Chapter–4

Research Methodology: The Design and Techniques

S K Acharya, A Biswas, and K Reang

The discussion on the methodology has been made to understand the concepts, methods and techniques, which are utilized to design the study, collect the information, analyze the data and interpret the findings for revelation of truth and formulation of theories. This chapter deals with the research methodology that had been adopted for the purpose of the present study. However, the entire discussion has been made under the following sub-themes:

i. Locale of study
ii. Sampling design
iii. Pilot study
iv. Variables and measurements
v. Method of data collection
vi. Statistical tools used for analysis and interpretation of data.

i) Locale of the study (research location):

Keeping in view the paucity of time, resources and current socio-political situation the present study was conducted at Manu block of Dhalai district

ISBN: 978-93-85822-09-4

(Tripura). The district, block and village were selected purposively due to the following reasons:

a. The research's close familiarity with respect to area, people, their culture and local dialects.
b. The concerned areas are easily accessible to the researcher in term of transportation and place of residence.

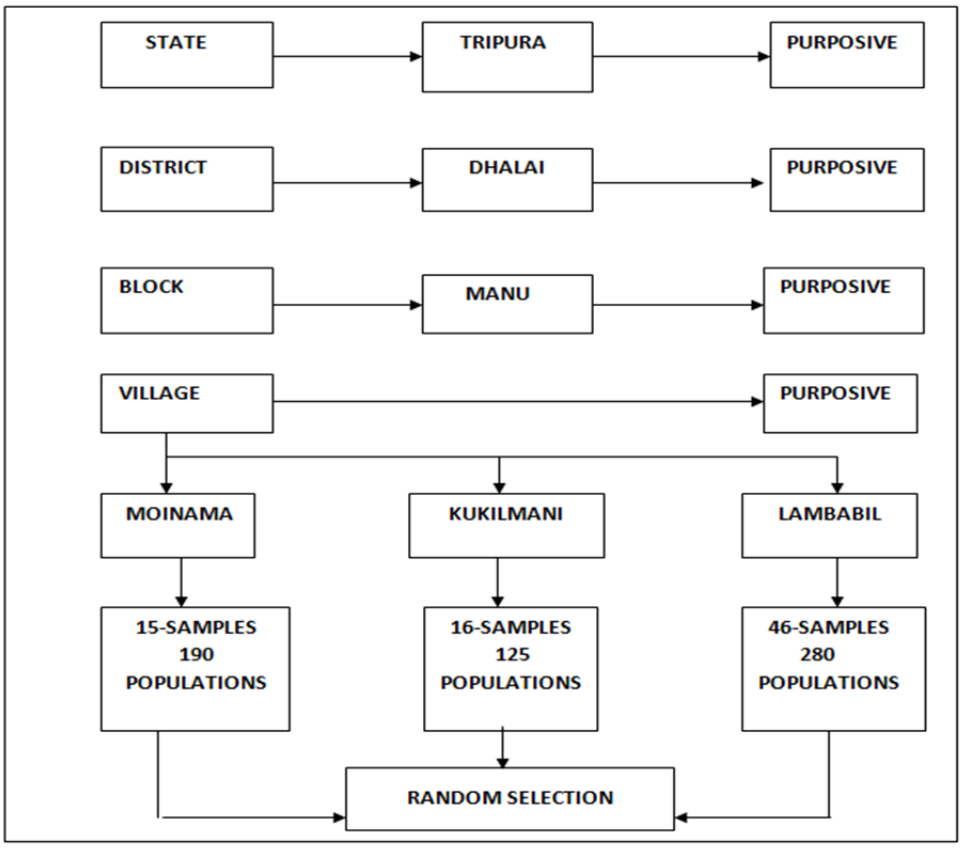
c. The opportunity of getting good response from the respondents

d. The study will help the researcher in future for conducting various extension programme and activities if he serves the agricultural department of state in future.

ii) Sampling design

The purposive as well as simple random sampling techniques were adopted for the present study. It may be termed as multistage random sampling procedure. The districts, blocks and villages were purposively selected for the study. The Dhalai district and the block Manu were considered. Under the Manu block Moinama, Lambabil and Kukilmani villages was selected. From Moinama village 14 samples have been selected out of 190 populations; from Kukilmani village 16 samples have been selected out of 125 populations, and from Lambabil 46 samples have been selected out of 280 populations. This way the total size of respondent stood 77. This was rather a disproportionate random selection of respondent.

FLOW CHART ON SAMPLING METHOD



iii) Pilot study

A pilot study was conducted in the selected villages before constructing the data collecting devices. In course of this survey informal discussion was carried out with some rubber growers and other resource persons of the localities. An outline of socio-economic background of the farmers of the concerned villages, their opinion towards different types of technology socialization process, adoption, discontinuance, rejection and re-invention were obtained that helped in the construction of reformative working tools.

