# PREDICTION OF PROBATIONARY FRESHMEN AND EFFECTIVENESS OF A SUPPLEMENTARY COUNSELING PROGRAM AT DRAKE UNIVERSITY

by

Eli Alan Zubay

A Dissertation Submitted to the
Graduate Faculty in Partial Fulfillment of
The Requirements for the Degree of
DOCTOR OF PHILOSOPHY

Major Subject: Education

Approved:

Signature was redacted for privacy.

In Charge of Major Work

Signature was redacted for privacy.

Head of Major Department

Signature was redacted for privacy.

Dean of Graduate College

Iowa State College

1950

TABLE OF CONTENTS

# LIST OF TABLES

# I. INTRODUCTION

During the last quarter of a century, there has been a remarkably increased emphasis upon student counseling in higher education. From informal counseling carried on as an extracurricular activity by members of the faculty primarily interested in students, the trend has been in the direction of an ever expanding separate faculty group whose sole academic mission has been student personnel problems. The teaching faculty has not been relieved of student counseling but rather the responsibility of the faculty has increased although now this counseling is generally directed by policy and general advice of a centralized personnel organization.

Perhaps no problem is more acute for college administrators, personnel officials, and other faculty than that of student mortality particularly at the freshman level.

The causes of such mortality are numerous but perhaps the one of greatest concern is poor achievement. Students whose achievement is low, either from lack of ability, lack of proper motivation, or any number of other causes, find themselves placed on probation. Such administrative action usually does not occur prior to midquarter or midsemester at which time remedial measures are begun. By this time much damage to the academic progress of the student has already occurred.

Although the problem of mortality is not confined to any one

college or university, the present study is limited to Drake University.

The problem of freshman mortality was evident at Drake University. Bixler[1] reported that of 634 freshman students enrolled in the fall of 1947 only 390 were enrolled the following fall, representing a loss of 244 students or 38.5% of the original enrollment. Of the losses, 102, or 41.4% were due to scholastic difficulties. For the two school years, 1948-1950, approximately 21% of the freshmen were placed on probation.

The present study has been designed to anticipate probation at the time of entrance in order that remedial counseling techniques may be immediately applied. The question arises as to whether an effective predictive scheme could be devised using the data available in the Admission's Office and the Testing Bureau of Drake University.

If a predictive method can be found, then the problem consists of evaluating the effectiveness of early supplemental counseling. The usual experimental and control groups were established to evaluate the effectiveness of a supplemented counseling program. Since at Drake University the major portion of the counseling is carried on by the faculty, the supplemented counseling is also a function of the faculty. It was not the aim of this study to devise

---

[1]Bixler, Roy W. and Ady, Carolyn R. A study of the freshman class of 1947-1948. Unpublished mimeographed manuscript. Registrar's Office, Drake University, Des Moines, Iowa. 1949.

a counseling procedure for which the university director of testing and counseling and his staff would be responsible, as it would not be administratively feasible for them to execute such a program. The number of students involved would be too great since the counseling center also has responsibilities to the better students. In addition, it would seem that the function of the director and his staff should be of a specialized consultative nature.

The purpose of this study is two-fold--the prediction of probably probation and evaluation of a supplemented counseling program.

## II. REVIEW OF LITERATURE

As the principal statistical technique employed in this study was the discriminant function, the review of literature contains a brief, historical development of this technique. Applications of this function to psychological and educational problems are also cited.

The first published application of the discriminant function was made by Barnard[1] in 1935. Following a suggestion from Fisher, the technique was applied in a work on craniometry. Fisher's[2] first article on the discriminant function appeared in 1936. He gave a further example of the technique as it was used in a taxonomic problem involving species of iris.

The general problem of the discriminant function was closely related to a problem with which Mahalanobis[3] in 1927 and Hotelling[4]

---

[1]Barnard, M. M. The secular variations of skull characters in four series of Egyptian skulls. Annals of Eugenics. 6:352-371. 1935.

[2]Fisher, R. A. The use of multiple measurements in taxonomic problems. Annals of Eugenics. 7:179-188. 1936.

[3]Mahalanobis, P. C. Analysis of race mixture in Bengal. Jour. Asiat. Soc. Beng. 23:301-333. 1927.

[4]Hotelling, Harold. Generalization of "Student's" Ratio. Annals of Mathematical Statistics. 2:360-378. 1931.

in 1931 were concerned. In fact, in a later article Fisher[1] showed the relationship between his discriminant function approach and the latter two methods of attack. Another point of importance in his article was the demonstration that the coefficients for the discriminant function were proportional to the coefficients of a multiple regression equation involving the same independent variables. The implication of this relationship was that either "within" deviations or total deviations could be used in solving the simpler type of discrimination problem. Fisher also pointed out that the appropriate test for the significance of the discriminant function had actually been given by Hotelling as early as 1931.

There was a slight error in the formula proposed by Fisher[2] as a test of significance of a discriminant function which he corrected in a later article in 1940. The article indicated how Hotelling's $T^2$ was utilized in the test.

Most of the early applications of the discriminant function were in the field of eugenics. However, in 1938 Travers[3] tried to

---

[1]Fisher, R. A. The statistical utilization of multiple measurements. Annals of Eugenics. 8:376-386. 1938.

[2]Fisher, R. A. The precision of discriminant functions. Annals of Eugenics. 10:422-429. 1940

[3]Travers, R. M. W. The use of a discriminant function in the treatment of psychological group differences. Psychometrika. 4:25-32. 1938.

point out the importance of the discriminant function to type psychology. As an illustration, he developed a discriminant function based on twenty successful engineer apprentices and twenty successful air pilots. During the same time this function was used to classify individuals either as engineers or pilots, Wallace and Travers[1] were making an application to a problem in the business world. They studied groups of successful and unsuccessful salesmen for the purpose of devising a method for the successful selection of salesmen. The discriminant function was used to obtain weights for the various measures combined into the selective scheme. However, in spite of the early works of Travers, the discriminant function was not widely used by psychologists; for in 1943 Garrett[2] again stressed the value of the technique to psychologists. He thought that little use was made of the discriminant function, probably because no presentation of the function had been made which could be understood by investigators with little mathematical background. He attempted to give such a treatment, one with less emphasis on the mathematical approach.

One of the earliest educational studies noted utilizing the

---

[1]Wallace, Noel and Travers, R. M. W. A psychometric sociological study of a group of specialty salesmen. Annals of Eugenics. 8:266-302. 1938.

[2]Garrett, Henry E. The discriminant function and its use in psychology. Psychometrika. 8:65-79. 1943.

discriminant function was that of Selover[1]. He was concerned with aiding the students in a proper selection of their major field. Up to this time, guidance had been performed by the method of profile analysis. Selover felt that if the classification of a student could be based upon objective criteria rather than upon subjective judgments, the selection of a major field would be more reliable. His initial study indicated that the discriminant function was valuable in determining objective criteria.

In his text, Johnson[2] illustrates the use of the discriminant function in an educational situation. The objective was to distinguish between the students enrolled in two sections of physics, one section being more advanced than the other. At Iowa State College, Wert[3], MacRae[4], and Herdman[5], made use of the discriminant

---

[1]Selover, Robert B. A study of the sophomore testing program at the University of Minnesota. Journal of Applied Psychology. 26:456-467. 1942.

[2]Johnson, Palmer O. Statistical Methods in Research. New York. Prentice-Hall, Inc. 1949.

[3]Wert, James E. and others. The discriminant function and multiple biserial R. Unpublished memeographed manuscript. The Educational Research Laboratories, 315 Curtiss Hall, Iowa State College, Ames, Iowa. 1949.

[4]MacRae, John Murray. Usefulness of the Minnesota Personality Scale for predicting achievement of freshman engineering students. Unpublished M. S. Thesis. Ames, Iowa, Iowa State College Library. 1949.

[5]Herdman, Raymond Wain. Predicting pupil mortality among high school boys. Unpublished M. S. Thesis, Ames, Iowa, Iowa State College Library. 1949.

function in obtaining a multiple biserial R.

At Michigan State College, Jackson[1] reported on the selection of students for admission to the beginning chemistry course. Prior to their entering the course, remedial work was given to those students failing to meet admission standards. Minimum total scores on a chemistry pre-test and the A.C.E. were tried but found to be unsatisfactory for such selection. A discriminant function proved to be much more effective.

The concept of the discriminant function is no longer confined to two groups. Articles by Day and Sandomire[2] and by Brown[3] indicate the use of more than two groups. In addition, if a large number of independent variables were to be used, Jackson[4] and later Beall[5] suggest approximate methods for obtaining the coefficients for a discriminant function.

---

[1]Jackson, Robert. The selections of students for freshman chemistry by means of discriminant functions. Journal of Experimental Education. 18:209-215. 1950.

[2]Day, Besse B. and Sandomire, Marion M. Use of the discriminant function for more than two groups. Journal of the American Statistical Association. 37:461-472. 1942.

[3]Brown, George W. Discriminant function. Annals of Mathematical Statistics. 18:514-528. 1947.

[4]Jackson, Robert. Approximate multiple regression weights. Journal of Experimental Education. 11:221-225. 1943.

[5]Beall, Geoffrey. Approximate methods in calculating discriminant functions. Psychometrika. 10:205-217. 1945.

The foregoing remarks indicate that the discriminant function is a relatively new technique; and, although its value was recognized by some, it was not used extensively until the past few years.

Mortality studies indicating the number of freshmen drop-outs in various universities and colleges as well as those studies indicating the role that scholastic difficulties played in student withdrawals were also reviewed.

At Michigan State College[1] a study was made of the freshmen students over a three-year period of time, 1937-1940. Of 4,007 freshmen enrolled during this three-year period, 1,389 or 35 per cent withdrew from the college during their first year. Students were classified according to their reasons for withdrawing. Those who withdrew for low grades constituted, by far, the largest group. For the three-year period, the withdrawals for low grades constituted 39 per cent of the entire group.

The most comprehensive general study of student mortality was made by McNeeley[2] in 1937. His report included 25 representative universities throughout the United States. He reported a nation-wide freshmen mortality of 34 per cent and a total mortality of

---

[1]Mitchell, Fred T. Why freshmen leave college. Journal of Higher Education. 13:95-100. 1942.

[2]McNeeley, John H. College student mortality. U. S. Department of the Interior, Office of Education, Bulletin 1937, No. 11, Washington, D.C. U. S. Government Printing Office. 1938.

62 per cent. For individual universities the freshmen losses ranged from 22 per cent to 49 per cent. Approximately 22 per cent of the students from all classes left the universities because of failure in their class work. This percentage is lower than the 38 per cent reported by Michigan State College; however, a greater percentage of freshmen would suggest more scholastic difficulties.

Both the general survey and the specific college study indicate that 34 per cent of the members of a freshman class would not be enrolled the following year. Of this group, over one-third of them withdrew because of poor grades.

Because this study was concerned with students placed on probation and this status is closely related to college achievement, several surveys concerned with the prediction of college achievement are included in this review of literature.

Garrett[1] made a survey of studies of scholastic success in colleges of arts and science and teacher colleges. He found that the factors which yielded the highest correlations with college achievement were high school average and high school rank. The median correlation between high school rank and college achievement was 0.56 whereas that for high school marks and college achievement was 0.55.

---

[1]Garrett, H. E. Review and interpretation of investigations of factors related to scholastic success in colleges of arts and science and teachers' colleges. Journal of Experimental Education 12:184-186. 1944.

Achievement tests, intelligence tests, and aptitude tests gave median correlations with college achievement of 0.49, 0.47, and 0.43, respectively. Character and personality tests did not seem to be too beneficial as the median coefficient of correlation was found to be 0.09 although the correlation did vary from -0.30 to 0.64 in the 63 studies included. Using combinations of these variables yielded a median multiple coefficient of 0.58 for two variables and 0.61 for three variables.

Another comprehensive survey of college achievement studies was that made by Durflinger[1]. The conclusions reached were practically identical with those attained by Garrett. Correlations between college scholarship and high school scholarship varied from 0.50 to 0.60 with a median of approximately 0.55. The summaries indicated that an intelligence test, a good achievement test, and some measure of high school achievement usually yielded the highest multiple r's. The multiple correlation coefficients rarely exceeded 0.80 regardless of the number and type of variables used. The median multiple r's were found to be between 0.60 and 0.70. The survey also indicated that of all of the personality traits, interest yielded a higher and more stable relationship than any other trait.

