Chapter - 4

Methodology

The deliberation on the methodology has been made to understand to concept, methods and techniques which utilized to design the study, collect on of information, analysis the data and interpretation of the findings for revelation of truths and formulation of theories. This present chapter deals with the method and a procedure used in the study and consists of six main parts.

A. Locale of research
B. Pilot study.
C. Methods of sampling
D. Variables and Measurements
E. Tools and techniques of data collection.
F. Statistical analysis and interpretation of data.

A. Locale of research

Selection of District

Keeping in view agriculturally and socio-economically backward area having their major source of income from agriculture enterprise, Nalanda district was selected for the study.

Selection of Block

Chandi block of Nalanda district was purposively selected for the study. The reasons behind such a selection are a follows:

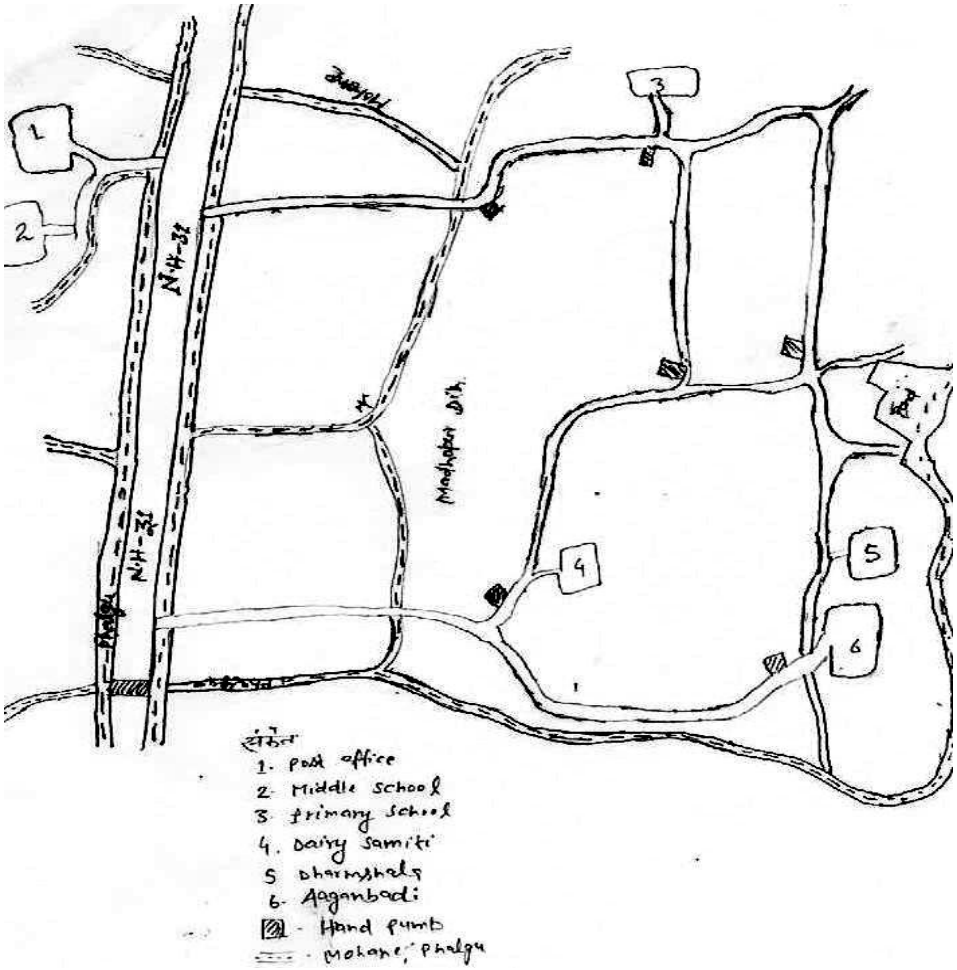
1. The researcher's close familiarity with the area, the people, their culture and the locale dialect provided added weight age facilitating the study and the process of collecting data.
2. The concern area was easily accessibility to the researcher in terms of transportation and place of residence.

Selection of village

Madhopur Dih village was under Madhopur gram panchayat selected purposively for study. The main reason behind the selection of Madhopur Dih village was due to the change dynamics of livelihood generation over decades there.

