62

ATTRITION RATE AT THE FACULTY OF SCIENCE, ADDIS ABABA UNIVERSITY

A Thesis

Presented to the

School of Graduate Studies

Addis Ababa University

In Partial Fulfilment of the

Requirements for the Degree of Master of Science

in Statistics

By

TEMESGEN ZEWOTIR

November, 1991.

ACKNOWLEDGEMENT

I wish to express my foremost heartfelt gratitude to my advisor, Professor Aynew Ejigou, for his unlimited assistance, keen encouragement and indispensable advice throughout the study.

My gratitude also goes to Ato Debalkew Azale who contributed his limitless assistance in writing the computer program.

Many thanks are also due to all my friends and my family for their special assistance and encouragement.

Finally, I would like to thank W/o Bogalech H/Mariam for typing this manuscript so quickly and efficiently.

ABSTRACT

In this paper, the attrition rates of students from the period 1980/81 to 1989/1990, in the Faculty of Science of Addis Ababa University, were examined. We found that freshmen attrition rate linearly decrease over time. The dampening effect of the attrition rate as the students progress from semester to semester was assessed using two attrition models: the Guttman and Olkin (1989) model and a reparameterized model that we formulated. Computer programs were prepared to obtain the maximum likelihood estimates for the underlying parameters in each model. Consistent results were found in the two models; the progress of Students in consecutive semesters appeared to show a difference from batch to batch.

TABLE OF CONTENTS

1 INTRODUCTION

The success of any development depends upon human resources and upon the level of qualitative considerations which they can develop. In this context, they can be described as the skill, knowledge, attitude, motivation and resourcefulness that constitute the human catalyst in bringing about improvements in production, services, technology and management. To inculcate these human attributes in individuals, it goes without saying that the education system is the main agent. Within any given economy, it is not the unskilled or semi-skilled who are the front agents in the high value industrial and modern agricultural production. Human resources for higher economic productivity are derived from middle and high level professionals.

In particular, third level educational institutions are set up to produce such human resources that development critically demands. Higher education increases man's capacity for bringing about socio-economic transformation. Besides, it imparts specific knowledge, creates a generalized capacity and receptivity to new ideas, instills disciplines, self-confidence, civic responsibility and organizational skill in the learner.

In spite of the fact that higher education is exceptionally important, it is open only to a fraction of the young talented citizens. Even then, unfortunately, not all of this group successfully goes through our four-year university system of education. Because students are selected on the basis of their ability, it does not necessarily mean they wish to take advantage of such selection. More importantly learning ability and study skill are all affected by motivational factors.

Therefore, the proportion of students who discontinue their studies on account of academic dismissal or voluntary dropout, which is referred to as an attrition rate, is substantial and it remains a major problem. The magnitude of the attrition rate in higher education institutions is an important research problem of great concern and it is an important and relevant issue for the assessment of the overall learning environment and its operational efficiency.

Baumgart and Johnstone (1977) observed that the magnitude of the attrition rate demands remedial action or intervention strategies of some kind. What kind of remedial action or intervention strategies are possible? They noted three possibilities: the first possibility is by refining selection procedures; the second possibility is to devise procedures to assist students to integrate both socially and academically into the university environment and the third possibility is that an institution may identify its own procedure and environment to meet the desire to bring education closer to the need and life of the community and fulfil individual aspirations more effectively.

Of course, to identify which remedial action or intervention strategy need to be taken, further research should be done. However, one thing that we observe here is that the specific finding in the magnitude of the attrition rate reflects selection procedures and/or social and academic integration and/or the quality of the instructional process.

Spady (1970), in a review and synthesis of literature on dropouts from higher education, noted two major factors believed to affect the magnitude of the attrition rate. According to this view, each student enters a higher education institution with a pattern of dispositions, interests, expectations, goals, and values shaped by his family background and high school experience. These interact with the particular college or university and lead to a certain level of integration into the academic and social systems of the institution. Thus, it may influence students overall ability to accommodate the influence and pressures that the student encounters in his new environment. This can be reflected through the magnitude of the attrition rate. The magnitude of attrition rate may have implications of a pattern of development of family background and the relation between secondary level education and third level education.

It is clear from the above paragraphs that complex designs can be adopted and developed. Such a study guides later research to be more efficiently planned and conducted. A detailed study of the magnitude of the attrition rate is a prerequisite for creating a more attractive and efficient academic environment. However, research in this regard is rare.

Most of the time, in the university, student budget planners and administrators use students' grade point average (GPA) to determine the number of possible admissions. The same technique is also used to predict the future working force. However, since students with high grade point average also left their study, the technique has not been reliable. As discussed by Guttman and Olkin (1989), and Spady (1970), the study of attrition rate is a central issue in such predictive studies. It is fairly obvious to say that it is a more reliable technique than that based on the grade point average alone.

In our Country, in any given year, some proportion of students leave their study for any of a number of academic or non-academic reasons. This happens in almost all higher education institutions. However, except for statements similar to the one above there is no officially published evidence that quantifies the extent of attrition.

Actually some research has been done about the probable cause of attrition among freshman students but there is none on the magnitude of attrition. Most of this research focused on the selection. The grade point average in the Ethiopian School Leaving Certificate Examination (ESLCE) is used as the chief criterion for admission to universities or colleges. Many of the references and summaries of the inconsistency of the research results can be obtained from a paper by Habte (1988). The findings in these studies vary from year to year, but no one knows the effect of attrition rates.

It is apparent that there is a dampening effect of the rates of attrition as the students progress from semester to semester or year to year. This is due to the fact that students entering at a later semester or year have completed more prerequisites and are also generally more motivated and acquainted with the environment. Based on this consideration, a study of the dampening effect of attrition rate is the core idea for our understanding of how students progress in our institutions.

Some academic departments are more successful than others in creating a good learning environment by making courses interesting and relevant and this motivates students to seek advanced training. Thus it also seems to be important to examine the variation in the attrition rate between departments within the Faculty of Science. So far, no study has been carried out in this regard. The present study is concerned with statistical analysis and examination of students' attrition rates in the 1980s'. The purpose of this study is, therefore, to:

1. Investigate the stability of freshmen attrition rate over time.
2. Study the dampening effect of attrition rate as the students progress to their senior year.

If administrators and planners in other faculties can perceive similarities between their own environment and the one described here, the results may be useful beyond the Faculty of Science. Moreover, the results of the study may serve educational researchers as a first step for designing detailed attrition studies. In addition, it may also help in budgetary and manpower planning by providing estimates of the work force at a future date in different fields.

It was originally planned to collect more data than those used in this study. However, as in most third-world countries, the system of file management in the Records Office of Addis Ababa University is not well-organized. For example, in fourth year, the attrition report makes no clear distinction between dropouts and students who have failed to graduate. On account of this ambiguity in the data, the coverage of the study includes students up to third year. Besides this, due to special admission considerations applied to war veterans in the period June to August 1985, the freshmen attrition report of that period has not been found suitable for our study. Thus, the 1984/85 batch of freshmen admissions is also excluded from this study.

2 THE DATA AND METHODS OF ANALYSIS

2.1 THE DATA

Data were collected on the total number of regular undergraduate degree program students enroled at the beginning of each semester in the Faculty of Science of Addis Ababa University as freshman, second year, and third year students. Correspondingly, the number of students who were either dropouts or academic dismissals at the end of the semester were collected as attrition.

The data were collected from the Record Office of the Faculty of Science and the Record Office of the Registrar of Addis Ababa University. Since two of these Record Offices had no complete enrolment and attrition records for the period prior to the academic year 1982/83, the data from 1982/83 to 1989/90 were used in the study. Fortunately, 1980/81 and 1981/82 freshmen enrolment and attrition records were available in the Record Offices. Hence, unlike data for the other groups, the data for our analysis of freshmen attrition rate is based on ten academic years. The data in the two sources were thoroughly reviewed and organized in a suitable way. Disagreement between the two sources, which was rare, was discussed and solved with concerned officers.

2.2 METHODS OF ANALYSIS

2.2.1 Trend Analysis

The systematic change of freshmen attrition rate through time was examined by means of trend analysis. In this analysis the following steps were followed.

i. Examining the scatter plot of the attrition rate against time.

ii. Formulating a general trend that relates the freshmen attrition rate to time.

iii. Fitting the trend.

iv. Examining the adequacy of the fitted trend.

v. Making inference based on the fitted trend.

The detailed theory that justifies these is found in Draper and Smith (1981), Nurminen (1986). Chapman and Nam (1960), Cook and Weisberg (1982).

2.2.2 Attrition Modelling

Suppose that an institution offers a bachelor's degree program of study over a specified number of semesters, say K+2, excluding the semester immediately after admission. Hence, each degree student must complete K+3 semester courses so that students are expected to progress through these K+3 semesters. However, some students may decide to leave the program whereas others who have an equivalent background may re-enter the program. We will, therefore, follow the progress of each cohort, including those readmitted, through the first K semesters after the semester following admission into the faculty.

Hence, the data required in the study are: (i) the total number of students enroled each semester and these consist of the new students plus those carrying over from the previous semester; (ii) the number of attrition at the end of each semester. We shall lay out these as follows:

Table 2.2.2.1

A lay out and notation for the attrition data

Semester	1	2	. . .	K
Students enroled	N_1	N_2		N_k
Number of attrition	a_1	a_2		a_k
Number of students remaining	b_1	b_2		b_k
Attrition rate per semester	p_1	p_2		p_k

Clearly,

$$b_i = N_i - a_i$$

$$p_i = a_i/N_i, \ i = 1,2,\ldots,k.$$

We assume that there is the dampening effect of the rates of attrition as we progress from one semester to another. Because of this dampening effect of the attrition rates, the proportions will generally decrease with time. That is,

$$p_1 \geq p_2 \geq \ . \ . \ . \ \geq p_k$$

The interest in this study is to model the above process. In the next two sections, two different models will be assessed. One is due Guttman and Olkin (1989) and the other is new.