---

[1]Durflinger, Glenn W. The prediction of college success - a summary of recent finding. Journal of the American Association of Collegiate Registrars. 19:68-78. 1943.

Review of studies of specific tests does not reveal results contrary to the conclusions drawn from the foregoing general surveys. It would seem that any predictive scheme for college achievement should contain at least some measure of high school scholarship and intelligence test score.

The purpose of this review of literature has been to give some indication of the proportion of freshman losses in universities, the role that scholastic difficulties play in these withdrawals, the effectiveness of available measures in predicting college achievement, and the development and applications made of the discriminant function.

## III. METHOD OF PROCEDURE

Freshman students at Drake University for the two year period 1948-1950 were used in this study. Complete data were available for 666 freshmen from the 1948 class and for 476 first semester freshmen of the 1949 class from four of the colleges, College of Liberal Arts, College of Commerce and Finance, College of Fine Arts, and College of Education.

The data available for each of the students were as follows: test scores on the American Council on Education Psychological Examination, Nelson-Denny Reading Test, the Kuder Preference Record, and the high school percentile rank. These tests were given during the week of registration preceding each semester. In addition, the American Institute of Accountants' Orientation Test which was given during the second week of each semester was used for the freshman students enrolled in the College of Commerce and Finance.

In the statistical treatment of the data the two parts of the American Council on Education Psychological Examination, the Quantitative score, Q, and the Linguistic score, L, were used as two of the variables. They were designated as $X_1$ and $X_2$, respectively. The two parts of the Nelson-Denny Reading Test, the Vocabulary, V, and the Paragraph Reading Comprehension, P, were also used separately. They were designated as $X_4$ and $X_5$, respectively. Six of the nine Interest scores available from the Kuder Preference Record were

used as variables. They were: Computational Interest, Persuasive Interest, Artistic Interest, Literary Interest, Social Service Interest, and Clerical Interest. In the statistical treatment they were designated as $X_8$, $X_9$, $X_{10}$, $X_{11}$, $X_{12}$, and $X_{13}$ respectively.

Of the 666 students for whom the data were complete, 234 were enrolled in the College of Liberal Arts, 187 in the College of Commerce and Finance, 117 in the College of Education, and 128 in the College of Fine Arts.

One of the objectives of the study was to predict which of the 1949 freshmen might be placed on probation. Since this means forming the dichotomy consisting of probation and non-probation students, the discriminant function was used. In Liberal Arts 64 freshman students were placed on probation during 1948-1949. To obtain data for determining the discriminant function applicable to Liberal Arts, an equal number of non-probation freshman students also were selected at random from the remaining 170 Liberal Arts freshmen. The same procedure was followed in the College of Commerce and Finance in which 42 freshman students had been placed on probation during 1948-1949. However, since only 25 freshman Education students and 12 freshman Fine Arts students had been placed on probation during 1948-1949, it was felt that the number of students was insufficient to determine discriminant functions for those colleges based solely on data from students in those colleges. Therefore an all-university discriminant function was developed based on the 143 freshman students who were placed on

probation in the four colleges during 1948-1949 and an equal number of non-probation students selected from those colleges. Of the 143 non-probation students, the number of students selected at random from each of the colleges was equal to the number of probation students from the respective colleges. This all-university discriminant function was used for predicting probation students in the Colleges of Fine Arts and Education.

It was assumed that in the various colleges the same proportion of freshman students would be placed on probation in 1949-1950 as were placed on probation in 1948-1949. Under this assumption and through the use of the discriminant functions developed, a list of potential, probationary students was developed for each of the colleges. This was done in the fall of 1949 as soon as the test scores were available. The group of 143 freshman students included 50 from Liberal Arts, 36 from Commerce, 14 from Education, and 14 from Fine Arts. In each college students were selected at random to form two groups. One group from each college was designated as the experimental group and the other as the control group. None of the students was aware of the fact that he had been placed on a potential probation list. The students in the control group went through the normal counseling procedures. Their faculty advisers were not aware of their classification as potential, probationary students.

The faculty counselors of the students in the experimental group were informed of the study. They knew that their advisees were pre-

dicted probationary students. The Mooney Problem Check List was filled out by each student and was available to each counselor before his first contact with the student. Meetings of these counselors and the university director of counseling and testing were held throughout the year. At these meetings the various guidance and testing facilities available to the student were explained and adequate use of them was stressed. The meetings also offered an opportunity for the counselors to discuss problems which arose in working with these students.

At the end of the 1949-1950 school year the following data were collected on all predicted, probationary students: (1) average marks earned for the first semester; (2) average marks earned for the school year; (3) whether or not the student had been placed on probation at the end of either the first or second semester; (4) whether or not the student had withdrawn from school sometime during the school year, and (5) whether or not the student had been dropped for low grades at the end of the school year by his college dean.

A second objective of this study was to determine whether the counseling given the experimental group was more effective than that of the control group. For purposes of this study, college achievement was defined in five ways: (1) as average marks for the first semester, (2) as average marks for the school year, (3) as successful-unsuccessful at the end of the first semester, and (4) as successful-unsuccessful at the end of the school year, and (5) survivor-mortality status at the end of the school year.

Thus, five criteria were available for the evaluation of the effectiveness of the counseling. For the first two criteria the test of significance used was the test of the difference between two means assuming a common variance. The last three criteria were tested for significant differences among the experimental and control groups by chi square.

## IV. PREDICTION OF PROBATION

For purposes of this study, probation was defined as the status of any student placed on the college freshman probation list by his dean. A student was usually placed on probation if his grade point average was below 2.00. The grade point average was the ratio of the number of grade points earned and the number of credit hours carried. The number of grade points earned was determined on the basis of A = 6, B+ = 5, B = 4, C+ = 3, C = 2, D+ = 1, D = 0, and F = -2 grade points per credit hour.

One objective of the study was to determine whether or not the battery of tests given new freshman students was effective in predicting which of them would be placed on probation during their first year at Drake University. The probation and non-probation group from the 1948 class were compared in achievement. Table 1 shows that there were significant differences between the two groups with respect to the scores on the American Council on Education Psychological Examination, the Nelson-Denny Reading Test, and high school achievement. For the Commerce and Finance students, the difference between the two groups was also significant on the American Institute of Accountant's Orientation Test. These results were not surprising in view of the apparent relationship between college achievement and probation status. It should be noted that none of the interest scores on the Kuder Preference Record was significantly different between the two groups.

Table 1

Mean Scores of Probation and Non-Probation Freshmen

| Variable | Symbol | Commerce and Finance (N = 84) | | | Liberal Arts (N = 128) | | | All University (N = 286) | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | Prob. | Non-prob. | t | Prob. | Non-prob. | t | Prob. | Non-prob. | t |
| Quantitative | $X_1$ | 41.40 | 46.55 | 2.019* | 39.14 | 43.83 | 2.664** | 38.13 | 43.16 | 3.898** |
| Linguistic | $X_2$ | 58.83 | 68.10 | 3.291** | 60.34 | 67.63 | 3.018** | 57.87 | 66.50 | 6.204** |
| Vocabulary | $X_4$ | 31.62 | 38.17 | 2.496* | 30.69 | 38.36 | 3.912** | 30.58 | 38.10 | 5.535** |
| Para.Read.Comp. | $X_5$ | 39.26 | 46.38 | 3.282** | 39.41 | 45.09 | 3.141** | 38.24 | 44.81 | 5.407** |
| High school | $X_7$ | 39.88 | 66.21 | 5.319** | 38.45 | 66.72 | 6.672** | 41.45 | 66.05 | 8.951** |
| Computational | $X_8$ | 42.90 | 42.12 | .264 | 33.45 | 30.09 | 1.515 | 34.63 | 33.27 | .929 |
| Persuasive | $X_9$ | 87.55 | 90.79 | .829 | 75.86 | 76.53 | .208 | 78.31 | 78.88 | .268 |
| Artistic | $X_{10}$ | 43.17 | 42.79 | .128 | 46.16 | 49.02 | 1.141 | 46.87 | 47.24 | .221 |
| Literary | $X_{11}$ | 50.57 | 50.33 | .069 | 53.22 | 54.91 | .524 | 51.84 | 53.18 | .670 |
| Social Service | $X_{12}$ | 60.60 | 62.12 | .369 | 70.11 | 71.42 | .427 | 69.67 | 71.31 | .756 |
| Clerical | $X_{13}$ | 67.29 | 64.14 | 1.093 | 55.61 | 55.16 | .155 | 58.08 | 57.15 | .545 |
| Orientation | $X_{14}$ | 41.31 | 51.57 | 2.616* | | | | | | |

* significant at the 5% level

** significant at the 1% level

More pertinent than the tests of significance was some measure of the ability of the various variables to predict the probation tendency. For this purpose biserial r was computed for each of the variables for which the "t" test indicated significant differences.

In making these computations it was assumed that the probation tendency was a variable which was normally distributed and existed only in the dichotomy of probation and non-probation. It was further assumed that there was a linear rather than a curvilinear relationship between probation tendency and the various test scores used. Biserial r was computed from the formula

$$r_{bis} = \frac{(\bar{X}_A - \bar{X}_B)}{\sigma} \frac{pq}{z}$$

where $\bar{X}_A$ and $\bar{X}_B$ are the means of the variable X in the dichotomous categories, $\sigma$ is the standard deviation in the total sample, p and q are the proportion of the cases in the probation and non-probation groups, and z is the height of the ordinate in a normal curve of unit area dividing the area into the proportions p and q.

As indicated in Table 2, the highest biserial correlation existed between high school achievement and the probation tendency. The correlations were comparable to what is found in college achievement studies.

To improve the prediction of probation, the natural step was to use several variables. The most appropriate technique was the discriminant function developed by Fisher.[1]

Table 2

Biserial for Variables and Probation

| Variable | Biserial r | | |
|---|---|---|---|
| | Liberal Arts | Commerce and Finance | University |
| Q score | .2892 | .2722 | .2706 |
| L score | .3250 | .4282 | .3699 |
| V score | .4127 | .3341 | .3890 |
| P score | .3382 | .4271 | .4115 |
| High school achievement | .6409 | .6355 | .5804 |

To obtain the linear function which would best discriminate between the two groups, Fisher determined the weights which would maximise the ratio obtained by dividing the variance between the composite means of the groups by the variance within the groups. The linear relationship between the dichotomous variable and the independent variables in deviation form was designated as $V = b_1x_1 + b_2x_2 + \ldots + b_kx_k$. It was assumed that the dichotomy consisted of two groups, Group A with $m_1$ variates and Group B with $m_2$

[1]Fisher, R. A. The use of multiple measurement in taxonomic problems. Annals of Eugenics. 7:179-188. 1936.

variates. The difference between the means of the two groups for any variable $X_i$ was $d_i = \bar{X}_{Ai} - \bar{X}_{Bi}$. Then the difference between composite means was symbolized by D, where

$$D = b_1 d_1 + b_2 d_2 + \ldots + b_k d_k = \bar{V}_A - \bar{V}_B$$

The within sum of squares is equal to $S = \sum_{q=1}^{k} \sum_{p=1}^{k} b_p b_q S_{pq}$

where $S_{pq}$ represents the sum of squares and products from the specific means within classes. If S is divided by the appropriate numbers of degrees of freedom, the variance within groups is obtained which is the denominator of the ratio specified by Fisher.[1]

The between means sum of squares is equal to

$$\frac{(\Sigma V_A)^2}{m_1} + \frac{(\Sigma V_B)^2}{m_2} - \frac{[\Sigma V_A + \Sigma V_B]^2}{m_1 + m_2} = \frac{m_1 m_2}{m_1 + m_2}(\bar{V}_A - \bar{V}_B)^2 = \frac{m_1 m_2}{m_1 + m_2} D^2.$$

Thus the variance between means is $\frac{m_1 m_2}{k(m_1 + m_2)} D^2$. Therefore the ratio to be maximized is $C \frac{(D^2)}{S}$ where

$$C = \frac{m_1 m_2 (m_1 + m_2 - k - 1)}{k(m_1 + m_2)} .$$

However, since the constant coefficient C, does not affect the value of the weights, $b_i$, the simplified function,

$F(b_1 b_2 \ldots b_K) = \frac{D^2}{S}$ is maximized to obtain the appropriate weights.

