Graphical Evaluation of Probabilities of Misclassifications for Normal and Edgeworth Series Distributions

Awogbemi, Clement Adeyeye

Department of Statistics, National Mathematical Centre, Abuja, Nigeria

Abstract:- **Failure to transform to normality before classification affects probabilities of misclassification while comparing the distribution of errors of misclassification. A uniformly distributed random variable generated by a varied and repeated method was employed to generate the errors of misclassification for normal and Edgeworth Series Distributions. The proposed method was proved and on the basis of this, an algorithm was developed. There is a non linear dependence of the total probabilities of misclassification on the skewness factor. It was observed that there is a disordered relationship between the probabilities of misclassication for normal and Edgeworth Series Distributions.**

Keywords: **Graphical Evaluation, Probabilities of Misclassification, Normal Distribution, Edgeworth Series Distribution, Skewness Factor**

I. INTRODUCTION

Identifying an appropriate region for classification has been a challenge to researchers. For an experimenter who does not recognize a region to be non-normal, he proceeds to use normal region for classification [1]. The problem that emanates from this scenario is "how does the failure to transform to normality, prior to classification, affect the probabilities of misclassification when there is a need to compare the distribution of errors of misclassification" [2],[3],[4].

Errors of misclassification for classification problems, with two classes of univariate gamma distribution, were studied by [5]. The gamma density functions used were reparameterized. The effects of applying the normal classificatory rule to gamma distribution were studied and assessed theoretically by comparing optimum and conditional probabilities of misclassification.

Errors of misclassification associated with the Inverse Gaussian Distribution (IGD) were worked upon by [6], focusing on the classification cases that are in line with the univariate form of the distribution considered. The effects of applying the Linear Discriminant Function (LDF) to IGD were utilized on the basis of normality. This was done by comparing the optimum and conditional probabilities based on the LDF and likelihood ratio for the distribution under consideration.

Errors of misclassifications of Normal and Edgeworth Series Distributions in a tabular form, were generated by [2], but in this work, probabilities of misclassification for Normal and Edgeworth Series Distributions are being interpreted graphically and discussed.

Preliminaries

Suppose X_{ij}, $i = 1, 2,\ j = 1, 2, \ldots, n_i$, denote two independent random samples from populations, π_i, $i = 1, 2,$ respectively. The observations X_{ij} emanate from the common distribution defined by the density function

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

The parameter, λ_3, $\mu_i (i = 1, 2)$, satisfies the conditions: $-\infty < \lambda_3 < \infty$, $-\infty < \mu_i < \infty$ and

$\sigma > 0$, where D denotes the differential operator, $\frac{d}{dx}$ *and*

$$\phi\left(\frac{x - \mu_i}{\sigma}\right) = \frac{(2\pi)^{-\frac{1}{2}}}{\sigma} \exp\left[\frac{-(x - \mu_i)^2}{2\sigma^2}\right]$$

and λ_3 is the skewness factor [7].

Let $X_1, X_2, \ldots X_k$ be independent and identically distributed random variables with mean $\theta_0 = \mu$ and finite variance, σ^2. If $\hat{\theta}_k$ is constructed from a sample of size, n, and $k^{-\frac{1}{2}}(\hat{\theta}_k - \theta_0)$ is asymptotically and normally distributed, then Edgeworth Series expansions are developed as approximations to distribution of estimates $\hat{\theta}_k$ of unknown quantities, θ_0. Thus the distribution functions of $k^{-\frac{1}{2}}(\hat{\theta}_k - \theta_0)$ is expanded as a power series in $k^{-\frac{1}{2}}$ *so that*

$$P\left(k^{\frac{1}{2}}\frac{\left(\hat{\theta}_k - \theta_0\right)}{\sigma} \le y\right) = \Phi(x) + k^{-\frac{1}{2}}P_1(x)\phi(x) + k^{-1}P_2(x)\phi(x) + \ldots + k^{-\frac{j}{2}}P_j(x)\phi(x) + \ldots, \quad (1.3)$$

where $\phi(x) = \sigma^{-1}\left(2\pi\right)^{-\frac{1}{2}} e^{-\frac{x^2}{2\sigma^2}}$ is the standard normal density function, $\hat{\theta}_k$ is the estimate of θ_k, θ_0 is the true value of unknown parameter θ and $\Phi(x) = \int_{-\infty}^{x} \phi(x)du$ is the standard normal distribution function. The functions P_j are polynomials with coefficients, depending on cumulants of $\hat{\theta}_k - \theta_0$. In particular, P_j is a polynomial of degree 3j -1 and is odd for even j and even for odd j.

Equation (1.3) is the Edgeworth Series or expansion, and the term of order $k^{-1/2}$ in the same equation corrects the basic normal approximation for the main effect of skewness, while the term of order k^{-1} corrects the main effect of kurtosis.

II. METHOD OF GENERATING PROBABILITIES OF MISCLASSIFICATION

The simulation of experiment for the generation of probabilities of misclassification is anchored on the work of [2]. However, further attempt is made in this work for the proposition of inverse transformation method. Let $N \sim U(0,1)$ be a uniformly distributed continuous random variable X. For any cumulative distribution F(x) that is strictly increasing over all x, we have 0 < F(x) < 1, Suppose the random variable X is defined by X= $F^{-1}(U)$, then the random variable X has the distribution F. This implies that $F^{-1}(x)$ is defined to be equal to the value of x for which F(x) = x [8].

Proof

$$F_y(a) = P(X \le a)$$
$$= P[F^{-1}(U) \le a]$$

Since F(x) is a monotone function, it follows that $F^{-1}(U) \le a$ if and only if

$F_y(a) = P[U \le f(a)] = F(a).$ From the above proposition, a random variable X with a continuous function F is simulated by generating a number U and then setting $X = F^{-1}(U)$

Algorithm

1. Set $N \sim U(0,1)$ where U is the generator of the random observation
2. Set F(x) = U
3 Solve F(x) = U for X such that F^{-1} (U) = X, X is the generated random observation.

