(iii) Slope/Mathematical Method

Mathematically, the price elasticity of demand is calculated by using the concept of differential calculus. By differential calculus, the slope of a demand curve at any point is calculated by the formula $\frac{dP}{dQ}$.

Where,

P – Price of the commodity

Q – Quantity demanded of the commodity.

If the slope (dP/dQ) is reversed i.e., the reciprocal of the slope (dQ/dP) is taken and multiplied with the ratio (P/Q) then, we get price elasticity of demand (ε_P).

i.e.,

$$\frac{dP}{dQ} = \frac{1}{dP/dQ}$$

$$\therefore \quad \varepsilon_P = \frac{dQ}{dP} \cdot \frac{P}{Q}$$

(iv) Total Outlay Method

This method is also called as 'Total Revenue Method'. Marshall provided the easiest method of ascertaining the nature of demand i.e., whether the demand is elastic or not. According to him, it can be found out that demand is elastic or not by examining the change/variation in the total outlay of the consumer or total revenue of the seller corresponding to the variations in the price of the product. The total revenue or total outlay is ascertained by the formula,

Total revenue/Total outlay = Price × Quantity purchased or sold

This can also be written as,

Total revenue = Price × Quantity sold

Total outlay = Price × Quantity purchased.

4.5 LEONTIEF'S METHOD OF DETERMINING DEMAND CURVE FROM TIME SERIES DATA AND THEIR LIMITATIONS

26. Explain Leontief's method of determining demand curve from time series data and its limitations.

Answer : **Model Paper-II, Q12(a)**

Leontief's Method

Leontief's method of determining demand curve from time series data is based on the following assumptions,

Every market transaction denotes the intersection of rapid supply and demand curves, which tend to shift from one position to another at times. This means that the range at which the curves tend to shift from time to time should also be studied, along with the determination of demand and supply elasticities.

2. The variations in demand and supply curves are independent of each other. This means that the shape (elasticities) of the curve stays unaffected by the shifting of them, which means that a variation in the shifting of the demand curve to the right is more likely to be related with the variation of the shifting of supply curve to the left as to the right.

3. Every demand and supply curve consists of constant elasticity, which implies that the supply and demand curves should be straight lines when plotted on a double logarithmic scale.

Let X_t and Y_t represent the consumption and price logarithms of a product respectively, at time t $(t = 1,2,3, \ldots, n)$,

We get,

$$\text{Demand Curve} = X_t = \eta_1 Y_t + P_t \qquad \ldots (1)$$

$$\text{Supply Curve} = X_t = \eta_2 Y_t + Q_t \qquad \ldots (2)$$

Where η_1 and η_2 denote demand elasticity and supply elasticity respectively and P_t and Q_t are two random variables distributed independently with,

$$E(P_t) = 0 = E(Q_t) \qquad \ldots (*)$$

It can be observed in (1) and (2), that X_t is taken for supply as well as consumption, as market equilibrium has,

$$d = S \Rightarrow \log d = \log S = X_t \text{ (Say)}$$

(1) and (2) can also be written as,

$$X_t - \eta_1 Y_t = P_t \qquad \ldots (3)$$

and $X_t - \eta_2 Y_t = Q_t \qquad \ldots (4)$

Multiply (3) and (4), we get,

$$X_t^2 + \eta_1 \eta_2 Y_t^2 - (\eta_1 + \eta_2) Y_t X_t = P_t Q_t \qquad \ldots (5)$$

As P_t and Q_t are independently distributed with,

$$E(P_t) = 0 \text{ and } E(Q_t) = 0,$$

$$\text{Cov}(P_t, Q_t) = 0 \Rightarrow E(P_t, Q_t) - E(P_t).E(Q_t) = 0$$

By using eq. (*), we get,

$$E(P_t, Q_t) = 0 \qquad \ldots (**)$$

By adding time t and using eq (**), we get,

$$\Sigma X_t^2 + \eta_1 \eta_2 \Sigma Y_t^2 - (\eta_1 + \eta_2) \Sigma Y_t X_t = 0 \qquad \ldots (6)$$

Equation (6) cannot be solved with two unknown quantities of η_1 and η_2 that's why the time range t : [1, n] is divided into two equal halves,

$$t_1 : \left[1, \frac{n}{2}\right] \text{ and } t_2 : \left[\frac{n}{2} + 1, n\right]$$

Adding (5) w.r.t. time for each half separately and using eq. (**), we get,

$$\left.\begin{aligned} &\sum_{t=1}^{n/2} X_t^2 + \eta_1\eta_2 \sum_{t=1}^{n/2} Y_t^2 - (\eta_1+\eta_2)\sum_{t=1}^{n/2} Y_t X_t = 0 \\ \text{and } &\sum_{t=\frac{n}{2}+1}^{n} X_t^2 + \eta_1\eta_2 \sum_{t=\frac{n}{2}+1}^{n} Y_t^2 - (\eta_1+\eta_2)\sum_{t=\frac{n}{2}+1}^{n} Y_t X_t = 0 \end{aligned}\right\} \ldots (7)$$

Equation (7) can be solved for η_1 and η_2 and ΣX_t^2, ΣY_t^2 and $\Sigma X_t Y_t$ can be calculated from the time series data.

Limitations

The following are the limitations of Leontief's method from statistical and economic point of view,

1. The assumption that the variation in shifting of the demand and supply curves are independent of each other and that the Cournot-Marshall demand and supply curve may shift in any possible direction, violates the principle of general theory of equilibrium's fundamental principle, which states that "the demand of any one commodity is a function not only of its price but also of all other prices".
2. The assumption of Leontief's method that the demand and supply curves are simultaneous is not a valid assumption in case of agricultural commodity. For example, on an average, low (high) price for a commodity in a specific year is related with low (high) production of the commodity in the following year.
3. The assumption which states that demand elasticity and supply elasticity are constant might not stand true.