(i) The Guttman and Olkin Method

Guttman and Olkin (1989) model the process by using π to denote the initial probability of attrition for the first semester. Further, they use ρ, $0<\rho\leq 1$, to represent a dampening effect, so that the probability of attrition for the second semester is $\pi\rho$. If we assume a constant dampening effect of the attrition rate, the probability of attrition for the third semester is $\pi\rho^2$, and so on. Thus, the model of Guttman and Olkin for a given batch, becomes:

$$p_i = \pi\rho^{i-1}, \; i = 1,2,\ldots,k. \qquad (1)$$

The problem is to estimate the initial probability π and the dampening effect ρ. The analysis of the model begins with the joint likelihood, L, where L is given by

$$\prod_{i=1}^{k}\binom{N_i}{a_i}p_i^{a_i}(1-p_i)\,b_i \qquad (2)$$

The assumption in (2) is that there is no carry-over effect from semester to semester, and the observations from semester to semester are independent, which is not, of course, quite true.

After taking the logarithm of L and using (1) in (2), setting the derivatives, with respect to π and ρ, equal to zero yields:

$$\partial\frac{logL}{\partial\pi}=\sum_{i=1}^{k}\frac{a_i}{\pi}-\sum_{i=1}^{k}\frac{b_i\rho^{i-1}}{(1-\pi\rho^{i-1})}=0 \qquad (3)$$

$$\partial\frac{logL}{\partial\rho}=\frac{1}{\rho}\left(\sum_{i=1}^{k}(i-1)a_i-\sum_{i=1}^{k}\frac{(i-1)b_i\pi\rho^{i-1}}{(1-\pi\rho^{i-1})}\right)=0 \qquad (4)$$

From these two likelihood equations, we can find the maximum likelihood estimates of π and ρ. To obtain the asymptotic distribution of the maximum likelihood estimates, MLE(Π,ρ), the information matrix is given as:

$$I=\begin{pmatrix} -E(\frac{\partial^2 logL}{\partial\pi^2}) & -E(\frac{\partial^2 logL}{\partial\pi\partial\rho}) \\ -E(\frac{\partial^2 logL}{\partial\pi\partial\rho}) & -E(\frac{\partial^2 logL}{\partial\rho^2}) \end{pmatrix}$$

$$=\begin{pmatrix} \frac{1}{\pi^2}\sum_{i=1}^{k}\frac{N_i\pi\rho^{i-1}}{(1-\pi\rho^{i-1})} & \frac{1}{\pi\rho}\sum_{i=1}^{K}\frac{(i-1)N_i\pi\rho^{i-1}}{(1-\pi\rho^{i-1})} \\ \frac{1}{\pi\rho}\sum_{i=1}^{k}\frac{(i-1)N_i\pi\rho^{i-1}}{(1-\pi\rho^{i-1})} & \frac{1}{\rho^2}\sum_{i=1}^{k}\frac{N_i\pi\rho^{i-1}}{(1-\pi\rho^{i-1})} \end{pmatrix} \qquad (5)$$

So that

$$(\hat{\pi}, \hat{\rho})' \sim N((\pi,\rho)', I^{-1}). \qquad (6)$$

That is, $(\hat{\pi}, \hat{\rho})'$ is asymptotically bivariate normal with dispersion matrix, I^{-1}.

All the details of theses derivations are available in Guttman and Olkin (1989). Our main task and problem is to solve for Π and ρ from equations (3) and (4). But the estimates of Π and ρ cannot be easily obtained from these two non-linear simultaneous equations. However, it is not too difficult to estimate them using an iteration method (Singer, 1964). Unless a computer program is developed the solution cannot be easily found manually. Hence a computer program was written to estimate these two parameters and the inverse of information matrix. This program is attached herewith as Appendix 1. The validity of the model is examined using the X^2 goodness of fit test.

(ii) The Reparameterized Model

When the initial probability is not only the effect of first semester considered but when the initial probability is assumed to be in a moderately rapid drift downward from one semester to the next, this method might be an appropriate representation.

Here the probability of attrition for a given batch, for the ith semester is p_i, where:

$$p_i = pe^{-ir}, \; i = 1,2,\ldots,k \qquad (7)$$

p = initial probability

r = dampening effect parameter, r>0.

To obtain the maximum likelihood estimates, we again start with the likelihood (2), where the p_i's are now given by (7). The likelihood equations are:

$$\frac{\partial logL}{\partial p} = \sum_{i=1}^{k} \frac{a_i}{p} - \sum_{i=1} \frac{b_i e^{-ir}}{1-pe^{-ir}} = 0 \qquad (8)$$

$$\frac{\partial logL}{\partial r} = \sum_{i=1}^{k} \frac{ib_i pe^{-ir}}{1-pe^{-ir}} - \sum_{i=1}^{k} ia_i = 0 \qquad (9)$$

To obtain the asymptotic distribution of the MLE, $(\hat{p},\hat{r})$, we evaluate the information matrix

$$I = \begin{pmatrix} -E(\frac{\partial^2 logL}{\partial p^2}) & -E(\frac{\partial^2 logL}{\partial p \partial r}) \\ -E(\frac{\partial^2 logL}{\partial r \partial p}) & -E(\frac{\partial^2 logL}{\partial r^2}) \end{pmatrix}$$

$$= \begin{pmatrix} \sum_{i=1}^{k} \frac{N_i e^{-2ir}}{1-pe^{-ir}} + \sum_{i=1}^{k} \frac{N_i e^{-ir}}{p}) & \sum_{i=1}^{k} \frac{iN_i e^{-ir}}{1-pe^{-ir}} \\ \sum_{i=1}^{k} \frac{iN_i e^{-ir}}{1-pe^{-ir}} & \sum_{i=1}^{k} \frac{i^2 pe^{-ir}}{1-pe^{-ir}} \end{pmatrix} \qquad (10)$$

from which we get the asymptotic distribution:

$$(\hat{p},\hat{r})' \sim N((p,r)', I^{-1}) \qquad (11)$$

As in the method above, the solution of the likelihood equations cannot be obtained in closed form. Hence, we use an iterative method (Singer, 1964) and the computer program for the solution of the above two non-linear simultaneous equations, (8) and (9), is attached at the back as Appendix 2. The chi-square goodness of fit statistic for the validity of the model has also been computed.

Most of the standard methods of statistical inference are based on the familiar assumption that the random variables have a normal distribution, or perhaps equivalently the sample size is sufficiently large to justify asymptotic normality. In such cases tests of hypothesis may be simple. In our case, however, testing hypothesis about students' attrition rates may be far from optimum. However, when we pooled observations across several batches, the size of the sample was fairly adequate to construct a likelihood ratio test.

3 ANALYSES AND RESULTS

3.1 ANALYSIS OF TREND IN RATES

Here we wish to study a series of freshman attrition rates. In the data shown in Table 3.1.1, our first concern of analysis is to test the consistency (or homogeneity) of the underlying attrition rates.

Table 3.1.1

Freshmen Attrition Rates

Academic Year	Total number of students enroled	Number of attrition	Attrition rate
1980/81	1475	485	0.3788
1981/82	906	242	0.2671
1982/83	1405	345	0.2477
1983/84	1073	325	0.3029
1984/85	1337	287	0.2147
1985/86	1077	249	0.2312
1986/87	806	126	0.1563
1987/88	793	139	0.1753
1988/89	770	121	0.1571
1989/90	732	96	0.1311

To test the hypothesis of homogeneity of attrition rate in the past ten years, i.e.

H_{01} : $P_i = P$

H_{11} : $P_i \neq P$ for at least one i, for i = 1,2,...,10,

where P_i is the attrition rate in year i.

The test statistic is:

$$\chi^2=\sum_{i=1}^{n} N_i \frac{(p_i-\bar{p})^2}{p_i(1-P_i)}$$

where,

N_i = the total number of students in year i

P_i = the attrition rate in year i

$$\bar{p}=\sum_{i=1}^{n} \frac{N_i p_i}{\sum_{i=1}^{n} N_i}$$

which has a chi-square distribution with degrees of freedom n-1. Using the above formula for our data we get 251.0245. The magnitude of chi-square which has nine degree of freedom leads to rejection of the null hypothesis. Therefore, a significant difference appears to exist among the rates.

Since the hypothesis of constant rates is not found to be tenable, we may ask whether there is a trend over time. Thus, it appears reasonable to test for any progressive relation between the attrition rate, P, and the academic years, t. To test the hypothesis of no association between t and P as in Nurminen (1986),

$$H_{o2} : \rho = 0$$

$$H_{12} : \rho \neq 0$$

We first compute the point biserial correlation coefficient, ρ, as a measure of progressive relationship; ρ is then estimated by

$$\hat{\rho} = \frac{\sum_{i=1}^{n} \sqrt{N}(t_i - \bar{t})\, a_i}{\left[a(N-a)\sum_{i=1}^{n} N_i (t_i - \bar{t})^2\right]^{\frac{1}{2}}}$$

where

a_i = the number of attrition in year i

t_i = year i

$a = \sum a_i$

$N = \sum N_i$

So that for our data $\hat{\rho} = -0.2374$. Nurminen's test statistic under the null hypothesis is:

$$F_0 = \frac{N(N-2)\left[\sum_{i=1}^{n} (t_i - \bar{t})\, a_i\right]^2}{a(N-a)(1-\hat{\rho}^2)\sum_{i=1}^{n} N_i (t_i - \bar{t})^2}$$

which is distributed as F with 1 and N-2 degrees of freedom. Thus, the basic F value under the null hypothesis on the basis of our data becomes $F_0 = 197.107$. With degrees of freedom 1 and 10373, the point biserial correlation is found highly significant. As the significance of the association is clearly evident, there is interest in extending the analysis.

