TITLE PAGE

CONSTRUCTION OF DISCRIMINANT ANALYSIS FOR REMEDIAL PROGRAMME (A CASE OF FEDERAL UNIVERSITY OF TECHNOLOGY, MINNA)

BY

ALAYA AYODELE MUH'D SANNI
PGD/MSC/2000/2001/1053.

A PROJECT SUBMITTED TO THE DEPARTMENT OF MATHEMATICS/ COMPUTER SCIENCE, FEDERAL UNIVERSITY OF TECHNOLOGY, MINNA IN PARTIAL FUFILMENT OF THE REQUIREMENTS FOR THE AWARD OF POST GRADUATE DIPLOMA IN COMPUTER SCIENCE.

APRIL 2002.

APPROVAL PAGE

This is to certify that this project has been read and approved, meeting the requirement for the award of Post Graduate Diploma in Computer Science in the Department of Mathematics/Computer Science, Federal University of Technology, Minna.

....

Mr. Isah Audu Date

Project Supervisor

....

Head of Department Date

L. N. Ezeako

....

External Examiner Date

DEDICATION

To Umar for his passion, caring and love.

ACKNOWLEDGEMENT

All praises and thanks be to Allah for his mercies on me and for making me completed this programme. May his blessing be upon his exalted prophet, his household and his Companions.

I wish to acknowledge my able and dedicated supervisor, for encouragement and constructive criticism during the course of my writing this project. May Almighty Allah be with him and his family. I wish to thank the entire lecturers of Mathematics/Computer Science department particularly Mr. Hakimi Danladi for their effort in this noble objective.

My profound gratitude and acknowledgement goes to my brother Umar Zanna Ashgar. I am gratefully indebted to him for his support morally and financially throughout this programme. May Almighty Allah Continue to be with him and his family in this world and hereafter.

I cannot forget my family whom through them I came to this world. My sincere appreciation and gratitude goes to them particularly my mother Alhaja Batully Alaya. My sisters and brothers for their backing through prayers.

I am pleased to thank my sister – Chinyere Chukwudome Augusta for her kindly assistance in writing this project. My colleagues Sheilk Alhaji Mohammed and Basiru Sanni for their friendly, accommodative and understanding. I wish to extend gratitude to my sister Halimat .B. Zubairu for her contribution toward this noble goal. I am also pleased to thank my good friend Mr. Yusuf Shawai whom through him I appreciated friendships. Last but not the least, all those who have contributed one way

or the other. It is not possible to mention everybody here but I thank everyone in whatever way he or she has contributed to this project.

ABSTRACT

Since the inception of this world, everything was created mathematically. That is, world was planned numerically. Nothing in this world that cannot describe or expressed numerically. Everything was built numerically.

This representation of things numerically was to make things in their rightful and appropriate place so that world can better place to live. With this building of things numerically, everything will be in control and checked for appropriate organisation and proper administration. Therefore, life without number that life will not be easy or worth living.

However, in order to have well organized and planned environment that we intend to construct the Fisher's discriminant analysis function to classify remedial student of Federal university Technology, Minna, to the appropriate (suited) school in order to have better and well organized school based on good administration.

The classification of these students will be based on their numerical characteristics which are their examination's score (marks) obtained. It is on these characteristics that we reclassified them to appropriate schools. The topic of this project work- construction of discriminant analysis-has been widely used statistical technique in many areas of applications in many fields. It is a very useful good technique in classification analysis.

TABLE OF CONTENTS

CHAPTER ONE

INTRODUCTION

1.1. GENERAL INTRODUCTION ON DISCRIMINANT ANALYSIS AND CLASSIFICATION.

Discriminant analysis is concerned with the problem of identifying the population from which (a vector (x) have come from one or two or more population) was Drawn. Discriminant analysis deals with the assignment of an observation say x of an unknown origin to one of two (or more) distinct populations on the basis of the value of the observation. According to PETER A. LACHENBRSUCH (1975) discriminant analysis was viewed as the problem of assigning an unknown observation to a group with low error rate. COOLEY and LOHNES (1962) defined discriminant analysis as a technique for description and testing of between group difference"

Linear discriminant analysis (LDA) is per haps the most widely used method for classification because of it simplicity and optimal properties. Linear discriminant analysis is known to be optimal (in the sense of minimizing the expected cost of classification for two multivariate normal group with equal covariance matrices). Estimation of linear discriminant function is considered for the problem of discriminant between P-dimensional normal population and with mean vectors and common covariance matrix.

Discriminant and classification one quite separate, though closely related, concepts. Classification consist of an attempt to discover, usually with no prior information, the number of groups that exit within a given data-set, together with details of group membership. If the number of group is known, as well as the characteristic of each group, the problem of assigning unclassified observation to their most likely group becomes logically possible. Discriminant involves the assignment of newly acquired or previously uncategorized samples to one of the existing classes. Classification is the identification of the category or group to which an individual or object belongs on the basis of its observed characteristics. When the characteristic are a number of numerical measurements, the assignment of group is called by some statistician DISCRIMINATION and the combination of measurement used is called DISCRIMINANT FUNCTION. The discriminant function discriminates between two (or more) completely specified populations. Therefore, the procedure of assigning an observation, X of unknown origin to one of two (or more) distinct group on the basis of the value of the observation is called DISCRIMINANT ANALYIS.

The main assumptions of discriminant analysis are that group have homogenous variance-covariance matrices, and that the variables are normally distributed. The second assumption is necessary of significance test are to be applied. In discriminant analysis often a two step procedure is to be followed. First, training samples are obtained to set up a discriminant rule, and then individuals are classified using the sample-base rule. However, if the criterion for assigning the training samples for their true classes is imperfect, some training samples may be classified. To classify an individual into either of the few groups, (or more) we need a criterion of goodness of classification. This implies that we need a rule that will lead to an optimum classification rule. We need a discriminant function to construct our assignment rule. The classification procedure is to regard the observed value x as coming from either the first population, or the second population (or other population) according to discriminant value.

Although, two-group discriminant function analysis has been widely used, it is more often the case that several groups are thought to exist in a given set of data. While it is true that each pair of group could be analysis separately. Therefore, two-group discriminant analysis has been generalized to deal with k-group case called Multiple Discriminant Analysis (MDA). The number of comparison grows rapidly with k, the number of groups. It is assumed that each of the "n" samples is drawn from a separate population. The "n" population will differ in their means (otherwise there would be no point in discriminating between them) but their variance-covariance matrices should be equal and the variables on which measurement are made should be normally distributed. It is considered that moderate departure from these ideal conditions do not have a serious effect on the results.

Multiple discriminant analysis (MDA) provides for the simultaneous comparison of several group, every member of each group being measured on a number of variables. The assumptions of multiple discriminant analysis are the same as those of two-group discriminant analysis, namely, Multivariate normality and equality of the within group covariance matrices. A logical pre-requisite of multiple discriminant analysis is that the group are, infant separate.

1.2 EXAMPLES OF DISCRIMINANT ANALYSIS PROBLEM.

(i) A geologist has obtained the mean, variance, skewness and kurtosis of the size of particular deposited in a beach. How can these statistics be used to determine if the beach is wave laid or colin in origin? Of course these statistics through the measurement can be used to determine whether the beach is wave laid or colin in origin.

(ii) Prospective students applying for admission into college are given a battery of test, the vector of scores is a set of measurements x. The prospective student may be a member of one population consisting of those students who will successfully complete the training or rather have potentialities for successfully completing the training or he may be a member of the other population, those who will not complete the college course successfully. The problem is to classify a student applying for admission on the basis of his scores on entrance examination.

(iii) A patient is admitted into a hospital with a diastolic of myocardinal infection, a systolic blood pressure, diagnose blood pressure, stroke index heart rate and mean curterial are obtained. Is it possible to predict whether the patient will survive? Can we use these measurements to compute the probability of survival for the patient? The answer is yes, it is possible to predict whether the patient will survive and also that measurement can be used to compute the probability of survival.

(iv) In routine banking or commercial finance an officer or analyst may wish to classify loan applicants as low or high credit risk on the basis of the elements of certain accounting statements.

(v) Indian man have been classified into three centres on the basis of stature, sitting, height and nasal depth and height (RAO 1984).

(vi) Six measurements on a skull found in England were used to determine whether it belongs to the Bronze Age or the iron age (RAO 1952).

1.3. THE PROBLEMS OF CLASSIFICATION.

The problem of Classification in its most basic form arises when it is required to allocate an individual to one or other of two population on the basis of a measurement of a p-dimensional random variables on the individual. It is presumed that the random variable has a different distribution for each of the populations. The problem of classification also arises when an investigator makes a number of measurements on individual and wishes to classify the individual into one of two (or more) distinct groups on the basis of those measurements.

Similarly, the problem of classification arises when the investigator cannot associate the individual directly with a category but must infer the category from the individual's measurement, response or other characteristics. In many cases, it can be assumed that there are a finite number of populations from which the individual may have come and that each Population is described by a statistical distribution of the characteristics of individuals. The individual to be classified is considered as a random observation from one of the populations. The question is, given an individual with certain measurements, from which population did he arises?

1.4. HISTORICAL BACKGROUND OF THE CASE STUDY

The Federal University of Technology, Minna established on 1st February, 1983 was the last of the seven Federal Universities of technology established by the defunct civilian administration of the second republic in Nigeria.

1.4.1 THE GOALS AND OBJECTIVES OF THE UNIVERSITY

i. To encourage the advancement of learning and to hold out all persons, without distinction of race, religion, creed, sex, or political conviction the opportunity of acquiring a higher education in Technology.

ii. To develop and offer academic and professional programmes leading to the award of certificates, diplomas, first degree, post graduate research and higher degrees, which emphasize planning, adaptive, technical, maintenance, development and productive skills in the engineering scientific, agricultural, medical and allied professional discipline with the aim of producing socially matured men and women with the capability not only to understand, use and adapt existing technology but also to develop new ones.

iii To act as agents and catalyst, through post graduate training, research and innovation for the effective and economic utilization, exploitation and conservation of the country's national economic and human resources.

iv To offer to the general public, as a forum of public service, the result of training and research and to foster the practical application of these results.

v. To establish appropriate relationship with other national institutions involved in training, research and development of technology.

vi. To identify technological problems based on the needs of society and to find a solution to them within the context of overall national development.

vii. To provide and promotes sound basic scientific training as a foundation for the development of technology and the applied science, taking into account the indigenous culture and the need to enhance national unity.

viii. To undertake another activities appropriate for a University of Technology of the highest standard.

1.4.2 ACADEMIC PROGRAMMES

The University presently runs different types of academic programmes. These are-pre-degree, full degree and postgraduate degree programmes.

