International Journal of Advances in Engineering and Management (IJAEM)
Volume 7, Issue 02 Feb. 2025, pp: 442-456 www.ijaem.net ISSN: 2395-5252

A New and Efficient Approach for Estimating Probabilities of Misclassification and Discrimination

Nwankwo-obi, J.A and Okoli , C.N
Department of Statistics, Chukwuemeka Odumegwu Ojukwu University, Anambra State, Nigeria.

Date of Submission: 05-02-2025 Date of Acceptance: 15-02-2025

ABSTRACT
This study proposed a new and efficient approach for estimating probabilities of misclassification and discrimination.The study explores the estimation of apparent and optimum probabilities of misclassification for three populations using real-life anthropometric data. It also compares the effectiveness of Linear Discriminant Analysis (LDA) and Quadratic Discriminant Analysis (QDA) in classifying data from three distinct populations, assessing their accuracy and reliability. Real life anthropometric datasets were utilized, and both stratified random and simple random sampling techniques were employed, comprising three populations of school students with four variables.A computer programming language codes were written via R-Studio package to solve the numerical problems posed in the study. The model misclassified around 4.86% of observations for Linear Discriminant Analysis (LDA) and approximately 3.71% for Quadratic Discriminant Analysis (QDA). QDA showed higher accuracy (98.76% vs. 98.38%) and lower error rate (1.24% vs. 1.62%) compared to LDA. Additionally, QDA demonstrated excellent discriminatory power with a perfect AUC-ROC score.The study shown that QDA outperformed LDA in terms of accuracy and error rates, demonstrating superior discriminatory power. This study provided valuable insights for those working with datasets involving multiple populations and variables with potential applications in various fields such as multivariate methods, data science, machine learning, business, healthcare and finance. Furthermore, the study offers a practical approach to classifying observations into distinct populations using LDA and QDA, achieving high accuracy rates for real-life data scenarios. It establishes a foundation for future research endeavours andpresents a comprehensive framework for comparing LDA and QDA performance in ESD data, highlighting the effectiveness of QDA in handling skewed data for multiple populations .The research recommended further exploration into developing a generalized model for estimating probabilities of misclassification via ESD with flexible distribution assumptions and robust estimation methods.

Keywords: Edgeworth Series Distribution, Optimal probability, Quadratic discriminant analysis, Linear discriminant analysis, AUC-ROC

I. INTRODUCTION

Error can be defined as an act or condition of ignorant or imprudent deviation from a code of behavior or an act involving an unintentional deviation from truth or accuracy (Venkatesan, 2014). An error is an action which is inaccurate or incorrect. In some usages, an error is synonymous with a mistake (Bruno et al., 2015). The etymology derives from the Latin term 'errare', meaning 'to stray'. In statistics, 'error' refers to the difference between the value which has been computed and the correct value (Metsämuuronen, 2022).

A classification problem occurs when one makes a number of measurements on objects (observations) and wishes to classify the observations into one of several groups on the basis of the measurements. The objects (observations) cannot be identified with a group directly without recourse to the measurements (Awogbemi and Onyeagu, 2019).

Fisher (1936) illustrated this classification issue by classifying iris flower from unknown group (specie) to any of the three known species (Iris setosa red, iris versicolour green, and iris virginica black) with regards to their attribute (Septal length in cm, septal width in cm, petal length in cm and petal width in cm) as recorded by (Awogbemi and Onyeagu, 2019). The general procedure for classifying an observation, x with p observed characters ($x_1,\dots x_p$) consists of

International Journal of Advances in Engineering and Management (IJAEM)
Volume 7, Issue 02 Feb. 2025, pp: 442-456 www.ijaem.net ISSN: 2395-5252

determining a function of $(x_1,\ldots x_p)$ approximately, and assigning x to one of two populations depending on the value of the discriminant function (Ruiz, 2019). Since the observation vector is random and the parameters for determining this function are often unknown, the procedure could result into two types of errors defined by errors of misclassification. Errors of misclassification occur when there is selection of criteria that is not suitable for classification (John, 2010).

When constructing a classification procedure, it is important to minimize on the average, the bad effects of misclassification since a good classification procedure results to few misclassifications (Hand, 2012).

When an experimenter fails to recognize an observation to be non-normal, and proceeds to use the normal regions for classification the question that emanates is "how does this failure to transform to normality, prior to classification affect the probability of misclassification"? This problem was investigated by comparing the errors of misclassification associated with Johnson system distributions in the appropriate transformable non-normal case with that of normal distribution (Awogbemi & Onyeagu, 2019). Errors of misclassification associated with Gamma distribution were also examined by Mahmoud and Mustafa (1995). A lot of work has been done by researchers in connection with errors of misclassification when the underlying distribution is transformable non-normal distribution, but the errors of misclassification associated with persistent non-normal distribution remain unresolved (Morgan et al., 2016).

Awogbemi and Onyeagu (2019) studied on errors of misclassification associated with Edgeworth series distribution survey on two populations using small sample sizes. However, this work majors on large sample sizes from three populations which none of the researchers sighted had written on. This justified the need for this work

II. REVIEW OF RELATED LITERATURE

Gasana et al. (2024) conducted a study on the moments of the likelihood-based discriminant function, which led to quadratic discriminant functions. They separately considered classification into one of two known multivariate normal populations with: known covariance matrix; unknown covariance matrix. The two cases depended on the sample size and an unknown squared Mahalanobis distance. Since the exact distributions were complicated to obtain, the researchers established moments for the likelihood-based discriminant functions to express the basic characteristics of the respective distributions. The study's results could be utilized in various applications, such as: Edgeworth expansion, which provided alternative approximations of the distribution of misclassification errors. By examining the moments of the likelihood-based discriminant function, they contributed to a deeper understanding of the underlying distributions and paved the way for further research in discriminant analysis.

Olusola and Onyeagu (2020) conducted a research study on binary classification problems in discriminant analysis using linear programming methods. The study focused on assigning a new object with multivariate features to one of two distinct populations based on historical sample sets from both populations. The researchers proposed a linear discriminant analysis framework called Minimised Sum of Deviations by Proportion (MSDP) to model the binary classification problem. In the MSDP formulation, they minimised the sum of proportion of exterior deviations subject to: group separation constraints; normalisation constraint; upper bound constraints on proportions of exterior deviations, sign unrestriction and non-negativity constraints. They adopted the two-phase method in linear programming to generate the discriminant function and constructed the decision rule for group-membership prediction using the apparent error rate. The performance of MSDP was compared to existing linear discriminant models using a previously published dataset on road casualties. The results showed that MSDP was more promising and well-suited for the imbalanced dataset on road casualties.