Criticism

The Leontief's method received criticism for the method which seems statistically defective to calculate $\eta_1 + \eta_2$. He splits up his series into the ellipse of the two curves formed from the mathematical solution are not similar and the respective axes are not parallel to each other. In addition to this, there is a noticeable difference between the first and the second half period, which shows that the data is not homogeneous. It is in such a way that every period must be studied separately.

4.6 PIGOUS METHOD OF DETERMINING DEMAND CURVE FROM TIME SERIES DATA AND THEIR LIMITATIONS

Q27. Explain Pigou's method of determining demand curve from time series data and its limitations.

Answer :

Pigou's Method

Pigou's method of determining demand curve from time series data is based on the following assumptions,

1. The first assumption of Pigou is that, at each time interval the demand curve has a smooth appearance, which implies that at each time interval, the demand curve has constant elasticity.

$$d = cp^{-\alpha} \Rightarrow \log d = \log c - \alpha \log p$$

i.e., $\log d = \log c + a \log p$; $(a = -\alpha)$

$$\therefore \quad X = aY + b$$

Where X are Y are the logarithms of consumption and price respectively.

2. The shifting of the demand curve over various time periods is steady and the extent to which the curves have shifted equal for every two successive time intervals.

It can also be said that the extent of the shifting of the demand curve is such that the distance between the i^{th} position $(i+1)^{th}$ position (on a logarithmic scale) is equal to the distance between $(i+1)^{th}$ position and $(i+2)^{th}$ position.

From the above assumptions, at different time points,

$t = 1, 2, \ldots$, we have,

$$\left.\begin{aligned} X_1 &= aY_1 + b \\ X_2 &= aY_2 + b + r \\ &\vdots \quad\quad \vdots \\ X_i &= aY_i + b + (i-1)r \\ X_{i+2} &= aY_{i+1} + b + ir \\ X_{i+2} &= aX_{i+2} + b + (i+1)r \\ &\vdots \quad\quad \vdots \end{aligned}\right\}$$

According to the above equation, we have,

$$(X_{i+1} - X_i) - a(Y_{i+1} - Y_i) = (X_{i+2} - X_{i+1}) - a(Y_{i+2} - Y_{i+1})$$

$$\therefore \quad a = \frac{X_{i+2} - 2X_{i+1} + X_i}{Y_{i+2} - 2Y_{i+1} + Y_i} = a_i \text{ (say) } (i = 1, 2, \ldots) \quad \ldots(1)$$

The steps involved in Pigou's method are as follows,

Step-1: A table of logarithms consisting of time-series values of price (Y) and consumption i.e., log p and log d.

Step-2: Calculate a_i (i = 1, 2, ...) from eq. (1). As demand is a diminishing price's functions, if the value of 'a_i' at any interval is said to be positive, then it cannot be considered as a measure of demand elasticity for appropriate set of times. But, if it is said to be negative, then it can be considered as a measure of the elasticity.

Step-3: Every negative a_i can be considered as an observation on the unknown elasticity belonging to the demand curve, if the negative values of a_i's are exceeding the positive values and if they are closely grouped into a given value and also if the data is not suspected otherwise.

Limitations

The following are the limitations of Pigou's method,

1. The assumption that the demand curve for a commodity is $x = f(p, t)$

   Where x denotes the quality of the demand commodity, p denotes the commodity's price and t denotes time. It implies that the prices of commodities which are closely related to that specific commodity have zero or no impact on the commodity or all the effecting factors are considered are ineffective while studying the variations occurring in x due to the variations occurring in y. But this assumption does not stand true as it is not possible to put a stop on all the other factors without consideration.

2. The method proves to be a complete failure if there are three collinear price-quantity points in three successive sets of observations. This means that, Pigou's method is not applicable to collinear successive observations but, it is applicable to functions which are non-linear and change in directions.

3. This method is not useful in some cases where the underlying theoretical demand curves tends to change its motion's direction.

4.7 PARETO LAW OF INCOME

Q28. Explain in detail Pareto's law of income distribution.

Answer :

During 19th and 20th century Vilfredo Pareto proposed the law of distribution of income. He is the first analyst who have conducted the study from the statistical perspective for solving the problem of income distribution among a group of citizens. This interpretation of his law of income distribution was based on the empirical study of the income data of different countries at different times.

The distribution of income in a stable economy is derived approximately by the following formula,

$$P = A(x - a)^{-v}$$

Where,

P = Number of people having income x or greater

a = Lowest income at which the curve starts

A and V = parameters

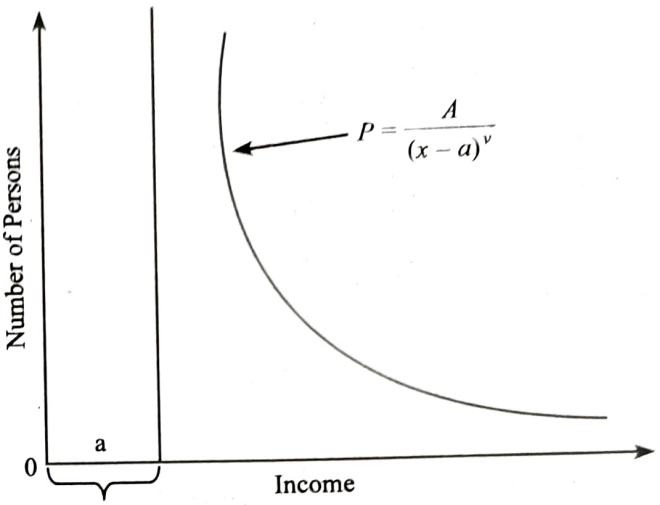


Figure: Pareto Curve

The curve which is derived with the help of above formula is extremely asymmetrical, similar to a hyperbola.