The Federal University of Technology is divided into four schools. These are:

i School of Agriculture and Agriculture Technology.

ii School of Engineering and Engineering Technology.

iii School of Science and Science Education.

iv School of Environmental Technology.

1.4.3. PRE – DEGREE PROGRAMMES

The University offers a one-year remedial programme designed to prepare candidates for admission into the five – year, full – time degree programme. Such candidates when admitted will be attached to the school of science and science education and take courses in the following five compulsory subjects – Mathematics, English Language, Physics, Chemistry, Biology.

1.4.4. ADMISSION REQUIREMENTS

Admission into the pre – degree programmes will usually require:

i. At least four-credit level passes at West Africa School Certificate in Science subjects.

ii At least four passes at General certificate of Education, ordinary level.

iii. At least four-credit/Merit level passes at Grade II Teachers' Certificate, passes in English Language and Mathematics will be an added advantage.

1.5 OBJECTIVES OF THE STUDY

The main objectives of this study include:

i. To construct the discriminant analysis for remedial programme.

ii To use a classification rule using Fisher's criterion to classify students into appropriate (suited) schools.

iii. To examine an apparent error rate from such a classification rule.

iv. To advice future remedial programme entrants on the appropriate (suited) schools.

1.6. SCOPE OF THE STUDY

i Analysis is bare on two – group fisher's discriminant analysis function due to lack of materials on multiple discriminant analysis on the part of the researcher.

ii Analysis is based between two schools.

iii Study only based on three compulsory subjects.

iv Study does not concerned about how admission were given to the students.

v. Only limited to the randomly selected students offer a one year remedial programme admission into university.

1.7 SOURCES OF THE DATA

The data (Secondary data) used for the analysis of this project work were collected from the office of Dean of school of science and Science Education, Federal University of Technology, Minna. The data are the 2000/2001 session Examination results of candidates who applied for admission into school of Science and Science Education and School of Environmental Technology.

CHAPTER TWO

LITERATURE REVIEW

2.1. INTRODUCTION

The aim of discriminant analysis is to find the line which best separate the groups in form of the projection of the group characteristics. The characteristics of the group can then be used to assign individual to their most probable class. Discriminant analysis is equivalent to the regression of inter – group mean different on the P – variable KENDALL (1965).

AFIFI and CLARK (1990) describe discriminant analysis as a technique, which are used to classify individual into one of two or more alternative groups (or population) on the basis of a set of measurement. The populations are known to distinct and each individual belongs to one of them.

FISHER (1936) introduced the method of discriminant analysis to deal with the problem of correctly assigning fossil remains to come of two classes (homimaid and ape) on the basis of measurement of several variables. Since then the topic has been discussed in the context of psychology, geography and geology.

THOMAS (1969) used two – group discriminant analysis in a study of glacial and periglacial sediments taken from slop deposited of the northern uplands of the Isle of Man. Group I consisted of colifcuxion deposited, and Group II was made up of interbedded gravels. The purpose of the analysis was to determine whether there was a statistically significant different between the two groups in term of the four variables X_1 (mean), X_2 (sorting), X_3 (skewness) and X_4 (kurtosis) computed according to the methods of Folk and Ward. Secondly, if it was found that there was a significant difference between the groups on the basis of these variables, an allocation procedure was required whereby future samples of sediment from this particular area of the Isle of man could be placed in group 1 or group 2 on the basis of their scores on the four variables.

KLOVAN (1966) employed Q- mode factor analysis (i.e analysis of a similarity matrix of individuals) to classify a group of 69 mean shore sediment samples from Baratria bay, on the Mississippi delta.

Studies on discrete data were performed by GILBERT (1969) and MOORE (1973) using multivariate Bernoulli distributions. LACHENBRUEH et al., (1973) considered the robustness to certain types of continuous non- normality. They transformed p-independent normal variables to the lognormal, and the Sinh^{-1} normal distributions. As a result of these non- linear transformations, the underlying observations no longer have the same covariance matrices. Therefore, the total misclassification rate was often greatly increased and the total individual misclassification rate were distorted in such a way that are error rate was increased and other was decreased. The smallest effect was for the logit – normal distribution, which are highly skewed.

CHERNOFF (1972, 1973) suggested some measures that indicates how well one can discriminate between Multivariate normal population with unequal covariance matrices using a linear discriminant functions. Chernoff used such criteria to compare the performance of linear discriminant functions based on balanced and unbalanced design.

ANDERSON and BAHADUR (1962) studies procedures for classifying two multivariate distributions with unequal covariance matrices. They showed how to construct a discriminant function that minimizes one probability of misclassification given the other and how to obtain a minimax discriminant procedure. But these discriminant procedures are non- linear. The best linear discriminant for these unequal covariance matrix context was found by CLUNIES –ROSS and RIFFENBERG (1960) and ANDERSON and BAHADUR (1962).

The effect of unequal covariance matrices on the linear discriminant analysis were studied by GILBERT (1969) for the large –sample case and by MARKS and DUN (1974) for small –sample case. They both concluded that the linear function quite satisfactory provided that the covariance matrices are not too different. An important result on non-parametric estimation of linear classification was suggested by GREER (1979, 1984). Greer considered algorithms designed to produce hyperplanes on a completely non- parametric manner for a large set of loss function. The estimation produce associated with a suitable loss function is consistent for any two underlying distributions, whether these are continuous, discrete or mixture of the two.

The principal features of discriminant analysis is that it allows several characters common to the two group to the examined but collapses these multiple characters into one for purpose of testing. The robustness of linear discriminant analysis and the effect of failure of assumptions to hold have been studies by GILBERT (1969), MOORE (1973), MARKS and DUNN (1974), LACHENBRUCH, SNEERINGER and REVO (1973) and LACHENBRUCH (1975).
BREIMAN et al (1984) proposed a non- parametric method that yields a perfect selection of two groups, when possible. Their method, however, has no continuos scoring system.

JEAN (1988) used a non- parametric discriminant analysis based on the construction of a binary decision tree procedure. He concluded that the discriminant tree procedure is a non- parametric method of discrimination of qualitative variables (binary, normal or ordinal).
Its prediction rule, given in the form of a binary decision tree, is easy to understand, use, explain, interpret and close to the physician reasoning. The method takes into account interaction between variables, is able to handle missing data, provides the possibilities of selecting splits to include Boolean combination of variables and is also able to deal with different cost of misclassification. Unlike the standard classical method of classification problems discriminant analysis and logistic regression but these cases, the prediction rule are given in the form of algebraic expression that are sometimes, difficult to understand and interpret.

AKIHIKO (1987) studied experiment comparison between the optimal discriminant plane based on samples and general discriminant analysis where he obtained optimal discriminant function (O. D. F) from the ratio between the population distributions functions of two group in which the boundary is determined differently by the likelihood, Bayes, risk methods.

He stated that when two populations are normal, optimal discriminant function is of quadratic form and it coincides with Anderson – Bahodour linear discriminant function. When their variance covariance matrices are equal it is linear form and it coincides with optimal discriminant function.

In comparison with the standard discriminant analysis, the linear co-efficient of which maximizes the ratio of the between –class variance to the within –class variance is frequently used as standard linear discriminant function- Fisher linear discriminant function (F.L.D.F).
He concluded that it is doubtless that fisher linear discriminant function is every useful, especially efficient at small sample size. Nevertheless, when population distribution, Fisher linear discriminant function does not converge to the optimal linear discriminant function for the population distribution as sample size be comes infinitely large.

On the other hand, the sample optimal linear discriminant functions needs vast calculation to obtain it, but it logical is very clear and

always converges to the optimal linear discriminant function for any population.
Occasionally, this non- parametric method may be very useful.

In some situation quantitative measurement of the attributes of interest are not possible. Instead, binary (presence-absence) measures may be considered.

RAMSAYER and BONHAM-CARTER (1974) describe a technique, which they term "adaptive pattern-recognition" which allows the discrimination between two groups on the basis of binary attributes alone. The adaptive pattern – recognition model can be represented by a linear equation except that the X's denote binary variables. The procedure begins with a trial set of coefficients, which are progressively modified until maximum discrimination is attained.

One drawback of the method is that the order in which the individual are presented may have an effect on the resulting coefficients. It can be shown that if a linear function exists which can separate the groups then the coefficients vector will converge to a solution. This solution will not necessarily be unique, for if the individuals had been listed in a different order then the vector could have converge to a different and equally feasible solution.

Ramsayer and Bonham – carter report their experience with the algorithm using actual geological data. For the sedimentological data recorded and analysed by Purdy, they find that the adaptive pattern –

recognition algorithm was approximately as successful as the discriminant function. In a second application to the study of breachiopod diversity patterns in Permian marine roots, Ramsayer and Bonham – carter carried out pair – by – pair analyses for every possible pair out of seven groups. Some difficulty was experienced in that situations where an individual could not be allocated to any particular group with any degree of certainty there is no clear basis for making probability statements about the likelihood of the individual belonging to any particular group. But, this is possible with both two – group and multiple group discriminant analysis.

FISHER (1936) suggested using a linear combination of observations and choosing the coefficients so that the ratio of the difference of the means of the linear combination in the two groups to its variance is maximized.

Fisher's linear discriminant function is frequently used for the two group discriminant analysis problems. It minimizes the expected loss in the case of unknown prior probabilities and it is an admissible procedure when prior probabilities are not known (ANDERSON 1958). However, in most practical situations, only sample data (one set from each population) are available and the parameters as well as the shapes of the two distributions are not known.

In this case, use of the linear discriminant function is defended by the argument that it maximizes the sample Mahalanobi's (squared) distance between the two data sets (Fisher 1936). This argument nonetheless does not imply that the linear discriminant function (L.D.F) is the best procedure for these situations.

Although, Fisher's linear discriminant function has been used in many practical applications, its statistical properties under non – optimal conditions have met received much attention until recently.

In research paper (project work), reviewing of some literatures on the subject matter (topic) is very important because of some advantages attached to it. How ever, we would like to point – out (emphasize) from this paper some advantages derived from reviewing this (Statistical tool) topic – discriminant analysis that are widely applied (used) in many areas or fields of research and these as follows;

i. In reviewing, we are able to understand the detail concept of the subject matter. Insight into the meanings presented with examples able to give us clear concise and precise understand of the topic. Different view points put forward through definitions so as to let us know what this topic is all about and how important in the context of our finding (research).

ii. From the literature review of this paper one is able to get acquainted with different methodology of application to different situations. Alternative methods were proffered in a situation where failure of assumption could not longer be held of a certain technique. Testing different methods to different situation whether it is applicable to be employ to that situation or not were given so that we would be able to know which method is applicable to our situation in terms of optimal method. We know how feasible certain methodology in terms of flexibility, analysis, interpretation, explanation and cost to suite our problem at hand

iii. We able to known or learned the limitation of hypothesis or theory put forward in some areas of findings in this particular topic. Its shortfall, problems and effect in comparing different methodology to different situations whether it is appropriate in terms of applicability.

iv. Another advantage is that it open – up further area of research in some fields. From the literature review we are able to detent some area of further study. Which open – up to carry out research on it, in terms of its principles and laws.

v. Exposure and awareness we are able to achieved from this literature review. Exposure to the latest research work or some new studies to various application of areas or fields. Dynamic aspects of the subject matter as relating to the contemporary issue were fully aware about.

