Chapter 5

Methods, Tools and Techniques

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

A. Locale of Research.

B. Sampling Design.

C. Pilot Study.

D. Variables and their measurement.

E. Preparation of Interview Schedule.

F. Pre-testing of Interview Schedule.

G. Techniques of Data Collection.

H. Statistical Tools used for Analysis of Data

5.1. Locale of research

BANTIKA-BOINCHI Gram Panchayat of the PANDUA block of HOOGHLY district in WestBengal was purposively selected for the study. The village namely BOINCHIGRAM wasselected by random sampling. The area had been selected for the study because of-

(a) There is ample scope for collecting relevant data for the present study.

(b) Acquaintance with the local people as well as the local language.

(c) The concern area was easily accessible to the researcher in terms of place ofresidence.

(d) The area was very easily accessible to the researcher in terms of transportation

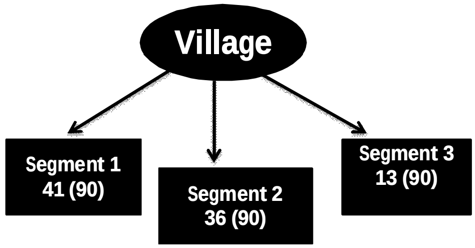
(e) The closure familiarities of the student researcher with the area, people,officials and local dialects.

5.2. Sampling Design

Purposive as well as simple random sampling techniques were adopted for thestudy. For selection of state, district, block and gram panchayat purposive samplingtechniques was adopted because the area was ideal for OccupationalHealth Management study,convenient for researcher and having the infrastructural facilities and in case ofSelection of villages and respondents simple random sampling technique was taken up.

Table 5.2.1: Sampling Technique and Sampling Design

Step	Items	Level	Approach
1	State	West Bengal	Purposive
2	District	Hooghly	Purposive
3	Block	Pandua	Purposive
4	Gram -Panchayat	Bantika- Boinchi	Purposive
5	Village	Boinchigram	Purposive
6	Respondents	90	Random
Total number of respondents : 90			



41 + 36 + 13 = 90

Fig . Selection of Respondents from villages

Fig.5.2.1.1: Selection of respondents from different segments of village

5.3.Pilot study

Before taking up actual fieldwork a pilot study was conducted to understand thearea, its people, institution, communication and extension system and the knowledge,perception and attitude of the people towards climate change concept. An outline of thesocio-economic background of the farm women of the concerned villages, theirperception on health issues, natural resources, ecology, nutritional aspects etc. helped in the construction of reformative working tools.The components of pilot study

are:

- General information;
- Specific information;
- Prevalence of variables;
- Body languages of the prospective respondents;
- Access to physical location;
- The type, level and intensity of responsiveness;

- Related information including Agriculture.

5.4. Variables and theirempiricalmeasurement

Several researchers pointed out that the behaviour of an individual has been understood more in depth if one has the knowledge of some variables, which comprisedthe constructed world of reality within which an individual received the stimuli and acts.

The socio personal, agro economic, socio-psychological and communication variablesare such type of variables, which determine the behaviour of an individual. Appropriateoperationalization and measurement of the variables help the researcher to land upon theaccurate conclusion. Therefore, the selected variables for this study had beenoperationalized and measured in following manner.

Variables in the present study have been categorized into two main categories.

1) Independent variables.

2) Dependent variables.

5.4.1. Independentvariables:-

The variables, andempirical measurements.

5.4.1.1. Age (x1) :

In all societies, age is one of the most important determinants of social status andsocial role of the individual. In the present study, the number of years rounded in thenearest whole number the respondent lived since birth at the time of interview, wastaken as a measure of age of the respondent.

5.4.1.2.Number of children (x2): Farm women have to play both the role of farm workers and the mother of their children. Children are the future who will hold the family baton. But, the number of children matters, because, with the help of this data one can assess that whether family planningaspect has adopted or not.

5.4.1.3 Number of farm work(x3): Total number of farm work is calculated for each farm women by summing up individual farm works.

W=(w1 + w2+w3+......+wn)

5.4.1.4.Working hour per day (x4) : This the total time a farm woman spends in her farm work. It has calculated in hours .