Kanuti and Ngaruye (2024) conducted a research on asymptotic results for expected probability of misclassifications in linear discriminant analysis with repeated measurements. They proposed approximations for the misclassification probabilities in linear discriminant analysis when the group means had a bilinear regression structure. They checked the accuracies of the proposed approximations numerically by conducting a Monte Carlo simulation. The key contributions were: they gave a unified location and scale mixture expression of the standard normal distribution for the linear discriminant function; they obtained estimated approximations of misclassification for the three cases: unweighted case, weighted known covariance matrix, and weighted unknown

International Journal of Advances in Engineering and Management (IJAEM)
Volume 7, Issue 02 Feb. 2025, pp: 442-456 www.ijaem.net ISSN: 2395-5252

covariance matrix. The findings were: they found that larger p (number of repeated measurements) was better for classification when the covariance matrix was known, also in the unweighted case; they discovered that in the case where the covariance matrix was unknown, they gained more information if fewer repeated measurements were used compared to when many repeated measurements closer to the number of included sample size were used. The research provided valuable insights into the behavior of LDA with repeated measurements and offered practical guidelines for improving classification accuracy.

III. RESEARCH METHODOLOGY

3.1 Data Collection

A cross- sectional study was conducted on the Anthropometric status of school learners in selected schools in Orumba North Local Government Area of Anambra State.This study used stratified random and simple random sampling techniques respectively, designed for school learners. Firstly, the schools were selected randomly by stratified sampling method according to socio-economic levels from among schools in Orumba North Local Government Area which represents one of the largest LGA in Anambra State. Secondly, simple random sampling was conducted in each strata of high socio-economic of interest. A total of 350 school learners were examined and equal allocation was maintained according to gender that is 175 males and 175 females. With a rich cultural heritage, the area is predominantly inhabited by the Igbo people, with a population of approximately 170,000 according to 2006 census. The heights of the learners were measured with the help of calibrated meter rule to the nearest 0.1cm. The learners were positioned with their feet closed together and stand uprightly, barefooted against a vertical measuring meter rule. Once the correct position was achieved the interviewer lowered the head plate until it just touched the top of the learners head and while maintaining this position, he/she were asked to stand upright without lifting the heels. Other variables (head circumference, shoulder width, elbow height) were also measured, and recorded in the nearest 0.1cm. The variables considered are defined for the three populations (Nursery, Primary, and Secondary) as follows:

Z_1: Height
Z_2: Head Circumference
Z_3: Shoulder Width
Z_4: Elbow Height

The data obtained from the schools are presented in Table 3.1 (See Appendix A).

3.2 Proposed Method of Estimating Probabilities of Misclassification Via LDA

Let x_{ijk} i= 1, 2, 3; j = 1,2,3; k = 1,2,3 are to be independent samples of sizes n_1, n_2 and n_3 from population π_1, π_2 and π_3. To estimate the apparent probabilities of misclassification, we define (1)

$$E_{12E} = \sum_{j=1}^{n_1} \frac{Y_j}{n_1}$$

Where $Y_j = 1$ if x_{ij} is classified as belonging to π_2 and $Y_j = 0$, if x_{ij} is classified as belonging to π_1, j = 1, 2, 3 ….. n_1

Similarly,

$$E_{21E} = \sum_{j=1}^{n_2} \frac{\delta_j}{n_2} \quad (2)$$

Where E_{21E} is the apparent probability of misclassification when an observation from population π_2 is misclassified by ESD (Awogbemi and Onyeagu, 2019).

Where $\delta_j = 1$ if x_{2j} is classified as belonging to π_1 and $\delta_j = 0$, if x_{2j} is classified as belonging to π_2, j = 1, 2, 3 ….. n_2 (3)

$$E_{13E} = \sum_{j=1}^{n_1} \frac{A_j}{n_1}$$

Where $A_j = 1$ if x_{1j} is classified as belonging to π_3 and $A_j = 0$, if A_{Ij} is classified as belonging to π_1, j = 1, 2, 3 ….. n_1

$$E_{23E} = \sum_{j=1}^{n_2} \frac{B_j}{n_2} \quad (4)$$

International Journal of Advances in Engineering and Management (IJAEM)
Volume 7, Issue 02 Feb. 2025, pp: 442-456 www.ijaem.net ISSN: 2395-5252

Where $B_j = 1$ if x_{2j} is classified as belonging to π_3 and $B_j = 0$, if x_{2j} is classified as belonging to π_2, j = 1, 2, 3 ….. n_2

$$E_{31E} = \sum_{j=1}^{n_3} \frac{C_j}{n_3}$$

Where $C_j = 1$ if x_{3j} is classified as belonging to π_1 and $C_j = 0$, if x_{3j} is classified as belonging to π_2, j = 1, 2, 3 ….. n_3 (5)

$$E_{32E} = \sum_{j=1}^{n_3} \frac{D_j}{n_3}$$

Where $D_j = 1$ if x_{3j} is classified as belonging to π_2 and $D_j = 0$, if x_{3j} is classified as belonging to π_3, j = 1, 2, 3 ….. n_3 (6)

Following the same procedure, for normal distribution classification rule for the purpose of comparison thus:

$$E_{12N} = \sum_{j=1}^{n_1} \frac{Y_j}{n_1} \quad (7)$$

$$E_{21N} = \sum_{j=1}^{n_2} \frac{\sigma_j}{n_2} \quad (8)$$

$$E_{23N} = \sum_{j=1}^{n_2} \frac{B_j}{n_2} \quad (9)$$

$$E_{13N} = \sum_{j=1}^{n_1} \frac{A_j}{n_1} \quad (10)$$

$$E_{31N} = \sum_{j=1}^{n_3} \frac{C_j}{n_3} \quad (11)$$

$$E_{32N} = \sum_{j=1}^{n_3} \frac{D_j}{n_3} \quad (12)$$

Where: E_{12N} is the apparent probability of misclassification when observation from population π_1 is misclassified by normal distribution (ND) classificatory rule.

E_{21N} is the apparent probability of misclassification when observation from population π_2 is misclassified by ND classificatory rule (Awogbemi and Onyeagu, 2019).

E_{23N} is the apparent probability of misclassification when observation from population π_2 is misclassified by ND classificatory rule.

Also E_{31N} is the apparent probability of misclassification when observation from population π_3 is misclassified by ND classificatory rule.

E_{32N} is the apparent probability of misclassification when observation from population π_3 is misclassified by ND classificatory rule.

3.2.1MODIFIED CLASSIFICATION RULES FOR NORMAL DISTRIBUTION (UNIVARIATE)VIA LDA

Let the probability density function of x in π_i (i = 1, 2, 3) be

$$f_i(x) = \frac{1}{\sigma\sqrt{3\pi}} = exp\left[-\frac{1}{3}\left(\frac{x-\mu_i}{\sigma}\right)^2\right], -\infty < x < \infty, i = 1, 2, 3 \quad (3.2)$$

If θ is the mean of the observation x and H_o: $\theta = \mu_1$ vs H_a: $\theta = \mu_2 = \mu_3$, then the likelihood when $\mu_1 < \mu_2 < \mu_3$.

$$L = \frac{f_1(x)}{f_2(x).f_3(x).} = exp[-\frac{1}{3}(\frac{x-\mu_1}{\sigma})^2 + \frac{1}{3}(\frac{x-\mu_2}{\sigma})^2] + \frac{1}{3}(\frac{x-\mu_3}{\sigma})^2] \quad (3.2.0a)$$

$$L^1 = \frac{1}{3}(\frac{x-\mu_1}{\sigma})^2 + \frac{1}{3}(\frac{x-\mu_2}{\sigma})^2 + \frac{1}{3}(\frac{x-\mu_3}{\sigma})^2 =$$

$$\frac{-1}{3\sigma^2}[3x - (\mu_1 + \mu_2 + \mu_3)] \ (\mu_3 - \mu_2 - \mu_1)$$

$$= \left[x - \frac{1}{3}(\mu_1 + \mu_2 + \mu_3)\right)(\frac{\mu_3 - \mu_2 - \mu_1}{\sigma}) \quad (3.2.0b)$$

The result in equation (3.20a) is the discriminant function from adjusted Anderson's classification statistic (w) when the distributions in the three populations are univariate normal with equal variance but different means (Sedransk, and Okamato, 1971).

