Journal of Natural Sciences Research
ISSN 2224-3186 (Paper) ISSN 2225-0921 (Online)
Vol.3, No.5, 2013
www.iiste.org

An Application of Discriminant Analysis On University Matriculation Examination Scores For Candidates Admitted Into Anamabra State University

Okoli C.N
Department of Statistics
Anambra State University, Uli

ABSTRACT:

The study was carried out on the university matriculation examination (UME) scores of candidates admitted in the department of Industrial Chemistry for 2009/2010 session with the aim of using discriminant function to achieve a sharper discrimination between those "accepted" and those not accepted" in the department. The data for this study was collected from the Anambra State University admission office. The data collected were analysed using average scores, Hotellings T^2 distribution and discriminant analysis.

The result of the analysis showed that the average scores of those candidates accepted using the four university matriculation examination UME) subjects in higher compared to the average score of not accepted candidates. The hotellings T^2 distribution used showed that the population mean vectors of the two groups (accepted and not accepted candidates) are different. Discriminant function found for 'accepted' and 'not accepted' candidates and classification rule also used showed that they are candidates that are wrongly classified or misclassified.
Keywords UME scores, mean, hotellings T^2 distribution, Discriminant analysis, discriminant function, classification rule.

INTRODUCTION

The university Matriculation examination (UME) was introduced by the joint Admission and matriculation Board (JAMB) in 1978 to be an avenue through which candidates wishing to get into the University can pass through. The university matriculation Examination is based on four subjects. The joint Admission and Matriculation Board (JAMB) admits candidates into different institution they applied for based on their performance. Candidates that score higher in different subject with higher aggregate will have better chances of admission than those with lower scores in different UME subjects. The admission is based on this trend till the number of candidates required for admission is complete while the other candidates that applied will not be admitted except through other means like supplementary admission which is only considered if the number of candidates in a given department is not up to the quota they needed.

The need for this study is that some of candidates admitted may not be on merit, using discriminant function and classification rule will help to fish out those candidates that are wrongly classified or misclassified. That is to say that it will successfully help discriminate between those accepted and those in accepted. Ogum (2002) used the method of multivariate analysis in analyzing the scores of candidates admitted into the university of Nigeria medical school in the 1975/1976 academic session and constructed a discriminant function that successfully discriminate between those "admitted" and those not "admitted"

Okpara (2001) Applied the method of discriminant analysis in analyzing the scores of candidates admitted into school of physical sciences in 2000/2001 session and constructed a discriminant function that successfully discriminate between those admitted and those not admitted.

Wagle, B (1968) applied the method of multivariate analysis to a study of examination scores at the London school of economics and arrived at a procedure for predicting the result of the part two examination based on marks obtained by candidates in the part one examination.

MATERIALS AND METHODS

The data for this research was secondary data collected from the Anambra State University admission office and the methods adopted for the analysis were average score; hotellings T^2 distribution discriminant analysis and classification rule.

THE ARITHMETIC MEAN

The arithmetic mean of a set of n observation is defined as.

The arithmetic mean = the sum of all the value in the population

Number of items in the population

That is `

$$\bar{X} = \frac{\sum_{i=1}^{n} x_i}{N}$$

HOTELLINGS T^2 DISTRIBUTION

Let x ~ Np (μ,Σ) and D~ Wp (Σ, V), D>o), and X, D independent. Then $T^2 = VX^1D^{-1}X$, V>p is known as the Hotellings T^2 based on V degree of freedom Onyeagu (2003). Hotellings T^2 is the multivariate generalisation of student's t. Hotellings T^2 is useful in all problems in multivariate analysis where one would use the t statistic in univariate analysis. It is also used in some situations for which there is no univariate counterpart.

PROCEDURE FOR THE TEST OF HYPOTHESIS

Ho: $\mu_1 = \mu_2$ (the mean vectors of the two groups are equal)

Hi: $\mu_1 = \mu_2|$(the mean vectors of the two groups are not equal)

α Is the level of significance using = 0.05 α

THE TEST STATISTIC OF HOTELLINGS T^2 DISTRIBUTION FOR TWO SAMPLE IS GIVEN AS

$$\frac{n_1 n_2}{n_1+n_2}\left(\bar{X}_1 - \bar{X}_2\right)^1 Sp^{-1}\left(\bar{X}_1 - \bar{X}_2\right)$$

$$T^2 \sim \frac{Pv}{V-P+1} \; F_{p,\, v-P+1}$$

Where n_1 = sample size of population I

n_2 = sample size of population 2

S_P = Spooled sample

V = Degree of freedom (n_1+n_2-2)

P = Number of variables

Decision Rule

The null hypothesis is rejected at level significance $\propto$

If $T^2 > F_{p,}\ v\text{-}P+1$, otherwise accept.