5.4.1.5.Incidence level of miscarriage(x5): As the farm women have to work hard in both the field and household and some of them have poor nutritional status. And the teenage marriage and then pregnancy is prevalent, the farm women sometimes face miscarriage problem. This is calculated by taking the number of times they have faced miscarriage upto the date of questioning to the respondents.

5.4.1.6.Number of animals reared (x6) : This is the total number of animals reared by the farm household in their yard. This includes both the livestock and the poultry birds .

5.4.1.7.Height(x7) : Height is measured in ft with the help of a measuring tape .

5.1.4.8.Weight(x8) : Weight is measured in kg with the help of a weight machine .

5.4.1.9.BMI(x9) : BMI is the abbreviated form of Body Mass Index. BMI is defined as body mass divided by the square of body height .It is calculated by using theformula :

kg/(m^2 * 0.305)^2,where 0.305 is used to convert ft^2 into m^2

BMI has categorized mainly in 5 groups viz. Very severely underweight (up to 15 kg/m2), Severely underweight (15 -16 kg/m2), Underweight (16 -18.5 kg/m2), Normal (18.5 – 25 kg/m2) and overweight (25-30 kg/m2) .

5.4.1.10.Cereals consumed per day(x10) : This the total amount of cereals consumed by the farm women per day. It is expressed in gram per day. Cereals are the staple food in this area .

5.4.1.11.Protein consumed per day(x11) : This the total amount of protein consumed by the farm women per day. It is expressed in gram per day.

5.4.1.12.Fruits consumed per day(x12) : This the total amount of fruits consumed by the farm women per day. It is expressed in gram per day .

5.4.1.13.**Vegetables consumed per day(x13) :** This the total amount of vegetables consumed by the farm women per day. It is expressed in gram per day .

5.4.1.14.**Total carbohydrate consumed per day (x14) :** This the total amount of carbohydrate consumed by the farm women per day. This calculated by summing up the total cereals and fruits consumed per day as these two are the main sources of carbohydrate to the respondents. It is expressed in gram per day .

5.4.1.15.**Fat consumed per day(x15) :** This the total amount of fat consumed by the farm women per day. It is expressed in gram per day .

5.4.1.16.**Breakfast time (a.m.) (x16) :** This is time when the respondents have their breakfast. It is mainly taken in the morning i.e. a.m.

5.4.1.17.**Lunch time (p.m.) (x17) :**This is time when the respondents take their breakfast. It is mainly taken in the afternoon i.e. p.m.

5.4.1.18.**Dinner time (p.m.) (x18) :** This is time when the respondents take their breakfast. It is mainly taken from the late evening to night i.e. p.m.

5.4.1.19. **Calorie in carbohydrate per day (x19) :**Calorie means the energy required to maintain one’s daily work. And the total calorie which is taken by the respondentper day through carbohydrate is calculated by multiplying each gram of carbohydrate consumed by 4. It is expressed in kcal .

5.4.1.20. **Calorie in protein per day(x20) :**Calorie means the energy required to maintain one’s daily work. And the total calorie which is taken by the respondent per day through protein is calculated by multiplying each gram of protein consumed by 4. It is expressed in kcal .

5.4.1.21. **Calorie in fat per day(x21) :**Calorie means the energy required to maintain one’s daily work. And the total calorie which is taken by the respondent per day through fat is calculated by multiplying each gram of fat consumed by 9. It is expressed in kcal .

5.4.1.22.**Total calorie per day(x22) :** This the total amount of calorie taken by a respondent per day. This is calculated as follows –

Total calorie per day = (calorie in carbohydrate /day + calorie in protein /day +calorie in fat/day)

It is expressed in kcal .

5.4.1.23.**Size of holding (x23) :**This is the total size of land one family has. It is taken as total size of both farm and homestead land but who have no agricultural land of their own i.e. they are share cropper or landless agricultural labourers, the size of only homestead land has taken. This is taken in Katta. So, the size of holding can depict the land status and main source of income of the respected farm family .

5.4.1.24.**Family income per annum (x24) :**Family income per annum is calculated as the earnings of the family from primary and secondary sources in a year in rupees. Family income boosts the participatory attitude of the respondents and determines their family expenditure per year .

