Journal of Economics and Allied Research Vol. 10, Issue 1, pp.282-321 (Mar. 2025) Print ISSN: 2536-7447 and E-ISSN: 3043-6591

AN OBJECTIVE APPRAISAL OF EDGEWORTH SERIES DISTRIBUTION AND NORMAL DISTRIBUTION FOR THREE POPULATIONS

NWANKWO-OBI .A. JANEFRANCES
Department of Statistics, Chukwuemeka Odumegwu Ojukwu University. Anambra State, Nigeria
Email: janenwankwo04@gmail.com, 08064172959

OKOLI CECILIA N.
Department of Statistics, Chukwuemeka Odumegwu Ojukwu University. Anambra State, Nigeria
Email:ceciliaokoli2@gmail.com, 08064063629

ABSTRACT
The study evaluates the optimum probabilities of misclassification using the Edgeworth Series Distribution (ESD) and compares the misclassification errors of ESD with the Normal Distribution (ND) for three populations using simulated data. It equally examined the adequacy of distribution performance between ESD and ND techniques and evaluates the performance of LDA and QDA in classifying ESD averaged over various sample sizes for three distinct populations. The optimal probabilities of misclassification for the Edgeworth Series Distribution (ESD) were computed with specific parameters ($\mu_1 = 0$, $\mu_2 = 1$, $\mu_3 = 1$ and $\sigma = 1$ with λ_4 being the skewness factor) within defined intervals (0.00625, 0.4 being in 14 intervals). The study also examined the apparent probabilities of misclassification for ESD and ND when means (μ_1, μ_2 and μ_3) are known or estimated from samples.. The findings of the study also revealed that QDA tends to have higher accuracy and AUC-ROC values than LDA across all the skewed levels. The study concluded that QDA outperformed LDA in terms of accuracy and error rates, demonstrating superior discriminatory power. This study provides 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. The research contributes to the advancement of robust classification methods and provides programming code for evaluation, enhancing the methodological toolkit in the field. It establishes a foundation for future research endeavours and presents 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: Optimal probability, Edgewoth Series, Discrimination, Quadratic, discriminant analysis

1. INTRODUCTION

In this work, we investigated the Edgeworth series distribution classification rule/technique and normal distribution classification rule/technique with regards to errors of misclassification for three populations. Error can be defined as an act or condition of ignorant or imprudent deviation from a code of behaviour 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).
Misclassification occurs when individuals are assigned to a different category than the one they should be in. This can lead to incorrect associations being observed between the assigned categories and the outcomes of interest (Fox et al., 2022). Discrimination and classification are

multivariate techniques that are based on a multivariate observations (Sharma et al., 2018). The aim of discrimination is to describe the differential features of observations that can separate the known populations. The aim of classification is to allocate a new observation to formerly defined groups (Wang, 2020). In practice, when we want to discriminate the known observations, first of all, we need to allocate them. Contrarily, a discriminator will be needed to allocate the observation, so the aims of discrimination and classification are regularly over lapped (as cited by Wang, 2020). 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). Awogbemi and Onyeagu (2019) studied on errors of misclassification associated with Edgeworth series distribution survey on two populations using small sample sizes. But this work majors on large sample sizes from three populations which none of the researchers sighted had written on. Also comparison on LDA and QDA on classification//misclassification with regards to Edgeworth series distribution have not been done by any researcher in history, hence the justification for this work.

This study is primarily concerned with evaluating objectively Edgeworth Series Distribution and Normal distribution for three populations. In specific terms, the researcher also seeks: to estimate the optimum probabilities of misclassification by ESD and errors of misclassification of Edgeworth series distribution (ESD) with Normal Distribution (ND) for three populations using simulated data; to investigate the distribution performance adequacy of ESD and ND Techniques; to compare the performance of Linear Discriminant Analysis (LDA) and Quadratic Discriminant Analysis (QDA) in classifying Edgeworth series distribution data averaged over different sample sizes for three distinct populations.

2. REVIEW OF RELATED LITERATURE

2.1. Empirical 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. Been well understood previously.

Mardia (2024) revisited Fisher's pioneering work on discriminant analysis and its significant impact on Artificial Intelligence. The study re-examined the famous iris data used in Fisher's 1936 study, testing the hypothesis of multivariate normality that Fisher had assumed. Mardia provided a deeper insight into Fisher's construction of the genetic discriminant, which had not been well understood previously. The study also explored how the field of discriminant analysis evolved with the computer revolution, highlighting newer methods such as: kernel classifiers, classification trees, support vector machines, neural networks and deep learning. Mardia noted that while computational power had shifted the emphasis of Multivariate Analysis,

Ngailo and Chuma (2023) investigated the classification of observations from repeated measurements using linear discriminant analysis. This common practice in fields like medicine, psychology, and environmental studies involves classifying data collected over time or under varying conditions. The researchers used an extended growth curve model to analyze repeated measurements and developed an approximation for misclassification probabilities in linear discriminant analysis. They derived the approximation for both known and unknown covariance matrices using specific statistical relationships. To evaluate the accuracy of their results, Ngailo and Chuma conducted a Monte Carlo simulation study. Their work provided a valuable contribution to the field by offering a reliable method for approximating misclassification probabilities in linear discriminant analysis with repeated measurements."
Xue et al. (2023) addressed the challenge of classifying high-dimensional functional data, where each observation is associated with multiple functional processes. Unlike existing methods that handle a single process or a few processes, this work tackled the complex inter-correlation structures among multiple processes. The researchers proposed a penalized classifier that achieves near-perfect classification and discriminant set inclusion consistency. This means that the classification-responsible functional predictors include those of the underlying optimal classifier. The challenges addressed by Xue et al. included: complex inter-correlation structures among multiple functional processes, truncation needed for approximation in functional data, difference in discriminant sets between infinite-dimensional and truncated optimal classifiers. Through simulation studies and real data applications, the researchers demonstrated the favourable performance of their proposed method."

Kanuti and Ngaruye (2022) investigated the misclassification probabilities in linear discriminant analysis (LDA) with repeated measurements. They proposed approximations for LDA misclassification probabilities when group means follow a bilinear regression structure. The researchers: derived a unified location and scale mixture expression for the standard normal distribution in LDA; obtained estimated approximations of misclassification probabilities for three cases: - Un-weighted case, weighted known covariance matrix and weighted unknown covariance matrix. The key findings revealed that larger p (number of repeated measurements) was beneficial for classification when the covariance matrix is known or in the un-weighted case. Again, when the covariance matrix is unknown, using fewer repeated measurements provided more information than using many measurements close to the sample size. The researchers validated their approximations through Monte Carlo simulations, confirming their accuracy."

Nikita and Nikitas (2020) conducted a study comparing seven techniques for sex estimation using ordinal variables: Binary logistic regression (BLR); Probit regression (PR); Cumulative probit regression (CPR); Linear discriminant analysis (LDA); Quadratic discriminant analysis (QDA); Artificial neural networks (ANN); and Naïve Bayes classification (NBC). They evaluated the performance of these methods using cranial and pelvic traits from the Athens Collection, a modern documented skeletal dataset. The researchers implemented an R package for cross-validated sex classification and discriminant function analysis. Additionally, they proposed a simple algorithm combining two discriminant functions. The results showed: small differences in classification performance among the methods; LDA was simpler, more flexible, and slightly outperformed BLR, NBC, and PR; combining pelvic and cranial traits via BLR or LDA discriminant functions: removed population-specificity, improved prediction accuracy above 97%. The study suggested that LDA might be the preferred method for skeletal sex estimation due to its simplicity, flexibility, and performance. The combination of traits and methods also demonstrated high accuracy and potential for practical applications.

Awogbemi and Onyeagu (2019) investigated the errors of misclassification associated with Edgeworth Series Distribution (ESD), focusing on the impact of non-normality on classification accuracy. They examined the effects of applying a normal classificatory rule to persistent non-normal distributions, comparing errors of misclassification between ESD and Normal Distribution (ND) across various small sample sizes and skewness levels. The study employed numerical inverse interpolation in R to generate uniformly distributed random variables and simulated 1000 configurations for each training sample, varying the skewness factor (λ_3) from 0.00625 to 0.4. The results showed that: as skewness increases, ESD's optimum misclassification probability (E_{12E}) decreases, while (E_{21E}) increases; the total probability of misclassification remains stable with increasing skewness; ESD's misclassification probabilities (E_{12E} and E_{21E}) are consistently higher than ND's (E_{12N} and E_{21N}) across all skewness levels. The findings suggested that the normal classification procedure was robust against departures from normality, maintaining stable total misclassification probabilities despite increasing skewness. The research provided valuable insights into the effects of non-normality on classification accuracy and the reliability of normal classificatory rules in real-world applications.

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 covariance matrix. The findings were: they found that larger p (number of repeated measurements) were better classified 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.

3. METHODOLOGY

3.1. Theoretical Framework:

Central Limit Theorem (CLT)

The central limit theorem says that the sampling distribution of the mean will always be normally distributed, as long as the sample size is large enough. Regardless of whether the population has a normal, poison, binomial, or any other distribution, the sampling distribution of the mean will be normal. (https://www.scribbr.com, 6 July 2022)

Normal distribution also known as Gaussian distribution is a probability distribution about the mean showing that data near the mean are more frequent in occurrence than data from the mean, The normal distribution appears as a' bell curved' when graphed

ESD and ND are both based on **central limit theorem** (CLT) which states that under certain condition,the sum of many independent random variables regardless of their original distribution will tend towards a normal distribution as the number of variables increases.

Essentially ESD acts as a way to approximate a distribution using the normal distribution as a base incorporating connections based on distribution 's moments like skewness and kurtosis through a series expansion.

Key elements of CLT are summarized below;

Population mean; The average value of the population

Population standard deviation; The measure of how spread out the population data is.

Sample mean; The average value of a given sample.

Standard error; The standard deviation of the sampling distribution, decreasing as sample size increasing.

Conditions for CLT to apply includes that samples should be randomly sampled, each sample should not influence the other (Independent sample) and finally samples should be sufficiently large enough to be considered adequate.

Application of CLT enables using normal distribution properties to make inferences about population parameters even when the original population distributions unknown.

It also underpins many statistical tests like hypothesis testing and confidence interval that rely on the normal distribution.

3.2 Methods

Examining the effects of non-normality in a three population discriminatory problem on errors of misclassification when Edgworth series distribution defined by Anderson's statistic (W) is used for classifying an observation as emanating from populations π_1, π_2 and π_3. The effects would be studied for varying values of Skewness factor based on the boundary of unimodal region for Edgeworth series distribution. Optimum probabilities of misclassification (OPM) by ESD would be computed from known parameter as well as estimating the probabilities of misclassification by ESD, where the apparent probabilities of misclassification (APM) in respect of ESD for known and estimated parameters are generated. To generate random variables from the ESD, the study used the method of numerical inverse interpolation. Schmidt, and Taylor, (1970) described this method in details. This work would also be analysed using RStudio programming package that gave in-depth analysis of the study.

3.3 Edgeworth Series Distribution (ESD)

The Edgeworth series distribution is a continuous probability distribution that approximates a probability distribution in terms of its cumumulants and Hermite polynomials. It relates the probability density function (PDF) to a standard normal distribution PDF. It is sometimes seen in statistical asymptotic theory, where approximations to sample statistic distributions of order greater than $n^{-\frac{1}{2}}$ are calculated (Adeyeye, 2020).

The ESD has been used for some practical purposes, including the study of nonlinear gust loading factors (used in the design of structures exposed to extreme winds).

Note that we are treating the effect of non-normality in a three population discrimination problem. So, we assume the distributions in the three populations to be univariate Edgworth series with different means and the variance are equal.

Here also, we only consider non-normality due to Skewness, regardless of the fact that some authors/writers have written on non-normality which considers both Kurtosis and Skewedness in their standard forms.

Let $x_{1j},\ x_{2j}\ and\ x_{3j}$ denote three independent random samples from three populations, $\pi_1, \pi_2\ and\ \pi_3$ respectively where (j = 1,2,3, …n_1), $(j = 1,2,3 \dots, n_2)$ for x_{2j} and $j = 1, 2, 3, \dots n_3$ for x_{3j}.

Then the density function of x_{ij}

becomes $f(x) = \left(1 - \frac{\Lambda_3}{6} D^3\right) \emptyset \frac{(x-\mu_1)}{\sigma} - \infty < x < \infty$ (1)

and that of x_{2j} becomes

$$f(x) = \left(1 - \frac{\Lambda_3}{6} D^3\right) \emptyset \frac{(x-\mu_2)}{\sigma}, -\infty < x < \infty \quad (2)$$

and that of x_{3j} becomes

$$f(x) = \left(1 - \frac{\Lambda_3}{6} D^3\right) \emptyset \frac{(x-\mu_3)}{\sigma}, -\infty < x < \infty \quad (3)$$

Where $\Lambda_3, \mu_i (i = 1,\ 2, 3) and\ \delta$ satisfy the conditions, $-\infty < \Lambda_3 < \infty$, and $\sigma > 0$

Here D represents the operator $\frac{\partial}{\partial x}$ $and\ \emptyset(\frac{x-\mu_1}{\sigma})$ is the density function

$$(3\pi)^{-\frac{1}{2}} \sigma^{-1} \exp[\frac{(x-\mu_i)}{3\sigma^2} \quad (4)$$

and Λ_3 is the Skewness factor

It is equally to note that all terms involving powers of Λ_3 higher than the first are ignored.

If x is a new observation, obtained independently of observation $x_{1j}, x_{2j}\ and\ x_{3j}$ drawn from either population π_1, π_2.or π_3. In other to do this, a classification rule is needed, this implies that the discriminant function has to be obtained; so in practice, one could use the univariate analogue of the w discriminant function which is defined as;

$$w = D\ (x; \overline{x}_1, \overline{x}_2, \overline{x}_3, \sigma^2) = [x - \frac{1}{3}(\overline{x}_1 + \overline{x}_2\ + \overline{x}_3)] \frac{(\overline{x}_1 - \overline{x}_2 - \overline{x}_3\ )}{\delta^2} \quad (5)$$

When σ^2 is known and

When σ^2 is estimated by $S_,^2$ the pooled sample variance of the observation in population π_1, π_2 and population π_3;

$$w = D\ (x; \overline{x}_1, \overline{x}_2,\ \overline{x}_{3.} S^2) = [x - \frac{1}{3}(\overline{x}_1 + \overline{x}_2\ + \overline{x}_3)] \frac{(\overline{x}_1 - \overline{x}_2 - \overline{x}_3)}{S^2} \quad (6)$$

3.4 Optimum Probability of Misclassification of ESD

The probability of misclassification is said to be optimum, when all parameters of the distributions in the populations are known. It is optimal by implication that we cannot improve upon it. When an observation from population π_1 is misclassified, its probability of the misclassification becomes;

$$\alpha_1[R, F] = P_r\left[x \geq \left(\frac{\mu_1 + \mu_2\ + \mu_3}{3}\right)\right] \quad (7)$$

$$= \int_\sigma^\infty \left[1 - \frac{\Lambda_3}{6} D^3\right] \varphi\,(\frac{\mathrm{x} - \mu_1}{\sigma}) dx$$

$$= \int_\sigma^\infty \left[1 + \frac{\Lambda_3}{6\sigma^3} H_3\left(\frac{x - \mu_1}{\sigma}\right)\right] \varphi\,(\frac{x - \mu_1}{\sigma}) dx$$

$$= \int_\sigma^\infty \varphi\left(\frac{x-\mu_1}{\sigma}\right) dx + \frac{\Lambda_3}{6\sigma^3} \int_\sigma^\infty H_3(\frac{x-\mu_1}{\sigma}) \varphi(\frac{x-\mu_1}{\sigma}) dx \quad (8)$$

where $\sigma = (\frac{\mu_1+\mu_2\ +\mu_3}{\sigma})$ and $H_n(x) is$ Chebyshev's Hermite polynomial of degree r and defined by the identity:

$$H_n(x)\ \varphi\ (x) = (-D)^n\ \varphi\ (x) \qquad (9)$$

(see Kendall and Stuart, 1958)