III. SIMULATION RESULTS

Table 1: Optimum Probabilities of Misclassification at Different Values of Skewness for ESD

	Optimum Probability of Misclassification		
Skewness Factor (λ_3)	**E_{12E}**	**E_{21E}**	**Total**
0.00625	0.3082	0.3088	0.6170
0.0125	0.3079	0.3091	0.6170
0.025	0.3074	0.3096	0.6170
0.05	0.3063	0.3107	0.6170
0.10	0.3041	0.3129	0.6170
0.15	0.3019	0.3151	0.6170
0.20	0.2997	0.3173	0.6170
0.25	0.2975	0.3195	0.6170
0.30	0.2953	0.3217	0.6170
0.35	0.2931	0.3239	0.6170
0.40	0.2909	0.3261	0.6170

Source: [2]

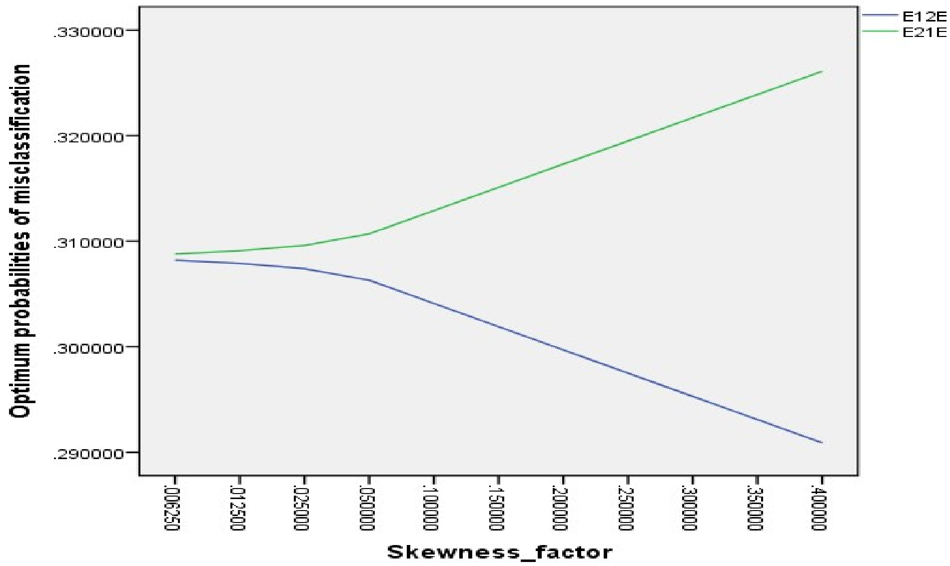

Figure 1.1: Graph showing optimum probabilities of misclassification at different values of skewness for ESD (all parameters known)

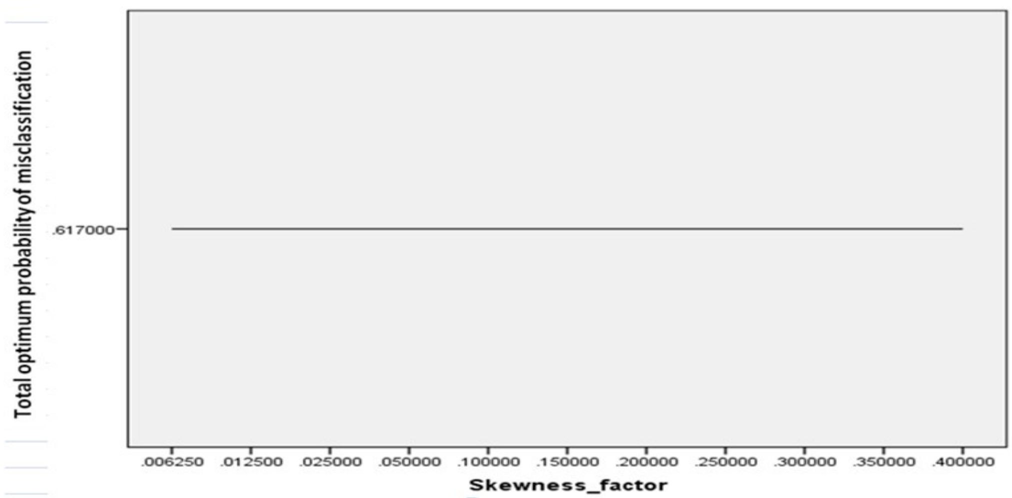

Figure1. 2: Graph showing total optimum probabilities of misclassification at different values of skewness for ESD

Table 2: Comparison of Errors of Misclassification for Means unknown and Estimated by Average Values over 5 Samples.

	ESD			**ND**		
Skewness Factor (λ_3)	**E_{12E}**	**E_{21E}**	**Total**	**E_{12N}**	**E_{21N}**	**Total**
0.00625	0.140	0.400	0.540	0.140	0.400	0.540
0.0125	0.220	0.410	0.630	0.220	0.410	0.630
0.025	0.225	0.465	0.690	0.220	0.475	0.695
0.05	0.210	0.395	0.605	0.205	0.400	0.605
0.10	0.205	0.475	0.680	0.175	0.495	0.670
0.15	0.260	0.285	0.545	0.230	0.320	0.550
0.20	0.305	0.365	0.670	0.295	0.395	0.690
0.25	0.455	0.185	0.640	0.420	0.230	0.650
0.30	0.195	0.465	0.660	0.115	0.545	0.660
0.35	0.225	0.465	0.660	0.125	0.520	0.645
0.40	0.440	0.180	0.610	0.360	0.250	0.610

Source: [2]