2.2 FISHER'S CRITERION AND DERIVATION

In Fisher's approach, let the linear combination be denoted by $Y=\lambda x$. Then, the mean of Y in the first population π_1 is $\lambda^1\mu_1$; and the mean of Y in the second population π_2 is $\lambda^1\mu_2$; its variance is $\lambda^1\Sigma\lambda$ in either of the groups if we assume that the covariance matrices is equal.

i.e. $\Sigma_1 = \Sigma_2 = \Sigma$

Then we wish to choose to maximize

$$\phi = \frac{\lambda^1\mu_1 - \lambda^1\mu_2}{\lambda^1\Sigma\lambda}$$

Differentiating ϕ with respect to λ^1, we have

$$\frac{\partial\phi}{\partial\lambda} = \frac{2(\mu_1 - \mu_2)\lambda^1\Sigma\lambda - 2\Sigma\lambda(\lambda^1\mu_1 - \lambda^1\mu_2)}{(\lambda^1\Sigma\lambda)^2}$$

Equating

$$\frac{\partial\phi}{\partial\lambda} = 0 \qquad \text{we have } \mu_1 - \mu_2 = \frac{\Sigma\lambda(\lambda^1\mu_1 - \lambda^1\mu_2)}{\lambda^1\Sigma\lambda}$$

Since λ is used only to separate the two population then λ may be multiplied by any constant. Thus, λ is proportional to $\Sigma^{-1}(\mu_1 - \mu_2)$

i.e. $\lambda \propto \Sigma^{-1}(\mu_1 - \mu_2)$

Therefore $\lambda = k\Sigma^{-1}(\mu_1 - \mu_2)$ where k is a constant

Then

$$Y = \lambda'\underline{X} = (\mu_1 - \mu_2)\Sigma^{-1}\underline{X}$$

If the parameters $\mu_1\mu_2$ and Σ are unknown, it is the usual practice to estimate them by $\overline{X}_1, \overline{X}_2$ and S respectively. The discriminant function for known parameters is $Y = \lambda'\underline{X} = (\mu_1 - \mu_2)\Sigma^{-1}X$. The discriminant function for unknown parameters is $\hat{Y} = (\overline{X}_1 - \overline{X}_2)^1 S^{-1}X$

The mean of Y in π_1 is $\overline{Y}_1 = (\overline{X}_1 - \overline{X}_2)^1 S^{-1}\overline{X}_1$

The mean of Y in π_2 is $\overline{Y}_2 = (\overline{X}_1 - \overline{X}_2)^1 S^{-1}X_2$

The mid – point of the interval between mean

$\overline{Y}_1$ and $\overline{Y}_2$ is $\frac{1}{2}(\overline{Y}_1 + \overline{Y}_2)$.

The assignment procedure is to assign an individual to π_1

if $Y > \frac{1}{2}(\overline{Y}_1 + \overline{Y}_2)$ and assign to π_2 if $Y \leq \frac{1}{2}(\overline{Y}_1 + \overline{Y}_2)$.

The mid – point ½ $(\overline{Y}_1 + \overline{Y}_2)$ is used as the cut off point for the assignment procedure. The difference between the means $\overline{Y}_1$ and $\overline{Y}_2$ is $\overline{Y}_1 - \overline{Y}_2 = (\overline{X}_1 - \overline{X}_2)^1 S^{-1}X - (\overline{X}_1 - \overline{X}_2)^1 S^{-1}X_2$

$$= (\overline{X}_1 - \overline{X}_2)^1 S^{-1} (\overline{X}_1 - \overline{X}_2)$$

$= D^2$ which is called the Mahalanobi's (squared) distance for unknown parameters. The distribution of D^2 is used to test if there are significant differences between the two groups.

In general, P – measurements made on sample of size n_1 and n_2 from populations π_1 and π_2 respectively are discussed as follows:

The mean vector of measurement in π_1 is given by

$$\bar{X}_1 = \left(\bar{X}_1 \bar{X}_2 \ldots\ldots\ldots\ldots\ldots \bar{X}_p \right) 1p$$

The mean vector of measurement in π_2 is given by

$$\bar{X}_2 = \left(\bar{X}_1 \bar{X}_2 \ldots\ldots\ldots\ldots\ldots \bar{X}_p \right) 1p$$

The covariance between measurement in π_1 is given by P x P matrix.

$$S_1 = \frac{1}{n_1 - 1} \sum_{i=1}^{n} (X_{ii} - \overline{X}_i)^1 (X_{ii} - \overline{X}_i)$$

The covariance between measurement in π_2 is given by P x P matrix.

$$S_2 = \frac{1}{n_2 - 1} \sum_{i=1}^{n} (X_{ii} - \overline{X}_2)^1 (X_{ii} - \overline{X}_2)$$

The covariance between measurement in π_j is given by

$$S_j = \frac{1}{n_j - 1} \sum_{i=1}^{n} (X_{ji} - \overline{X}_j)^1 (X_{ji} - \overline{X}_j)$$

Where $i = 1,2,\ldots.. N_j$ $\quad j = 1,2 \ldots\ldots$

The discriminant score for n_1 and n_2 individuals in π_1 and π_2 respectively are given below.

$$\pi_1$$

$$Y_1 = \lambda_1 X_{11} + \ldots\ldots + \lambda_p X_{p1}$$

$$Yn_1 = \lambda_1 X_{n1} + \ldots\ldots + \lambda_p X_{pn1}$$

$$\pi_2$$

$$Y_1 = \lambda_1 X_{11} + \ldots\ldots + \lambda_p X_{p1}$$

$$Y_{n2} = \lambda_1 X_{n2} + \ldots\ldots + \lambda_p X_{pn2}$$

2.3. TEST BETWEEN – GROUP DIFFERENCES

The distribution of Mahanlanobi's (squared) distance, D^2 is used to test if there are significant differences between the two groups.

$$F = \frac{n_1 n_2 (n_1 + n_2 - K - 1) D^2}{n_1 n_2 (n_1 + n_2 - 2)k}$$

Where:

n_1 is the sample size in the first population

n_2 is the sample size in the second population.

K is the number of variables, has an F – distribution with k and $n_1 + n_2 - 1$ degree of freedom.

HYPOTHESIS: The null hypothesis, Ho States that there are no significant differences between the two populations. While the alternative hypothesis, H_1 States that there are significant differences between the two populations.

TEST STATISTICS: Let the probability of committing type I error, $(1 - \alpha) = 95\%$, that is, α is at 5% significant level ($\alpha = 0.05$).

DECISION RULE: We reject the null hypothesis if the calculated F is greater than F_0, k, $n_1 + n_2 - k - 1$. Otherwise we do not reject null hypothesis, Ho.

CONCLUSION: If the null hypothesis is rejected, we conclude that there are significant differences between the two population at 5% significant level.

2.4. PROCEDURE FOR CALCULATING PROBABILITY OF MISCLASSIFICATION USING FISHER'S CRITERION.

STATISTICIANS DECISION TABLE

		π_1	π_2	TOTAL
TRUE POPULATION	π_1	CORRECT POPULATION	M_1	N_1
	π_2	M_2	CORRECT POPULATION	N_2
	TOTAL			N

Therefore,

$$P_1 = \frac{M_1}{n_1}\ ;\ P_2 = \frac{M_2}{n_2}\ ;\ \overline{P} = \frac{M_1 + M_2}{n_1 + n_2}$$

Where

P_1 is the probability of misclassification in π_1.

P_2 is the probability of misclassification in π_2.

$\bar{P}$ is the probability of error rate.

N is the grand total

M_1 is the number of the misclassification in π_1.

M_2 is the number of the misclassification in π_2.

2.5. OTHER CRITERIONS.

i. WELCH'S CRITERION

This is a discriminant function which follows a likelihood ratio approach. This linear discriminant function was discovered by Welch (1936), although it follows the operation of Fisher's linear discriminant function.

Suppose the random variable X has either the density function $f_1 (X_1\theta_1)$ or $f_2 (X_2\theta_2)$ where parameters θ_i are unknown and mathematical forms of the densities are specified. If the likelihood $f_1 (X_1\theta_1)$ is large relative to $f_2 (X_2\theta_2)$, we could be inclined to believe that x come from the first population, if $f_2 (X_2\theta_2)$ has the large value, the second population would seem more likely. This rule may be written in terms of the likelihood ratio

$$\lambda = \frac{f_1 (X_1\theta_1)}{f_2 (X_2\theta_2)}$$

i.e. classify x as from first population if $\lambda = 1$ and from second population if $\lambda \leq 1$. Welch obtained a discriminant function by minimizing the total probability of misclassification.

Let $f_1 (x)$ be the density function of Xi

Let P_1 be the proportion of π_1 in the general population and $P_2 = 1 - P$, be the proportion of π_2 in the second population.

We think of an observation as a point on a k – dimensional space. We divide this space into two regions R_1 and R_2 where R_1 and R_2 are mutually exclusive and these Union includes the entire space R. If the observation X falls in R_1, we classify it as coming from π_1 and if it falls in R_2, we classify it as coming from π_2.

In the following any given classification procedure, there are two kinds of error in classification. The first type of misclassification is

$$P_1 = \int R_1 \, f1 \, (x) \, dx \text{ and the}$$

Second type is

$$P_2 = \int R_2 \, f_2 \, (x) dx$$

Welch suggested minimizing the Total probability of misclassification.

$$T(R_1 f) = P_1 \, SR_1 \, f_1 \, (x) \, dx + P_2 \, SR_2 \, f_2 \, (x) \, dx$$
$$= P_1 \, \{1 - SR_1 \, f_1 \, (x) \, dx\} + P_2 \, SR_2 \, (x) \, dx$$
$$= P_1 + SR_1 \, \{P_2 \, f_2 \, (x) - P_1 \, f_1 \, (1)\} dx.$$

This quantity is minimized if R_1 is chosen so that $P_2 \, f_2 \, (x) \, P_1 \, f_1 \, (x) < 0$, for all points in R_1. Thus, the classification rule is

Assign x to π_1 if $\frac{f_1(x)}{f_2(x)} > \frac{P_2}{P_1}$ and

Assign x to π_1 if $\frac{f_1(x)}{f_2(x)} \leq \frac{P_2}{P_1}$

An important special case on Fisher's criterion is when π_1 and π_2 are multivariate normal with meanμ_1, μ_2 and common covariance matrixΣ.