If φ (x) denotes the standard normal density function, then we define the Hermite Polynomial $H_n(x)$ for any integral n by

$$\frac{(-1)^n\ d^n}{\sqrt{2\pi}\ dx^n} \int_{-\infty}^{\infty} e^{itx}\ e\text{-}t^{2/2}dt = (-1)^n \frac{d^n}{dx^n}\ \varphi(x) = H_n(x)\ \varphi(x)$$

and putting $Z = \frac{x-\mu_1}{\sigma}$ in (3.47)

we get

$$\alpha_1[R,F] = \int_{\frac{\sigma-\mu_1}{\sigma}}^{\infty} \varphi(z)dz + \frac{\kappa_3}{6\sigma^2}\int_{\frac{\sigma-\mu_1}{\sigma}}^{\infty} H_3(z)\varphi\,(z)dz$$

$$= 1 - \varphi\left[(\frac{\sigma-\mu_1}{\sigma}) + \frac{\kappa_3}{6\sigma^2}H_2[\left(\frac{\sigma-\mu_1}{\sigma}\right)\right]\varphi\left(\frac{\sigma-\mu_1}{\sigma}\right)$$

$$= 1 - \varphi\left[\left(\frac{\mu_3-\mu_2-\mu_1}{3\sigma}\right) + \frac{\kappa_3}{6\sigma^2}\left(\frac{\mu_3-\mu_2-\mu_1}{3\sigma}\right) - 1\right]\varphi\,(\frac{\mu_3-\mu_2-\mu_1}{3\sigma}) \qquad (10)$$

When an observation from population π_2 is misclassificatied, the optimum probability of its misclassification becomes

$$\alpha_2[R,F] = P_r\{x < \left(\frac{\mu_1+\mu_2\ \ +\mu_3}{3}\right)\}$$

$$= \int_{-\infty}^{\sigma}\left[1 - \frac{\kappa_3}{\sigma}D^3\right]\varphi\,(\frac{x-\mu_2}{\sigma})dx$$

$$= \int_{-\infty}^{\sigma}\varphi\left(\frac{x-\mu_2}{\sigma}\right) + \frac{\kappa_3}{6\sigma^3}\int_{-\infty}^{\sigma}H_3\left(\frac{x-\mu_2}{\sigma}\right)\varphi(\frac{x-\mu_2}{\sigma})dx \qquad (11)$$

Putting $\sigma = (\frac{\mu_1+\mu_2\ +\mu_3}{2})$ and $z = \frac{x-\mu_2}{\sigma}$

we get

$$\alpha_2[R,F] = P_r\int_{-\infty}^{\frac{\sigma-\mu_2}{\sigma}} \emptyset(z)dz + \frac{\kappa_3}{6\sigma^2}\int_{-\infty}^{\frac{\sigma-\mu_2}{\sigma}} H_3(z)\varphi(z)dz$$

$$= \varphi\left[\frac{\sigma-\mu_2}{\sigma}\right] - \frac{\kappa_3}{6\sigma^2}H_2\left[\frac{\sigma-\mu_2}{\sigma}\right]\varphi[\frac{\sigma-\mu_2}{\sigma}]$$

$$= \varphi\left[\frac{\mu_1-\mu_2\ -\mu_3}{3\sigma}\right] - \frac{\kappa_3}{6\sigma^2}\left[(\frac{\mu_1-\mu_2\ -\mu_3}{3\sigma})^2 - 1\right]\varphi\,(\frac{\mu_1-\mu_2\ -\mu_3}{3\sigma}) \qquad (12)$$

When an observation from population π_3 is misclassificatied, the optimum probability of misclassification is given by

$$\alpha_3[R,F] = P_r\{x \leq \left(\frac{\mu_1+\mu_2\ \ +\mu_3}{3}\right)\}$$

$$= \int_{-\infty}^{\sigma} \left[1 - \frac{\kappa_3}{\sigma} D^3\right] \varphi\left(\frac{x-\mu_3}{\sigma}\right) dx \quad (13)$$

$$= \int_{-\infty}^{\sigma} \varphi\left(\frac{x-\mu_3}{\sigma}\right) + \frac{\kappa_3}{6\sigma^3} \int_{-\infty}^{\sigma} H_3\left(\frac{x-\mu_3}{\sigma}\right) \varphi\left(\frac{x-\mu_3}{\sigma}\right) dx \quad (14)$$

Putting $\sigma = \left(\frac{\mu_1+\mu_2+\mu_3}{3}\right)$ and z = $\frac{x-\mu_3}{\sigma}$ in equation 84

we get

$$\alpha_3[R, F] = P_r \int_{-\infty}^{\frac{\sigma-\mu_3}{\sigma}} \emptyset(z) dz + \frac{\kappa_3}{6\sigma^2} \int_{-\infty}^{\frac{\sigma-\mu_3}{\sigma}} H_3(z)\varphi(z) dz$$

$$= \varphi\left[\frac{\sigma-\mu_3}{\sigma}\right] - \frac{\kappa_3}{6\sigma^2} H_2 \left[\frac{\sigma-\mu_3}{\sigma}\right] \varphi\left[\frac{\sigma-\mu_3}{\sigma}\right]$$

$$= \varphi\left[\frac{\mu_1-\mu_2-\mu_3}{3\sigma}\right] - \frac{\kappa_3}{6\sigma^2}\left[\left(\frac{\mu_1-\mu_2-\mu_3}{3\sigma}\right)^2 - 1\right] \varphi\left(\frac{\mu_1-\mu_2-\mu_3}{3\sigma}\right) \quad (15)$$

The optimum probability of misclassification is of important for comparison purposes and it is very useful in this work.

3.5. Model Specifications:

Model Adequacy for the Difference ESD and ND Techniques

Wilcoxon rank sum test was employed to examine the relationship of errors of misclassification values averaged over small samples between ESD and ND techniques.

3.6 Wilcoxon Rank Sum Test

Wilcoxon rank sum test (WRST) was developed by an American statistician, Frank Wilcoxon, who worked in the chemical industry in 1945 (Bangdiwala, 2013). The statistic claims that given two sets of data say Z and Y from independent continuous distributions, the ranks of the $Z's$ in the combined ordered arrangement of the two sets would generally be larger than the ranks of the $Y's$ if the median of the Z population exceeds that of the Y population. Following the argument, he proposed a test where the location alternative hypothesis, $H_1: \theta \neq 0$ is not rejected if the sum of ranks of the $Z's$ is either too large or too small (Solaro et al., 2021). In other words, in the WRST, the values of the data for both samples Z and Y are combined and then ranked. If the null hypothesis (H_0: θ = 0) is true, then there is no difference in the population distributions – and the values in each set should be ranked approximately the same. Therefore, when the ranks are summed for each set, the sums should be approximately equal, and the null hypothesis (H_0) will not be rejected. If there is a large difference in the sums of the ranks, then the distributions are not identical, and H_0 will be rejected.

For large samples, the normal approximation to the distribution or rejection regions for W can be used because of the asymptotic normality of the general linear rank statistic (Beasley et al., 2009). This approximation is shown to be accurate enough for most practical applications for combined sample sizes $N \geq 12$ (Bellera et al., 2010). The normal distribution approximates the Wilcoxon rank sum statistic T as (Harris & Hardin, 2013):

$$Z = \frac{|R - \mu| - 0.5}{\sigma} \quad (16)$$

0.5 in Equation (92) is the continuity correction term required since T is not a continuous random variable. When ties are included in the ranking, the mid-rank method is easily applied

to handle the problem of ties. The presence of a moderate number of tied observations seems to have little effect on the probability distribution (Gibbons, 2003).
Given H_0,

$$\mu = \frac{m(N+1)}{2} \tag{17}$$

and

$$\sigma^2 = \frac{mn(N+1)}{12} \tag{18}$$

respectively. R is the sum of ranks for smaller sample size (n), m is the larger of sample sizes, $N = m + n.$

A table of critical values corresponding to WRST is contained in any standard text. The table gives the rejection regions for level of significance α, and the sample sizes m and n. Of course, if Z is less than or greater than the critical values, the decision is to reject the null hypothesis in favour of the alternative.

3.7 Choice of skewness Factor Value

The choice of the value of the Skewness factor λ_4, lay its emphasis on the boundary of the unimodal region for Edgeworth series distribution, and this is where the probability density function is only cogent. With this reason, the Skewness factor is chosen to be in the range (0.00625, 0.4) or (6.25x10^{-3}, 4x10^{-1}) (Barton, D. E., & Benin, N. (1952)), Draper, N. R., & Tierney, D. E. (1972)

3.8 Simulated Data from ESD (Generation of Data from ESD)

The optimum probabilities of misclassification for the Edgeworh Series Distribution (ESD) are computed with $\mu_1 = 0$, $\mu_2 = 1$, $\mu_3 = 1$ and $\sigma = 1$ with λ_4 being the skewness factor within the interval (0.00625, 0.4).

The apparent probabilities of misclassification for the (ESD) and Normal Distribution (ND) were also examined when the means (μ_1, μ_2, and μ_3) are known and when the parameters are estimated from the samples. Three independent samples of simulation size of 200 each were configured at each value of the skewness factor (λ_4) from three populations (π_1, π_2 and π_3) whose distributions are of ESD with the respective parameters: ($\mu_1 = 0$, $\sigma_1 = 1$), ($\mu_2 = 1, \sigma_2 = 1$) and ($\mu_3 = 1$, $\sigma_3 = 1$).
Employing the ESD and ND classification rules, the proportion misclassified in π_1, π_2 and π_3 were obtained and repeated for small samples (n = 4, 8, 12, 16, 20, 24, 28).

4. RESULTS OF DATA ANALYSIS AND DISCUSSION

4.1. Results of the Simulation Experiments

The optimum probabilities of misclassification for the Edgeworh Series Distribution (ESD) are computed with $\mu_1 = 0$, $\mu_2 = 1$, $\mu_3 = 1$ and $\sigma = 1$ with λ_4 being the skewness factor within the interval (0.00625, 0.4).The apparent probabilities of misclassification for the (ESD) and

Normal Distribution (ND) were also examined when the means (μ_1, μ_2, and μ_3) are known and when the parameters are estimated from the samples. Three independent samples of simulation size of 200 each were configured at each value of the skewness factor (λ_4) from three populations (π_1, π_2 and π_3) whose distributions are of ESD with the respective parameters: ($\mu_1 = 0$, $\sigma_1 = 1$), ($\mu_2 = 1, \sigma_2 = 1$) and ($\mu_3 = 1$, $\sigma_3 = 1$).

Employing the ESD and ND classification rules, the proportion misclassified in π_1, π_2 and π_3 were obtained and repeated for small samples (n = 4, 8, 12, 16, 20, 24, 28). The random numbers were generated using RStudio program and simulation results were obtained and displayed in Tables 4.1- 4.3

Table 4.1: Optimum Probabilities of Misclassification at Various Skewness Values for ESD

	Optimum Probabilities of Misclassification			
Skewness Factor (λ_4)	E_{1E}	E_{2E}	E_{3E}	Total
6.25×10^{-3}	0.285	0.305	0.270	0.860
1.25×10^{-2}	0.305	0.320	0.305	0.930
0.025	0.295	0.310	0.300	0.905
0.050	0.300	0.280	0.295	0.875
0.085	0.285	0.295	0.305	0.885
0.120	0.300	0.285	0.300	0.885
0.155	0.305	0.300	0.315	0.920
0.190	0.320	0.315	0.305	0.940
0.225	0.320	0.305	0.300	0.925
0.260	0.295	0.295	0.315	0.905
0.295	0.310	0.295	0.315	0.920
0.330	0.310	0.280	0.305	0.895
0.365	0.285	0.275	0.290	0.850
0.400	0.310	0.305	0.290	0.905

Source; IDE, R-Version 4.4.1, R-studio

The result in Table 4.1 shows the optimum probability of misclassification for each population at various skewness levels for Edgeworh Series Distribution. At skewness 0.00625, the optimum probabilities of misclassification for populations one, two and three are 0.285, 0.305 and 0.270 respectively, whereas its sum of optimum probabilities is 0.860. At skewness 0.01250, the optimum probabilities of misclassification for populations 1, 2 and 3 are 0.305, 0.320 and 0.305 respectively, whereas its sum of optimum probabilities is 0.930. Also, the values of the optimum probabilities of misclassification for populations one, two and three, as well as their sum of optimum probabilities for different skewness factors (0.025, 0.050, 0.085, 0.120, 0.155, 0.190, 0.225, 0.260, 0.295, 0.330, 0.365 and 0.400) are also presented. At lower skewness levels (0.00625-0.025), population 2 has the highest probability of misclassification (0.305-0.310), followed by Population 1 (0.285-0.295) and Population 3 (0.270-0.300). In general, the optimum probabilities of misclassification vary across populations and skewness levels, but their variations are relatively close, indicating similarities in the populations.

Table 4.2: Comparison of Errors of Misclassification of ESD with ND Averaged over 4 Samples for all Known Parameters with Simulation Size of 200

Skewness Factor (λ_4)	E_{1E}	E_{2E}	E_{3E}	Total	E_{1N}	E_{2N}	E_{3N}	Total
6.25×10^{-3}	0.30625	0.29875	0.30250	0.90750	0.29125	0.30000	0.29625	0.88750
1.25×10^{-2}	0.30875	0.30250	0.29000	0.90125	0.29375	0.30375	0.30375	0.90125
0.025	0.30375	0.30375	0.30375	0.91125	0.30250	0.30500	0.30125	0.90875
0.050	0.31500	0.29250	0.30125	0.90875	0.28875	0.28875	0.30625	0.88375
0.085	0.30500	0.30000	0.31375	0.91875	0.29750	0.29625	0.29000	0.88375
0.120	0.31250	0.29750	0.30750	0.91750	0.30125	0.30375	0.30000	0.90500
0.155	0.29750	0.28750	0.30750	0.89250	0.30000	0.29875	0.29375	0.89250
0.190	0.30250	0.29750	0.30125	0.90125	0.30625	0.30000	0.29500	0.90125
0.225	0.31250	0.30375	0.30125	0.91750	0.30875	0.30000	0.30375	0.91250
0.260	0.30000	0.28750	0.29500	0.88250	0.30125	0.30000	0.30250	0.90375
0.295	0.30375	0.29875	0.30000	0.90250	0.30625	0.29625	0.30625	0.90875
0.330	0.29625	0.31375	0.30625	0.91625	0.29625	0.31375	0.31375	0.92375
0.365	0.30750	0.31250	0.28750	0.90750	0.30750	0.29875	0.29375	0.90000
0.400	0.29500	0.30125	0.30500	0.90125	0.30125	0.29125	0.28625	0.87875

Source: IDE, R-Version 4.4.1, R-studio

Table 4.2 shows results of the simulation size of 200, which compares the performance of the Edgeworth Series Distribution (ESD) and Normal Distribution (ND) methods averaged over 4 samples for estimating probabilities of misclassification across different populations and skewness levels. The probabilities of misclassification vary across populations and skewness levels, but their variations are relatively close between the two methods, indicating that both methods perform similarly. However, the ESD method tends to have slightly higher probabilities of misclassification compared to the ND method, especially for Population 1 at skewness levels (6.25×10^{-3} – 0.12).