5.4.1.25.**Per capita income per annum(x25) :**Per capita income of a farm women per annum can revel her status in the society and her access to the total family income as well as resources. The gross income is constituted with the income from farming, wage of agricultural labourers or part time work as maid servant. It is expressed in rupees. In the present study it is calculated as follows :

(Family income per annum / Family size) = Per capita income per annum

5.4.1.26.**Family expenditure per annum(x26):** Family expenditure per annum is the household expenses in different activities in a year. This is the output of family income .

5.4.1.27.**Per capita expenditure per annum(x27):**This is theoutcomeof the per capita annual income. It expresses respondent's power in buying decisions within the family .

5.4.1.28.**Functional literacy (x28) :** Functional literacy is a term was defined for UNESCO by Willam S. Gray (The Teachimg of Reading and Writing, 1956,p.21) as the training of adults to meet independently the reading and writing demands placed on them. The judgements were given on a 5-point scale (1-very week, 2-week ,3-normal, 4-strong, 5- very strong) by assessing mastery over the reading and writing capability and some other functional tasks.

5.4.2. Predictedvariables :

5.4.2.1. **Perceived physical problems (y1) :** This is the physical problems faced by the individual farm womantill the date of interviewing. This has calculated by the following way –

Perceived physical problems={ (p1/ r1) + (p2/r2) + ... +(pn /rn) }/ Total no. of physicalproblems

Where, p = physical problem, r = respective rank in matrix ranking

5.4.2.2.**Psycho-social hazards (y2) :** Psycho-social hazard is any occupational hazard that affects psychosocial well-being of the workers ,here farm women, including their ability to cope up in their work environment. It is related to the way of the work designed, it's organization and management as well as the economic and social context. In the present it was recorded in the following method –

Psycho-social hazard =pse * 5 + psf * 4 + psw * 3

Where ,pse = Psycho-social hazard due to economic stress, it was multiplied by 5 as this contributes highest to the Psycho-social hazard .

psf = Psycho-social hazard due to family problem, it was multiplied by 4 as this has moderate impact on the Psycho-social hazard .

psw = Psycho-social hazard due to work related problem, it was multiplied by 3 as this has comparatively less impact on the Psycho-social hazard .

5.4.2.3. **Frequency of visit to doctor (y3) :**This states that how often a person pays visit to doctors. Here, it was calculated by taking the total number of times a farm woman goes to doctors clinic or hospital in ayear .

5.5. Preparation of interview schedule

On the basis of the findings of pilot study a preliminary interview schedule wasformed with the help of literature and by the assistance of Chairman of

AdvisoryCommittee. The interview schedule consisted of three major parts according to thespecific objectives of the study.

5.6. Pre-testing of Interview Schedule

Pretesting or preliminary testing is the process of an advance testing of the studydesign after the schedule/questionnaire has been prepared. The object of pretesting is to detect the discrepancies that have emerged and to remove them after necessarymodification in the schedule. It also helps to identify whether the questions are logicallyorganized, the replies could properly recorded in the space provided for or there is anyscope for further improvement. After conducting pretesting appropriate changes andmodification of the interview schedule have been made. The individuals who respondedin pretesting have been excluded in the final sample selected for the study.

5.7. Techniques of field data collection

The respondents were personally interviewed from October 2016 to June 2017 and October 2017. The items were asked in Bengali version in a simple term so that the members could understand easily. The entries were done in the schedule by student investigator himself at the time of interview.

5.7.1.Some photographs taken during field Data collection

Photograph with NamitaHembramandSumiHembram

Photograph with PutulKshetrapal

Photograph with a farm family

5.8.Statistical tools for Analysis and Interpretation of Data

The statistical methods used for analysis and interpretation of raw data were –

1. Mean
2. Standard deviation
3. Coefficient of Variance
4. Correlation of coefficient
5. Multiple regression analysis
6. Path analysis
7. Factor analysis
8. Canonical correlation Analysis
9. Discriminant Analysis

5.8.1.Mean

The mean is the arithmetic average and is the result obtained when the sum of thevalue 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 reflectsand is affected by every score in the distribution. We can work it out as follows

Mean or (x) =

$$\overline{X} = \frac{\sum_{i=1}^{n} X_i}{n} = \frac{X_1 + X_2 + \ldots.. + X_n}{n}$$

Where,

(x) = The symbol we use for mean (pronounced as x bar)

Σ = Symbol for summation

xi = Value of the ith item X, I = 1, 2, ……………………….n

N = Total number of items.