---

[1] Ibid.

Now Log F. = 2 log D - log S

$$\frac{\partial \text{Log } F}{\partial b_i} = \frac{2}{D}\frac{\partial D}{\partial b_i} - \frac{1}{S}\frac{\partial S}{\partial b_i} = 0$$

$$\frac{\partial S}{\partial b_i} = \frac{2S}{D}\frac{\partial D}{\partial b_i} \qquad (1)$$

In particular,

$$\frac{\partial S}{\partial b_i} = \frac{\partial\left(\sum_{q=1}^{k}\sum_{p=1}^{k} b_p b_q S_{pq}\right)}{\partial b_i}$$

$$= \sum_{q=1}^{k} b_q S_{iq} + \sum_{p=1}^{k} b_p S_{ip} = 2\sum_{i=1}^{k} b_i S_{ii} \qquad \text{when } i = 1.$$

Also, $\frac{\partial D}{\partial b_1} = d_1$

Substituting these values into equation (1) the first normal equation is $\sum_{i=1}^{k} b_i S_{1i} = \frac{S}{D} d_1$ .

Following this procedure the set of normal equations obtained are:

$$b_{11}S_{11} + b_2 S_{12} + \cdots + b_k S_{1k} = \frac{S}{D} d_1$$

$$b_1 S_{21} + b_2 S_{22} + \cdots + b_k S_{2k} = \frac{S}{D} d_2$$

$$- \quad - \quad - \quad - \quad - \quad - \quad - \quad -$$

$$b_1 S_{k1} + b_2 S_{k2} + \cdots + b_k S_{kk} = \frac{S}{D} d_k$$

Since each of the right hand members has a common factor $\frac{S}{D}$, the the equations are divided by $\frac{S}{D}$ to obtain a new set of normal equations

in which the coefficients are proportional to the original set of b's. The new set of normal equations obtained are the ones usually given in the literature. They are as follows:

$$C_1S_{11} + C_2S_{12} + \cdots + C_kS_{1k} = d_1$$
$$C_1S_{21} + C_2S_{22} + \cdots + C_kS_{2k} = d_2$$
$$- \quad - \quad - \quad - \quad - \quad - \quad - \quad -$$
$$C_1S_{k1} + C_2S_{k2} + \cdots + C_kS_{kk} = d_k$$

The solutions of this set of normal equations give the coefficients of the linear function $V = C_1x_1 + C_2x_2 + \cdots + C_kx_k$

If the original set of normal equations are multiplied by $b_1$, $b_2$, $\cdots$ $b_k$ respectively, and then added, the sum is equal to S. This is easily seen as the sum of the right hand sides of the equations becomes $\frac{S}{D}(b_1d_1 + b_2d_2 + \cdots + b_kd_k) = S$

since the quantity in parentheses is equal to D. This is an easier method of obtaining the value of S than by following the original definition for S.

The following analysis of variance shows the separation of the total variance into its two component parts. $F = \frac{B}{W}$ with k and $N - k - 1$ degrees of freedom is an exact test of significance of the difference between the two composite means $\overline{V}_A$ and $\overline{V}_B$. If the equations in the simplified set of normal equations are multiplied by $C_1$, $C_2$, $C_k$ respectively and then added, the sum is equal to $D = S$.

The sum of the right hand sides is $c_1d_1 + c_2d_2 + \ldots + c_kd_k = D$. The sum of the left hand sides of the equations gives the expression for S in terms of the original definition of $S = \sum_{q=1}^{k} \sum_{p=1}^{k} C_pC_qS_{pq}$.

This means that for the simplified set of normal equations the simplest way to obtain the "within" sum of squares is to evaluate D.

Table 3

Analysis of Variance

| Source | Degrees of Freedom | Sum of Squares | Mean Squares |
|---|---|---|---|
| Between composite test score means | k | $\frac{m_1m_2}{N} D^2$ | $\frac{m_1m_2}{kN} D^2 = B$ |
| Within composite test scores | N-k-1 | S | $\frac{S}{N-k-1} = W$ |
| Total | N-1 | $\Sigma v^2 = \Sigma(V - \bar{V})^2$ | |

In some of the later work it was necessary to find the total sum of squares, $\Sigma v^2$. The conventional method was to obtain the sum of squares by squaring and summing the discriminant function as follows:

$$\Sigma v^2 = \Sigma(b_1x_1 + b_2x_2 + \ldots + b_kx_k)^2$$

The deviation scores used here were deviations from the general mean rather than from the group means. This method of obtaining the total

sum of squares is convenient if just a few variables are involved. However, if 5 independent variables are used, for example, $\Sigma v^2$ consists of fifteen terms.

Now in the general discussion of the discriminant function it was shown that the "between" sum of squares was equal to $\frac{n_1 n_2}{N} D^2$ and that the "within" sum of squares was equal to D. Therefore, the total sum of squares was equal to $\frac{n_1 n_2}{N} D^2 + D$. Many times after a discriminant function has been determined it is modified in form by equating one of the coefficients to unity and making the appropriate changes in the rest of the coefficients. Even if the equations were modified in this manner, this method could be used by first determining the new values for the right hand sides of the normal equations, that is, the d's.

An even shorter method was developed by noting that the only difference between the original discriminant function,

$V = b_1 x_1 + b_2 x_2 + \ldots + b_k x_k$, and the modified function,

$V' = x_1 + \frac{b_2}{b_1} x_1 + \ldots + \frac{b_k}{b_1} x_k$ was the factor $\frac{1}{b_1}$.

Therefore the total sum of squares was found for the modified function by first finding the total sum of squares for the original equation and then multiplying by $\left(\frac{1}{b_1}\right)^2$. That is, the total sum of squares for the function $V' = x_1 + \frac{b_2}{b_1} x_2 + \ldots + \frac{b_k}{b_1} x_k$ was equal to $\left(\frac{1}{b_1}\right)^2 \left[ \frac{n_1 n_2}{N} D^2 + D \right]$ where again $D = b_1 d_1 + b_2 d_2 + \ldots + b_k d_k$.

A. College of Commerce and Finance

Reference to Table 1 shows that as individual measures of predicting probation, the Kuder Preference Record Scores were of no significant value. However, it was conceivable that in combination with some of the other variables, the contribution of these Interest Scores might be significant. To test the contribution made by the Kuder Preference Record Scores, discriminant functions were used. First, a discriminant function was computed using two independent variables that apparently would give a significant result. The two variables selected were the Vocabulary Score from the Nelson-Denny Reading Test, $X_4$, and high school achievement, $X_7$. Then a series of discriminant functions were formed using three independent variables, $X_4$, $X_7$, and one of the Interest Scores. Tests were made to determine if there was a significant loss in dropping from three to two variables by removing the Interest Score.

In the case of the two independent variables the normal equations were as follows:

$$11{,}832.7881b_1 + 4{,}799.5953b_2 = 6.5476$$

$$4{,}799.5953b_1 + 42{,}067.4763b_2 = 26.3333$$

When the coefficients were determined, the discriminant function was found to be $V = .00031394x_4 + .00059016x_7$. For convenience, a modified V was determined by equating the coefficient of $x_4$ to unity. The modified discriminant function was $V = x_4 + 1.87985x_7$.

A biserial r was computed between these variables and the probation tendency. This biserial coefficient, $R_{bis}$, is in reality a multiple biserial r and was computed from the formula

$$R_{bis} = \frac{D}{\sigma_v}\left(\frac{pq}{z}\right)$$

As was noted earlier, $D = b_1d_1 + b_2d_2 + \ldots + b_kd_k$. This appeared to be the most expedient method for computing D.

The standard deviation of the V scores for the total group was found by the formula $\sigma_v = \sqrt{\frac{\Sigma v^2}{N}}$

Methods for computing the total sum of squares have already been discussed. The multiple biserial R was found to be 0.6510.

The results obtained in computing the discriminant functions and multiple biserial R's for the three variables, $X_4$, $X_7$, and one of the Interest Scores, are shown in Table 4. The highest multiple biserial R of 0.6879 was obtained when the Clerical Interest Score, $X_{13}$, was combined with $X_4$ and $X_7$.

To determine whether or not the loss was significant when $X_{13}$ was eliminated and the multiple biserial R was reduced from 0.6879 to 0.6510, a t-test was made by using the formula

$$t = \sqrt{\frac{(R_k^2 - R_{k-1}^2)(N - k - 1)}{\frac{pq}{z^2} - R_k^2}}$$ where $R_k$ is the multiple biserial R determined for k variables. A value for t of 1.897 was obtained which was not significant at the five per cent

Table 4

Discriminant Functions with Three Variables
for Commerce and Finance

| Independent Variables | | Original Coefficient | Modified Coefficient | Multiple biserial R |
|---|---|---|---|---|
| $x_4$ | | .00033135 | 1.86760 | |
| $x_7$ | | .00059812 | 3.37121 | 0.6804 |
| $x_8$ | Computational | -.00017742 | -1.00000 | |
| $x_4$ | | .00026675 | 1.00000 | |
| $x_7$ | | .00066879 | 2.50718 | 0.6770 |
| $x_9$ | Persuasive | .00034577 | 1.29623 | |
| $x_4$ | | .00029391 | 1.31186 | |
| $x_7$ | | .00061933 | 2.76437 | 0.6670 |
| $x_{10}$ | Artistic | -.00022404 | -1.00000 | |
| $x_4$ | | .00036276 | 2.67937 | |
| $x_7$ | | .00059035 | 4.36037 | 0.6812 |
| $x_{11}$ | Literary | -.00013539 | -1.00000 | |
| $x_4$ | | .00033901 | 2.19296 | |
| $x_7$ | | .00060086 | 3.88680 | 0.6601 |
| $x_{12}$ | Social Service | .00015459 | 1.00000 | |
| $x_4$ | | .00034508 | 1.27251 | |
| $x_7$ | | .00058958 | 2.17413 | 0.6879 |
| $x_{13}$ | Clerical | -.00027118 | -1.00000 | |

level for 80 degrees of freedom. Thus the loss of information due to dropping $X_{13}$ was not considered significant. Since the highest value of $R_3$ had been tested, it was concluded that none of the Interest Scores made a significant contribution. Thus all parts of the Kuder Preference Record were dropped from further consideration for the predictive function for the College of Commerce and Finance.

The following method was used to obtain the discriminant function actually used for the prediction of probationary students in Commerce and Finance. The six remaining variables were used to determine a discriminant function. A multiple biserial of 0.7075 was obtained. A variable was dropped and a t test was made to determine if the loss was significant at the five per cent level. This process was continued until only $X_2$ and $X_7$ remained. The multiple biserial between these two variables and the probation tendency was 0.6797.

Table 5

Summary for six variables

| Variable | Original Coefficient | Modified Coefficient |
|---|---|---|
| $x_1$ | -0.00013440 | -1.000000 |
| $x_2$ | 0.00102947 | 7.659747 |
| $x_4$ | -0.00045834 | -3.410268 |
| $x_5$ | 0.00050889 | 3.786384 |
| $x_7$ | 0.00062129 | 4.622693 |
| $x_{14}$ | -0.00034746 | -2.585268 |

The normal equations necessary for determining the coefficients of the discriminant function with six variables appear on page 33. A value of 0.7075 was obtained for multiple biserial R. The coefficients for the successive discriminant functions determined by dropping variables, the multiple biserial R's obtained, and the value of t found for making the tests of significance are shown in Table 6.

It was necessary next to determine whether the reduction from a multiple biserial R of 0.6797 for the remaining two variables to a biserial r of 0.6355, obtained between high school achievement and probation, was significant. The value of t was found to be 2.061 which is significant at the five per cent level. This result indicated that the two variables, $X_2$ and $X_7$, should be used for prediction of probation.