The ESD and ND classification procedures have similar total probability of misclassification at all λ_4 values. The total probability of misclassification values shows that using a small sample of 4 to estimate μ_1, μ_2, and μ_3, results is either underestimation or overestimation for each value of λ_4. The skewness component (λ_4) has minimal effect on the overall probability of misclassification, indicating that it is not affected by deviations from normality. Based on these values, the probabilities of misclassification across all populations can be considered relatively high, as they exceed 0.2 (20%) and are close to 0.3 (30%)

Table 4.3: Comparison of Errors of Misclassification of ESD with ND Averaged over 8 Samples for all Known Parameters with Simulation Size of 200

	E_{1E}	E_{2E}	E_{3E}	Total	E_{1N}	E_{2N}	E_{3N}	Total
6.25×10^{-3}	0.30625	0.29813	0.29438	0.89876	0.29313	0.30500	0.30250	0.90063
1.25×10^{-2}	0.31188	0.29063	0.30500	0.90751	0.29875	0.30063	0.30313	0.90251
0.025	0.31000	0.31000	0.30500	0.92500	0.29625	0.29688	0.29938	0.89251
0.050	0.30188	0.29125	0.30813	0.90126	0.30188	0.30188	0.29438	0.89814
0.085	0.30375	0.30438	0.29813	0.90626	0.30938	0.30313	0.30063	0.91314
0.120	0.30188	0.29188	0.29688	0.89064	0.29750	0.29938	0.30563	0.90251
0.155	0.30313	0.30500	0.29500	0.90313	0.29938	0.30438	0.30813	0.91189
0.190	0.29875	0.30250	0.29688	0.89813	0.29938	0.28875	0.29313	0.88126
0.225	0.30375	0.30313	0.30188	0.90876	0.30625	0.30938	0.30188	0.91751
0.260	0.30063	0.29875	0.29375	0.89313	0.29813	0.30250	0.30125	0.90188
0.295	0.30125	0.29875	0.30500	0.90500	0.29438	0.29375	0.29750	0.88563
0.330	0.30125	0.29438	0.30438	0.90001	0.29813	0.29688	0.29813	0.89314
0.365	0.30250	0.29813	0.29875	0.89938	0.29563	0.29625	0.29875	0.89063
0.400	0.30188	0.28688	0.29438	0.88314	0.30250	0.30188	0.30250	0.90688

Source: IDE, R-Version 4.4.1, R-studio

Table 4.3 shows results of the simulation size of 200, which compares the performance of the Edgeworth Series Distribution (ESD) and Normal Distribution (ND) methods averaged over 8 samples for estimating probabilities of misclassification across different populations and skewness levels. The probabilities of misclassification vary across populations and skewness levels, but their variations are relatively close between the two methods, indicating that both methods perform similarly. However, the ESD method tends to have slightly higher or equal probabilities of misclassification compared to the ND method, especially for Population 1 at skewness levels (6.25×10^{-3} – 0.05).

The ESD and ND classification procedures have similar total probability of misclassification at all λ_4 values. The total probability of misclassification values shows that using a small sample of 8 to estimate μ_1, μ_2, and μ_3, results is either underestimation or overestimation for each value of λ_4. The skewness component (λ_4) has minimal effect on the overall probability of misclassification, indicating that it is not affected by deviations from normality. Based on these values, the probabilities of misclassification across all populations can be considered relatively high, as they exceed 0.2 (20%) and are close to 0.3 (30%).

Comparison of Errors of Misclassification of ESD with ND Averaged over 12 , 16,20,24, and 28 for all known Parameters with Simulation Size of 200 for estimating probabilities of misclassification and skewness level vary across population indicating that both methods perform similarly.

Table 4.4: Summary of Decision for Testing Errors of Misclassification Values Averaged over 4 Samples with Computations: ESD vs. ND

			Errors of Misclassification		Ranks		Z	Decision
S/N	Population	SKN	ESD	ND	ESD	ND		
1		1	0.30625	0.29125	20.0	2.0		
2		2	0.30875	0.29375	24.5	3.0		

3	1	3	0.30375	0.30250	16.5	14.5	1.68	Do not Reject H_0
4		4	0.31500	0.28875	28.0	1.0		
5		5	0.30500	0.29750	18.0	7.5		
6		6	0.31250	0.30125	26.5	12.0		
7		7	0.29750	0.30000	7.5	9.5		
8		8	0.30250	0.30625	14.5	20.0		
9		9	0.31250	0.30875	26.5	24.5		
10		10	0.30000	0.30125	9.5	12.0		
11		11	0.30375	0.30625	16.5	20.0		
12		12	0.29625	0.29625	5.5	5.5		
13		13	0.30750	0.30750	22.5	22.5		
14		14	0.29500	0.30125	4.0	12.0		
15	2	1	0.29875	0.30000	11.5	16.0	0.05	Do not Reject H_0
16		2	0.30250	0.30375	20.0	22.5		
17		3	0.30375	0.30500	22.5	25.0		
18		4	0.29250	0.28875	5.0	3.0		
19		5	0.30000	0.29625	16.0	6.5		
20		6	0.29750	0.30375	8.5	22.5		
21		7	0.28750	0.29875	1.5	11.5		
22		8	0.29750	0.30000	8.5	16.0		
23		9	0.30375	0.30000	22.5	16.0		
24		10	0.28750	0.30000	1.5	16.0		
25		11	0.29875	0.29625	11.5	6.5		
26		12	0.31375	0.31375	27.5	27.5		
27		13	0.31250	0.29875	26.0	11.5		
28		14	0.30125	0.29125	19.0	4.0		
29	3	1	0.30250	0.29625	16.5	9.0	0.85	Do not Reject H_0
30		2	0.29000	0.30375	3.5	19.0		
31		3	0.30375	0.30125	19.0	13.5		
32		4	0.30125	0.30625	13.5	23.0		
33		5	0.31375	0.29000	27.5	3.5		
34		6	0.30750	0.30000	25.5	10.5		
35		7	0.30750	0.29375	25.5	5.5		
36		8	0.30125	0.29500	13.5	7.5		
37		9	0.30125	0.30375	13.5	19.0		
38		10	0.29500	0.30250	7.5	16.5		
39		11	0.30000	0.30625	10.5	23.0		
40		12	0.30625	0.31375	23.0	27.5		
41		13	0.28750	0.29375	2.0	5.5		
42		14	0.30500	0.28625	21.0	1.0		
43	Total	1	0.90750	0.88750	17.5	5.0	1.40	Do not Reject H_0
44		2	0.90125	0.90125	11.0	11.0		
45		3	0.91125	0.90875	22.0	20.0		
46		4	0.90875	0.88375	20.0	3.5		
47		5	0.91875	0.88375	27.0	3.5		
48		6	0.91750	0.90500	25.5	16.0		
49		7	0.89250	0.89250	6.5	6.5		
50		8	0.90125	0.90125	11.0	11.0		
51		9	0.91750	0.91250	25.5	23.0		
52		10	0.88250	0.90375	2.0	15.0		
53		11	0.90250	0.90875	14.0	20.0		
54		12	0.91625	0.92375	24.0	28.0		
55		13	0.90750	0.90000	17.5	8.0		
56		14	0.90125	0.87875	11.0	1.0		

Source: IDE, R-Version 4.4.1, R-studio

Table 4.5: Summary of Decision for Testing Errors of Misclassification Values Averaged Over 8 Samples with Computations: ESD vs. ND

			Errors of Misclassification		Ranks		Z	Decision
S/N	Population	SKN	ESD	ND	ESD	ND		
1	1	1	0.30625	0.29313	24.5	1.0	2.69	Reject H_0
2		2	0.31188	0.29875	28.0	8.5		
3		3	0.31000	0.29625	27.0	4.0		
4		4	0.30188	0.30188	16.5	16.5		
5		5	0.30375	0.30938	22.5	26.0		
6		6	0.30188	0.29750	16.5	5.0		
7		7	0.30313	0.29938	21.0	10.5		
8		8	0.29875	0.29938	8.5	10.5		
9		9	0.30375	0.30625	22.5	24.5		
10		10	0.30063	0.29813	12.0	6.5		
11		11	0.30125	0.29438	13.5	2.0		
12		12	0.30125	0.29813	13.5	6.5		
13		13	0.30250	0.29563	19.5	3.0		
14		14	0.30188	0.30250	16.5	19.5		
15	2	1	0.29813	0.30500	11.5	25.5	0.67	Do not Reject H_0
16		2	0.29063	0.30063	3.0	16.0		
17		3	0.31000	0.29688	28.0	9.5		
18		4	0.29125	0.30188	4.0	17.5		
19		5	0.30438	0.30313	23.5	21.5		
20		6	0.29188	0.29938	5.0	15.0		
21		7	0.30500	0.30438	25.5	23.5		
22		8	0.30250	0.28875	19.5	2.0		
23		9	0.30313	0.30938	21.5	27.0		
24		10	0.29875	0.30250	13.5	19.5		
25		11	0.29875	0.29375	13.5	6.0		
26		12	0.29438	0.29688	7.0	9.5		
27		13	0.29813	0.29625	11.5	8.0		
28		14	0.28688	0.30188	1.0	17.5		
29	3	1	0.29438	0.30250	4.0	19.5	0.44	Do not Reject H_0
30		2	0.30500	0.30313	24.0	21.0		
31		3	0.30500	0.29938	24.0	14.0		
32		4	0.30813	0.29438	27.5	4.0		
33		5	0.29813	0.30063	10.5	15.0		
34		6	0.29688	0.30563	7.5	26.0		
35		7	0.29500	0.30813	6.0	27.5		
36		8	0.29688	0.29313	7.5	1.0		
37		9	0.30188	0.30188	17.5	17.5		
38		10	0.29375	0.30125	2.0	16.0		
39		11	0.30500	0.29750	24.0	9.0		
40		12	0.30438	0.29813	22.0	10.5		
41		13	0.29875	0.29875	12.5	12.5		
42		14	0.29438	0.30250	4.0	19.5		
43		1	0.89876	0.90063	11.0	14.0		
44		2	0.90751	0.90251	23.0	17.5		

45	Total	3	0.92500	0.89251	28.0	6.0	0.25	Do not Reject H_0
46		4	0.90126	0.89814	15.0	10.0		
47		5	0.90626	0.91314	21.0	26.0		
48		6	0.89064	0.90251	5.0	17.5		
49		7	0.90313	0.91189	19.0	25.0		
50		8	0.89813	0.88126	9.0	1.0		
51		9	0.90876	0.91751	24.0	27.0		
52		10	0.89313	0.90188	7.0	16.0		
53		11	0.90500	0.88563	20.0	3.0		
54		12	0.90001	0.89314	13.0	8.0		
55		13	0.89938	0.89063	12.0	4.0		
56		14	0.88314	0.90688	2.0	22.0		

Source: IDE, R-Version 4.4.1, R-studio

Table 4.6: Summary of Multiple Metrics Statistics between LDA and QDA from ESD Averaged over 4 Samples with Simulation Size of 200

Skew			LDA			QDA		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.00625	**Statistics by Class**	Sensitivity	0.7600	0.540	0.3300	0.7500	0.495	0.4250
		Specificity	0.8075	0.735	0.7725	0.8175	0.765	0.7525
	Accuracy		0.5750			0.5841		
	AUC-ROC		0.7709			0.7823		
			LDA			**QDA**		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.0125	**Statistics by Class**	Sensitivity	0.765	0.4850	0.3700	0.775	0.600	0.3600
		Specificity	0.815	0.7225	0.7725	0.815	0.695	0.8575
	Accuracy		0.5658			0.5625		
	AUC-ROC		0.7641			0.7706		
			LDA			**QDA**		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.0025	**Statistics by Class**	Sensitivity	0.7700	0.5200	0.3950	0.745	0.5250	0.425
		Specificity	0.8475	0.7375	0.7575	0.860	0.7325	0.755
	Accuracy		0.5521			0.5795		
	AUC-ROC		0.7686			0.7801		
			LDA			**QDA**		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.05000	**Statistics by Class**	Sensitivity	0.8200	0.540	0.375	0.8200	0.555	0.4150
		Specificity	0.8475	0.745	0.775	0.8525	0.745	0.7975
	Accuracy		0.5579			0.5742		
	AUC-ROC		0.7644			0.7735		
			LDA			**QDA**		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.08500	**Statistics by Class**	Sensitivity	0.8100	0.475	0.41	0.8000	0.4500	0.570
		Specificity	0.8325	0.745	0.77	0.8525	0.8125	0.745
	Accuracy		0.5513			0.5679		
	AUC-ROC		0.7686			0.7800		
			LDA			**QDA**		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.120	**Statistics by Class**	Sensitivity	0.77	0.370	0.5150	0.755	0.530	0.4200
		Specificity	0.81	0.795	0.7225	0.825	0.715	0.8125
	Accuracy		0.5579			0.5671		
	AUC-ROC		0.7662			0.7747		
			LDA			**QDA**		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III

0.155	**Statistics by Class**	Sensitivity	0.8200	0.4600	0.42	0.8100	0.4450	0.5400
		Specificity	0.8275	0.7625	0.76	0.8425	0.8325	0.7225
	Accuracy		0.5492			0.5671		
	AUC-ROC		0.7623			0.7714		
			LDA			**QDA**		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.190	**Statistics by Class**	Sensitivity	0.8400	0.56	0.3400	0.8250	0.470	0.5350
		Specificity	0.8425	0.73	0.7975	0.8525	0.835	0.7275
	Accuracy		0.5654			0.5829		
	AUC-ROC		0.7641			0.7784		
			LDA			**QDA**		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.225	**Statistics by Class**	Sensitivity	0.7900	0.3250	0.555	0.7800	0.550	0.4600
		Specificity	0.8125	0.7875	0.735	0.8275	0.745	0.8225
	Accuracy		0.5625			0.5900		
	AUC-ROC		0.7704			0.7823		
			LDA			**QDA**		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.260	**Statistics by Class**	Sensitivity	0.805	0.4650	0.3900	0.78	0.515	0.3800
		Specificity	0.800	0.7525	0.7775	0.81	0.725	0.8025
	Accuracy		0.5579			0.5796		
	AUC-ROC		0.7643			0.7742		
			LDA			**QDA**		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.295	**Statistics by Class**	Sensitivity	0.8000	0.5050	0.3800	0.795	0.4750	0.4500
		Specificity	0.8325	0.7325	0.7775	0.830	0.7575	0.7725
	Accuracy		0.5504			0.5821		
	AUC-ROC		0.7669			0.7796		
			LDA			**QDA**		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.330	**Statistics by Class**	Sensitivity	0.8050	0.42	0.43	0.79	0.62	0.3450
		Specificity	0.8075	0.78	0.74	0.81	0.69	0.8775
	Accuracy		0.5496			0.5725		
	AUC-ROC		0.7585			0.7710		
			LDA			**QDA**		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.36500	**Statistics by Class**	Sensitivity	0.7950	0.3200	0.575	0.785	0.3350	0.555
		Specificity	0.8125	0.7825	0.750	0.810	0.7675	0.760
	Accuracy		0.5613			0.5771		
	AUC-ROC		0.7672			0.7788		
			LDA			**QDA**		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.400	**Statistics by Class**	Sensitivity	0.7950	0.46	0.4350	0.7750	0.5000	0.4300
		Specificity	0.8175	0.77	0.7575	0.8125	0.7625	0.7775
	Accuracy		0.5488			0.5758		
	AUC-ROC		0.7583			0.7723		

Source: IDE, R-Version 4.4.1, R-studio

The result in Table 4.6 compares the performance multiple metrics statistics of Linear Discriminant Analysis (LDA) and Quadratic Discriminant Analysis (QDA) averaged over 4 samples with simulation size of 200 across various skewness levels. QDA tends to have higher accuracy values in all skewness levels than LDA except for skewness level 0.01250, whereas QDA tends to have higher AUC-ROC values than LDA across all the skewness levels. QDA's average accuracy (0.576) is higher than LDA's (0.558), whereas QDA's average AUC-ROC (0.776) is higher than LDA's (0.765). QDA tends to have higher sensitivity for Pop I and Pop

III, whereas QDA tends to have higher specificity for Pop I and Pop II. QDA tends to outperform LDA across various skewness levels, especially in terms of accuracy and AUC-ROC. QDA's robustness to skewness makes it a better choice for classification tasks with skewed data.