B .Pilot study

Before going to collect sample area of work or investigation, pilot study was conducted to understand the area, its people, institution, communication, extension system and attitude of people. Basic situational and background information was collected during period of pilot study from the different sources including panchayat office, Block office, ADO office.



Map of the village: Madhopur Dih

C. Methods of sampling

Purposive as well as simple random techniques were adopted for the study. For selection of district, block, village purposive sampling techniques was employed for selection of respondent. There are 325 families in the village Madhopur Dih which constitute the total population of the study. Out of 325

families only 75 farmers have been randomly selected. In the present investigation out of 325 farmers only 75 entrepreneurs have been interviewed.

STEP	LEVEL	APPROCH
·I	District Nalanda	Purposive
·II	Block Chandi	Purposive
III	Village Madhopur Dih	Purposive
İV	75 respondents	Randomly selected

D .Variables and Measurements

After reviewing various literatures related to the field of present study and consultation with respected of chairman of advisory committee and other experts, a list of variables was prepared. On the basis of selected variables, a schedule was formed.

1. **Age (X_1):** It refers to Chronological age of respondent at the time of interview has been considered.
2. **Education(X_2):** Structured schedule developed it had been measured according to the educational class attainment of the respondent. The education had been divided into ten categories that is Illiterate, can read only, can read and write, primary, middle school, secondary, higher secondary, graduate, post-graduate and above. It had been measured with the help of scale developed by Pareek and Trivedi (1964) with

modification. Scale is socio-economic status (rural) and the weightages had been given as Illiterate –(0), Can read only –(1), Can read and write –(2), Primary –(3), Junior High – (4),High school –(5), Graduate-(6) Post graduate-(7).

3. **Family size (X_3):** It refers to the total number of members in farmer's family.
4. **Communication index (X_4):** The attribute material possession had been operationalized as the material possession of the respondent in the social system. It had been measured with the scale developed by the Pareek and Trivedi (1964) with modification and weightages as the motor cycle, radio, cycle, Television, Computer, Car, power tiller, Refrigerator and Mobile.
5. **Distance Matrix(x_5):** : Critical point = (market + health center +school + bank)/n
6. **Doctors visited(x_6):** How many times visited in a year.
7. **House type (X_7):** The attribute house type had been operational zed as the house type of the respondent in the social system. It had been measured with the scale developed by the Pareek and Trivedi (1964)with modification and weight ages as the no house, hut, kutcha house, mixed house, pucca house,. Socio-economic status (rural) and the weightages had been given as No house –(0), Hut–(1), Kutcha house–(2), Mixed house –(3) and Pucca house–(4).

8. **Sanitation(x_8):** Structured schedule was developed and score assigned to each respondent on the basis of 10 point scale.
9. **Drinking water(x_9):** Structured schedule was developed and score assigned to each respondent on the basis of 10 point scale.
10. **Drudegry(x_{10}):** Structured schedule was developed and score assigned to each respondent on the basis of 10 point scale.
11. **Home Innovation(x_{11})** The attribute material possession had been operationalized as the material possession of the respondent in the social system. It had been measured with the scale developed by the Pareek and Trivedi (1964) with modification and weightages as the motor cycle, radio, cycle, Television, Computer, Car, power tiller, Refrigerator and Mobile.
12. **Land holding(x_{12}):)**: the attribute of land holding had been operationalized as the land holding of the respondent in the social system.
13. **Yield(x_{13}):** Actual yield of different crops (Rice,Wheat,pulses,Potato and vegetables), were obtained in q/land holding.It has been measured with Structured schedule was developed and score assigned to each respondent on the basis of certain scale point.
14. **Adoption level(x_{14}):** It refers to adoption behaviour of farmers towards the fertilizer, pesticides. It has been measure with Structured schedule

was developed and score assigned to each respondent on the basis of certain scale point.

15. **Annual Income(x_{15}):** Income from primary and secondary sources of occupation in rupees per year was taken in to account.

16. **Expenditure(x_{16}):** (total exp./family size)

Dependent variable:

1. Wage (Y_1): It has been defined in term of daily remuneration paid against the service render by an individual. Wage has been calculated in terms of money value earn by an individual through services.

2. Mandays (y_2): Number of days engaged in wage generating activities in a month by a person male or female.

3. Change dynamics of rural livelihood (Y): Y_1+Y_2

E. Preparation of schedule

After quantifying the variables (both dependent and independent) for measurement, Interview schedule were prepared with the help of respected chairman.