The equation used for predicting probationary students from the 1949 class was $V = X_2 + 1.142696X_7$. V scores were computed for the 177 freshmen for whom complete data were available in the fall of 1949. The V scores ranged from 42 to 187. Under the assumption that the same proportion of freshman students would be placed on probation in 1949-1950 as were placed on probation in 1948-1949, a cut was made in the V scores just below the score of 92. This cut was also influenced by a material break in the distribution of the scores at this point. All of those whose scores were below 92 were considered predicted, probationary students. Out of the 137 students, 36 were placed in this category. This was the group from which

Table 6

Summary for Deleting Variables from College of Commerce and Finance Discriminant Functions

| Variable | Original Coefficient | Modified Coefficient | Multiple biserial R | t |
|---|---|---|---|---|
| | | Dropping $x_{14}$ | | |
| $x_1$ | -0.00016891 | -1.000000 | | |
| $x_2$ | 0.00079532 | 4.708543 | | |
| $x_4$ | -0.00056864 | -3.366527 | 0.6958 | 1.089 |
| $x_5$ | 0.00044983 | 2.663134 | | |
| $x_7$ | 0.00057825 | 3.423421 | | |
| | | Dropping $x_5$ | | |
| $x_1$ | -0.00007289 | -1.000000 | | |
| $x_2$ | 0.00084623 | 11.609686 | | |
| $x_4$ | -0.00040574 | -5.566470 | 0.6866 | 0.956 |
| $x_7$ | 0.00059272 | 8.131705 | | |
| | | Dropping $x_4$ | | |
| $x_1$ | 0.00003370 | 1.000000 | | |
| $x_2$ | 0.00048649 | 14.435905 | 0.6798 | 0.817 |
| $x_7$ | 0.00057635 | 17.102374 | | |
| | | Dropping $x_1$ | | |
| $x_2$ | 0.00050436 | 1.000000 | 0.6797 | 0.099 |
| $x_7$ | 0.00057633 | 1.142696 | | |

## Normal Equations

$$
\begin{aligned}
11{,}218.5239x_1 + 7{,}225.6430x_2 + 3{,}546.6429x_4 + 4{,}566.7858x_5 + 2{,}168.0954x_7 + 8{,}154.5953x_{14} &= 5.142857 \\
7{,}225.6430x_1 + 13{,}631.4525x_2 + 10{,}354.6668x_4 + 6{,}968.3096x_5 + 4{,}141.3096x_7 + 14{,}888.8810x_{14} &= 9.261905 \\
3{,}546.6429x_1 + 10{,}354.6668x_2 + 11{,}832.7881x_4 + 6{,}368.5238x_5 + 4{,}799.5953x_7 + 12{,}763.9524x_{14} &= 6.547619 \\
4{,}566.7858x_1 + 6{,}968.3096x_2 + 6{,}368.5238x_4 + 8{,}100.0239x_5 + 4{,}022.8811x_7 + 9{,}046.4525x_{14} &= 7.119048 \\
2{,}168.0954x_1 + 4{,}141.3096x_2 + 4{,}799.5953x_4 + 4{,}022.8811x_5 + 42{,}067.4763x_7 + 10{,}424.4049x_{14} &= 26.333333 \\
8{,}154.5953x_1 + 14{,}888.8810x_2 + 12{,}763.9524x_4 + 9{,}046.4525x_5 + 10{,}424.4049x_7 + 26{,}477.2620x_{14} &= 10.261905
\end{aligned}
$$

students were selected at random to form the experimental and control group. The control group received the traditional counseling while the experimental group received the supplemented counseling as discussed later.

Of the 36 students, 17 had been placed on probation at the end of the first semester and 6 had withdrawn from the university mainly because of scholastic difficulties. Thirteen of them were successful during their first semester.

In addition to the 17 students who had been predicted as probationary students and actually had been placed on probation there were 13 students whose V scores were greater than 92 who also were placed on probation. Thus, of the 30 students who were placed on probation, 17 had been predicted as probation. The percentage predicted, 57 per cent, in contrast to the 22 per cent which would have been expected if the potential, probationary group were selected at random, is a conservative measure of the effectiveness of the discriminant function to predict probationary students. Six students had withdrawn during the first semester. Their grades at the time that they withdrew indicated that they would have been placed on probation had they remained. Also, 18 of the 36 were members of the experimental group to whom supplemented counseling was given to avoid their being placed on probation. This counseling factor tended to lower the percentage of successful prediction of probation. If only the students in the control group were considered, 78 per cent of those students

were placed on probation or withdrew during the first semester.

Actually, this percentage is a function of the cut made to form the dichotomy. For example, if the cut had been made at 41 the percentage would have been zero as none of the probationary students had V scores below 41; whereas if the cut had been made at 150, the ratio would have increased to unity as all probationary students had scores below 150.

For the predicted non-probation students the V scores varied from 101 to 149. Their grade-point average was 1.370 for the first semester compared to an average of 1.904 for the experimental group and 1.070 for the control group. None of the differences was significant at the five per cent level.

When a V score is calculated for a particular student, he is classified as probationary or non-probationary according to some predetermined cut. It would be of interest to obtain some measure of the probability that the student was classified properly.

If these probabilities are to be determined, it is convenient to have the discriminant function in raw score form. In the general equation for the discriminant function, the x values are deviations from group means. The first step in getting the equation in raw score form is changing from "within" deviations to deviations from a total mean.

A general expression for one of the normal equations for obtaining the coefficients, $b_j$, for the discriminant function may be written

as

$$\sum_{j=1}^{k} S_{ij} b_j = d_i \qquad (1)$$

where, as before, $S_{ij}$ is equal to the "within" sum of squares or products.

If total deviations are used, the normal equations are of the form

$$\sum_{j=1}^{k} S'_{ij} b'_j = d_i \qquad (2)$$

where $S'_{ij}$ represents sum of squares or products from the general means.

Now $S'_{ij} = S_{ij} + \frac{m_1 m_2}{N} d_i d_j$ or

$$S_{ij} = S'_{ij} - \frac{m_1 m_2}{N} d_i d_j$$

Substituting in equation (1)

$$\sum_{j=1}^{k} (S'_{ij} - \frac{m_1 m_2}{N} d_i d_j) b_j = d_i \quad \text{or}$$

$$\sum_{j=1}^{k} S'_{ij} b_j = d_i + d_i \sum_{j=1}^{k} \frac{m_1 m_2}{N} b_j d_j$$

$$= d_i [1 + \frac{m_1 m_2}{N} \sum_{j=1}^{k} b_j d_j]$$

$$= d_i [1 + \frac{m_1 m_2}{N} D]$$

as $D = b_1 d_1 + b_2 d_2 + \ldots + b_k d_k = \sum_{j=1}^{k} b_j d_j$ .

If $C = 1 + \frac{m_1 m_2}{N} D$, the equation becomes

$$\sum_{j=1}^{k} S'_{ij} b_j = C d_i$$

Dividing by C, $\sum_{j=1}^{k} s' \frac{b_1}{C} = d_1$

Comparison with equation (2) shows that

$$b_j' = \frac{b_1}{C} = \frac{b_1}{1 + \frac{m_1 m_2}{N} D}$$

If $m_1 = m_2 = m$, $b_j' = \frac{b_1}{1 + \frac{m}{2} D}$

This indicates that a coefficient for the total deviation equation is found by dividing the corresponding coefficient in the "within" deviation equation by $1 + \frac{m_1 m_2}{N} D$.

A general expression is also needed for the distance between two means for groups which when combined form a normal distribution.

Assume a normal distribution with unit area and unit standard deviation. Suppose the area underneath the curve is divided into the proportion of p to q by an ordinate erected at $V_1$. Let $z_1$ represent the height of the ordinate at $V_1$. It is desired to find the average deviation of a truncated portion of the normal distribution. Following the general definition of a mean, the sum of all the deviations will be divided by their number. That is, the average deviation

$$= \int_{V_1}^{\infty} \frac{Vz}{q} dV = \frac{1}{q} \int_{V_1}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{V^2}{2}} V dV \qquad \text{as } z = \frac{1}{\sqrt{2\pi}} e^{-\frac{V^2}{2}}$$

$$= \frac{1}{q}\left(\frac{1}{\sqrt{2\pi}} e^{-\frac{V_1^2}{2}}\right) = \frac{z_1}{q}$$

In the same manner, the average deviation for the other portion of the curve is $\frac{z_1}{p}$. Therefore the total distance between the two means of the truncated portions of the normal curve is $\frac{z_1}{q} + \frac{z_1}{p} = \frac{z_1}{pq}$ as $p + q = 1$.

The reciprocal, $\frac{pq}{z}$, is easily obtained from a table such as that furnished by Peatman.[1]

To obtain the probability of proper classification for any student in Commerce College, the following data are needed:

$V = 0.00050436x_2 + .00057633x_7$

$D = 0.01984802$

$m_1 = m_2 = 42$

$1 + \frac{m}{2} D = 1 + 21\ D = 1.41680851$

$\sigma = 0.0182967$

$\bar{X}_2 = 63.4643$

$\bar{X}_7 = 53.0476$

$p = \frac{36}{137} = .2628$, the proportion of the 1949 commerce freshmen predicted as probationary

---

[1]Peatman, John Gray. Descriptive and sampling statistics. Harper and Brothers, New York. 1947. p. 519.

$$q = \frac{101}{137} = .7372$$

$$\frac{pq}{z} = .5938$$

L = .6347, the sigma distance from the mean, $\overline{V}$, to the cut.

The within deviation equation is divided by 1 + 21D, 1.41680851, to put it in total deviation form. Next, the equation is divided by $\sigma$, 0.0182967, so that the deviations from the general mean are expressed in standard units. The results of these two operations give $V = 0.01945614x_2 + 0.02223245x_7$.

The equation is then converted to raw score form by adding the constant $L - b_2\overline{X}_2 - b_7\overline{X}_7 = -1.77944842$. The assumption is made that the sigma unit used in computing the probabilities of correct classification should be the sigma distance between the two group means. This has been found to be $\frac{z}{pq}$. Therefore the final step is dividing the last equation by $\frac{z}{pq}$, or multiplying by $\frac{pq}{z}$, a tabulated value.

Following these two last steps, the equation becomes $v = 0.01155306X_2 + 0.01320163X_7 - 1.05663647$. For any values of $X_2$ and $X_7$, the above equation gives the deviation in standard units which is used as an argument for entering a normal curve table to obtain the probability of a student belonging to a particular group.

For example, a student has a $X_2$ score of 20 and $X_7$ of 30. Using the foregoing equation, the value of v is found to be -0.4295. Reference to a normal curve table gives a probability of 0.666 that

this student belongs in the probation group. For another student who has scores of 70 and 90, respectively, for $X_2$ and $X_7$, the value of v obtained and is 0.9403, which leads to a probability of 0.174 that this student is a member of the probation group.

## B. College of Liberal Arts

As was found for Commerce and Finance, reference to Table 1 indicates that as individual measures of predicting probation, the Kuder Preference Record Scores were of no significant value. A preliminary study was made to determine if they made a significant contribution in combination with some of the other variables. As before, a discriminant function was determined using $X_4$ and $X_7$ as the independent variables. Then a series of discriminant functions were formed using three independent variables, $X_4$, $X_7$, and one of the Interest Scores. Tests were made to determine if there was a significant loss in dropping from three to two variables.

The normal equations for the discriminant function with $x_4$ and $x_7$ as variables were as follows:

$$15{,}480.4844b_1 + 8{,}051.5313 = 7.671875$$

$$8{,}051.5313b_1 + 72{,}180.7969 = 28.265625$$

The solution of these equations gave the following discriminant function. $V = 0.00030990x_4 + 0.00035703x_7$

For convenience the equation was modified by equating the coefficient of $x_4$ to unity. The modified discriminant function was

$V = x_4 + 1.152049x_7$. The multiple biserial R computed between these variables and the probation tendency was found to be 0.6693, a value similar to that of 0.6510, obtained for Commerce and Finance.

Table 7 gives the coefficients for the discriminant functions and the multiple biserial R's obtained between the three variables and probation tendency. The highest multiple biserial R's were obtained when the Artistic Interest Score, $X_{10}$, or the Computational Interest Score, $X_8$, were used in combination with $X_4$ and $X_7$.

The t test was used again to determine whether the loss was significant when $X_{10}$ was eliminated. On substituting in the formula used previously, t was found to be 1.271 which was not significant. Since the largest value of $R_3$ obtained had been tested, it was concluded that none of the Kuder Interest Scores made a significant contribution and they were dropped from further consideration for developing a predictive function for the College of Liberal Arts.