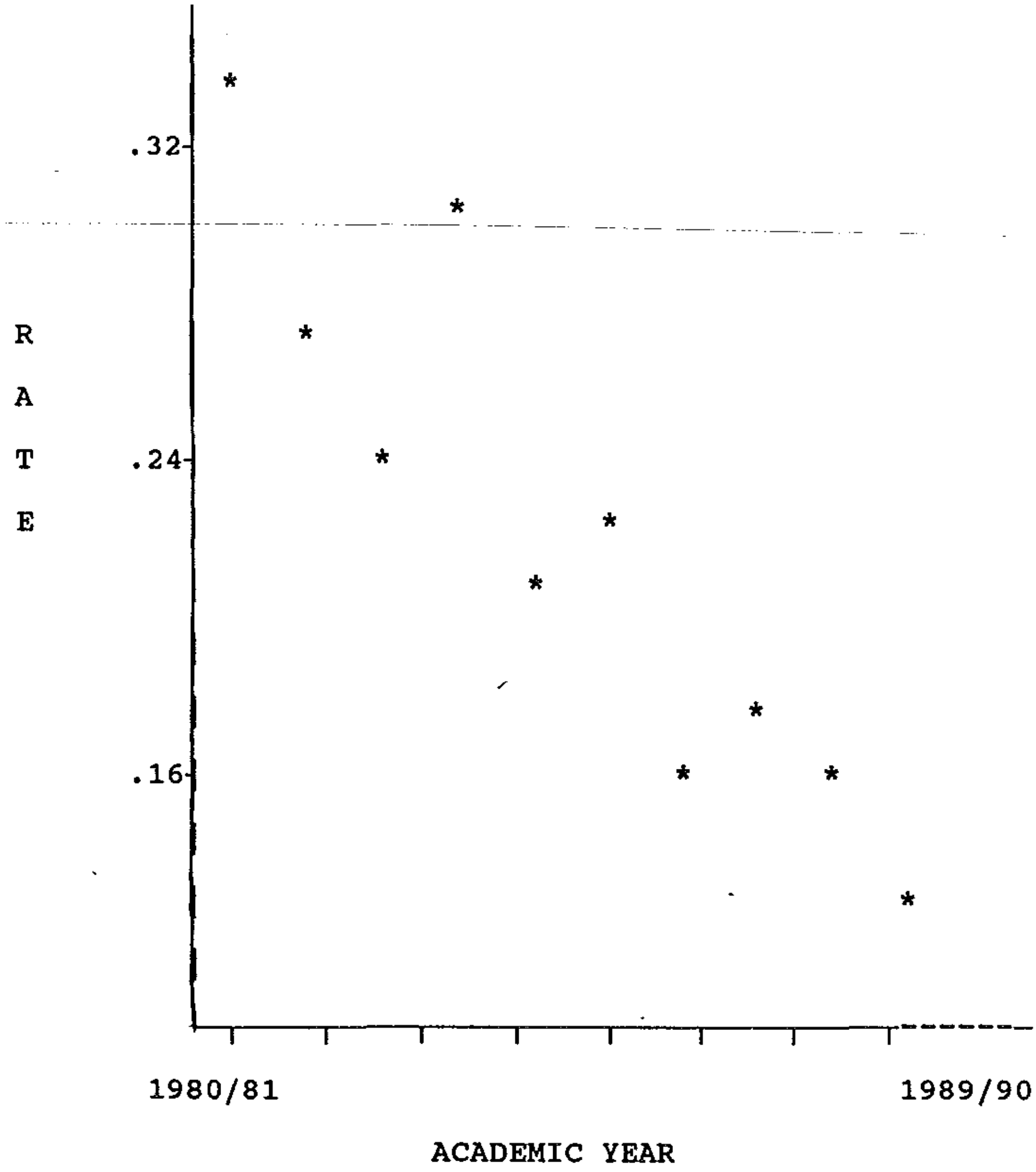

Figure 3.1.1 Scatter Plot of Attrition Rate Versus Academic Year.

To do this, the first step is to plot the rates against time in order to get some idea of a particular model in which the observations can be explained. A scatter plot of the data in Table 3.1.1 is given in Fig. 3.1.1. Where we see that a straight line model appears plausible, although cases 4 and 7 appear to dominate our perception of this plot. Therefore, we restrict consideration to a linear relationship between the two variables P and t. Our model is then:

$$P = \alpha + \beta t + e \qquad (1)$$

Where α and β are unknown constants and e is a random disturbance term. A method for estimating α and β is the so-called the weighted least square method. The weight to be assigned to p_i, as used by Chapman and Nam (1960), is W_i, where

$$W_i = N_i/(P_i(1-P_i)), \quad i = 1,2,\ldots,10.$$

So that, the fitted line becomes

$$\hat{P}_i = 0.33121-0.02022t \qquad (2)$$

As can be seen in Table 3.1.2, the actual and estimated rates are close in magnitude. Actually, R^2 (the proportion of total variation about the mean of p explained by the regression) is 0.852. This shows that the fitted equation (2) explains 85.2% of the total variation about the average of P.

Table 3.1.2

Observed and Predicted Attrition Rates

t	W	P	$\hat{P}$	$\hat{e}$
1	6683.57	0.3288	0.3110	0.0178
2	4628.17	0.2671	0.2908	-0.0237
3	7539.79	0.3029	0.2705	-0.0228
4	5081.66	0.3029	0.2503	0.0526
5	7929.83	0.2147	0.2301	-0.0154
6	6059.19	0.2312	0.2099	0.0213
7	6112.07	0.1563	0.1896	-0.0333
8	5485.24	0.1753	0.1694	0.0059
9	5814.85	0.1571	0.1492	0.0079
10	6425.97	0.1311	0.1290	0.0021

This regression model is based on certain assumptions about the distribution of the value of e. They are crucial for the estimates of the parameters. In order to have faith in the conclusion, we must be convinced that our assumptions are not seriously violated. So, after estimation of parameters with the method of weighted least square, we should establish how trustworthy these estimates are. Thus, in order to attach importance to R^2 we should first make sure that the basic assumptions are satisfied in any particular case.

We start by examining the assumption of normality which is necessary for conducting the statistical tests of significance of the parameter estimates and for constructing confidence intervals. The normal probability (P-P) plot is used to test the assumption of normality. Fig.3.1.2 is a normal plot of studentized residuals for the model (1). With only 10 cases, the observed plot have no strong evidence against normality of the error term. Hence normality is satisfied (Cook and Weisberg, 1982).

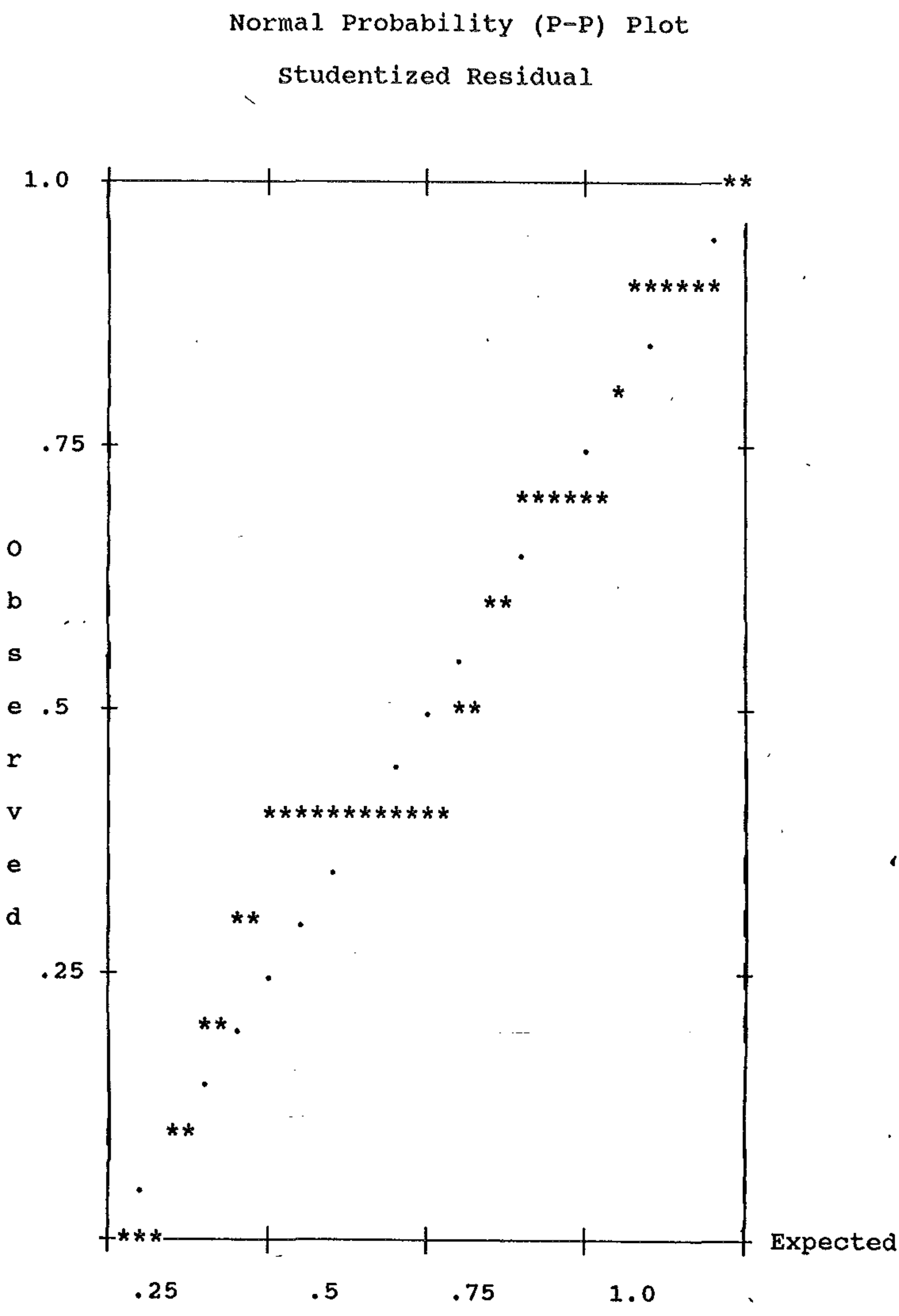

Figure 3.1.2 Normal probability plot

The second concern is in the examination of the status of suspected outlier and influential points. Here Cook's formulation is used to assess the influence of the ith point on estimators of the parameters. The indication of the influence of each point using Cook's distance is very small. For instance, the largest value of Cook's distance appears in case 4, d_4 = 0.00004, which is very small when compared with theoretical F value with degrees of freedom 2 and 8. This implies none of the observations has a special influence on the estimated regression parameters.

