sample proportions are consistent estimators, since from their formulas as **n** get big, the standard errors gets small.

5. State and explain Central Limit Theorem.

Answer:
The theorem which explains this sort of relationship between the shape of the population distribution and the sampling distribution of the mean is known as the Central Limit Theorem. It ensures that the sampling distribution of the mean approaches Normal distribution as the sample size increases. In formal terms we may say that the Central Limit Theorem states that "The distribution of means of random sample taken from a population having mean μ and finite variance σ^2 approaches Normal distribution with mean μ and variance σ^2/n as n goes to infinity".

The significance of the central limit theorem lies in the fact that it permits us to use sample estimators to make inference about the population parameters without knowing anything about the shape of the frequency distribution of that population other than what we can get from the sample.

6. Write a note on efficiency of an estimator.

Answer:
Another desirable property of a good estimator is that it be efficient. Efficiency refers to the size of the Standard error of the statistic. If we compare two statistics from a sample of same size and try to decide which one is more efficient estimator, we would pick a statistic that has a smaller standard error or standard deviation of sampling distribution. It makes sense that a statistic of smaller standard error (with less variation) will have more chance of producing an estimate nearer to the population parameter under consideration. Hence, an estimator of least variance is said to be more efficient among the class of all unbiased estimators.

Example: The median is an unbiased estimator of μ when the sample distribution is normally distributed; but its standard error is 1.25 greater than that of the sample mean, so the sample mean is a more efficient estimator than the median.

7. What cautions to be taken while selecting a best estimator?

Answer:
A given sample statistic is not always the best estimator of its analogous population parameter. Consider a symmetrically distributed population in which the values of the median and then mean coincide. In this instance a sample mean would be an unbiased and consistent estimator of a population median. Because as the sample size increase the value of the sample mean would tend to come very close to the population median. And sample mean would be more efficient estimator of the population median than the sample median itself because in large samples the sample mean has a smaller standard error than the

sample median. At the same time the sample median in a symmetrically distributed population would be an unbiased and consistent estimator of the population mean but not the most efficient estimator because in large samples its standard error is larger than that of the sample mean. In this case the best estimator would be a sample mean than sample median. Hence, there may be more than one estimator for the same population parameter but we should be careful while selecting the best estimator.

8. Write a note on Bias in estimators.

Answer:
The amount of bias is B = Estimated value – true value of the parameter.

As a working rule the effect of bias on the accuracy of the value of an estimator is negligible if the Bias is less than one tenth of the Standard deviation of the estimate.

That is $B < 0.1/\sigma$ where B is the absolute value of the bias. Anyhow if the sample is large enough we can be confident that B/σ will not exceed 0.1.

In sample survey theory it is necessary to consider biased estimators for two reasons:

1) In some of the most common problems , particularly in the estimation of

   ratios, estimators that are otherwise convenient and suitable are found to be biased

2) Even with estimators that are unbiased in probability sampling errors of

   measurement and nonresponsive may produce biases in the numbers that we are able to compute from the data.

9. What is the role of a sample mean in Statistical estimation?

Answer:

In general, we use statistics as a means of characterizing the nature of some sample on the basis of a few key indicators.

The first indicator is known as The Sample Mean:

- The mean quantity in some sample represents the average value or the most probable value in the sample.
  - A sample mean is calculated by summing up the individual measurements and dividing by the number of measurements, usually denoted as N.
  - All samples can be characterized by a mean value regardless of the shape of the distribution.

- Also according to Central Limit Theorem: The distribution of means of random sample taken from a population having mean μ and finite variance σ^2 approaches Normal distribution with mean μ and variance σ^2/n as n goes to infinity.

10. Briefly explain Standard error.

Answer

The standard error is the standard deviation of an estimator. The standard error is important because it is used to compute other measures, like confidence intervals and margins of error.

The variability of a statistic is measured by its standard error.

The values of population parameters are often unknown, making it impossible to compute the standard deviation of a statistic. When this occurs, use the standard error.

The standard error is computed from known sample statistics, and it provides an unbiased estimate of the standard deviation of the statistic.

Hence Standard error is the standard deviation of the sampling distribution of the estimator.

For instance: The Standard error of sample mean $\bar{y}$ is:

S.D ($\bar{y}$) = $\sqrt{v(\bar{y})}$.

11. What do you mean by a sampling distribution?

Answer:

It is possible to draw more than one sample from the same population and the value of an estimator will in general vary from sample to sample. For example, the average value in a sample is a statistic. The average values in more than one sample, drawn from the same population, will not necessarily be equal. The probability distribution of these estimators is known as a sampling distribution.

The sampling distribution describes probabilities associated with an estimator when a random sample is drawn from a population.

The sampling distribution is the probability distribution or probability density function of an estimator.

12. What is the role of variability in estimation?

Answer:

The ***variability*** of a statistic is determined by the spread of its sampling distribution. In general, larger samples will have smaller variability. This is because as the sample size increases, the chance of observing extreme values decreases and the observed values for the statistic will group more closely around the mean of the sampling distribution. Furthermore, if the population size is significantly larger than the sample size, then the size of the population will not affect the variability of the sampling distribution (i.e., a sample of size 100 from a population of size 100,000 will have the same variability as a sample of size 100 from a population of size 1,000,000).

In general, the dispersion is a more important quantity than the sample mean. The dispersion represents the range of the data about the mean value. Understanding the role of dispersion is the most critical aspect of understanding and interpreting statistical sampling data.

The calculation of dispersion in a distribution is very important because it represents a uniform way to determine probabilities and therefore to determine if some event in the data is expected (i.e. probable) or is significantly different than expected (i.e. improbable).

13. Distinguish between accuracy and precision of an estimator.

Answer:
Due to the difficulty of ensuring that no unsuspected bias enters into estimates we will usually speak of the precision of an estimator instead of the accuracy. Accuracy refers to the size of the deviation from the true mean μ whereas precision refers to the size of the deviations from the (estimated) mean $\bar{y}$ obtained by repeated applications of the sampling procedure.

14. What do you mean by Mean Square error of an estimator?

Answer:
In order to compare a biased estimator with an unbiased estimator or two estimators with different amounts of bias, a useful criterion is the Mean Square Error (M.S.E) of the estimator measured from the population value that is being estimated. Hence Mean Square Error is an average of the squared deviations of actual value of the parameter from its estimated value.

$$\text{M.S.E}(\hat{\mu}) = E(\hat{\mu} - \mu)^2$$

15. What are the two important points to be considered to get a best estimator of the population parameter?

Answer:

1. The sampling distribution of the point estimator should be centered over the true value of the parameter to be estimated.

2. The spread (as measured by the variance) of the sampling distribution should be as small as possible.