The remaining five variables, the two scores from the ACE Test, the two parts of the Nelson-Denny Reading Test, and high school achievement, were used to form a discriminant function. A multiple biserial R of 0.6751 was obtained. The same method was used as before. Variables were deleted one at a time. Each time a variable was dropped a t test was made to determine if the loss was significant. None of the losses proved to be significant so that the final discriminant function was one of two variables, $X_4$ and $X_7$.

The normal equations necessary for determining the coefficients of the discriminant function with five variables are given on page 43.

Table 7

Discriminant Functions with Three Variables for Liberal Arts

| Variables | Original Coefficients | Modified Coefficients | Multiple Biserial R |
|---|---|---|---|
| $x_4$ | 0.00026068 | 1.54550 | |
| $x_7$ | 0.00036892 | 2.18723 | 0.6792 |
| $x_8$ | -0.00016867 | -1.00000 | |
| $x_4$ | 0.00031252 | 8.51089 | |
| $x_7$ | 0.00035741 | 9.73338 | 0.6703 |
| $x_9$ | 0.00003672 | 1.00000 | |
| $x_4$ | 0.00033142 | 2.47717 | |
| $x_7$ | 0.00035342 | 2.64160 | 0.6800 |
| $x_{10}$ | 0.00013379 | 1.00000 | |
| $x_4$ | 0.00031421 | 33.82238 | |
| $x_7$ | 0.00035654 | 38.37890 | 0.6693 |
| $x_{11}$ | -0.00000929 | -1.00000 | |
| $x_4$ | 0.00031055 | 10.58453 | |
| $x_7$ | 0.00035665 | 12.15576 | 0.6699 |
| $x_{12}$ | 0.00002934 | 1.00000 | |
| $x_4$ | 0.00028775 | 3.92458 | |
| $x_7$ | 0.00036786 | 5.01718 | 0.6725 |
| $x_{13}$ | -0.00007332 | -1.00000 | |

## Normal Equations for Liberal Arts

$$12{,}498.8438x_1 + 6{,}402.7813x_2 + 4{,}222.7657x_4 + 4{,}381.3751x_5 + 4{,}883.8282x_7 = 4.687500$$

$$6{,}402.7813x_1 + 23{,}525.4375x_2 + 15{,}501.5000x_4 + 12{,}509.3125x_5 + 4{,}146.2813x_7 = 7.281250$$

$$4{,}222.7657x_1 + 15{,}501.5000x_2 + 15{,}480.4844x_4 + 9{,}706.9688x_5 + 8{,}051.5313x_7 = 7.671875$$

$$4{,}381.3751x_1 + 12{,}509.3125x_2 + 9{,}706.9688x_4 + 13{,}180.8750x_5 + 5{,}586.9063x_7 = 5.687500$$

$$4{,}883.8282x_1 + 4{,}146.2813x_2 + 8{,}051.5313x_4 + 5{,}586.9063x_5 + 72{,}180.7969x_7 = 28.265625$$

The coefficients for the successive discriminant functions, the multiple biserial R's obtained, and the values of t obtained on making the tests of significance are shown in Table 8.

The multiple biserial R of 0.6693 for the two variables was compared next with the biserial r of 0.6409 obtained between high school achievement and probation. A t value of 2.035 was obtained which is significant. This result indicated that it would be advantageous to use the two variables, $X_4$ and $X_7$ for prediction of probation.

The equation used for predicting probationary students from the 1949 class was $V = X_4 + 1.152049X_7$. The V scores ranged from 32 to 182 for the 177 students for whom V scores were computed. The same basic assumption was made to determine the cut as was used for Commerce and Finance. All students whose V scores were below 75 were predicted as probationary students thus forming a group of 50 students. Of this group, 26 were placed on probation at the end of their first semester and four withdrew from the university during the semester. In addition to these 26 students, there were 14 students placed on probation whose V scores were above 75. Thus, of the 40 students who were placed on probation and for whom complete data were available, 26 or 65 per cent were predicted as probation. If only the control group was considered, 72 per cent of them were placed on probation or withdrew during the first semester.

For the non-predicted students the V scores ranged from 79 to

Table 8

Summary for Deleting Variables from College of Liberal Arts Discriminant Functions

| Variable | Original Coefficients | Modified Coefficients | Multiple biserial R | t |
|---|---|---|---|---|
| $x_1$ | 0.000123852 | 3.484960 | | |
| $x_2$ | 0.000078231 | 2.201272 | | |
| $x_4$ | 0.000175835 | 4.947663 | 0.6751 | |
| $x_5$ | 0.000035539 | 1.000000 | | |
| $x_7$ | 0.000356356 | 10.027181 | | |
| | | Dropping $x_5$ | | |
| $x_1$ | 0.00012677 | 1.404809 | | |
| $x_2$ | 0.00009024 | 1.000000 | | |
| $x_4$ | 0.00018484 | 2.048316 | 0.6749 | 0.172 |
| $x_7$ | 0.00035722 | 3.958555 | | |
| | | Dropping $x_2$ | | |
| $x_1$ | 0.00014547 | 1.000000 | | |
| $x_4$ | 0.00027320 | 1.878050 | 0.6738 | 0.405 |
| $x_7$ | 0.00035128 | 2.414793 | | |
| | | Dropping $x_1$ | | |
| $x_4$ | 0.00030990 | 1.000000 | | |
| $x_7$ | 0.00035703 | 1.152049 | 0.6693 | 0.819 |

129. The first semester grade point average for this group was 1.514 compared to an average of 1.757 for the probation group and 1.552 for the control group. None of these differences was significant at the five per cent level.

A measure of the probability that a student is properly classified can be obtained by following the procedure outlined for the Commerce and Finance group. The following data are needed:

$V = 0.00030990x_4 + 0.00035703x_7$

$D = 0.0126919$

$1 + 32D = 1.39901408$

$\sigma = 0.0116742$

$\bar{X}_4 = 34.5234$

$\bar{X}_7 = 52.5859$

$p = 0.2825$

$\frac{pq}{s} = 0.5996$

$L = 0.5755$

The equation used for determining the probabilities is found to be $v = 0.01011377X_4 + 0.01310742X_7 - 0.73697395$

As before, v is the deviation in standard units which is used as an argument for entering a normal curve table to obtain the probability.

For example, a student has a $X_4$ score of 20 and a $X_7$ score of 25. Substituting into the foregoing equation the value of v is found to be -0.2070 and reference to a normal curve table gives a probability of 0.581 that this student belongs to the probationary group.

## C. Education and Fine Arts

Since there were so few students on which to base a discriminant function for the College of Education or for the College of Fine Arts, an all-university discriminant function was developed based on the 143 freshman students who had been placed on probation in the four colleges during 1948-1949. An equal number of non-probation students was selected from the colleges to correspond to the number of probation students which the particular college contributed. The only significant change in the procedure was that the "within deviations" used were "deviations within groups within colleges" since the predictions would be made for students from known colleges.

First, a discriminant function using the five independent variables, $X_1$, $X_2$, $X_4$, $X_5$, and $X_7$, was found. The normal equations are given on page 49. Again the same process was adopted of deleting a variable and making a t test to determine if the loss was significant. A summary of the results is given in Table 9. Again $X_4$ and $X_7$ were the two variables which made the most significant contribution.

The equation used for prediction was $V = X_4 + 1.037029X_7$. V scores were computed for 70 freshman students in the College of Education and 92 freshman students from Fine Arts for whom complete data were available. Under the basic assumption used previously, cuts were determined for the two distributions. By coincidence, both cuts were made at 57 and both prediction groups contained 14 students.

Table 9

Summary for All-university Discriminant Functions

| Variable | Original Coefficients | Modified Coefficients | Multiple biserial R | t |
|---|---|---|---|---|
| $x_1$ | 0.00002354 | 1.000000 | | |
| $x_2$ | 0.00014204 | 6.033984 | | |
| $x_4$ | 0.00000030 | .002112 | 0.6340 | |
| $x_5$ | 0.00003237 | 1.375106 | | |
| $x_7$ | 0.00014661 | 6.228122 | | |
| | | Dropping $x_5$ | | |
| $x_1$ | 0.00002142 | 28.560000 | | |
| $x_2$ | 0.00016607 | 221.426666 | | |
| $x_4$ | -0.00000075 | -1.000000 | 0.6299 | 1.114 |
| $x_7$ | 0.00014710 | 196.133333 | | |
| | | Dropping $x_1$ | | |
| $x_2$ | 0.00018368 | 18.043222 | | |
| $x_4$ | -0.00001018 | -1.000000 | 0.6257 | 1.124 |
| $x_7$ | 0.00014777 | 14.515717 | | |
| | | Dropping $x_2$ | | |
| $x_4$ | 0.00014151 | 1.000000 | | |
| $x_7$ | 00.00014675 | 1.037029 | 0.6237 | 0.773 |

**Normal Equations**

$$33{,}809.6844x_1 + 19{,}126.0376x_2 + 9{,}343.2619x_4 + 11{,}810.1609x_5 + 7{,}710.1137x_7 = 5.027972$$

$$19{,}126.0376x_1 + 39{,}341.2442x_2 + 30{,}849.0868x_4 + 27{,}133.1154x_5 + 11{,}667.8476x_7 = 8.636364$$

$$9{,}343.2619x_1 + 30{,}849.0868x_2 + 37{,}457.0592x_4 + 21{,}308.2393x_5 + 15{,}154.9999x_7 = 7.524476$$

$$11{,}810.1609x_1 + 27{,}133.1154x_2 + 21{,}308.2393x_4 + 29{,}959.2722x_5 + 9{,}947.0640x_7 = 6.566434$$

$$7{,}710.1137x_1 + 11{,}667.8476x_2 + 15{,}154.9999x_4 + 9{,}947.0640x_5 + 152{,}980.5366x_7 = 24.594406$$

Both groups were combined for the purpose of appraising the effectiveness of the discriminant function. Out of the combined group of 28 students, 10 were placed on probation at the end of the first semester and 5 withdrew from the university during the semester. In addition to the predicted students there were 4 from Education and 4 from Fine Arts with V scores above 75 who were also placed on probation. Thus, of the 18 students who were placed on probation, 10, or 56 per cent, had been predicted probation. If only the control groups were considered, 71 per cent had been placed on probation or had withdrawn during their first semester.

The range of V scores for the predicted non-probationary students was from 58 to 84. Their grade point average for the first semester was 1.625 compared to 2.685 for the experimental group and 1.536 for the control group.

## V. APPLICATION OF PREDICTION

From the point of view of counseling, the most important question was what could be done to help these predicted, probationary students to avoid their being placed on probation or withdrawing from the university. At Drake University, counseling is the responsibility of the general faculty. None of the members of the faculty has his teaching load decreased so that he can handle a greater portion of the counseling. Thus, a procedure had to be adopted which would be administratively feasible year after year. Also, it was felt that it would be worthless from a practical point of view to promote a counseling program in which the faculty members would be reluctant to participate each year. The main question was whether with the present available facilities and policies, it would be possible to give significant aid to the predicted group of probationary students.

None of the students was aware of the fact that he had been predicted a probationary student. The students in the control group went through the usual counseling procedures. Their faculty advisers did not know of their classification as potential probation students. The study was, however, thoroughly discussed with the faculty counselors of students in the experimental group. They knew that these advisees were potentially poor students and had been predicted as probation students. They were well aware of the challenge in working with this low ability group.

To aid the counselors for the experimental group, the Mooney Problem Check List, College Form, was filled out by each student of this group. This check list is a listing of 330 troublesome problems which students in college often face. For convenience in summarizing results on an individual case the 330 problems are classified in eleven areas:

(1) Health and Physical Development
(2) Finances, Living Conditions, and Employment
(3) Social and Recreational Activities
(4) Social-Psychological Relations
(5) Personal-Psychological Relations
(6) Courtship, Sex, and Marriage
(7) Home and Family
(8) Morals and Religion
(9) Adjustment to College Work
(10) The Future: Vocational and Educational
(11) Curriculum and Teaching Procedures

In addition to checking problems which are troublesome to him, a student is given the opportunity to indicate additional problems, to summarize in his own words his chief problems, and to indicate his reaction to the check list. These check lists were filled out by the students after the first six weeks of the fall semester and at the completion of the first semester before marks had been released. The Mooney Problem Check List seemed to be worthwhile for two reasons.

First, it forced a student to consciously think of his problems and realize that he might need outside help. Second, it was an expedient through which a counselor might quickly learn of his advisee's difficulties.