Mean is the simplest measurement of central tendency and is a widely usedmeasure. Its chief use consists in summarizing the essential features of a series and inenabling data to be compared. It is a relatively stable measure of central tendency. But itsuffers from some limitations *viz.* it is unduly affected by extreme; it may not coincidewith actual value of an item in a series, and it may lead to strong impressions,particularly when the item values are not given with the average. However, mean isbetter than other average, especially in economic and social studies where directquantitative measurements are possible.

5.8.2.Standard Deviation

Standard deviation is the most widely used measure of dispersion of a series and iscommonly denoted by the symbol **σ** (pronounced as sigma) Standard deviation is thesquare root of the arithmetic mean of the square of the deviations, the deviations beingmeasured from the arithmetic mean of distribution.

It is less affected by sampling errors and is more stable measure of dispersion.

worked out as follows,

$$S = \sqrt{\frac{\Sigma(X - \overline{X})^2}{N}}$$

where S = the standard deviation of a sample,
Σ means "sum of,"
X = each value in the data set,
$\overline{X}$ = mean of all values in the data set,
N = number of values in the data set.

5.8.3.Coefficient of Variation

A measure of variation which is independent of the unit of measurement isprovided by Coefficient of variation. Being unit free, this is useful for computation ofvariability between different populations. The Coefficient of variation is standarddeviation expressed as percentage of the mean and is measured by the formula.

$$Coefficient\ of\ Variation = \frac{Standard\ Deviation}{Mean} \times 100$$

$$Coefficient\ of\ Variation = \frac{\sigma}{\mu} \times 100$$

5.8.4. Coefficient of Correlation

When an increase or decrease in one variable is accompanied by an increase ordecrease in other variable, the two are said to be correlated and the phenomenon isknown as correlation. Correlation coefficient (r) is a measure of the relationship betweentwo variables, which are at the interval or ratio level or measurement and are linearlyrelated. A Karl Pearson's coefficient of correlation also known as product moment 'r' iscomputed by the formula :

$$r = \frac{n(\Sigma xy) - (\Sigma x)(\Sigma y)}{\sqrt{[\,n\Sigma x^2 - (\Sigma x)^2\,]\,[\,n\Sigma y^2 - (\Sigma y)^2\,]}}$$

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 x squared, then summed

($\sum x$) 2 = Sum of x scores, squared

$\sum y$= Sum of y scores

$\sum y^2$= each y squared, then summed

($\sum y$)2 = Sum of y scores, squared

This coefficient assumes the following;

- That there is linear relationship between the two variables;
- That the two variables are causally related which means that one of the variable isindependent and other is dependent and;
- A large number of independent causes are operating in both variables so as toproduce a normal distribution.

The value of ‘r’ lies between +1 to -1. Positive values of r indicate that positivecorrelation between the two variables (i.e. changes in both variables take place in thesame direction), whereas negative values of ‘r’ indicate negative correlation i.e. changesin the two variables taking place in opposite direction. A zero value of ‘r’ indicates thatthere is no association between the two variables. When r (+) 1, it indicates perfectpositive correlation and when it is (-) 1, it indicates perfect negative correlation, meaningthereby that variations in independent variable (x) explain 100 per cent of the variationsin the dependent variable (y). We can also say that for a unit change in independentvariable, if there happens to be constant change in the dependent variable in the samedirection, the correlation will be termed as perfect positive. But if such change occurs inthe opposite direction, the correlation will be termed as perfect negative. The value of ‘r’nearer to +1 or -1 indicates high degree of correlation between the two variables.

5.8.5. Regression

The correlation coefficient only expresses association and by itself tells nothingabout the causal relationships of the variables. Thus, purely from the knowledge that twovariables x and y are correlated, we cannot say whether variation in x is the cause or theresults from mutual dependence of the two variables or from common causes affectingboth of them. Similarly, the mere existence of a high value of correlation coefficient isnot necessarily of an underlying relationship between the two variables.

The underlying relation between y and x in a bi-variate population can be expressedin the form of a mathematical equation known as regression equation and is said torepresent the regression of the variable y on the variable x. (Panse and Sukhatme, 1967)

If y is the dependent variable and x is the independent variable, then the linear regressionequation can be written as

y = a + bx

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

Σ(yi – a –bxi)2 with respect to a and b.