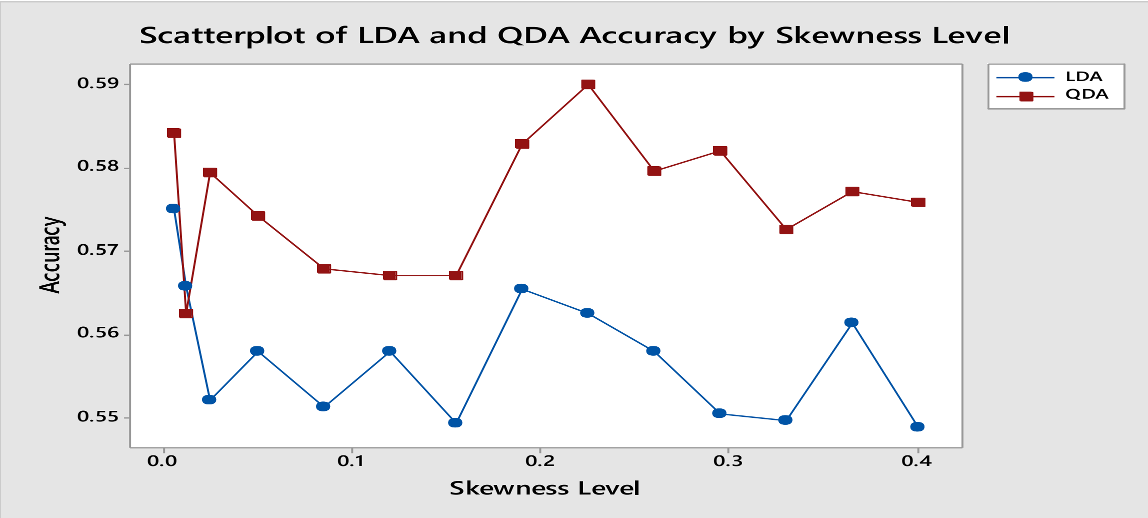


Fig. 4.1: Graph Displaying LDA and QDA Accuracy by Skewness Level Averaged over 4 Samples with Simulation Size of 200 for all Known Parameters

Figure 4.1 compares the accuracy of Linear Discriminant Analysis (LDA) and Quadratic Discriminant Analysis (QDA) across various skewness levels averaged over 4 samples with simulation size of 200 for all known parameters. QDA outperforms LDA, as QDA achieves higher accuracy than LDA across most skewness levels. Hence, QDA is more suitable for classification tasks with skewed data.

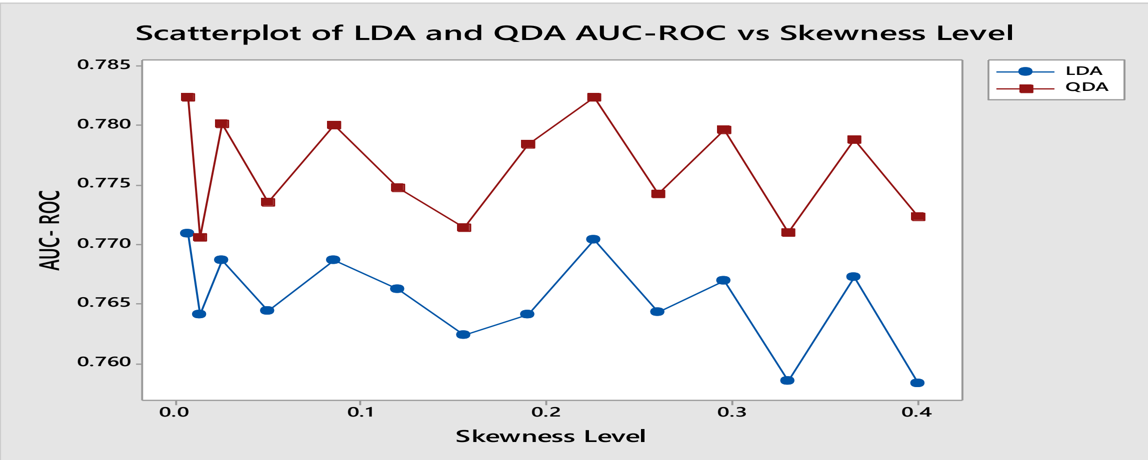


Fig. 4.2: Graph Displaying LDA and QDA AUC-ROC by Skewness Level Averaged over 4 Samples with Simulation Size of 200 for all Known Parameters

Figure 4.2 compares the Area under the Receiver Operating Characteristic Curve (AUC-ROC) of LDA and QDA across various skewness levels averaged over 4 samples with simulation size of 200 for all known parameters. QDA outperforms LDA, as QDA achieves higher AUC-ROC than LDA across most skewness levels. Hence, QDA is more suitable for classification tasks with skewed data.

Table 4.7: Summary of Multiple Metrics Statistics between LDA and QDA from ESD Averaged over 8 Samples with Simulation Size of 200

Skew			LDA			QDA		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.00625	**Statistics by Class**	Sensitivity	0.7800	0.445	0.460	0.780	0.45	0.46
		Specificity	0.8025	0.775	0.765	0.815	0.78	0.75
	Accuracy		0.5681			0.5725		
	AUC-ROC		0.7675			0.7768		
			LDA			**QDA**		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.01250	**Statistics by Class**	Sensitivity	0.7650	0.3250	0.4950	0.76	0.3200	0.6400
		Specificity	0.7925	0.7725	0.7275	0.81	0.8725	0.6775
	Accuracy		0.5548			0.5750		
	AUC-ROC		0.7661			0.7768		
			LDA			**QDA**		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.00250	**Statistics by Class**	Sensitivity	0.775	0.4500	0.430	0.7650	0.4200	0.5100
		Specificity	0.785	0.7575	0.785	0.7975	0.7925	0.7575
	Accuracy		0.5590			0.5763		
	AUC-ROC		0.7726			0.7829		
			LDA			**QDA**		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.05000	**Statistics by Class**	Sensitivity	0.760	0.56	0.3350	0.735	0.555	0.390
		Specificity	0.805	0.74	0.7825	0.820	0.735	0.785
	Accuracy		0.5558			0.5735		
	AUC-ROC		0.7626			0.7717		
			LDA			**QDA**		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.08500	**Statistics by Class**	Sensitivity	0.7900	0.3250	0.555	0.7800	0.550	0.4600
		Specificity	0.8125	0.7875	0.735	0.8275	0.745	0.8225
	Accuracy		0.5675			0.5927		
	AUC-ROC		0.7710			0.7852		
			LDA			**QDA**		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.12000	**Statistics by Class**	Sensitivity	0.805	0.4900	0.33	0.805	0.6300	0.3550
		Specificity	0.850	0.7125	0.75	0.850	0.7025	0.8425
	Accuracy		0.5556			0.5760		
	AUC-ROC		0.7638			0.7754		
			LDA			**QDA**		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.15500	**Statistics by Class**	Sensitivity	0.770	0.425	0.4200	0.745	0.4300	0.4950
		Specificity	0.805	0.745	0.7575	0.815	0.7875	0.7325
	Accuracy		0.5540			0.5727		

	AUC-ROC		0.7619			0.7733		
			LDA			**QDA**		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.19000	**Statistics by Class**	Sensitivity	0.8000	0.5100	0.2950	0.7900	0.435	0.5250
		Specificity	0.8025	0.7275	0.7725	0.8125	0.830	0.7325
	Accuracy		0.5573			0.5785		
	AUC-ROC		0.7642			0.7775		
			LDA			**QDA**		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.22500	**Statistics by Class**	Sensitivity	0.8050	0.45	0.4150	0.8050	0.390	0.5450
		Specificity	0.8225	0.76	0.7525	0.8375	0.815	0.7175
	Accuracy		0.5544			0.5710		
	AUC-ROC		0.7638			0.7751		
			LDA			**QDA**		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.26000	**Statistics by Class**	Sensitivity	0.825	0.450	0.4000	0.8200	0.56	0.295
		Specificity	0.805	0.765	0.7675	0.8025	0.69	0.845
	Accuracy		0.5483			0.5652		
	AUC-ROC		0.7623			0.7704		
			LDA			**QDA**		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.29500	**Statistics by Class**	Sensitivity	0.7950	0.4450	0.40	0.7750	0.465	0.465
		Specificity	0.8275	0.7525	0.74	0.8225	0.780	0.750
	Accuracy		0.5573			0.5750		
	AUC-ROC		0.7656			0.7759		
			LDA			**QDA**		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.33000	**Statistics by Class**	Sensitivity	0.765	0.4750	0.445	0.765	0.475	0.4850
		Specificity	0.830	0.7625	0.750	0.835	0.775	0.7525
	Accuracy		0.5525			0.5713		
	AUC-ROC		0.7656			0.7759		
			LDA			**QDA**		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.36500	**Statistics by Class**	Sensitivity	0.760	0.345	0.5500	0.755	0.550	0.455
		Specificity	0.825	0.770	0.7325	0.840	0.715	0.825
	Accuracy		0.5569			0.5825		
	AUC-ROC		0.7659			0.7792		
			LDA			**QDA**		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.40000	**Statistics by Class**	Sensitivity	0.7850	0.420	0.38	0.795	0.340	0.5000
		Specificity	0.8075	0.725	0.76	0.800	0.795	0.7225
	Accuracy		0.5667			0.5860		
	AUC-ROC		0.7715			0.7790		

Source: IDE, R-Version 4.4.1, R-studio

The result in Table 4.7 compares the performance multiple metrics statistics of Linear Discriminant Analysis (LDA) and Quadratic Discriminant Analysis (QDA) averaged over 8 samples with simulation size of 200 across various skewness levels. QDA tends to have higher accuracy and AUC-ROC values than LDA across all the skewness levels. QDA's average accuracy (0.576) is higher than LDA's (0.558), whereas QDA's average AUC-ROC (0.777) is higher than LDA's (0.766). QDA tends to have higher sensitivity for Pop I and Pop III, whereas QDA tends to have higher specificity for Pop I and Pop II. QDA tends to outperform LDA across various skewness levels, especially in terms of accuracy and AUC-ROC. QDA's robustness to skewness makes it a better choice for classification tasks with skewed data.

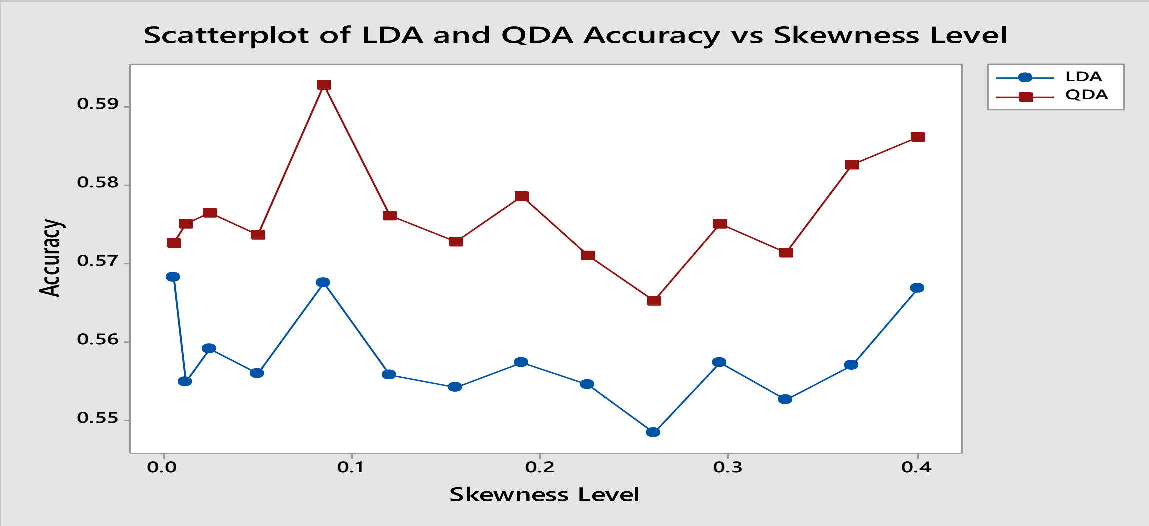


Fig. 4.3: Graph Displaying LDA and QDA Accuracy by Skewness Level Averaged over 8 Samples with Simulation Size of 200 for all Known Parameters

Figure 4.3 compares the accuracy of Linear Discriminant Analysis (LDA) and Quadratic Discriminant Analysis (QDA) across various skewness levels averaged over 8 samples with simulation size of 200 for all known parameters. QDA outperforms LDA, as QDA achieves higher accuracy than LDA across all skewness levels. Hence, QDA is more suitable for classification tasks with skewed data.

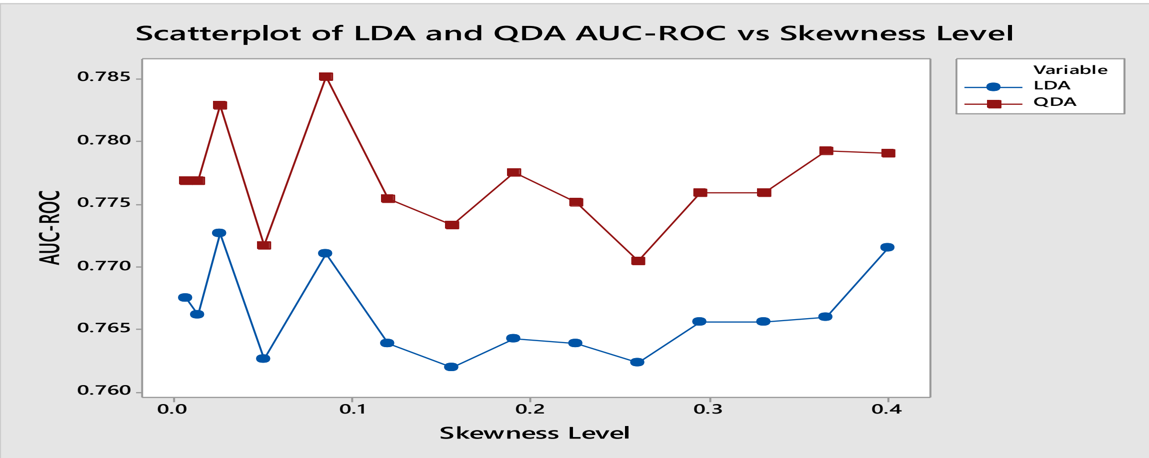


Fig. 4.4: Graph Displaying LDA and QDA AUC-ROC by Skewness Level Averaged over 8 Samples with Simulation Size of 200 for all Known Parameters

Figure 4.4 compares the Area under the Receiver Operating Characteristic Curve (AUC-ROC) of LDA and QDA across various skewness levels averaged over 8 samples with simulation size of 200 for all known parameters. QDA outperforms LDA, as QDA achieves higher AUC-ROC than LDA across all skewness levels. Hence, QDA is more suitable for classification tasks with skewed data.