DISCRIMINANT ANALYSIS

Morrison, Donald F. (1967) Discriminant analysis is a powerful statistical tools that is concerned with the problem of classification. This problem of classification arises when an investigator makes a number of measurements on an individual and wishes to classify the individual into one of severe categories or population groups on the basis of these measurements.

Johnson and Wichern (1992) defined discriminant analysis and classification as multivariate techniques concerned with separating distinct set of objects (or observations) and with allocating new objects (observations) to previously defined groups.

A discriminant function has property that is better than any other linear function it will discriminate between any two chosen class, such as those candidate that qualified for admission and those not qualified for admission.

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

THE FISHER'S LINEAR DISCRIMINANT FUNCTION IS GIVEN AS.

$$\bar{Y} \quad = \quad (\bar{X}_1 - \bar{X}_2)^1 \; S_p^{-1} X$$

$$\bar{Y} \quad = \quad (\bar{X}_1 - \bar{X}_2)^1 \; S_p^{-1} \bar{X}_1$$

$$\bar{Y} \quad = \quad (\bar{X}_1 - \bar{X}_2)^1 \; S_p^{-1} \bar{X}_2$$

Then the midpoint for the interval between $\bar{Y}_1 + \bar{Y}_2$ is

$$\bar{Y}_c = \frac{\bar{Y}_1 + \bar{Y}_2}{2}$$

This is known as the critical value. This is what is used as the cut off point for the assignment

CLASSIFICATION RULE:
The classification rule is as follows; Assign the individual to group 1 if the discriminant function of Y of the individual is greater than the critical value of Y and to group 2 otherwise.

THE RESULTS OF THE ANALYSIS
From table 1 and 2 (see appendix), the following results were obtained.

$$\left.\begin{aligned} \bar{X}_1 E &= 56.05 \\ \bar{X}_2 M &= 53.82 \\ \bar{X}_3 C &= 56.24 \\ \bar{X}_4 P &= 59.11 \end{aligned}\right\} \text{----------------------- (1)}$$

$$\left.\begin{aligned} \bar{X}_1 E^1 &= 47.08 \\ \bar{X}_2 M^1 &= 47.52 \\ \bar{X}_3 C^1 &= 48.37 \\ \bar{X}_4 P^1 &= 54.68 \end{aligned}\right\} \text{-------------------- (2)}$$

$$Sp^{-1} = \begin{pmatrix} 0.01421 & -0.00080 & -0.00499 & -000042 \\ -0.00080 & 0.01351 & -0.00483 & 0.00170 \\ -0.00499 & -0.00483 & 0.01789 & -0.00439 \\ -0.00042 & -0.000617 & -0.0439 & -0.01640 \end{pmatrix} \text{-------(3)}$$

Where X_1E is the mean scores of candidates in use of English of group 1, X_2M is the mean scores of calculates in mathematics of group 1, X_3C is the mean scores of candidates in chemistry of group 1 and x_4p is the mean scores of candidates in physics of group 1

Also

X_1E^1, X_2M^1, X_3C^1 and $X_4P^1$is the mean scores of candidates in English, mathematics, chemistry and physics respectively of group 2.

TESTING FOR EQUALITY OF TWO MEAN VECTORS OF THE GROUPS USING HOTELLINGS T^2 DISTRIBUTION TEST STATISTIC.

$$\frac{n_1 n_2}{n_1 + n_2} \left(\bar{X}_1 - \bar{X}_2 \right)^1 S_P^{-1} \left(\bar{X}_1 - \bar{X}_2 \right)$$

$$T^2 \sim \frac{Pv}{V-P+1} F^{\forall}_{p,\, v-P+1}$$

$$T^2_{(12.220)} \sim F^{0.05}_{9.989,}$$

We reject the null hypothesis that the population mean vectors of the groups are equal and conclude that the population mean vectors of the groups are not equal.

THE FISHER LINEAR DISCRIMINANT FUNCTION

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

$$\text{i.e } Y = 0.0813x_1 + 0.0475x_2 + 0.0460x_3 + 0.0450x_4$$

$$\bar{Y}_1 = 12.3562; \quad \bar{Y}_2 = 10.7704$$

$$\bar{Y}_c = 11.5633$$

CLASSIFICATION RULE

Since the discriminant function cut if point is 11.5633, Assign an individual to group 1 (i.e accepted candidates) if the discriminant function is greater than 11.5633, and group 2 (ie not accepted candidates) if the discriminant function equals to 11.5633 and below

From group 1 of table 3 those that are misclassified or wrongly classified are candidates numbers 11, 14,15,16,17,18,19,20,21,22,25,27,33,34,37 and 54. From group 2 of table 3 those that are misclassified are candidates numbers 20,21,22, 23,24,25,26 and 46.