The values of a and b are

a= y – bx

b $= \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2}$

The regression equation can now be written as

Where b is the regression coefficient

5.8.6. Stepwise Multiple Regression

Stepwise regression is a variation of multiple regressions which provides a meansof choosing independent variables that yield the best prediction possible with the fewest independent variables. It permits the user to solve a sequence of one or more multiple linear regression problems by stepwise application of the least square method.

At each step in the analysis, a variable is added or removed which results in the greatest production in the error sum of squares (Burroughs Corporation, 1975).

According to Draper and Smith (1981), the method of stepwise multiple regression analysis is to insert variables in turn until the regression equation in satisfactory. The order of insertion is determined by suing the partial correlation coefficient as a measure of the importance of variables not yet in the equation.

The program, according to Burroughs Corporation (1975), first forms a correlation matrix, finds the best predictor (the independent variable having the highest correlation with criterion variable) and performs a regression analysis with this predictor. Then, the second best predictor (independent), and so on. At any given step, the group of predictors being used is not necessarily the best group of that size (i.e. the particular group independent variables does not necessarily have the highest multiple correlation with the criterion that any group of this size does). Rather, this group contains the variables that have the highest individual correlation with the criterion.

Significance of variable that is being considered for entrance into the regression equation is measured by theF-statistic. If F is too small (less than F 'include'), the variable is not added to the regression equation. Include statement establishes the minimum value of the F-statistic required for the inclusion of a variable in the regression equation. In the example which follows, the F-value for inclusion was 0.01.

Significance of variables already in the regression equation may change as new variables are entered. This significance of the variables currently in the equation is also measured by the F-statistic. If F is too small (less than F 'delete'), the variable is not added to the equation. Delete establishes the value of the F-statistic below which the variable is deleted from the regression equation. Here, the F-value for deletion was

0.005.

The 'tolerance' level specified is used as control of degeneracy occurs when variable entered into the equation is a linear combination of variables already present inthe equation. Tolerance statement establishes the maximum value a pivoted element may attain while still allowing its associated variable to be brought into equation. A variableis not brought into the regression equation if its associated pivoted element is below the specified tolerance level, which was 0.001 in the present example

5.8.7. Factor Analysis

Factor analysis is a very useful and popular method of multivariate research technique, mostly used in social and behavioural sciences. This technique is applicable when there is a systematic interdependence among a set of observed or manifest variables, and the researcher is interested in finding out something more fundamental or latent which creates this communality (commonness).

For example, we may have data on farmers' education, occupation, land, house,farm power, material possession, social participation etc. and want to infer from these some factor relating to social status, which shall summarize the communality of all the variables.

According to Kothari (1996), Factor analysis seeks to resolve a large set of measured variables in terms of relatively few categories, known 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 grouped into the factor. The meaning and name ofsuch new variable is subjectively determined by the researcher.Since the factors happen to be linear combinations of data, the coordinates of each observation or variables is measured to obtain what are called factor loadings. Such factors loadings represent the correlation between the particular variable and the factor,and are usually placed in a matrix of correlations between the variable and the factors.

5.8.7.1.Concepts Used In Factor Analysis

Some important concepts used in factor analysis are explained, following Kothari (1996).

A. **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.

B. **Factor loadings**: Factor loadings are those values which explain how closely the variables are related to each one of the factors discovered. Factor loadings work as key to understanding what the factors mean. It is the absolute size (rather the signs, plus or minus) of the loadings that is important in the interpretation of a factor.

C. **Communality (h2)** Communality, represented by h2, shows how much of each variable is accounted for by the underlying factor taken together. A high value of communalitymeans that not much of the variable is left over after whatever the factors represent istaken into consideration.

D. **Eigen value (or latent root):** The sum of squared values of factor loadings relating to afactor is referred to as eigenvalue or latent root. Eigen value indicates the relativeimportance of each factor in accounting for the particular set of variables being analysed.

E. **Rotation:** Rotation reveals different structures in the data and provides meaning to the results of factor analysis. There are different types of rotations such as orthogonal rotations, oblique rotations, varimax rotation etc. One has to select a rotation appropriateto the study. For the present study varimax rotation has been used.