The Pareto curve becomes,

$P = A x^{-v}$ when origin is shifted to $x = a$

Logarithm of both sides are taken as,

$\log P = \log A - v \log x$

Therefore, the graph will be a straight line with its slope equal to –v when Pareto curve is graphically represented on a double logarithmic scale.

Some of the important consideration of Pareto's law of income distribution are as follows,

1. Generally, simplified form of Pareto curve is used due to lack of information regarding number of persons with less incomes.

2. Pareto curve is best fitted for higher incomes as compared to lower incomes.

3. According to Pareto value of 'v' varies from 1.2 to 1.9 in several countries. Thus, the value of 'v' can be taken as 1.5 on an average.

4. If $\log P = \log A - v. \log x$ is differentiated.

   We get,

$$\frac{d \log P}{d \log x} = -v$$

Therefore, V is interpreted as the income elasticity of P.

5. If $P = Ax^{-v}$ is differentiated, with reference to x, we get,

$$dP = -Avx^{-v-1}\,dx$$

Hence, corresponding to an increment dx in income of class with lower limit x, the relative increment in P would be derived by,

$$\frac{dP}{P} = -\frac{v}{x}\,dx$$

6. Pareto's law received various criticism by scientific investigators for its rigid and uncompromising form.

Formulation of Problem of Income Distribution

The problem of distribution of income can be formulated as,

(i) The form of frequency function $\phi(x)$ is determined for total income distribution of society members from poorest to the richest.

(ii) Checking whether $\phi(x)$ is governed by the type of society from which income is derived or is its form is unavoidable.

(iii) Checking any deductive reason be assigned for the form of $\phi(x)$.

Q41. Explain the criterion for testing the consistency of index numbers.

Answer : Model Paper-II, Q12(b)

The various tests of consistency of index numbers are as follows,

1. Unit Test

According to unit test, the formula of index number should be independent of the units under which prices and quantities are quoted. All formulae satisfy this test except simple aggregative test.

2. Time Reversal Test

Time reversal test is basically used for checking whether the selected method would work for both forward and backward. According to this test, the formula should give exact ratio when we compare one point with the another i.e., for example,

$$P_{01} = \frac{1}{P_{01}} \text{ or } P_{01} \times P_{10} = 1$$

$$Q_{01} \times Q_{10} = 1$$

Only two methods, Laspeyre's and Paasche's do not satisfy the time reversal test. Besides these two, the other methods of index numbers satisfy the time reversal method.

Fishers index method satisfies time reversal test.

3. Factor Reversal Test

According to factor reversal test, when change in price is multiplied with change in quantity, it should give total change in value. That is, if the price of a commodity is increased by 3 times and its quantity has also increased by 4 times, then the total change in value would be 12 times than that of the former value. Thus, the formula for entire commodity price and quantity is p_0 and p_1 and q_0 and q_1 for base and current year is:

$$P_{01} \times Q_{01} = \frac{\Sigma p_1 q_1}{\Sigma p_0 q_0}$$

Except Fisher's ideal index, no other method satisfies the factor reversal test.

4. Circular Test

According to circular test, if P_{ab} is the price index whose base year is 'a' and 'b' is the period, P_{bc} is the price index with base year 'b' and 'c' is the period and P_{ca} is the price index whose base year is 'c' and 'a' is the period then,

$$P_{ab} \times P_{bc} \times P_{ca} = 1$$

The following methods satisfy the circular test,

(a) Simple geometric mean of price relatives

(b) Simple aggregative index

(c) Weighted aggregative index.

Apart from these methods, no other method of price index satisfies the circular test.

4.15 BASE SHIFTING OF INDEX

Q49. What is base shifting? Explain it with an illustration.

Answer :

Base Shifting

Base shifting refers to the process of shifting a base period of an index. It helps in calculating the in new base. As the old base gets outdated, it is required to shift or change such old base to new-base.

The base shifting can be calculated by using the following formula,

$$\text{Base Shifting} = \frac{\text{Current year's old index number}}{\text{New base year's old index number}} \times 100$$

Illustration

From the following information of index numbers calculate base shifting.

Consider 2001 base year of old index number and 2004 as new base year.

Years	2001	2002	2003	2004	2005
Index Numbers	100	120	150	180	225

Solution :

The calculation of index number based on new base year is,

$$= \frac{\text{Current year's old index number}}{\text{New base year's old index number}} \times 100$$

Years	Old Index Numbers (Base Year 2001 = 100)	New Index Numbers (Base Year 2004) (2004 = 180)
2001	100	$\frac{100}{180} \times 100 = 55.55$
2002	120	$\frac{120}{180} \times 100 = 66.66$
2003	150	$\frac{150}{180} \times 100 = 83.33$
2004	180	$\frac{180}{180} \times 100 = 100$
2005	225	$\frac{225}{180} \times 100 = 125$

Q53. What do you mean by deflating? Explain with an example.

Answer :

Deflating

Deflating refers to the process of making allowances for the impact of changing prices. A rise in price level means a reduction in the purchasing power of money.

Example

The price of rice in 2010 was ₹ 28 per kg but in 2013 the price increased to ₹ 45 per kg. It means a person who could buy 1 kg rice for ₹ 28 in 2010 would be able to buy only half kg rice in 2013 as their income level remains same. During the period of inflation, the purchasing power of money is the reciprocal of the price index. This reciprocal relationship can be shown in the form of formula as follows,

$$\text{Purchasing power of money} = \frac{1}{\text{Price index}}$$

value of money decreases with the increase in prices wage workers or salaried people show more interest in real wage rather n money wage. Real waged can be obtained by wing the following formulae,

$$\text{Real wage} = \frac{\text{Money wage}}{\text{price index}} \times 100$$

$$\text{Real wage or income index No.} = \frac{\text{Index of money wages}}{\text{Consumer price index}}$$

PART-B

ESSAY QUESTIONS WITH SOLUTIONS

4.1 DEMAND ANALYSIS – INTRODUCTION

Q8. Define Demand Analysis. Discuss the decisions made under demand analysis.