According to the Neyman and Pearson lemma cited by Rao, (1965), we reject H_o if L<k where K is a constant.

Following equation (3.20) and the decision rule made, we specify the classification rule as follows:

Classify x as members of $\pi_1$if $w > 0$ and or classify x as member of π_2 if w ≤ 0 and or classify x E π_2 if w ≤ 0
Classify x as members of π_3 if $w < 0$ and or classify $XE\pi_3$ if $w \leq 0$ (3.2.1)

The rule stated in equation (3.21) reduces to;
Classify x E π_1 if $x < \frac{1}{3}(\mu_1 + \mu_2 + \mu_3)$
Classify x E π_2 if $x \geq \frac{1}{3}(\mu_1 + \mu_2 + \mu_3)$
Classify x E π_3 if $x > \frac{1}{3}(\mu_1 + \mu_2 + \mu_3)$
(3.2.2)
Similarly, when $\mu_1 > \mu_2 > \mu_3$ the classification rule becomes
Classify x E π_3 if $x \leq \frac{1}{3}(\mu_1 + \mu_2 + \mu_3)$
Classify x E π_2 if $x < \frac{1}{3}(\mu_1 + \mu_2 + \mu_3)$
Classify x E π_1 if $x \geq \frac{1}{3}(\mu_1 + \mu_2 + \mu_3)$
(3.2.3)

The rules in equations (3.2) is made when $\mu_1 > \mu_2 > \mu_3$ and are known. But when the parameters μ_1, μ_2, μ_3 are unknown, they are to be estimated from the sample sizes of n_1 for π_1, n_2 for π_2 and n_3 for π_3 by $\bar{x}_1, \bar{x}_2$and $\bar{x}_3$. The classification rule becomes:
Classify $x \in \pi_1$ if $x < \frac{\bar{x}_1 + \bar{x}_2 + \bar{x}_3}{3}$, $\bar{x}_1 < \bar{x}_2 < \bar{x}_3$
(3.2.4)
Similarly;
Classify $x \in \pi_2$ if $x \geq \frac{\bar{x}_1 + \bar{x}_2 + \bar{x}_3}{3}$, $\bar{x}_1 < \bar{x}_2 < \bar{x}_3$
(3.2.5)
In the similar way;
Classify $x \in \pi_3$ if $x > \frac{\bar{x}_1 + \bar{x}_2 + \bar{x}_3}{3}$, $\bar{x}_1 < \bar{x}_2 < \bar{x}_3$
(3.2.6)

3.2.2 MODIFIED CLASSIFICATION RULE FOR EDGEWORTH SERIES DISTRIBUTION (MESD) (UNIVARIATE) VIA LDA

Let the probability density function of population $\pi_i$be;

$$f_i(x) = \left[1 - \frac{\kappa_3}{6} D^3\right] \varphi\left(\frac{x-\mu_i}{\sigma}\right), -\infty < x < \infty, i = 1,2,3 \quad (3.2.7)$$

When $\mu_1 < \mu_2 < \mu_3$, the likelihood Ratio (LR) is now

$$L = \frac{f_1(x)}{f_2(x) f_3(x)}$$

$$L = \frac{=\{1 - [\frac{\kappa_3}{2\sigma^3}](\frac{x-\mu 1}{\sigma}) + [\frac{\kappa_3}{6\delta^3}](\frac{x-\mu 1}{\sigma})^3\}\varphi(\frac{x-\mu 1}{\sigma})}{\{1 - [\frac{\kappa_3}{2\sigma^3}](\frac{x-\mu 2}{\sigma}) + [\frac{\kappa_3}{6\sigma^3}](\frac{x-\mu 2}{\sigma})^3\}\varphi(\frac{x-\mu 2}{\sigma}).\ [1 - [\frac{\kappa_3}{2\sigma^3}](\frac{x-\mu 3}{\sigma}) + [\frac{\kappa_3}{6\sigma^3}](\frac{x-\mu 3}{\sigma})^3\}\varphi(\frac{x-\mu 3}{\sigma})} \quad (3.2.8)$$

By implication, equation (3.28) becomes;

$$L = \frac{A\ exp\ [-\frac{1}{2}(\frac{x-\mu 1}{\sigma})^2]}{B\ exp\ [-\frac{1}{2}(\frac{x-\mu 2}{\sigma})^2]\ R\ exp\ [-\frac{1}{2}(\frac{x-\mu 3}{\sigma})^2]} \quad (3.2.9)$$

$$where\ A = [1 - \left(\frac{\kappa_3}{2\sigma^3}\right)\left(\frac{x-\mu_1}{\sigma}\right) + \left(\frac{\kappa_3}{6\sigma^3}\right)\left(\frac{x-\mu_1}{\sigma}\right)^3] \quad (3.3.0)$$

$$B = [1 - \left(\frac{\kappa_3}{2\sigma^3}\right)\left(\frac{x-\mu_2}{\sigma}\right) + \left(\frac{\kappa_3}{6\sigma^3}\right)\left(\frac{x-\mu_2}{\sigma}\right)^3] \quad (3.3.1)$$

$$R = [1 - \left(\frac{\kappa_3}{2\sigma^3}\right)\left(\frac{x-\mu_3}{\sigma}\right) + \left(\frac{\kappa_3}{6\sigma^3}\right)\left(\frac{x-\mu_3}{\sigma}\right)^3] \quad (3.3.2)$$

According to Neyman and Pearson Lemma,
We reject H_0 if

$$L < K \equiv lnL < k \quad (3.3.3)$$

Taking $k = 1 \implies InL < 0$
Then we reject x ∈ π_1 if

$$\text{lnA} - \frac{1}{2}\{\frac{x-\mu_1}{\sigma}\}^2 - InB + \frac{1}{2}\{\frac{x-\mu_2}{\sigma}\}^2 . InR + \frac{1}{2}\{\frac{x-\mu_3}{\sigma}\}^2 < o \quad (3.3.4)$$

Equation (3.34) reduces to

$$\text{In}\left(\frac{A}{BR}\right) + \left[x - \left(\frac{\mu_1 + \mu_2 + \mu_3}{3}\right)\right]\left(\frac{\mu_1 - \mu_2 - \mu_3}{\sigma}\right) < o \quad (3.3.5)$$

From equation (3.35) the classification rule will now be: when $\mu_1 < \mu_2 < \mu_3$
Classify x as element of $\pi_1$if

$$Ln\left(\frac{A}{BR}\right) + W > 0$$

or classify $x \in \pi_1$ if $ln\left(\frac{A}{BR}\right) + W > 0$ (3.3.6)
classify x as element $\pi_2$if