The optimal rule is to assign x to π_1 if $D_T(x) = \log P_2/P_1$ and assign x to π_2 otherwise.

If parameters are unknown, the sampling analogue is

$D_s(x) = [x - ½ (\overline{x}_1 - \overline{x}_2)] S^{-1} (x_1 - x_2)$

This is called Welch's sample discriminant function while $D_T(x)$ is called true discriminant function due to Welch. The mean of $D_T(x)$ in X comes from the population, the variance is

$$E\{D_T(x) - D_T\}^2 = (\mu_1 - \mu_2)^1 \Sigma^{-1} (\mu_1 - \mu_2) = D^2$$

The quantity σ^2 is the Mahalanolities distance for unknown parameters. The probabilities of misclassification are

$$P_1 = p_r\left[D_T(x) < \log\frac{1-p}{p_1}\right] = \frac{\phi\left[\log\frac{1-p_1}{p_1} - \frac{\sigma^2}{2}\right]}{\sigma}$$

$$P_2 = \frac{\phi\left[\left(-\log\frac{1-p_2}{p_2} + \frac{\sigma^2}{2}\right)\right]}{\sigma^2}$$

ii. BAYE'S THEOREM APPROACH.

This is the assignment of an observation, x to the population with the largest posterior probability. The conditional density of x given π_i is $f_i(x)$. Some of the prior probability of π_i is P_i, the posterior probability by Baye's theory is

$$P_r(\pi_i / x) = \frac{P_r(\pi_i x)}{P_r(x)} = \frac{P_i f_i(x)}{P_1 f_1(x) + P_2 f_2(x)}$$

If an observation is assign to π_I when P_r (π_i /x) > P_r (π_l/x), this is equivalent to the rules that minimizes the total probability of misclassification. When estimating the risk of belonging to, the posterior probability is useful. The Baye's approach towards classification when all parameters are known and misclassification costs are equal will begin with an evaluation of the posterior probability that $X\Sigma\pi_j$ given for each j = 1, 2 ..... k. Then posterior odds might be computed for each pair of population, alternatively, with k > 2, the population with the greater posterior probability density can be selected. When the cost of misclassification is unequal, the Bayesian would select the population that produced a minimum cost when average with respect to the posterior distribution. Moreover, this result also holds for all k > 2 when all parameters are known.

CHAPTER THREE

ANALYSIS OF DATA

3.1. INTRODUCTION

In this chapter, we are going to employ fisher's linear discriminant analysis function procedure to construct an assignment rule in our classification. The choice of chosen Fisher's criterion for our analysis of data is due to the following reasons:

i. It is very efficient at small sample size.

ii. It is useful when the population is multivariate normal.

iii. It is a reasonable criterion for constructing a linear discriminant contribution.

iv. It is extremely simple to apply and interpret.

v. It maximizes the difference between groups relative to the standard deviation within the groups.

vi. It is optimal in its classification.

The Federal University of Technology, Minna, is divided into four schools and under each of these schools are different (various) departments. It is these departments that constitute or comprises of these schools. Out of these four schools, two schools were selected as our two populations (two - group). This is done in order to come in line with our states scope of study in chapter one of this paper-study will base on Two – group fisher's discriminant analysis function. From these two schools, fifty samples of students were randomly selected.

The data for this analysis were obtained from three compulsory subjects of examination's marks (scores) of each of fifty selected student from each schools. The students are pre-degree students of the University and the Three compulsory subjects they offer during the course of programme are physics, chemistry and Biology.

However, first population shall be referred to as School of Science and Science Education while the second population shall be referred to as school of Environmental Technology. We take each of these subjects as our variables, say X and each of these subjects shall be labeled as X_1, X_2 and X_3 variables accordingly with following order of arrangement as physics, Chemistry and Biology respectively.

3.2 COMPUTATION OF THE ESTIMATES OF PARAMETERS FOR THE SCHOOL OF SCIENCE AND SCIENCE EDUCATION.

The sums of measurements in the School of Science and Science Education are computed below:

$$\sum_{i=1}^{100} x_{1i} = 3878$$

$$\sum_{i=1}^{100} x_{2i} = 3898$$

$$\sum_{i=1}^{100} x_{3i} = 3662$$

This means measurements computed for the school

$$\bar{X}_1 = \frac{1}{50}\sum_{i=1}^{100} x_{1i} = \frac{3878}{50} = 77.56$$

$$\bar{X}_2 = \frac{1}{50}\sum_{i=1}^{100} x_{2i} = \frac{3898}{50} = 77.96$$

$$\bar{X}_3 = \frac{1}{50}\sum_{i=1}^{100} x_{3i} = \frac{3662}{50} = 73.24$$

The mean vector X_1 of the measurement for the school is given by

$$X_1 = \begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix} = \begin{pmatrix} 77.56 \\ 77.96 \\ 73.24 \end{pmatrix}$$

The sum of squares and cross product terms computed for the school is given below

$$\sum_{i=1}^{100} X_{1i}^2 = 302100$$

$$\sum_{i=1}^{100} X_{2i}^2 = 305334$$

$$\sum_{i=1}^{100} X_{3i}^2 = 269102$$

$$\sum_{i=1}^{100} X_{1i}X_{2i} = 301655$$

$$\sum_{i=1}^{100} X_{1i}X_{3i} = 284029$$

$$\sum_{i=1}^{100} X_{1i}X_{1i} = 284816$$

The sum of squares and cross product matrix for the school are calculated below and labeled V_1

$$V_1 = \begin{pmatrix} \sum_{i=1}^{100} X_{1i}^2 = n\bar{X}_1^2 & \sum_{i=1}^{100} X_{1i}^2 X_{2i} - n\bar{X}_1\bar{X}_2 & \sum_{i=1}^{100} X_{1i}^2 X_{3i} - n\bar{X}_1\bar{X}_3 \\ \sum_{i=1}^{100} X_{2i}^2 X_{1i} - n\bar{X}_2\bar{X}_1 & \sum_{i=1}^{100} X_{2i}^2 = n\bar{X}_2^2 & \sum_{i=1}^{100} X_{2i}^2 X_{3i} - n\bar{X}_2\bar{X}_3 \\ \sum_{i=1}^{100} X_{3i}^2 X_{1i} - n\bar{X}_3\bar{X}_1 & \sum_{i=1}^{100} X_{3i}^2 X_{2i} - n\bar{X}_3\bar{X}_2 & \sum_{i=1}^{100} X_{3i}^2 = n\bar{X}_3^2 \end{pmatrix}$$

$$V_1 = \begin{pmatrix} 1322.32 & -673.88 & 4.28 \\ -673.88 & 1445.95 & -693.52 \\ 4.28 & -673.52 & 897.12 \end{pmatrix}$$

3.3 COMPUTATION OF THE ESTIMATES OF PARAMETERS FOR THE SCHOOL OF ENVIRONMENTAL TECHNOLOGY

The sums of measurements in the School of Environmental Technology are computed below:

$$\sum_{i=1}^{100} x_{1i} = 2748$$

$$\sum_{i=1}^{100} x_{2i} = 2832$$

$$\sum_{i=1}^{100} x_{3i} = 2933$$

The means measurement computed for the school

$$\bar{X}_1 = \frac{1}{50}\sum_{i=1}^{100} x_{1i} = \frac{27488}{50} = 54.96$$

$$\bar{X}_2 = \frac{1}{50}\sum_{i=1}^{100} x_{2i} = \frac{2832}{50} = 56.64$$

$$\bar{X}_3 = \frac{1}{50}\sum_{i=1}^{100} x_{3i} = \frac{2933}{50} = 58.66$$

The mean vector X_2 of the measurement for the school is given by

$$X_2 = \begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix} = \begin{pmatrix} 54.96 \\ 56.64 \\ 58.66 \end{pmatrix}$$

The sum of squares and cross product terms computed for the school is given below:

$$\sum_{i=1}^{100} X_{1i}^2 = 157098$$

$$\sum_{i=1}^{100} X_{2i}^2 = 165984$$

$$\sum_{i=1}^{100} X_{3i}^{2} = 17483$$

$$\sum_{i=1}^{100} X_{1i} X_{2i} = 157166$$

$$\sum_{i=1}^{100} X_{1i} X_{3i} = 161991$$

$$\sum_{i=1}^{100} X_{1i} X_{1i} = 167169$$

The sum of squares and cross product matrix for the school are calculated below and labeled V_2

$$V_2 = \begin{pmatrix} \sum_{i=1}^{100} X_{1i}^{2} = n\overline{X}_1^2 & \sum_{i=1}^{100} X_{1i}^{2} X_{2i} - n\overline{X}_1\overline{X}_2 & \sum_{i=1}^{100} X_{1i}^{2} X_{3i} - n\overline{X}_1\overline{X}_3 \\ \sum_{i=1}^{100} X_{2i}^{2} X_{1i} - n\overline{X}_2\overline{X}_1 & \sum_{i=1}^{100} X_{2i}^{2} = n\overline{X}_2^2 & \sum_{i=1}^{100} X_{2i}^{2} X_{3i} - n\overline{X}_2\overline{X}_3 \\ \sum_{i=1}^{100} X_{3i}^{2} X_{1i} - n\overline{X}_3\overline{X}_1 & \sum_{i=1}^{100} X_{3i}^{2} X_{2i} - n\overline{X}_3\overline{X}_2 & \sum_{i=1}^{100} X_{3i}^{2} = n\overline{X}_3^2 \end{pmatrix}$$

$$V_2 = \begin{pmatrix} 6067.92 & 1519.28 & 793.32 \\ 1519.28 & 5579.52 & 1043.88 \\ 793.32 & 1043.88 & 3432.22 \end{pmatrix}$$

3.4 COMPUTATION OF POOLED SAMPLE COVARIANCE MATRIX

The pooled covariance matrix is calculated as follows

$$S = \frac{V_1 + V_2}{n_1 + n_2 - 2}$$

$$S = 1/98 \begin{pmatrix} 7390.24 & 845.40 & 797.60 \\ 845.40 & 7025.44 & 44370.36 \\ 797.60 & 370.36 & 4330.34 \end{pmatrix}$$

$$S = \begin{pmatrix} 75.41061224 & 8.626530612 & 813877551 \\ 8.626530612 & 71.68816327 & 3.779183673 \\ 8.13877551 & 3.779183673 & 44.18714296 \end{pmatrix}$$

The inverse of matrix is given below.

$$S^{-1} = \frac{adjoint\ of\ S}{determinant\ of\ S}$$

$$S^{-1} = \begin{pmatrix} 0.013692952 & -0.00151632 & -0.002391947 \\ -0.001521632 & 0.01481574 & -0.0009326636 \\ -0.002391947 & -0.0009326636 & 0.0231511352 \end{pmatrix}$$

3.5 COMPUTATION OF FISHER'S LINEAR DISCRIMINANT FUNCTION FOR TWO POPULATIONS

Fisher's linear discriminant function is given by the following formula

$$Y = (\bar{X}_1 - \bar{X}_2)' S^{-1} X$$

Where X is the three X vector

i.e. $\begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix}$

$$Y = [22.6\ 21.32\ 14.58] \begin{pmatrix} 0.0136929526 & -0.001521632 & -000239194 \\ -0.001521632 & 0.01481574 & -000932636 \\ -0.002391947 & -0.000932636 & 0.0231513527 \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix}$$

$$= [0.242144899\ \ 0.254364405\ \ 0.263604868] \begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix}$$

$$\hat{Y}_1 = 0.242144899X_{1i} + 0.254364405X_{2i} + 0.263604868X_{3i}$$

Which is the Fisher's Linear discriminant function.