Answer :

Demand Analysis

Demand analysis is very important in taking effective decisions, relating to the various business activities. further development of the business by taking decisions through the analysis of various factors, affecting the dema factors include price of the product, level of expenditure, product quality and customer service. The demand for a p be analyzed and estimated by determining the influence of each factor on the sale of the product.

Demand analysis helps in determining and formulating the optimum operational policies for the firm and als on effective way to respond immediately for the most competitive policies of the competitors so as to sustain the con

Decisions Under Demand Analysis

Demand analysis helps in taking various business decisions related to the following,

(i) Sales forecasting

(ii) Pricing of product

(iii) Distribution and marketing

(iv) Product decisions

(v) Financial aspects.

(i) Sales Forecasting

The demand estimation for a product plays an important role in forecasting the sales of a product. There exists relation between sales and demand for a product. If the demand increases then the sales will also increase and if the de a product reduces then its sales will also reduce.

(ii) Pricing of Product

Demand analysis plays a prominent role in taking decisions related to pricing policies. It acts as a base for decisions i.e., on the basis of the market demand the firms can fix the prices of their products. If the demand is more t usually fix high prices, otherwise they will fix lower prices in order to make use of the market opportunities. If the elasticity is more i.e., highly elastic then the firm will fix high price for that produced.

(iii) Distribution and Marketing

It plays an important role in taking decisions related to marketing and distributing the product.

(iv) Product Decisions

It is also important in taking product decisions. Firms can decide the levels of production process of a proc analyzing the demand for a particular product in the market. If the demand is more i.e., demand is highly elastic then t can go for producing more in order to meet the demand. Otherwise it would go for other alternative. Firms always prod products according to the demand for the product. It helps in reducing the costs of holding inventory.

(v) Financial Aspects

It is also important to take into consideration the decision related to the financial aspects. If the demand is high th firm is required to allocate more financial resources for that particular product or else the firm would distribute its resou produce other products in order to avoid the wastage of financial resources.

In case of business functions, the demand analysis play a prominent role. If the price elasticity of demand is '0 the firm need not reduce the prices as there is no effect of price change on the sales of the product or quality demanded managers should analyze the effect of factors that affect demand for a product in taking the decisions. Otherwise, it would in the huge loss for the firm.

Q12. State and explain the law of demand with the help of a graph.

Answer :

Law of Demand

The law of demand explains a consumer's behaviour, in demanding a commodity in relation to the variations in its prices. It expresses a functional relationship between two variables of demand relation i.e., the price and the quantity demanded of a commodity. It also states the inverse relation between the quantity demanded and the price of the commodity. This makes the law of demand one of the significant laws in economics.

Statement of Law of Demand

The law of demand states that, other things remaining constant, the higher the price of the commodity, the lower is the demand and lower the price, higher is the quantity demanded.

Assumptions of the Law of Demand

The following are the assumptions of law of demand,

(i) It assumes that, consumer's income tastes and preferences remains unchanged throughout the operation of law of demand.

(ii) It assumes that there will be no change in the fads, fashions and latest trends.

(iii) It assumes that, prices of the related goods remain unchanged or are equal.

(iv) It assumes that, consumer does not expect any changes in future prices and in his income.

(v) It assumes that, the demographic factors like age, composition and gender composition remains constant.

(vi) It assumes that, the climatic and weather conditions do not change.

(vii) It assumes that there will be no change in the Government policies like level of taxation, fiscal policy etc.

Therefore, it is concluded that, the law of demand mainly considers the changes in price of the commodity by assuming other things remains unchanged.

Explanation of Law of Demand

The conventional law of demand is given by,

$$Q_x = f(P_x)$$

Where,

Q_x = Quantity demanded of a commodity x

P_x = Price of the commodity x

The inverse relation between demand and price is given by the law of demand as,

$$D \propto \frac{1}{P}$$

Where,

D – Demand for a commodity

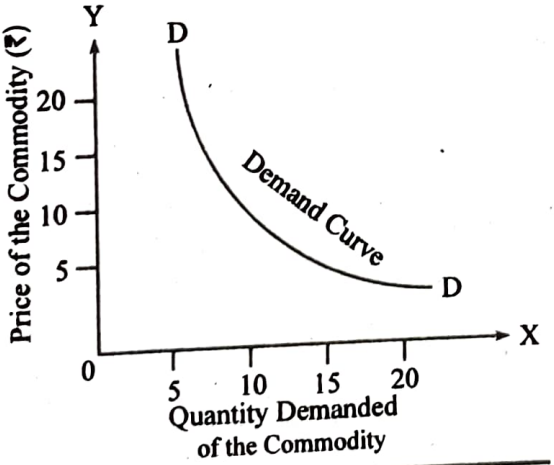
P – Price of the commodity.

Law of demand generally operates due to variations in substitution effect and income of the consumer.

Demand Curve Under Law of Demand

A demand curve is a diagrammatical or a graphical representation of a demand schedule which shows the inverse relationship between price and quantity of commodity. However, the demand curve has negative or downward slope which can be seen in the following figure of imaginary values i.e., price and quantity demanded of a commodity. A demand curve is drawn by taking the price of the commodity on x-axis and the quantity demanded on y-axis. The table below shows a demand curve for a demand schedule.

Price (₹)	Quantity (kgs)
5	15
8	14
10	12
12	10
15	8
20	5



PART-B
ESSAY QUESTIONS WITH SOLUTIONS

1.1 SAMPLE SURVEYS

1.1.1 Concepts of Population, Sample, Sampling Unit, Parameter, Statistic, Sample Frame and Standard Error

Q12. Discuss in brief about the following,

(a) Population

(b) Sample

(c) Sample unit

(d) Parameters.