Informal discussions were held between the university director of counseling and testing and the counselors. Not only were good principles of guidance and the various facilities of counseling and testing available to the student discussed, but the meetings afforded an opportunity for the counselors to discuss their specific counseling problems with the group. These counselors were encouraged to use all available facilities such as: Counseling and Testing Center, Reading and Study Skills Laboratory, Placement Services, Writing Laboratory, and Vocational Counseling for Veterans.

The disposition of the predicted probation students is shown in Table 10. Reference to the all-university tabulations indicates that the experimental group was more successful than the control group. More of them had satisfactory marks; fewer of them were placed on probation; and a smaller number of them withdrew from the university.

To evaluate the effectiveness of the counseling, five criteria were available. First, college achievement was defined as average marks for the semester. Within each college a test of significance was made of the difference between the means of the two groups. The hypothesis that the two groups were random samples from the same normal population was tested by a t test where

Table 10

Disposition of Predicted Probation Students

| | Commerce | | Liberal Arts | | Fine Arts | | Education | | All University | |
|---|---|---|---|---|---|---|---|---|---|---|
| | E | C | E | C | E | C | E | C | E | C |
| Initial No. | 18 | 18 | 25 | 25 | 7 | 7 | 7 | 7 | 57 | 57 |
| Non-Probation 1st semester | 9 | 3 | 10 | 7 | 5 | 3 | 4 | 1 | 28 | 14 |
| Probation 1st semester | 7 | 11 | 11 | 15 | 1 | 2 | 1 | 6 | 20 | 34 |
| Withdrew during semester | 2 | 4 | 4 | 0 | 1 | 2 | 2 | 0 | 9 | 6 |
| Withdrew at end of semester | 0 | 3 | 0 | 3 | 0 | 1 | 0 | 2 | 0 | 7 |
| Number at beginning of 2nd semester | 16 | 11 | 21 | 22 | 6 | 4 | 5 | 5 | 48 | 42 |
| Non-probation 2nd semester | 8 | 5 | 12 | 10 | 6 | 3 | 5 | 3 | 31 | 21 |
| Probation 2nd semester | 8 | 4 | 8 | 10 | 0 | 0 | 0 | 2 | 16 | 16 |
| Withdrew during semester | 0 | 2 | 1 | 2 | 0 | 1 | 0 | 0 | 1 | 5 |
| Dropped by the Dean of College at end of year | 2 | 1 | 1 | 2 | 0 | 0 | 0 | 0 | 3 | 3 |
| Number remaining at end of year | 14 | 8 | 19 | 18 | 6 | 3 | 5 | 5 | 44 | 34 |

$$t = \frac{\bar{X}_A - \bar{X}_B}{\sqrt{\frac{\Sigma(X_A - \bar{X}_A)^2 + \Sigma(X_B - \bar{X}_B)^2}{m_1 + m_2 - 2}\left(\frac{1}{m_1} + \frac{1}{m_2}\right)}} .$$

For these tests the College of Education and the College of Fine Arts were combined into a single group. Since the object of the analysis was not to test differences between colleges but to test differences between experimental and control groups the combination was beneficial in that it gave larger groups with which to work. Also, an all-university experimental group and control group were formed by combining the respective groups from the four colleges and a test of significance was made on an all-university basis. The results are reported in Table 11. Significant differences were obtained between the combined Fine Art and Education groups and also between the all-university groups. Although the t value of 1.707 obtained for Commerce and Finance was not significant at the conventional five per cent level, it was significant at the 10 per cent level.

Next, college achievement was defined as average marks for the school year, 1949-1950. Tests, similar to the foregoing, were made. Again the differences were significant between the experimental and control groups for both the all-university combination and the Education and Fine Arts combinations. Table 11 also includes a summary of these tests.

The objective of the counseling was not only to elevate the scholastic achievement of the predicted students, but to prevent, if

possible, their being placed on probation or their withdrawing from the university. Only if the counseling were effective in this direction could student mortality be decreased.

For purposes of evaluation, a student was classified as unsuccessful if he had been placed on probation or had withdrawn from the university. All others were designated as successful. Using these classifications, college achievement was defined in two ways: (1) as successful-unsuccessful status at the end of the first semester and (2) as successful-unsuccessful status at the end of the school year.

For a particular college, students were classified by two characteristics, experimental-control and successful-unsuccessful on the first semester basis. A four cell contingency table was formed. The hypothesis that the two characteristics were independent was tested by chi square for which expected frequencies were obtained from row and column totals. When the frequency of a cell was less than five, the Yates[1] correction was applied. The contingency tables and values of chi square obtained for the first semester are given in Table 12. The chi-square value for the all university combination was 7.389, a highly significant value indicating rejection of the hypothesis of independence. The chi square for Commerce was 3.125. Under the hypothesis stated previously, the probability of getting a value as large as 3.125 in repeated random sampling was less than eight in a

---

[1]Yates, F. Contingency tables involving small numbers and the chi square test. Supplement to the Journal of the Royal Statistical Society. 1:217-235. 1934.

Table 11

Summary of Average Marks for Experiment and Control Groups

| Group | Average Marks | |
|---|---|---|
| | 1st semester | Year |
| Commerce | | |
| Experimental | 1.904 | 1.750 |
| Control | 1.070 | 1.338 |
| t | 1.707 | 1.197 |
| Liberal Arts | | |
| Experimental | 1.757 | 1.652 |
| Control | 1.552 | 1.675 |
| t | .676 | .100 |
| Education and Fine Arts | | |
| Experimental | 2.685 | 2.693 |
| Control | 1.536 | 1.877 |
| t | 2.515* | 2.606* |
| All University | | |
| Experimental | 2.019 | 1.965 |
| Control | 1.416 | 1.633 |
| t | 2.550* | 1.964* |

* significant at five per cent level

hundred. Although not conclusive, there is some evidence for rejecting the hypothesis of independent classifications. Rejection of this hypothesis would indicate that there was a difference between the experimental and control groups with regard to the status of success.

Using the definition of college achievement on a school year basis, the effectiveness of the counseling procedure was evaluated. Contingency tables were formed with the same foregoing classifications, and chi square was used again as a test. Not only were significant values obtained between the all-university groups but also between the combined Education and Fine Arts groups. The value of chi square for Commerce was slightly higher than that found previously. These values are found in Table 13.

The number of students who withdrew from the university or were dropped by their dean for low grades was used as a simple measure of mortality. Of the 57 students in the experimental group and the 57 students in the control group, 13 and 23, respectively, withdrew from the university or were dropped by their dean for low, academic standing. Under the hypothesis that the survivor-mortality classification was independent of the experimental-control classification, a chi-square test was made. The value of chi square obtained was 4.060. Under the hypothesis stated the probability of obtaining a value as large as 4.060 with repeated random sampling was less than five in a hundred. Evidence indicates there was a significant difference between the experimental and control groups with respect to mortality. Mortality was

Table 12

Contingency Tables and Chi-square Values for First Semester

| Group | Classification | | Total |
|---|---|---|---|
| | Successful | Unsuccessful | |
| Commerce | | | |
| Experimental | 9 | 9 | 18 |
| Control | 3 | 15 | 18 |
| Total | 12 | 24 | 36 |
| Chi square | 3.125 | P = 0.08 | |
| Liberal Arts | | | |
| Experimental | 10 | 15 | 25 |
| Control | 7 | 18 | 25 |
| Total | 17 | 33 | 50 |
| Chi square | 0.802 | P = 0.39 | |
| Education and Fine Arts | | | |
| Experimental | 9 | 5 | 14 |
| Control | 4 | 10 | 14 |
| Total | 13 | 15 | 28 |
| Chi square | 2.297 | P = 0.14 | |
| All University | | | |
| Experimental | 28 | 29 | 57 |
| Control | 14 | 43 | 57 |
| Total | 42 | 72 | 114 |
| Chi square | 7.389** | P < 0.01 | |

Table 13

Contingency Tables and Chi-square Values for First Year

| Group | Classification | | Total |
|---|---|---|---|
| | Successful | Unsuccessful | |
| **Commerce** | | | |
| Experimental | 17 | 17 | 34 |
| Control | 8 | 21 | 29 |
| Total | 25 | 38 | 63 |
| Chi square | 3.285 | P = 0.07 | |
| **Liberal Arts** | | | |
| Experimental | 22 | 24 | 46 |
| Control | 17 | 30 | 47 |
| Total | 39 | 54 | 93 |
| Chi square | 1.297 | P = 0.26 | |
| **Education and Fine Arts** | | | |
| Experimental | 20 | 4 | 24 |
| Control | 10 | 14 | 24 |
| Total | 30 | 18 | 48 |
| Chi square | 7.200** | P < 0.01 | |
| **All University** | | | |
| Experimental | 59 | 45 | 104 |
| Control | 35 | 65 | 100 |
| Total | 94 | 110 | 204 |
| Chi square | 9.689** | P < 0.01 | |

concluded to be lower for the experimental group.

In summary, all of the evidence indicates that the supplemented counseling was beneficial to this incipient, probationary group. It would seem wise to include all such students under the extended, counseling program. If this were done, student mortality due to scholastic achievement might be appreciably decreased.

## VI. SUMMARY

Among factors observed upon examination of Drake University records, one of the most disturbing noted was the high student mortality, especially among freshmen groups. This observation is emphasized by similar conditions in other colleges and universities. The situation naturally results from the withdrawing of large numbers of students, due to the inability of the student to adjust to college-life difficulties of one type or another. This study is confined to one area of difficulty, that is, the scholastic problem which undoubtedly represents a large portion of such situations encountered by students. More specifically, this study is confined to the freshman probation group. It would seem that a favorable approach to the solution of this problem may be developed through a supplemented counseling program.

It was assumed that fewer students would be placed on probation if effective counseling could reach the student long before he was placed on probation. The need for some method by which potential probation students could be designated, is clearly seen. Data available in the Admission Office and Testing Bureau at Drake University were used to develop an effective, predictive scheme.

This predictive scheme must then be put to its best use in preparation for the most desirable counseling.

Complete data were available for 666 freshmen from the 1948 class.

The data available for each of the students were as follows: test scores on A.C.E., Nelson-Denny Reading Test, the Kuder Preference Record, and the high school percentile rank. The discriminant function was used to give maximum separation between the probation and non-probation groups so that identification of an individual with a group was more reliable. Using the scores on the two parts of A.C.E., the two sections of the Nelson-Denny Reading Test, six of the nine Interest Scores, and the high school percentile rank gave 11 independent variables to consider. For the Commerce students, the score of American Institute of Accountant's Orientation Test was used as an additional variable.

Of the 666 students with complete data, 187 were from the College of Commerce and Finance. Preliminary work was done first to determine if the Interest Scores from the Kuder Preference Record would make a significant contribution to a discriminant function formed. It was concluded that they made no significant contribution and could be eliminated from further consideration. A discriminant function was formed from the six remaining variables. Variables were deleted and the losses were tested until only two variables remained, $X_2$, linguistic score on the A.C.E., and $X_7$, high school achievement. The discriminant function actually used for prediction of probationary students in the College of Commerce was $V = X_2 + 1.142696X_7$. As a measure of the relationship between probation tendency and these two variables, a multiple biserial R was computed and found to be 0.6797.

The same procedure was followed to obtain a discriminant function based on the 234 students from Liberal Arts College. Again, after a preliminary analysis, the Kuder Interest Scores were dropped from further consideration. A discriminant function with the five remaining independent variables was then determined. Variables were deleted and the losses were tested until again only two variables remained, $X_4$, Paragraph Reading Comprehension of the Nelson-Denny Test, and, as before, $X_7$. $V = X_4 + 1.152049X_7$ was the discriminant function used for Liberal Arts College. A multiple biserial R of 0.6693 was found between probation tendency and the two variables.

Since the number of students involved from the College of Education and the College of Fine Arts respectively was too small on which to base a discriminant function, the predicted probationary students for these colleges were obtained from an all-university discriminant function. Following the foregoing procedure, the equation of the discriminant function finally used was $V = X_4 + 1.037029X_7$. A multiple biserial R of 0.6237 was obtained.