ISBN: 978-93-85822-09-4

iv) Variables and their measurement

Several researchers pointed out that the behaviour of an individual was understood more in depth if one has the knowledge of some variables, which comprised the constructed world of reality within which an individual received the stimuli and acts. The socio-personal, agro-economic, socio-psychological and communication variables are such type of variables, which determine the behaviour of an individual. Appropriate operationalisation and measurement of the variables help the researcher to land upon the accurate conclusion. Therefore, the selected variables for this study had been operationalised and measured in following manner:

INDEPENDENT VARIABLES

Socio-personal variables

Age (X_1):

It denotes the chronological age, years and months elapsed since birth of the respondent. It was measured through counting the chronological age.

Education (X_2):

Education is the factor that has been conceived in terms of acquisition of knowledge and skill formality in school. The values ascribed the class in which he or she studied or studying.

Family size (X_3):

It denotes the total number of persons living in the respondents' house under a single household.

ISBN: 978-93-85822-09-4

AGRO-ECONOMIC VARIABLES

Farm mechanisation(X_4):

It is measured by the percentage of under farm mechanization with respect of tools and implements used for rubber plantation.

Cropping intensity(X_5):

It refers to the proportion of acreage annually put under different crops to the total cropped area expressed in percentage. The cropping intensity is calculated by

$$\text{Cropping intensity (\%)} = \frac{\text{Gross Cropped Area}}{\text{Net Cropped Area}} \times 100$$

Farm size:

It is defined in term of area of the land in kani held under family ownership of Homestead + Cultivable land + leased in.

Homestead(X_6):

Mass of land congruent to dwell and generally used for household purpose. It is measured in term of kani (1 ha= 6.25 kani).

Own land(X_7):

The mass of land possessed by the individuals measured in term of kani.

Area under rubber(X_8):

The mass of land area under rubber cultivation measured in term of kani.

Annual income-farm (X_9):

The grossed income is constituted by the total income generated from the farming activities. It is measured in tem of rounded of rupees/kani/annum.

ISBN: 978-93-85822-09-4

Income from rubber (X_{10}):

The yearly income from rubber cultivation in term of rupees/kani/year.

Annual-income off farm (X_{11}):

The yearly income other than farm activity and it is measured in term of rupees.

Economic status (X_{12}):

The economic status of farmer was measured with the help of some components of socio-economic status (Pareek and Trivedi, 1964). The 3 items are land holding, house type and farm power. The score assigned as in the case mentioned in the interview scheduled.

Material possession (X_{13}):

The respondents were asked whether they possessed some of the household materials like cycle, radio, mobile etc .It is calculated by the following formula-

$$\frac{\text{Present value}}{\text{Family size}}$$

SOCIO-PSYCHOLOGICAL VARIABLES

Scientific orientation (X_{14}):

In the present study scientific orientation was operationalised as the characteristic of individual, which made him to trust and rely on ideas and practices developed through scientific research. It was measured with the help of scale developed by Supe (1969). His scale has five response categories ranging from “strongly agree” to strongly disagree.

ISBN: 978-93-85822-09-4

Orientation towards adoption (X_{15}):

Structured scheduled developed. The scale has five responses categories ranging from strongly agree to strongly disagree.

Orientation towards discontinuance (X_{16}):

Structured scheduled developed. The scale has five response categories ranging from strongly agree to strongly disagree.

Orientation towards rejection (X_{17}):

Structured scheduled developed. The scale has five response categories ranging from strongly agree to strongly disagree.

COMMUNICATION

Utilization of source of information (X_{18}):

It refers to the use of information source(s) both from outside and within the community for getting agricultural information. It was measured through scale developed by Singh (1993).

DEPENDENT VARIABLES

The appropriate operationalisation and measurement of the predicted variables help in concluding the study in a proper manner. This is a very interesting area of work in measuring the variables after conceptualizing them.

Adoption index (Y_1):

The adopted technologies of rubber cultivation by the respondents. It is measured as-

$$\text{Adoption index}(\%) = \frac{\text{Level of adoption}}{\text{Recommended level}} \quad X\ 100$$

ISBN: 978-93-85822-09-4

Re-invention (Y_2):

Structural schedule developed. The scale consists of few statements made by respondents and each statement was of 10 point scale and asked to the respondents to give preference score out of 10. The score for each respondent was obtained by summating the score for each statement divided by total statement.

Perceived constraints (Y_3):

Structured schedule developed on the basis on scale and each scale consists of few statements responded by the each individual Rubber growers in their own circumstances and were asked to give their preference score out of 10 scale. The score of each respondent was obtained by summating the score of each statements divided by the total statement.

Socialization level (Y):

Socialization level, the ultimate consequent variable has been measured as follows:

$$\frac{Y_1+Y_2+Y_3}{3}$$

v) Techniques of data collection

The primary data in the present study were collected directly from the farmers with the help of structured schedule through personal interview methods. Only the functional head of the household were taken as respondents for the study.