The formula for the mean of the function in the School of Science and Science Education is given by

$$\overline{Y}_1 = \left(\overline{X}_1 - \overline{X}_2\right)^1 S^{-1} \overline{X}_1$$

$$\overline{Y}_1 = [22.6 \quad 21.32 \quad 14.58] \begin{bmatrix} 0.136929526 & -0.001521632 & -0.002391947 \\ 0.0015216332 & 0.01418115743 & -000932636 \\ -0.002391947 & -0.000932636 & 0.0231513527 \end{bmatrix} \begin{bmatrix} 77.56 \\ 77.96 \\ 73.24 \end{bmatrix}$$

$$= [0.242144899 \quad 0.254364405 \quad 0.263604868] \begin{bmatrix} 77.56 \\ 77.96 \\ 73.24 \end{bmatrix}$$

$$\overline{Y}_1 = 57.91742791$$

The formula for the mean of the School of Environmental Technology is given by

$$\overline{Y}_2 = \left(\overline{X}_1 - \overline{X}_2\right)^1 S^{-1} \overline{X}_2$$

$$\overline{Y}_2 = [22.6 \quad 21.32 \quad 14.58] \begin{bmatrix} 0.136929526 & -0.001521632 & -0.002391947 \\ 0.0015216332 & 0.01418115743 & -000932636 \\ -0.002391947 & -0.000932636 & 0.0231513527 \end{bmatrix} \begin{bmatrix} 54.96 \\ 56.64 \\ 58.66 \end{bmatrix}$$

$$= [0.242144899 \quad 0.254364405 \quad 0.263604868] \begin{bmatrix} 54.96 \\ 56.64 \\ 58.66 \end{bmatrix}$$

$$\overline{Y}_2 = 43.17854511$$

The formula for the cut-off point is given by

$$1/2\left(\overline{Y}_1 + \overline{Y}_2\right) = 1/2\left(\overline{X}_1 - \overline{X}_2\right)^1 S^{-1}\left(\overline{X}_1 - \overline{X}_2\right)$$

$= ½(57.91742791 + 43.17854511)$

$= 50.5480$

The Maharianobi's (squared distance, D^2 between the two populations (schools) is given by the following formula

$$\bar{Y}_1 - \bar{Y}_2 = (\bar{X}_1 - \bar{X}_2)' S^{-1} (\bar{X}_1 - \bar{X}_2)$$

$$= 57.91742791 - 43.17854511$$

$$= 14.7389$$

Assignment Rule:

Assign an individual (student) to the first population (School of Science and Science Education) if $(\bar{X}_1 - \bar{X}_2)' S^{-1} (\bar{X}_1 - \bar{X}_2) > ½(\bar{Y}_1 + \bar{Y}_2)$ and assign an individual to the second population (School of Environmental Technology) if $(\bar{X}_1 - \bar{X}_2)' S^{-1} (\bar{X}_1 - \bar{X}_2) \leq ½(\bar{Y}_1 + \bar{Y}_2)$

i.e. assign to π_1 if $Y_i > 50.5480$ and assign to π_2 if $Y_i \leq 50.5480$

3.6 COMPUTATION OF DISCRIMINANT SCORES FOR THE TWO SCHOOLS

We shall use the discriminant function

$\hat{Y}_1 = 0.242144899X_{1i} + 0.254364405X_{2i} + 0.263604868X_{3i}$ to obtain discriminant scores in the two schools. See Appendix at back for generated scores for the two schools.

3.7 PROBABILITY OF MISCLASSIFICATION FOR THE TWO SCHOOLS USING FISHER'S LINEAR DISCRIMINANT FUNCTION

SCHOOL	SCIENCE EDUCATION	ENVIRONMENTAL	TOTAL
SCIENCE EDUCATION	50	0	50
ENVIRONMENTAL	6	44	50
TOTAL	56	44	100

The probability of misclassification into the School of Science Education

$P_1 = M_1/n_1 = 0/50 = 0$

The probability of misclassification into the School of Environment

$P_2 = M_2/n_2 = 6/50 = 0.12$

The total probability of misclassification is

$\overline{P} = M_1M_2/n_1n_2 = 0+6/100 = 0.06$

Where

P_1 is the probability of misclassification of the School of Science Education

P_2 is the probability of misclassification of the School of Environment

P is the probability of misclassification of the two schools

M_1 is the number of students who are misclassified into the School of Science Education

M_2 is the number of students who are misclassified into the School of Environment.

n_1 and n_2 are the number of students in the two schools respectively.

3.8 TESTS BETWEEN GROUP DIFFERENCES

We shall make use of F-distribution to find out the observed differences between the groups

Therefore

$$F = \frac{n_1 n_2 (n_1 + n_2 - k - 1)}{n_1 + n_2 (n_1 + n_2) k} + D^2 \text{ where } n_1 = n_2 = 50,\ k = 3,$$

$D^2 = Y_1 - Y_2 = 14.7389$

$$F = \frac{50\,x\,50(50 + 50 - 3 - 1)\,x\,14.7389}{50 + 50(50 + 50 - 2)\,x\,3}$$

$= 120.3176 \Rightarrow 120.32$

$F_{k,\ n_1+n_2-k-1,\ 0.10} = F_{3,\ 98,\ 0.10} = 2.68$

Hypothesis: H_0: $\mu_1 = \mu_2$ Vs H_1: $\mu_1 \neq \mu_2$

Decision: We reject the null hypothesis, H_0

Conclusion: We then conclude that there are significant differences between the two schools.

CHAPTER FOUR

SOFTWARE DEVELOPMENT AND IMPLEMENTATION.

4.1 INTRODUCTION

This Chapter will concentrate on software development implementation. The chapter will discuss the software to be used, its programming language and the programming detail.

4.2 CHOICE OF SOFTWARE PACKAGES AND PROGRAMMING LANGUAGE.

In selecting a software packages certain criteria needed to be considered. These criteria are;

1. The effectiveness and efficiency of the packages with regard to the functions of the developed programmes.
2. The facilities for different types of file processing.
3. The security of the records in the files.
4. The facilities for maintaining the files e.g. adding new records, easy retrieval, modifying of records, etc.
5. The flexibility of the packages
6. User's friendliness of the packages.

Based on the above criteria, the application software package that will be able to solve this problem will be – Statistical Package for the Social Science SPSS. And the programming language that will be developed for solving this problem will be Qbasic programming language.

4.3 FEATURES OF SOFTWARE / PROGRAMMING LANGUAGE

Statistical packages are well-defined integrated set of programs designed for performing various statistical analyses bore on user's specification using statistical techniques. There are lots of application packages that requires just data definition and specification of the desired operation.

The processes involves are usually;

i. Establishing the set of ideas or concepts
ii. Collecting data and analyzing them and establishing facts using the various statistical techniques contained in the packages.

Examples of Statistical packages include:

1. Statistical packages for the social (SPSS)
2. Statistical Analysis system (SAS) Professional Statistics and graphic packages (MYSTAT).

Basic programming language, which uses complier as translator. It can handle large data in which relative computation is involved. It is easy to learn and user friendly. And it is one of high level language that takes care of mathematical formula. The following are procedures for entering the program environment through Disk Operating system.

i. Booting system.
ii. C:\> ↵
iii. C:\>cd Qbasic ↵
iv. C:\>Qbasic > Qb. ↵

4.4. PROGRAM DEVELOPMENT

Program is defined as an instruction set describing the logical steps the computer will follow to solve a particular problem. The act of program writing is programming and it involved the following steps.

i. Problem definition: Before a program can be written to perform a particular task the nature and complexity of the problem or task must be known. That is, sub task procedures and routines that must be well defined and formulates using mathematical statements and operators. It is when problem are well defined that the problems can be processed.

ii. Problem Analysis: It involves analysis the various procedures or routines defined to find a method of solution. That involves manipulating the records in a file, establishing a relationship between the various data elements and the description of the medium storage.

iii. Algorithm is a step by step method or rules for solving a problem in a finite sequence of steps. It involves describing in literal terms the steps to be taken to solve a given problem.

iv. Flowchart can simply be defined as diagrammatic representation of algorithms. It is a pictorial representation of a complex procedure with considerable charity. Two major types of flowchart and they are system and program flowchart, system flowchart give a general pictorial representation of the system overview. It shows the various relationships among the input data, the processing and the desired output. While program flowchart is a pictorial representation of the logical steps the computer takes to solve a problem.

v. Coding program: is the actual writing of the instructions set the computer follows to solve the problem according to a specified rate.

vi. Testing program: this step is more to test data to see the effectiveness or adequacy of the program through run and debug the program.

vii. Program documentation: Entails giving a concise description of programs in form of user manual and operating instructions. it gives details of what the program can do and what it cannot do as well as simplifying the task of a maintenance programmer and making provision for future amendments.

CHAPTER FIVE

SUMMARY, CONCLUSION AND RECOMMENDATION

5.1 SUMMARY

From the analysis, Fisher's linear discriminant function was computed thus:

$$\hat{Y}i = 0.022144899X_{1i} + 0.2544364405X_{2i} + 0.263604868X_{3i}$$

Given the function, we shall see that there are no much differences in the coefficient of the three variables X1, X2, X3. This implies that nearly all three subjects contributed to the function. Since 0.263604868 is greater than the other two coefficients, we can say that Biology contributed more significantly to the function than Physics and Chemistry.

In test of between group-differences, we used F-statistics to find out the observed group-differences. The result is that the two population (School of Science Education and Environmental Technology) are entirely different, and that is why we want to further to discriminate between the two populations.