Answer :

(a) Population

Galton defines 'Population' as "totality of individual observations about which inferences are to be made, existing anywhere in the world, or within a definitely specified area limited by space and time".

When the term 'Population' is used in statistics, it corresponds to the observation of all the individuals that are used in giving conclusion at a specific time. The term 'Population' does not refer only to the human or animal population, instead it also refers to the biostatistical population. There are two different types of populations. They are,

(i) Finite population

(ii) Infinite population.

(i) Finite Population

A population whose values can be theoretically observed as it has fixed number is called a finite population. The size of the finite population is limited.

Example: Number of students in a school, number of rice plants in a field etc.

(ii) Infinite Population

A population whose values cannot be theoretically observed as it does not have a fixed number is called an infinite population. The size of the infinite population is unlimited.

Example: Number of White Blood Cells (WBC) in the human body.

The size of the population is always very large, even if it is finite or infinite. So, the collection of data from each of the individuals is very difficult. Because of this, samples are selected from the population and based on their results conclusions are made regarding that population.

(b) Sample

In biostatistical studies, the data collected depends on the observation of several individuals of a population. Here, sampling can also be performed i.e., few individuals can be selected from a population to represent the complete population. These selected individuals are together called as 'sample'.

Random samples is a sample selected from a population in such a way that each individual in the population has a fair chance of getting selected.

(c) Sample Unit

It can be defined as the distinguishable and observable part of the population which is used for gathering information.

Example

1. In a university (population) survey, each student can be a sample unit.
2. In a crop (population) estimation survey, a specific part of the land can be a sample unit.

1.1.2 Principal Steps in Sample Surveys

Q14. Discuss the principal steps involved in sample surveys.

Answer : Model Paper-I, Q9(a)

Principal Steps in a Sample Survey

The planning and execution of the sample survey incorporates the following steps.

1. Purpose and Scope of Survey

The purpose and scope of survey must be clearly defined so that the investigator has a clear picture about the type of data to be collected and the processing techniques to be applied. If the purpose of enquiry is not clearly mentioned then the investigator is more likely to collect some irrelevant information and may even miss some useful information, thereby resulting in wastage of resources.

2. Population to be Sampled must be Defined

There must not be any ambiguity in the population from which the samples are to be collected. This population must be clearly defined. However, the population (i.e., sampled population) is quite different from the target population.

3. Sampling Units

The sampling units are extracted from the sampled population. In other words, the sampled population is divided into certain sampling units based on the observations to be made. The sampling units must be specific, stable, relevant and free from ambiguity.

4. Type of Data to be Collected

The nature, purpose and scope of survey must be considered while making a decision about the type of data to be collected. In addition to this, the amount of time and money to be spent along with the desired level of accuracy in the final result must also be considered.

5. Questionnaire or Schedule

After making a decision about the type of data to be collected, the next step is to prepare a questionnaire or a schedule for requesting the information from the respondent.

6. Information Collection Method

The next step after preparation of a questionnaire is to select an appropriate method for data collection. Generally there are two methods available. They are,

(a) Direct personal investigation

(b) Mailed questionnaire method.

The selection of method depends upon the purpose of enquiry, budget, time available, and the desired level of accuracy.

7. **Non-respondents**

If the information from the sampled unit cannot be directly obtained in situations like for example, the respondent is unavailable even after multiple visits or the mailed questionnaire is not answered by all the respondents, then the replies received from some users are of no use to the investigator. Thus, certain efforts must be made so as to minimize the non-response.

8. **Sampling Design**

There are multiple number of sampling designs available such as simple random sampling, stratified random sampling etc. Selection of appropriate sampling design must be of top priority in planning and execution of sample survey. For selection of this sampling design, the purpose of survey, nature of sampled population, expenses involved, time to be consumed, etc., must be taken into consideration. The sampling design must be chosen with a great care because an ill-designed sampling plan will often result in an ambiguity.

9. **Organization of Field Work**

Field work must be properly organized so as to obtain reliable and efficient results in a sample survey. Apart from the field work organization, the personnel employed for conducting the survey must be skilled enough to handle the problems associated with the survey. Further, an adequate and a regular supervisory check must be carried out on field work at regular intervals.

10. **Pretest or Pilot Survey**

Before initiating the actual survey, a pilot survey must be carried out so as to examine the working of questionnaires and the field methods that are used for obtaining population information. This pilot survey helps in cost estimation of the actual survey, elimination of the faults and defects, improving questionnaire etc.

11. **Analysis of Data**

The next step after the execution of a sample survey is the analysis step. In this step, all the data obtained from the execution of sample survey is analyzed, scrutinized, edited, tabulated, and concluded.

Precautions to be Taken while Collecting the Sample

1. The sample must be drawn in a scientific way.
2. The sampling design must be carefully chosen.
3. The size of the sample must be adequate enough to describe the sampled population.

1.1.3 Need for Sampling, Census Versus Sample Surveys

Q15. Discuss in brief about the need for sampling.

Answer :

Need for Sampling

Sampling is used in practice for different reasons which are as follows,

(i) It saves both time and money.

(ii) It allows more precise measurements because trained and experienced investigators are hired to conduct sample survey.

(iii) It allows the estimation of the sampling errors. However, it helps to provide information regarding certain population characteristics.

(iv) It is used in the following cases,

(a) If a population contains infinite number of items.

(b) If a test involves elimination of item which is under study.

Q16. Discuss in detail about census versus sampling surveys.

Answer :

The term 'survey' refers to the investigation of opinions and behaviours of a particular group of people, entity, business or some other activities. In statistics, the survey about human activities such as business, economics or social sciences can be efficiently carried out using either of the following two methods

(i) The census method or complete enumeration.

(ii) The sample method or partial enumeration.