Pre-testing of schedule

Before going to field of work and starting final data collection, schedule was pre-tested for elimination, addition, and alternation with non-sample respondents of the study area for getting better result.

F. Tools for data collection

The major tool used for collection of primary data in the study was structured schedule. Data were collected be investigator personally from respondents.

Techniques of data collection

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

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

G. Statistical analysis and interpretation of data (Analytical tools)

After collection of data, data were processed and analyzed in accordance with 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-

1. **Range**
   It is the measure of maximum and minimum value of variables.
2. **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 simplest and relatively stable measure of central tendency. The mean reflects and is affected by every score in the distribution.

When the data are expressed in a frequency distribution (grouped), the mean calculated by the formula use was as follows –

$$\bar{X} = \frac{\sum_{i=1}^{N} fixi}{N}$$

Where,

x = Mean of the observation

fi = Frequency of the class

xi = Mid value of the class

N = Total number of observation

3. Standard deviation

Standard deviation is a measure of dispersion which implies the extent to which observation vary among themselves. Standard deviation (SD) of a set of observation is the square root of the arithmetic mean of squares or deviations from arithmetic mean. It is denoted by a (Sigma).

For frequency distribution standard deviation (SD) is measure as follows

$$\sigma = \sqrt{1/N \sum fi (x_i - \bar{x})^2}$$

Where,

σ = Standard deviation

N = Total No. of observation in a particular cells x = Value of observation in a particular cell

f = Frequency of observation

$\bar{x}$ = Mean number of observation

i = Any number (e.g., 1, 2, 3) denoting position

4. Coefficient of variation

A measure of variation which is independent of the unit of measurement is proved 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 percentage of the mean.

Coefficient of variation is measured by the formula use was as follows –

$$C.V. = \frac{S.D.}{Mean} \times 100$$

5. Correlation coefficient

When an increase or decrease in one variety is accompanied by an increase or decrease in other variety, the two are said to be correlated and the phenomenon is known as correlation. Correlation coefficient (r) is a measure of the relationship between two variables, which are at the interval or ration 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 scares in variables X and Y

N = Number of paired scores

∑XY = Each X multiplied by its corresponding Y, then summed

∑X = Sum of X scores

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

$(\sum X)^2$ = Sum of X score squared

∑Y = Sum of Y scores

$\sum Y^2$ = Each of Y squared, than summed

$(\sum Y)^2$ = Sum of Y score squared

The range of correlation coefficient is between -1 to + 1. This means that -1 is perfect negative correlation, + 1 is perfect positive correlation. A perfect correlation is, however, seldom achieved. An idea of positive and negative correlation is given here. If the numbers of errors increase with increase in typing speed, it indicates positive correlation. If the numbers of correct words decrease with increase in typing speed, it is indicative of negative correlation. A correlation coefficient to be acceptable should be statistically significant. Otherwise, we say that no significant relationship exist between the variables.

6. Path analysis:

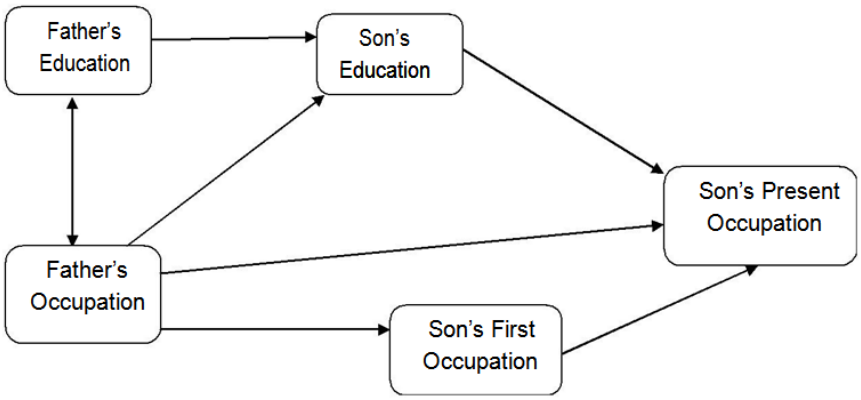
The term was first introduced by the biologist Sewal Wright in 1934 in connection with decomposing the total correlation between any two variables in a causal system. The technique is based on a service of multiple regression analysis with the added assumption of the causal relationship between independent and dependent variable.