The third assumption is that the successive values of the random variable e are mutually independent; that is, the value which e assumes in any one period is independent of the value which it assumed in any previous period. In short, we assume that, there is no autocorrelation or serial correlation of the error terms. We examined this using the Durbin-Watson test. The test measures the extent of first-order autoregressive scheme ($e_t = \rho e_{t-1} + u_t$). The test may be outlined as follows (Johnston, 1984).

The null hypothesis is

$$H_{03} : \rho = 0$$

or H_{03}: the e's are not autocorrelated with first-order scheme. This hypothesis is tested against the alternative hypothesis

H_{13}: $\rho \neq 0$

or H_{13}: the e's are autocorrelated with first-order scheme. To test the null hypothesis the Durbin-Watson statistic

$$d=\frac{\sum_{t=2}^{n}(e_t-e_{t-1})^2}{\sum_{t=1}^{n}e_t^2}$$

The test compares the empirical d value, calculated from the regression residual, with the d_L (lower) and d_U (upper) limit in the Durbin-Watson table at test size α. The testing procedure is as follow

1. If $d < d_L$ or $4-d < d_L$, reject the null hypothesis and accept that there is autocorrelation.

2. If $d > d_U$ or $4-d > d_U$, do not reject the null hypothesis

3. Otherwise, the test is inconclusive.

The empirical d value calculated from the regression residuals in Table 3.1.2, is 2.956. From the Durbin-Watson table, with 2 percent level of significance, n = 10 observations, and K′=1 independent variable the significance points of d_L and d_U are

$$d_L = 0.604$$

$$d_U = 1.001$$

Since $d = 2.956 > d_U$ and $4-d > d_U$, we conclude there is no first order autocorrelation. All these assure us of the validity of the basic assumptions we use in the weighted regression model of (1).

We now return to the analysis of variance in Table 3.1.3. Where the observed F ratio is compared with the theoretical F value with 1 and 8 degrees of freedom. The scope is found to be significant. That is, time, t, is a significant explanatory factor for the variation of P.

Table 3.1.3

Analysis of Variance for the Regression

Source of variation	Sum of squares	Degree of freedom	Mean square	F
Regression	205.90307	1	205.90307	47.457
Residual	34.71010	8	4.33876	
Total	240.61317	9		

Based on the regression analysis derived above, we have already remarked that the fitted regression is highly significant. In addition, it is necessary to measure the standard errors of the estimates of the parameters; they are used to form confidence intervals and to test hypothesis concerning the parameters. The estimates of the intercept and the regression coefficient, their standard errors, the t statistic and the p value after comparing t's with t(8) are summarized in Table 3.1.4.

Table 3.1.4

Coefficient Analysis

Coefficient	Estimate	S.E.	t	Sig t
β	-0.02022	2.9356E-3	-6.889	0.0001
α	0.33121	0.01813	18.271	0.000

Thus $H_{04} : \beta = 0$ is not accepted as is $H_{05} : \alpha = 0$ at the 5% level. We, therefore, have some reliable evidence in the data to state that $H_{14} : \beta \neq 0$ and $\alpha \neq 0$ are more plausible, and that the freshman attrition rate has been linearly related to academic year as:

$$\hat{P} = 0.33121 - 0.02022t.$$

This fitted straight line implies a decrease of 2% in the freshmen attrition rate per year. Why this decrease came about and how it come about would have been interesting to investigate, but this would be outside the scope of this research.

3.2 MODELLING ATTRITION RATES

To estimate the dampening effect of the attrition rate as students progress through five consecutive semesters, we shall use the procedure discussed in Section 2.2.2.

Table 3.2.1 displays the attrition rate of different batches of students, where Batch 1 is the group of 1982/83 entries, including readmitted students, to the Faculty of Science after successfully completing first semester course. Similarly, with the same characterization Batch 2, Batch 3, Batch 4 and Batch 5 are 1983/84, 1985/86, 1986/87 and 1987/88 entries, respectively.

Table 3.2.1

Observed Attrition Rates by Batch

Batch	Semester 1	2	3	4	5
1	0.2940	0.1892	0.0938	0.0613	0.0377
2	0.1975	0.1264	0.0862	0.0812	0.0803
3	0.2325	0.1354	0.0677	0.0529	0.0483
4	0.1605	0.1156	0.0988	0.0358	0.0312
5	0.1353	0.1187	0.0922	0.0627	0.0167

As we can see from the Table above, the attrition rate decreases as we progress from semester to semester. Assuming the rate decreases from semester to semester by a constant factor, we will estimate this constant dampening effect of the attrition rate. To do this first let us apply the Guttman and Olkin (1989) method to our data.

3.2.1 Guttman and Olkin Model

Table 3.2.1.1 presents the estimates of π and ρ, and their standard deviations. These approximate standard deviations are obtained from the inverse of the information matrix. All these estimates are obtained using the computer program in Appendix 1.

Table 3.2.1.1

Estimates of Parameters of Guttman and Olkin Model

Batch	$\hat{\pi}$	$\hat{\rho}$	S.E.($\hat{\pi}$)	S.E.($\hat{\rho}$)
1	0.2978	0.5911	1.543E-2	2.383E-2
2	0.1817	0.7666	1.248E-2	2.912E-2
3	0.2234	0.6265	1.510E-2	2.922E-2
4	0.1683	0.6696	1.470E-2	0.362E-2
5	0.1518	0.7072	1.463E-2	0.395E-2

From the estimates presented above we can estimate the attrition rates using:

$$\hat{P}_i = \hat{\pi} \; \hat{\rho}^{\,i-1} \quad \text{for } i = 1,2,\ldots 5.$$

The goodness of fit test statistic value for the model and the estimated attrition rate values are given in Table 3.2.1.2. Here, the chi-square goodness of fit statistics, each with 4 degrees of freedom, are all small. Hence the model fits our data.

Table 3.2.1.2

Estimated Attrition Rates and chi-squares for Guttman and Olkin Method

Batch	Semester 1	2	3	4	5	χ^2
1	0.2978	0.1760	0.1040	0.0615	0.0364	0.002
2	0.1817	0.1393	0.1068	0.0819	0.0628	0.011
3	0.2324	0.1400	0.0877	0.0549	0.0344	0.011
4	0.1683	0.1127	0.0755	0.0505	0.0338	0.012
5	0.1518	0.1073	0.0759	0.0537	0.0379	0.018

When we compare the estimated proportion in Table 3.2.1.2 with the actual proportion, Table 3.2.1, the corresponding values are close in magnitude for all or most categories. That is why the chi-square values are very small.

It is interesting to draw some inferences about the parameter. We prefer to consider carrying out tests on the dampening effect parameters. Let ρ_j be the dampening effect of the attrition for the jth Batch. We consider the hypothesis

$$H_{06} : \rho_1 = \rho_2 = \rho_3 = \rho_4 = \rho_5 = \rho$$

against the alternative

H_{16} : At least one ρ_j, j = 1,2,...,5 is different. The likelihood function is

$$L=\prod_{j=1}^{5}\prod_{i=1}^{5}\binom{N_{ij}}{a_{ij}}(\pi_j\rho_j^{i-1})^{a_{ij}}(1-\pi_j\rho_j^{i-1})^{b_{ij}}$$

where i and j represent semester and Batch, respectively. We will compare the highest value of the likelihood under H_{06}, i.e.

Sup L/H_{06}

with its highest value under the model, i.e.

Sup L

A problem that we may solve is that the maximum likelihood estimates of the parameter $\pi_1, \pi_2, \pi_3, \pi_4, \pi_5$ and ρ. That is, ML estimates of the parameters under H_{06}. As usual we will solve the log likelihood equations, which are as follows:

$$\frac{\partial logL/H_{06}}{\partial\rho}=\frac{1}{\rho}\left(\sum_{j=1}^{5}\sum_{i=1}^{5}(i-1)\left(a_{ij}-\frac{b_{ij}\Pi_j\rho^{i-1}}{1-\pi_j\rho^{i-1}}\right)=0\right.$$

$$\frac{\partial logL/H_{06}}{\partial\pi_j}=\sum_{j=1}^{5}\sum_{i=1}^{5}\left(\frac{a_{ij}}{\pi_j}-\frac{b_{ij}\rho^{i-1}}{1-\pi_j\rho^{i-1}}\right)=0$$

for j = 1, 2, . . ., 5.

The solution of these six non-linear simultaneous equations cannot be obtained easily. However, using the iteration method of Demidovich and Maron (1987) a computer program is developed. This program is also attached as Appendix 3. Applying the program to our data, the estimates are presented in Table 3.2.1.3.

Table 3.2.1.3

Estimates of the Parameters under H_{06}

$\hat{\pi}_1$	$\hat{\pi}_2$	$\hat{\pi}_3$	$\hat{\pi}_4$	$\hat{\pi}_5$	ρ
0.2700	0.2129	0.2186	0.1692	0.1623	0.6609

The likelihood ratio for testing H_{06} against H_{16} is, therefore

$$\Lambda = \frac{\text{Sup } L/H_{06}}{\text{Sup } L}$$

Thus, asymptotically, $-2\log\Lambda$ is distributed, under H_{06}, as a chi-square with 4 degrees of freedom. Using the above results, we obtain a value of 48.8. This value is in favour of the alternative hypothesis. Therefore, our data suggests that the dampening effect of the attrition rate, as the students progress through five consecutive semesters, differs from Batch to Batch.