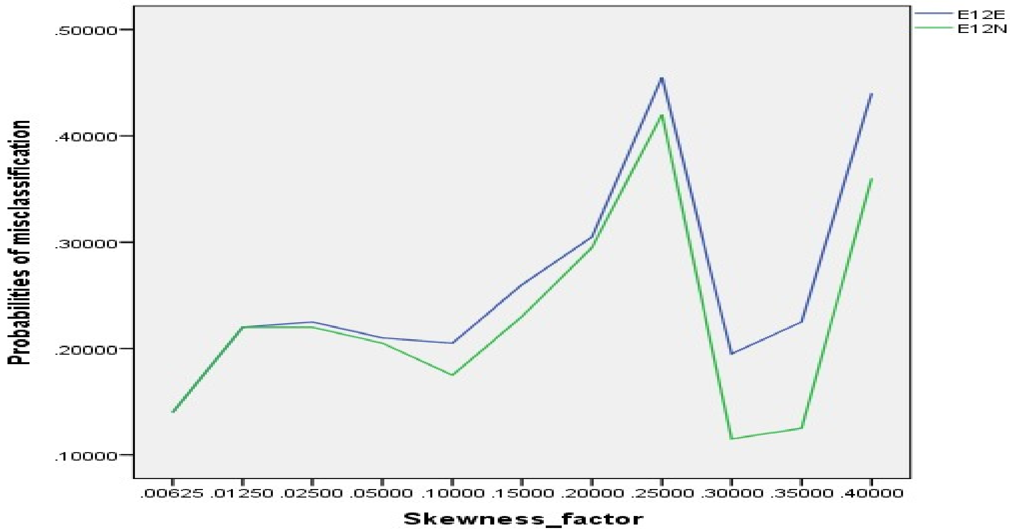

Figure 2.1: Graph showing probabilities of misclassification for unknown parameters averaged over 5 samples.

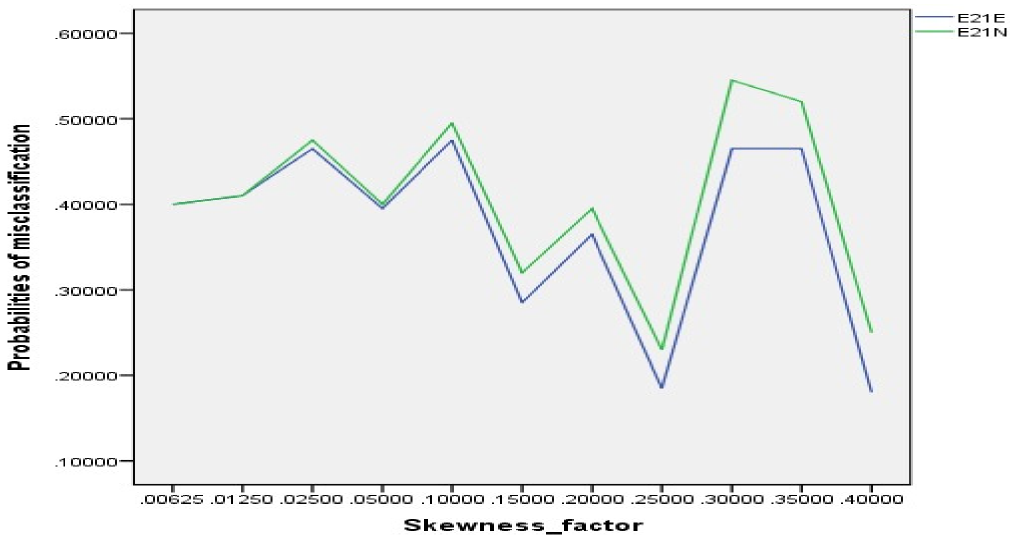

Figure 2.2: Graph showing probabilities of misclassification for unknown parameters averaged over 5 samples.

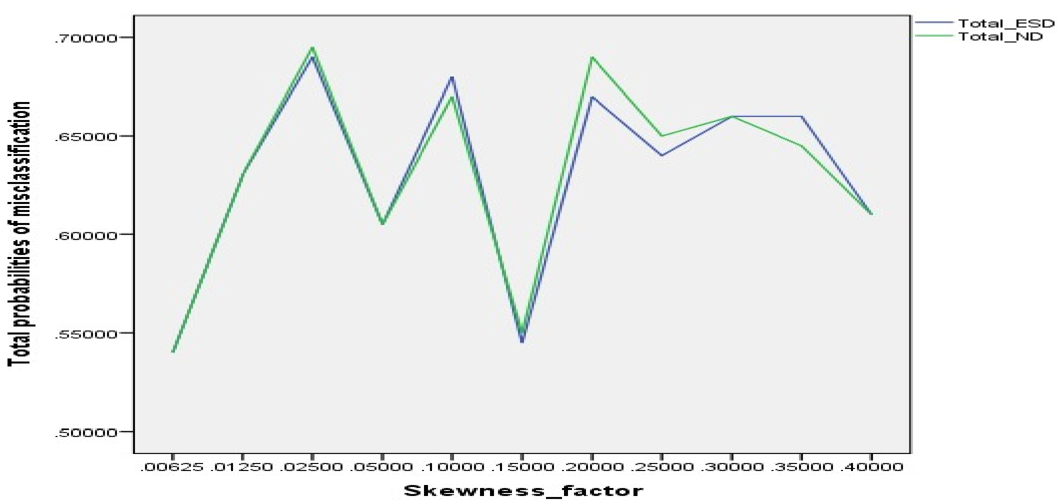

Figure 2.3: Graph showing total probabilities of misclassification for unknown parameters average over 5 samples.

Table 3: Comparison of Errors of Misclassification for Means unknown and Estimated by Average Values over 10 Samples.

