# Frequently Asked Questions

1. Write a note on estimation of parameters.

**Answer:**

Estimation is the process of determining a likely value for a variable in the survey population, based on information collected from the sample. Researchers are usually interested in looking at estimates of many statistics—totals, averages and proportions being the most frequent—for different variables.

For example, a sample survey could be used to produce any of the following statistics: estimates for the proportion of smokers among all people aged 15 to 24 in the population; the average earnings of men and women with a university degree; or the total number of cars possessed by the whole survey population.

2. List the desirable properties of estimators.

**Answer:**

Some desirable properties for estimators are:

- Unbiased or nearly unbiased.
- Have a low MSE (mean square error) or a low variance when the estimator is unbiased. [MSE measures how far the estimate is from the parameter of interest whereas variance measures how far the estimate is from the mean of that estimate. Thus, when an estimator is unbiased, its MSE is the same as its variance.]
- Robust - so your answer does not fluctuate too much with respect to extreme values.

3. Name the situations where we may have to estimate of population parameters.

**Answer:**

Estimation of Average household income in a country is important during tough economic times or assessment of taxes and property values.

We might want to estimate the total pyrite content of the rocks

We might want to estimate the total market size for electrical cars.

We might want to estimate failure rate of a certain electronic component.

In the above situations, we have to go for the estimation of either population mean or total.

4. Under SRSWR prove that sample mean is an unbiased estimator of population mean.

**Answer:**

**Proof:** Let $y_1, y_2, \ldots, y_n$ be SRSWR sample of size n drawn from a population containing N units then we know that population mean

$$\overline{Y} = \frac{\sum_{i=1}^{N} Y_i}{N}$$ and sample mean $$\overline{y} = \frac{\sum_{i=1}^{n} y_i}{n}$$

Consider $E(\overline{y}) = E\left(\frac{\sum_{i=1}^{n} y_i}{n}\right) = \frac{\sum_{i=1}^{n} E(y_i)}{n}$ $\longrightarrow$ (1)

$$E(y_i) = \sum_{i} y_i p_i$$

Consider $E(y_i)$, since each $y_i$ takes values $Y_i$ with probability 1/N we have $E(y_i) = \sum_{i=1}^{N} Y_i \bullet \frac{1}{N} = \overline{Y}$

By substituting the above in equation (1 ), we get

$$E(\overline{y}) = \frac{\sum_{i=1}^{n} \overline{Y}}{n} = \frac{n\overline{Y}}{n} = \overline{Y}$$

Hence, $E(\overline{y}) = \overline{Y}$ i.e. sample mean is unbiased estimator of population mean.

5. Obtain an unbiased estimator of the population total under SRSWR.

**Answer:**

**Proof:** let $y_1, y_2, \ldots, y_n$ be SRSWR sample of size n drawn from a population of size N.

A population total $Y = \sum_{i=1}^{N} Y_i = N\overline{Y}$

$\hat{Y} = N\hat{\overline{Y}} = N\overline{y}$ [because $E(\overline{y}) = \overline{Y} \Rightarrow \hat{\overline{Y}} = \overline{y}$]

6. Show that $E(\bar{y}) = \bar{Y}$ , population mean under SRSWOR.

**Answer:**

**Proof:** Let $y_1, y_2, \ldots, y_n$ be SRSWOR sample of size n drawn from a population containing N units then we know that population mean

$$\bar{Y} = \frac{\sum_{i=1}^{N} Y_i}{N}$$ and sample mean $$\bar{y} = \frac{\sum_{i=1}^{n} y_i}{n}$$

Consider $E(\bar{y}) = E\left(\frac{\sum_{i=1}^{n} y_i}{n}\right) = \frac{\sum_{i=1}^{n} E(y_i)}{n}$ $\longrightarrow$ 1

$$E(y_i) = \sum y_i p_i$$

Consider $E(y_i)$, since each $y_i$ takes values $Y_i$ with probability 1/N we have $E(y_i) = \sum_{i=1}^{N} Y_i \bullet \frac{1}{N} = \bar{Y}$

By substituting the above in equation 1, we get

$$E(\bar{y}) = \frac{\sum_{i=1}^{n} \bar{Y}}{n} = \frac{n\bar{Y}}{n} = \bar{Y}$$

Hence, $E(\bar{y}) = \bar{Y}$ i.e. sample mean is unbiased estimator of population mean.

7. Deduce an unbiased estimator of the population total under SRSWOR

**Answer:**

**Proof:** let $y_1, y_2 \ldots y_n$ be SRSWOR sample of size n drawn from a population of size N.

A population total $Y = \sum_{i=1}^{N} Y_i = N\bar{Y}$

$\hat{Y} = N\hat{\bar{Y}} = N\bar{y}$ [because $E(\bar{y}) = \bar{Y} \Rightarrow \hat{\bar{Y}} = \bar{y}$ ]

8. Explain Simple Random sampling of attributes

**Answer:**

A qualitative characteristic, which cannot be measured quantitatively, is known as an attribute. For e.g. Honesty, beauty, intelligence etc. Quite often, we come across the situations where it may not be possible to measure the characteristic under study but it may be possible to classify the whole population into various classes with respect to the attributes under study.

We consider the cases where the population is divided into two classes say C and C dash with respect to an attribute. Such a classification is termed as dichotomous classification. Hence, any sampling unit in the population may be placed in class C or C dash respectively, depending whether it possess or does not possess the given attribute. In the study of attributes, we are interested in the estimate of total number of proportion.

