# COMPARATIVE PERFORMANCE OF STUDENTS FROM ADDIS ABABA AND OTHER PARTS OF THE COUNTRY

A THESIS

PRESENTED TO THE SCHOOL OF GRADUATE STUDIES

ADDIS ABABA UNIVERSITY



IN PARTIAL FULFILLMENT

OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE IN STATISTICS

BY

Mohammed Omer

Addis Ababa
June, 1989

ACKNOWLEDGEMENT

I would like to express my deep gratitude to my Advisor Professor Ayenew Ejigou who suggested the research project and spent his valuable time on advising in every stage of this study.

Many thanks are due to Mr.J.D. Marshal, Assistant Dean of the Faculty of Science, who allowed me to enter into students' records after a series of correspondence with the concerned authorities. My thanks are also due to Dr. Asmerom Kidane ,Head, Department of Statistics, for his support in this respect.

I am indebted to Ato Gebre-Egziabher Kiros of the N.U.P.I Computing Center for his assistance in data analysis. I am also indebted to Ato Genene Zewge of ECA for his cooperation in producing the photocopies of the graphs.

Finally, my thanks also go to W/Zenebech W.Tsadik, Secretary of the Department of Statistics, for typing the manuscript quickly and efficiently.

TABLE OF CONTENTS

# ABSTRACT

Some multivariate statistical methods have been employed in this study. In particular, the method of profite analysis have been used to compare the performance of students who originated from Addis Ababa with that of the remaining students who originated from other parts of Ethiopia. The data, four semester grade point averages (GPA's), collected from the record office of the Faculty of Science of Addis Ababa University have been used to test whether difference in origin resulted in any real difference in performance between the two groups of students after the first semester of the freshman program.

The study is primarily based on the students from the Department of Statistics who were enrolled at the begining of the 1981/82 academic year and graduated by the end of the 1984/85 academic year for various reasons. Based on the emporical findings some tentative conclusions have been made.

# 1. INTRODUCTION

## 1.1 Objective of the Study

Some studies (King and King, 1972; Mekonnen, 1987; Habte, 1988) have been conducted at Addis Ababa University to evaluate whether the Ethiopian School Leaving Certificate Examination (ESLCE) grade point average (GPA) is a good index for screening candidates for admission. Other studies included those on ranking of faculties/departments according to students' performance and the selection of the best performing faculty/department (Mekonnen, 1987) and on factors affecting academic performance of freshmen (Habte, 1988).

These studies considered freshmen, in most cases, at Addis Ababa University. However, no follow-up study has been attempted to study the pattern or trend of students' performance after the freshman year.

Habte (1988) has discussed factors such as dormitory facility, ESLCE grade point average (GPA) and student's region of origin in relation to the performance of freshmen in the Faculty of Science. These factors may also influence students' performance during the remaining years of the study program. In addition, there may also exist other factors that may affect students' progress. These factors may include, among others:

i) failure to get first choice

ii) lack of student interest

iii) departmental inadequacy relating to staff and facilities

iv) heavy semester work loads.

Due to the possibly differing degrees of impact of these and other factors, one may expect a non-uniform trend in the performance of students in different departments and from different regions of origin. The objective of the present study is to compare the four-year performance of the 1985 degree graduates of Addis Ababa origin with that of the remaining graduates who originated from other parts of the Country. The study focussed on students from the Department of Statistics of the Faculty of Science, for various reasons.

## 1.2 Problems and Limitations

The primary purpose of this study was to carry out a profile analysis on the GPA of students from Addis Ababa and Other Regions and to also apply a multivariate regression approach to identify the factors which may possibly affect students' performance after the freshman year.

The multivariate regression approach was envisaged with the anticipation that the computing facility which was expected to be available to our department would arrive as expected. Unfortunately, the facility has not yet arrived. Besides, all relevant pieces of information expected to have been filled in by the students during their application for admission into the freshman program were not there. Hence, the regression approach had to be put aside. The profile analysis would also require computing support, but this requirement could be kept small.

Furthermore, upon subdividing the study population, it was found out that the maximum number of graduates, equal to 6, was observed in the Department of Statistics for Addis Ababa (See Table 2.2.1 below). However, the numbers corresponding to Addis Ababa for the remaining departments were observed to be even smaller. Thus, this study is primarily based on the GPA's of 19 graduates from the Department of Statistics, of which 13 were from parts of the Country other than Addis Ababa.

Moreover, it is worthwhile mentioning the time constraint encountered; because of some administrative problems data were not collected on the sheduled time.

## 2. MATERIALS AND METHODS USED IN THE STUDY

### 2.1 Methods of Data Collection and Source

The subjects of the study were the 1985 degree graduates of the Department of Statistics of the Faculty of Science of Addis Ababa University.

The data (four semester GPA's, one at the end of each academic year, together with some background information on each graduate) were collected on a form prepared for this purpose (See Appendix A) aided by the list of 25 graduates, who got the approval of the Faculty's Academic Commission and graduated at the end of the 1984/85 academic year. Unlike other batches, the graduation day was in November 1985, after the rehabilitation campaign of Gambella in what was then called Illubabor Administrative Region. The list and files of graduates' record were obtained from the Faculty's record office.

Later on, we sorted out the graduates belonging to other batches, drop-out cases and dismissed students inorder to have GPA's with identical semester work loads, the same instructions and complete set of GPA's. Finally, we came up with a total of 19 graduates in this Department. This was regarded as a generation sample.

.teristics of the Data

ie objective of this Section is to show the general racteristics of the graduates coming from each origin nd to point out the possible implication of these characteristics on the analytical results.

Background information was collected on sex, age, marital status, origin, type of study at the last high school, ESLCE GPA, and status of admission to the Faculty; this was all obtained from the graduates' application forms filled in upon their enrollement. This was done so as to get some understanding of the extent of the heterogeneity of the study groups. Based on this information, the following characterizations were made.

Table 2.2.1 presents the break-down of the number of Science graduates of 1985 from Addis Ababa and the remaining regions of the Country, classified by departments. One will observe in Table 2.2.1 that the largest numbers of graduates were from Physics and Biology, consisting of 39 and 38 graduates, respectively, and the smallest were from Geology and Statistics, each having 19 graduates. The largest number of graduates from Addis Ababa region were in Statistics and the smallest in Biology. However, the proportions belonging to the regions under each department are different in magnitude. The closest proportions belong to the Depart-

Table 2.2.1: Graduates of 1985 According to Region of Origin and Department

| Source Region | Biology | Chem. | Geol. | Math. | Phys. | Stat. | Total |
|---|---|---|---|---|---|---|---|
| Addis Ababa | 1 | 5 | 3 | 4 | 5 | 6 | 24 |
| Other Regions | 36 | 22 | 15 | 21 | 34 | 13 | 141 |
| Not Specified | 1 | 1 | 1 | - | - | - | 3 |
| Total | 38 | 28 | 19 | 25 | 39 | 19 | 168 |

ment of Statistics. Taking into consideration the profile analysis of interest (with $p = 4$) and these differing constituent proportions we elected to carry out the analysis for the graduates from the Department of Statistics. Thus, the descriptive analyses refer to this Department.

Table 2.2.2 shows the age, upon enrollement, of graduates from Statistics. Most of them were at the ages of 18 and 19, and the least numbers were at the ages of 17 and 20. Moreover, there were two graduates who did not specify their age, upon enroll ment, one from each region. There are differences at the ages 18, 19 and 20 between the two regions, but the extent of youthfulness of the two groups appears to be about the same so that age is not expected to show any differential effect in the two groups.

Table 2.2.2: Age Distribution upon Enrollement in Statistics

| Age | Addis Ababa | Other Regions | Total |
|---|---|---|---|
| 17 | 1 | 1 | 2 |
| 18 | 3 | 2 | 5 |
| 19 | 1 | 7 | 8 |
| 20 | - | 2 | 2 |
| No specification | 1 | 1 | 2 |
| T o t a l | 6 | 13 | 19 |

Table 2.2.3 presents the sex distribution of graduates. As shown in this Table, most of the graduates were male in both regions and there were only two female graduates, one from each region. Again there is a similarly oriented imbalance in the number of males in both groups.

Table 2.2.3: Sex Distribution of Graduates of Statistics

| Sex | Addis Ababa | Other Regions | Total |
|---|---|---|---|
| Male | 5 | 12 | 17 |
| Female | 1 | 1 | 2 |
| T o t a l | 6 | 13 | 19 |

Table 2.2.4 presents the distribution of ESLCE grade point average, upon enrollement. As shown in this Table, the maximum observed GPA is 3.2 (all from Addis Ababa) and

the minimum is 2.2 (a graduate outside Addis Ababa enrolled as a quota student). The highest frequency corresponds to the GPA 2.8 (three from Addis Ababa and four from Other Regions) and the least are at the GPA's 2.2 and 3.0 (all from other regions). All graduates from Addis Ababa had grades above 2.4 and those from Other Regions had below 3.2. There was no specification for a graduate outside Addis Ababa. There is also one graduate who has been enrolled as a private applicant from Addis Ababa. Moreover, graduates from Addis Ababa had a higher mean GPA than those from Other Regions (the graduate with no specification of origin was not considered for mean calculation).

Table 2.2.4: ESLCE Grade Point Average Distribution of Statistics Graduates

| ESLCE GPA | Addis Ababa | Other Regions | Total |
|---|---|---|---|
| 2.2 | - | 1 | 1 |
| 2.4 | - | 2 | 2 |
| 2.6 | 1 | 4 | 5 |
| 2.8 | 3 | 4 | 7 |
| 3.0 | - | 1 | 1 |
| 3.2 | 2 | - | 2 |
| No specification | - | 1 | 1 |
| Total | 6 | 13 | 19 |
| Mean | 2.90 | 2.63 | - |

In addition, the marital status upon enrollement was predominantly single in both regions with the exception of one married student (outside Addis Ababa) and the "no specification" for a graduate (outside Addis Ababa). Regarding the type of study at the last highschool, most of the graduates were from academic science in both regions, three graduates outside Addis Ababa were from agriculture stream and no specification was obtained for four graduates (one from Addis Ababa and three from Other Regions). Eventhough no information was obtained, it is theoretically known that a dormitory facility is provided to any student from outside Addis Ababa but not to one from Addis Ababa.