3.2.2 Our Reparameterization

Now let us see the results obtained using our own reparameterization. Using the computer program in Appendix 2, the estimates of the parameter are presented in Table 3.2.2.1.

Table 3.2.2.1

Estimates of the Parameters of our model

Batch	$\hat{P}$	$\hat{r}$	S.E.($\hat{P}$)	S.E.($\hat{r}$)
1	0.5038	0.5258	1.102E-3	0.434E-2
2	0.2369	0.2658	0.573E-3	0.622E-2
3	0.3566	0.4676	1.031E-3	0.594E-2
4	0.2514	0.4011	0.953E-3	0.781E-2
5	0.2153	0.3478	0.869E-3	0.837E-2

From these estimates, the estimated attrition rates, using

$$\hat{P}_i = \hat{P}e^{-i\hat{r}} \qquad i = 1,2,\ldots,5,$$

and the goodness of fit test statistic for the model are given in Table 3.2.2.2. The chi-square goodness of fit statistics are all found to be very small each with 4 degrees of freedom.

Table 3.2.2.2

Estimated Attrition Rate and chi-square values for Our Reparameterization

Batch	Semester 1	2	3	4	5	χ^2
1	0.2978	0.1760	0.1040	0.0615	0.0364	0.002
2	0.1816	0.1392	0.1068	0.0818	0.0627	0.011
3	0.2234	0.1400	0.0877	0.0549	0.0344	0.011
4	0.1683	0.1127	0.0755	0.0505	0.0338	0.012
5	0.1520	0.1074	0.0758	0.0535	0.0378	0.012

Now, again, it is of interest to test the equality of the dampening effect parameters. That is

H_{07} : $r_1 = r_2 = r_3 = r_4 = r_5 = r$

H_{17} : all r_j are not equal, $j = 1,2,..,5.$

We accomplish this by applying the procedure as one above except here

$$L=\prod_{j=1}^{5}\prod_{i=1}^{5}\binom{N_{ij}}{a_{ij}}(p_j e^{-ir_j})^{a_{ij}}(1-P_j e^{-ir_j})^{b_{ij}}$$

The computer program is attached as Appendix 4. Using this program, the estimates are presented in Table 3.2.2.3.

Table 3.2.2.3

Estimates of the Parameters Under H_{07}

$\hat{P}_1$	$\hat{P}_2$	$\hat{P}_3$	$\hat{P}_4$	$\hat{P}_5$	$\hat{r}$
0.4052	0.3197	0.3166	0.2541	0.2438	0.4063

Using the results, twice the negative of log likelihood ratio value is 27.6. This result also suggests that the attrition rate dampening parameters are different from Batch to Batch. Thus, the two models leads to the same conclusion.

Actually, it can be easily seen from Table 3.2.1.2 and Table 3.2.2.2 that the estimated values are identical in most of the categories. In some batches the estimated values are identical up to eight digits throughout the progress. In general, the two models can be used equivalently for fitting attrition rates that drift rapidly downward with time.

4 DISCUSSION AND CONCLUSION

In this study, several general conclusions can be derived from the findings. Correspondingly, several explanations can also be suggested for the specific findings. We, therefore, discuss situations that may reflect the specific changes and developments related to our study.

Although all persons have the right of access to education, higher education at this stage of development is not a personal right. It goes without saying that the secondary school graduates that should enter higher education institutions must naturally reach the necessary academic standard, which is often arbitrarily defined. Two questions naturally arises: how many or what proportion should qualify? What specialization should they be directed to pursue?

Though all are anxious to proceed to higher education institutions, secondary school graduates are usually larger in number than the available space. Many are inadequately prepared and some have a misplaced choice, and consequently, they fail to pass the necessary academic attainment tests. The sole criterion for college or university admission is the GPA in the Ethiopian School Leaving Certificate Examination (ESLCE). The minimum score, in ESLCE, required for admission to higher education institutions increases from one period to another so that the quality of students progressively increases. The minimum GPA in ESLCE to enter the Faculty of Science has increased from 2.2 in 1980/81 to 3.2 by 1989/90.

The number of incoming freshman students to the Faculty of Science has been reduced from 1475 in 1980/81 to 732 in 1989/90. This may have upgraded the academic performance of freshman students since a decrease of intake may have made available a more balanced share of facilities in the faculty. Consequently, it may have partly led to an improved level of academic performance over the academic years: 1980/81 to 1989/90.

Students' performance and persistence is influenced to some extent by wider social, economic, and political factors. Of course, this changes slowly through time. These include, among others, students' opportunities for securing local employment. For example, a change in local unemployment rate could produce a change in students attitudes towards education.

Given these major reasons, there is little room for surprise in the results of our analysis. Many of the changes we discussed strictly increase or decrease from year to year.

In the model approach, the tests about the dampening effect in the two models gave consistent results. The test we have described for attrition rate dampening effect was the likelihood ratio test. The test assumes that the sample size is fairly large. In a similar study, Gross and Lam (1981) used a still smaller sample: a sample size of ten. Here, as Armitage (1955, P.377) points out, its distribution is strictly valid asymptotically and this holds true as the N_i (the total number of case in category i) increase. Thus, for our sample, the test is likely to be meaningful and appropriate in leading to valid conclusions.

In both models, throughout the five consecutive semesters the estimated values for Batch 1 and Batch 3 are identical up to eight digits. Also the estimates in Batch 4 are identical up to five digits. An inspection of the actual data indicates that in these three batches the probability of attrition decreases at a rapid rate. Hence, either of the two models can readily be used for such data to attain essentially the same result. The reparameterized model provides a more conventional representation of the process of attrition rate decay. Its goodness of fit is, however, the same as for Guttman and Olkin model.

To summarize the discussion so far, our study reveals a decrease of attrition rate of freshman students over the academic year 1980/81 to 1989/91. However, no firm and general conclusion can be reached or concrete suggestions proposed as to why it linearly decreased. Nevertheless, we feel justified in concluding that the progress of students in consecutive semesters varies from Batch to Batch. Moreover, the two models can be used equivalently for rapid downward drift data.

REFERENCES

Armitage, P. (1985). Test for linear trends in proportion and frequencies. *Biometrics* 11, 375-385.

Baumgart, L.N. and Johnstone, N.J. (1977). Attrition at Australian University: A case study. *The Journal of Higher Education* 48, 553-570.

Chapman, D.G. and Nam, J. (1968). Asymptotic power of chi-square tests for linear trends in proportions. *Biometrics* 24, 315-327.

Cook, R.D. and Weisberg, S. (1982). *Residual and Influence in Regression*. New York: Chapman and Hall.

Demidovich, B.P. and Maron, I.A. (1987). *Computational Mathematics*. 4th ed., Moscow: Mir Publishers.

Draper, N. and Smith, H. (1981). *Applied Regression Analysis*. 2nd ed., New York: John Wiley.

Gross A.J. and Lam, C.F. (1981). Paired observations from survival distribution. *Biometrics* 37, 505-511.

Guttman, I. and Olkin, I. (1989). Retention and attrition models. *Journal of Educational Statistics* 14(1). 1-20.

Habte Tewoldeberhan (1988). *An Investigation into Some Factors Affecting Academic Performance of First Year Regular Science Degree Students 1987/88.* An M.Sc. Thesis, Addis Ababa University.

Johnson, J. (1984). *Econometric Methods.* 3rd ed., Tokyo: McGraw-Hill.

Nurminen, M. (1986). Analysis of trends in proportion with an ordinally scaled determinant. *Biometrical Journal* 28, 965-974.

Singer, J. (1964). *Elements of Numerical Analysis.* New York: Academic Press

Spady, G.W. (1970). Dropout from higher education: An interdisciplinary review and synthesis. *Interchange* 1 (3), 64-85.