	ESD			ND		
Skewness Factor (λ_3)	**E_{12E}**	**E_{21E}**	**Total**	**E_{12N}**	**E_{21N}**	**Total**
0.00625	0.252	0.249	0.501	0.252	0.315	0.567
0.0125	0.236	0.236	0.472	0.236	0.236	0.472
0.025	0.266	0.219	0.485	0.231	0.295	0.526
0.05	0.224	0.282	0.506	0.216	0.314	0.530
0.10	0.290	0.278	0.568	0.208	0.336	0.544
0.15	0.387	0.203	0.590	0.215	0.220	0.435
0.20	0.277	0.320	0.597	0.270	0.337	0.607
0.25	0.255	0.245	0.500	0.230	0.292	0.522
0.30	0.248	0.334	0.582	0.182	0.394	0.576
0.35	0.216	0.339	0.555	0.175	0.354	0.529
0.40	0.253	0.209	0.462	0.170	0.196	0.366

Source: [2]

Figure 3.1: Graph showing probabilities of misclassification for unknown parameters averaged over 10 samples

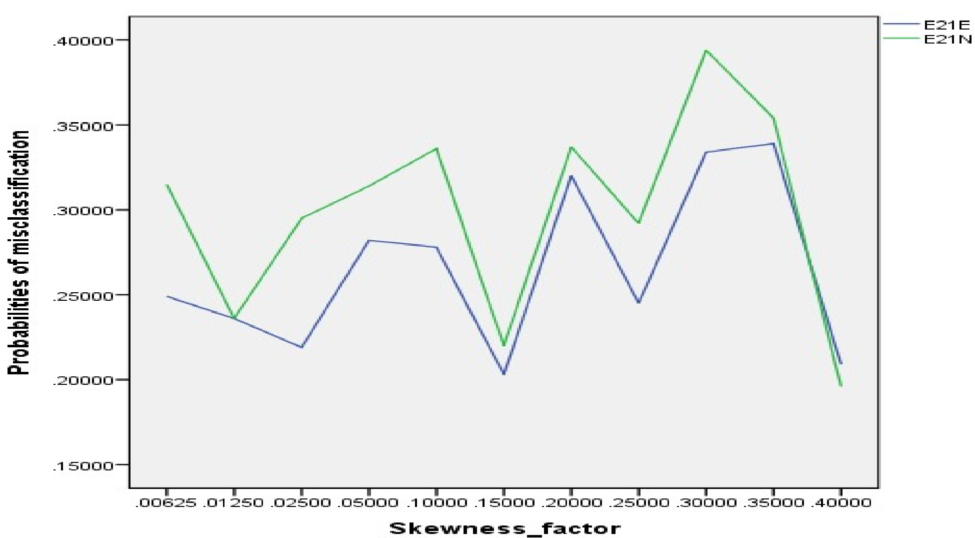

Figure 3.2: Graph showing probabilities of misclassification for unknown parameters averaged over 10 samples

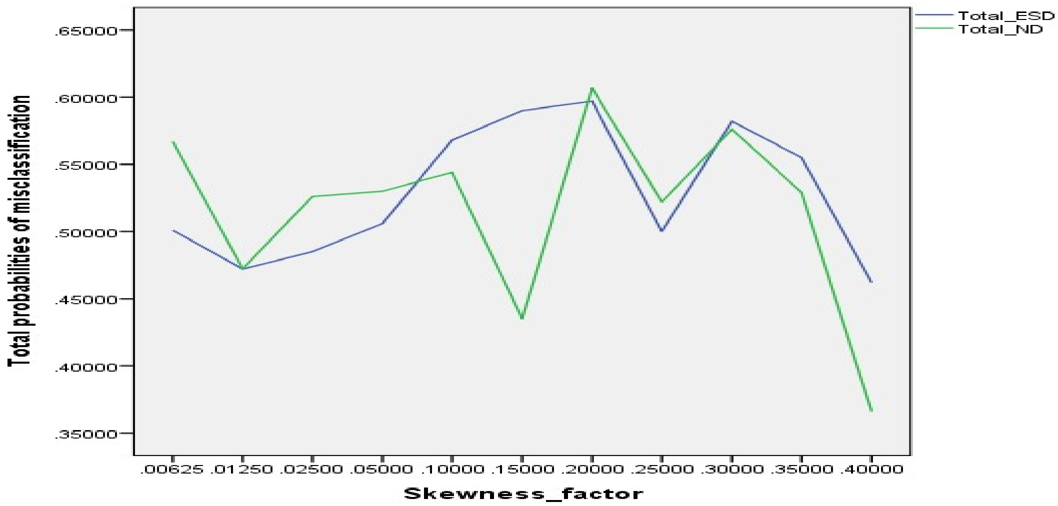

Figure 3.3: Graph showing probabilities of misclassification for unknown parameters averaged over 10 samples

Table 4: Comparison of Errors of Misclassification for Means unknown and Estimated by Average Values over 15 Samples.

Skewness Factor (λ_3)	E_{12E}	E_{21E}	Total	E_{12N}	E_{21N}	Total
0.00625	0.345	0.145	0.490	0.345	0.150	0.495
0.0125	0.310	0.310	0.620	0.310	0.310	0.620
0.025	0.405	0.280	0.685	0.400	0.285	0.685
0.05	0.230	0.390	0.620	0.225	0.395	0.620
0.10	0.375	0.305	0.680	0.350	0.315	0.665
0.15	0.405	0.180	0.585	0.360	0.225	0.585
0.20	0.355	0.325	0.680	0.320	0.355	0.675
0.25	0.295	0.340	0.635	0.235	0.395	0.630
0.30	0.320	0.350	0.670	0.230	0.385	0.615
0.35	0.260	0.345	0.605	0.200	0.430	0.630
0.40	0.315	0.375	0.690	0.145	0.415	0.560

Source: [2]

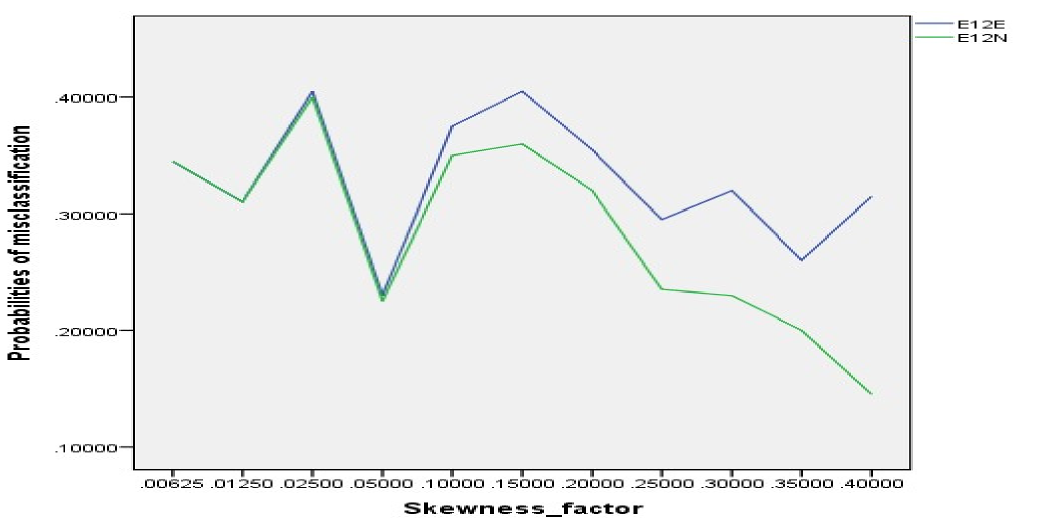

Figure 4.1: Graph showing probabilities of misclassification for unknown parameters averaged over 15 samples