5.8.7.2. Factor Analysis is used:

- To reduce the dimensionality of large number of variables to a fewer number of factors.
- To confirm the hypothesized factor structure by way of testing of hypothesis about the structure of variables in terms of expected number ofsignificant factor loading .

Hence in factor analysis specific and error variablesare excluded and only the common variables are taken into account. There are some steps in factor analysis :

*We have to collect data then we have to work out the correlation between the variables .the variables.

*It is to explore the possibility of data reduction i.e. initial steps of factor are to be explored.

The common method of extraction of factors isPrincipleComponent Analysis (P.C.A.) .

5.8.8. Principal Component Analysis

There are several methods of factor analysis. The method of Principal ComponentAnalysis which is widely used is discussed here.

The principal component analysis extracts m-eigenvectors (principal component axes)and corresponding m-eigenvalues (the variance measured along the eigenvector)from m x m symmetrical matrix of correlation.

The eigenvectors obtained from this principal component analysis are allorthogonal (i.e. inter-column correlations are near zero). The eigenvalues account for allof the original data variances in decreasing order such that each has variance oreigenvalue less than the previous ones. The total of the eigenvalues

$$(\lambda_1 + \lambda_2 + \cdots \ldots \ldots \ldots . + \lambda_m ,)$$

This is the same as the sum of the variances constituting the diagonal or trace of thecorrelation matrix before transformation. The principal components are then convertedinto factors by multiplying each element of the principal components or eigenvectors (v)by the square – root of the corresponding eigenvalues ($\lambda^{\frac{1}{2}}$. v). Factors, thus, besides thedirection also represent the variances.Kaiser (1958) and others have recommendedretaining all those eigenvalues, which have values more than one.

Next step is to remove the noise imposed by (m-p) unnecessary axes. Toaccomplish this, p-orthogonal reference axes or factors are routed about the origin topositions such that the variance of the loading from each variable onto each factor axis iseither extreme (±1) or zero. This maximization of the range of the loadings wasperformed by using Kaiser's Varimax criterion. Scanning through each factor column for large absolute values in the varimax matrix will reveal a few variables with significantly highloadings and many others with insignificantly loadings. The column showing communality is the total amount of variance of each variable retained in the factors, and is computed by summing the squares of the elements of the factors in each row of varimax matrix. Fairly high communality of each variable implies the appropriateness of the model adopted, for the study. The last step involved interpretation of the factors .

5.8.9. Canonical correlation analysis:

In statistics, canonical-correlation analysis (CCA) is a way of making sense of cross-covariance matrices. If we have two vectors X = (X1, ..., Xn) and Y = (Y1, ..., Ym) of random variables, and there are correlations among the variables, then canonical-correlation analysis will find linear combinations of the Xi and Yj which have maximum correlation with each other. Virtually all of the commonly encountered parametric tests of significance can be treated as special cases of canonical-correlation analysis, which is the general procedure for investigating the relationships between two sets of variables. The method was first introduced by Harold Hotelling in 1936.

Given two column vectors $X = (x_1, \dots, x_n)'$ and $Y = (y_1, \dots, y_m)'$ of random variables with finite second moments, one may define the cross-covariance $\Sigma_{XY} = \operatorname{cov}(X, Y)$ to be the $n \times m$ matrix whose (i,j) entry is the covariance $\operatorname{cov}(x_i, y_j)$. In practice, we would estimate the covariance matrix based on sampled data from X and Y (i.e. from a pair of data matrices).

Canonical-correlationanalysisseeksvectorsa'andb'suchthattherandom

Variables $a'X$and $b'Y$maximize the correlation $\rho = \operatorname{corr}(a'X, b'Y)$. The random variables$U = a'X$ and $V = b'Y$ are the first pair of canonical variables. Then one seeks vectors maximizing the same correlation subject to the constraint that they are to be uncorrelated with the first pair of canonical variables; this gives the second pair of canonical variables. This procedure may be continued up to $\min\{m, n\}$times.

5.8.10. Discriminant Analysis :

Discriminant analysis is a useful statistical technique to classify an observation into one or several apriori groups that is dependent upon the individual's characteristics. To distinguish between the groups, the researcher selects a collection of discriminating variables that measures characteristics on which the groups are expected to differ. Discriminant function analysis is Multivariate Analysis of Variance (MANOVA) reversed. In discriminant analysis, the independent variables are the predictors and the dependent variables are the groups.