Whereas age, sex and marital status may not have any serious effect on student performance, it is expected that ESLCE GPA and dormitory facility may affect post-freshman GPA; the effect of ESLCE GPA is expected to be positive, suggesting a higher profile for Addis Ababa, but the absence of a dormitory facilty, on the other hand, may show a depressing effect. The opposite will be true for the performance of students who come from areas outside Addis Ababa. The net effect, in both cases, may be to create a situation whereby positive and negative effects have a tendency to neutralize each other.

Furthermore, failure to get first choice is expected to have no effect on the performance of students in the Department of Statistics, because every student is assigned to this Department according to his or her first choice. However, lack of student interest, departmental inadequacy relating to staff and facilities and heavy semester work loads may affect students' performance but they are expected to show no differential effect on the two groups.

## 2.3 Methods of Data Analysis

### 2.3.1 Assessment of Multivariate Normality

The assumption of multivariate normality underlies much of the standard "classical" multivariate statistical methodology. Thus in real life application of this methodology, the first step that must be taken before going into analysis is the assessment of data for any serious violation of the underlying assumption. This assumption is thus, a basis for the methods of analysis described in Sections 2.3.2, 2.3.3 and 2.3.4.

There are various methods suggested by different authors (See Gnanadesikan, 1977) which have been in use so far for this purpose. For its simplicity and ease of interpretation we used the chi-square plot method described in Johnson and Wichern (1982).

This method of judging the joint normality of a data set is based on the squared generalized distances $d_j^2$ given by

$$d_j^2 = (X_j-\bar{X})'S^{-1}(X_j-\bar{X}), \quad j=1,2,\ldots,n$$

where $X_j$ are column vectors of P measurements on the $j^{th}$ element or individual, $\bar{X}$ is a column vector of P means computed from treatment totals and S is the sample covariance matrix. The procedure to construct a chi-square plot is as follows:

i) Compute and order the squared distances in an ascending order as $d^2_{(1)} \leq d^2_{(2)} \leq \cdots \leq d^2_{(n)}$.

ii) Compute the values of $\chi^2_p \left(\frac{j-\frac{1}{2}}{n}\right)$, where $\chi_p\left(\frac{j-\frac{1}{2}}{n}\right)$ is the 100 $\left(\frac{j-\frac{1}{2}}{n}\right)$ percentile of the chi-square distribution with P degrees of freedom, and finally,

iii) plot the pairs $\left(d_{(j)}, \chi^2_p\left(\frac{j-\frac{1}{2}}{n}\right)\right)$.

The resulting plot should resemble a straight line through the origin, if multivariate normality holds. Any systematic curved pattern suggests lack of normality and any point to the right of the line indicate large distance or an outlier that needs further attention.

### 2.3.2 Assessment of Equality of Covariance Matrices

The generalized Bartlett test for the equality of k covariance matrices, given in Morrison (1976), is used for this purpose.

The hypothesis

$$H_0: \Sigma_1 = \Sigma_2 = \cdots = \Sigma_k$$

of the equality of the covariance matrices of k p-dimensional multinormal populations can be tested against the alternative of general positive definite matrices by a modified generalized likelihood-ratio statistic. Let $S_i$ be the unbiased estimate of $\Sigma_i$ based on $f_i$ degrees of freedom, where $f_i = n_i - 1$ for the usual case of a random sample of $n_i$ observation vectors from the $i^{th}$ population. When $H_o$ is true

$$S = \frac{1}{\sum_{i=1}^{k} f_i} \sum_{i=1}^{k} f_i S_i$$

is the pooled estimate of the common covariance matrix. The test statistic is

$$M = \left(\sum_{i=1}^{k} f_i\right) \ln|S| - \sum_{i=1}^{k} f_i \ln|S_i|$$

and if the scale factor

$$D^{-1} = 1 - \frac{2p^2 + 3p - 1}{6(p+1)(k-1)} \left( \sum_{i=1}^{k} \frac{1}{f_i} - \frac{1}{\sum_{i=1}^{k} f_i} \right)$$

is introduced the quantity $MD^{-1}$ is approximately distributed as a chi-square variate with degrees of freedom $\frac{1}{2}(k-1)\,p(p+1)$ as the $f_i$ becomes large. A decision rule with level of significance $\alpha$ is: Reject $H_o$ if $\chi^2 > \chi^2_\alpha(\frac{1}{2}(k-1)p(p+1))$, where $\chi^2_\alpha(\frac{1}{2}(k-1)p(p+1))$ is the $100(1-\alpha)$ percentile point of the $\chi^2$-distribution with $\frac{1}{2}(k-1)p(p+1)$ degrees of freedom.

### 2.3.1 The Method of Testing the Equality of Two Mean Vectors

This is a multivariate analog of the test on difference between two means. Suppose one draws independent random samples of sizes $n_1$ and $n_2$ from p-variate normal populations $N(\mu_1, \Sigma)$ and $N(\mu_2, \Sigma)$, respectively.

From data in the given samples, one may compute the following sample mean vectors:

$$\bar{X}_1 = \begin{bmatrix} \bar{X}_{11} \\ \bar{X}_{12} \\ \vdots \\ \bar{X}_{1p} \end{bmatrix} \quad \text{and} \quad \bar{X}_2 = \begin{bmatrix} \bar{X}_{12} \\ \bar{X}_{22} \\ \vdots \\ \bar{X}_{2p} \end{bmatrix}$$

and sample covariance matrices $S_1$ and $S_2$ and the pooled sample covariance matrix $S$ where

$S_1$ = covariance matrix obtained from sample 1,

$S_2$ = covariance matrix obtained from sample 2,

and $S$ = pooled covariance matrix $= \dfrac{(n_1-1)S_1+(n_2-1)S_2}{n_1+n_2-2}$

To test the p-variate hypothesis $H_o$: $\mu_1 = \mu_2$ against the alternative hypothesis $H_1$: $\mu_1 \neq \mu_2$ , an appropriate test statistic $T^2$ (proposed by Hotelling, 1931) is given by

$$T^2 = \frac{n_1 n_2}{n_1+n_2} (\bar{X}_1-\bar{X}_2)'S^{-1}(\bar{X}_1-\bar{X}_2) \qquad (1)$$

When the hypothesis, $H_o$, is true, the statistic

$$F = \frac{n_1 + n_2 - p-1}{(n_1 + n_2 - 2)\, p} T^2 \qquad (2)$$

is distributed as the F-distribution with degrees of freedom p and $n_1 + n_2 - P-1$. A decision rule with level of significance $\alpha$ is: Reject $H_o$ if $F > F_\alpha(p, n_1+n_2-p-1)$ , where $F_{\alpha}(p, n_1+n_2-p-1)$ is the $100(1-\alpha)$ percentile point of the F distribution with degrees of freedom p and $n_1+n_2-p-1$.

The test in the next section provide a more detailed analysis.

### 2.3.4 Profile Analysis

The detailed theory of profile analysis is found in Morrison (1976) and Johnson and Wichern (1982).

Profile analysis, in general, pertains to situations where a battery of p treatments (tests, questions, and so forth) are administered to two or more groups of subjects. Since our primary interest is to consider two groups we restrict the discussion to the method to be used under this situation.

The following two assumptions underly the model for profile analysis:

i) The responses are described by a p-dimensional multivariate normal random variable X.

ii) Both populations have a common, though unknown, covariance matrix $\Sigma$.

In addition, all responses must be expressed in similar units inorder to carry out a complete profile analysis, otherwise it will be meaningless to test $H_{o3}$ (given below).

To employ this method of analysis we might ordinarily pose the question of whether the population mean vectors are the same. In profile analysis, the question of equality of mean vectors is divided into several specific possibilities. Let

$$\mu_1' = (\mu_{11}, \mu_{12}, \ldots, \mu_{1p}) \text{ and } \mu_2' = (\mu_{21}, \mu_{22}, \ldots, \mu_{2p})$$

be the mean vectors from populations 1 and 2, respectively. Broken-line graphs constructed from these means are called population profiles for the two groups or populations. The profiles can also be estimated using the sample means so that then they are refered to as sample profiles (see Fig. 3.2 below for illustration).

In terms of the population profiles, we can formulate the question of equality in a stagewise fashion.

a) Are the profiles parallel?

Equivalently: Is $H_{o1}$: $\mu_{1i} - \mu_{1i-1} = \mu_{2i} - \mu_{2i-1}$, $i=2,3,\ldots,p$ acceptable?

b) Assuming the profiles are parallel, are the profiles coincident or are they at the same level?

Equivalently: Is $H_{o2}$: $\mu_{1i} = \mu_{2i}$, $i=1,2,\ldots,p$ acceptable?

c) Assuming the profiles are parallel, are the population mean responses the same?

Equivalently: Is $H_{03}$: $\mu_{11} = \mu_{12} = \cdots = \mu_{1p} = \mu_{21} = \mu_{22} = \cdots = \mu_{2p}$ acceptable?

i) <u>Parallelism of Profiles</u>

This refers to the hypothesis in stage (a). It can be written as $H_{o1}$: $C\mu_1 = C\mu_2$

where C is the contrast matrix

$$C_{(P-1)XP} = \begin{bmatrix} -1 & 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & -1 & 1 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & -1 & 1 \end{bmatrix}$$

For independent samples of sizes $n_1$ and $n_2$ from the two populations, the null hypothesis can be tested by constructing the transformed observations

$$CX_{1j}, \quad j = 1,2,\dots,n_1$$

and

$$CX_{2j}, \quad j = 1,2,\dots,n_2$$

These have sample mean vectors $C\bar{\bar{X}}_1$ and $C\bar{\bar{X}}_2$, respectively, and estimators of covariance matrices, $\frac{1}{n_1} CSC'$ and $\frac{1}{n_2} CSC'$, respectively, where S is the unbiased estimator of $\Sigma$, the common covariance matrix.

