International Journal of Innovation and Scientific Research
ISSN 2351-8014 Vol. 17 No. 2 Aug. 2015, pp. 241-245

http://www.ijisr.issr-journals.org/

ON THE USE OF DISCRIMINANT ANALYSIS IN CLASSIFICATION OF THE MODE OF DELIVERY OF AN EXPECTANT MOTHER

U.I. Usman[1] and O.S. Balogun[2]

[1]Monitoring and Compliance Unit, Abuja Enterprise Agency, Entrepreneurial Complex, Plot 74 Sector Centre A, Cadastral Zone B15, Jahi, Abuja, Nigeria

[2]Department of Statistics and Operations Research, Modibbo Adama University of Technology, P.M.B. 2076, Yola, Adamawa State, Nigeria

ABSTRACT: The aim of the study was to obtain a discriminant function that can be used to classify the mode of delivery of pregnant women using some variables. Data from Health Records of 184 Pregnant women who delivered at the General Hospital, Wuse were used. The data consist of Mother's Weight, Height, Age and Baby's Weight, Baby's Gender and mode of delivery (Natural birth and Caesarian Section). This indicates that the Mother's Age and Mother's weight significantly affects the discrimination between the two groups. The Discriminant function D(X), which can be used in classifying the mode of delivery of women was obtained and used. The discriminant analysis gave a correct classification rate of 64.7% and misclassification rate of 35.3%.

KEYWORDS: Natural birth, Caesarian section, Discriminant function, Classification.

1 INTRODUCTION

Child birth poses considerable risk to the lives of both mother and child particularly in situations where complication arises. Child birth is defined as the complete expulsion or extraction of a fetus from its mother.

Child birth is preceded by a period known as the Gestation period. It has been of interest to researchers to know the mode of delivery a mother is likely to use. Under normal conditions, a mother is expected to give birth by natural birth otherwise known as safe delivery, but in certain cases complications may arise leading to the use of Caesarian section. Caesarian section poses considerable risk.

West/Central Africa accounts for more than 30% of global maternal deaths, and 162, 000 women died of pregnancy or childbirth related causes in 2005. The maternal mortality ration is substantially higher here than in any other region, at 1100 maternal deaths per 100, 000 live births. Furthermore, no discernible progress has been made in reducing the ratio since 1990. Of the 23 countries in the region with comparable estimates every country but Cape Verde has an MMR of at least 500, and a third of these countries have an MMR of 1, 000 or greater. Almost two thirds of maternal deaths in the region occur in the Democratic Republic of Congo, Niger and Nigeria, which together account for approximately 20 per cent of all maternal deaths worldwide [1]. Several factors influence the high rate of maternal mortality in Nigeria, but the most common causes are lack of access to ante – natal care, inadequate access to skilled birth attendees, delays in the treatment of complications of pregnancy, poverty and harmful traditional practices.

Discriminant analysis has had its earliest and most widespread educational research applications in the areas of vocational and careers development. Because education prepares people for a variety of positions in the occupational structures prevalent in their societies, an important class of education research studies is concerned with testing of theories about the causes of occupational placements and/or the estimation of production equations for allocating positions or anticipating such allocation. Discriminant analysis is a descriptive procedure of separation in which linear functions of the variables are used to describe or elucidate the differences between the two or more groups. That is, the aim of this analysis includes identifying the relative contribution of say, p variables to separation of groups and finding the optimal plane on which the points can be projected to best illustrate the configuration of the groups [2]. The classification of objects to groups is usually thought of as partition of the objects into subsets in which the members are more similar. Classifying individuals into groups such that there is a relative homogeneity between the groups and heterogeneity between the groups is a problem which has been considered for many years [3].

[4] used discriminant function in classifying drug offenders into groups (Drug Peddlers and Non-Drug Peddlers) using some variables (Type of Exhibit, Weight of Exhibit, Age and Gender) instead of the oral evidences used by National Drug Enforcement Agency to classify drug offenders into drug Peddlers and Non-Peddlers in Kwara State.

2 THE DISCRIMINANT MODEL

The elements of the discriminant models are given as

$$Z = a + W_1X_1 + W_2X_2 + \ldots + W_kX_k$$

where

$$\text{Z} = \text{discriminant score}$$

$$\text{a} = \text{discrimanat constant}$$

$$\text{W}_\text{k} = \text{an independent variable or predictor variable}$$

Discriminant analysis uses ordinary least squares to estimate the values of the parameters 'a' and W_k that minimize the within Group sum of Squares. Discriminant analysis involves deriving linear combination of the independent variables that will discriminate between the prior defined groups in such a way that the misclassification error rates are minimized [5]. The function is given as

$$D(X) = b'\underline{X}$$

$$where\, b' = S^{-1}\left(\underline{\bar{X}}_1 - \underline{\bar{X}}_2\right)'$$

$$X = \left(\underline{\bar{X}}_1 - \underline{\bar{X}}_2\right)$$

Where, $D(X)$ *is a* $1 \times n$ vector of discriminant scores, b' *is a* $1 \times p$ vector of discriminant weights, and X *is a* $p \times n$ matrix containing the values of the n individuals on the p independent variables. S^{-1} the inverse of the pooled sample variance-covariance matrix of the independent variable. The Mahalanobis generalized distance D^2 Statistic is used to determine whether the between group differences in mean score profiles are statistically significant. Large value of D^2 would lead us to believe that the groups are sufficiently spread in terms of mean separation [5]. It is given as:

$$D^2 = \left(\underline{\bar{X}}_1 - \underline{\bar{X}}_2\right)' S^{-1} \left(\underline{\bar{X}}_1 - \underline{\bar{X}}_2\right)$$

The test can be constructed by for $\min g$:

$$Z = \frac{n_1 n_2 \left(n_1 + n_2 - p - 1\right) D^2}{n_1 + n_2 \left(n_1 + n_2 - 2\right) p}$$

3 DATA ANALYSIS AND RESULT

The analysis is done to compute the discriminant weights, to examine the associated significance and assumption tests based on linear combination of the predictor variables and also to classify each case into one of the two groups it closely resembles. The variables used in this analysis are Dependent variable: group 1 (Natural birth), group 2 (Caesarian section)and Independent variables: Mothers Height (X_1), Mothers Weight (X_2), Mothers Age (X_3), Baby's Weight (X_4) and Baby's Gender (X_5).