APPENDIX 1

A PROGRAM FOR APPROXIMATION OF MLE OF GUTTMAN AND OLKIN MODEL

```
 const
   max_itt = 200;
type
    pi_type = array[0..max_itt] of real;
var
    i,k,j,l,l1          : integer;
    loop_itt,ex         : integer;
    pi,rhow,n,a,b       : pi_type;
    a1,a2,b1,b2         : real;
    a11,a12,a21,a22         : real;
    b11,b12,b21,b22         : real;
    file1                   : text;

    c1,c2               : real;
    eps, chi            : real;
    round, sum          : real;
    conv_ok,close_ok: boolean;
    sing_ok             : boolean;
function expo(a,x:real):real;
begin
    expo  := exp(x*ln(a));
end;
function f(pi:real;rhow:real): real;
var   i : integer;
    first_term  : real;
    second_term : real;
function first_comp(ai:real) : real;
begin
       first_comp := ai/pi;
end;
function second_comp(bi:real) : real;
begin
     second_comp :=  (bi*expo(rhow,i-1))/(1-pi*expo(rhow,i-1));
end;
begin
    first_term := 0.0;
    second_term:= 0.0;
    for i := 1 to k do begin
       first_term := first_term + first_comp(a[i]);
       second_term:= second_term + second_comp(b[i]);
    end;
       f           := first_term - second_term;
end;

function g(pi:real;rhow:real): real;
var   i : integer;
    first_term  : real;
    second_term : real;
function first_comp(ai:real) : real;
```

```
begin
      first_comp := (i-1)*ai;
end;
function second_comp(bi:real) : real;
begin
     second_comp          :=     ((i-1)*bi*pi*expo(rhow,i-1))/(1-
pi*expo(rhow,i-1));
 end;
begin
    first_term := 0.0;
    second_term:= 0.0;
    for i := 1 to k do begin
       first_term := first_term + first_comp(a[i]);
       second_term:= second_term + second_comp(b[i]);
    end;
       g          := (first_term - second_term)/rhow;
end;
(* Partial derivatives *)
function f_pi(pi:real;rhow:real): real;
var   i : integer;
    first_term  : real;
    second_term : real;
function first_comp(ai:real) : real;
begin
       first_comp := ai/sqr(pi);
end;
function second_comp(bi:real) : real;
begin
      second_comp          :=     (bi*expo(rhow,2*i-2))/sqr(1-
pi*expo(rhow,i-1)); end;
begin
    first_term := 0.0;
    second_term:= 0.0;
    for i := 1 to k do begin
       first_term := first_term + first_comp(a[i]);
       second_term:= second_term + second_comp(b[i]);
    end;
    f_pi       := -(first_term + second_term);
end;
function f_rhow(pi:real;rhow:real): real;
var   i : integer;
    first_term  : real;
    second_term : real;
function second_comp(bi:real) : real;
begin
     second_comp          :=     (bi*(i-1)*expo(rhow,i-2))/sqr(1-
pi*expo(rhow,i-1));
 end;
begin
    second_term:= 0.0;
    for i := 1 to k do begin
       second_term:= second_term + second_comp(b[i]);
    end;
    f_rhow        := -second_term;
end;
function g_pi(pi:real;rhow:real): real;
```

```
var   i : integer;
    second_term : real;
function second_comp(bi:real) : real;
begin
     second_comp          :=    (bi*(i-1)*expo(rhow,i-1))/sqr(1-
pi*expo(rhow,i-1)); end;
begin
    second_term:= 0.0;
    for i := 1 to k do begin
       second_term:= second_term + second_comp(b[i]);
    end;
    g_pi        := -second_term;
end;
function g_rhow(pi:real;rhow:real): real;
var   i : integer;
    first_term  : real;
    second_term : real;
    third_term  : real;
function first_comp(ai:real) : real;
begin
       first_comp := ai*(i-1)/sqr(pi);
end;
function second_comp(bi:real) : real;
begin
     second_comp   := ((i-1)*(i-2)*bi*pi*expo(rhow,i-3))/sqr(1-
pi*expo(rhow,i-1));
 end;
function third_comp(bi:real):real;
begin
     third_comp    :=  ((i-1)*bi*sqr(pi)*expo(rhow,i-4))/sqr(1-
pi*expo(rhow,i-1)); end;
begin
    first_term := 0.0;
    second_term:= 0.0;
    third_term := 0.0;
    for i := 1 to k do begin
       first_term := first_term + first_comp(a[i]);
       second_term:= second_term + second_comp(b[i]);
       third_term := third_term  + second_comp(b[i]);
    end;
    g_rhow         := -(first_term + second_term + third_term);
end;

function det(a1,a2,b1,b2 : real ):real;
begin
     det := a1*b2 - a2*b1;
end;
function norm(x1,x2 :real ) : real;
begin
     norm := sqrt(sqr(x1) + sqr(x2));
end;
function pi_jbar(i:integer;pi,rhow : real ):real;
begin
    pi_jbar := 1 - pi*expo(rhow,i-1);
end;
function pi_j(i:integer;pi,rhow : real ):real;
```

```
begin
    pi_j := pi*expo(rhow,i-1);
end;
function inf_11(pi,rhow : real): real;
var
   j    : integer;
   term : real;
function pi_jbar(i:integer;pi,rhow : real ):real;
begin
    pi_jbar := 1 - pi*expo(rhow,i-1);
end;
function pi_j(i:integer;pi,rhow : real ):real;
begin
    pi_j := pi*expo(rhow,i-1);
end;
begin
   term := 0.0;
   for j:=1 to k do begin
       term := term + n[j]*pi_j(j,pi,rhow)/pi_jbar(j,pi,rhow);
  end;
   inf_11 := term/sqr(pi);
end;
function inf_12(pi,rhow : real): real;
var
   j    : integer;
   term : real;
function pi_jbar(i:integer;pi,rhow : real ):real;
begin
    pi_jbar := 1 - pi*expo(rhow,i-1);
end;
function pi_j(i:integer;pi,rhow : real ):real;
begin
    pi_j := pi*expo(rhow,i-1);
end;
begin
   term := 0.0;
   for j:=1 to k do begin
                          t  e  r  m    : =    t  e  r  m    +
n[j]*(j-1)*pi_j(j,pi,rhow)/pi_jbar(j,pi,rhow);
   end;
   inf_12 := term/(pi*rhow);
end;
function inf_22(pi,rhow : real): real;
var
   j    : integer;
   term : real;
function pi_jbar(i:integer;pi,rhow : real ):real;
begin
    pi_jbar := 1 - pi*expo(rhow,i-1);
end;
function pi_j(i:integer;pi,rhow : real ):real;
begin
    pi_j := pi*expo(rhow,i-1);
end;
begin
   term := 0.0;
```

```
   for j:=1 to k do begin
                        t e r m  : =  t e r m  +
n[j]*sqr(j-1)*pi_j(j,pi,rhow)/pi_jbar(j,pi,rhow);
    end;
   inf_22 := term/sqr(rhow);
end;
procedure inv_22;
function det(a1,a2,b1,b2 : real ):real;
begin
     det := a1*b2 - a2*b1;
end;
begin
      b11 := a22/det(a11,a12,a21,a22);
      b12 := -a21/det(a11,a12,a21,a22);
      b21 := -a12/det(a11,a12,a21,a22);
      b22 := a11/det(a11,a12,a21,a22);
end;
begin
     assign(file1,'e:\tez\result5');
     rewrite(file1);
     close_ok := false;
     while not close_ok do begin
       writeln('Give the maximum no. of itteration to be
          carried out');
       readln(loop_itt);
       writeln('Give the periscribed epsilon(sufficiently
          small!!)');
       readln(eps);
       writeln('ENTER INITIAL PARAMETERS');
          writeln('Enter the summation index...k');
       readln(k);
       writeln('Give n[i],a[i]...i=1,..k');
       for i:=1 to k do read(n[i]);
       readln;
       for i:=1 to k do read(a[i]);
       readln;
          for i:=1 to k do
         b[i] := n[i] - a[i];

          l := 0;conv_ok := false;sing_ok := false;
          pi[0]   := a[1]/n[1];
          rhow[0] := (a[2]*n[1])/(a[1]*n[2]);
          writeln(file1,'Results of each itteration ');

(*     writeln(pi[0],rhow[0],f(pi[0],rhow[0]),g(pi[0],rhow[0]));
*)     while ( not conv_ok ) and ( l < loop_itt ) do begin
          a1 := f_pi(pi[l],rhow[l]);a2 := g_pi(pi[l],rhow[l]);
           b1   :=   f_rhow(pi[l],rhow[l]);b2   :=
          g_rhow(pi[l],rhow[l]);
     if det(a1,a2,b1,b2) <> 0 then begin
          c1 := f(pi[l],rhow[l]);c2 := g(pi[l],rhow[l]);
              pi[l+1]   :=   pi[l]   -   (   b2*c1-b1*c2
               )/det(a1,a2,b1,b2);
          rhow[l+1]   :=   rhow[l]   +   (   a2*c1-
          a1*c2)/det(a1,a2,b1,b2);
          l := l + 1;
```

```
(* write(pi[l],rhow[l]); *)
if ( norm(f(pi[l],rhow[l]),g(pi[l],rhow[l])) < eps )
   then              conv_ok := true;
  end else begin
sing_ok := true;ll := l;
l := loop_itt;
  end;
end;
   if conv_ok then begin
writeln('Convergency  attained[Up  to  the  given
prescision] ......!!');
write(file1,'The  approximate  solution  vector  [
Pi,Rhow].. ',pi[l],rhow[l]);
writeln(file1,f(pi[l],rhow[l]),g(pi[l],rhow[l])); a11
:= inf_11(pi[l],rhow[l]);
a12 := inf_12(pi[l],rhow[l]);
a21 := a12;
a22 := inf_22(pi[l],rhow[l]);
inv_22;
   writeln(file1,'The  information  matrix-inverse');
writeln(file1,b11,b12); writeln(file1,b21,b22);
writeln(file1,'THE ESTIMATED & ACTUAL VALUE');
chi :=0.0;
for j := 1 to k do begin
          chi    :=    chi    +    sqr(a[i]/n[i]    -
pi_j(j,pi[l],rhow[l]))/
pi_j(j,pi[l],rhow[l]);
    writeln(file1,pi_j(j,pi[l],rhow[l]),'
            ',a[j]/n[j]);
end;
     writeln(file1,'THE CHI-SQUARE VALUE IS', chi);
   sum :=0.0;
     for j :=1 to k do begin

sum :=sum+a[j]*ln(pi_j(j,pi[l],rhow[l]))+
   b[j]*ln(1-pi_j(j,pi[l],rhow[l]));
     end;
    writeln(file1,'THE RATIO ESTIMATE',sum);
end else begin
if sing_ok then begin
writeln('Singularity-faced.........');
   write(file1,'The approximate solution vector [
     Pi,Rhow].. ');
writeln(file1,pi[ll],rhow[ll],f(pi[l],
rhow[l]), g(pi[l],rhow[l] ));
a11 := inf_11(pi[l],rhow[l]);
a12 := inf_12(pi[l],rhow[l]);
a21 := a12;
a22 := inf_22(pi[l],rhow[l]);
inv_22;
   writeln(file1,'The  information  matrix-inverse');
writeln(file1,b11,b12); writeln(file1,b21,b22);
writeln(file1,'THE ESTIMATED & ACTUAL VALUE');
chi :=0.0;
for j := 1 to k do begin
```

```
                chi    :=    chi    +    sqr(a[i]/n[i]    -
        pi_j(j,pi[1],rhow[1]))/
        pi_j(j,pi[1],rhow[1]);
            writeln(file1,pi_j(j,pi[1],rhow[1]),'
                       ',a[j]/n[j]);
               end;
              writeln(file1,'THE CHI-SQUARE VALUE IS', chi);
             sum :=0.0;
              for j := 1to k do begin

          sum :=sum+a[j]*ln(pi_j(j,pi[1],rhow[1]))+
                 b[j]*ln(1-pi_j(j,pi[1],rhow[1]));
               end;
             writeln(file1,'THE RATIO ESTIMATE',sum);
             end else begin
        writeln('The required convergency not
                  attained.....!!');
           write(file1,'The approximate solution vector [
                    Pi,Rhow]..    ',pi[1],rhow[1]);
        writeln(file1,f(pi[1],rhow[1]),g(pi[1],rhow[1]));
          a11 := inf_11(pi[1],rhow[1]);
        a12 := inf_12(pi[1],rhow[1]);
        a21 := a12;
        a22 := inf_22(pi[1],rhow[1]);
        inv_22;
           writeln(file1,'Inverse   information   matrix');
        writeln(file1,b11,b12); writeln(file1,b21,b22);
        writeln(file1,'THE ESTIMATED & ACTUAL VALUE');
       chi :=0.0;
        for j := 1 to k do begin
                  chi := chi + sqr(a[i]/n[i] -
          pi_j(j,pi[1],rhow[1]))/
              pi_j(j,pi[1],rhow[1]);
             writeln(file1,pi_j(j,pi[1],rhow[1]),' '
                  ,a[j]/n[j]);
               end;
              writeln(file1,'THE CHI SQUARE VALUE IS', chi);
             sum :=0.0;
              for j := 1 to k do begin

          sum :=sum+a[j]*ln(pi_j(j,pi[1],rhow[1]))+
                   b[j]*ln(1-pi_j(j,pi[1],rhow[1]));
                   end;
             writeln(file1,'THE RATIO ESTIMATE',sum);
           end;
        end;
     writeln('Want to run again If yes__1 else 2 ');
     readln(ex);
     if ex = 2 then
       close_ok := true;
   end;
close(file1);
end.
```