International Journal of Advances in Engineering and Management (IJAEM)
Volume 7, Issue 02 Feb. 2025, pp: 442-456 www.ijaem.net ISSN: 2395-5252

$Ln\left(\frac{A}{BR}\right) + W \leq 0$ or classify $x \in \pi_2$ $if\ ln\left(\frac{A}{BR}\right) + W \leq 0$ (3.3.7)

and

classify x ∈ $\pi_3$if

$Ln\left(\frac{A}{BR}\right) + W < 0$ (3.3.8)

When$\mu_1 > \mu_2 > \mu_3$

Classify $x \in \pi_1\ if\ ln\left(\frac{A}{BR}\right) - W > 0$ (3.3.9)

Classify $x \in \pi_2\ if\ ln\left(\frac{A}{BR}\right) - W \leq 0$ (3.4.0)

and

Classify $x \in \pi_3\ if\ ln\left(\frac{A}{BR}\right) - W < 0$ (3.4.1)

When the parameters μ_1, μ_2, μ_3 are unknown, they are to be estimated by $\bar{x}_1, \bar{x}_2, \bar{x}_3$ respectively and substituted in equation (3.41) before classification starts.

In the method of the comparisons of errors of misclassification using MESD and MND classification rules, and data generated from the MESD, we would investigate by empirical method, the effect of applying normal classification rule (likelihood ratio) when the distribution is MESD. Thus the classification rule for MESD is left in the form; when $\mu_1 < \mu_2, < \mu_3$ (Chun'g'anda 1976)\

Classify $x \in \pi_1\ if\ \frac{A\ exp\left[-\frac{1}{2}\left(\frac{x-\mu_1}{\sigma}\right)^2\right]}{B\ exp\left[-\frac{1}{2}\left(\frac{x-\mu_2}{\sigma}\right)^2\right].\ R\ exp\left[-\frac{1}{2}\left(\frac{x-\mu_3}{\sigma}\right)^2\right]} < 1$ (3.4.2)

also

classify $x \in \pi_2\ if\ \frac{A\ exp\left[-\frac{1}{2}\left(\frac{x-\mu_1}{\sigma}\right)^2\right]}{B\ exp\left[-\frac{1}{2}\left(\frac{x-\mu_2}{\sigma}\right)^2\right].\ R\ exp\left[-\frac{1}{2}\left(\frac{x-\mu_3}{\sigma}\right)^2\right]} \geq 1$ (3.4.3)

and

classify $x \in \pi_3\ if\ \frac{A\ exp\left[-\frac{1}{2}\left(\frac{x-\mu_1}{\sigma}\right)^2\right]}{B\ exp\left[-\frac{1}{2}\left(\frac{x-\mu_2}{\sigma}\right)^2\right].\ R\ exp\left[-\frac{1}{2}\left(\frac{x-\mu_3}{\sigma}\right)^2\right]} > 1$ (3.4.4)

Where A and B remain as defined earlier in equation (3.30), (3.31) and (3.32)

The normal classificatory rule when $\mu_1 < \mu_2 < \mu_3$, is

classify $x \in \pi_1\ if\ x < \left(\frac{\mu_1+\mu_2+\mu_3}{3}\right)$

also classify $x \in \pi_2\ if\ x > \left(\frac{\mu_1+\mu_2+\mu_3}{3}\right)$

and

Classify $x \in \pi_3\ if\ x \geq \left(\frac{\mu_1+\mu_2+\mu_3}{3}\right)$ (3.4.5)

3.3 PROPOSED METHOD OF ESTIMATING PROBABILITIES OF MISCLASSIFICATION VIA QDA

Let x_{ijk} i= 1, 2, 3; $j = 1, 2, 3$; $k = 1, 2, 3$be independent samples of sizes n_1, n_2 *and* n_3 from population π_1, π_2 *and* π_3. To estimate the apparent probabilities of misclassification, we define.

QDA Classification Rule

$$E_{12Q} = \sum_{j=1}^{n_1} \frac{Z_j}{n_1}$$

Where $Z_j = 1$ if x_{ij} is classified as belonging to π_2 and $Z_j = 0$, if x_{ij} is classified as belonging to π_1, j = 1, 2, 3 ….. n_1

Similarly,

$$E_{21Q} = \sum_{j=1}^{n_2} \frac{\theta_j}{n_2}$$

where $E_{21Q,}$ is the apparent probability of misclassification when an observation from population π_2 is misclassified by QDA. (3.4.6)

Where $\theta_j = 1$ if x_{2j} is classified as belonging to π_1 and $\theta_j = 0$, if x_{2j} is classified as belonging to π_2, j = 1, 2, 3 ….. n_2

$$E_{13Q} = \sum_{j=1}^{n_1} \frac{\alpha_j}{n_1}$$

Where $\alpha_j = 1$ if x_{1j} is classified as belonging to π_3 and $\alpha_j = 0$, if α_j is classified as belonging to π_1, j = 1, 2, 3 ….. n_1

$$E_{23Q} = \sum_{j=1}^{n_2} \frac{\emptyset_j}{n_2}$$

Where $\emptyset_j = 1$ if x_{2j} is classified as belonging to π_3 and $\emptyset_j = 0$, if x_{2j} is classified as belonging to π_2, j = 1, 2, 3 ….. n_2

$$E_{31Q} = \sum_{j=1}^{n_3} \frac{\delta_j}{n_3}$$

Where $\delta_j = 1$ if x_{3j} is classified as belonging to π_1 and $\delta_j = 0$, if x_{3j} is classified as belonging to π_3, j = 1, 2, 3 ….. n_3

$$E_{32Q} = \sum_{j=1}^{n_3} \frac{\omega_j}{n_3}$$

Where $\omega_j = 1$ if x_{3j} is classified as belonging to π_2 and $\omega_j = 0$, if x_{3j} is classified as belonging to π_3, j = 1, 2, 3 ….. n_3

Following the same procedure, for normal distribution classification rule for the purpose of comparison thus: (3.4.8)

$$E_{12N} = \sum_{j=1}^{n_1} \frac{Y_j}{n_1} \quad (3.5.2)$$

$$E_{21N} = \sum_{j=1}^{n_2} \frac{\sigma_j}{n_2} \quad (3.53)$$

$$E_{23N} = \sum_{j=1}^{n_2} \frac{B_j}{n_2} \quad (3.4.9) \quad (3.5.4)$$

$$E_{13N} = \sum_{j=1}^{n_1} \frac{A_j}{n_1} \quad (3.5.5)$$

$$E_{31N} = \sum_{j=1}^{n_3} \frac{C_j}{n_3} \quad (3.5.0) \quad (3.5.6)$$

$$E_{32N} = \sum_{j=1}^{n_3} \frac{D_j}{n_3} \quad (3.5.7)$$

Where:

$E_{12N}, E_{21N}, E_{23N}, E_{13N}, E_{31N}$ and E_{32N} are defined similarly to the QDA classification rule. (3.5.1)

It should be noted that QDA classification rule uses a quadratic discriminant function, which takes into account the covariance matrix of each population, whereas the LDA classification rule uses a linear discriminant function, which assumes equal covariance matrices across populations.