The value of the mean in the first population was computed as $\overline{Y}_1 = 57.9174$ while the value of the mean in the second population was computed as $\overline{Y}_2 = 43.1785$. We noticed that the mean in the School of Science and Science Education is greater than the mean of the school of Environmental Technology. This means that, mean increases with greater viability of scores in the populations.

The value of the Mahalanobics, D^2 is large i.e $D^2 = 14.7389$. This implies that we have few classifications. The probability of misclassification in the School of Science was 0 and probability of misclassification in the School of Environmental was 0.12. The total probability of misclassification was obtained as 0.06. The total probability of misclassification very small, this is because the value of Mahanalobics distance is very large.

As for the performance of the discriminant function, since the probability of misclassification is very small, then the discriminant function performed very well. Another factor that influence performance of the function is the fact that data used construct the function was re-used to classify the observations.

The Fisher's linear discriminant scores for the students from the two schools were computed. The discriminant scores for the students in the school of science were greater than those of School of Environmental.

5.2 CONCLUSION

We have been able to construct a classification rule using Fisher's linear discriminant function. Using the function, individual students from the two schools were reclassified into either of the two schools or the basis of the cut off point and their discriminant scores.

The apparent error rate for the two schools was obtained from the Fisher's linear discriminant function.

Since all the students in the school of Science were correctly classified and only six were wrongly classified into the school of Environment, then it is appropriate to use the Fisher's discriminant function for constructing classification rate.

5.3 RECOMMENDATION

Base on the data collected and the analysis out a student should be advised to be in the school of Science Education if his discriminant score is above the cut-off point, otherwise he should be advised to be in the School of Environmental Technology if his discriminant score falls below the cut off point.

Moreover, we can use the discriminant function for the subsequent years. This can be achieved by reviewing the function sequentially every year until the estimated coefficients are constant

APPENDIX I

THE STUDENTS SCORES FOR THE SCHOOL OF SCIENCE AND SCIENCE EDUCATION

	PHYSICS	CHEMISTRY	BIOLOGY
1.	89	78	72
2.	87	82	70
3.	82	78	78
4.	82	76	79
5.	82	78	76
6.	75	82	78
7.	83	74	78
8.	76	82	76
9.	87	80	68
10.	80	83	70
11.	80	76	76
12.	78	80	74
13.	80	76	74
14.	83	71	76
15.	87	67	76
16.	75	85	70
17.	76	76	78
18.	75	76	78
19.	75	80	74
20.	76	85	68
21.	68	85	76
22.	85	65	79
23.	82	72	74
24.	76	74	78
25.	80	69	79
26.	76	80	72
27.	80	74	74
28	75	80	72
29	75	85	67
30	71	82	74
31	73	82	72
32	69	80	78
33	82	74	72
34	83	80	63
35	75	83	68
36	78	78	70
37	76	85	65
38	69	85	72
39	73	87	65
40	75	85	65
41	78	67	79
42	80	72	72
43	76	74	74
44	82	74	68
45	76	74	74
46	76	69	79
47	73	76	74
48	69	82	72
49	71	78	72
50	68	81	74

APPENDIX II

THE STUDENTS SCORES FOR THE SCHOOL OF ENVIRONMENTAL TECHNOLOGY

	PHYSICS	CHEMISTRY	BIOLOGY
1	83	65	76
2	85	80	58
3	62	80	78
4	73	72	70
5	75	71	65
6	69	78	61
7	66	71	67
8	68	61	68
9	59	60	68
10	69	40	70
11	55	65	58
12	46	72	59
13	52	51	72
14	55	58	61
15	52	67	54
16	60	56	56
17	57	65	50
18	55	63	52
19	62	40	67
20	45	51	72
21	53	61	54
22	53	54	61
23	60	54	52
24	48	60	58
25	46	65	54
26	20	71	74
27	50	60	54
28	50	54	59
29	52	56	54
30	48	43	70
31	59	43	58
32	60	52	47
33	50	60	47
34	60	49	47
35	52	43	61
36	48	54	54
37	52	52	50
38	53	58	43
39	55	45	54
40	43	54	56
41	41	54	58
42	50	45	58
43	45	54	54
44	59	45	48
45	60	43	48
46	52	43	56
47	48	43	59
48	38	52	59
49	45	52	52
50	50	47	52

APPENDIX III

CALCULATED FISHERS DISRIMINANT SCORES FOR THE SCHOOL OF SCIENCE AND SCIENCE EDUCATION

1. 0.242144899(89) + 0.254364405(78) + 0.26304868(72) = 60.3708
2. 0.242144899(87) + 0254364405(82) + 0.263604868(70) = 6.3768
3. 0242144899(82) + 0.254364405(78) + 0.263604868(78) =60.3575
4. 0.242144899(82) + 0.254364405(76)+ 0.26304868(79) = 60.0124
5. 0.242144899(82) + 0.254364405(78)+ 0.26304868(76)= 59.7303
6. 0.242144899(75)+ 0.254364405(82)+ 0.26304868(78)= 59.5800
7. 0.242144899(83)+ 0.254364405(74)+ 0.26304868(78)= 59.4822
8. 0.242144899(76)+ 0.254364405(83)+ 0.26304868(76)=59.5492
9. 0.242144899(87)+ 0.254364405(80)+ 0.26304868(68)= 59.3409
10.0.242144899(80)+ 0.254364405(83)+ 0.26304868(70)=58.9362
11.0.242144899(80)+ 0.254364405(76)+ 0.26304868(76)=58.7373
12.0.242144899(78)+ 0.254364405(80)+ 0.26304868(74)=58.7432
13.0.242144899(80)+ 0.254364405(76)+ 0.26304868(74)=58.2101
14.0.242144899(83)+ 0.254364405(71)+ 0.26304868(76)=58.1919
15.0.242144899(87)+ 0.254364405(67)+ 0.26304868(76)=58.1430
16.0.242144899(75)+ 0.254364405(85)+ 0.26304868(70)=58.2342
17.0.242144899(76)+ 0.254364405(76)+ 0.26304868(78)= 58.2959
18.0.242144899(75)+ 0.254364405(76)+ 0.26304868(78)=58.0537
19.0.242144899(75)+ 0.254364405(80)+ 0.26304868(74)=58.0168
20.0.242144899(76)+ 0.254364405(85)+ 0.26304868(68)=57.9491
21.0.242144899(68)+ 0.254364405(85)+ 0.26304868(76)=58.1208
22.0.242144899(85)+ 0.254364405(65)+ 0.26304868(79)=57.9408
23.0.242144899(82)+ 0.254364405(72)+ 0.26304868(74)=57.6769
24.0.242144899(76)+ 0.254364405(74)+0.26304868(78)=57.7872
25.0.242144899(80)+ 0.254364405(69)+0.26304868(79)=57.7475
26.0.242144899(76)+ 0.254364405(80)+0.26304868(72)=57.7317

27.0.242144899(80)+ 0.254364405(74)+0.26304868(74)=57.7013

28.0.242144899(75)+ 0.254364405(80)+0.26304868(72)=57.4896

29.0.242144899(75)+ 0.254364405(85)+0.26304868(67)=57.4434

30.0.242144899(71)+ 0.254364405(82)+0.26304868(74)=57.5569

31.0.242144899(73)+ 0.254364405(82)+0.26304868(72)=57.5140

32.0.242144899(69)+ 0.254364405(80)+0.26304868(78)=57.6183

33.0.242144899(82)+ 0.254364405(74)+0.26304868(70)57.1312

34.0.242144899(83)+ 0.254364405(80)+0.26304868(63)=57.0543

35.0.242144899(75)+ 0.254364405(83)+0.26304868(68)=57.1982

36.0.242144899(78)+ 0.254364405(78)+0.26304868(70)=57.1801

37.0.242144899(76)+ 0.254364405(85)+0.26304868(65)=57.1583

38.0.242144899(69)+ 0.254364405(85)+0.26304868(72)=57.3085

39.0.242144899(73)+ 0.254364405(87)+0.26304868(65)=56.9406

40.0.242144899(75)+ 0.254364405(85)+0.26304868(65)=56.9162

41.0.242144899(78)+ 0.254364405(67)+0.26304868(79)=56.7545

42.0.242144899(80)+ 0.254364405(72)+0.26304868(72)=56.6654

43.0.242144899(76)+ 0.254364405(74)+0.26304868(74)=56.7327

44.0.242144899(82)+ 0.254364405(74)+0.26304868(68)=56.6040

45.0.242144899(76)+ 0.254364405(74)+0.26304868(74)=56.7327

46.0.242144899(76)+ 0.254364405(69)+0.26304868(79)=56.7789

47.0.242144899(73)+ 0.254364405(76)+0.26304868(74)=56.5150

48.0.242144899(69)+ 0.254364405(82)+0.26304868(72)=56.5454

49.0.242144899(71)+ 0.254364405(78)+0.26304868(74)=56.5395

50.0.242144899(68)+ 0.254364405(81)+0.26304868(74)=56.5761

APPENDIX IV

CALCULATED FISHERS DISRIMINANT SCORES FOR THE SCHOOL OF ENVIRONMENTAL TECHNOLOGY

1. 0.242144899(83) + 0.254364405(65) + 0.26304868(76) = 56.6657
2. 0.242144899(85) + 0254364405(80) + 0.263604868(58) = 22.2206
3. 0242144899(62) + 0.254364405(80) + 0.263604868(78) =55.9233
4. 0.242144899(73) + 0.254364405(72)+ 0.26304868(70) = 54.4432
5. 0.242144899(75) + 0.254364405(71)+ 0.26304868(65)= 53.3551
6. 0.242144899(69)+ 0.254364405(78)+ 0.26304868(61)= 52.6283
7. 0.242144899(66)+ 0.254364405(71)+ 0.26304868(67)= 51.7030
8. 0.242144899(68)+ 0.254364405(61)+ 0.26304868(68)=49.9072
9. 0.242144899(59)+ 0.254364405(60)+ 0.26304868(68)= 47.4735