(i) Census Method or Complete Enumeration

This method deals with gathering of information for all individual units. A unit might represent a household, industry, land, river, store, supermarkets or other similar units in any specified geographical region to be covered under study. All these units together constitutes *'population'* or the '*universe*'.

Example

Information about the income of each individual need to be collected for the purpose of income-tax assessment, a list of all voters has to be prepared for election purposes, applications of all the candidates have to be scrutinized for the purpose of recruitment.

Advantages of Census Method

1. The data obtained through census survey is more reliable and accurate as the information is gathered from each individual item that exists in the universe.

2. Census enumeration involves extensive study about the items. It brings out details about the complexities associated with each and every unit in the universe.

For example, population census gathers information not only about the number of individuals, but also collects information about their ages, marital status, education, income, employment, etc.

Disadvantages of Census Method

1. It demands much time, workers, administrative personnel and money.
2. It does not produce immediate results. In fact, by the time results are available, the conditions might have changed entirely.
3. Generally, this method is not practically useful. It might help huge organizations and government agencies who have great deal of money, power and various resources to carry out the survey.

(ii) Sample Method or Partial Enumeration

Business surveys usually do not require information to be gathered from each and every element that exists in the universe. It can involve certain elements from among the total elements based on a particular procedure, and form a sample. Therefore, the procedure of gathering information from the elements constituting the sample is known as '*Sample survey*'.

Here, the enquiry is based on a small part which acts as representative of the entire population ensuring a fairly reliable estimate of the characteristics of the universe. The main goal of a sample enquiry as compared to census enquiry is to collect information regarding the characteristics of population to a maximum extent, with minimal while minimizing, cost, time and effort. It is also used to set the limits of accuracy for these estimates.

Advantages of Sample Method

The advantages of carrying out a sample survey over a complete enumeration or census survey are as follows.

1. **Saving in Time and Labour**

It involves less amount of time and labour for its collection and processing. Since, only a part of the population is required to be inspected and examined,

2. **Administrative Convenience**

Less number of employees are required for the administration of sample survey.

3. **Enquiry is Made at Lower Costs**

Sample survey usually results in reduction of cost in terms of money and manpower. The total cost of the sample survey is expected to be much lower than that of a census survey.

4. **Sometimes the only Method Possible**

When the investigation involves destruction of material e.g., the strength of a bullet, the life of a lamp etc., sample survey is the only practical way of assessing the quality.

5. **Detailed Enquiry**

When small data is to be surveyed, it is possible to have greater precision and in-depth study, because more detailed information can be obtained from a small group of respondents.

6. **Results are More Reliable**

Conclusions and results obtained by sampling enumeration are generally more reliable than those obtained by census enumeration.

7. **More Scientific**

The method of sampling is scientific and is independent of expediency or tradition. There is an easy follow-up in case of nonresponse or incomplete response.

Disadvantages of Sample Method

1. If a sample survey is not thoroughly planned and executed, it may lead to imprecise and misleading results.
2. An efficient sampling scheme requires the services of qualified, skilled and experienced staff, better supervision, more sophisticated equipments and statistical techniques for planning and execution of the survey as well as for collection, processing and analysis of the sampled data. If any of these are absent, the results of the survey may not be reliable.
3. Sampling method is exposed to personal biases and prejudices of the investigator especially with regard to choice of technique and drawing of sampling units.
4. It may fail to indicate the true characteristics of the population. When the size of the sample is inadequate,
5. Sample survey is not useful in the situations where all the units of population need to be observed. So, complete enumeration is more efficient when the population is small or when time and money are not considered.

1.1.4 Sampling and Non-sampling Errors, Sources and Treatment of Non-sampling Errors

Q17. Write about sampling and non-sampling errors.

Answer : Model Paper-II, Q9(a)

Sampling Errors

When the sample is used to make conclusions about the population, then the errors can occur. These errors are called sampling errors.

It is also defined as the difference between the result of the census of the whole population. This error can occur as the estimation depends only on small part of population and not on the complete population.

There are two different types of sampling errors,

- ❖ Biased errors
- ❖ Unbiased errors.

Non-Sampling Errors

When the data is observed and processed while performing complete enumeration or sample survey, few errors can occur. These errors are called as non-sampling errors. This type of errors can occur only when the biostatistical data is being acquired, recorded or tabulated.

Several reasons due to which non-sampling errors can occur are listed below.

1. If the definition of population is not appropriate.
2. If the investigator has his own bias.
3. If the definitions of all the variables is not appropriate.
4. If the statistical methods used are not appropriate.
5. If errors occur while processing data.
6. If errors are committed when the tabulated results are presented and printed.
7. If errors arise due to irresponsibility.

Example: When units are not covered completely.

Q18. Discuss about the sources and treatment of non-sampling errors.

Answer : *(V.V.Imp)*

Sources of Non-sampling Errors

The sources of non sampling errors are,

1. Errors in planning or definition
2. Errors in response
3. Errors in non response biases
4. Errors in coverage
5. Errors in compilation
6. Errors in publication.

1. Errors in Planning or Definition

Planning mainly focuses on aims or objectives of the survey that are clearly stated in the planning phase. These objectives are converted into the following,

(a) A set of definitions

(b) A set of specifications.

Non sampling errors in planning occur due to the following reasons,

(i) Incomplete and inconsistent data specification.

(ii) Improper designing of questionnaires.

(iii) Errors in recording the observed measurements.

(iv) Errors in location of units.

(v) Insufficient skilled investigators and less number of supervisory staff.

2. Errors in Response

These errors occur while providing the responses given by the people (respondents) of certain category. Some of the other reasons are discussed below,

(i) Accidental Errors

These errors occur when a person misunderstands a question and gives irrelevant response to it. This irrelevant information which is unintentionally provided by the respondent leads to accidental errors.