3.3.1 MODIFIED CLASSIFICATION RULES FOR NORMAL DISTRIBUTION (MULTIVARIATE) USING QDA

Let the probability density function of X in π_i (i = 1, 2, 3) be:

$$f_i(x) = \frac{1}{\sqrt{\left((2\pi)^p |\Sigma_i|\right)}} \exp\left[-\frac{1}{2}(x-\mu_i)^T \Sigma_i^{-1}(x-\mu_i)\right] \quad (3.5.8)$$

where X is a p-dimensional vector, μ_i is the mean vector, Σ_i is the covariance matrix, and $|\Sigma_i|$ is the determinant of Σ_i.

If θ is the mean of the observation X and H_0: $\theta = \mu_1$ vs. H_1: $\theta = \mu_2 = \mu_3$, then the likelihood ratio when $\mu_1 < \mu_2 < \mu_3$:

$$L = \frac{f_1(x)}{f_2(x) \times f_3(x)}$$

$$L = \exp\left[-\frac{1}{2}(x-\mu_1)^T \Sigma_1^{-1}(x-\mu_1) + \frac{1}{2}(x-\mu_2)^T \Sigma_2^{-1}(x-\mu_2) + \frac{1}{2}(x-\mu_3)^T \Sigma_3^{-1}(x-\mu_3)\right] \quad (3.68)$$

Taking the logarithm and simplifying:

International Journal of Advances in Engineering and Management (IJAEM)
Volume 7, Issue 02 Feb. 2025, pp: 442-456 www.ijaem.net ISSN: 2395-5252

$$\log(L) = -\frac{1}{2}\left[(x-\mu_1)^T \Sigma_1^{-1}(x-\mu_1) - (x-\mu_2)^T \Sigma_2^{-1}(x-\mu_2) - (x-\mu_3)^T \Sigma_3^{-1}(x-\mu_3)\right]$$

The result in equation (3.27) is the discriminant function from QDA.

Classification Rule:

Classify x as members of π_1 if w > 0 (3.5.9)

Classify x as members of π_2 if $w \leq 0$ and $\Delta_1 > \Delta_2$ (3.6.0)

Classify x as members of π_3 if $w < 0$ and $\Delta_2 > \Delta_3$ (3.6.1)

where:

$$w = (x-\mu_1)^T \Sigma_1^{-1}(x-\mu_1) - (x-\mu_2)^T \Sigma_2^{-1}(x-\mu_2) \quad (3.6.2)$$

$$\Delta_1 = (x-\mu_1)^T \Sigma_1^{-1}(x-\mu_1) - (x-\mu_3)^T \Sigma_3^{-1}(x-\mu_3) \quad (3.6.3)$$

$$\Delta_2 = (x-\mu_2)^T \Sigma_2^{-1}(x-\mu_2) - (x-\mu_3)^T \Sigma_3^{-1}(x-\mu_3) \quad (3.6.4)$$

When the parameters: μ_1, μ_2, μ_3, Σ_1, Σ_2, and Σ_3 are unknown, they are estimated from the sample sizes of n_1 for π_1, n_2 for π_2, and n_3 for π_3.

Classification Rule becomes:

Classify $x \in \pi_1$ if

$$x^T \Sigma_1^{-1} x - 2\mu_1^T \Sigma_1^{-1} x + \mu_1^T \Sigma_1^{-1} \mu_1 > x^T \Sigma_2^{-1} x - 2\mu_2^T \Sigma_2^{-1} x + \mu_2^T \Sigma_2^{-1} \mu_2 \quad (3.6.5)$$

Classify $x \in \pi_2$ if

$$x^T \Sigma_2^{-1} x - 2\mu_2^T \Sigma_2^{-1} x + \mu_2^T \Sigma_2^{-1} \mu_2 \geq x^T \Sigma_1^{-1} x - 2\mu_1^T \Sigma_1^{-1} x + \mu_1^T \Sigma_1^{-1} \mu_1 \quad (3.6.6)$$

Classify $x \in \pi_3$ if

$$x^T \Sigma_3^{-1} x - 2\mu_3^T \Sigma_3^{-1} x + \mu_3^T \Sigma_3^{-1} \mu_3 > x^T \Sigma_2^{-1} x - 2\mu_2^T \Sigma_2^{-1} x + \mu_2^T \Sigma_2^{-1} \mu_2 \quad (3.6.7)$$

Replace μ_i with x_i and Σ_i with S_i (sample covariance matrix) for estimated parameters.

Note: QDA assumes different covariance matrices for each population, whereas LDA assumes equal covariance matrices.

3.3.2 MODIFIED CLASSIFICATION RULE FOR EDGEWORTH SERIES DISTRIBUTION (MESD) USING QDA

Let the probability density function of population π_i be:

$$f_i(x) = \left[1 - \frac{\lambda_3}{6} D^3\right] \varphi\left(\frac{x-\mu_i}{\sigma}\right), -\infty < x < \infty, i = 1,2,3 \quad (3.6.8)$$

$$f_i(x) = \left[1 - \frac{\lambda_3}{6} D^3\right] \varphi\left(\frac{x-\mu_i}{\sigma}\right), -\infty < x < \infty, i = 1,2,3 \quad (3.6.9)$$

where φ is the standard normal density function.

When $\mu_1 < \mu_2 < \mu_3$, the likelihood ratio (LR) is:

International Journal of Advances in Engineering and Management (IJAEM)
Volume 7, Issue 02 Feb. 2025, pp: 442-456 www.ijaem.net ISSN: 2395-5252

$$L = \frac{f_1(x)}{f_2(x) \times f_3(x)}$$

$$L = \frac{\left[\left\{1 - \frac{(x-\mu_1)}{\sigma} + \frac{(x-\mu_1)}{\sigma^3}\right\}\varphi\left(\frac{(x-\mu_1)}{\sigma}\right)\right]}{\left[\left\{1 - \frac{(x-\mu_2)}{\sigma} + \frac{(x-\mu_2)}{\sigma^3}\right\}\varphi\left(\frac{(x-\mu_2)}{\sigma}\right)\right] \cdot \left[\left\{1 - \frac{(x-\mu_3)}{\sigma} + \frac{(x-\mu_3)}{\sigma^3}\right\}\varphi\left(\frac{(x-\mu_3)}{\sigma}\right)\right]}$$

(3.7.0)

By implication, equation (3.80) becomes:

$$L = \frac{\left(A\exp\left[-\frac{1}{2}\left(\frac{(x-\mu_1)}{\sigma}\right)^2\right]\right)}{\left(B\exp\left[-\frac{1}{2}\left(\frac{(x-\mu_2)}{\sigma}\right)^2\right]\right)\left(\mathrm{Re}\,xp\left[-\frac{1}{2}\left(\frac{(x-\mu_3)}{\sigma}\right)^2\right]\right)} \quad (3.7.1)$$

where A, B, and R are defined in equations (3.82), (3.83), and (3.84).

$$A = \left[1 - \left(\frac{\lambda_3}{2\sigma^3}\right)\left(\frac{x-\mu_1}{\sigma}\right) + \left(\frac{\lambda_3}{6\sigma^3}\right)\left(\frac{x-\mu_1}{\sigma}\right)^3\right] \quad (3.7.2)$$

$$B = \left[1 - \left(\frac{\lambda_3}{2\sigma^3}\right)\left(\frac{x-\mu_2}{\sigma}\right) + \left(\frac{\lambda_3}{6\sigma^3}\right)\left(\frac{x-\mu_2}{\sigma}\right)^3\right] \quad (3.7.3)$$

$$R = \left[1 - \left(\frac{\lambda_3}{2\sigma^3}\right)\left(\frac{x-\mu_3}{\sigma}\right) + \left(\frac{\lambda_3}{6\sigma^3}\right)\left(\frac{x-\mu_3}{\sigma}\right)^3\right] \quad (3.7.4)$$

These expressions represent the coefficients of the Edgeworth Series Distribution (ESD) for each population π_i, where:

- λ_3 is the skewness parameter
- σ is the standard deviation
- μ_i is the mean of population π_i
- x is the observation

These coefficients are used in the likelihood ratio and classification rules for ESD.