Table 4.8: Summary of Multiple Metrics Statistics between LDA and QDA from ESD Averaged over 12 Samples with Simulation Size of 200

Skew			**LDA**			**QDA**		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.00625	**Statistics by Class**	Sensitivity	0.83	0.360	0.4850	0.815	0.540	0.485
		Specificity	0.84	0.775	0.7225	0.845	0.767	0.807
	Accuracy		0.5640			0.5729		
	AUC-ROC		0.7648			0.7740		
			LDA			**QDA**		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.01250	**Statistics by Class**	Sensitivity	0.825	0.455	0.3650	0.825	0.360	0.465
		Specificity	0.825	0.730	0.7675	0.830	0.785	0.710
	Accuracy		0.5558			0.5740		
	AUC-ROC		0.7687			0.7790		
			LDA			**QDA**		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.00250	**Statistics by Class**	Sensitivity	0.820	0.4200	0.5200	0.8200	0.4100	0.60
		Specificity	0.825	0.7925	0.7625	0.8325	0.8425	0.74
	Accuracy		0.5561			0.5729		
	AUC-ROC		0.7618			0.7724		
			LDA			**QDA**		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.05000	**Statistics by Class**	Sensitivity	0.795	0.4100	0.4750	0.8000	0.4700	0.440
		Specificity	0.795	0.7625	0.7825	0.8025	0.7375	0.815
	Accuracy		0.5619			0.5850		
	AUC-ROC		0.7678			0.7803		
			LDA			**QDA**		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.08500	**Statistics by Class**	Sensitivity	0.7950	0.3200	0.575	0.785	0.3350	0.555
		Specificity	0.8125	0.7825	0.750	0.810	0.7675	0.760
	Accuracy		0.5522			0.5782		
	AUC-ROC		0.7631			0.7759		
			LDA			**QDA**		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III

0.12000	**Statistics by Class**	Sensitivity	0.790	0.4550	0.410	0.7900	0.5200	0.4250
		Specificity	0.825	0.7475	0.755	0.8125	0.7475	0.8075
	Accuracy		0.5506			0.5757		
	AUC-ROC		0.7631			0.7759		
			LDA			**QDA**		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.15500	**Statistics by Class**	Sensitivity	0.775	0.425	0.4750	0.7700	0.540	0.3900
		Specificity	0.830	0.760	0.7475	0.8325	0.695	0.8225
	Accuracy		0.5518			0.5647		
	AUC-ROC		0.7623			0.7708		
			LDA			**QDA**		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.19000	**Statistics by Class**	Sensitivity	0.820	0.320	0.5450	0.8150	0.455	0.455
		Specificity	0.805	0.805	0.7325	0.8225	0.740	0.800
	Accuracy		0.5532			0.5731		
	AUC-ROC		0.7652			0.7757		
			LDA			**QDA**		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.22500	**Statistics by Class**	Sensitivity	0.760	0.345	0.5500	0.755	0.550	0.455
		Specificity	0.825	0.770	0.7325	0.840	0.715	0.825
	Accuracy		0.5569			0.5796		
	AUC-ROC		0.7672			0.7797		
			LDA			**QDA**		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.26000	**Statistics by Class**	Sensitivity	0.7900	0.475	0.44	0.7850	0.470	0.52
		Specificity	0.8075	0.785	0.76	0.8125	0.815	0.76
	Accuracy		0.5606			0.5778		
	AUC-ROC		0.7692			0.7766		
			LDA			**QDA**		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.29500	**Statistics by Class**	Sensitivity	0.790	0.3700	0.495	0.78	0.440	0.480
		Specificity	0.795	0.7875	0.745	0.81	0.785	0.755
	Accuracy		0.5615			0.5694		
	AUC-ROC		0.7665			0.7743		
			LDA			**QDA**		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.33000	**Statistics by Class**	Sensitivity	0.790	0.4450	0.470	0.7750	0.445	0.5000
		Specificity	0.805	0.7725	0.775	0.8175	0.790	0.7525
	Accuracy		0.5554			0.5746		
	AUC-ROC		0.7657			0.7758		
			LDA			**QDA**		

			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.36500	**Statistics by Class**	Sensitivity	0.8200	0.2650	0.5100	0.820	0.190	0.6650
		Specificity	0.8125	0.7625	0.7225	0.795	0.885	0.6575
	Accuracy		0.5610			0.5763		
	AUC-ROC		0.7679			0.7777		
			LDA			**QDA**		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.40000	**Statistics by Class**	Sensitivity	0.775	0.4300	0.44	0.7800	0.535	0.42
		Specificity	0.815	0.7575	0.75	0.8275	0.730	0.81
	Accuracy		0.5633			0.5788		
	AUC-ROC		0.7676			0.7781		

Source: IDE, R-Version 4.4.1, R-studio

The result in Table 4.8 compares the performance multiple metrics statistics of Linear Discriminant Analysis (LDA) and Quadratic Discriminant Analysis (QDA) averaged over 12 samples with simulation size of 200 across various skewness levels. QDA tends to have higher accuracy and AUC-ROC values than LDA across all the skewness levels. QDA's average accuracy (0.575) is higher than LDA's (0.559), whereas QDA's average AUC-ROC (0.776) is higher than LDA's (0.767). QDA tends to have higher sensitivity for Pop I and Pop III, whereas QDA tends to have higher specificity for Pop I and Pop II. QDA tends to outperform LDA across various skewness levels, especially in terms of accuracy and AUC-ROC. QDA's robustness to skewness makes it a better choice for classification tasks with skewed data.

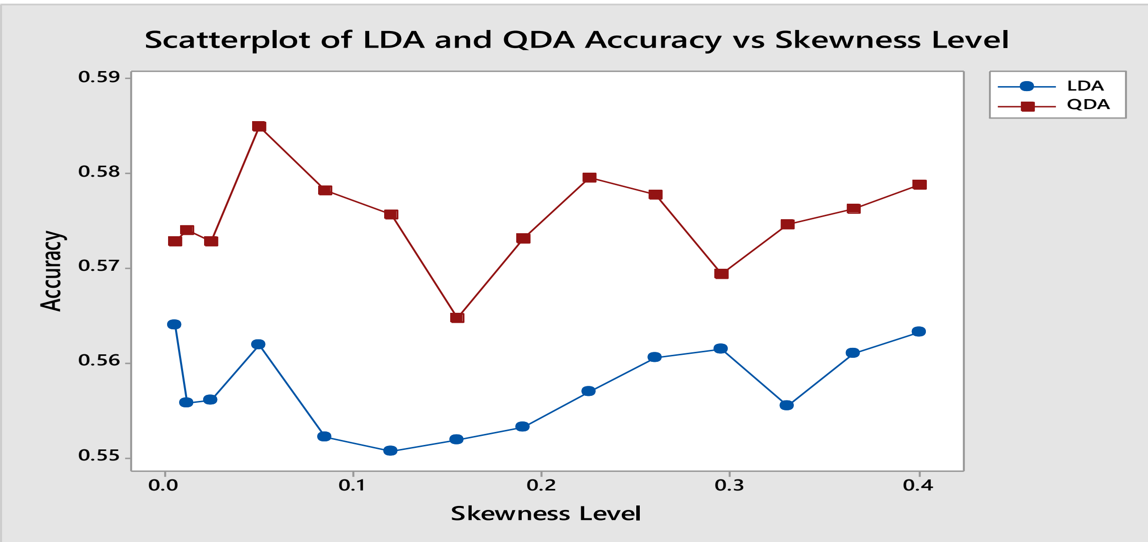


Fig. 4.5: Graph Displaying LDA and QDA Accuracy by Skewness Level Averaged over 12 Samples with Simulation Size of 200 for all Known Parameters

Figure 4.5 compares the accuracy of Linear Discriminant Analysis (LDA) and Quadratic Discriminant Analysis (QDA) across various skewness levels averaged over 12 samples with simulation size of 200 for all known parameters. QDA outperforms LDA, as QDA achieves higher accuracy than LDA across all skewness levels. Hence, QDA is more suitable for classification tasks with skewed data.

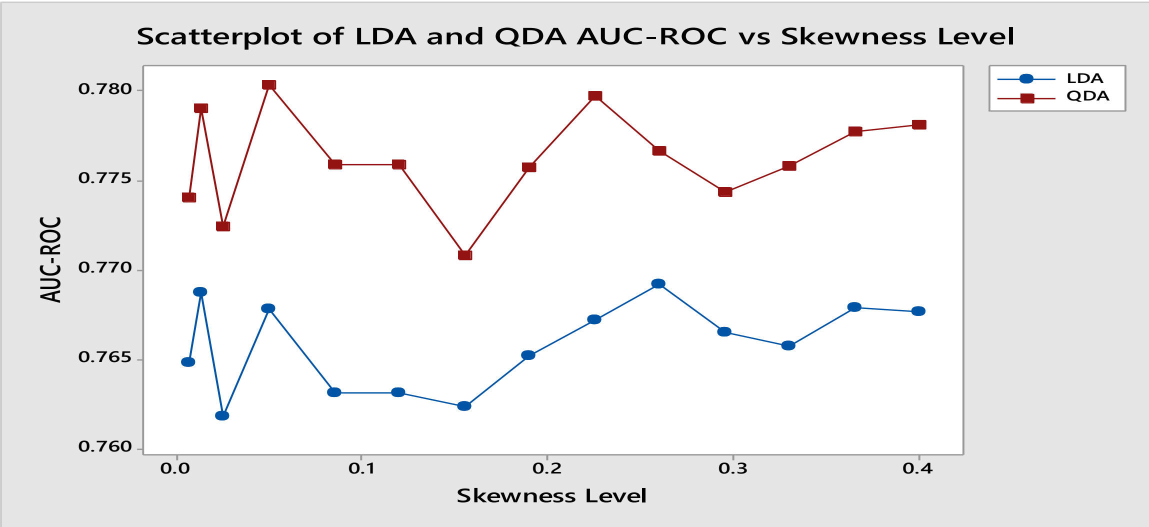


Fig. 4.6: Graph Displaying LDA and QDA AUC-ROC by Skewness Level Averaged over 12 Samples with Simulation Size of 200 for all Known Parameters

Figure 4.6 compares the Area under the Receiver Operating Characteristic Curve (AUC-ROC) of LDA and QDA across various skewness levels averaged over 12 samples with simulation size of 200 for all known parameters. QDA outperforms LDA, as QDA achieves higher AUC-ROC than LDA across all skewness levels. Hence, QDA is more suitable for classification tasks with skewed data.

Table 4.9: Summary of Multiple Metrics Statistics between LDA and QDA from ESD Averaged over 16 Samples with Simulation Size of 200

Skew			LDA			QDA		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.00625	Statistics by Class	Sensitivity	0.800	0.470	0.4550	0.8000	0.460	0.4650
		Specificity	0.835	0.765	0.7625	0.8325	0.752	0.7775
	Accuracy		0.5606			0.5730		
	AUC-ROC		0.7659			0.7760		
			LDA			QDA		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.01250	Statistics by Class	Sensitivity	0.805	0.505	0.3250	0.8050	0.375	0.425
		Specificity	0.790	0.730	0.7975	0.8025	0.775	0.725
	Accuracy		0.5554			0.5752		
	AUC-ROC		0.7670			0.7772		
			LDA			QDA		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.00250	Statistics by Class	Sensitivity	0.795	0.4400	0.48	0.8000	0.5050	0.425
		Specificity	0.820	0.7675	0.77	0.8425	0.7225	0.800
	Accuracy		0.5625			0.5858		
	AUC-ROC		0.7665			0.7795		
			LDA			QDA		

			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.05000	Statistics by Class	Sensitivity	0.8200	0.420	0.395	0.800	0.3500	0.6200
		Specificity	0.7875	0.765	0.765	0.805	0.8725	0.7075
	Accuracy		0.5524			0.5739		
	AUC-ROC		0.7619			0.7740		
			LDA			QDA		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.08500	Statistics by Class	Sensitivity	0.7950	0.490	0.3950	0.79	0.565	0.365
		Specificity	0.8175	0.755	0.7675	0.82	0.730	0.810
	Accuracy		0.5545			0.5743		
	AUC-ROC		0.7646			0.7759		
			LDA			QDA		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.12000	Statistics by Class	Sensitivity	0.795	0.400	0.50	0.80	0.485	0.43
		Specificity	0.840	0.777	0.73	0.83	0.737	0.79
	Accuracy		0.5502			0.5685		
	AUC-ROC		0.7632			0.7731		
			LDA			QDA		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.15500	Statistics by Class	Sensitivity	0.7600	0.445	0.4050	0.7550	0.485	0.410
		Specificity	0.8275	0.735	0.7425	0.8275	0.732	0.765
	Accuracy		0.5563			0.5795		
	AUC-ROC		0.7669			0.7788		
			LDA			QDA		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.19000	Statistics by Class	Sensitivity	0.860	0.505	0.3550	0.8600	0.440	0.45
		Specificity	0.825	0.747	0.7875	0.8375	0.787	0.75
	Accuracy		0.5655			0.5784		
	AUC-ROC		0.7701			0.7777		
			LDA			QDA		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.22500	Statistics by Class	Sensitivity	0.79	0.425	0.4050	0.79	0.515	0.3500
		Specificity	0.79	0.752	0.7675	0.80	0.715	0.8125
	Accuracy		0.5539			0.5716		
	AUC-ROC		0.7620			0.7723		
			LDA			QDA		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.26000	Statistics by Class	Sensitivity	0.8050	0.490	0.405	0.80	0.530	0.455
		Specificity	0.8375	0.747	0.765	0.84	0.767	0.785
	Accuracy		0.5602			0.5749		
	AUC-ROC		0.7663			0.7753		
			LDA			QDA		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III

0.29500	Statistics by Class	Sensitivity	0.8600	0.555	0.30	0.8500	0.385	0.4600
		Specificity	0.7975	0.730	0.83	0.8075	0.807	0.7325
	Accuracy		0.5601			0.5782		
	AUC-ROC		0.7672			0.7776		
			LDA			QDA		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.33000	Statistics by Class	Sensitivity	0.8650	0.540	0.3850	0.8550	0.48	0.4750
		Specificity	0.8275	0.755	0.8125	0.8325	0.80	0.7725
	Accuracy		0.5577			0.5798		
	AUC-ROC		0.7670			0.7790		
			LDA			QDA		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.36500	Statistics by Class	Sensitivity	0.7900	0.37	0.5100	0.7800	0.430	0.5050
		Specificity	0.8175	0.77	0.7475	0.8175	0.757	0.7825
	Accuracy		0.5592			0.5751		
	AUC-ROC		0.7679			0.7788		
			LDA			QDA		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.40000	Statistics by Class	Sensitivity	0.78	0.490	0.395	0.7900	0.490	0.4450
		Specificity	0.82	0.752	0.760	0.8175	0.772	0.7725
	Accuracy		0.5656			0.5768		
	AUC-ROC		0.7696			0.7800		

Source IDE, R-Version 4.4.1, R-studio

The result in Table 4.9 compares the performance multiple metrics statistics of Linear Discriminant Analysis (LDA) and Quadratic Discriminant Analysis (QDA) averaged over 16 samples with simulation size of 200 across various skewness levels. QDA tends to have higher accuracy and AUC-ROC values than LDA across all the skewness levels. QDA's average accuracy (0.576) is higher than LDA's (0.559), whereas QDA's average AUC-ROC (0.778) is higher than LDA's (0.766). QDA tends to have higher sensitivity for Pop I and Pop III, whereas QDA tends to have higher specificity for Pop I and Pop II. QDA tends to outperform LDA across various skewness levels, especially in terms of accuracy and AUC-ROC. QDA's robustness to skewness makes it a better choice for classification tasks with skewed data.

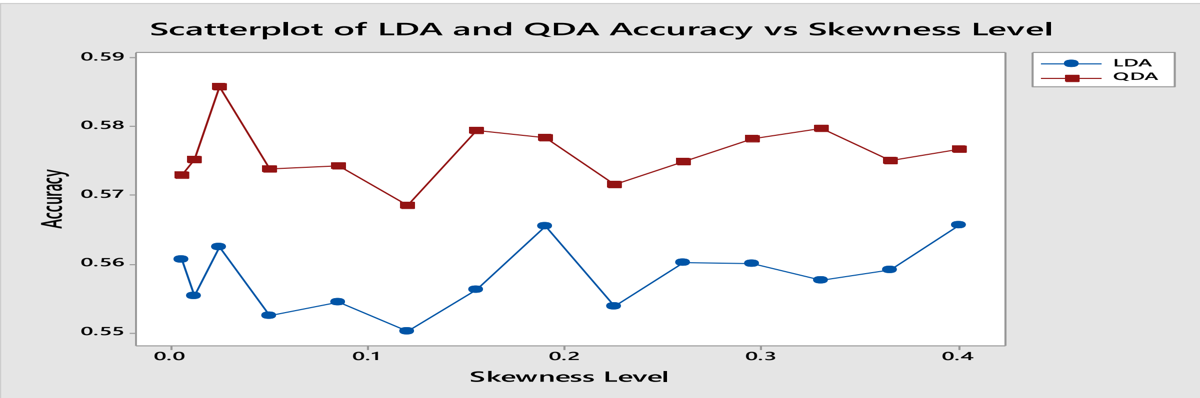


Fig. 4.7: Graph Displaying LDA and QDA Accuracy by Skewness Level averaged over 16 Samples with Simulation Size of 200 for all Known Parameters

Figure 4.7 compares the accuracy of Linear Discriminant Analysis (LDA) and Quadratic Discriminant Analysis (QDA) across various skewness levels averaged over 16 samples with simulation size of 200 for all known parameters. QDA outperforms LDA, as QDA achieves higher accuracy than LDA across all skewness levels. Hence, QDA is more suitable for classification tasks with skewed data.