10.0.242144899(69)+ 0.254364405(40)+ 0.26304868(70)=45.3349

11.0.242144899(55)+ 0.254364405(65)+ 0.26304868(58)=45.1407

12.0.242144899(46)+ 0.254364405(72)+ 0.26304868(59)=45.0060

13.0.242144899(52)+ 0.254364405(51)+ 0.26304868(72)=44.5437

14.0.242144899(55)+ 0.254364405(58)+ 0.26304868(61)=44.1510

15.0.242144899(52)+ 0.254364405(67)+ 0.26304868(54)=43.8686

16.0.242144899(60)+ 0.254364405(56)+ 0.26304868(56)=43.5350

17.0.242144899(57)+ 0.254364405(65)+ 0.26304868(50)= 43.5162

18.0.242144899(55)+ 0.254364405(63)+ 0.26304868(52)=43.0504

19.0.242144899(62)+ 0.254364405(40)+ 0.26304868(67)=42.8499

20.0.242144899(45)+ 0.254364405(51)+ 0.26304868(72)=42.8487

21.0.242144899(53)+ 0.254364405(61)+ 0.26304868(54)=.42.5846

22.0.242144899(53)+ 0.254364405(54)+ 0.26304868(61)=42.6493

23.0.242144899(60)+ 0.254364405(54)+ 0.26304868(52)=41.9718

24.0.242144899(48)+ 0.254364405(60)+0.26304868(58)=42.1739

25.0.242144899(46)+ 0.254364405(65)+0.26304868(54)=41.9070

26.0.242144899(20)+ 0.254364405(71)+0.26304868(74)=42.4095

27.0.242144899(50)+ 0.254364405(60)+0.26304868(54)=41.0638

28.0.242144899(50)+ 0.254364405(54)+0.26304868(59)=41.3956

29.0.242144899(52)+ 0.254364405(56)+0.26304868(54)=41.0706

30.0.242144899(48)+ 0.254364405(43)+0.26304868(70)=41.0130

31.0.242144899(59)+ 0.254364405(43)+0.26304868(58)=40.5133

32.0.242144899(60)+ 0.254364405(52)+0.26304868(47)=40.1451

33.0.242144899(50)+ 0.254364405(60)+0.26304868(47)=39.7585

34.0.242144899(60)+ 0.254364405(49)+0.26304868(47)=39.3820

35.0.242144899(52)+ 0.254364405(43)+0.26304868(61)=39.6091

36.0.242144899(48)+ 0.254364405(54)+0.26304868(54)=39.5933

37.0.242144899(52)+ 0.254364405(52)+0.26304868(50)=38.9987

38.0.242144899(53)+ 0.254364405(58)+0.26304868(43)=38.9218

39.0.242144899(55)+ 0.254364405(45)+0.26304868(54)=38.9990

40.0.242144899(43)+ 0.254364405(54)+0.26304868(56)=38.9098

41.0.242144899(41)+ 0.254364405(54)+0.26304868(58)=38.9527

42.0.242144899(50)+ 0.254364405(45)+0.26304868(58)=38.8427

43.0.242144899(45)+ 0.254364405(54)+0.26304868(54)=38.8669

44.0.242144899(59)+ 0.254364405(45)+0.26304868(48)=38.3860

45.0.242144899(60)+ 0.254364405(43)+0.26304868(48)=38.1194

46.0.242144899(52)+ 0.254364405(43)+0.26304868(56)=38.2911

47.0.242144899(48)+ 0.254364405(43)+0.26304868(59)=38.1133

48.0.242144899(38)+ 0.254364405(52)+0.26304868(59)=37.9811

49.0.242144899(45)+ 0.254364405(72)+0.26304868(52)=37.8309

50.0.242144899(50)+ 0.254364405(47)+0.26304868(52)=37.7698

REFERENCES

1. AYO C.K. (1994); Computer Literacy (operations and appreciations) By Oniboye, Press, Ibadan

2. CHERNOFF, H. (1973); Some measure for Discriminating between Normal multivariate Distribution with Unequal Covariance Matrices, in Multivariate. By Macmillan Publishing Co. Inc. New York.

3. COOLEY AND LOHNES (1963); Multivariate Procedures for the Behavioral Sciences, By John Wiley and Sons Inc. New York.

4. LACHENBRUCH P.A (1975); Discriminant Analysis, New York, By Halner Press London.

5. MATHER P.M. (1976); Computational Methods of Multivariate Analysis in Physical Geography. By John Wiley & Sons, Nothingham London.

6. FISHER R.A. (1936); The Use of Multiple Measurements in Taxonomic Problems, Annals of Eugenics. Vol. 9 page 179 –188.

7. GILIBERT, E.S. (1969); The Effects of Unequal Variance – Covariance Matrices on Fisher's Linear Discriminant Functions, Journal of Biometrics Vol. 25 Pg. 505 –516.

8. JEAN-PIERE N. AND AKIHIHO M. (1988); Recent Developments in Clustering and Data Analysis: Proceedings of the Japanese-French Scientific Seminar.

9. MARKS, S. AND DUNN O.J (1974); Discriminant Functions When Covariance Matrices are Unequal, J.A.S.A. Vol. 69.

10. MOORE D.H. (1973); Evaluation of Five Discriminant Procedures for Binary Variables" J.A.S.A. Vol. 68.

11. WILLIAMS H.K. AND JUDITH M. T. (1978); International Encyclopedia of Statistics Vol. 1 Judith M. T. (1978). By Collier Macmillan Publishing Co. Inc. New York.