All three of these multiple biserial R's seemed to indicate a fair degree of success in predicting probationary students. To determine which of the 476 freshmen were potential probationary students, V scores were computed for each of them using the appropriate discriminate function. The predicted probation list of 114 students consisted of 36 students from Commerce College, 50 from College of Liberal Arts, 14 from College of Education, and 14 from College of

Fine Arts. Of the 114 students, 68, or 60 per cent, were on probation or withdrew from the university at the end of the first semester. In addition to the 53 predicted students there were 33 other freshman students also on probation at the end of the first semester. Thus, of the 86 freshmen placed on probation, 53, or 62 per cent, had been predicted as such.

Naturally there are many other factors, such as social maladjustment, too many hours of outside work, or poor study habits, which will affect a student's academic achievement. The use of the discriminant function was an attempt to use variables which were available before a student began his college experience to predict probation tendency.

Within each college, students from the predicted probationary group were selected at random to form experimental and control groups. The students in the control group were exposed to the ordinary counseling while the experimental group received supplemental counseling.

Five criteria available for the evaluation of the effectiveness of the counseling were: (1) average marks for the first semester, (2) average marks for the school year, (3) successful-unsuccessful status at the end of the first semester, (4) successful-unsuccessful status at the end of the school year, and (5) survivor-mortality status at the end of the school year.

For purposes of analysis, the Colleges of Commerce and Liberal Arts were treated separately; the Colleges of Education and Fine Arts

were combined; and an all-university group was formed combining all colleges.

Under the hypothesis that the experimental and control groups were random samples from the same population, a t test was used to determine if differences in average marks between groups were significant.

At the end of the first semester, significant differences were found between groups for the all-university combination and for the College of Education and College of Fine Arts combination. Again, at the end of the year, significant differences between groups were obtained for the same combinations.

If a student had been placed on probation or had withdrawn from the university he was designated as unsuccessful. All others were classified as successful. Four cell contingency tables were formed by classifying the students according to the two characteristics, experimental-control and successful-unsuccessful. Under the assumption that the two characteristics were independent, chi-square tests were performed. At the end of the first semester there was some evidence to indicate a difference between the experimental and control groups for the all-university combination. At the end of the year, significant differences were indicated between groups for the all-university and College of Education and College of Fine Arts combinations.

For the final evaluation, a student was classified as a mortality

if he withdrew from the university or had been dropped from the university by his dean for low academic standing. All others were designated as survivors. A four cell contingency table was formed and a significant value of chi square, 4.060, was obtained. It was concluded that there was a significant difference between the experimental and control groups with respect to mortality.

In summary, all of the evidence indicates that the supplemented counseling was beneficial to this incipient, probationary group. It would seem wise to include all such students under the extended, counseling program. If this were done, student mortality due to scholastic achievement might be appreciably decreased.

## VII. SELECTED REFERENCES

Barnard, M. M. The secular variation of skull characters in four series of Egyptian skulls. Annals of Eugenics. 6:352-371. 1935.

Beall, Geoffrey. Approximate methods in calculating discriminant functions. Psychometrika. 10:205-217. 1945.

Bixler, Roy W. and Ady, Carolyn R. A study of the freshman class of 1947-1948. Unpublished mimeographed manuscript. Registrar's Office, Drake University, Des Moines, Iowa. 1949.

Brown, George W. Discriminant function. Annals of Mathematical Statistics. 18:514-528. 1947.

Day, Bessie B. and Sandomire, Marion M. Use of the discriminant function for more than two groups. Journal of the American Statistical Association. 37:461-472. 1942.

Durflinger, Glenn W. The prediction of college success - a summary of recent findings. Journal of the American Association of Collegiate Registrars. 19:68-78. 1943.

Fisher, R. A. The use of multiple measurements in taxonomic problems. Annals of Eugenics. 7:179-188. 1936.

_____ The statistical utilization of multiple measurements. Annals of Eugenics. 8:376-386. 1938.

_____ The precision of discriminant functions. Annals of Eugenics. 10:422-429. 1940.

Garrett, Henry E. The discriminant function and its use in psychology. Psychometrika. 8:65-79. 1943.

_____ Review and interpretation of investigations of factors related to scholastic success in colleges of arts and science and teachers' colleges. Journal of Experimental Education. 12:184-186. 1944.

Harris, Daniel. Factor affecting college grades - a review of the literature 1930-1937. Psychological Bulletin. 37:125-166. 1940.

Herdman, Raymond Wain. Predicting pupil mortality among high school boys. Unpublished M. S. Thesis. Ames, Iowa, Iowa State College Library. 1949.

Hotelling, Harold. Generalization of "student's" ratio. Annals of Mathematical Statistics. 2:360-378. 1931.

Jackson, Robert. Approximate multiple regression weights. Journal of Experimental Education. 11:221-225. 1943.

_____ The selection of students for freshman chemistry by means of discriminant function. Journal of Experimental Education. 18:209-215. 1950.

Johnson, Palmer O. Statistical methods in research. New York, Prentice-Hall, Inc., 1949.

MacRae, John Murray. Usefulness of the Minnesota Personality Scale for predicting achievement of freshman engineering students. Unpublished M. C. Thesis. Ames, Iowa, Iowa State College Library. 1949.

Mahalanobis, P. C. Analysis of race mixture in Bengal. Journal Asiat. Soc. Beng. 23:301-333. 1927.

McNeeley, John H. College student mortality. U. S. Department of the Interior, Office of Education, Bulletin 1937, No. 11, Washington, D.C. U. S. Government Printing Office. 1938.

Mitchell, Fred T. Why freshmen leave college. Journal of Higher Education. 13:95-100. 1942.

Peatman, John Gray. Descriptive and sampling statistics. Harper and Brothers, New York, 1947. p. 519.

Remmers, H. H. and others. Curricular differences in predicting scholastic achievement - application to counseling. Journal of Educational Psychology. 39:183-186. 1948.

Ross, C. C. Should low ranking college freshmen be told their scores on intelligence tests? School and Society. 70:23-25. 1949.

Selover, Robert B. A study of the sophomore testing program at the University of Minnesota. Journal of Applied Psychology. 26:456-467. 1942.

Travers, R. M. W. The use of the discriminant function in the treatment of psychological group differences. Psychometrika. 4:25-32. 1938.

Wallace, Noel and Travers, R. M. W. A psychometric sociological study of a group of specialty salesmen. Annals of Eugenics. 8:266-302. 1938.

Wert, James E. and others. The discriminant function and the multiple biserial R. Unpublished mimeographed manuscript. The Educational Research Laboratories, 315 Curtiss Hall, Iowa State College, Ames, Iowa. 1949.

Yates, F. Contingency tables involving small numbers and the chi square test. Supplement to the Journal of the Royal Statistical Society. 1:217-235. 1934.

# VIII. APPENDIX

Raw and Deviation Scores for 1948-1949 Commerce and Finance Students

| Symbol | Raw Scores | | | Deviation Scores | | | |
|---|---|---|---|---|---|---|---|
| | Probation | Non-Probation | Both | Probation | Non-Probation | Within Deviation | Total Deviation |
| $\Sigma x_1$ | 1,739 | 1,955 | 3,694 | | | | |
| $\Sigma x_2$ | 2,471 | 2,860 | 5,331 | | | | |
| $\Sigma x_4$ | 1,328 | 1,603 | 2,931 | | | | |
| $\Sigma x_5$ | 1,649 | 1,948 | 3,597 | | | | |
| $\Sigma x_7$ | 1,675 | 2,781 | 4,456 | | | | |
| $\Sigma x_8$ | 1,802 | 1,769 | 3,571 | | | | |
| $\Sigma x_9$ | 3,677 | 3,813 | 7,490 | | | | |
| $\Sigma x_{10}$ | 1,813 | 1,797 | 3,610 | | | | |
| $\Sigma x_{11}$ | 2,124 | 2,114 | 4,238 | | | | |
| $\Sigma x_{12}$ | 2,545 | 2,609 | 5,154 | | | | |
| $\Sigma x_{13}$ | 2,826 | 2,694 | 5,520 | | | | |
| $\Sigma x_{14}$ | 1,735 | 2,166 | 3,901 | | | | |
| $\Sigma x_1^2$ | 77,087 | 97,135 | 174,222 | 5,084.1191 | 6,134.4048 | 11,218.5239 | 11,773.9524 |
| $\Sigma x_2^2$ | 153,577 | 200,184 | 353,761 | 8,199.8334 | 5,431.6191 | 13,631.4525 | 15,432.8929 |
| $\Sigma x_4^2$ | 48,338 | 66,603 | 114,941 | 6,347.9048 | 5,484.8833 | 11,832.7881 | 12,670.0358 |
| $\Sigma x_5^2$ | 69,009 | 94,184 | 163,193 | 4,266.1191 | 3,833.9048 | 8,100.0239 | 9,164.3215 |
| $\Sigma x_7^2$ | 88,765 | 204,245 | 293,010 | 21,964.4048 | 20,103.0715 | 42,067.4763 | 56,629.8096 |
| $\Sigma x_8^2$ | 82,324 | 84,457 | 166,781 | 5,009.6191 | 9,948.4048 | 14,958.0239 | 14,970.9881 |
| $\Sigma x_9^2$ | 333,027 | 361,249 | 701,766 | 11,114.4048 | 15,083.0715 | 26,197.4763 | 33,907.6667 |
| $\Sigma x_{10}^2$ | 84,667 | 85,565 | 170,232 | 6,405.8334 | 8,679.0715 | 15,084.9049 | 15,087.9524 |
| $\Sigma x_{11}^2$ | 118,708 | 116,134 | 234,842 | 11,294.2858 | 9,729.3334 | 21,023.6192 | 21,024.8096 |
| $\Sigma x_{12}^2$ | 170,185 | 175,231 | 345,416 | 15,970.1191 | 13,162.4048 | 29,132.5239 | 29,181.2858 |
| $\Sigma x_{13}^2$ | 197,464 | 179,792 | 377,256 | 7,314.5715 | 6,991.1429 | 14,305.7144 | 14,513.1429 |
| $\Sigma x_{14}^2$ | 83,759 | 126,094 | 209,853 | 12,086.9762 | 14,390.2858 | 26,477.2620 | 28,688.7024 |