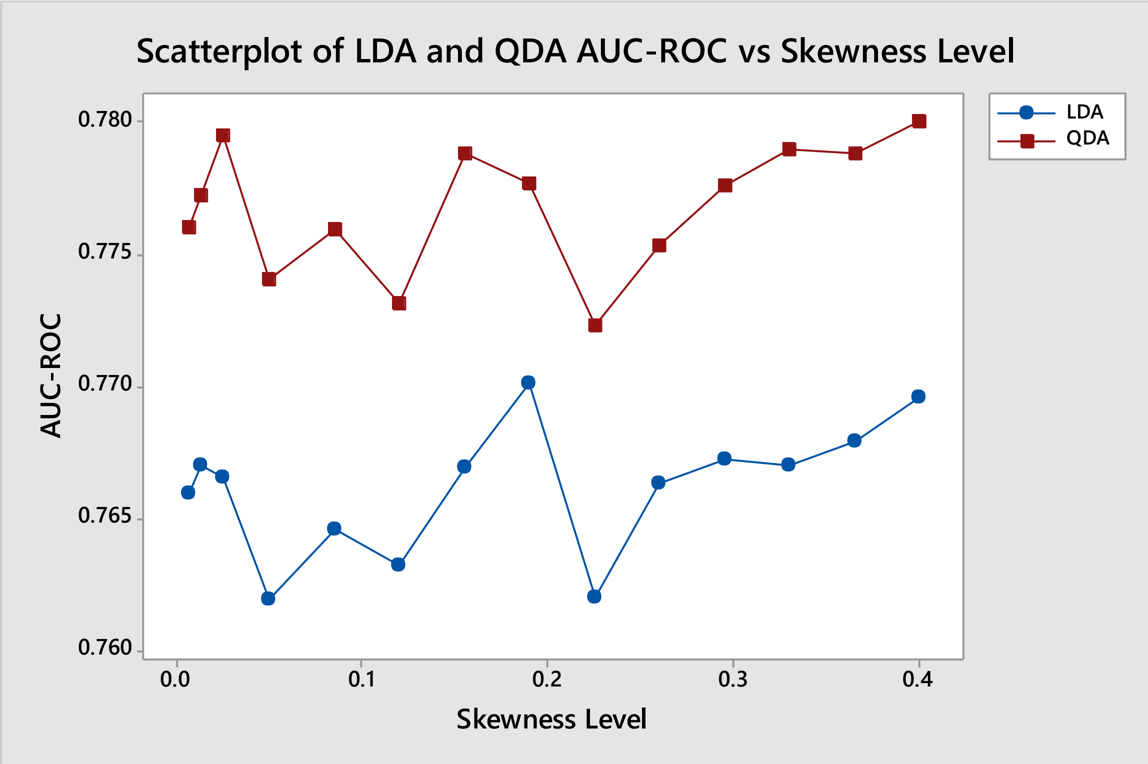


Fig. 4.8: Graph Displaying LDA and QDA AUC-ROC by Skewness Level Averaged over 16 Samples with Simulation Size of 200 for all Known Parameters

Figure 4.8 compares the Area under the Receiver Operating Characteristic Curve (AUC-ROC) of LDA and QDA across various skewness levels averaged over 16 samples with simulation size of 200 for all known parameters. QDA outperforms LDA, as QDA achieves higher AUC-ROC than LDA across all skewness levels. Hence, QDA is more suitable for classification tasks with skewed data.

Table 4.10: Summary of Multiple Metrics Statistics between LDA and QDA from ESD Averaged over 20 Samples with Simulation Size of 200

Skew			LDA			QDA		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.00625	**Statistics by Class**	Sensitivity	0.815	0.3650	0.4700	0.8050	0.5250	0.3750
		Specificity	0.795	0.7775	0.7525	0.8025	0.7275	0.8225
	Accuracy		0.5613			0.5749		
	AUC-ROC		0.7665			0.7762		
			LDA			**QDA**		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.01250	**Statistics by Class**	Sensitivity	0.8300	0.435	0.4450	0.8050	0.5650	0.4350
		Specificity	0.8175	0.775	0.7625	0.8325	0.7475	0.8225
	Accuracy		0.5563			0.5755		
	AUC-ROC		0.7659			0.7769		
			LDA			**QDA**		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III

0.00250	**Statistics by Class**	Sensitivity	0.770	0.425	0.4200	0.745	0.4300	0.4950
		Specificity	0.805	0.745	0.7575	0.815	0.7875	0.7325
	Accuracy		0.5568			0.5793		
	AUC-ROC		0.7645			0.7763		
			LDA			**QDA**		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.05000	**Statistics by Class**	Sensitivity	0.820	0.4200	0.4900	0.8300	0.2800	0.640
		Specificity	0.835	0.7625	0.7675	0.8425	0.8675	0.665
	Accuracy		0.5539			0.5743		
	AUC-ROC		0.7632			0.7756		
			LDA			**QDA**		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.08500	**Statistics by Class**	Sensitivity	0.805	0.395	0.4550	0.810	0.430	0.48
		Specificity	0.825	0.725	0.7775	0.835	0.735	0.79
	Accuracy		0.5517			0.5694		
	AUC-ROC		0.7645			0.7740		
			LDA			**QDA**		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.12000	**Statistics by Class**	Sensitivity	0.7850	0.420	0.38	0.795	0.340	0.5000
		Specificity	0.8075	0.725	0.76	0.800	0.795	0.7225
	Accuracy		0.5608			0.5828		
	AUC-ROC		0.7683			0.7789		
			LDA			**QDA**		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.15500	**Statistics by Class**	Sensitivity	0.775	0.4850	0.415	0.7750	0.5550	0.370
		Specificity	0.810	0.7525	0.775	0.8075	0.7175	0.825
	Accuracy		0.5588			0.5694		
	AUC-ROC		0.7658			0.7736		
			LDA			**QDA**		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.19000	**Statistics by Class**	Sensitivity	0.785	0.370	0.5200	0.785	0.3800	0.5450
		Specificity	0.820	0.765	0.7525	0.825	0.7825	0.7475
	Accuracy		0.5599			0.5775		
	AUC-ROC		0.7663			0.7769		
			LDA			**QDA**		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.22500	**Statistics by Class**	Sensitivity	0.81	0.39	0.44	0.8100	0.495	0.36
		Specificity	0.82	0.76	0.74	0.8325	0.700	0.80
	Accuracy		0.5591			0.5773		
	AUC-ROC		0.7670			0.7770		
			LDA			**QDA**		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.26000	**Statistics by Class**	Sensitivity	0.8050	0.4700	0.415	0.785	0.4400	0.470
		Specificity	0.8375	0.7475	0.760	0.840	0.7725	0.735
	Accuracy		0.5618			0.5834		
	AUC-ROC		0.7691			0.7808		
			LDA			**QDA**		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.29500	**Statistics by Class**	Sensitivity	0.8100	0.455	0.49	0.8100	0.4050	0.4800
		Specificity	0.8075	0.780	0.79	0.7975	0.7875	0.7625
	Accuracy		0.5597			0.5745		
	AUC-ROC		0.7670			0.7779		
			LDA			**QDA**		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.33000	**Statistics by Class**	Sensitivity	0.770	0.4650	0.4050	0.76	0.4650	0.5150
		Specificity	0.785	0.7475	0.7875	0.79	0.8025	0.7775
	Accuracy		0.5583			0.5718		

	AUC-ROC		0.7637			0.7741		
			LDA			**QDA**		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.36500	**Statistics by Class**	Sensitivity	0.800	0.4750	0.4450	0.78	0.4350	0.5150
		Specificity	0.835	0.7625	0.7625	0.84	0.7975	0.7275
	Accuracy		0.5566			0.5711		
	AUC-ROC		0.7661			0.7761		
			LDA			**QDA**		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.40000	**Statistics by Class**	Sensitivity	0.81	0.385	0.445	0.8050	0.5600	0.385
		Specificity	0.83	0.750	0.740	0.8375	0.7075	0.830
	Accuracy		0.5537			0.5728		
	AUC-ROC		0.7639			0.7735		

Source: IDE, R-Version 4.4.1, R-studio

The result in Table 4.10 compares the performance multiple metrics statistics of Linear Discriminant Analysis (LDA) and Quadratic Discriminant Analysis (QDA) averaged over 20 samples with simulation size of 200 across various skewness levels. QDA tends to have higher accuracy and AUC-ROC values than LDA across all the skewness levels. QDA's average accuracy (0.576) is higher than LDA's (0.559), whereas QDA's average AUC-ROC (0.777) is higher than LDA's (0.766). QDA tends to have higher sensitivity for Pop I and Pop III, whereas QDA tends to have higher specificity for Pop I and Pop II. QDA tends to outperform LDA across various skewness levels, especially in terms of accuracy and AUC-ROC. QDA's robustness to skewness makes it a better choice for classification tasks with skewed data.

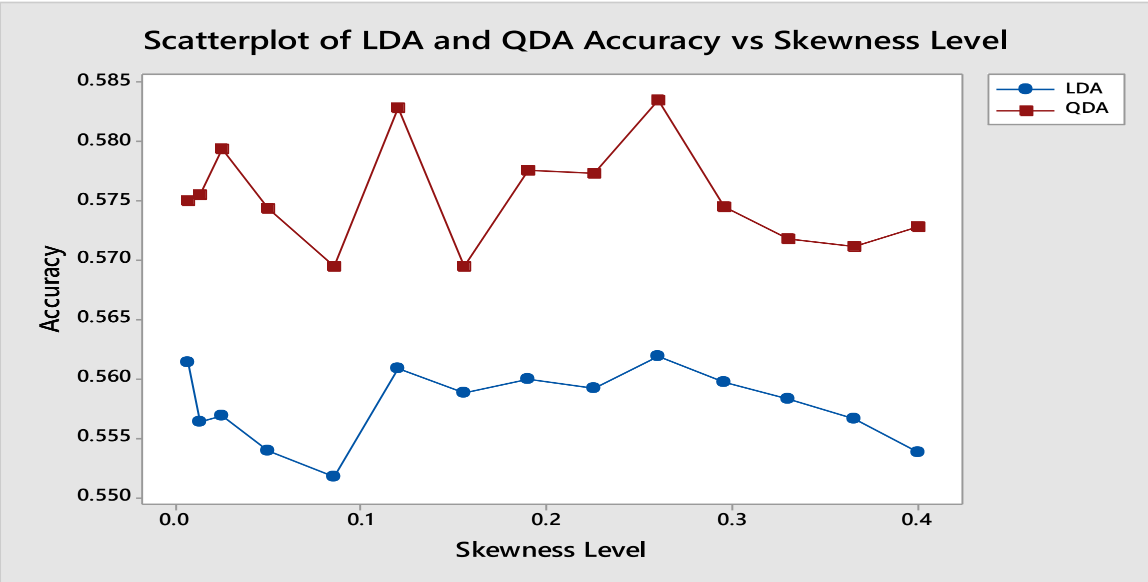


Fig. 4.9: Graph Displaying LDA and QDA Accuracy by Skewness Level Averaged over 20 Samples with Simulation Size of 200 for all Known Parameters

Figure 4.9 compares the accuracy of Linear Discriminant Analysis (LDA) and Quadratic Discriminant Analysis (QDA) across various skewness levels averaged over 20 samples with simulation size of 200 for all known parameters. QDA outperforms LDA, as QDA achieves higher accuracy than LDA across all skewness levels. Hence, QDA is more suitable for classification tasks with skewed data.

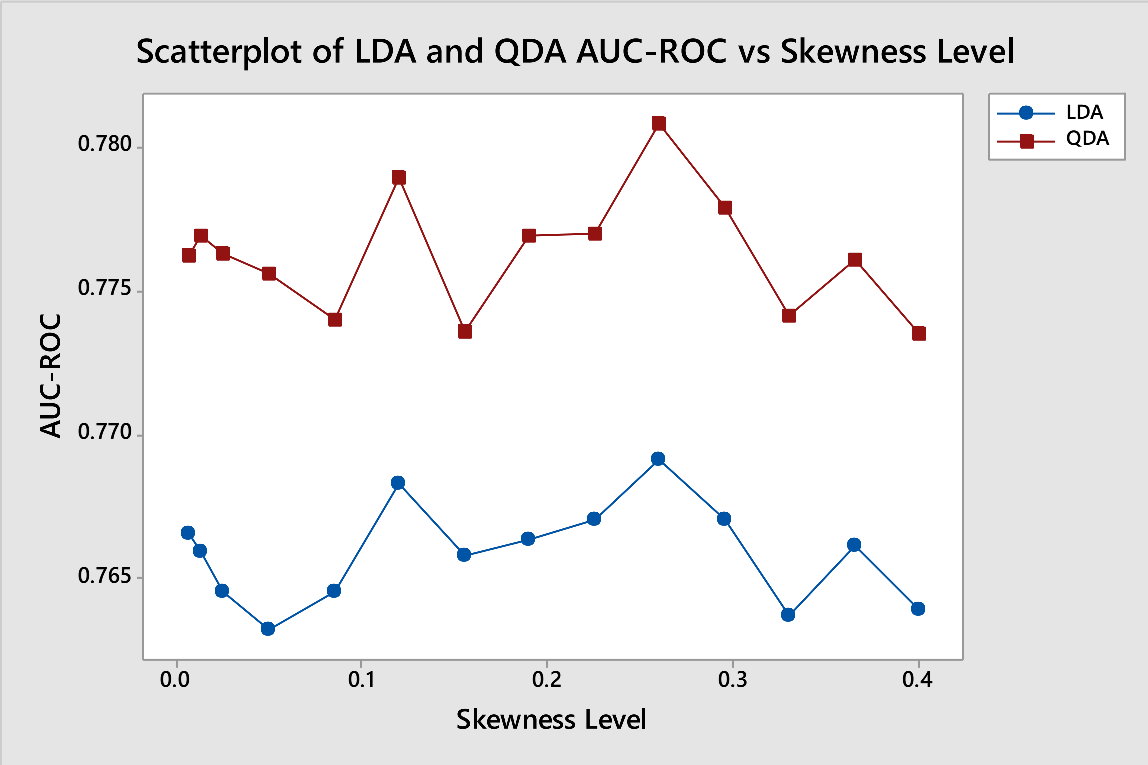


Fig. 4.10: Graph Displaying LDA and QDA AUC-ROC by Skewness Level Averaged over 20 Samples with Simulation Size of 200 for all Known Parameters

Figure 4.10 compares the Area Under the Receiver Operating Characteristic Curve (AUC-ROC) of LDA and QDA across various skewness levels averaged over 20 samples with simulation size of 200 for all known parameters. QDA outperforms LDA, as QDA achieves higher AUC-ROC than LDA across all skewness levels. Hence, QDA is more suitable for classification tasks with skewed data.