12. FEDERAL UNIVERSITY OF TECHNOLOGY (1985); Information Booklet by information and Publications Division. Registry Department, FUT, Minna.

```
 REM PROGRAM TO CONSTRUCT DISCRIMINANT ANALYSIS
 CLS
   DIM SA(50, 3), SB(50, 3)
ATA
9,78,72,87,82,70,82,78,78,82,76,79,82,78,76,75,82,78,83,74,78,76,83,76,87,80,68,80,83
70,80,76,76,78,80,74,80,76,74,83,71,76,87,67,76,75,85,70,76,76,78,75,76,78,75,80,74,7
,85,68,68,85,76,85,65,79,82,72,74,76,74,78,80,69,79,76,80,72,80,74,74,75,80
ATA
2,75,85,67,71,82,74,73,82,72,69,80,78,82,74,70,83,80,63,75,83,68,78,78,70,76,85,65,69
85,72,73,87,65,75,85,65,78,67,79,80,72,72,76,74,74,82,74,68,76,74,74,76,69,79,73,76,7
,69,82,72,71,78,74,68,81,74
ATA
3,65,76,85,80,58,62,80,78,73,72,70,75,71,65,69,78,61,66,71,67,68,61,68,59,60,68,69,40
70,55,65,58,46,72,59,52,51,72,55,58,61,52,67,54,60,56,56,57,65,50,55,63,52,62,40,67,4
,51,72,53,61,54,53,54,61,60,54,52,48,60,58,46,65,54,20,71,74,50,60,54,50,54,59,52,56,
4,48,43,70,59,43,58,60,52,47,50,60,47,60,49,47,52,43,61,48,54,54,52,52,50,53,58,43,55
45,54,43,54,56,41,54,58,50,45,58
ATA 45,54,54,59,45,48,60,43,48,52,43,56,48,43,59,38,52,59,45,52,52,50,47,52
RINT " DATA FOR GROUP 1"
RINT
OR K = 1 TO 3
  SUMSA(K) = 0
EXT K

OR J = 1 TO 50

   FOR K = 1 TO 3
      READ SA(J, K)
       SUMSA(K) = SUMSA(K) + SA(J, K)
   NEXT K
EXT J


OR K = 1 TO 3
UMSB(K) = 0
EXT

RINT " DATA FOR GROUP 2"
OR J = 1 TO 50
   FOR K = 1 TO 3
       READ SB(J, K)
       SUMSB(K) = SUMSB(K) + SB(J, K)
   NEXT K
EXT J
LS
RINT
RINT TAB(20); "SCORES FOR GROUP 1"
RINT
 A = 20
FOR J = 1 TO 50
    PRINT J;
    FOR K = 1 TO 3
        PRINT TAB(A); SA(J, K);
         A = A + 10
    NEXT K
    A = 20
    PRINT
NEXT J
PRINT
```

```
PRINT "TOTAL";
 A = 20
FOR K = 1 TO 3
     PRINT TAB(A); SUMSA(K);
     A = A + 10
NEXT K
PRINT : PRINT
ANS$ = INPUT$(1)
PRINT TAB(20); "SCORES FOR GROUP 2"
PRINT
 A = 20
FOR J = 1 TO 50
      PRINT J;
     FOR K = 1 TO 3
      PRINT TAB(A); SB(J, K);
      A = A + 10
     NEXT K
     PRINT
     A = 20
NEXT J
PRINT
PRINT "TOTAL";
A = 20
FOR K = 1 TO 3
     PRINT TAB(A); SUMSB(K);
     A = A + 10
NEXT K
ANS$ = INPUT$(1)
REM COMPUTATION OF THE MEANS OF THE POPULATIONS
REM MEAN OF GROUP 1
PRINT : PRINT
PRINT TAB(20); "THE MEAN OF INDIVIDUAL VARIABLES IN GROUP 1"
PRINT
A = 20
n = 50
PRINT "MEAN FOR GROUP 1";
FOR K = 1 TO 3
     MEANSA(K) = SUMSA(K) / n
     PRINT TAB(A); MEANSA(K);
     A = A + 10
NEXT K
PRINT : PRINT
PRINT TAB(20); "THE MEAN OF INDIVIDUAL VARIABLES IN GROUP 2"
PRINT
A = 20

PRINT "MEAN FOR GROUP 2";
FOR K = 1 TO 3
     MEANSB(K) = SUMSB(K) / n
     PRINT TAB(A); MEANSB(K);
     A = A + 10
NEXT K
PRINT
ANS$ = INPUT$(1)
REM COMPUTATION OF THE SUM  OF SQUARES
PRINT
PRINT TAB(20); "THE SUM OF SQUARES OF VARIABLES IN GROUP 1"
PRINT
FOR K = 1 TO 3
```

```
     SUMXASQ(K) = 0
NEXT K
FOR J = 1 TO 50
     FOR K = 1 TO 3
          XSQA(K) = SA(J, K) ^ 2
          SUMXASQ(K) = SUMXASQ(K) + XSQA(K)
     NEXT K
NEXT J
PRINT
 A = 20
FOR J = 1 TO 50
      PRINT J;
     FOR K = 1 TO 3
      PRINT TAB(A); XSQA(K);
      A = A + 10
     NEXT K
     PRINT
     A = 20
NEXT J
PRINT
PRINT "SUM OF SQUARES";
A = 20
FOR K = 1 TO 3
     PRINT TAB(A); SUMXASQ(K);
     A = A + 10
NEXT K
PRINT
PRINT TAB(20); "THE SUM OF SQUARES OF VARIABLES IN GROUP 2"
PRINT
FOR K = 1 TO 3
     SUMXBSQ(K) = 0
NEXT K

FOR J = 1 TO 50
     FOR K = 1 TO 3
          XSQB(K) = SB(J, K) ^ 2
          SUMXBSQ(K) = SUMXBSQ(K) + XSQB(K)
     NEXT K
NEXT J
PRINT
 A = 20
FOR J = 1 TO 50
      PRINT J;
     FOR K = 1 TO 3
      PRINT TAB(A); XSQB(K);
      A = A + 10
     NEXT K
     PRINT
     A = 20
NEXT J
PRINT
PRINT "SUM OF SQUARES";
A = 20
FOR K = 1 TO 3
     PRINT TAB(A); SUMXBSQ(K);
     A = A + 10
NEXT K
PRINT
ANS$ = INPUT$(1)
```

```
PRINT
REM COMPUTATION OF THE SUM OF CROSS PRODUCT  FOR THE FIRST POPULATION

SUMAB = 0
FOR J = 1 TO 50
    X(J) = SA(J, 1) * SA(J, 2)
    SUMAB = SUMAB + X(J)
NEXT J

SUMAC = 0
FOR J = 1 TO 50
    X(J) = SA(J, 1) * SA(J, 3)
     SUMAC = SUMAC + X(J)
NEXT J

SUMBC = 0
FOR J = 1 TO 50
    X(J) = (SA(J, 2)) * (SA(J, 3))
     SUMBC = SUMBC + X(J)
NEXT J
PRINT
PRINT "CROSS PRODUCT AB";
PRINT TAB(20); SUMAB
PRINT : PRINT
PRINT "CROSS PRODUCT AC";
PRINT TAB(20); SUMAC
PRINT : PRINT
PRINT "CROSS PRODUCT BC";
PRINT TAB(20); SUMBC

REM COMPUTATION OF THE SUM OF SQUARES AND CROSS PRODUCT MATRICES FOR SCHOOL A
REM CPAA = SUM OF SQUARES AND CROSS PRODUCT OF VARIABLE A SQUARED
CPAA = SUMXASQ(1) - (n * (MEANSA(1)) ^ 2)
CPAB = SUMAB - (n * MEANSA(1) * MEANSA(2))
CPAC = SUMAC - (n * MEANSA(1) * MEANSA(3))
CPBB = SUMXASQ(2) - (n * (MEANSA(2)) ^ 2)
CPBC = SUMBC - (n * MEANSA(2) * MEANSA(3))
CPCC = SUMXASQ(3) - (n * (MEANSA(3)) ^ 2)
PRINT
ANS$ = INPUT$(1)
PRINT "THE SUM OF SQUARES AND CROSS PRODUCT MATRIX OF THE FIRST POPULATION"
PRINT CPAA, CPAB, CPAC
PRINT CPAB, CPBB, CPBC
PRINT CPAC, CPBC, CPCC


REM COMPUTATION OF THE SUM OF CROSS PRODUCT  FOR THE SECOND POPULATION

SUMAB2 = 0
FOR J = 1 TO 50
    X(J) = (SB(J, 1)) * (SB(J, 2))
     SUMAB2 = SUMAB2 + X(J)
NEXT J

SUMAC2 = 0
FOR J = 1 TO 50

    X(J) = SB(J, 1) * SB(J, 3)
     SUMAC2 = SUMAC2 + X(J)
```

```
NEXT J

SUMBC2 = 0
FOR J = 1 TO 50
    X(J) = (SB(J, 2)) * (SB(J, 3))
      SUMBC2 = SUMBC2 + X(J)
NEXT J
PRINT
PRINT "CROSS PRODUCT FOR THE SECOND POPULATION"
PRINT "CROSS PRODUCT AB";
PRINT TAB(20); SUMAB2
PRINT : PRINT
PRINT "CROSS PRODUCT AC";
PRINT TAB(20); SUMAC2
PRINT : PRINT
PRINT "CROSS PRODUCT BC";
PRINT TAB(20); SUMBC2

REM COMPUTATION OF THE SUM OF SQUARES AND CROSS PRODUCT MATRICES FOR SCHOOL B
REM CPAA = SUM OF SQUARES AND CROSS PRODUCT OF VARIABLE A SQUARED
CPAA2 = SUMXBSQ(1) - (n * (MEANSB(1)) ^ 2)
CPAB2 = SUMAB2 - (n * MEANSB(1) * MEANSB(2))
CPAC2 = SUMAC2 - (n * MEANSB(1) * MEANSB(3))
CPBB2 = SUMXBSQ(2) - (n * (MEANSB(2)) ^ 2)
CPBC2 = SUMBC2 - (n * MEANSB(2) * MEANSB(3))
CPCC2 = SUMXBSQ(3) - (n * (MEANSB(3)) ^ 2)

PRINT "THE SUM OF SQUARES AND CROSS PRODUCT MATRIX OF THE SECOND POPULATION"
PRINT CPAA2, CPAB2, CPAC2
PRINT CPAB2, CPBB2, CPBC2
PRINT CPAC2, CPBC2, CPCC2
PRINT
REM ANS$ = INPUT$(1)

REM COMPUTATION OF POOLED SAMPLE COVARIANCE
PCAA = (CPAA + CPAA2) / (n + n - 2)
PCAB = (CPAB + CPAB2) / (n + n - 2)
PCAC = (CPAC + CPAC2) / (n + n - 2)
PCBB = (CPBB + CPBB2) / (n + n - 2)
PCBC = (CPBC + CPBC2) / (n + n - 2)
PCCC = (CPCC + CPCC2) / (n + n - 2)
PRINT " THE POOLED SAMPLE COVARIANCE MATRIX S"
PRINT PCAA, PCAB, PCAC
PRINT PCAB, PCBB, PCBC
PRINT PCAC, PCBC, PCCC
PRINT
REM THE INVERSE OF THE MATRIX S i.e S^-1 IS ADJOINT OF S/DETERMINANT OF S
REM THE ADJOINT OF MATRIX S
CFA = (PCBB * PCCC) - (PCBC * PCBC)
CFB = -((PCAB * PCCC) - (PCAC * PCBC))
CFC = (PCAB * PCBC) - (PCAC * PCBB)
CFD = -((PCAB * PCCC) - (PCBC * PCAC))
CFE = (PCAA * PCCC) - (PCAC * PCAC)
CFF = -((PCAA * PCBC) - (PCAC * PCAB))
CFG = (PCAB * PCBC) - (PCBB * PCAC)
CFH = -((PCAA * PCBC) - (PCAB * PCAC))
CFI = (PCAA * PCBB) - (PCAB * PCAB)
PRINT
PRINT "THE COFACTORS OF MATRIX S"
```

```
PRINT CFA, CFB, CFC
PRINT CFD, CFE, CFF
PRINT CFG, CFH, CFI
PRINT
PRINT "THE ADJOINT OF MATRIX S"
REM TRANSPOSE THE ABOVE COFACTOR
PRINT CFA, CFD, CFG
PRINT CFB, CFE, CFH
PRINT CFC, CFF, CFI
PRINT
REM COMPUTATION OF THE DETERMINANT OF MATRIX S
DETERMS = PCAA * ((PCBB * PCCC) - (PCBC * PCBC)) - PCAB * ((PCAB * PCCC) - (PCAC *
PCBC)) + PCAC * ((PCAB * PCBC) - (PCAC * PCBB))

PRINT "THE DETERMINANT OF MATRIX S = "; DETERMS
REM COMPUTATION OF INVERSE OF MATRIX S
INVA = CFA / DETERMS
INVB = CFB / DETERMS
INVC = CFC / DETERMS
INVD = CFD / DETERMS
INVE = CFE / DETERMS
INVF = CFF / DETERMS
INVG = CFG / DETERMS
INVH = CFH / DETERMS
INVI = CFI / DETERMS

PRINT
PRINT "THE INVERSE OF MATRIX S = S^-1 = "
PRINT INVA,
PRINT TAB(30); INVD;
PRINT TAB(60); INVG
PRINT
PRINT INVB,
PRINT TAB(30); INVE; ;
PRINT TAB(60); INVH
PRINT
PRINT INVC,
PRINT TAB(30); INVF; ;
PRINT TAB(60); INVI
PRINT
REM COMPUTATION OF FISHER'S LINEAR DISCRIMINANT FUNCTION
REM FOR THE TWO POPULATIONS
REM FISHER'S LINEAR DISCRIMINANT FUNCTION IS GIVEN BY THE FOLLOWING FORMULAR
REM Y= (MEANSA-MEANSB)*(S^-1)*X, WHERE X IS A 3X VECTOR
REM          X1
REM      X = X2
REM          X3
REM COMPUTATION OF MEANDIFF (MEANSA-MEANSB)
A = 20
FOR K = 1 TO 3
     MEANDIFF(K) = MEANSA(K) - MEANSB(K)
     PRINT TAB(A); MEANDIFF(K);
     A = A + 20
NEXT K
 MEAN1 = MEANDIFF(1)
 MEAN2 = MEANDIFF(2)
 MEAN3 = MEANDIFF(3)
 E1 = (MEAN1 * INVA) + (MEAN2 * INVB) + (MEAN3 * INVC)
 E2 = (MEAN1 * INVD) + (MEAN2 * INVE) + (MEAN3 * INF)
```

```
 E3 = (MEAN1 * INVG) + (MEAN2 * INVH) * (MEAN3 * INVI)
 PRINT
 PRINT "E1="; E1,
 PRINT "E2="; E2,
 PRINT "E3="; E3

 Y1 = E1 * MEANSA(1) + E2 * MEANSA(2) + E3 * MEANSA(3)
 PRINT Y1
Y2 = (E1 * MEANSB(1)) + (E2 * MEANSB(2)) + (E3 * MEANSB(3))
PRINT Y2
REM THE FORMULAR FOR CUT-OFF POINT
COP = (Y1 + Y2) / 2
ANS$ = INPUT$(1)
PRINT "THE CUT-OFF POINT ="; COP
 FOR J = 1 TO 50
       FOR K = 1 TO 3
YSA(J) = (E1 * SA(J, 1)) + (E2 * SA(J, 2)) + (E3 * SA(J, 3))
       NEXT K
       PRINT YSA(J)

NEXT J
PRINT

 FOR J = 1 TO 50
       FOR K = 1 TO 3
YSB(J) = (E1 * SB(J, 1)) + (E2 * SB(J, 2)) + (E3 * SB(J, 3))
       NEXT K
       PRINT YSB(J)

NEXT J
PRINT
FOR J = 1 TO 50
     IF YSA(J) > COP THEN
          SAY(J) = YSA(J)
     ELSE
          SBY(J) = YSA(J)
     END IF

NEXT J

PRINT
FOR J = 1 TO 50
     IF YSB(J) > COP THEN
          SAY(J) = YSB(J)
     ELSE
          SBY(J) = YSB(J)
     END IF

NEXT J
PRINT
PRINT "GROUP 1"
FOR J = 1 TO 50
     PRINT SAY(J)
NEXT J
PRINT
PRINT "GROUP 2"
FOR J = 1 TO 50
     PRINT SBY(J)
NEXT J
```