Taking the logarithm and simplifying:

$$\ln(L) = \ln\left(\frac{A}{BR}\right) - \frac{1}{2}\left(\frac{(x-\mu_1)}{\sigma}\right)^2 + \frac{1}{2}\left(\frac{(x-\mu_2)}{\sigma}\right)^2 + \frac{1}{2}\left(\frac{(x-\mu_3)}{\sigma}\right)^2 - \left[\frac{\lambda_3}{6\sigma^3}\right]\left[\frac{(x-\mu_1)}{\sigma}\right]^3 - \left[\frac{\lambda_3}{6\sigma^3}\right]\left[\frac{(x-\mu_2)}{\sigma}\right]^3 - \left[\frac{\lambda_3}{6\sigma^3}\right]\left[\frac{(x-\mu_3)}{\sigma}\right]^3 \quad (3.7.5)$$

Classification Rule:

Classify x as an element of π_1 if $\ln(L) > 0$ or

International Journal of Advances in Engineering and Management (IJAEM)
Volume 7, Issue 02 Feb. 2025, pp: 442-456 www.ijaem.net ISSN: 2395-5252

Classify $x \in \pi_1$ if $\ln(A/BR)+W>0$ (3.7.6)

Classify x as an element of π_2 if $\ln(L)\le 0$ or
Classify $x \in \pi_2$ if $\ln(A/BR)+W\le 0$ (3.7.7)

Classify x as an element of π_3 if $\ln(L)<0$ or
Classify $x \in \pi_3$ if $\ln(A/BR)+W<0$ (3.7.8)

where $W=[x-((\mu_1+\mu_2+\mu_3)/3)]((\mu_1-\mu_2-\mu_3)/\sigma)$

When $\mu_1 > \mu_2 > \mu_3$:
Classify x $\in \pi_1$ if $\ln(A/BR)-W>0$ (3.7.9)
Classify x $\in \pi_2$ if $\ln(A/BR)-W\le 0$ (3.8.0)
Classify x $\in \pi_3$ if $\ln(A/BR)-W<0$ (3.8.1)

When the parameters: μ_1, μ_2, μ_3 are unknown, they are estimated by $\bar{x}_1$, $\bar{x}_2$, $\bar{x}_3$ respectively.

Comparison with Normal Classification Rule:
Classify x $\in \pi_1$ if $x<((\mu_1+\mu_2+\mu_3)/3)$ (3.8.2)
Classify x $\in \pi_2$ if $x\ge((\mu_1+\mu_2+\mu_3)/3)$ (3.8.3)
Classify x $\in \pi_3$ if $x>((\mu_1+\mu_2+\mu_3)/3)$ (3.8.4)
It should be noted that QDA takes into account the covariance matrix of each population, whereas LDA assumes equal covariance matrices. MESD is used to model non-normal data.

IV. RESULTS OF ANALYSIS AND DISCUSSION

Table 4.1: Optimum Probabilities of Misclassification and Errors of Misclassification for LDA

Optimum Probabilities of Misclassification			
Population I	Population II	Population III	Total
0.00185	0.01130	3.71786×10^{-49}	0.01315
Errors of Misclassification			
Population I	Population II	Population III	Total
0	0.04857	0	0.04857

The result in Table 4.1 shows the optimum probability of misclassification for each population as well as the errors of misclassification for LDA. The optimum probability of misclassification for population Iis very low (0.00185), indicating that the model is highly accurate in classifying population I observations. The error of misclassification for population I is 0, which means that the model correctly classified all population I observations.The optimum probability of misclassification for population II is slightly higher (0.01130), indicating that the model is still accurate but slightly less so than for Population I. The error of misclassification is 0.04857, which means that the model misclassified approximately 4.86% of Population II observations.

The optimum probability of misclassification of Population III is extremely low (3.71786×10^{-49}), indicating that the model is highly accurate in classifying Population III observations.The error of misclassification is 0, which means that the model correctly classified all Population III observations. Hence, the results suggest that the model is highly accurate in classifying observations for all three populations, with Population I and Population III, having virtually no errors and Population II, having a small error rate. In addition, the total optimum

International Journal of Advances in Engineering and Management (IJAEM)
Volume 7, Issue 02 Feb. 2025, pp: 442-456 www.ijaem.net ISSN: 2395-5252

probability of misclassification (0.01315) suggests that the model has a relatively low probability of misclassifying observations across all populations. The total error of misclassification (0.04857) indicates that the model misclassifies approximately 4.86% of the observations across all populations. Hence; in overall, these metrics suggest that the model performs well in classifying observations, with a low probability of misclassification and a low error rate.

Table 4.2: Optimum Probabilities of Misclassification and Errors of Misclassification for QDA

Optimum Probabilities of Misclassification			
Population I	Population II	Population III	Total
0.00687	0.02395	2.13445×10^{-37}	0.03082
Errors of Misclassification			
Population I	Population II	Population III	Total
0	0.03714	0	0.03714

The result in Table 4.2 shows the optimum probability of misclassification for each population as well as the errors of misclassification for QDA. The optimum probability of misclassification for population Iis very low (0.00687), indicating that the model is highly accurate in classifying population I observations. The error of misclassification for population I is 0, which means that the model correctly classified all population I observations.The optimum probability of misclassification for population II is slightly higher (0.02395), indicating that the model is still accurate but slightly less so than for Population I. The error of misclassification is 0.03714, which means that the model misclassified approximately 3.71% of Population II observations.

The optimum probability of misclassification of Population III is extremely low (2.13445×10^{-37}), indicating that the model is highly accurate in classifying Population III observations.The error of misclassification is 0, which means that the model correctly classified all Population III observations. Hence, the results suggest that the model is highly accurate in classifying observations for all three populations, with Population I and Population III, having virtually no errors and Population II, having a small error rate. Furthermore, the total optimum probability of misclassification (0.03082) suggests that the model has a relatively low probability of misclassifying observations across all populations. The total error of misclassification (0.03714) indicates that the model misclassifies approximately 3.71% of the observations across all populations. Hence; in overall, these metrics suggest that the model performs well in classifying observations, with a low probability of misclassification and a low error rate.