Table 4.11: Summary of Multiple Metrics Statistics between LDA and QDA from ESD Averaged over 24 Samples with Simulation Size of 200

Skew			**LDA**			**QDA**		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.00625	**Statistics by Class**	Sensitivity	0.8100	0.475	0.41	0.8000	0.4500	0.570
		Specificity	0.8325	0.745	0.77	0.8525	0.8125	0.745
	Accuracy		0.5597			0.5738		
	AUC-ROC		0.7668			0.7769		
			LDA			**QDA**		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.01250	**Statistics by Class**	Sensitivity	0.805	0.4650	0.3900	0.78	0.515	0.3800
		Specificity	0.800	0.7525	0.7775	0.81	0.725	0.8025
	Accuracy		0.5594			0.5800		
	AUC-ROC		0.7658			0.7771		
			LDA			**QDA**		

			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.00250	**Statistics by Class**	Sensitivity	0.795	0.5250	0.265	0.7900	0.5450	0.3700
		Specificity	0.815	0.7025	0.775	0.8325	0.7375	0.7825
	Accuracy		0.5533			0.5763		
	AUC-ROC		0.7628			0.7756		
			LDA			**QDA**		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.05000	**Statistics by Class**	Sensitivity	0.7800	0.45	0.3900	0.7700	0.5050	0.435
		Specificity	0.8075	0.75	0.7525	0.8225	0.7525	0.780
	Accuracy		0.5542			0.5698		
	AUC-ROC		0.7650			0.7745		
			LDA			**QDA**		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.08500	**Statistics by Class**	Sensitivity	0.85	0.4600	0.4600	0.8550	0.4250	0.57
		Specificity	0.85	0.7575	0.7775	0.8575	0.8275	0.74
	Accuracy		0.5580			0.5789		
	AUC-ROC		0.7673			0.7777		
			LDA			**QDA**		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.12000	**Statistics by Class**	Sensitivity	0.760	0.4150	0.41	0.7600	0.555	0.400
		Specificity	0.825	0.7375	0.73	0.8425	0.710	0.805
	Accuracy		0.5568			0.5697		
	AUC-ROC		0.7649			0.7736		
			LDA			**QDA**		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.15500	**Statistics by Class**	Sensitivity	0.7850	0.50	0.3350	0.7850	0.615	0.3100
		Specificity	0.8125	0.74	0.7575	0.8125	0.705	0.8375
	Accuracy		0.5637			0.5785		
	AUC-ROC		0.7684			0.7779		
			LDA			**QDA**		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.19000	**Statistics by Class**	Sensitivity	0.8050	0.4000	0.425	0.8000	0.575	0.3850
		Specificity	0.8175	0.7525	0.745	0.8425	0.700	0.8375
	Accuracy		0.5590			0.5769		
	AUC-ROC		0.7661			0.7768		
			LDA			**QDA**		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.22500	**Statistics by Class**	Sensitivity	0.7450	0.40	0.530	0.7350	0.4800	0.485
		Specificity	0.8225	0.78	0.735	0.8475	0.7375	0.765
	Accuracy		0.5612			0.5781		
	AUC-ROC		0.7687			0.7796		
			LDA			**QDA**		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
		Sensitivity	0.7700	0.39	0.5200	0.74	0.5600	0.375

0.26000	**Statistics by Class**	Specificity	0.8025	0.80	0.7375	0.80	0.6975	0.840
	Accuracy		0.5580			0.5712		
	AUC-ROC		0.7647			0.7750		
			LDA			**QDA**		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.29500	**Statistics by Class**	Sensitivity	0.8050	0.505	0.285	0.8000	0.455	0.4850
		Specificity	0.7875	0.745	0.765	0.8125	0.830	0.7275
	Accuracy		0.5583			0.5730		
	AUC-ROC		0.7664			0.7767		
			LDA			**QDA**		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.33000	**Statistics by Class**	Sensitivity	0.7850	0.51	0.370	0.7600	0.4750	0.4700
		Specificity	0.8175	0.74	0.775	0.8425	0.7675	0.7425
	Accuracy		0.5535			0.5733		
	AUC-ROC		0.7644			0.7745		
			LDA			**QDA**		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.36500	**Statistics by Class**	Sensitivity	0.7950	0.470	0.420	0.775	0.505	0.4850
		Specificity	0.8425	0.745	0.755	0.865	0.765	0.7525
	Accuracy		0.5520			0.5746		
	AUC-ROC		0.7644			0.7763		
			LDA			**QDA**		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.40000	**Statistics by Class**	Sensitivity	0.8500	0.5150	0.335	0.835	0.4300	0.510
		Specificity	0.8125	0.7375	0.800	0.835	0.8275	0.725
	Accuracy		0.5608			0.5737		
	AUC-ROC		0.7673			0.7777		

Source: IDE, R-Version 4.4.1, R-studio

The result in Table 4.11 compares the performance multiple metrics statistics of Linear Discriminant Analysis (LDA) and Quadratic Discriminant Analysis (QDA) averaged over 24 samples with simulation size of 200 across various skewness levels. QDA tends to have higher accuracy and AUC-ROC values than LDA across all the skewness levels. QDA's average accuracy (0.575) is higher than LDA's (0.559), whereas QDA's average AUC-ROC (0.777) is higher than LDA's (0.766). QDA tends to have higher sensitivity for Pop I and Pop III, whereas QDA tends to have higher specificity for Pop I and Pop II. QDA tends to outperform LDA across various skewness levels, especially in terms of accuracy and AUC-ROC. QDA's robustness to skewness makes it a better choice for classification tasks with skewed data.

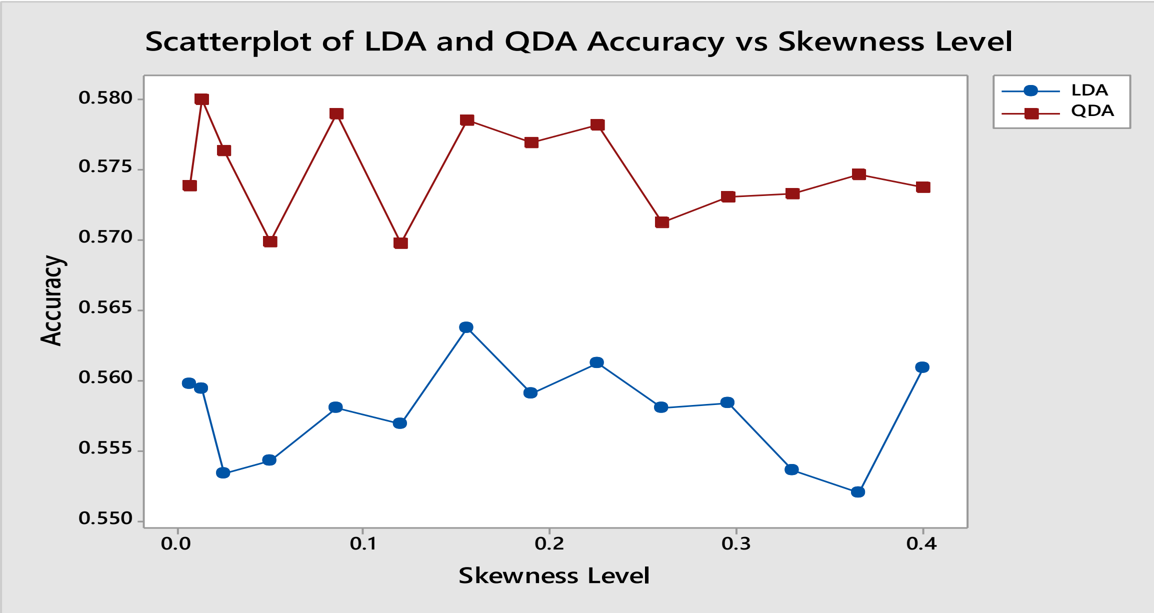


Fig. 4.11: Graph Displaying LDA and QDA Accuracy by Skewness Level Averaged over 24 Samples with Simulation Size of 200 for all Known Parameters

Figure 4.11 compares the accuracy of Linear Discriminant Analysis (LDA) and Quadratic Discriminant Analysis (QDA) across various skewness levels averaged over 24 samples with simulation size of 200 for all known parameters. QDA outperforms LDA, as QDA achieves higher accuracy than LDA across all skewness levels. Hence, QDA is more suitable for classification tasks with skewed data.

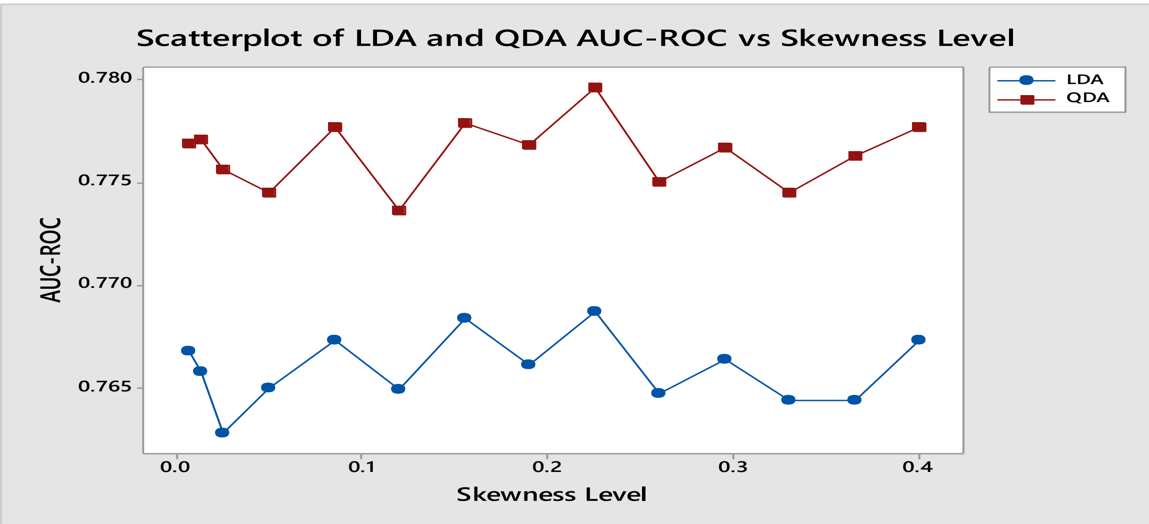


Fig. 4.12: Graph Displaying LDA and QDA AUC-ROC by Skewness Level Averaged over 24 Samples with Simulation Size of 200 for all Known Parameters

Figure 4.12 compares the Area under the Receiver Operating Characteristic Curve (AUC-ROC) of LDA and QDA across various skewness levels averaged over 24 samples with simulation size of 200 for all known parameters. QDA outperforms LDA, as QDA achieves higher AUC-ROC than LDA across all skewness levels. Hence, QDA is more suitable for classification tasks with skewed data.

Table 4.12: Summary of Multiple Metrics Statistics between LDA and QDA from

Skew			**LDA**			**QDA**		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.00625	**Statistics by Class**	Sensitivity	0.8050	0.4000	0.545	0.8000	0.3650	0.615
		Specificity	0.8025	0.8025	0.770	0.7975	0.8575	0.735
	Accuracy		0.5588			0.5733		
	AUC-ROC		0.7668			0.7768		
			LDA			**QDA**		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.01250	**Statistics by Class**	Sensitivity	0.740	0.3150	0.5150	0.735	0.3700	0.58
		Specificity	0.795	0.7675	0.7225	0.800	0.8225	0.72
	Accuracy		0.5588			0.5805		
	AUC-ROC		0.7653			0.7774		
			LDA			**QDA**		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.00250	**Statistics by Class**	Sensitivity	0.7750	0.300	0.4850	0.7800	0.265	0.54
		Specificity	0.8025	0.775	0.7025	0.8125	0.830	0.65
	Accuracy		0.5534			0.5743		
	AUC-ROC		0.7631			0.7746		
			LDA			**QDA**		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.05000	**Statistics by Class**	Sensitivity	0.7900	0.4850	0.425	0.800	0.455	0.515
		Specificity	0.8325	0.7575	0.760	0.835	0.810	0.740
	Accuracy		0.5550			0.5740		
	AUC-ROC		0.7661			0.7769		
			LDA			**QDA**		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.08500	**Statistics by Class**	Sensitivity	0.760	0.465	0.485	0.745	0.480	0.455
		Specificity	0.835	0.770	0.750	0.835	0.735	0.770
	Accuracy		0.5599			0.5743		
	AUC-ROC		0.7667			0.7749		
			LDA			**QDA**		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.12000	**Statistics by Class**	Sensitivity	0.815	0.465	0.4350	0.7950	0.4700	0.475
		Specificity	0.845	0.750	0.7625	0.8425	0.7625	0.765
	Accuracy		0.5602			0.5779		
	AUC-ROC		0.7672			0.7773		
			LDA			**QDA**		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.15500	**Statistics by Class**	Sensitivity	0.785	0.4300	0.4650	0.7700	0.5250	0.4300
		Specificity	0.805	0.7675	0.7675	0.8125	0.7175	0.8325
	Accuracy		0.5569			0.5773		
	AUC-ROC		0.7662			0.7773		
			LDA			**QDA**		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.19000 W2	**Statistics by Class**	Sensitivity	0.7550	0.405	0.4600	0.740	0.5200	0.405
		Specificity	0.7875	0.760	0.7625	0.795	0.7125	0.825
	Accuracy		0.5626			0.5765		
	AUC-ROC		0.7688			0.7792		
			LDA			**QDA**		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.22500	**Statistics by Class**	Sensitivity	0.770	0.41	0.520	0.765	0.495	0.4750
		Specificity	0.805	0.79	0.755	0.815	0.750	0.8025
	Accuracy		0.5575			0.5736		
	AUC-ROC		0.7651			0.7758		
			LDA			**QDA**		

			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.26000	**Statistics by Class**	Sensitivity	0.840	0.360	0.54	0.8350	0.3300	0.615
		Specificity	0.845	0.785	0.74	0.8425	0.8425	0.705
	Accuracy		0.5538			0.5701		
	AUC-ROC		0.7646			0.7745		
			LDA			**QDA**		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.29500	**Statistics by Class**	Sensitivity	0.755	0.3600	0.530	0.7700	0.330	0.585
		Specificity	0.810	0.7875	0.725	0.8325	0.805	0.705
	Accuracy		0.5532			0.5748		
	AUC-ROC		0.7643			0.7763		
			LDA			**QDA**		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.33000	**Statistics by Class**	Sensitivity	0.7850	0.4350	0.4150	0.785	0.41	0.4600
		Specificity	0.8175	0.7625	0.7375	0.815	0.78	0.7325
	Accuracy		0.5589			0.5749		
	AUC-ROC		0.7671			0.7779		
			LDA			**QDA**		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.36500	**Statistics by Class**	Sensitivity	0.825	0.3250	0.5300	0.83	0.34	0.5450
		Specificity	0.800	0.7925	0.7475	0.81	0.82	0.7275
	Accuracy		0.5574			0.5743		
	AUC-ROC		0.7677			0.7777		
			LDA			**QDA**		
			Pop I	Pop II	Pop III	Pop I	Pop II	Pop III
0.40000	**Statistics by Class**	Sensitivity	0.7700	0.3950	0.515	0.7650	0.4350	0.5250
		Specificity	0.8225	0.7475	0.770	0.8325	0.7725	0.7575
	Accuracy		0.5599			0.5764		
	AUC-ROC		0.7653			0.7758		

Source: IDE, R-Version 4.4.1, R-studio

The result in Table 4.12 compares the performance multiple metrics statistics of Linear Discriminant Analysis (LDA) and Quadratic Discriminant Analysis (QDA) averaged over 28 samples with simulation size of 200 across various skewness levels. QDA tends to have higher accuracy and AUC-ROC values than LDA across all the skewness levels. QDA's average accuracy (0.574) is higher than LDA's (0.559), whereas QDA's average AUC-ROC (0.777) is higher than LDA's (0.766). QDA tends to have higher sensitivity for Pop I and Pop III, whereas QDA tends to have higher specificity for Pop I and Pop II. QDA tends to outperform LDA across various skewness levels, especially in terms of accuracy and AUC-ROC. QDA's robustness to skewness makes it a better choice for classification tasks with skewed data.