Since the two sets of transformed observations have (P-1) dimensional normal distributions with means $C\mu_1$ and $C\mu_2$, respectively, and common covariance matrix $C\Sigma C'$, the statistic for testing $H_{01}$ is the two sample $T^2$ computed from the (P-1) differences of the successive mean responses:

$$T^2 = \frac{n_1 n_2}{n_1+n_2} (\bar{X}_1-\bar{X}_2)' C' (CSC')^{-1} C (\bar{X}_1-\bar{X}_2) \qquad (3)$$

We refer

$$F = \frac{n_1+n_2-P}{(n_1+n_2-2)(P-1)} T^2 \qquad (4)$$

to a table of the F-distribution with degrees of freedom $(P-1)$ and $(n_1+n_2-P)$ and reject $H_{o1}$ at the $\alpha$ level if the observed F exceeds the ciritical value $F_\alpha(P-1,\ n_1+n_2-P)$, where $F_\alpha(P-1,\ n_1+n_2-P)$ is the $100(1-\alpha)$ percentile point of the F-distribution with degrees of freedom $P-1$ and $n_1+n_2-P$.

ii) Equality of Levels or Coincident Profiles

This refers to the hypothesis in stage (b). If the hypothesis of parallel profiles, $H_{o1}$, is tenable, we may test $H_{o2}$, the hypothesis of coincident profiles. Under $H_{o1}$, the profiles will be coincident only if the total heights $\mu_{11}+\mu_{12}+\ldots+\mu_{1P}= 1'\mu_1$ and $\mu_{21}+\mu_{22}+\ldots+\mu_{2P}=1'\mu_2$ are equal. Therefore, $H_{o2}$ can be written in matrix notation as:

$$H_{o2}:\ 1'\mu_1 = 1'\mu_2$$

where $1' = (1,\ldots,1)$ is the P-component vector with unity in each position. This hypothesis is tested by computing the usual two-sample t-statistic from the sums of the observations on all responses in each sampling unit, given by:

$$t = \frac{1'(\bar{X}_1-\bar{X}_2)}{\sqrt{(1'S1)\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} \qquad (5)$$

If $H_{o2}$ is true, the statistic has the t-distribution with $n_1+n_2-2$ degrees of freedom or its square has the F-distribution with 1 and $n_1+n_2-2$ degrees of freedom. The usual rules for acceptance or rejection of the null hypothesis apply.

iii) Equality of Mean Responses

This refers to the hypothesis in stage (C). The hypothesis that all response means are equal, $H_{o3}$, can be tested as follows. As in the preceding test, we must require that the parallelism hypothesis, $H_{o1}$, is acceptable. If it is not, the test must be carried out separately for the two groups. Under the assumption of parallel mean profiles, the hypothesis is:

$$H_{o3}: C(\mu_1+\mu_2) = 0$$

For its test, we compute the grand mean vector

$$\bar{X} = \frac{n_1}{n_1+n_2}\bar{X}_1 + \frac{n_2}{n_1+n_2}\bar{X}_2 \qquad (6)$$

and from it the single-sample $T^2$ statistic:

$$T^2 = (n_1+n_2)\bar{X}'C'(CSC')^{-1}C\bar{X} \qquad (7)$$

where S is the unbiased estimator of $\Sigma$, the common covariance matrix. When $H_{o3}$ is true, the quantity

$$F = \frac{n_1+n_2-p}{(n_1+n_2-2)(p-1)}T^2 \qquad (8)$$

has the F-distribution with degrees of freedom $p-1$ and $n_1+n_2-p$, and we reject $H_{o3}$ using the same rule as under $H_{o1}$.

# 3. RESULTS

## 3.1 Multivariate Normality Plots

The method discussed in Section 2.3.1 above was applied on the data sets for Addis Ababa and Other Regions. The ordered squared distances and the corresponding chi-square percentiles for Addis Ababa and for the Other Regions, given in Appendix D, were obtained. The plot of the pairs $(d^2_{(j)}, \chi^2_4(\frac{j-\frac{1}{2}}{6}))$, for Addis Ababa, is shown in Fig. 3.1a and of the pairs $(d^2_{(j)}, \chi^2_4(\frac{j-\frac{1}{2}}{13}))$, for Other

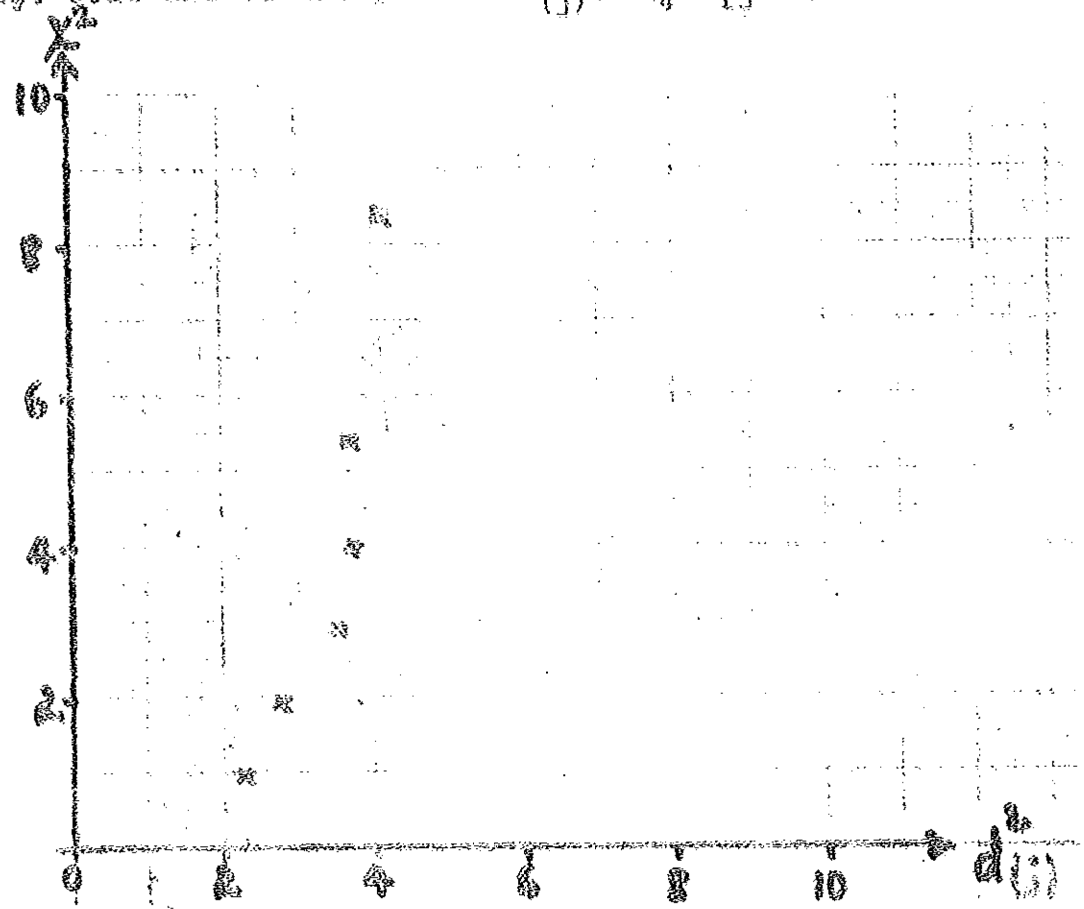


Fig. 3.1a: A chi-square plot of the ordered squared distances versus the chi-square percentile points for Addis Ababa

Regions, is shown in Fig. 3.1b. The plot for Addis Ababa appears to show some deviation from linearity, through the origin, and this may well be a reflection of the small sample size. On the other hand, the plot for the Other Regions suggests that there is no serious violation of the multivariate normality assumption.

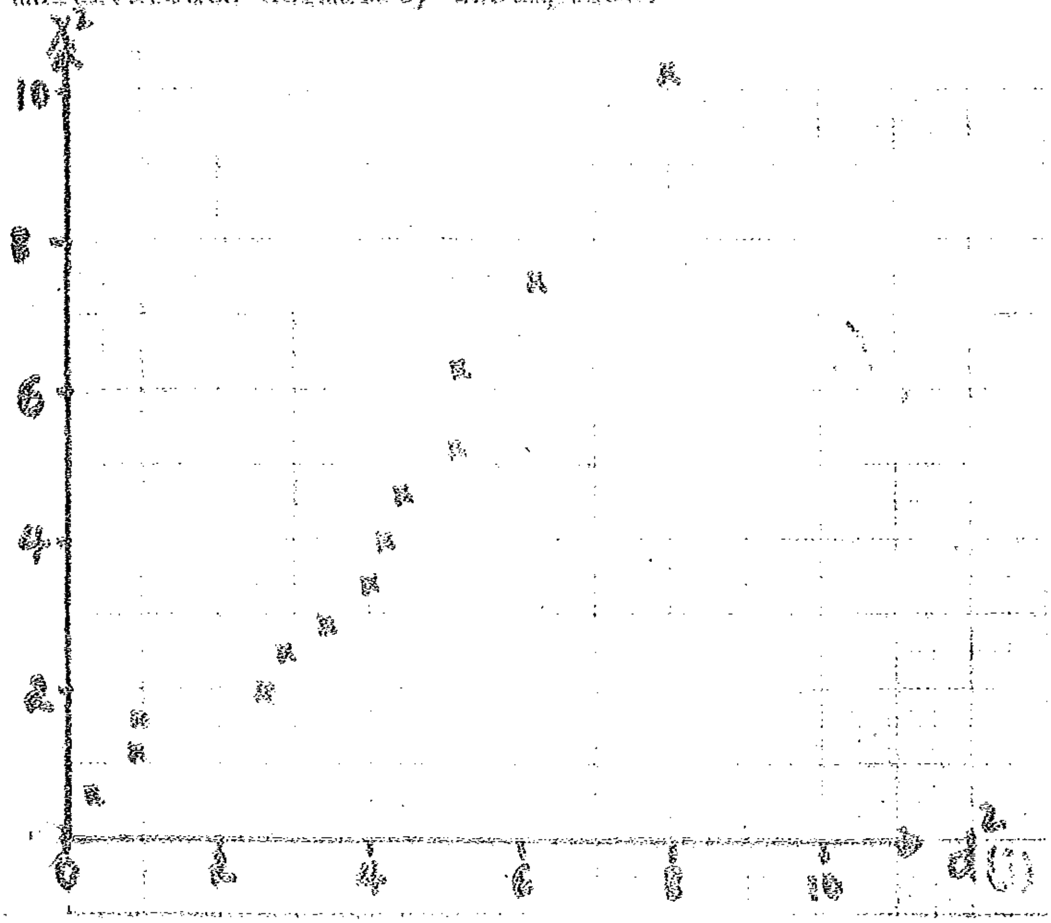


Fig. 3.1b: A chi-square plot of the ordered squared distances versus the chi-square percentile points for Other Region

Taking into account the similarity of the structures of the sample covariance matrices and the profiles for the two groups, a plot was also made by pooling the two samples. The ordered squared distances and the corresponding chi-square percentiles for p=4 and n=19, given in Appendix D, were obtained.