RESULTS AND DISCUSSION

The discriminant function found for accepted and not accepted candidates successfully discriminated those candidates accepted from those not accepted. It agrees with the result of the study by Ogum (2002) that analysed the scores of candidates admitted into the University of Nigeria medical schools in the 1975/1976 academic session in which a discriminant function constructed successfully discriminate between those 'admitted' and those not 'admitted.' It also agrees with the result of Okpara (2001) that analysed the scores of candidates admitted into school of physical science in 2000/2001 session using discriminant function that successfully discriminated between those admitted and those not admitted. Hotellings T^2 distribution used reject the null hypothesis that the population mean vectors of the two groups are equal and conclude that the population mean vectors are different since T^2 (12.220) > F9.989. Average scores of those 'accepted" in the four subjects is higher compared to average scores of those not accepted.

CONCLUSION

The fundamental finding of this study is that discriminant function successfully discriminated those candidates "accepted" from those candidates "not accepted". Therefore, those who were misclassified need to be reclassified into the appropriate group they rightly belong to if the post UME is not considered.

REFERENCES

Fisher R.A. (1936): "The use of Multiple measurements in Taxonomic problems" Annals of Eugenies page 179-188

Johnson R.A. and Wichern D.W. (1992): Applied Multivariate Statistical Analysis. Englewood cliffs new jersey.

Morrison, Donald F. (1967). Multivarate stastical methods. New York McGraw-Hill Book Company

Ogum G.E.O. (2002) An application of discriminant analysis in university Admission (A case of the university of medical school, 1975/76), Introduction to methods of multivariate Analysis page 119-134

Okpara P.I. (2001), A statistical discriminant Approach (A case of school of physical) B.Sc thesis, department of statistics, IMSU.

Onyeagu S.I. (2003): A first course in Multivariate statistical Analysis. Page 207-265.

Wagle, B (1968): "An Application of Discriminiant Analysis: Predicting the part II classification based on part 1 Marks" Ph.D Thesis, London University.

APPENDIX Table 1: U.M.E SCORES FOR THE ACCEPTED CANDIDATES (GROUP 1)

S/N	English (X_1E)	Mathematics (X_2M)	Chemistry (X_3C)	Physics (X_4P)
1.	60	72	69	70
2.	74	68	62	62
3.	41	73	65	74
4.	56	49	64	74
5.	73	42	66	57
6.	41	61	63	68
7.	65	56	52	49
8.	59	51	53	66
9.	56	55	53	52
10.	64	51	48	52
11.	51	37	53	62
12.	48	61	57	68
13.	50	58	52	56
14.	51	58	49	49
15.	47	48	48	52
16.	42	56	44	52
17.	40	45	49	55
18.	42	40	48	58
19.	54	38	42	49
20.	40	45	35	62
21.	48	37	36	60
22.	43	44	45	49
23.	52	68	55	53
24.	53	43	69	60
25.	35	42	52	57
26.	58	59	61	61
27.	40	46	48	69
28.	50	58	56	52
29.	60	72	70	69
30.	73	42	57	66
31.	53	51	59	66
32.	61	48	57	68

TABLE 1 CONTINUES

Table 1: U.M.E SCORES FOR THE ACCEPTED CANDIDATES (GROUP 1)

S/N	English (X_1E)	Mathematics (X_2M)	Chemistry (X_3C)	Physics (X_4P)
33.	42	56	44	52
34.	40	45	49	55
35.	69	43	53	60
36.	58	61	61	59
37.	42	38	49	54
38.	56	55	52	53
39.	70	60	55	67
40.	69	55	70	65
41.	60	69	70	72
42.	73	41	64	74
43.	61	63	68	41
44.	58	51	52	56
45.	59	66	48	52
46.	62	64	53	51
47.	60	55	60	67
48.	52	61	53	52
49.	58	61	59	60
50.	60	43	69	53
51.	58	50	56	52
52.	60	72	60	51
53.	69	43	59	60
54.	50	50	58	52
55.	72	60	70	66
56.	59	51	53	66
57.	48	57	56	48
58.	69	53	43	60
59.	65	56	70	65
60.	64	51	53	66
61.	64	65	56	57
62.	68	62	74	62

Table 2: U.M.E SCORES FOR THE CANDIDATES NOT ACCEPTED (GROUP 2)