Table 4.3:Summary of Multiple Metrics Statistics between LDA and QDA

		LDA			**QDA**		
		Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
Confusion Matrix	Pop I	350	17	0	350	13	0
	Pop II	0	333	0	0	337	0
	Pop III	0	0	350	0	0	350
Statistics by Class	Sensitivity	1.0000	0.9514	1.0000	1.0000	0.9629	1.0000
	Specificity	0.9757	1.0000	1.0000	0.9814	1.0000	1.0000
Accuracy		0.9838			0.9876		
AUC-ROC		0.99992925170068			1		

The result in Table 4.3 shows that for accuracy and error rates, both LDA and QDA achieve high accuracy rates, with QDA slightly outperforming LDA (98.76% vs. 98.38%). This indicates that both models are effective in classifying the data. However, QDA's lower error rate (1.24% vs. 1.62%) suggests it is more reliable. For confusion matrices, QDA's confusion matrix shows fewer misclassifications (13 vs. 17) compared to LDA. Specifically, QDA reduces misclassifications between Population 1 and Population 2, which is a common source of error.

International Journal of Advances in Engineering and Management (IJAEM)
Volume 7, Issue 02 Feb. 2025, pp: 442-456 www.ijaem.net ISSN: 2395-5252

For ROC Curve and AUC-ROC, QDA's perfect AUC-ROC score (1) indicates exceptional discriminatory power. LDA's AUC-ROC score (0.99992925170068) is also excellent but slightly lower. This suggests QDA is better at distinguishing between classes. For statistics by class, both models demonstrate high sensitivity and specificity for all classes. However, QDA shows improved sensitivity for Population 2 (0.9629 vs. 0.9514), indicating better detection of this class. Hence, QDA's superior performance across multiple metrics suggests it may be more robust. The results based on the extensive interpretation, concludes that QDA appears to be the better model due to its: Higher accuracy rate (98.76% vs. 98.38%); lower error rate (1.24% vs. 1.62%); improved misclassification reduction; exceptional discriminatory power (AUC-ROC = 1) and enhanced sensitivity for population 2.

4.1 DISCUSSION OF FINDINGS

This study investigated the apparent and optimum probabilities of misclassification for three populations from real-life anthropometric datasets. For LDA, the optimum probability of misclassification for population Iis very low (0.00185), indicating that the model is highly accurate in classifying population I observations. The error of misclassification for population I is 0, which means that the model correctly classified all population I observations.The optimum probability of misclassification for population II is slightly higher (0.01130), indicating that the model is still accurate but slightly less so than for Population I. The error of misclassification is 0.04857, which means that the model misclassified approximately 4.86% of Population II observations.The optimum probability of misclassification of Population III is extremely low $(3.71786 \times 10^{-49})$, indicating that the model is highly accurate in classifying Population III observations.The error of misclassification is 0, which means that the model correctly classified all Population III observations. Hence, the results suggest that the model is highly accurate in classifying observations for all three populations, with Population I and Population III, having virtually no errors and Population II, having a small error rate. Again, the total optimum probability of misclassification (0.01315) suggests that the model has a relatively low probability of misclassifying observations across all populations. The total error of misclassification (0.04857) indicates that the model misclassifies approximately 4.86% of the observations across all populations. Hence; in overall, these metrics suggest that the model performs well in classifying observations, with a low probability of misclassification and a low error rate. The results of this study support the findings of Kanuti and Ngaruye (2024) on asymptotic results for expected probability of misclassifications in linear discriminant analysis with repeated measurements and Gasana et al. (2024) and on moments of the likelihood-based discriminant function.The result of this study is in disagreement with the findings of Olusola and Onyeagu (2020) on binary classification problems in discriminant analysis using linear programming methods, and Nikita and Nikitas (2020) on sex estimation using various classification methods.

For QDA, the optimum probability of misclassification for population Iis very low (0.00687), indicating that the model is highly accurate in classifying population I observations. The error of misclassification for population I is 0, which means that the model correctly classified all population I observations.The optimum probability of misclassification for population II is slightly higher (0.02395), indicating that the model is still accurate but slightly less so than for Population I. The error of misclassification is 0.03714, which means that the model misclassified approximately 3.71% of Population II observations.The optimum probability of misclassification of Population III is extremely low $(2.13445 \times 10^{-37})$, indicating that the model is highly accurate in classifying Population III observations.The error of misclassification is 0, which means that the model correctly classified all Population III observations. Hence, the results suggest that the model is highly accurate in classifying observations for all three populations, with Population I and Population III, having virtually no errors and Population II, having a small error rate. Furthermore, the total optimum probability of misclassification (0.03082) suggests that the model has a relatively low probability of misclassifying observations across all populations. The total error of misclassification (0.03714) indicates that the model misclassifies approximately 3.71% of the observations across all populations. Hence; in overall, these metrics suggest that the model performs well in classifying observations, with a low probability of misclassification and a low error rate. The results of this study support the findings of Kouamo et al. (2020) who investigated QDA's performance in image classification tasks, highlighting its effectiveness and Li et al. (2020) whose work explored QDA's application in medical diagnosis, demonstrating its accuracy.

The findings from objective two conclude that for accuracy and error rates, both LDA and QDA achieve high accuracy rates, with QDA slightly outperforming LDA (98.76% vs. 98.38%). This indicates that both models are effective in classifying the data. However, QDA's lower error rate (1.24% vs. 1.62%) suggests it is more reliable. For confusion matrices, QDA's confusion matrix shows fewer misclassifications (13 vs. 17) compared to LDA. Specifically, QDA reduces misclassifications between Population 1 and Population 2, which is a common source of error. For ROC Curve and AUC-ROC, QDA's perfect AUC-ROC score (1) indicates exceptional discriminatory power. LDA's AUC-ROC score (0.99992925170068) is also excellent but slightly lower. This suggests QDA is better at distinguishing between classes. For statistics by class, both models demonstrate high sensitivity and specificity for all classes. However, QDA shows improved sensitivity for Population 2 (0.9629 vs. 0.9514), indicating better detection of this class. Hence, QDA's superior performance across multiple metrics suggests it may be more robust. The results based on the extensive interpretation, concludes that QDA appears to be the better model due to its: Higher accuracy rate (98.76% vs. 98.38%); lower error rate (1.24% vs. 1.62%); improved misclassification reduction; exceptional discriminatory power (AUC-ROC = 1) and enhanced sensitivity for population 2. The result of this study is in line with the result of Kouamo et al. (2020) who found that QDA outperformed LDA in image classification tasks; Li et al. (2020) whose work demonstrated QDA's superior performance in medical diagnosis, mirroring the current study's error rates (1.24% vs. 1.62%); Singh et al. (2022) whose findings showed QDA's exceptional discriminatory power, consistent with the current study's AUC-ROC scores (1 vs. 0.99992925170068); and Huang et al. (2020) who compared LDA and QDA performance in classification tasks and found QDA's superiority in accuracy and AUC-ROC. On the other hand, Wang et al. (2020) found LDA performed better in high-dimensional data, contrasting with the current study's findings.