A plot of the pairs $(d^2_{(j)}, \chi^2_4(\frac{j-\frac{1}{2}}{19}))$ is shown in Fig. 3.1c. The plot shows no serious departure from a straight line through the origin, again suggesting that the two samples arise from multivariate normal population.

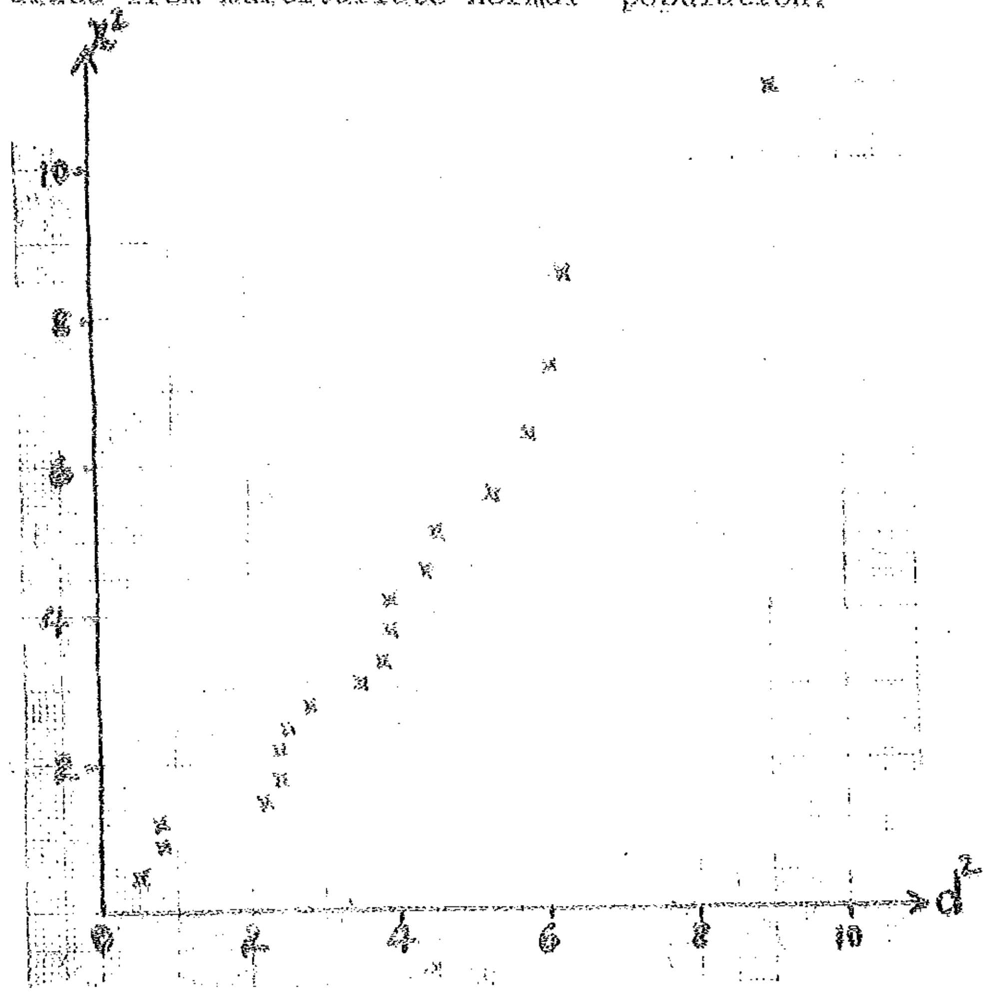


Fig. 3.1c: A chi-square plot of the ordered squared distances versus the chi-square percentile points for the combined sample

## 3.2 Test for Equality of Covariance Matrices

The samples of $n_1 = 6$ (from Addis Ababa) and $n_2 = 13$ (from Other Regions) gave the sample covariance matrices $S_1$ (for Addis Ababa) and $S_2$ (for Other Regions), respectively, and the pooled sample covariance matrix S (given in Appendix C). These matrices have the following determinants:

$$|S_1| = 2.9271 \times 10^{-5},$$

$$|S_2| = 3.4381 \times 10^{-5},$$

and $$|S| = 9.5289 \times 10^{-5}, \text{ respectively.}$$

Using the formulae for M and $D^{-1}$ given in Section 2.3.2 we get

$$M = 18.1345$$

$$D^{-1} = 0.6782$$

Thus, $MD^{-1} = \chi^2 = 12.30$, with 10 degrees of freedom. This indicates that the data do not show evidence $(p > 0.25)$ for any difference between the covariance matrices for the two groups.

## 3.3 Test for the Equality of Mean Vectors of the two Groups

The method of testing the equality of two mean vectors discussed in Section 2.3.3 will now be applied on the data. The samples gave the mean vectors,

$$\bar{X}_1 = \begin{bmatrix} 2.54 \\ 2.77 \\ 2.49 \\ 2.82 \end{bmatrix} \quad \text{and} \quad \bar{X}_2 = \begin{bmatrix} 2.70 \\ 2.78 \\ 2.50 \\ 2.76 \end{bmatrix}$$

and hence,

$$(\bar{X}_1 - \bar{X}_2) = \begin{bmatrix} -0.16 \\ -0.01 \\ -0.01 \\ 0.06 \end{bmatrix}$$

The inverse of the pooled sample covariance matrix S (given in Appendix C) becomes

$$S^{-1} = \begin{bmatrix} 17.8803 & -16.4731 & 13.0225 & -11.3079 \\ -16.4731 & 23.9642 & -19.7785 & 13.5980 \\ 13.0225 & -19.7785 & 32.2503 & -24.5099 \\ -11.3079 & 13.5980 & -24.5099 & 23.9180 \end{bmatrix}$$

Therefore,

$$(\bar{X}_1 - \bar{X}_2)' S^{-1} (\bar{X}_1 - \bar{X}_2) = 0.7647$$

Thus, the use of equations (1) and (2) of Section 2.3.3 give $T^2 = 3.1393$ and $F = 0.65$. This value of F gives no sufficient evidence for any difference ($p>0.25$), for 4 and 14 degrees of freedom, between the two population mean vectors.

4 Profile Analysis Results

3.4.1 Graduates of Statistics

The results obtained in Sections 3.1 and 3.2 of this Chapter allow us to carry out the profile analysis of interest on the two groups. The sample mean vectors, given in Section 3.3, are plotted as sample profiles in Fig. 3.2. It appears that the

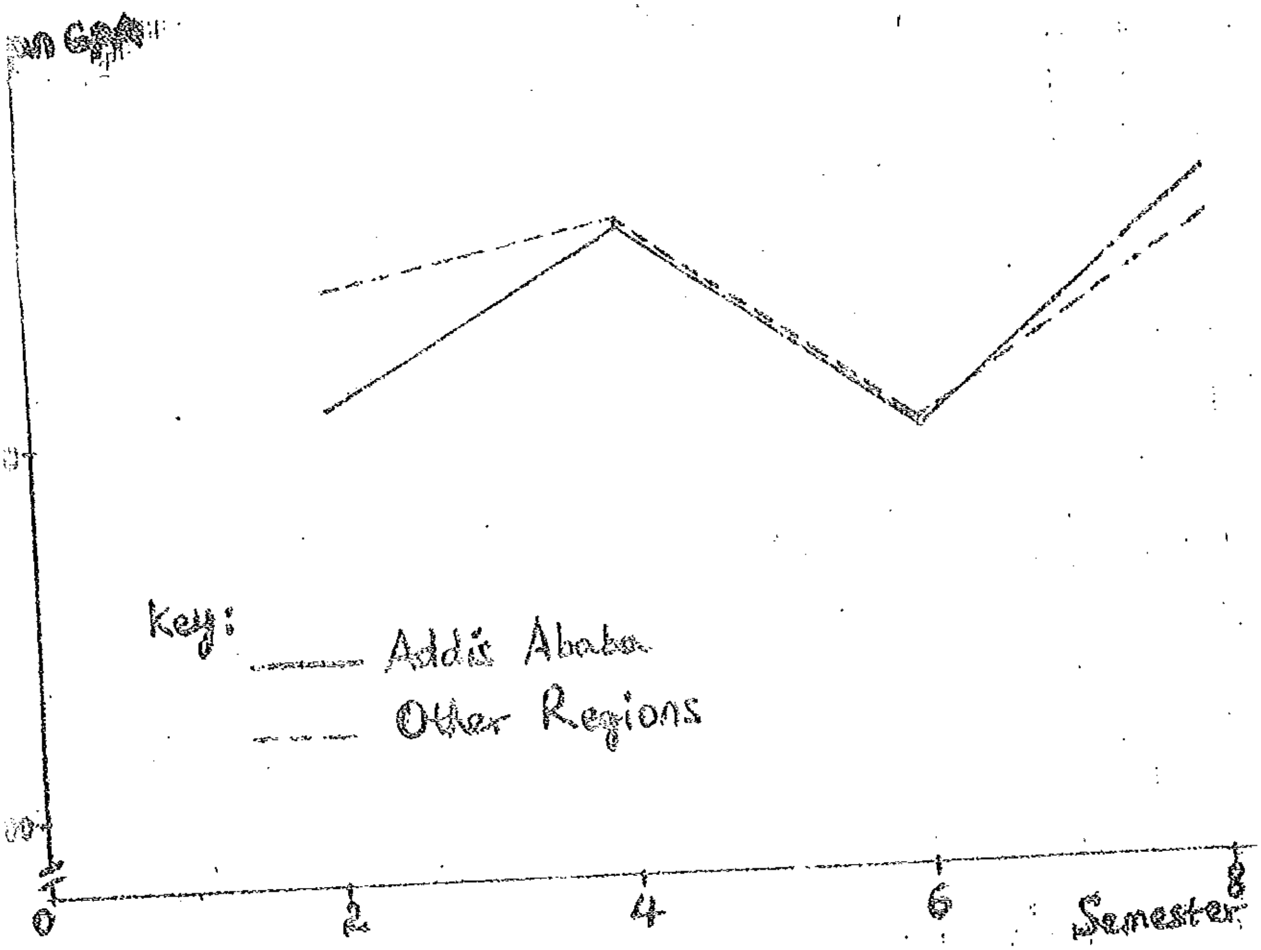


Fig. 3.2: Sample Profiles for the Graduates from the Department of Statistics.

graduates outside Addis Ababa had larger GPA's (on the average) as compared to that of the graduates from Addis Ababa at the end of first year, second semester and the converse is true at the end of fourth year, second semester. But at the two intermediate semesters the differences are not magnificient. (In Fig. 3.2 semesters 2, 4, 6 and 8 represent the second semesters in each year.)