ISBN: 978-93-85822-09-4

The personal interview method was followed during the month of January and May, 2012 to collect the relevant information from targeted respondents.

vi) Statistical tools used for analysis and interpretation of data

After collection of data, data were processed and analyzed in accordance with the outline laid down for the purpose at the time of developing the research plan. Processing implies editing, coding, classification, and tabulation of collected data. The main Statistical techniques and tool used in the present study-

Mean

The mean is the arithmetic average and is the result obtained when the sum of the value of individual in the data is divided by the number of individuals in the data. Mean is the simplest and relatively stable measure of central tendency. The mean reflects and is affected by every score in the distribution. Thus, extreme scores affect the mean.

Calculation of mean from grouped data

When the data are expressed in a frequency distribution (grouped), the mean calculated by the formula.

$$\bar{x} = \frac{\sum \int x}{N}$$

Where,

$\bar{x}$ = mean of the distribution

f = frequency of the class

x = class value of midpoint of the class internal

N = number of observations

ISBN: 978-93-85822-09-4

Standard deviation

Standard deviation is the square root of the arithmetic mean of the squares of the deviations, the deviations being measured from the arithmetic mean of distribution. It is commonly denoted by the symbol sigma. It is less affected by sampling errors and is more stables measure of dispersion. The Standard deviation of the data grouped in the form of frequency distribution is computed by the formula-

$$a=\sqrt{\frac{\sum f.d^2}{N}}$$

Where,

I = frequency of the class

d = deviation of the mid-value of the class from the population mean

N = total number of observations.

Coefficient of variation

A measure of variation which is independent of the unit of measurement is provided by the Coefficient of variation. Being unit free, this is useful for comparison of variability between different populations. The Coefficient of variation is standard deviation expressed as per centage of the mean and is measured by the formula-

$$CV=\frac{Standard\ deviation\ (\sigma) x\ 100}{Mean}$$

Correlation

When an increase or decrease in one variety is accompanied by an increase or decrease in the other variety, the two are said to be correlated and the phenomenon is known as correlation. Correlation coefficient (r) is a

ISBN: 978-93-85822-09-4

measure of the relationship between two variables, which are at the interval or ratio level of measurement and are linearly related. A Person product-moment 'r' is computed by the formula.

$$r_{xy} = \frac{N\sum XY - (\sum X)(\sum Y)}{\sqrt{[N\sum x^2 - (\sum x^2)][N\sum Y^2 - (\sum Y)^2]}}$$

Where,

X and Y = original scores in variables X and Y

N = number of paired scores

$\sum XY$ = each X multiplied by its corresponding Y, then summed

$\sum X$ = sum of X scores

$\sum X^2$ = each of X squared, then summed

$(\sum X)^2$ = sum of X scores squared

$\sum Y$ = sum Y scores

$\sum Y^2$ = each of Y squared, then summed

Regression

The correlation coefficient only expresses association and by itself tells us nothing about the causal relationships of the variables. Thus, purely from the knowledge that two variables x and y are correlated, we cannot say whether variation in x is the cause or the results from mutual dependence of the two varieties or form common causes affecting both of them. Similarly, the mere existence of a high value of correlation coefficient is not necessarily of an underlying relationship between the two variables.

The underlying relation between y and x in a bi variant population can be expressed in the form of a mathematical equation known as regression

ISBN: 978-93-85822-09-4

equation and is said to represent the regression of the variety y on the variety x.

If Y is the dependent variable and X is the independent variable, then the linear regression equation can be written as -

$$Y = a + bx$$

The values of a and b can be obtained by the method of least squares which consists of minimizing the expression.

$(\sum y_i - a - bx_i)^2$ with respect to a and b. The value of a and b are

$$b = \frac{\sum XY - \frac{(\sum Xi)(\sum Yi)}{n}}{\sum Xi^2 - \frac{(\sum Xi)^2}{n}}$$

$$a = \bar{y} + \overline{bx}$$

The regression line can now be written as

$$Y = \bar{y} + b\bar{x} + bx \text{ or } y - \bar{y} = b(x - \bar{x})$$

Where, b is the regression coefficient.

Path analysis

The terms 'path analysis' was first introduced by the biologist Sewall Wright in 1934 in connection with decomposing the total correlation between two variables in the casual system. The technique of path analysis is based on a series of multiple regression analysis with the added assumption of causal relationship between independent and dependent variables. This technique lays relatively heavier emphasis on the heuristic use of visual diagram, technically described as a path diagram. An illustrative path diagram showing interrelationship between father's education, father's occupation, son's education, son's first and son's first and son's present occupation can be shown.