S/N	English (X_1E^1)	Mathematics (X_2M^1)	Chemistry (X_3C^1)	Physics (X_4P^1)
1.	39	48	44	62
2.	38	58	45	52
3.	51	39	54	47
4.	40	37	42	68
5.	49	61	38	62
6.	46	61	38	41
7.	47	51	38	47
8.	47	44	40	47
9.	39	35	47	65
10.	46	38	44	49
11.	41	42	49	53
12.	45	39	45	62
13.	45	42	34	57
14.	45	54	36	61
15.	56	50	54	52
16.	45	51	43	49
17.	47	42	51	62
18.	45	38	42	62
19.	51	42	47	57
20.	70	56	41	62

21.	50	63	55	62
22.	44	58	72	62
23.	54	48	60	57
24.	60	63	60	60
25.	63	67	36	55
26.	73	59	65	63
27.	47	39	63	62
28.	41	57	67	68
29.	43	44	45	68
30.	42	44	55	61
31.	42	54	62	65

TABLE CONTINUES

Table 2: U.M.E SCORES FOR THE CANDIDATES NOT ACCEPTED (GROUP 2)

S/N	English (X_1E^1)	Mathematics (X_2M^1)	Chemistry (X_3C^1)	Physics (X_4P^1)
32.	41	48	58	57
33.	54	51	49	49
34.	35	56	56	55
35.	41	47	42	70
36.	47	42	45	65
37.	55	43	47	54
38.	52	47	44	55
39.	43	47	55	57
40.	45	46	46	57
41.	47	40	51	47
42.	41	53	45	42
43.	44	47	44	47
44.	44	44	35	39
45.	42	36	42	45
46.	61	45	54	54
47.	41	51	41	42
48.	45	68	57	41
49.	42	43	62	65
50.	41	58	42	41
51.	56	35	49	49
52.	47	56	42	55
53.	51	47	56	41
54.	47	42	45	65
55.	55	43	47	44
56.	52	47	47	44
57.	45	46	56	46
58.	39	62	44	48
59.	58	38	45	52
60.	39	31	51	47
61.	38	49	38	41
62.	40	37	42	68

TABLE 3: DISCRIMINANT FUNCTION SCORES FOR GROUP 1 AND GROUPS 2

S/N	ACCEPTED CANDIDATES (GROUP 1)	NOT ACCEPTED CANDIDATES (GROUP 2)
1.	14.6220	10.2647
2.	14.8882	10.2544
3.	13.1208	10.2544
4.	13.1543	10 5978
5.	13.5309	10.3267
6.	12.1888	10.2303
7.	I2.5415	10.1986
8.	1 2.6272	10.1881
9.	11.94133	9.7822
10.	12.1737	10.0038
11.	11.1318*	9.7833
12.	12 4819	9.8650
13.	1 1.7320	9.8745
14.	1 1.3603*	1 1.4525
15.	10.64 91*	1 1.2458
16.	10.4386*	10.6320
17.	10.1 185*	10.5381
18.	10.1326*	10.4155
19.	10.3322*	10.5923
20.	9.7895 *	13.6710*
21.	10.0159*	13.1595*
22.	9.8609*	1 1.8822*
23.	12.3726	12.2202*
24.	12.2254	12 2265*
25.	9.8435*	13.7694 *
26.	13.0689	14.4704*
27.	1 0.7500*	11.5456
28.	11.7360	11.1708

TABLE 3 CONTINUES

S/N	ACCEPTED CANDIDATES (GROUP 1)	NOT ACCEPTED CANDIDATES (GROUP 2)
29.	14.6230	1 1.1759
30.	13.5219	10.1016
31.	12.4154	1 1.1 126
32.	12.9213	10.8463
33.	10.4386*	1 1.2717
34.	10.1 185*	10.5565
35.	12.7902	10.6478
36.	13.0739	10.8111
37.	9.9036*	11.1060
38.	11.9423	10.9590
39.	14.0860	10.8234
40.	14.3672	10.5245
41.	14.6155	10.1821
42.	12.1564	9.8108
43.	12.9248	9.9487
44.	12.0499	9.0322
45.	12.4797	9.0816
46.	12.8136	12.0108*
47.	13.2655	9.5318
48.	11.9031	11.3555
49.	13.0269	11.2341
50.	12.4795	10.2253
51.	12.0064	10.6743
52.	13.3530	10.8881
53.	13.0662	10.7998
54.	1 1.4480*	10.8111

TABLE 3 CONTINUES

S/N	**ACCEPTED CANDIDATES (GROUP 1)**	**NOT ACCEPTED CANDIDATES (GROUP 2)**
55.	14.8936	10.6560
56.	12.6272	10.6021
57.	11.8979	10.4895
58.	12.8052	10.2997
59.	14.0895	10.9304
60.	13.0337	9.1042
61.	13.4317	9.0099
62.	14.6674	10.0015

Where those with * mark are the misclassified once.