Now let us consider the three tests discussed in Section 2.3.4 step by step. In testing the hypotheses we need $\bar{\bar{X}}_1$, $\bar{\bar{X}}_2$ and $(\bar{\bar{X}}_1 - \bar{\bar{X}}_2)$ given in Section 3.3, and S, the pooled sample covariance matrix, given in Appendix C. Moreover, the contrast matrix to be used in our case is the $3 \times 4$ matrix:

$$C = \begin{bmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \end{bmatrix}$$

i) <u>Test for Parallelism</u>

To test $H_{o1}: C\mu_1 = C\mu_2$, we compute

$$CSC' = \begin{bmatrix} 0.0968 & -0.0015 & -0.0497 \\ -0.0015 & 0.1647 & 0.0094 \\ -0.0497 & 0.0094 & 0.0685 \end{bmatrix}$$

which gives

$$(CSC')^{-1} = \begin{bmatrix} 16.5373 & -0.5384 & 12.0725 \\ -0.5384 & 6.1470 & -1.2342 \\ 12.0725 & -1.2342 & 23.5506 \end{bmatrix}$$

and

$$C'(CSC')^{-1}C = \begin{bmatrix} 16.5373 & -17.0757 & 12.6109 & -12.9725 \\ -17.0757 & 23.7611 & -19.9921 & 13.3067 \\ 12.6109 & -19.9921 & 32.1660 & -24.7848 \\ -12.0725 & 13.3067 & -24.7348 & 23.5506 \end{bmatrix}$$

Using the vector $(\bar{\bar{X}}_1 - \bar{\bar{X}}_2)$, we get

$$(\bar{\bar{X}}_1 - \bar{\bar{X}}_2)'C'(CSC')^{-1}C(\bar{\bar{X}}_1 - \bar{\bar{X}}_2) = 0.7410$$

Thus, using equation (3) of Section 2.3.4, $T^2 = 3.0420$. Again, using equation (4) of Section 2.3.4, $F = 0.89$. Since this value of F is too small as compared to any F-distribution value for 3 and 15 degrees of freedom at any common significance level, we conclude the hypothesis of parallel profiles for the two groups is tenable. This is not a surprising result given the plot in Fig. 3.2. This appears to suggest that the trend in performance (on the average) of all graduates is similar so that the origin of the graduates does not seem to play a role.

ii) Test for Coincident Profiles

Assuming the profiles are parallel, we test for coincident profiles, $H_{o2}$: $1'\mu_1 = 1'\mu_2$.

Using $1' = (1\ 1\ 1\ 1)$, $(\bar{X}_1 - \bar{X}_2)$ and S, we get

$$1'S1 = 2.3399$$

and $$1'(X_1 - X_2) = -0.12$$

Thus, using equation (5) of Section 2.3.4, we get $T^2 = 0.0253$ which has the F-distribution with 1 and 17 degrees of freedom. Again this value falls short of the theoretical F values for the specified degrees of freedom at any common significance level. Thus, we accept the hypothesis of coincident profiles for the two groups. Therefore, we conclude that the two groups show one and the same trend in performance on the average.

iii) Test for Equality of Response Means

Assuming the parallelism hypothesis, the test of the hypothesis of equal response means, $H_{o3}$: $C(\mu_1 + \mu_2) = 0$, will now be carried out.

Applying equation (6) of Section 2.3.4 on $\bar{X}_1$ and $\bar{X}_2$, given in Section 3.3, we get

$$\bar{Y} = \begin{bmatrix} 2.65 \\ 2.78 \\ 2.50 \\ 2.78 \end{bmatrix}$$

and using the matrix $C'(CSC')^{-1}C$ given above, we get

$$\bar{X}'C'(CSC')^{-1}C\bar{X} = 3.7194.$$

Then, using equations (7) and (8) of Section 2.3.4, $T^2 = 70.6686$ and $F=20.78$, respectively. This value of F is a highly significant value ($p<0.005$) for 3 and 15 degrees of freedom. Thus, we reject $H_{03}$ and conclude that all the means are not identical; that is, there is some trend in average performance for both groups. These differences could be due to the nature of courses offered and number of credit hours taken by the graduates at the respective semesters (they took 16, 17, 20 and 19 credit hours at the second semesters of each year), among others.

### 3.4.2 Profiles for Other Graduates of the Faculty

Attempts were also made to have some crude impression of the situations in the remaining departments of the Faculty by constructing sample profiles from the collected GPA's. The sample profiles for the Departments of Chemistry, Geology, Mathematics and Physics are depicted below. Taking into account the treatment differences, the GPA's of graduates from Chemistry minoring in Physics, from Geology minoring in Chemistry, from Physics minoring in Geology and GPA's of graduates who did not specify their origin were ignored because of the absence of Addis Ababa graduates in some of them or presence of only one graduate in the others. Moreover, the GPA's of graduates from Biology were not considered due to the presence of only one graduate from Addis Ababa. Thus, only the GPA's of the graduates presented in Table 3.1 are considered for this purpose. In addition to the four semester GPA's (one at the second semester of each year) we also considered the first year, first semester GPA in order to have some additional information.

Table 3.1: Graduates from Addis Ababa and Other Regions According to Department and Minor Field of Study

| Department | Minor Field of Study | Addis Ababa | Other Regions |
|---|---|---|---|
| Chemistry | Mathematics | 4 | 11 |
| Geology | Physics | 3 | 13 |
| Mathematics | Physics | 4 | 21 |
| Physics | Mathematics | 5 | 25 |

Fig. 3.3 depicts the sample profiles of the groups from the Department of Chemistry. From this Figure, one can observe that some profile difference appears to exist between the two groups. The graduates from Addis Ababa seem to have been performing less than those from Other Regions except in third year.

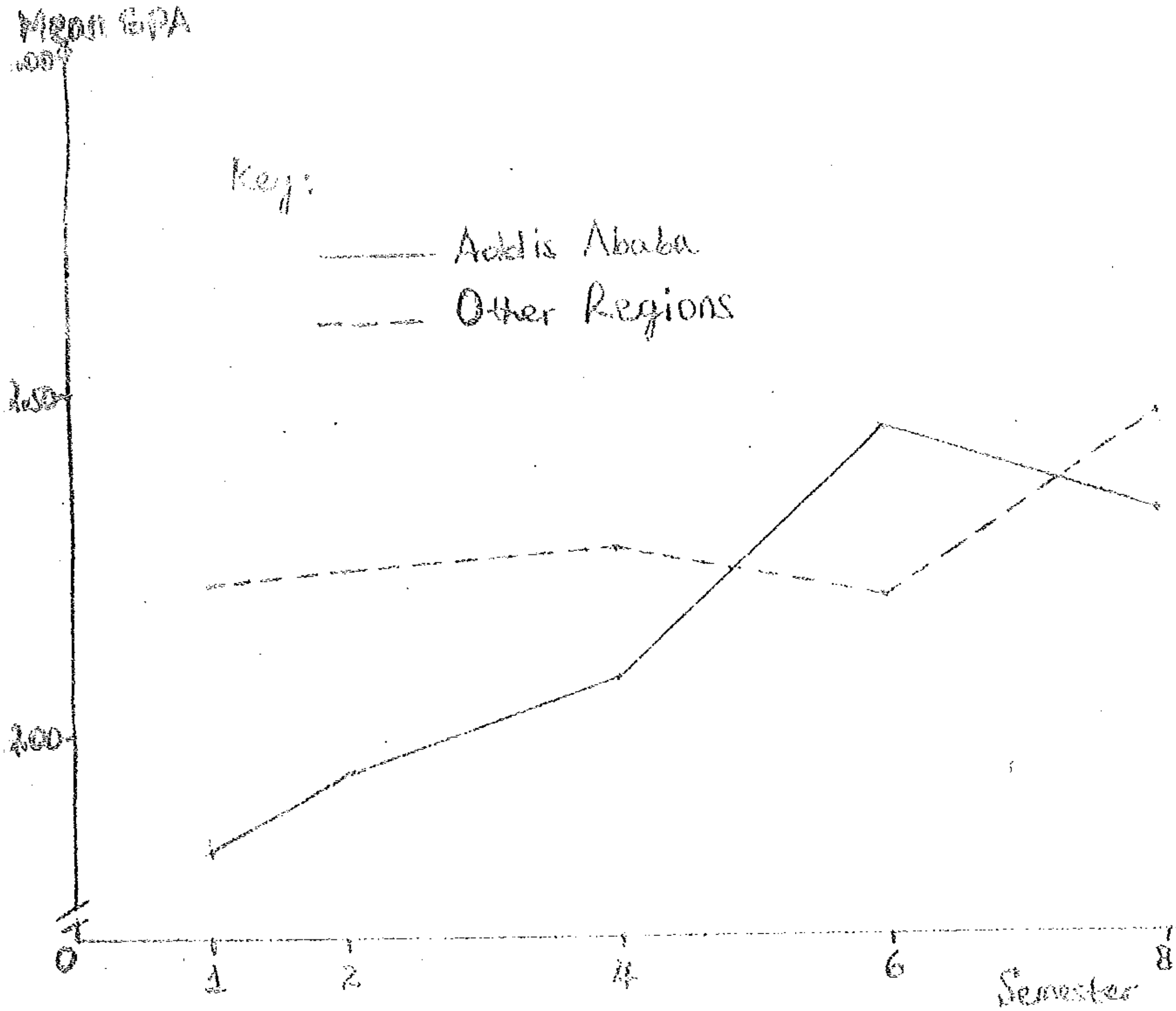


Fig. 3.3: Sample Profiles for the Graduates from the Department of Chemistry.

The profiles for the Department of Geology are shown in Fig. 3.4. Unlike those in the Department of Chemistry, the performance of graduates from Addis Ababa appears to be generally better than that of the Other Regions. There is also an upward trend in the perform-ance of graduates from Addis Ababa relative to that of

Other Regions, starting from second year, second semester.