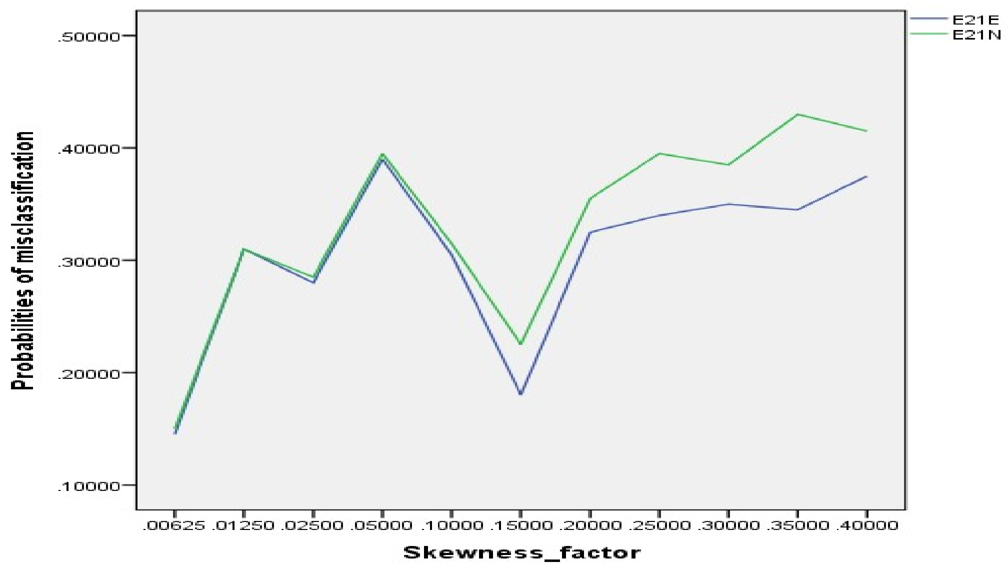

Figure 4.2: Graph showing probabilities of misclassification for unknown parameters averaged over 15 samples

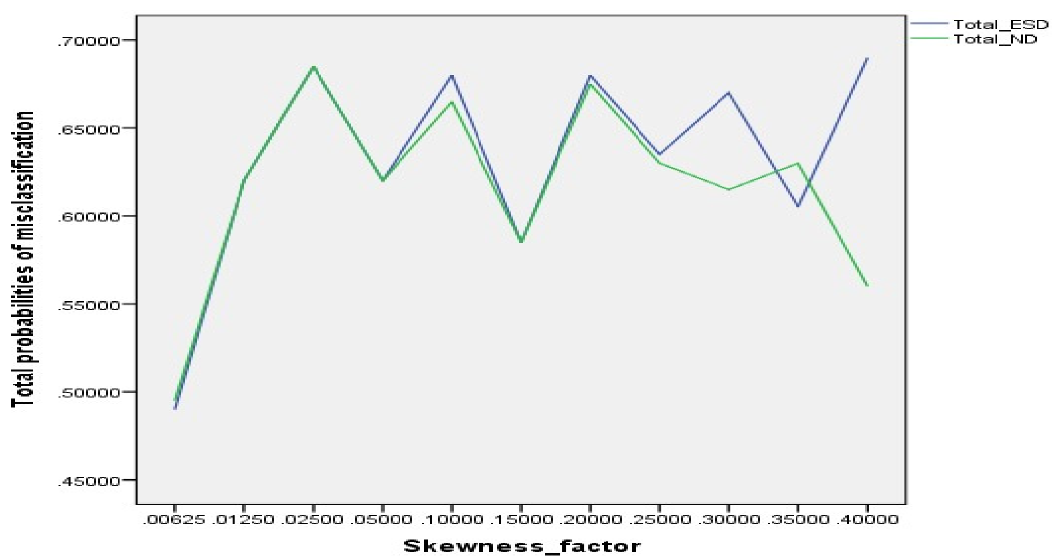

Figure 4.3: Graph showing probabilities of misclassification for unknown parameters averaged over 15 samples

Table 5: Comparison of Errors of Misclassification for Means unknown and Estimated by Average Values over 20 Samples.

Skewness Factor (λ_3)	E_{12E}	E_{21E}	Total	E_{12N}	E_{21N}	Total
0.00625	0.220	0.206	0.426	0.220	0.206	0.426
0.0125	0.280	0.280	0.560	0.192	0.295	0.487
0.025	0.330	0.210	0.540	0.290	0.230	0.520
0.05	0.345	0.205	0.550	0.295	0.250	0.545
0.10	0.265	0.300	0.565	0.230	0.390	0.620
0.15	0.340	0.350	0.690	0.330	0.375	0.705
0.20	0.350	0.240	0.590	0.320	0.255	0.575
0.25	0.295	0.270	0.565	0.270	0.295	0.565
0.30	0.300	0.195	0.495	0.265	0.200	0.465
0.35	0.310	0.350	0.660	0.270	0.360	0.630
0.40	0.405	0.285	0.690	0.380	0.400	0.780

Source:[2]