APPENDIX 2

A PROGRAM FOR APPROXIMATION OF MLE OF THE REPARAMETERIZED MODEL

```
uses crt;
type
   vectorNAB = array [0..5] of integer;
   vectorp =   array [0..5] of real;
   matrix   =   array [1..2,1..2] of real;
var
   t,i: integer;
   A,B,N : vectorNAB;
   Dp,F,G,Fp,R,Fr,Gp,Gr,P:vectorp;

det,chi_square,c,e,efr,efp,egp,egr,ratio,sum2,sum3:real;
   filevar :text;
      procedure inputdata;
         begin
           writeln('Enter the values of N and A');
           for i := 1 to 5 do read(N[i]);
           for i := 1 to 5 do read(A[i]);
           for i := 1 to 5 do B[i] := N[i] - A[i];
         end;

       function power(r:real; i:integer):real;
        var j:integer;
            product:real;
            begin
               product := 1.0;
               if i = 0 then product := 1.0
               else if i > 0 then for j:= 1 to i do product :=
product*r                    else
                   begin
                     i:= -i;
                     for j:= 1 to i do product := product*r;
                  product := 1/product;
                   end;
                 power:=product;
                end;

      function calculatef(PP,RR:real):real;
         var pr1,pr2:real;
        begin
           F[t-1] := 0.0;
            for i := 1 to 5 do
               begin
                  pr1 :=B[i]/exp(i*RR);
                  pr2 := 1-PP/exp(i*RR);
                  F[t-1] := F[t-1]+ (A[i]/PP - pr1/pr2);
               end;
            calculatef := F[t-1];
          end;

   function calculateg(PP,RR:real):real;
```

```
var pr:real;
begin
   G[t-1] := 0.0;
   for i := 1 to 5 do
      begin
        pr:= PP/exp(i*RR);
        G[t-1] := G[t-1] - i*A[i] +
                     (i*B[i]*pr/(1-pr));
       calculateg := G[t-1];
     end;
    end;
function calculatefp(PP,RR:real):real;
  var pr1,pr2:real;
  begin
     Fp[t-1] := 0.0;
     for i := 1 to 5 do
        begin
           pr1:=B[i]/exp(2*i*RR);
           pr2:=1-PP/exp(i*RR);
           Fp[t-1] := Fp[t-1] - (A[i]/(PP*PP) +
                                   pr1/(pr2*pr2));
            end;
       calculatefp:= Fp[t-1];
     end;
  function calculatefr(PP,RR:real):real;
      var pr1,pr2:real;
      begin
      Fr[t-1] := 0.0;
      for i := 1 to 5 do
         begin
            pr1 := i*B[i]/exp(i*RR);
            pr2 := 1-PP/exp(i*RR);
            Fr[t-1] := Fr[t-1] + pr1/(pr2*pr2)
         end;
        calculatefr:= Fr[t-1];
     end;
 function calculategr(PP,RR:real):real;
    var pr1,pr2,pr3:real;
    begin
       Gr[t-1] := 0.0;
       for i := 1 to 5 do
          begin
             pr1 := i*i*B[i]*PP/exp(i*RR);
             pr2 := 1-PP/exp(i*RR);
             Gr[t-1] := Gr[t-1]+pr1/(pr2*pr2);
           end;
         calculategr := -Gr[t-1];
       end;
   function calculategp(PP,RR:real):real;
       var pr1,pr2:real;
        begin
          Gp[t-1] := 0.0;
          for i := 1 to 5 do
            begin
              pr1 := i*B[i]/exp(i*RR);
              pr2 := 1-PP/exp(i*RR);
```

```
            Gp[t-1] := Gp[t-1]+pr1/(pr2*pr2);
           end;
           calculategp := Gp[t-1];
          end;

        procedure finalloop;
          begin
            P[0] := A[1]/N[1];
            R[0] := ln(P[0]) - ln(A[2]/N[2]);
            t:= 0;
               repeat
                  t:= t+1;
                  F[t-1] := calculatef(P[t-1],R[t-1]);
               G[t-1] := calculateg(P[t-1],R[t-1]);
            Fr[t-1] :=calculatefr(P[t-1],R[t-1]);
          Fp[t-1] :=calculatefp(P[t-1],R[t-1]);
        Gr[t-1] :=calculategr(P[t-1],R[t-1]);
      Gp[t-1] :=calculategp(P[t-1],R[t-1]);
    Dp[t-1] :=1/(Fp[t-1]*Gr[t-1] - Fr[t-1]*Gp[t-1]);
             P[t]:=    P[t-1]-Dp[t-1]*(Gr[t-1]*F[t-1]-
Fr[t-1]*G[t-1]);                                    R[t]:=
R[t-1]+Dp[t-1]*(Gp[t-1]*F[t-1]-Fp[t-1]*G[t-1]);
       until abs(P[t]-P[t-1]) + abs(R[t]-R[t-1]) <
                            0.000000001;
             F[t]:= calculatef(P[t],R[t]);
                G[t] := calculateg(P[t], R[t]);
               end;
          begin {main program}
            inputdata;
            clrscr;
            finalloop;
            assign(filevar,'out');
            rewrite(filevar);
            writeln(filevar,'THE RESULTS OF ITERATION
                             P, R, F AND G ');
             writeln(filevar,'-----------------');
       writeln(filevar,P[t],  '  ',  R[t],  '   ' ,
                               F[t],  '  ', G[t]);
        writeln;
            Fp[t] :=calculatefp(P[t],R[t]);
            Gr[t] :=calculategr(P[t],R[t]);
            Gp[t] :=calculategp(P[t],R[t]);
            Fr[t] := calculatefr(P[t],R[t]);
            efp := 0.0;
            for i := 1 to 5 do
              begin
                sum2 :=(N[i]*exp(-r[t]*i))/p[t];
                sum3 := (N[i]*exp(-2*r[t]*i))/(1-
                         p[t]*exp(-r[t]*i));

                 efp:= sum3+sum2;
              end;
            efr:= 0.0; egp := 0.0;egr := 0.0;
            for i := 1 to 5 do
                 efr:= efr + (i*N[i]*exp(-r[t]*i))/(1-
                                  p[t]*exp(-r[t]*i));
```

```
                  for i := 1 to 5 do
                      egp:= egp + (i*N[i]*exp(-r[t]*i))/(1-
                                    p[t]*exp(-r[t]*i));
                  for i := 1 to 5 do
                     egr := egr + (i*i*p[t]*exp(-
                          r[t]*i))/(1-p[t]*exp(-r[t]*i));
                  det :=egr*efp -egp*efr;
                 writeln(filevar);
                 writeln(filevar);
                 writeln(filevar,'THE INFORMATION MATRIX
                                            INVERSE ');

           WRITELN(filevar);
                 writeln(filevar,-egr/det,  '    ',
                                  efr  /det);
        writeln(filevar,egp/det ,  '    ',   -efp/det);
          writeln(filevar);
                 writeln(filevar,'THE ESTIMATED VALUE');
                     for i := 1 to 5 do
writeln(filevar,P[t]/exp(i*R[t]));
                  chi_square := 0.0;
                 for I := 1 to 5 do
                    begin
                      c:= P[t]/exp(i*R[t]);
                      e:= A[i]/N[i] - c;
                      {chi_square := chi_square + (e*e)/c;}
                   end;
                  ratio:= 0.0;
                  for i := 1 to 5 do
                     begin
                       sum2 := ln(p[t])-(r[t]*i);
                       sum3 := ln(1-p[t]*exp(-r[t]*i));
                 ratio := ratio + A[i]*sum2 + B[i]*sum3;
                end;
                  writeln(filevar);
                 writeln(filevar,'CHI-SQUARE  GOODNESS-OF-FIT
                   =', ' ',chi_square);
            writeln(filevar);
                 writeln(filevar,'The likelyhood ratio = ',
                                 ratio );
          close(filevar);
              end.
```