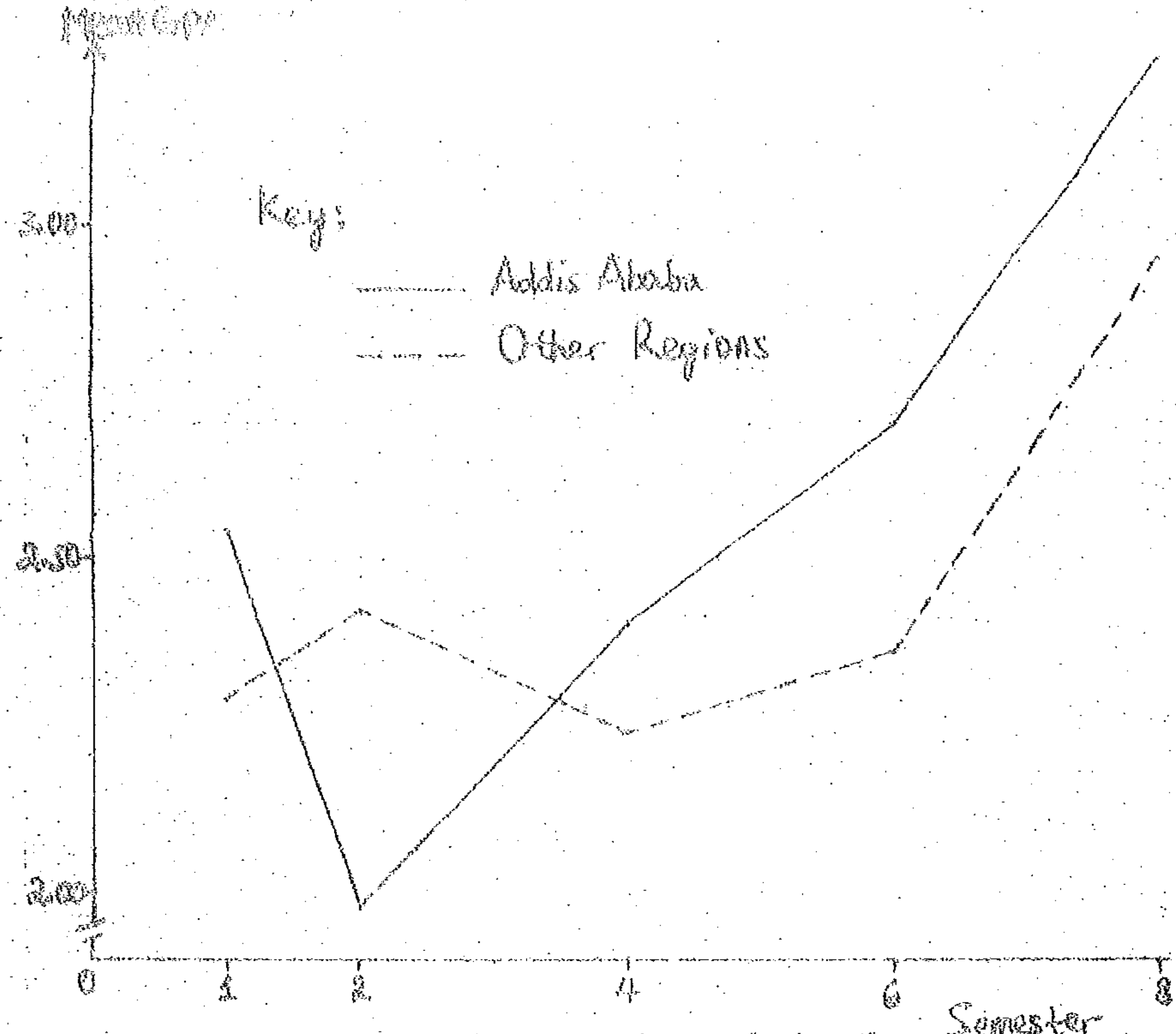


Fig. 3.4: Sample Profiles for the Graduates from the Department of Geology.

The profiles for the Department of Mathematics are shown in Fig. 3.5. Unlike the preceding profiles, these ones do not intersect at all. The performance of the graduates from Other Regions appears to be superior to that of the graduates from Addis Ababa, eventhough the differences of the first two semesters seem to be marginal. Moreover, it appears that there is less consistency in performance among graduates from Addis Ababa than those from Other Regions.

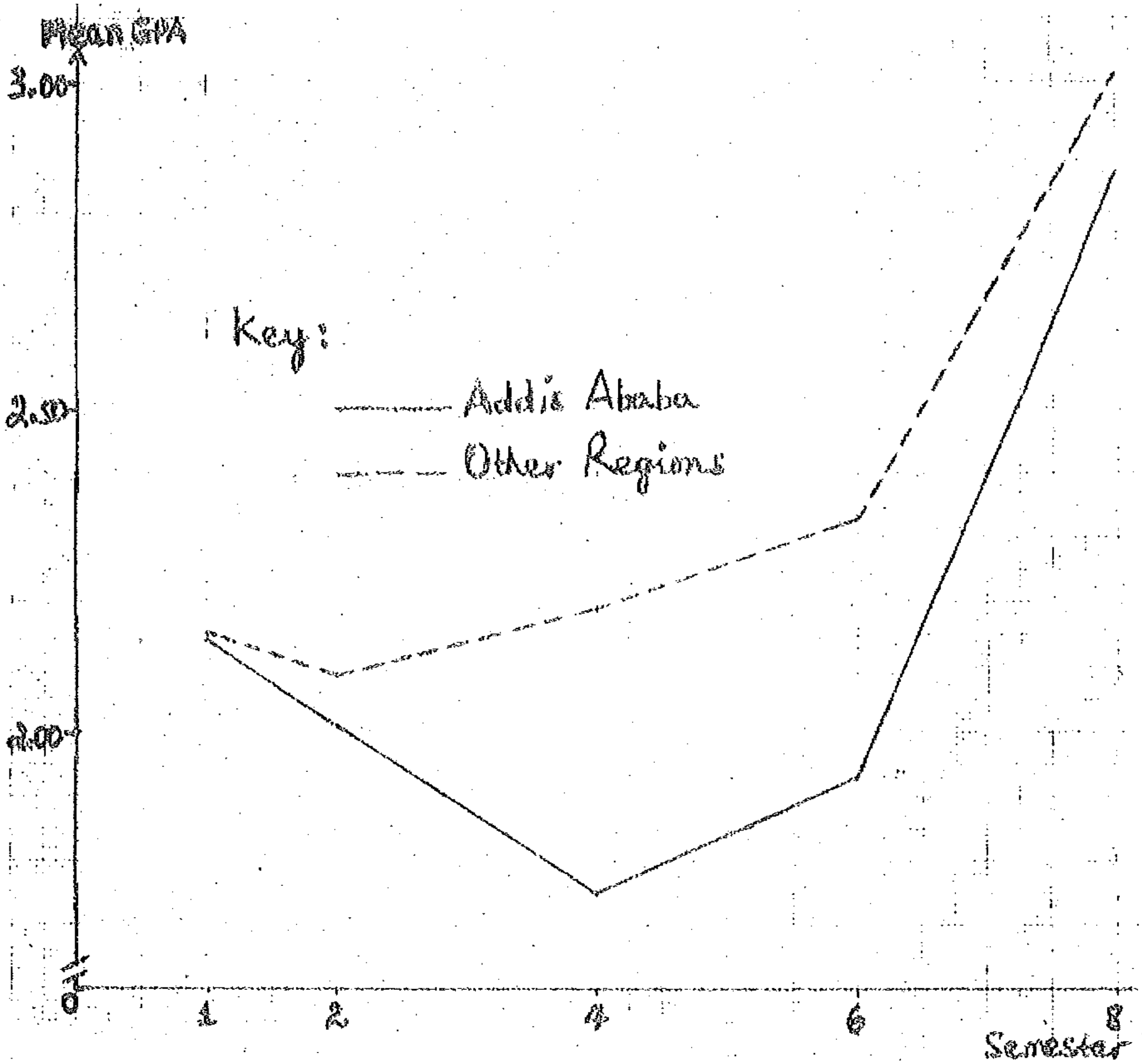


Fig. 3.5: Sample Profiles for the Graduates from the Department of Mathematics

Lastly, Fig. 3.6 show the profiles for the Department of Physics. In this case, the graduates from Other Regions seem to have taken some three years to catch up with the graduates from Addis Ababa. This may be a reflection of the generally inferior Science laboratory facilities available in schools outside Addis Ababa. Unlike the situation in the previous three departments, the performance trend of the graduates from Addis Ababa has remained almost constant.