(ii) Prestige Bias Errors

These errors are related to the prestige of a person who is interviewed. For instance, a person may provide incorrect information regarding his/her education, income, age etc., inorder to maintain their pride in the society.

(iii) Self-interest Errors

These errors occur when some people give wrong information inorder to satisfy their self interests. For instance, a person may overrate his/her salary and underrate expenses.

(iv) Biasing Errors due to the Interviewer

An interviewer's beliefs, prejudices and way of questioning or recording can influence the responses and its accuracy.

(v) Failure Error (Respondents Fail to Remember)

'Recall' is one of the most common source of error which is used for collecting information.

The questions in the survey are related to the past events and a responder may not remember the information and recall it at that particular instant.

3. Errors in Non Response Biases

These errors are usually found in house to house survey in the following cases,

(i) Respondent is unavailable.

(ii) Respondent refuses to answer certain questions.

(iii) Respondent is unable to answer each and every question.

So, due to this non response, population belonging to a certain category is excluded and therefore it results in some bias.

4. Errors in Coverage

These errors occur when the objectives of the survey are not clearly mentioned which results in the following,

(i) Includes certain topics which are not necessary.

(ii) Excludes the topics which are needed.

Example

A survey is conducted to find the number of people belonging to a age group of 10 to 30. However, if area and time are not defined errors occur.

5. Errors in Compilation

These errors occur during the data processing operations such as editing, coding, tabulating, summarizing the actual observation in the survey. They are key sources for the occurrence of compilation errors. However, these type of errors can be reduced by conducting data verification and checking data consistency.

6. Errors in Publication

These errors occur during the presentation and printing of calculated results. The two main sources for such errors are,

(i) Mechanics of publication-proofing errors.

(ii) Failing to identify the limitations of the statistics during the organizational survey.

Treatment for Non-sampling Errors

Methods of Reducing/Controlling Non-sampling Error

1. Employing qualified and trained personnel for the survey.
2. Providing adequate supervisory checks on the field work.
3. Pretesting or conducting a pilot survey.
4. Thorough editing and scrutiny of the results.
5. Effective checking of all the steps in the processing and analysis of data.
6. More effective follow up of non response cases.

1.1.5 Advantages and Limitations of Sampling

Q19. Discuss in detail about sampling along with advantages and limitations.

Answer :

Sampling

According to Kothari, sampling can be defined as "The selection of some part of an aggregate or totality on the basis of which a judgement or inference about the aggregate or totality is made".

Advantages/Merits of Sampling

For answer refer Unit-I, Page No. 9, Q.No. 16, Topic: Advantages of Sample Method.

Limitations/Demerits of Sampling

For answer refer Unit-I, Page No. 9, Q.No. 16, Topic: Disadvantages of Sample Method.

1.2 SAMPLE

1.2.1 Types of Sampling: Subjective, Probability and Mixed Sampling Methods

Q20. Discuss in detail about the different types of sampling.

Answer :

The different types of sampling includes,

(a) Subjective sampling

(b) Probability sampling

(c) Mixed sampling.

(a) Subjective Sampling

This method is also called *judgement* or *deliberate* or purposive sampling. In this method selection of individual units is completely based on the judgement of investigator. Results obtained by purposive sampling majority depends upon the experience and capability of the sampling person. Only those individual units which the investigator feels to be appropriate are selected . So the question of probability or chance is completely eliminated. For instance, if 10 employee have to be selected from a company of 100 employees, inorder to investigate the work status of the company, then the investigation choose those employees who are the good representatives in his/her perspective. Hence, valuable results can be obtained if the investigator is well experienced, highly skilled and proper sampling is applied.

(b) Probability Sampling

This scientific method allows the selection of samples through some laws of chance of probability theory. The basis of probability sampling is that, the population contains the members having equal and known probability for selection. It supports the following types of sampling,

(a) Sampling which is based on equal chance of selecting a unit.

(b) Sampling which is based on unequal probabilities.

(c) Sampling where the selection probability of a unit is equals to the sample size.

(c) Mixed Sampling

This method involves both, the samples that are chosen according to some laws of chance and fixed sampling rule (i.e., no probability assignments). Therefore it is essential to choose right method for selection of sample, depending on the nature of the sample. Some of the types of mixed sampling are listed below,

1. Stratified random sampling
2. Simple cluster sampling
3. Multistage cluster sampling
4. Quata sampling
5. Simple random sampling
6. Quasi random sampling
7. Systematic sampling
8. Area sampling
9. Multistage sampling.

Q31. In simple random sampling without replacement (srswor), the sample mean is an unbiased estimate of the population mean i.e., $E(\bar{y}_n) = \bar{Y}_N$.

Solution : *(Imp)*

In simple random sampling without replacement (srswor), the sample mean is given by,

$$(\bar{y}_n) = \frac{1}{n}\sum_{i=1}^{n} a_i Y_i \qquad \text{Where } a_i = \begin{cases} 1, \text{ when the sample contains } i^{th} \text{ unit} \\ 0, \text{ when the sample donot contains } i^{th} \text{ unit} \end{cases}$$

Then, $$E(\bar{y}_n) = E\left[\frac{1}{n}\sum_{i=1}^{n} a_i Y_i\right]$$

$$= \frac{1}{n}\sum_{i=1}^{n} E(a_i) Y_i$$

$$E(a_i) = 1 \times P(a_i = 1) + 0 \times P(a_i = 0)$$

$$= 1\frac{n}{N} + 0\left(1 - \frac{n}{N}\right)$$

$$= \frac{n}{N}$$

$$E(\bar{y}_n) = \frac{1}{n}\sum_{i=1}^{N} \frac{n}{N} Y_i \qquad \left[\because E(a_i) = \frac{n}{N}\right]$$

$$= \frac{1}{N}\sum_{i=1}^{N} Y_i$$

$$\therefore E(\bar{y}_n) = \bar{Y}_N$$

Hence proved.