| Symbol | Raw Scores | | | Deviation Scores | | | |
|---|---|---|---|---|---|---|---|
| | Probation | Non-Probation | Both | Probation | Non-Probation | Within Deviation | Total Deviation |
| $\Sigma x_1 x_2$ | 106,113 | 136,550 | 242,663 | 3,801.8334 | 3,423.8096 | 7,225.6430 | 8,225.9286 |
| $\Sigma x_1 x_4$ | 56,792 | 76,356 | 133,148 | 1,806.4762 | 1,740.1667 | 3,546.6429 | 4,253.7858 |
| $\Sigma x_1 x_5$ | 70,578 | 92,940 | 163,518 | 2,301.5477 | 2,265.2381 | 4,566.7858 | 5,335.6429 |
| $\Sigma x_1 x_7$ | 68,802 | 132,168 | 200,970 | -550.9761 | 2,719.0715 | 2,168.0954 | 5,012.0953 |
| $\Sigma x_1 x_{14}$ | 74,861 | 105,953 | 180,814 | 3,023.7381 | 5,130.8572 | 8,154.5953 | 9,262.8810 |
| $\Sigma x_2 x_4$ | 84,008 | 113,634 | 197,642 | 5,877.3334 | 4,477.3334 | 10,354.6668 | 11,628.1786 |
| $\Sigma x_2 x_5$ | 100,804 | 135,830 | 236,634 | 3,787.8334 | 3,180.4762 | 6,968.3096 | 8,352.9643 |
| $\Sigma x_2 x_7$ | 98,674 | 193,386 | 292,060 | 128.1667 | 4,013.1429 | 4,141.3096 | 9,263.1429 |
| $\Sigma x_2 x_{14}$ | 109,710 | 154,749 | 264,459 | 7,634.1667 | 7,254.7143 | 14,888.8810 | 16,884.8215 |
| $\Sigma x_4 x_5$ | 55,475 | 77,382 | 132,857 | 3,335.1905 | 3,033.3333 | 6,368.5238 | 7,347.3929 |
| $\Sigma x_4 x_7$ | 54,339 | 109,564 | 163,903 | 1,377.0953 | 3,422.5000 | 4,799.5953 | 8,420.4286 |
| $\Sigma x_4 x_8$ | 58,432 | 67,443 | 125,875 | 1,454.4762 | -73.8333 | 1,380.6429 | 1,272.6072 |
| $\Sigma x_4 x_9$ | 116,115 | 146,201 | 262,316 | -148.2380 | 671.5000 | 523.2620 | 968.5000 |
| $\Sigma x_4 x_{10}$ | 56,549 | 68,927 | 125,476 | -776.3333 | 341.5000 | -434.8333 | -487.2142 |
| $\Sigma x_4 x_{11}$ | 70,311 | 81,809 | 152,120 | 3,152.1429 | 1,124.6667 | 4,276.8096 | 4,244.0715 |
| $\Sigma x_4 x_{12}$ | 79,069 | 98,729 | 177,798 | -1,401.4761 | -847.8333 | -2,249.3094 | -2,039.7857 |
| $\Sigma x_4 x_{13}$ | 91,675 | 101,848 | 193,523 | 2,319.5715 | -973.0000 | 1,346.5715 | 914.4286 |
| $\Sigma x_4 x_{14}$ | 62,007 | 88,285 | 150,292 | 7,147.9524 | 5,616.0000 | 12,763.9524 | 14,174.9643 |
| $\Sigma x_5 x_7$ | 66,714 | 132,058 | 198,772 | 950.3096 | 3,072.5715 | 4,022.8811 | 7,959.7143 |
| $\Sigma x_5 x_{14}$ | 72,357 | 105,270 | 177,627 | 4,237.5953 | 4,808.8572 | 9,046.4525 | 10,580.6072 |
| $\Sigma x_7 x_8$ | 71,658 | 119,699 | 191,357 | -207.4761 | 2,565.9286 | 2,358.4525 | 1,923.9524 |
| $\Sigma x_7 x_9$ | 142,171 | 248,035 | 390,206 | -4,471.2619 | -4,440.0714 | -8,911.3333 | -7,120.6666 |
| $\Sigma x_7 x_{10}$ | 70,407 | 122,138 | 192,545 | 1,897.1667 | 3,150.9286 | 5,048.0953 | 1,043.0953 |
| $\Sigma x_7 x_{11}$ | 84,590 | 141,884 | 226,474 | -117.1428 | 1,907.0000 | 1,789.8572 | 1,658.1905 |
| $\Sigma x_7 x_{12}$ | 100,962 | 169,598 | 270,560 | -535.0238 | -3,155.0714 | -3,690.0952 | -2,847.4285 |
| $\Sigma x_7 x_{13}$ | 111,888 | 179,658 | 291,546 | -815.5714 | 1,276.7143 | 461.1429 | -1,276.8571 |
| $\Sigma x_7 x_{14}$ | 73,611 | 149,427 | 223,038 | 4,417.5477 | 6,006.8572 | 10,424.4049 | 16,099.2381 |

Raw and Deviation Scores for 1948-1949 Liberal Arts Students

| Symbol | Raw Scores | | | Deviation Scores | | | |
|---|---|---|---|---|---|---|---|
| | Probation | Non-Probation | Both | Probation | Non-Probation | Within Deviation | Total Deviation From Gen. Mean |
| $\Sigma x_1$ | 2,505 | 2,805 | 5,310 | | | | |
| $\Sigma x_2$ | 3,862 | 4,328 | 8,190 | | | | |
| $\Sigma x_4$ | 1,964 | 2,455 | 4,419 | | | | |
| $\Sigma x_5$ | 2,522 | 2,886 | 5,408 | | | | |
| $\Sigma x_7$ | 2,461 | 4,270 | 6,731 | | | | |
| $\Sigma x_8$ | 2,141 | 1,926 | 4,067 | | | | |
| $\Sigma x_9$ | 4,855 | 4,898 | 9,753 | | | | |
| $\Sigma x_{10}$ | 2,954 | 3,137 | 6,091 | | | | |
| $\Sigma x_{11}$ | 3,406 | 3,514 | 6,920 | | | | |
| $\Sigma x_{12}$ | 4,487 | 4,571 | 9,058 | | | | |
| $\Sigma x_{13}$ | 3,559 | 3,530 | 7,089 | | | | |
| $\Sigma x_1^2$ | 103,435 | 130,049 | 233,484 | 5,387.7344 | 7,111.1094 | 12,498.8438 | 13,201.9687 |
| $\Sigma x_2^2$ | 243,434 | 305,820 | 549,254 | 10,386.4375 | 13,139.0000 | 23,525.4375 | 25,221.9687 |
| $\Sigma x_4^2$ | 65,704 | 104,219 | 169,923 | 5,433.7500 | 10,046.7344 | 15,480.4844 | 17,363.9297 |
| $\Sigma x_5^2$ | 105,252 | 137,452 | 242,704 | 5,869.4375 | 7,311.4375 | 13,180.8750 | 14,216.0000 |
| $\Sigma x_7^2$ | 120,971 | 330,732 | 451,703 | 26,337.8594 | 45,842.9375 | 72,180.7969 | 97,747.0547 |
| $\Sigma x_8^2$ | 80,757 | 68,628 | 149,385 | 9,133.8594 | 10,667.4375 | 19,801.2969 | 20,162.4297 |
| $\Sigma x_9^2$ | 387,225 | 397,544 | 784,769 | 18,927.7344 | 22,693.9375 | 41,621.6719 | 41,636.1172 |
| $\Sigma x_{10}^2$ | 150,492 | 164,919 | 315,411 | 14,146.4375 | 11,156.9844 | 25,303.4219 | 25,565.0547 |
| $\Sigma x_{11}^2$ | 200,236 | 215,930 | 416,166 | 18,972.9375 | 22,989.4375 | 41,962.3750 | 42,053.5000 |
| $\Sigma x_{12}^2$ | 329,665 | 349,391 | 679,056 | 15,084.2344 | 22,921.6094 | 38,005.8438 | 38,060.9687 |
| $\Sigma x_{13}^2$ | 213,167 | 213,312 | 426,479 | 15,253.2344 | 18,610.4375 | 33,863.6719 | 33,870.2422 |

## Liberal Arts

| Symbol | Raw Scores | | | Deviation Scores | | | |
|---|---|---|---|---|---|---|---|
| | Probation | Non-Probation | Both | Probation | Non-Probation | Within Deviation | Total Deviation From Gen. Mean |
| $\Sigma x_1 x_2$ | 154,283 | 192,969 | 347,252 | 3,121.9063 | 3,280.8750 | 6,402.7813 | 7,494.9687 |
| $\Sigma x_1 x_4$ | 78,469 | 110,224 | 188,693 | 1,596.8125 | 2,625.9532 | 4,222.7657 | 5,373.5469 |
| $\Sigma x_1 x_5$ | 101,098 | 128,484 | 229,582 | 2,385.3438 | 1,996.0313 | 4,381.3751 | 5,234.5000 |
| $\Sigma x_1 x_7$ | 96,003 | 192,352 | 288,355 | -322.0781 | 5,205.9063 | 4,883.8282 | 9,123.6719 |
| $\Sigma x_2 x_4$ | 124,373 | 175,663 | 300,036 | 5,857.8750 | 9,643.6250 | 15,501.5000 | 17,289.0469 |
| $\Sigma x_2 x_5$ | 156,930 | 202,932 | 359,862 | 4,743.0625 | 7,766.2500 | 12,509.3125 | 13,834.5000 |
| $\Sigma x_2 x_7$ | 145,339 | 296,072 | 441,411 | -3,166.9687 | 7,313.2500 | 4,146.2813 | 10,732.1719 |
| $\Sigma x_4 x_5$ | 80,784 | 117,022 | 197,806 | 3,390.1250 | 6,316.8438 | 9,706.9688 | 11,103.2500 |
| $\Sigma x_4 x_7$ | 75,227 | 172,141 | 247,368 | -294.9375 | 8,346.4688 | 8,051.5313 | 14,990.7422 |
| $\Sigma x_4 x_8$ | 63,569 | 72,064 | 135,633 | -2,132.9375 | -1,816.1562 | -3,949.0937 | -4,773.8203 |
| $\Sigma x_4 x_9$ | 147,221 | 188,424 | 335,645 | -1,766.8125 | 539.7813 | -1,227.0312 | -1,062.0859 |
| $\Sigma x_4 x_{10}$ | 91,316 | 117,390 | 208,706 | 665.1250 | -2,943.3593 | -2,278.2343 | -1,576.2578 |
| $\Sigma x_4 x_{11}$ | 108,283 | 137,773 | 246,056 | 3,761.3750 | 2,978.1563 | 6,739.5313 | 7,153.8125 |
| $\Sigma x_4 x_{12}$ | 135,640 | 177,182 | 312,822 | -2,054.8125 | 1,841.2969 | -213.5156 | 108.7031 |
| $\Sigma x_4 x_{13}$ | 107,180 | 133,969 | 241,149 | -2,036.8125 | -1,439.5937 | -3,476.4062 | -3,587.6484 |
| $\Sigma x_5 x_7$ | 98,874 | 198,242 | 297,116 | -104.7812 | 5,691.6875 | 5,586.9063 | 12,731.2500 |
| $\Sigma x_7 x_8$ | 83,827 | 129,739 | 213,566 | 1,498.8594 | 1,238.6875 | 2,737.5469 | -301.0078 |
| $\Sigma x_7 x_9$ | 187,272 | 324,883 | 512,155 | 582.0782 | -1,905.4375 | -1,323.3593 | -715.6484 |
| $\Sigma x_7 x_{10}$ | 114,662 | 207,577 | 322,239 | 1,071.4688 | -1,719.7187 | 648.2499 | 938.0547 |
| $\Sigma x_7 x_{11}$ | 131,837 | 233,471 | 365,308 | 865.6563 | -978.6875 | -113.0312 | 1,413.3125 |
| $\Sigma x_7 x_{12}$ | 170,471 | 307,779 | 478,250 | -2,068.1718 | 2,807.5938 | 739.4220 | 1,926.5781 |
| $\Sigma x_7 x_{13}$ | 138,312 | 242,297 | 380,609 | -1,457.3282 | 6,779.8125 | 8,237.1407 | 7,827.2991 |

Raw and Deviation Scores for 1948-49 All University Students

| Symbol | Raw Scores | | | Deviation Scores | |
|---|---|---|---|---|---|
| | Probation | Non-Probation | Both | Within group Within College | Within College |
| $\Sigma x_1$ | 5,453 | 6,172 | 11,625 | | |
| $\Sigma x_2$ | 8,275 | 9,510 | 17,785 | | |
| $\Sigma x_4$ | 4,373 | 5,449 | 9,822 | | |
| $\Sigma x_5$ | 5,469 | 6,408 | 11,877 | | |
| $\Sigma x_7$ | 5,928 | 9,445 | 15,373 | | |
| $\Sigma x_1^2$ | 225,637 | 285,644 | 511,281 | 33,809.6844 | 35,629.1411 |
| $\Sigma x_2^2$ | 510,611 | 656,576 | 1,167,187 | 39,341.2442 | 58,473.0156 |
| $\Sigma x_4^2$ | 151,057 | 228,271 | 379,328 | 37,457.0592 | 41,502.4188 |
| $\Sigma x_5^2$ | 224,881 | 302,480 | 527,361 | 29,959.2722 | 33,089.9748 |
| $\Sigma x_7^2$ | 313,682 | 714,249 | 1,027,931 | 152,980.5366 | 200,274.0027 |
| $\Sigma x_1 x_2$ | | | | 19,126.0376 | 22,275.4173 |
| $\Sigma x_1 x_4$ | | | | 9,343.2619 | 12,041.2527 |
| $\Sigma x_1 x_5$ | | | | 11,810.1609 | 14,186.6229 |
| $\Sigma x_1 x_7$ | | | | 7,710.1137 | 16,541.9072 |
| $\Sigma x_2 x_4$ | | | | 30,849.0868 | 35,465.6122 |
| $\Sigma x_2 x_5$ | | | | 27,133.1154 | 31,250.4776 |
| $\Sigma x_2 x_7$ | | | | 11,667.8476 | 26,806.9918 |
| $\Sigma x_4 x_5$ | | | | 21,308.2393 | 24,846.5896 |
| $\Sigma x_4 x_7$ | | | | 15,154.9999 | 27,994.0441 |
| $\Sigma x_5 x_7$ | | | | 9,947.0640 | 23,285.0576 |