Path analysis makes use of standardized partial regression co-efficient (known as beta weights) was effect co-efficient. In linear additive affects are assumed, then through path analysis simple set of equations can be built up showing how each variable depend on preceding variable. The main principle of path analysis is that a correlation coefficient between two variables, or a gross or overall measure of empirical relationship can be decomposed in a series of parts: separate parts of influence leading through

chronologically intermediate variable to which both the correlated variable 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 or explanatory variables on the consequent or criterion variables by first making explicit the assumption, underlying the causal connections and then by elucidation the direct affect the explanatory variables.

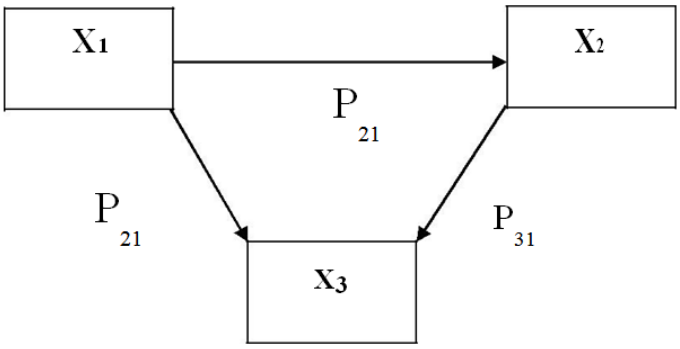
An illustrative path diagram showing inter relationship between father's education, father's occupation, son's first and son's present occupation can be shown as;



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:



Path diagram (with three variables)

The structural equation for the above can be written as:

$X_1 = e_1$

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

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

X_1 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 X_1 (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 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 co-efficient b i.e.

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

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

7. Factor Analysis

Factor analysis is a very useful and popular method of multivariate research technique, mostly used in social and behavioural sciences. According to Kothari (1996), factor analysis seeks to resolve a large set of measured variables in terms of relatively few categories, knows as factors. This technique allows the researcher to group variables into factors (based on correlation between variables), and the factors so derived may be treated as new variables (often termed as latent variables) and their value derived by summing the values of the original variables, which had been grouped into the factor. The meaning and name of such new variable is subjectively determined by the researcher. Since the factors happen to be linear

combinations of data, the coordinates of each observation or variable is measured to obtain what are factor loadings. Such factor loading represent the correlation between the variable and the factor, and are usually placed in a matrix of correlations the variable and the factors. In the Factor Analysis the "Principle Component Method" was followed.

12. Stepwise Multiple Regression analysis

Generally a number of antecedent variables simultaneously contribute or influence the consequent variable, as in the case under study. It is of immense practical value to know the extent to which the antecedent variables, individually or jointly, could predict or contribute towards the consequent variable. This was done by computing multiple regressions. If Y is the consequent variable and X_1, X_2, X_3… are the antecedent; the multiple regression equation is given by –

$Y = a + b_1X_1 + b_2X_2 + b_3X_3$ ……

Or, $Y = a + \Sigma bX$

Where,

a = Intercept constant

b = Regression coefficient

The significance of the b – values was judged by calculating their respective t – values and comparing them to the table values, given by Fisher and Yates (1963) with n – p – 1 degree of freedom (where, n = number of respondents and p = number of antecedent variables) at 5 per cent and 1 per cent levels of significance.

The square root of the ratio of the regression sum of squares to the total sum of squares is known as multiple correlation coefficients and is denoted by R. The square of the multiple corre1ation coefficient R^2 is called the multiple coefficient of determination and represents the fraction of the variation of y accounted for by its joint association with the varieties X_1, X_2, X_3, Central to the application of multiple regression analysis is the interpretation of the final fitted mode. A significant F-value for R means that the fitted model is adequate. The significance of the F-value was judged by comparing it to the table value given by Fisher and Yates (1963) with p and n – p – 1 degrees of freedom (where, p = number of antecedent variables and n = number of respondents) at 5 per cent and 1 per cent levels.

H. Analysis of data

The data of the present investigation with the help of the above mentioned statistical tools has been analyzed taking the support of the package SPSS (Ver. 19).