Q32. In srswor, the sample mean square is an unbiased estimate of the population mean square i.e., $E(s^2) = S^2$.

Solution : *(V.Imp)*

$$s^2 = \frac{1}{n-1}\left[\sum_{i=1}^{n} y_i^2 - n\bar{y}_n^2\right]$$

$$= \frac{1}{n-1}\left[\sum_{i=1}^{n} y_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n} y_i\right)^2\right] \qquad \left[\because \bar{y}_n = \sum_{i=1}^{n} y_i\right]$$

$$= \frac{1}{n-1}\left[\sum_{i=1}^{n} y_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n} y_i^2 + \sum_{i\neq j=1}^{n} y_i y_j\right)\right] \qquad \left[\because \left(\sum_{i=1}^{n} y_i\right)^2 = \sum_{i=1}^{n} y_i^2 + \sum_{i\neq j=1}^{n} y_i y_j\right]$$

$$= \frac{1}{n-1}\left[\sum_{i=1}^{n} y_i^2 - \frac{1}{n}\sum_{i=1}^{n} y_i^2 - \frac{1}{n}\sum_{i\neq j=1}^{n} y_i y_j\right]$$

$$= \frac{1}{n-1}\left[\sum_{i=1}^{n} y_i^2\left(1 - \frac{1}{n}\right) - \frac{1}{n}\sum_{i\neq j=1}^{n} y_i y_j\right]$$

$$= \frac{1}{n-1}\sum_{i=1}^{n} y_i^2\left(1 - \frac{1}{n}\right) - \frac{1}{n(n-1)}\sum_{i\neq j=1}^{n} y_i y_j$$

$$= \frac{1}{n-1}\sum_{i=1}^{n} y_i^2\left(\frac{n-1}{n}\right) - \frac{1}{n(n-1)}\sum_{i\neq j=1}^{n} y_i y_j$$

$$= \frac{1}{n}\sum_{i=1}^{n} y_i^2 - \frac{1}{n(n-1)}\sum_{i\neq j=1}^{n} y_i y_j$$

$$\text{i.e., } s^2 = \frac{1}{n-1}\left[\sum_{i=1}^{n}\left[y_i^2 - n\bar{y}_n^2\right] = \frac{1}{n}\sum_{i=1}^{n} y_i^2 - \frac{1}{n(n-1)}\sum_{i\neq j=1}^{n} y_i y_j\right] \quad \ldots (a)$$

Therefore,

$$E(s^2) = \frac{1}{n}E\left(\sum_{i=1}^{n} y_i^2\right) - \frac{1}{n(n-1)}E\left(\sum_{i\neq j=1}^{n} y_i y_j\right) \quad \ldots(1)$$

But we know that,

$$E\left(\sum_{i=1}^{n} y_i^2\right) = E\left[\sum_{i=1}^{N} a_i Y_i^2\right]$$

$$= \sum_{i=1}^{N} E(a_i) Y_i^2$$

$$\therefore E\left(\sum_{i=1}^{n} y_i^2\right) = \frac{n}{N}\sum_{i=1}^{N} Y_i^2 \qquad \left[\because E(a_i) = \frac{n}{N}\right] \quad \ldots(2)$$

Also,

$$E\left[\sum_{i\neq j=1}^{n} y_i y_j\right] = E\left[\sum_{i\neq j=1}^{n} a_i a_j Y_i Y_j\right]$$

$$\Rightarrow \quad E\left[\sum_{i\neq j=1}^{n} y_i y_j\right] = \sum_{i\neq j=1}^{N} E(a_i a_j) Y_i Y_j$$

Where, $E(a_i a_j) = 1 \times P(a_i a_j = 1) + 0 \times P(a_i a_j = 0)$

$= P[(a_i = 1) \cap (a_j = 1)]$

$= P(a_j = 1) \times P(a_j = 1 | a_i = 1)$

$$E(a_i a_j) = \frac{n}{N} \times \frac{n-1}{N-1}$$

(Since, $P(a_i = 1)$ means probability that the i^{th} unit is included in the sample of size n which is equal to $\frac{n}{N}$ and P[$1/a_i = 1$] means probability that the j^{th} unit is included in the sample such that the i^{th} unit is already present in the sample whi equals to $\frac{n-1}{N-1}$).

Now, substituting equation (4) in equation (3)

i.e., $$E\left(\sum_{i\neq j=1}^{n} y_i y_j\right) = \sum_{i\neq j=1}^{N} \frac{n}{N} \times \frac{n-1}{N-1} Y_i Y_j \quad \ldots(5)$$

Now, substituting equations (2) and (5) in equation (1), we get,

$$E(s^2) = \frac{1}{n}\left[\frac{n}{N}\sum_{i=1}^{N} Y_i^2\right] - \frac{1}{n(n-1)}\left[\sum_{i\neq j=1}^{N} \frac{n}{N} \times \frac{n-1}{N-1} Y_i Y_j\right]$$

$$= \frac{1}{N}\sum_{i=1}^{N} Y_i^2 - \frac{1}{n(n-1)} \times \frac{n(n-1)}{N(N-1)} \sum_{i\neq j=1}^{N} Y_i Y_j$$

$$= \frac{1}{N}\sum_{i=1}^{N} Y_i^2 - \frac{1}{N(N-1)} \sum_{i\neq j=1}^{N} Y_i Y_j$$

$$= \frac{1}{N-1}\left[\sum_{i=1}^{N} Y_i^2 - N\bar{Y}_N^2\right]$$

$E(s^2) = S^2$ [$\because$ from equation ... (a)]

Hence proved.

Q33. In srswor, the variance of the sample mean is given by, $\text{Var}(\bar{y}_n) = \left(\frac{1}{n} - \frac{1}{N}\right)S^2 = \left(\frac{N-n}{nN}\right)S^2$.

Solution :