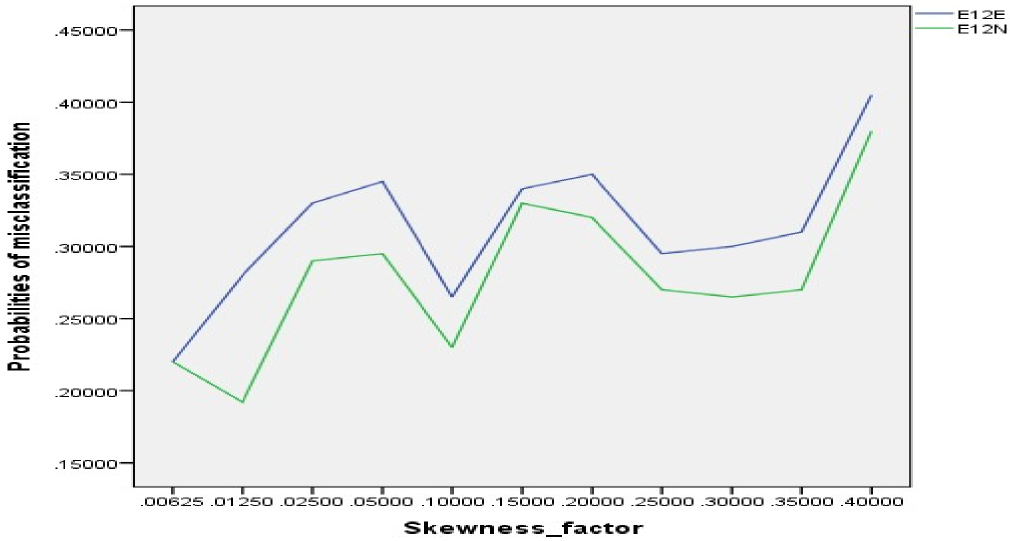

Figure 5.1: Graph showing probabilities of misclassification for unknown parameters averaged over 20 samples

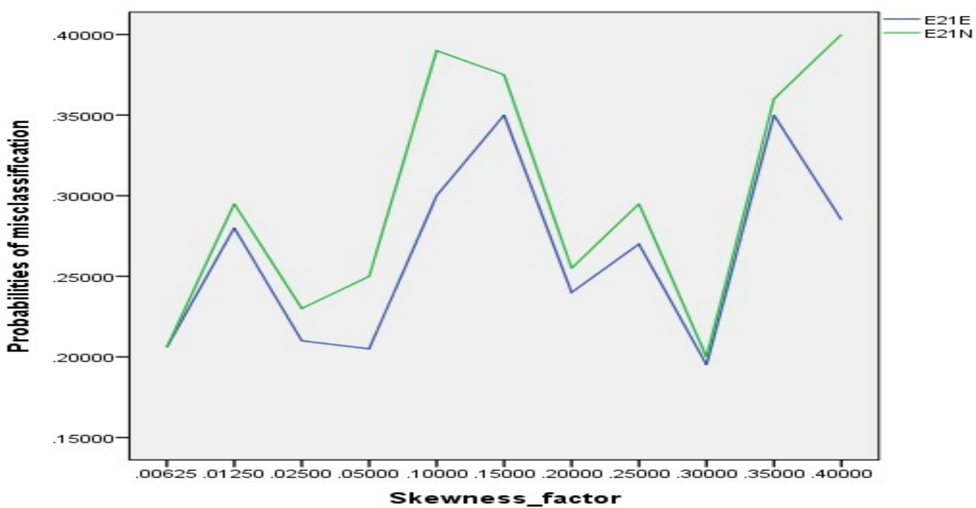

Figure 5.2: Graph showing probabilities of misclassification for unknown parameters averaged over 20 samples.

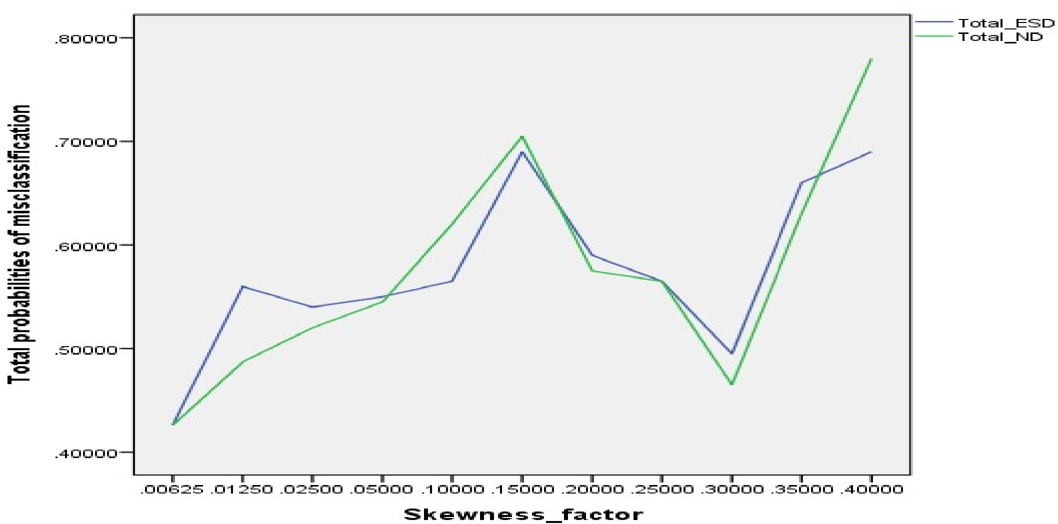

Figure 5.3: Graph showing probabilities of misclassification for unknown parameters averaged over 20 samples

Table 6: Comparison of Errors of Misclassification for Means unknown and Estimated by Average Values over 25 Samples.