9. Give some examples where we use the estimation of population proportions.

**Answer:**

- Proportion of defective items in a large consignment of items
- Proportion of the literates or the breadwinners in a town
- Proportion of sales a particular product accounts for
- Proportion of the viewing public of a particular program
- Proportion of trees with a diameter of eleven inches or more

10. Under Simple Random sampling for attributes show that Population mean is equal to population proportion and the sample mean is the sample proportion

**Answer:**

Let us suppose that a population with N units $Y_1$, $Y_2$, …., $Y_N$ is classified into two disjoint and exhaustive classes C and C' respectively with respect to a given attribute. Let the number of individuals in the classes C and C' be A and a' respectively such that A+A' = N
Then
P = The Proportion of units possessing the given attribute = A/N
Q= The Proportion of units does not possess the given attribute = A'/N = 1-P.
In statistical language, P and Q are the proportion of successes and failures respectively in the population.
Let us consider SRS sample of size n from this population. Let 'a' be the number of units in the sample possessing the given attribute
Then p = Proportion of sampled units possessing the given attribute = a/n
and q= Proportion of sampled units which do not possess the given attribute = 1 – p
With the $i^{th}$ sampling unit, let us associate a variate $Y_i$ ( i=1,2,…,N) defined as follows
$Y_i$ = 1, if it belongs to the class C i.e., if it possess the given attribute
And $Y_i$ = 0, if it belongs to the class C' i.e., if it does not possess the given attribute.
Similarly let us associate a variable $y_i$ ( i=1,2,….,n) with the $i^{th}$ sample unit defined as follows.

$y_i$ = 1 , if the $i^{th}$ sampled unit possess the given attribute

And $Y_i = 0$ , if the $i^{th}$ sampled unit does not possess the given attribute.
Then $\sum Y_i = A$, the number of units in the population possessing the given attribute

$\sum y_i = a$, the number of sample units possessing the given attribute

Thus $\overline{Y} = \frac{\sum_{i=1}^{N} Y_i}{N} = \frac{A}{N} = P$ and $\overline{y} = \frac{\sum_{i=1}^{n} y_i}{n} = \frac{a}{n} = p$

11. Under Simple Random Sampling prove that Sample proportion 'p' is an unbiased estimate of the population Proportion 'P'.

**Answer:**
Proof: we know that in simple Random sampling the sample mean provides an unbiased estimate of the population mean.

$$E(\overline{y}) = \overline{Y}$$

$$\overline{Y} = \frac{\sum_{i=1}^{N} Y_i}{N} = \frac{A}{N} = P \quad \text{-- -----(1)}$$

and $\overline{y} = \frac{\sum_{i=1}^{n} y_i}{n} = \frac{a}{n} = p$ -----------------(2)

But from Equations (1 ) and (2) we have
$\overline{Y} = P$ *and* $\overline{y}$ = p
*Hence* E (p) = P

12. Show that an unbiased estimate of the under units possessing an attribute in the population, A which we denote by $\widehat{A}$ = Np.

**Answer:**

We have E (p) =P

Which implies E (Np) = N E(p) = NP = A

Hence, an unbiased estimate of A which we denote by $\hat{A}$ = Np

13. From a population of size 484 a simple random sample of size 9 is drawn. Estimate an average amount of money

| Accounts | Y1 | Y2 | Y3 | Y4 | Y5 | Y6 | Y7 | Y8 | Y9 |
|---|---|---|---|---|---|---|---|---|---|
| Dollars | 33.5 | 32 | 52 | 43 | 40 | 41 | 45 | 42.5 | 39 |

**Answer:**

Estimate of average amount of money is the average of the sample. That is a sample mean

$$\bar{y} = \frac{\sum_{i=1}^{n} y_i}{n} = \frac{368}{9} = 40.89$$

Hence, estimated average amount is equal to 40.89 dollars

14. At the end of every school year, the state administers a reading test to a Simple Random Sample Drawn Without Replacement from a population of 20,000 third graders. This year, the test was administered to 36 students selected via simple random sampling. The test score from each sampled student is shown below:

50, 55, 60, 62, 62, 65, 67, 67, 70, 70, 70, 70, 72, 72, 73, 73, 75, 75, 75, 78, 78, 78, 78, 80, 80, 80, 82, 82, 85, 85, 85, 88, 88, 90, 90, 90

Using sample data, estimate the mean reading achievement level in the population.

**Answer:**

Identify a sample statistic. Since we are trying to estimate a population mean, we choose the sample mean as the sample statistic. The sample mean is:

$\bar{x}$ = Σ ( $x_i$ ) / n

$\bar{x}$ = ( 50 + 55 + 60 + ... + 90 + 90 + 90 ) / 36 = 75

Therefore, based on data from the simple random sample, we estimate that the mean reading achievement level in the population is equal to 75.

15. A simple random sample of size 100 is drawn from a population of students of size 300 who have two wheelers. Estimate the proportion of students in a college who possess two wheelers.

**Answer:**

Given n = 100 and N = 300
A fraction of students in a college who have two wheelers
Xi = 0 if she/ he do not possess two wheeler
Xi = 1 if she/ he possess a two wheeler

Student: 1,2,3, ,4, ……………………………., 99,100
Xi : 0,1,1,………………………………...,1,1,1
$\sum Xi = 65$
Estimate of the proportion of students in a college who have two wheelers

$$\hat{P} = \bar{x} = \frac{\sum_{i=1}^{n} x_i}{100} = \frac{65}{100} = 0.65$$

65% of the students have two wheelers in a college.