3.1 TEST FOR EQUALITY OF MEANS

Test of function(s)	Wilks' lamda	Sig
1	0.936	0.037

Table 1. Wilks' Lamda

$$Hypothesis: H_0: \underline{\mu}_1 = \underline{\mu}_2 \ vs \ H_1: Not \ H_0$$

$$Test \ Statistic: \lambda = \frac{|BSS|}{|BSS + WSS|}$$

$$\alpha = 0.05$$

$$Decision: since \ p-value(0.037) < 0.05, \ Reject \ H_0.$$

Under the hypothesis $H_0 : \underline{\mu}_1 = \underline{\mu}_2$ and common variance covariance matrix the test statistic F is distributed as a F-distribution with $n_1 + n_2 - p - 1df \qquad i.e.$

$$F: F_\alpha : (p, n_1 + n_2 - p - 1)$$

Using the above relation H_0 is rejected at the significant level α , if $F : F_\alpha : (p, n_1 + n_2 - p - 1)$

CLASSFICATION RULE

Assign an individual with realized score $\underline{X}$ on the p independent variables to $G_1 \ if \ D(X) \geq C$ otherwise,

To $G_2 \ if \ D(X) < C$ where

$C = (\underline{\bar{X}} - \underline{\bar{X}})' S^{-1} (\underline{\bar{X}} - \underline{\bar{X}})$ we assume that in each group observed scores on the p independent variables as multivariate normal with mean $\mu_i, i = 1,2$ and variance-covariance matrix S^{-1}, and if we can further assume that the prior probabilities of group membership and costs of misclassification of an individual that actually belongs to group 1(2) into 2(1) are equal.

Conclusion: The vector of the 2 group means are not the same.

3.2 TEST FOR EQUALITY OF GROUP MEANS INDIVIDUAL VARIABLE

$$Hypothesis: H_0: t_1 = t_2 \ vs \ H_1: Not \ H_0$$

$$Test \ Statistic: t = \frac{\bar{X}_1 - \bar{X}_2}{S_p \sqrt{\frac{1}{N_1} + \frac{1}{N_2}}}$$

$$Where \ S_p = \sqrt{\frac{N_1 S_1^2 + N_2 S_2^2}{N_1 + N_2 - 2}}$$

Variable	t	Df	Sig
Mothers Height	1.345	182	0.180
Mothers Weight	2.208	182	0.028
Mothers Age	2.053	182	0.042
Baby's Weight	0.077	182	0.940

Table 2. t-test for equality of mean

(i) Mothers Height

Decision: since p-value (0.180) > 0.05. Accept H_0.

Conclusion: There is no difference in the mean Mother's Height for group 1 and group 2.

(ii) Mothers Weight

Decision: since p-value (0.028) < 0.05. Reject H_0.

Conclusion: There is a difference in the mean Mother's Weight for group 1 and group 2.

(iii) Mothers Age

Decision: since p-value (0.042) < 0.05. Reject H_0.

Conclusion: There is a difference in the mean Mother's Ages for group 1 and group 2.

(iv) Baby's Weight

Decision: since p-value (0.942) > 0.05. Accept H_0.

Conclusion: There is a difference in the mean Baby's Ages for group 1 and group 2.

3.3 Canonical Discriminant Function Coefficient

Variable	Function
	1
Mothers Height	-0.074
Mothers Weight	0.047
Mothers Age	0.083
Baby's Weight	-0.199
Baby's Gender	-0.783
Constant	8.089

Table 3. Canonical Discriminant Function Coefficient

Hence,

$$8.089 - 0.074X_1 + 0.047X_2 + 0.083X_3 - 0.199X_4 - 0.783X_5$$

3.4 Classification of Results

		Predicted group Membership		Total
		Natural birth	Caesarian section	
Original Count	Natural birth	64	35	99
	Caesarian section	30	55	85
%	Natural birth	64.6	35.4	100
	Caesarian section	35.3	64.7	100

Table 4. Classification of Results

From the table 64.7% of the group cases were correctly classified while 35.3% were wrongly classified.

4 Conclusion

In general, we were able to construct a Discriminant score: for detecting the variables (Mothers Height, Mothers Weight, Mothers Age, Baby's Weight and Baby's Gender) that allow the researcher to discriminate between Natural birth and

Caesarian section and we have shown that the group differs with regards to the mean of variables, and the variables to predict group membership.

REFERENCES

[1] UNICEF, (2008): Progress for Children Report.
[2] Rencher, A.C. (2002), Method of Multivariate Analysis, Second Edition, John Wiley and Sons Inc.
[3] Ganesalingam, S. (1989), "Classification and Mixture Approaches to Clustering via Maximum Likelihood". Applied Statistics, 38, No. 3, Page 445-466.
[4] Balogun, O.S., Oyejola, B.A. and Akingbade, T.J. (2012), Use of Discriminant analysis in Classification of Drug Peddlers and Non-Drug Peddlers in Kwara State. International Journal of Engineering Research and Application, Vol. 2, Issue 5, Page 936-938.
[5] Dillion, W.R. and Goldstein, S. (1984), "Multivariate Analysis Method and Applications. John Wiley Sons, Inc. Canda.