Skewness Factor (λ_3)	E_{12E}	E_{21E}	Total	E_{12N}	E_{21N}	Total
0.00625	0.270	0.220	0.490	0.270	0.220	0.490
0.0125	0.290	0.330	0.620	0.290	0.235	0.525
0.025	0.390	0.295	0.685	0.375	0.310	0.685
0.05	0.340	0.270	0.610	0.335	0.280	0.615
0.10	0.375	0.305	0.680	0.360	0.315	0.675
0.15	0.360	0.230	0.590	0.345	0.245	0.590
0.20	0.275	0.430	0.705	0.225	0.480	0.705
0.25	0.375	0.255	0.630	0.320	0.290	0.610
0.30	0.390	0.240	0.630	0.300	0.330	0.630
0.35	0.290	0.300	0.590	0.240	0.345	0.585
0.40	0.405	0.225	0.630	0.305	0.290	0.595

Source:[2]

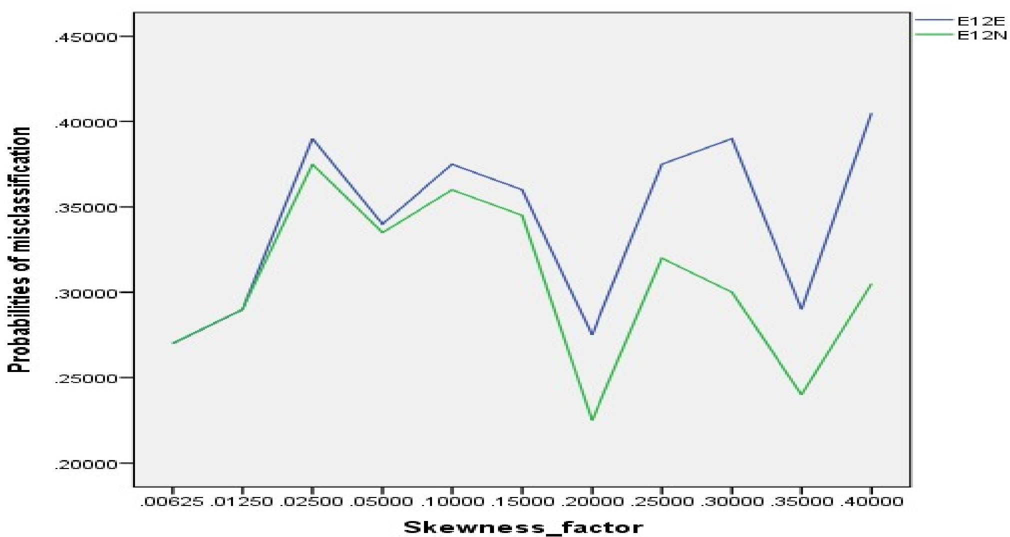

Figure 6.1: Graph showing probabilities of misclassification for unknown parameters averaged over 25 samples

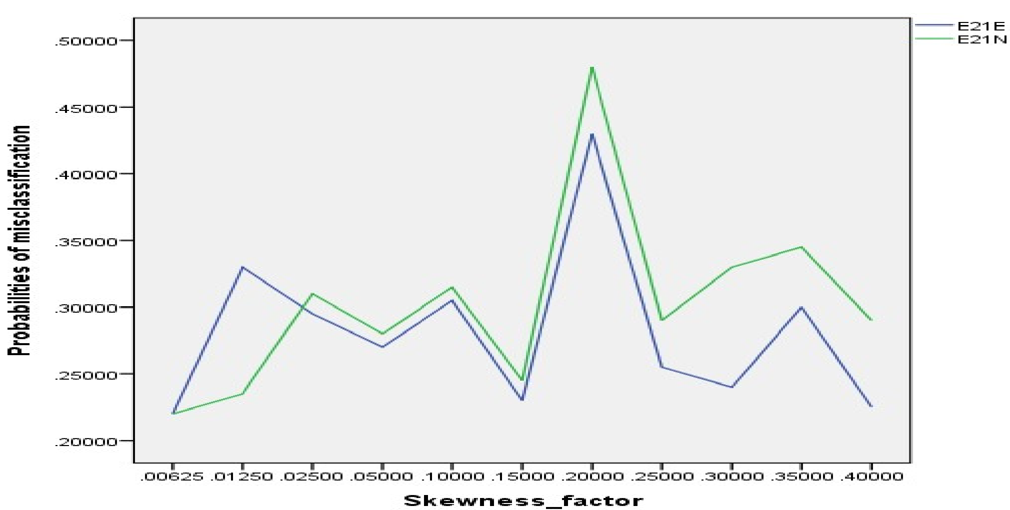

Figure 6.2: Graph showing probabilities of misclassification for unknown parameters averaged over 25 samples

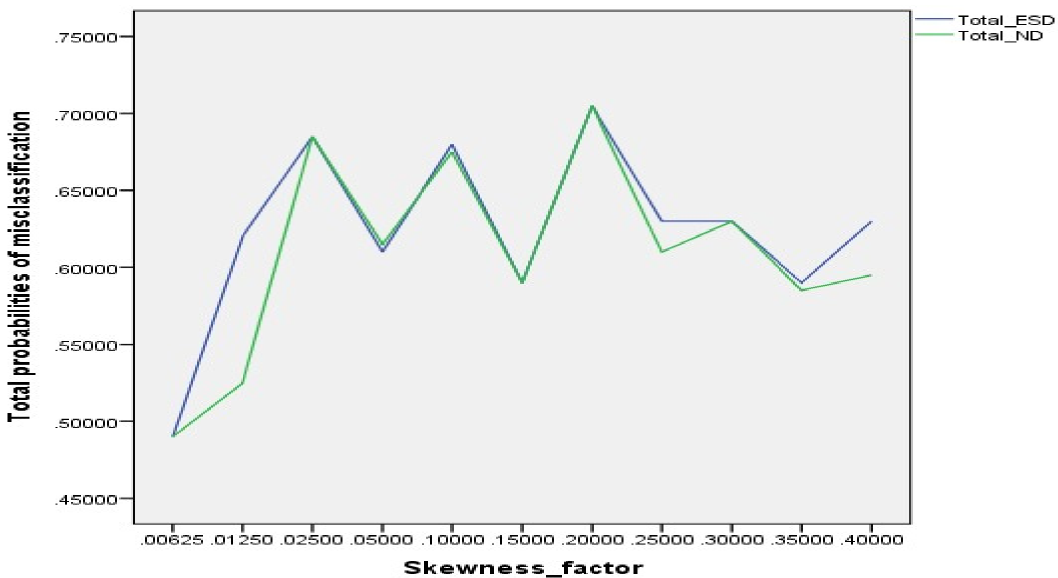

Figure 6.3: Graph showing probabilities of misclassification for unknown parameters averaged over 25 samples

IV. DISCUSSION OF RESULTS

The plots of the probabilities of misclassification in Tables 1-6 are all shown in Figures 1.1- 6.3.