ISBN: 978-93-85822-09-4

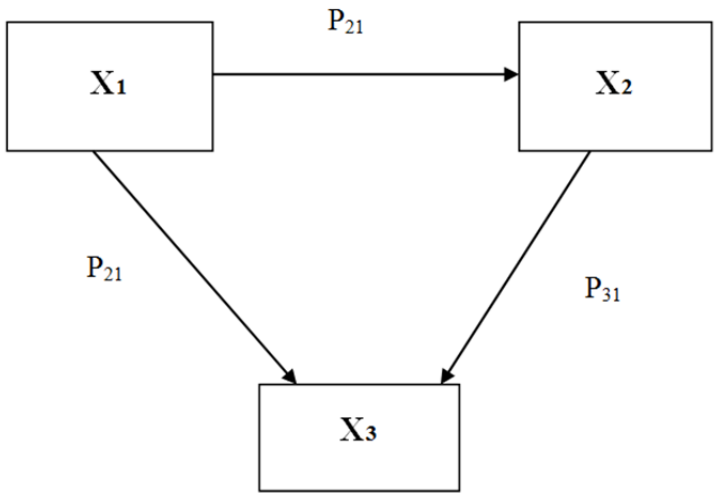
Path analysis makes use of standardized partial regression coefficients (known as beta weights) as effect coefficients. In linear additive effects are assumed, then through path analysis a simple set of equations can be build up showing how each variable depends on preceding variables. The main principle of path analysis is that any correlation coefficient between two variables, or a gross or overall measure of empirical relationship can be decomposed into a series a paths: separate path of influence leading through chronologically intermediate variable to which both the correlated variables have links.

The merit of path analysis in comparison to correlation analysis is that it makes possible the assessment of the relative influence of each antecedent of explanatory variable on the consequent or correlation variable by first making explicit assumptions underlying the causal connections and then by elucidating the indirect effect of the explanatory variables.

The use of the path analysis technique requires the assumption that there are linear additives, a symmetry relationship among a set of variables which can be measured at least on a quest interval scale. Each dependent variable is regarded as determined by the variable preceding it in the path diagram, and a residual variable defined as uncorrelated with other variables, is postulated to account for the unexplained portion of the variance in the dependent variable. The determining variables are summed for the analysis to be given (exogenous in the model).

We may illustrate the path analysis technique in connection with a simple problem of testing a causal model with three explicit variables as shown in the following path diagram:

ISBN: 978-93-85822-09-4



Path diagram (with three variables)

The structural equation from the previous can be written as:

$X_1 = e_1$

$X_2 = P_{21}X_1 + e_2 = px + e$

$X_3 = P_{53}X_2 + P_{32}X_2 + e_3$

X1 and X variable are measured as deviation from their respective means. P_{21} may be estimated from the simple regression of X_2 on X_1, i.e., $b_{21}X_1$ and P_{31}may be estimated from the regression of X on X_2 and X_1 as under:

$$X = P_{31}X_1 + b_{21}x_2$$

Where, $b_{21}x_2$ means the standardized partial regression coefficient for predicting variable 1 when the effect of variable 2 is held constant.

In path analysis the beta co-efficient indicates the direct of X1 (j=1, 2, 3,..p) on the dependent variable. Squaring the direct effect yields the proportion of variance on the dependent variable Y which is due each of the number of independent variable X_1 (j=1,2,3...p). After calculating the direct effect one

ISBN: 978-93-85822-09-4

may obtain a summary measure of the total indirect of X_1 on the dependent variable Y by subtracting from the correlation coefficient r_{yxj} the beta coefficient b i.e.

Indirect effect X_1 on $y=Cjy=r_{yxj}-b_1$

For all j=1, 2, 3...p

Factor analysis

Many statistical methods are used to study the relation between independent and dependent variables. Factor analysis is different; it is used to study the patterns of relationship among many dependent variables, with the goal of discovering something about the nature of the independent variables that affect them, even though those independent variables were not measured directly. Thus answers obtained by factor analysis are necessarily more hypothetical and tentative than is true when independent variables are observed directly. The inferred independent variables are called factors.

Concepts used in factor analysis

Some important concepts used in factor analysis are explained by Kothari (1996) as follows-

Factor-A factor is an underlying dimension that accounts for several observed variables. Factor is a hypothetical construct or classification. There may be one or more factors, depending upon the nature of the study and the number of variables involved in it.

Factor loading - Factor loadings are those values which explain how closely the variables are related to each one of the factor discovered. Factor loadings work as key to understanding what the factor mean. It is the

ISBN: 978-93-85822-09-4

absolute size (rather the sign, plus or minus) of the loading that is important in the interpretation of a factor.

Since the factors happen to be linear combinations of data, the co-ordinates of each observation or variables is measured to obtain what are called factor loading. Such factor loading represent the correlation between the particular variable and the factor, and are usually placed in a matrix of correlation between the variables and the factor.