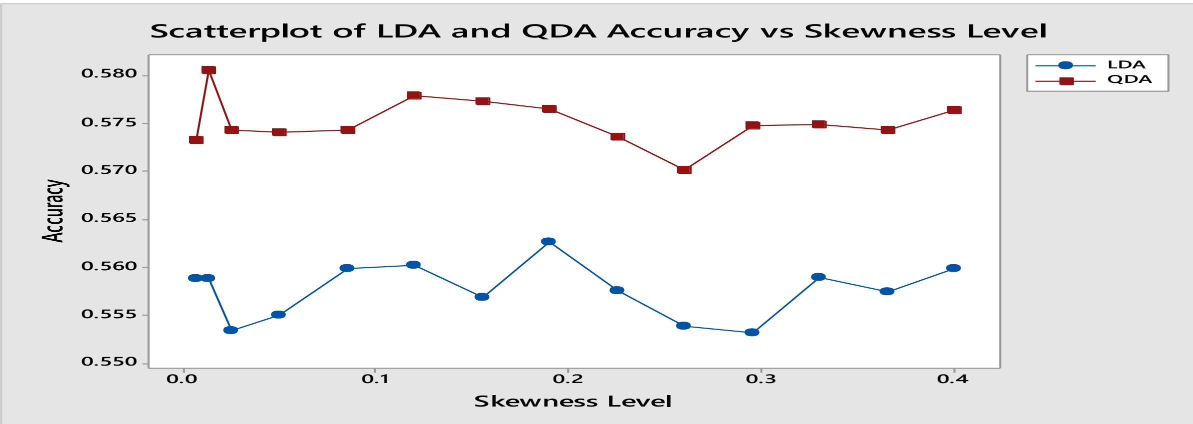
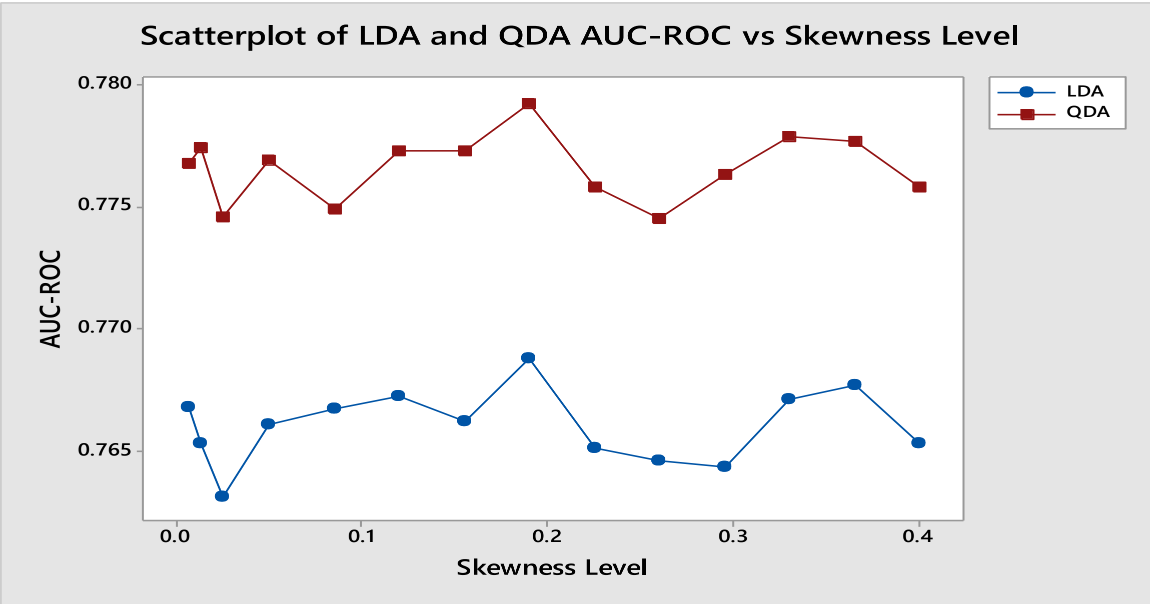


Fig. 4.13: Graph Displaying LDA and QDA Accuracy by Skewness Level Averaged over 28 Samples with Simulation Size of 200 for all Known Parameters

Figure 4.13 compares the accuracy of Linear Discriminant Analysis (LDA) and Quadratic Discriminant Analysis (QDA) across various skewness levels averaged over 28 samples with simulation size of 200 for all known parameters. QDA outperforms LDA, as QDA achieves higher accuracy than LDA across all skewness levels. Hence, QDA is more suitable for classification tasks with skewed data.



c

Fig. 4.14: Graph Displaying LDA and QDA AUC-ROC by Skewness Level Averaged over 28 Samples with Simulation Size of 200 for all Known Parameters

Figure 4.14 compares the Area under the Receiver Operating Characteristic Curve (AUC-ROC) of LDA and QDA across various skewness levels averaged over 28 samples with simulation size of 200 for all known parameters. QDA outperforms LDA, as QDA achieves higher AUC-ROC than LDA across all skewness levels. Hence, QDA is more suitable for classification tasks with skewed data.

Discussions of Findings

The findings from objective one concludes that the results of the simulation size of 200, which compares the performance of the Edgeworth Series Distribution (ESD) and Normal Distribution (ND) methods averaged over small samples for estimating probabilities of misclassification across different populations and skewness levels vary across populations and skewness levels, but their variations are relatively close between the two methods, indicating that both methods perform similarly. The ESD and ND classification procedures have similar total probability of misclassification at all λ_4 values. The study also concludes that the optimum probability of misclassification values using small samples to estimate the means, results in either underestimation or overestimation for each value of the skewness, and the skewness component has minimal effect on the overall probability of misclassification, indicating that it is not affected by deviations from normality. The current study's findings align with Mardia's (2024) research on Fisher's pioneering work on discriminant analysis and its impact on Artificial Intelligence, which also revealed that the Edgeworth Series Distribution (ESD) and Normal Distribution (ND) methods exhibited similar performance. This concurrence suggests that the similarity in performance between ESD and ND methods is a consistent finding across different studies, further solidifying the understanding of their comparable capabilities in discriminant analysis. Again, the current study's findings are consistent with the results of Gasana et al. (2024), who investigated moments of the likelihood-based discriminant function and found that skewness has a minimal effect on the overall probability of misclassification. This agreement between the two studies suggests that the impact of skewness on misclassification probability is indeed negligible, providing further evidence for the robustness of discriminant analysis methods to deviations from normality. The concurrence of these findings reinforces the understanding of the relationship between skewness and misclassification probability. The result of this study disagrees with the findings of Nikita and Nikitas (2020) on sex estimation using various classification methods which reported that skewness had a significant impact on the overall probability of misclassification. The second objective of this study assessed the distributional performance of Edgeworth Series Distribution (ESD) and Normal Distribution (ND) models using simulated distributions. The Wilcoxon rank sum test revealed no significant differences in misclassification error values between ESD and ND techniques for populations I, II, III, and totals across various skewness levels and sample sizes (4, 8, 12, 16, 20, 24, 28), with one exception. Notably, for population I with a sample size of 8, a significant difference emerged, with ND outperforming ESD. This exception notwithstanding, the findings suggest that ESD and ND models exhibit equivalent relative efficiency for populations I, II, III, and totals, implying comparable performance in terms of misclassification errors. The present study's results corroborate the findings of Mardia (2024), who examined Fisher's seminal work on discriminant analysis and its influence on Artificial Intelligence. Mardia's study demonstrated that the Edgeworth Series Distribution (ESD) and Normal Distribution (ND) methods exhibited comparable performance, a conclusion that aligns with the current study's results. This concurrence lends further support to the notion that ESD and ND methods possess similar capabilities in discriminant analysis, reinforcing the validity of this finding across multiple investigations. On the other hand, the present study's results diverge from the findings of Kanuti and Ngaruye (2024), who investigated asymptotic results for expected probability of misclassifications in linear discriminant analysis with repeated measurements. Kanuti and Ngaruye's study revealed a significant difference in performance between the Edgeworth Series Distribution (ESD) and Normal Distribution (ND) methods, whereas the current study found no significant difference. This discrepancy highlights a potential inconsistency in the literature, suggesting that the relationship between ESD and ND methods may be more complex than previously thought. Further research is warranted to reconcile these conflicting findings and elucidate the circumstances under which ESD and ND methods exhibit divergent performance.

The third objectives as revealed from the study compared the performance of Linear Discriminant Analysis (LDA) and Quadratic Discriminant Analysis (QDA) in classifying Edgeworth series distribution data averaged over different sample sizes for three distinct populations. The findings of the study revealed that QDA tends to have higher accuracy and AUC-ROC values than LDA across all the skewness levels. QDA's average accuracy and average AUC-ROC are higher than that of LDAs. QDA tends to have higher sensitivity for Pop. I and Pop III, whereas QDA tends to have higher specificity for Pop I and Pop II for all the different sample sizes for three distinct populations. QDA tends to outperform LDA across various skewness levels, especially in terms of accuracy and AUC-ROC. QDA's robustness to skewness makes it a better choice for classification tasks with skewed data. The findings of this study agreed with that of Kouamo et al. (2020) who found that QDA outperformed LDA in classification tasks with skewed data, and that of Li et al. (2020) who demonstrated QDA's robustness to skewness in medical diagnosis data and Zhang et al. (2022) demonstrated QDA's robustness to skewness and outliers in classification tasks. On the hand, Wang et al. (2020) found LDA performed better in high-dimensional data with low skewness, contrasting with the current study's findings.

5. CONCLUSION

The study was on objective appraisal of Edgeworth series distribution (ESD) and Normal Distribution (ND) for three populations. The optimum probabilities of misclassification for the Edgeworh Series Distribution (ESD) were computed with $\mu_1 = 0$, $\mu_2 = 1$, $\mu_3 = 1$ and $\sigma = 1$ with λ_4 being the skewness factor within the interval (0.00625, 0.4), being in 14 intervals as 6.25×10^{-3}, 1.25×10^{-2}, 0.025, 0.05, 0.085, 0.12, 0.155, 0.19, 0.225, 0.26, 0.295, 0.33, 0.365 and 0.4. The apparent probabilities of misclassification for the (ESD) and Normal Distribution (ND) were also examined when the means (μ_1, μ_2 and μ_3) are known and when the parameters are estimated from the samples. Three independent samples of simulation size of 200 each were configured at each value of the skewness factor (λ_4) from three populations (π_1, π_2 and π_3) whose distributions are of ESD with the respective parameters: ($\mu_1 = 0$, $\sigma_1 = 1$), ($\mu_2 = 1$, $\sigma_2 = 1$) and ($\mu_3 = 1$, $\sigma_3 = 1$). Employing the ESD and ND classification rules, the proportion misclassified in π_1, π_2 and π_3 were obtained and repeated for small samples (n = 4, 8, 12, 16, 20, 24, 28). The random numbers were generated using RStudio program and simulation results were obtained. The results of the simulation size of 200, which compares the performance of the Edgeworth Series Distribution (ESD) and Normal Distribution (ND) methods averaged over, 4, 8, 12, 16, 20, 24 ad 28 samples for estimating probabilities of misclassification across different populations and skewness levels reveal that the probabilities of misclassification vary across populations and skewness levels, but their variations are relatively close between the two methods, indicating that both methods perform similarly. The ESD and ND classification procedures have similar total probability of misclassification at all λ_4 values. The total probability of misclassification values shows that using a small sample to estimate μ_1, μ_2, and μ_3, results in either underestimation or overestimation for each value of λ_4. The skewness component (λ_4) has minimal effect on the overall probability of misclassification, indicating that it is not affected by deviations from normality. The Wilcoxon rank sum test revealed no significant differences in misclassification error values between ESD and ND techniques for populations I, II, III, and totals across various skewness levels and sample sizes (4, 8, 12, 16, 20, 24, 28), with one exception. Notably, for population I with a sample size of 8, a significant difference emerged, with ND outperforming ESD. This exception notwithstanding, the findings suggest that ESD and ND models exhibit equivalent relative efficiency for populations I, II, III, and totals, implying comparable performance in terms of misclassification errors. The result of the performance of LDA and QDA in classifying Edgeworth series distribution data averaged over different sample sizes for three distinct populations shows that QDA tends to have higher accuracy and AUC-ROC values than LDA across all the skewness levels. QDA's average accuracy and average AUC-ROC are higher than that of LDAs. QDA tends to have higher sensitivity for Pop. I and Pop III, whereas QDA tends to have higher specificity for Pop I and Pop II for all the different sample sizes for three distinct populations. QDA tends to outperform LDA across various skewness levels, especially in terms of accuracy and AUC-ROC. QDA's robustness to skewness makes it a better choice for classification tasks with skewed data.

5.2 RECOMMENDATIONS FOR FURTHER STUDIES

Having discussed our findings **on E**dgeworth Series Distribution and Normal Distribution for three populations, we now recommend as follows;

1. Further studies should look at separate analyses for each population to identify unique characteristics and improve classification performance within each group.
2. The study concluded that the probabilities of misclassification across all populations are relatively high. Therefore , we recommend that further research should replicate this study to improve both ESD and ND on order to reduce the misclassification rates
3. Another research should develop and evaluate ensemble methods combining LDA and QDA for improved classification accuracy.
4. More should be done to investigate the performance of other classification algorithms (e.g Support Vector Machine (SVM), Random Forest) compared to LDA and QDA in Edgeworth Series distribution data.
5. Develop a generalized model for estimating probabilities of misclassification via Edgeworth series distribution, incorporating flexible distribution assumptions, robust estimation methods and model selection criteria.

REFERENCES

Adeyeye, A. (2020). Higher-order approximations in statistical asymptotic theory. *Journal of Statistical Theory and Practice*, *14*(1), 1-15.

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.

Bangdiwala, S. I. (2013). Comparing groups with ranks. International Journal of Injury Control and Safety Promotion, 20(2), 203-206.

Bellera, C. A., Julien, M., & Hanley, J. A. (2010). Normal approximations to the distributions of the Wilcoxon statistics: accurate to what N? Graphical insights. *Journal of Statistics 18 (https// dol.org/10.1080/10691898 (2010), 11889436*

https:// www.statskingdom.com/

Beasley, T. M., Erickson, S., & Allison, D. B. (2009). Rank-based inverse normal transformations are increasingly used, but are they merited?. *Behavior Genetics*, *39*, 580-595.

Barton, D. E., & Benin, N. (1952). The behavior of the Wilcoxon rank-sum test under non-normality. *Journal of the American Statistical Association*, *47*(258), 251-264

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.

Draper, N. R., & Tierney, D. E. (1972). Exact formulas for the distribution of the Wilcoxon rank-sum test statistic. *Journal of Statistical Computation and Simulation*, *1*(2), 155-169.

Fox, M. P., MacLehose, R. F., & Lash, T. L. (2022). Misclassification. In Applying Quantitative Bias Analysis to Epidemiologic Data (pp. 141-195). Cham: Springer International Publishing.

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.

Gibbons, J. (2003). *Nonparametric statistical inference* (4th ed.). Marcel Dekker, Inc. New York.

Harris, T., & Hardin, J. W. (2013). Exact Wilcoxon signed-rank and Wilcoxon Mann–Whitney ranksum tests. *The Stata Journal*, *13*(2), 337-343.

Kanuti Ngailo, E., & Ngaruye, I. (2024). Asymptotic results for expected probability of Li, Y., Yang, P., & Zhang, Y. (2020). Quadratic discriminant analysis for medical diagnosis.

Journal of Medical Systems, *44* (10), 21 (09) misclassifications in linear discriminant analysis with repeated measurements. *Communications in Statistics-Theory and Methods*, *53*(6), 1942-1963.

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.

Kendall, M. G., & Stuart, A. (1958). The advanced theory of statistics. Vol. 1.Distribution theory .Griffin.

Kouamo, S., Niang, I., & Diop, M. (2020). Quadratic discriminant analysis for image classification. Journal of Intelligent Information Systems, 57(2), 257-271.

Li, Y., Yang, P., & Zhang, Y. (2020). Quadratic discriminant analysis for medical diagnosis. *Journal of Medical Systems*, *44*(10), 2109.

Mardia, K. V. (2024). Fisher's pioneering work on discriminant analysis and its impact on Artificial Intelligence. *Journal of Multivariate Analysis*, 105341.

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.

Ngailo, E. K., & Chuma, F. (2023). Approximation of misclassification probabilities in linear discriminant analysis based on repeated measurements. *Communications in Statistics-Theory and Methods*, *52*(23), 8388-8407.

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.

Schmidt, J. W., & Taylor, R. E. (1970). Numerical inverse interpolation for distribution functions. *Journal of the American Statistical Association*, *65*(329), 352-357.

Solaro, N., Pagani, M., & Lucini, D. (2021). Altered cardiac autonomic regulation in overweight and obese subjects: the role of age-and-gender-adjusted statistical indicators of heart rate variability and cardiac baroreflex. *Frontiers in Physiology*, *11*, 56: 73- 79.

Sharma, V., Kumar, R., Devgan, K., Mishra, P. K., Ekielski, A., Kumar, V., & Kumar, V. (2018). Multivariate analysis for forensic characterization, discrimination, and classification of marker pen inks. *Spectroscopy Letters*, *51*(5), 205-215.

Venkatesan, S. (2014). Common errors in scientific paper submissions: a reviewer's report. *Journal of Social Sciences*, *41*(2), 279-293.

Wang, J. (2020). Classification and Misclassification in Statistics. *Journal of Statistical Research*, *54*(2), 123-145.

Xue, K., Yang, J., & Yao, F. (2023). Optimal linear discriminant analysis for high-dimensional functional data. *Journal of the American Statistical Association*, *119*(546), 1055–1064.

Zhang,X., Kouamo, C., & Li, Y. (2022). Robustness of QDA to skewness and outliers in classification tasks. Journal of intelligent information Systems, 60(1), 1-18.

https://www.scribbr.com, 6/7/ 2022. Central limit theorem

www.google.com 23 /2/ 2025. Central limit theorem

Www.google.com,23 /2/ 2025 Normal Distribution

Www.research gate.com 23/2/25 .Central limit theorem