APPENDIX 3

A PROGRAM FOR LR TEST OF GUTTMAN AND OLKIN MODEL

```
type Matrix = array[1..5,1..5] of integer;
     Vector = array [1..5] of real;
var m,p,ratio,sum2,sum3:real;
    i,j: integer;
    A,B: Matrix;
    D,F,PH: Vector;
    datf: text;
function power (p:real;i:integer):real;
    var t:integer;
        pp:real;
    begin
     pp:= 1.0;
     if i = 0 then pp:= 1.0
     else if i > 0 then for t:= 1 to i do pp:= pp * p
     else
       begin
         i:= -i;
         for t:= 1 to i do pp:= pp * p;
         pp:= 1/pp;
       end;
       power:= pp;
     end;
 function ff (PH:vector; p:real; J:integer):real;
   var sum1, sum2, sum3: real;
   begin
     sum1:= 0.0;
     for i:= 1 to 5 do
         begin
           sum2:= B[i,j]*power(p,i-1);
           sum3:= 1-PH[j]*power(p,i-1);
           sum1:= sum1 + A[i,j]/PH[j]-sum2/sum3;
         end;
         ff:=sum1;
     end;
   Procedure calculateF5(var F:vector);
     begin
       for j:=1 to 5 do F[j]:= ff(PH,p,j);
     end;
   function DD(PH:vector;p:real;j:integer):real;
     var sum1,sum2,sum3,sum4:real;
     begin
       sum1:= 0.0;
       for i:= 1 to 5 do
         begin
           sum2:= B[i,j]*power(p,2*i-2);
           sum3:= 1-PH[j]*power(p,i-1);
           sum4:= sum2/(sum3*sum3);
           sum1:= sum1+(A[i,j]/(PH[j]*PH[j]))+sum4;
         end;
         DD:= -sum1;
```

```
      end;
   Procedure calculateD5(var D:vector);
      begin
        for j:= 1 to 5 do D[j]:= DD(PH,p,j);
      end;
   function sumf:real;
      var temp2,sum1,sum2:real;
      begin
        temp2:= 0.0;
        for i:= 1 to 5 do
          for j:= 1 to 5 do
            begin
              sum1:= PH[j]*power(p,i-1);
              sum2:= ((i-1)*B[i,j]*sum1)/(1-sum1);
              temp2:= temp2 + ((i-1)*A[i,j]-sum2);
            end;
            temp2:= temp2/p;
            sumf:=temp2;
          end;

      function sumf1:real;
        var temp1,sum1,sum2:real;
        begin
          temp1:=0.0;
          for i:= 1 to 5 do
            for j:= 1 to 5 do
              begin
                sum1:=PH[j]*power(p,i-2);
                sum2:=1-PH[j]*power(p,i-1);
                                                      t e m p 1 : =
temp1+((i-1)*(i-1)*B[i,j]*sum1)/(sum2*sum2);                       end;
              sumf1:=-temp1;
            end;

        procedure calculatePHj(var PH:vector);
          begin
            for j:=1 to 5 do PH[j]:= PH[j]-F[j]/D[j];
          end;
        procedure inputdata;
          begin
            assign(datf,'c:datf.pas');
            reset(datf);
            for i:= 1 to 5 do
              begin
               for j:= 1 to 5 do
                 read(datf,A[i,j]);
               readln(datf);
              end;
           readln(datf);
           for i:=1 to 5 do
            begin
              for j:= 1 to 5 do
                read(datf,B[i,j]);
              readln(datf);
            end;
            readln(datf);
```

```
          close(datf);
        end;

begin{main program}
  Inputdata;
  p:=0.66;
  for j:= 1  to 5 do PH[j]:=A[1,j]/(A[1,j]+B[1,j]);
  repeat
    write(m);
    calculateF5(F);
    calculateD5(D);
    calculatePHj(PH);
    p:=p - sumf/sumf1;
                                              m          :          =
(abs(F[1])+abs(F[2])+abs(F[3])+abs(F[4])+abs(F[5])+abs(sumf));
until m < 0.5*power(10,-7);
  writln;
  writeln('The final fvalue of p is:');
  writeln;
  write ('p = ',p);
  writeln;
  writeln('The values of the vector PH[j] are:');
  for j := 1 to 5 do writeln (PH[j]);
  ratio := 0.0;
  for i := 1 to 5 do
    for j := 1 to 5 do
       begin
         sum2 := ln(PH[j]*power(p,i-1));
         sum3 := ln(1-PH[j]*power(p,i-1));
         ratio := A[i,j]*sum2+B[i,j]*sum3;
     end;
  writeln('The likelyhood ratio = ', ratio );
end.
```

APPENDIX 4

A PROGRAM FOR LR TEST OF THE REPARAMETERIZED MODEL

```
type Matrix = array[1..5,1..5] of integer;
     Vector = array [1..5] of real;
var m,r,ratio,sum2,sum3:real;
    i,j: integer;
    A,B: Matrix;
    D,F,PH: Vector;
    datf: text;
function power (p:real;i:integer):real;
    var t:integer;
        pp:real;
    begin
     pp:= 1.0;
     if i = 0 then pp:= 1.0
     else if i > 0 then for t:= 1 to i do pp:= pp * p
     else
       begin
         i:= -i;
         for t:= 1 to i do pp:= pp * p;
         pp:= 1/pp;
       end;
       power:= pp;
     end;
 function ff (PH:vector; p:real; J:integer):real;
   var sum1, sum2, sum3: real;
   begin
     sum1:= 0.0;
     for i:= 1 to 5 do
         begin
           sum2:= B[i,j]*exp(-r*i);
           sum3:= 1-PH[j]*exp(-r*i);
           sum1:= sum1 + A[i,j]/PH[j]-sum2/sum3;
         end;
         ff:=sum1;
     end;
   Procedure calculateF5(var F:vector);
     begin
       for j:=1 to 5 do F[j]:= ff(PH,r,j);
     end;
   function DD(PH:vector;r:real;j:integer):real;
     var sum1,sum2,sum3,sum4:real;
     begin
       sum1:= 0.0;
       for i:= 1 to 5 do
         begin
           sum2:= B[i,j]*exp(-r*i*2);
           sum3:= 1-PH[j]*exp(-r*i);
           sum4:= sum2/(sum3*sum3);
           sum1:= sum1+(A[i,j]/(PH[j]*PH[j]))+sum4;
        end;
```

```
      DD:= -sum1;
    end;
Procedure calculateD5(var D:vector);
    begin
      for j:= 1 to 5 do D[j]:= DD(PH,r,j);
    end;
function sumf:real;
    var temp2,sum1,sum2:real;
    begin
      temp2:= 0.0;
      for i:= 1 to 5 do
        for j:= 1 to 5 do
          begin
            sum1:= PH[j]*exp(-r*i);
            sum2:= (i*B[i,j]*sum1)/(1-sum1);
            temp2:= temp2 + (i*A[i,j]-sum2);
          end;
          sumf:=-temp2;
        end;

    function sumf1:real;
      var temp1,sum1,sum2:real;
      begin
        temp1:=0.0;
        for i:= 1 to 5 do
          for j:= 1 to 5 do
            begin
              sum1:=PH[j]*exp(-r*i);
              sum2:=1-PH[j]*exp(-r*i);
              temp1:=
           temp1+((i*i)*B[i,j]*sum1)/(sum2*sum2);
             end;
            sumf1:=-temp1;
          end;

    procedure calculatePHj(var PH:vector);
        begin
          for j:=1 to 5 do PH[j]:= PH[j]-F[j]/D[j];
        end;
      procedure inputdata;
        begin
          assign(datf,'c:datf.pas');
          reset(datf);
          for i:= 1 to 5 do
            begin
             for j:= 1 to 5 do
               read(datf,A[i,j]);
             readln(datf);
            end;
         readln(datf);
         for i:=1 to 5 do
          begin
            for j:= 1 to 5 do
              read(datf,B[i,j]);
            readln(datf);
          end;
```

```
            readln(datf);
          close(datf);
        end;

begin{main program}
  Inputdata;
  r:=0.37;
  for j:= 1  to 5 do PH[j]:=A[1,j]/(A[1,j]+B[1,j]);
  repeat
    write(m);
    calculateF5(F);
    calculateD5(D);
    calculatePHj(PH);
    r:=r - sumf/sumf1;
                                                 m          :          =
(abs(F[1])+abs(F[2])+abs(F[3])+abs(F[4])+abs(F[5])+abs(sumf));
until m < 0.5*power(10,-7);
  writeln('The final falue of p is:');
  writeln;
  write ('r = ',r);
  writeln;
  writeln('The values of the vector PH[j] are:');
  for j := 1 to 5 do writeln (PH[j]);
  ratio:=0.0;
  for i := 1 to 5 do
    for j := 1 to 5 do
      begin
        sum2 := ln(PH[j]*exp(-r*i));
        sum3 := ln(1-PH[j]*exp(-r*i));
        ratio := A[i,j]*sum2+B[i,j]*sum3;
      end;
  writeln;
  writeln('The likelyhood ratio = ', ratio );
end.
```

APPENDIX 5

ORIGINAL DATA

1 BY BATCH

BATCH	SEMESTER 1	2	3	4	5
1	755	576	501	473	451
	222	109	47	29	17
2	714	625	580	542	498
	141	79	50	44	40
3	628	517	458	435	430
	146	70	31	23	21
4	511	450	417	391	385
	82	52	41	14	12
5	451	455	412	383	359
	61	54	38	24	6

Note: The first row within a batch shows the total number of students enrolled and the second shows attrition.

2 BY DEPARTMENT (the rate)

AC.YEAR	BIO	CHEM	GEOL	MATH	PHY	STAT
1982/83	.2393	.2431	.1401	.2384	.1866	.1898
1983/84	.2078	.1899	.1339	.1253	.1073	.1201
1984/85	.1080	.1495	.0363	.1270	.0950	.0620
1985/86	.1667	.1750	.1447	.1431	.1584	.0679
1986/87	.0870	.1300	.1023	.0782	.0964	.1071
1987/88	.1045	.1236	.0478	.0956	.0784	.0899
1988/89	.0735	.0923	.0646	.1063	.0744	.0772
1989/90	.0670	.1266	.0766	.0891	.0895	.0512