Chapter *III*

Research Methodology

The research methods and procedures adopted in the study are laid down in this chapter under the following heads-

1) Locale of study
2) Pilot study
3) Methods of sampling
4) Variables and Measurements.
5) Tools and techniques of data collection
6) Statistical analysis and interpretation of data.

1) Locale of the study (Research location)

Selection of District

Keeping in view the paucity of time, resources and current socio-political situation, NADIA district was selected purposively for the study.

Selection of Blocks

Chakdaha block of Nadia district was purposively selected for the study. The reasons behind such a selection is 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

The investigator has a dialogue with respected Chairman of Advisory Committee and concern experts of area of study. Goragachcha village was selected purposively for study.

2) Pilot study

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

3) Methods of sampling

Purposive as well as simple random sampling techniques were adopted for the study. For selection of district, block, village purposive sampling techniques was employed for selection of respondents. There are 134 families in the village Goragachcha which constitute the total population of the study. Out of 134 families only 53 families have been randomly

selected. In the present investigation only the head of the selected 53 families have been interviewed.

STEP	LEVEL	APPROACH
I	District Nadia	Purposive
II	Block Chakdah	Purposive
III	Village: Ghoraghacha	Purposive
IV	53 Respondents	Randomly selected

4) 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.

The socio-economic variables are described below

Independent variables and dependent variables and empirical measurement

Socio-personal variables

Independent variables

Age-X_1	:	It refers to Chronological age of a respondent at the time of interview rounded off to the nearest year.
Education-X_2	:	Formal education received by respondent, measured in terms of year of schooling+1.
Family size -X_3	:	The family size was measure with the help of number of family members of the respondent in terms of members below 14yrs. + members above 14yrs.
Family statement with adult person -X_4	:	The family statement with adult person was measure with the help of number of family members of the respondent in term
Functional Education Strata (FES) -X_5	:	It was measured by percentage of under Functional Education Strata with respect of training, vocational course, seminars, meeting, workshop (last 5yrs).
Cropping Intensity -X_6	:	It refers to the proportion of coverage annually put under different crops to the total cropped area expressed in percentage. The cropping intensity was calculated by the following formula- C.I=(Total annual cropped area x 100)/size of holding

Irrigation status -X_7	:	It was defined in terms of percentage of the cultivable land under irrigation by any conventional source for example: shallow tube-well, deep tube-well, and pond.
Animal/Bird number - X_8	:	It is digital count of animals/poultry birds nourished by the family.
Holding size -X_9	:	It was defined in terms of area land in bigha held under single family ownership.
Income(Rs) per cottah - X_{10}	:	The gross income was constituted by the total income generated from farming activities. It was measured in terms rounded of rupees/cottah/annum.
Spacing(Percentage)-X_{11}	:	It has been conceived in terms of level of adoption for fertilizer in light of; spacing of land. Spacing was calculated by the following formula: (adopted spacing x 100)/recommended spacing.
Fertilizer(Percentage)-X_{12}	:	It has been conceived in terms of level of adoption for fertilizer in light of; (adopted dose x 100)/recommended dose.
Irrigation (Percentage)-X_{13}	:	It was defined in terms of percentage of the cultivable land under irrigation by any conventional source for example: shallow tube-well, deep tube-well, and pond.(applied amount of water x 100)/recommended amount of water.
Pesticide(Percentage)-X_{14}	:	It has been conceived in terms of level of adoption for pesticide in light of; (adopted dose x 100)/recommended dose.
Yield(Percentage)-X_{15}	:	It has been measured in terms of volume of biological production in agriculture or non agriculture sector and measured as yield/unit area/year.
		Dependent Variables
Livelihood security-Y_1	:	The operational definition of livelihood here has been conceived in terms of number of men/women days generated by agril/non-agril area in a year.
Food intake value /g /day/head-Y_2	:	It has been measured by in terms of how much of food intake incurred/g/Day/head.
Health status-Y_3	:	Health status is here has been defined the status and the regime of health conceived in terms of following elements (frequency of disease, expenditure of health, immunization of babies). The elements have been quantified and standardized through befitting measure.
Wage-Y_4	:	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.
Sustainable Livelihood(Y_5)	:	It has been calculated through result of dependent variables. These variables named as sustainable livelihood. This variable has been derive by multiply in the digital values of the four preceding dependent variables (Y_1 x Y_2 x Y_3 x Y_4) and subsequently divided it by four.

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.

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 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.

The personal interview method was followed during the month of February-june, 2009 to collect the relevant information from targeted respondents.

Statistical analysis and interpretation of data (Analytical tools)

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

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-

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

Where,

I = frequency of the class

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

N = total number of observations.

Percentage (%)

Percentage was used for simple comparison for calculating percentages; the frequency of a particular set was divided by the total number multiplied by 100.

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 percentage 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

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

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 number 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.

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 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.

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 se 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.

Path analysis makes'' The use of the path analysis technique requires the assumption that there are linear additives, a symmetric relationship among a set of variables which can be measured at least on a quasi-interval scale. Each dependent variable is regarded as determined by the variables preceding it in the path diagram, and a residual variable, defined as un-correlated with the other variables, is postulated to account for the unexplained portion of the variance in the dependent variable.

Factor Analysis

Factor analysis is a very useful and popular method of multivariate research technique, mostly used in social science and behavioral science. This technique is applicable when there is a systematic interdependence among a set of observed or manifest variables and the researcher is interest in finding out something more fundamental or latent which create this communality (commonness).For example we may have data on farmers' education, land, homestead, material possession, etc. and want to infer from these some factor relating to scientific orientation, which will summarize the communality of all the variables.

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 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.