In Figure 1.1, there is a positive linear relationship between the skewness factor (λ_3) and E_{12E}, and negative linear relationship between skewness factor (λ_3) and E_{12N}. The upward trend of E_{12E} is more pronounced when $\lambda_3 = 0.00625$, and the down ward trend is more pronounced when $\lambda_3 = 0.05$.

In Figure 1.2, the total probabilities of misclassification remain constant irrespective of the increase in the skewness factor (λ_3). The relationship between (λ_3) and the total optimum probabilities of misclassification is unpredictable since there is non-linear dependence of the total optimum probabilities of misclassification on the skewness factor (λ_3).

In Figures 2.1, the up and down trends of E_{12E} and E_{12N} move towards the same direction with E_{12E} and E_{12N} meeting at points 0.140 when $\lambda_3 = 0.00625$ and point 0.220 when $\lambda_3 = 0.0125$.

The up and down movements of E_{21E} and E_{21N} in Figure 2.2 also follow the same direction. E_{21E} and E_{21N} meet at points 0.400 when $\lambda_3 = 0.00625$ and at point 0.410 when $\lambda_3 = 0.0125$.

In Figure 2.3, the trends of total probabilities of misclassification for ESD and ND also follow the same direction with three meeting points: 0.540 when $\lambda_3 = 0.00625$, 0.630 when $\lambda_3 = 0.0125$ and 0.610 when $\lambda_3 = 0.40$.

In Figure 3.1, there is disordered movements of trends of E_{12E} and E_{12N} which eventually coincide at point 0.252 when $\lambda_3 = 0.00625$ and at point 0.236 when $\lambda_3 = 0.0125$

In Figure 3.2, the up and down movements of the trends E_{21E} and E_{21N} meet at point 0.236 when $\lambda_3 = 0.0125$.

In Figure 3.3, the trends of the total probabilities of misclassification for ESD and ND are interwoven with the meeting point at 0.472 when $\lambda_3 = 0.0125$.

The plots of E_{12E} and E_{12N} in Figure 4.1 meet at point 0.345 when $\lambda_3 = 0.00625$, 0310 when $\lambda_3 = 0.0125$.

In Figure 4.2, the trends of E_{21E} and E_{21N} meet at point 0.310 when $\lambda_3 = 0.0125$.

In Figure 4.3, the trends of the plots of total probabilities of misclassification for ESD and ND meet at point 0.620 when $\lambda_3 = 0.0125$, 0.05 and at point 0.685 when $\lambda_3 = 0.025$.

In Figure 5.1, the plots of E_{12E} and E_{12N} meet at point 0.220 when $\lambda_3 = 0.00625$, with the plots also swinging in the same direction.

The plots of E_{21E} and E_{21N} in Figure 5.2 exhibit up and down movement in a chaotic form, with the trends of E_{21E} and E_{21N} - meeting at point 0.206 when $\lambda_3 = 0.00625$.

The plots of total probabilities of misclassification using ESD and ND in Figure 5.3 are in disordered form, but meeting at point = 0.426 when $\lambda_3 = 0.00625$.

In Figure 6.1, the up and down movements of trends E_{12E} and E_{12N} meet at point 0.270 when $\lambda_3 = 0.0625$ and at point 0.290 when $\lambda_3 = 0.0125$.

In Figure 6.2, the plots E_{21E} and E_{21N} are also not in ordered form. The trends of E_{12E} and E_{12N} meet at point 0.220 when $\lambda_3 = 0.00625$.

The plots of the total probabilities of misclassification using ESD and ND in Figure 6.3, at every level of skewness factor (λ_3), meet at point 0.490 when $\lambda_3 = 0.00625$ and point 0.630 when $\lambda_3 = 0.30$.

V. CONCLUSION

Graphical evaluation of the errors of misclassification using Normal and Edgeworth Series Distributions has been discussed in this work. An algorithm from a uniformly distributed random variable that was proved has also been developed. It is observed that there exists a disordered relationship between the probabilities of misclassification for Normal and Edgeworth Series distributions. The total probabilities of misclassification remain constant irrespective of the increase in the skewness factor.

REFERENCES

[1] Awogbemi, C.A. and Onyeagu, S.I. (2019). Errors of misclassification Associated with Edgeworth Series Distribution, American Journal of Theoretical and Applied Statistics, 8(6), pp. 203-213.

[2] Awogbemi, C.A. (2019). Errors of Misclassification Associated with Edgeworth Series Distribution, Unpublished Ph.D. Dissertation, Department of Statistics, Faculty of Physical Sciences, Nnamdi Azikiwe University, Awka, Nigeria.

[3] Kocherlakota, S., Kocherlakota, K, and Balakrishnan, N. (1987). Expansions of Errors of Misclassification, Advances in Multivariate Statistical Analysis, pp. 19 -21.

[4] Fred, S. (2010). The Refined Positive Definite and Unimodal Regions for the Gram-Charlier and Edgeworth Series Expansions. Advances in Decision Sciences, pp. 5-8.

[5] Mahmoud, M.A. and Moustafa, H.M. (1995). Errors of misclassification Associated with Gamma Distribution. Journal of Mathematical Computing and Modeling, 22(3), pp. 105-119.

[6] Amoh, R.K. and Kocherlakota, K. (1986). Errors of Misclassification Associated with Inverse Gaussian Distribution, Communications in Statistics, Theory and Methods, 15(2), pp. 589-612.

[7] Bhuyan, C.K. (2010). Probability Theory and Statistical Inference, New Central Book Agency Ltd., London

[8] Sheldon, R. (2013). A First Course in Probability, 9th ed., Pearson, University of Southern Califonia.