V. CONCLUSION

This study focussed on estimating probabilities of misclassification for three populations via Edgeworth series distribution.Real-life anthropometric dataset was used, consisting of three populations of school learners and four variates each **(**Height, Head Circumference, Shoulder Width and Elbow Height).The study concluded that, for the real-life dataset, the total optimum probability of misclassification suggests that the model has a relatively low probability of misclassifying observations across all populations whereas the total error of misclassification indicates that the model misclassifies approximately 4.86% of the observations across all populations for LDA, whereas it misclassifies approximately 3.71% of the observations across all populations for QDA. Hence; in overall, these metrics concluded that the model performs well in classifying observations with a low probability.

REFERENCES

[1]. Anderson, T. W. (1958). An introduction to multivariate statistical analysis. John Wiley & Sons.

[2]. Awogbemi, C. A. & Onyeagu, S. I. (2019). Errors of Misclassification Associated with Edgworth Series Distribution (EDS). American Journal of Theoretical and Applied Statistics, 8(6), 203-213..

[3]. Bruno, M. A., Walker, E. A., & Abujudeh, H. H. (2015). Understanding and confronting our mistakes: the epidemiology of error in radiology and strategies for error reduction. Radiographics, 35(6), 1668-1676.

[4]. Chinganda, E. F. (1976). Distribution of the probability of misclassification. Journal of Multivariate Analysis, 6(2), 247-255.

[5]. Chinganda, E. F., & Kocherlakota, K. (1978). Robustness of the linear discriminant function to non-normality: An empirical study using Johnson's systems of distributions. Journal of Statistical Computation and Simulation, 7(3), 231-245.

[6]. Ching'anda, F. (1976). Classification errors and costs. Journal of Statistical Research, 10(1), 12-20.

[7]. Das Gupta, S. (1965). Optimum classification rules for two multivariate normal populations. Annals of Mathematical Statistics, 36(3), 828-835.

[8]. Das Gupta, S. (1973). Theoretical advances in Bayes and empirical Bayes inference. Journal of the American Statistical Association, 68(342), 347-353.

[9]. Fisher, R. A. (1936). The use of multiple measurements in taxonomic problems. Annals of Eugenics, 7(2), 179-188.
[10]. John, D. (2010). Errors of misclassification in statistical analysis. Journal of Data Science,
[11]. Hand, D. J. (2012). Assessing the performance of classification methods. International Statistical Review, 80(3), 400-414.
[12]. 8(3), 537-553.
[13]. Hills, M. (1966). Allocation rules and their error rates. Journal of the Royal Statistical Society. Series B (Methodological), 28(1), 1-20.
[14]. Gasana, E. U., von Rosen, D., & Singull, M. (2024). Moments of the likelihood-based discriminant function. Communications in Statistics-Theory and Methods, 53(3), 1122-1134.
[15]. Kanuti Ngailo, E., & Ngaruye, I. (2024). Asymptotic results for expected probability of misclassifications in linear discriminant analysis with repeated measurements. Communications in Statistics-Theory and Methods, 53(6), 1942-1963.
[16]. Kanuti, N. E., & Ngaruye, I. (2022). Asymptotic results for expected probability of misclassifications in linear discriminant analysis with repeated measurements. Communications in Statistics - Theory and Methods, 53(6), 1942–1963.
[17]. Kendall, M. G. (1966). Discrimination and classification: in multivariate analysis (2nd ed., pp. 165-183). Charles Griffin & Company.
[18]. Kocherlakota, S., Kocherlakota, K., & Balakrishnan, N. (1985). The bivariate Edgeworth series distribution: An alternative to the bivariate normal distribution. Journal of Multivariate Analysis, 17(2), 251-267.
[19]. Kocherlakota, S., Kocherlakota, K., & Balakrishnan, N. (1987). The effect of non-normality on the classification error rates in discriminant analysis. Journal of Statistical Planning and Inference, 16(2), 149-162.
[20]. Kouamo, S., Niang, I., & Diop, M. (2020). Quadratic discriminant analysis for image classification. Journal of Intelligent Information Systems, 57(2), 257-271.
[21]. Lachenbruch, P. A. (1973). Discriminant analysis (pp. 12-15). Hafner Publishing Company.
[22]. Lachenbruch, P. A. (1973). Discriminant analysis when the initial probabilities are unknown. Journal of the American Statistical Association, 68(342), 362-365.
[23]. Lachenbruch, P. A. (1973). Discriminant analysis. Hafner Publishing Company.
[24]. Lachenbruch, P. A., Clarke, W. R., & Lin, S. P. (1977). The effects of non-normality on the quadratic discriminant function. Biometrika, 64(2), 239-244.
[25]. Li, Y., Yang, P., & Zhang, Y. (2020). Quadratic discriminant analysis for medical diagnosis. Journal of Medical Systems, 44(10), 2109.
[26]. Mahmoud, M. A. W., & Moustafa, H. M. (1995). Errors of misclassification associated with gamma distribution. Mathematical and computer modelling, 22(3), 105-119.
[27]. Metsämuuronen, J. (2022). The effect of various simultaneous sources of mechanical error in the estimators of correlation causing deflation in reliability: Seeking the best options of correlation for deflation-corrected reliability. Behaviormetrika, 49(1), 91-130.
[28]. Moran, M. A. (1975). On the optimum allocation of observations in classification problems. Journal of the American Statistical Association, 70(352), 659-665.
[29]. Morgan, G. B., Hodge, K. J., & Baggett, A. R. (2016). Latent profile analysis with non-normal mixtures: A Monte Carlo examination of model selection using fit indices. Computational Statistics & Data Analysis, 93, 146-161.
[30]. Nikita, E., & Nikitas, P. (2020). Sex estimation: a comparison of techniques based on binary logistic, probit and cumulative probit regression, linear and quadratic discriminant analysis, neural networks, and naïve Bayes classification using ordinal variables. International journal of legal medicine, 134(3), 1213-1225
[31]. Olusola, M. O. & Onyeagu, S. I. (2020). On the binary classification problem in discriminant analysis using linear programming methods. Operations Research and Decisions, 30(1), 119-130.

[32]. Rao, C. R. (1965). Linear statistical inference and its applications. John Wiley & Sons.
[33]. Ruiz, D. A. P. (2019). Discriminant analysis: a functional perspective. The University of Manchester (United Kingdom).
[34]. Sedransk, N., & Okamoto, M. (1971). Approximations for the distribution of the W-statistic for testing the hypothesis of equality of means. Annals of Mathematical Statistics, 42(5), 1814-1825.
[35]. Sedransk, N., & Okamoto, M. (1971). Estimation of the probability of misclassification for a linear discriminant function. Annals of Mathematical Statistics, 42(4), 1279-1292.
[36]. Venkatesan, S. (2014). Common errors in scientific paper submissions: a reviewer's report. Journal of Social Sciences, 41(2), 279-293.
[37]. Wang, J. (2020). Classification and Misclassification in Statistics. Journal of Statistical Research, 54(2), 123-145.
[38]. Welch, B. L. (1939). Note on a method for correcting the effects of heterogeneity in the comparison of means and regression coefficients. Annals of Mathematical Statistics, 10(2), 235-240.
[39]. WWW.R-IDE (R studio). Com 10/07/2024