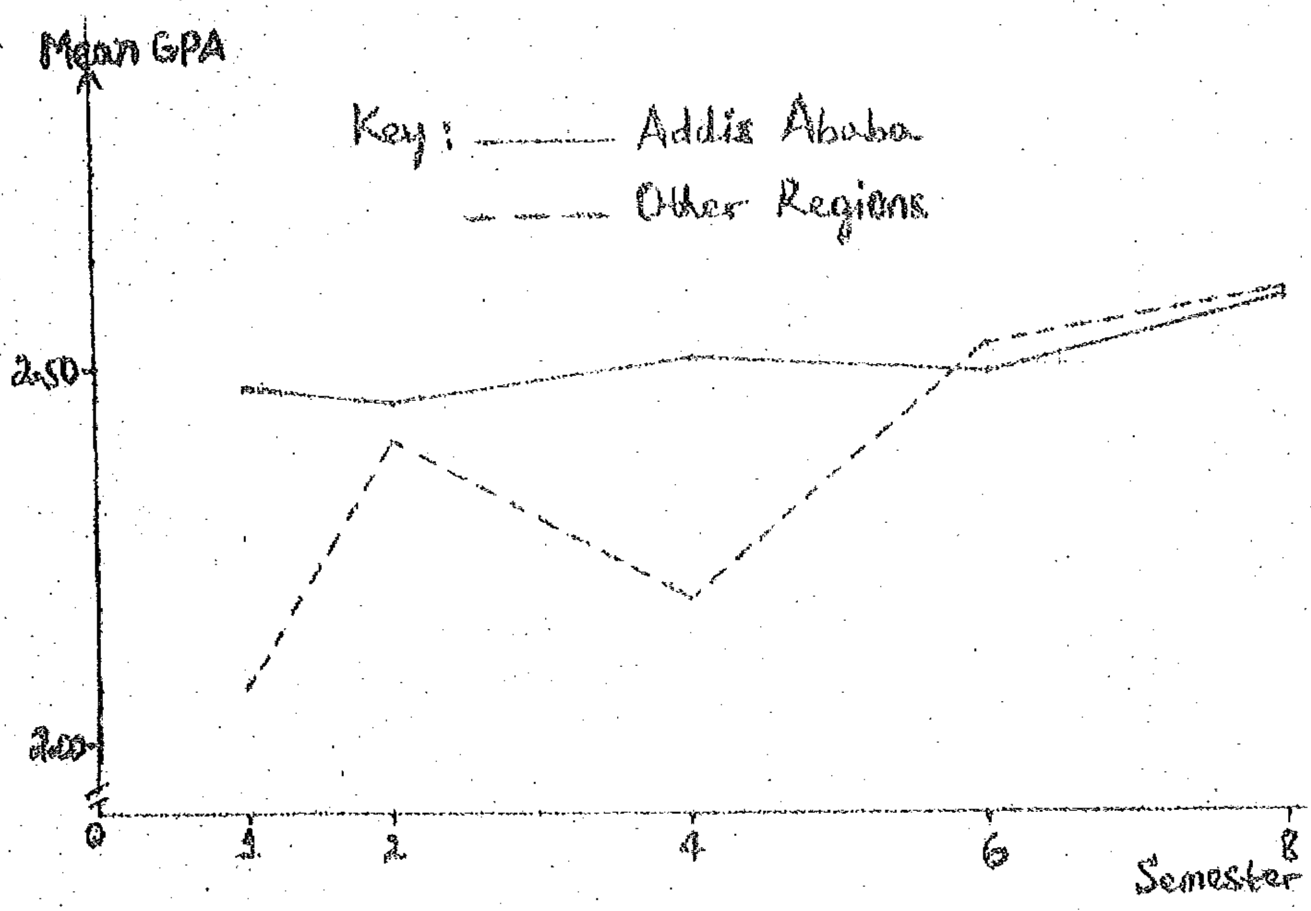


Fig. 3.6: Sample Profiles for the Graduates from the Department of Physics.

# 4. DISCUSSION AND CONCLUSION

Attempts were made to test whether the origin factor resulted in any real difference in performance between students originating from the two groups: those from Addis Ababa and the others which came from other parts of the Country. However, the results obtained are not in favour of the existence of any significant difference.

The hypotheses of parallelism of profiles and that of coincident profiles are accepted. These results appear to indicate that the performance trends are similar for the Statistics graduates from Addis Ababa and Other Regions of the Country. However, the hypothesis of equal response means is rejected. This finding seems to suggest that there is an overall performance trend regardless of origins.

The suggested fluctuation in average performance might be attributed to factors like students' personal problems, which may vary from semester to semester, nature of courses and difference in grading among instructors.

The profiles presented in Section 3.4 suggest that there are different features in different departments for the two groups under study. Such differences may arrise due to difference in background, departmental difference in relation to courses, teaching staff, availability of teaching and reference materials and grading.

Lastly, since the study primarily considered the graduates of the Department of Statistics, these findings have to be considered as informative but not conclusive.

REFERENCES

Gnanadesikan, R. (1977). Methods for Statistical Data Analysis of Multivariate Observations. John Wiley and Sons, New York.

Habte Tewelde-Berhan (1988). "An Investigation into some Factors Affecting Academic Performance of First Year Regular Science Degree Students (1987/88)." An M.Sc. Thesis, Department of Statistics, Faculty of Science, Addis Ababa University

Johnson, Richard A. and Wichern, Dean W. (1982). Applied Multivariate Statistical Analysis. Prentice Hall, New Jersey.

King, M. and King, J. (1972). "Some Correlates of University Performance in a Developing Country: The Case of Ethiopia." The Ethiopian Journal of Education, Vol. V(2), 20-30.

Mekonnen Tadesse (1987). "Ranking and Selection Procedures and their Applications to the Educational Performance of Students at Addis Ababa University." An M.Sc. Thesis, Department of Statistics, Faculty of Science, Addis Ababa University.

Morrison, D.F. (1976). Multivariate Statistical Methods, Second Edition. McGraw Hill, New York.

Appendix A: Form for Data Collection for Graduate Research

1. Name of student ____________________ I.D.No.________
2. Sex: Male __________ Female__________
3. Age when enrolled at the Faculty ___________
4. Marital status: Single _____ Married_____ Other ______
5. Last secondary school attended
   5.1 Name of school ________________________________
   5.2 Location: Town __________ Adm. Region _________
   5.3 Type of school attended:

| | | Code |
|---|---|---|
| Government | ____ | 1 |
| Public | ____ | 2 |
| Mission | ____ | 3 |
| Other | ____ | 4 |

   5.4 Type of study in school: (1 = academic science, 0 = otherwise)
   Academic Science____ Academic Art____ Commercial___
   Home Economics ____ Agriculture ____ Technique/
   Productive Technology _____ other____________________
6. ESLCE GPA ______________
7. Status of admission: Regular ____ Private____ Quota ____
8. Student's first choice ________________
9. Fields of study: Major ______________ Minor ____________
10. Number of credit hrs. taken each semester and GPA

| Year | Semester | Credit hrs. | GPA | CGPA |
|---|---|---|---|---|
| I | I | ________ | ______ | ______ |
| | II | ________ | ______ | ______ |
| II | I | ________ | ______ | ______ |
| | II | ________ | ______ | ______ |
| III | I | ________ | ______ | ______ |
| | II | ________ | ______ | ______ |
| | Total | ________ | ______ | ______ |

11. Dormitory facility: Provided ________ not provided ____
12. Date of graduation ________________

Appendix B: Basic Data

(In all tables that follow, column headings 2,4,6,8 are the second semester of freshman, sophomore, junior and senior years, respectively, and 1 is the first semester of freshman year)

GPA of Statistics Graduates from Addis Ababa

| Serial Number | Semester 2 | Semester 4 | Semester 6 | Semester 8 |
|---|---|---|---|---|
| 1 | 2.41 | 3.24 | 2.75 | 2.68 |
| 2 | 2.63 | 2.59 | 2.85 | 3.21 |
| 3 | 2.63 | 2.24 | 2.10 | 2.84 |
| 4 | 2.00 | 2.24 | 2.10 | 2.32 |
| 5 | 2.38 | 2.88 | 2.00 | 2.00 |
| 6 | 3.19 | 3.41 | 3.12 | 3.84 |

GPA of Statistics Graduates from Other Regions

| Serial Number | Semester 2 | Semester 4 | Semester 6 | Semester 8 |
|---|---|---|---|---|
| 1 | 2.63 | 2.65 | 2.45 | 2.68 |
| 2 | 2.94 | 2.88 | 2.15 | 2.68 |
| 3 | 2.44 | 2.41 | 2.15 | 2.53 |
| 4 | 3.06 | 3.06 | 2.95 | 2.84 |
| 5 | 2.00 | 2.29 | 2.70 | 3.00 |
| 6 | 2.81 | 2.82 | 3.00 | 3.16 |
| 7 | 1.81 | 2.18 | 2.15 | 2.39 |
| 8 | 3.25 | 3.41 | 2.35 | 2.53 |
| 9 | 3.06 | 3.47 | 3.65 | 3.84 |
| 10 | 2.38 | 2.41 | 2.06 | 2.74 |
| 11 | 3.19 | 2.65 | 2.00 | 2.42 |
| 12 | 2.94 | 3.29 | 2.75 | 2.68 |
| 13 | 2.63 | 2.65 | 2.15 | 2.39 |

GPA of Chemistry Graduates from Addis Ababa

| Serial Number | Semester | | | | |
|---|---|---|---|---|---|
| | 1 | 2 | 4 | 6 | 8 |
| 1 | 2.00 | 2.06 | 2.32 | 2.42 | 2.67 |
| 2 | 2.00 | 1.67 | 1.89 | 2.42 | 2.20 |
| 3 | 1.83 | 1.89 | 2.11 | 2.47 | 2.05 |
| 4 | 1.50 | 2.13 | 2.00 | 2.43 | 2.37 |

GPA of Chemistry Graduates from Other Regions

| Serial Number | Semester | | | | |
|---|---|---|---|---|---|
| | 1 | 2 | 4 | 6 | 8 |
| 1 | 2.50 | 2.44 | 3.21 | 3.05 | 3.29 |
| 2 | 1.89 | 2.00 | 1.95 | 1.84 | 2.06 |
| 3 | 2.50 | 2.44 | 2.00 | 2.00 | 2.60 |
| 4 | 2.50 | 2.72 | 2.11 | 2.58 | 2.40 |
| 5 | 2.06 | 2.22 | 2.32 | 2.05 | 2.60 |
| 6 | 1.61 | 2.17 | 2.00 | 1.89 | 2.25 |
| 7 | 2.22 | 1.78 | 1.84 | 1.95 | 2.00 |
| 8 | 2.61 | 2.50 | 3.21 | 2.79 | 2.56 |
| 9 | 2.06 | 1.89 | 2.39 | 2.00 | 2.30 |
| 10 | 2.22 | 1.89 | 2.11 | 2.21 | 3.12 |
| 11 | 2.22 | 2.56 | 1.79 | 1.82 | 1.88 |

GPA of Geology Graduates from Addis Ababa

| Serial Number | Semester 1 | 2 | 4 | 6 | 8 |
|---|---|---|---|---|---|
| 1 | 2.67 | 2.22 | 2.81 | 3.30 | 3.37 |
| 2 | 2.22 | 1.78 | 2.00 | 2.47 | 3.31 |
| 3 | 2.72 | 1.94 | 2.38 | 2.32 | 3.06 |

GPA of Geology Graduates from Other Regions

| Serial Number | Semester | | | | |
|---|---|---|---|---|---|
| 1 | 2.22 | 2.50 | 2.43 | 2.41 | 3.50 |
| 2 | 2.39 | 2.50 | 2.43 | 2.26 | 2.88 |
| 3 | 2.28 | 2.17 | 2.57 | 2.79 | 3.54 |
| 4 | 2.28 | 2.28 | 2.14 | 2.48 | 2.60 |
| 5 | 2.22 | 1.83 | 2.38 | 1.84 | 3.24 |
| 6 | 2.39 | 2.67 | 2.43 | 2.94 | 3.20 |
| 7 | 2.22 | 2.28 | 1.95 | 1.37 | 2.44 |
| 8 | 1.89 | 2.67 | 2.57 | 3.50 | 2.89 |
| 9 | 2.44 | 2.11 | 2.09 | 2.42 | 2.88 |
| 10 | 2.28 | 2.44 | 2.10 | 2.41 | 3.80 |
| 11 | 2.39 | 2.89 | 2.52 | 2.59 | 2.56 |
| 12 | 2.39 | 2.61 | 2.10 | 1.32 | 1.95 |
| 13 | 2.39 | 2.47 | 2.76 | 2.35 | 2.87 |

GPA of Mathematics Graduates from Addis Ababa

| Serial Number | Semester 1 | 2 | 4 | 6 | 8 |
|---|---|---|---|---|---|
| 1 | 2.11 | 2.18 | 1.76 | 2.17 | 2.93 |
| 2 | 2.28 | 1.76 | 1.78 | 2.00 | 3.43 |
| 3 | 2.11 | 1.76 | 1.94 | 2.20 | 3.06 |
| 4 | 2.06 | 2.35 | 1.50 | 1.33 | 2.00 |

GPA of Mathematics Graduates from Other Regions

| Serial Number | Semester 1 | 2 | 4 | 6 | 8 |
|---|---|---|---|---|---|
| 1 | 2.00 | 2.06 | 1.94 | 2.00 | 2.93 |
| 2 | 1.94 | 1.88 | 1.94 | 2.67 | 3.21 |
| 3 | 2.11 | 2.29 | 1.67 | 2.05 | 2.55 |
| 4 | 2.61 | 2.29 | 2.00 | 2.60 | 3.00 |
| 5 | 2.06 | 2.53 | 2.22 | 2.47 | 3.35 |
| 6 | 2.50 | 2.29 | 2.44 | 2.00 | 3.29 |
| 7 | 2.11 | 2.18 | 2.89 | 2.60 | 3.35 |
| 8 | 2.06 | 2.29 | 2.67 | 2.60 | 3.12 |
| 9 | 2.22 | 2.29 | 2.44 | 2.47 | 2.71 |
| 10 | 2.00 | 2.12 | 2.33 | 2.60 | 3.47 |
| 11 | 2.00 | 2.06 | 1.72 | 2.00 | 2.19 |
| 12 | 1.61 | 2.06 | 2.33 | 2.94 | 3.65 |
| 13 | 2.61 | 1.88 | 2.22 | 2.20 | 2.53 |
| 14 | 2.00 | 1.88 | 1.83 | 2.39 | 3.21 |
| 15 | 2.00 | 2.12 | 2.39 | 2.20 | 3.43 |
| 16 | 2.33 | 2.29 | 2.89 | 2.60 | 2.21 |
| 17 | 2.22 | 1.88 | 2.17 | 2.40 | 3.18 |
| 18 | 2.61 | 2.06 | 1.78 | 2.42 | 2.50 |
| 19 | 2.06 | 1.94 | 2.39 | 2.40 | 3.30 |
| 20 | 2.00 | 2.41 | 1.72 | 1.67 | 2.81 |
| 21 | 2.06 | 2.00 | 2.11 | 1.67 | 3.14 |

CPA of Physics Graduates from Addis Ababa

| Serial Number | Semester 1 | 2 | 4 | 6 | 8 |
|---|---|---|---|---|---|
| 1 | 2.67 | 2.47 | 2.33 | 2.00 | 2.00 |
| 2 | 2.67 | 3.05 | 3.67 | 3.53 | 3.13 |
| 3 | 2.22 | 2.41 | 1.94 | 2.44 | 2.82 |
| 4 | 2.67 | 2.24 | 2.94 | 2.53 | 2.38 |
| 5 | 2.11 | 2.06 | 1.67 | 1.90 | 2.60 |

CPA of Physics Graduates from Other Regions

| Serial Number | Semester 1 | 2 | 4 | 6 | 8 |
|---|---|---|---|---|---|
| 1 | 1.83 | 2.12 | 2.28 | 2.41 | 2.00 |
| 2 | 2.67 | 2.00 | 2.22 | 2.82 | 3.00 |
| 3 | 2.22 | 2.18 | 2.89 | 2.26 | 2.07 |
| 4 | 2.17 | 2.59 | 1.72 | 2.13 | 2.25 |
| 5 | 2.00 | 2.18 | 2.28 | 3.00 | 2.54 |
| 6 | 2.22 | 2.65 | 2.22 | 2.71 | 1.81 |
| 7 | 2.00 | 2.12 | 2.83 | 2.63 | 3.15 |
| 8 | 3.39 | 3.00 | 2.33 | 3.05 | 2.69 |
| 9 | 1.83 | 1.88 | 2.10 | 2.81 | 3.00 |
| 10 | 1.83 | 3.18 | 1.61 | 2.50 | 3.00 |
| 11 | 1.61 | 2.29 | 2.00 | 2.75 | 2.69 |
| 12 | 1.67 | 2.65 | 1.72 | 2.38 | 3.00 |
| 13 | 2.00 | 2.00 | 2.39 | 2.53 | 2.54 |
| 14 | 2.28 | 2.94 | 2.67 | 3.19 | 2.63 |
| 15 | 1.89 | 2.47 | 2.06 | 2.47 | 2.81 |
| 16 | 1.56 | 2.35 | 2.22 | 2.11 | 2.46 |
| 17 | 1.61 | 1.68 | 2.06 | 1.68 | 2.00 |
| 18 | 2.17 | 2.44 | 1.83 | 1.56 | 1.84 |
| 19 | 2.06 | 2.33 | 2.17 | 2.79 | 2.46 |
| 20 | 2.67 | 3.24 | 3.39 | 3.75 | 3.55 |
| 21 | 1.83 | 2.06 | 1.83 | 2.37 | 2.19 |
| 22 | 2.28 | 2.65 | 2.39 | 2.47 | 2.63 |
| 23 | 2.00 | 2.24 | 1.61 | 2.00 | 2.81 |
| 24 | 1.83 | 2.29 | 1.61 | 2.50 | 2.46 |
| 25 | 2.22 | 2.35 | 2.44 | 2.41 | 3.38 |

Appendix C: Sample covariance matrices, their inverses and the pooled sample covariance matrix for statistics.

$$S_1 = \begin{bmatrix} 0.1546 & 0.1119 & 0.1324 & 0.2239 \\ 0.1119 & 0.2473 & 0.1640 & 0.1362 \\ 0.1324 & 0.1640 & 0.2277 & 0.2671 \\ 0.2239 & 0.1362 & 0.2671 & 0.4270 \end{bmatrix}$$

$$S_1^{-1} = \begin{bmatrix} 89.6887 & -44.6010 & 69.3038 & -76.1538 \\ -44.6010 & 32.0396 & -47.2812 & 42.7428 \\ 69.3038 & -47.2812 & 86.7092 & -75.4975 \\ -76.1538 & 42.7428 & -75.4975 & 75.8657 \end{bmatrix}$$

$$S_2 = \begin{bmatrix} 0.1992 & 0.1577 & 0.0592 & 0.0324 \\ 0.1577 & 0.1790 & 0.1216 & 0.0741 \\ 0.0592 & 0.1216 & 0.2363 & 0.1742 \\ 0.0324 & 0.0741 & 0.1742 & 0.1588 \end{bmatrix}$$

$$S_2^{-1} = \begin{bmatrix} 22.6489 & -25.3658 & 10.7676 & -4.5966 \\ -25.3658 & 38.6274 & -20.2805 & 9.8647 \\ 10.7676 & -20.2805 & 34.5712 & -30.6574 \\ -4.5966 & 9.8647 & -30.6574 & 36.2625 \end{bmatrix}$$

$$S = \begin{bmatrix} 0.1861 & 0.1442 & 0.0807 & 0.0887 \\ 0.1442 & 0.1991 & 0.1341 & 0.0924 \\ 0.0807 & 0.1341 & 0.2338 & 0.2015 \\ 0.0887 & 0.0924 & 0.2015 & 0.2377 \end{bmatrix}$$

where $S_1$ is for Addis Ababa

$S_2$ is for Other Regions

$S$ is the pooled sample covariance matrix.

Appendix D: Ordered squared distances and chi-square percentile points

For Addis Ababa

| $j$ | $d^2_{(j)}$ | $\chi^2(\frac{j-\frac{1}{2}}{6})$ |
|---|---|---|
| 1 | 2.25 | 0.94 |
| 2 | 2.79 | 1.92 |
| 3 | 3.51 | 2.38 |
| 4 | 3.68 | 4.03 |
| 5 | 3.69 | 5.39 |
| 6 | 4.06 | 8.36 |

For Other Regions

| $j$ | $d^2_{(j)}$ | $\chi^2_4(\frac{j-\frac{1}{2}}{13})$ |
|---|---|---|
| 1 | 0.32 | 0.60 |
| 2 | 0.86 | 1.15 |
| 3 | 0.94 | 1.59 |
| 4 | 2.58 | 2.03 |
| 5 | 2.84 | 2.47 |
| 6 | 3.44 | 2.92 |
| 7 | 3.99 | 3.36 |
| 8 | 4.15 | 3.99 |
| 9 | 4.35 | 4.61 |
| 10 | 5.15 | 5.24 |
| 11 | 5.20 | 6.31 |
| 12 | 6.21 | 7.54 |
| 13 | 7.97 | 10.26 |

For the Combined Sample

| j | $d^2_{(j)}$ | $\chi^2_4(\frac{j-\frac{1}{2}}{19})$ |
|---|---|---|
| 1 | 0.48 | 0.49 |
| 2 | 0.79 | 0.91 |
| 3 | 0.81 | 1.24 |
| 4 | 2.16 | 1.54 |
| 5 | 2.38 | 1.84 |
| 6 | 2.40 | 2.15 |
| 7 | 2.46 | 2.45 |
| 8 | 2.82 | 2.76 |
| 9 | 3.52 | 3.06 |
| 10 | 3.84 | 3.36 |
| 11 | 3.87 | 3.79 |
| 12 | 3.90 | 4.21 |
| 13 | 4.38 | 4.64 |
| 14 | 4.54 | 5.07 |
| 15 | 5.31 | 5.60 |
| 16 | 5.82 | 6.44 |
| 17 | 6.12 | 7.27 |
| 18 | 6.34 | 8.50 |
| 19 | 9.06 | 11.04 |