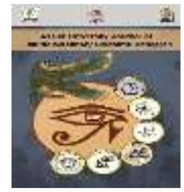
Assiut University Journal of Multidisciplinary Scientific Research (AUNJMSR)
Faculty of Science, Assiut University, Assiut, Egypt.
Printed ISSN 2812-5029
Online ISSN 2812-5037
Vol. 54 (3): 404- 415 (2025)
https://aunj.journals.ekb.eg



Estimation of Discriminant Functions from a Mixture of Two Weibull Population Mean Distributions

Omar M. Ahmed [1], Shereen A. Mohammed [2,*] and Ayman M. Abd-Elrahman [2]

[1]Department of Basic Sciences, Faculty of Computer Science and Information Technology, Ahram Canadian University, 6th of October City, Egypt.
[2]Department of Mathematics, Faculty of Science, Assiut University, Assiut, Egypt.

[*]Corresponding Author: shereen.elsayed@science.aun.edu.eg

ARTICLE INFO
Article History:
Received: 2025-04-23
Accepted: 2025-06-16
Online: 2025-08-28

Keywords:
Finite mixture, maximum likelihood estimation, discriminant function, bias, mean square error, relative efficiency.

ABSTRACT

This paper introduces the Weibull population mean distribution (WPMD), providing a method for the maximum likelihood estimates (MLEs) of the parameters in a finite mixture of Weibull population mean distributions. The methodology is investigated through both mixture and classified procedures. Measurement of nonlinear discriminant function estimates based on different sample sizes. that it presentation is examined by simulation experiments. The experiments result in estimating a nonlinear discriminant function using the maximum likelihood approach, based on mixture data obtained from a combination of the basic groups are proposed that the rate of errors can be greatly reduced for groups that are typically separated. Overall, the performance of the hybrid discriminant procedure (assessed in terms of overall probability) is good compared to the full classification approach.

1. Introduction

The development of mixture models has important applications on a lot of applied fields. The mixture model has been expansively studied by numerous authors to provide a more comprehensive overview of numerical methods, discussions and applications we looked at some features of a finite mixture of a life-cycle model. In addition, hypotheses examining that number of modules have been explored. This paper focuses on deriving and estimating nonlinear discriminant functions from WPMD using a common shape criterion based on different sampling schemes [1]. Finite mixtures of distributions have been used as models throughout the history of modern statistics. The mixing of two populations is typically associated with two main difficulties. The first problem is to estimate the parameters of both populations using unclassified data. Another issue with the mixture model is the estimation of a discriminant function from unclassified data and the analysis of its performance [2]. Studies in this area have been under taken by O'Neill [3] and Ganasalingam and Mclachlan [4,5]. In all these studies the underlying population are assumed to be normal. Amoh [6] estimated a discriminant function from a mixture of two inverse Gaussian distributions when sample size is small. Mahmoud and Moustafa [7] have estimated a discriminant function from a mixture of two gamma distributions when the sample size is small. Ahmad [8] has studied small-sample results for a nonlinear discriminant function estimated from a mixture of two Burr type-XII distributions. Also, Ahmad and Abd-Elrahman [9] have studied a nonlinear discriminant function estimated from a mixture of two Weibull distributions. Mahmoud and Moustafa [10] have studied the errors of misclassification associated with the gamma distribution. Ahmad [11] has studied the efficiency of a nonlinear discriminant function based on unclassified initial samples from a mixture of two Burr type-XII distributions. Moustafa and Ramadan [12] have estimated a discriminant function from a mixture of two Gompertz distributions when the sample size is small. Recently, Ahmad et al. [13] have estimated a discriminant function from a mixture of two Gumbel distributions when the sample size is small. In response to these limitations, researchers began to explore alternative component distributions. Abd-Elrahman [14] introduces a novel statistical distribution, the Weibull population mean distribution WPMD, as a flexible alternative to existing models. The WPMD is characterized by its ability to accommodate various shapes of the hazard rate function, including increasing, decreasing, and bathtub shaped. This flexibility makes it suitable for modeling a wide range of real-world phenomena, particularly in fields such as reliability engineering and survival analysis. We will refer to the Weibull population mean distribution as WPMD. To derive the WPMD, the probability density function (PDF) of the Weibull distribution is multiplied by the factor $(\frac{x}{\theta})^{\lambda}$, resulting in the following expression for the PDF of the WPMD:

$$f_X(x;\theta,\lambda) = \left(\frac{\lambda}{\theta}\right)\left(\frac{x}{\theta}\right)^{2\lambda-1} e^{-\left(\frac{x}{\theta}\right)^{\lambda}}, \quad x>0, \ (\theta,\lambda>0), \tag{1.1}$$

where θ and λ are the scale and shape parameters, respectively.

2. MLEs of the Parameters

Suppose that λ is a common parameter. Then, the probability density function of a mixture of two components of $WPMD(\theta_i,\ \lambda)$ is given by

$$f(x;p,\theta_1,\theta_2,\lambda) = pf((x;\theta_1,\lambda) + qf((x;\theta_2,\lambda) = S(x)\,[pg_1(x) + qg_1(x)\,], \tag{2.1}$$

where $S(x)=\lambda\, x^{2\lambda-1}e^{-x^{\lambda}}$ *and* $g_i\,(\theta) = \theta_i^{-2\lambda}e^{-\frac{1}{\theta_i}}$, i = 1 , 2
and ($0<p,q<1$) are the mixing proportions satisfy that $p+q=1$.

Let $\underline{x} = x_1, x_2, ..., x_n$ be a random sample of size n drawn from the mixture, given by (2.1). Then, the likelihood function generated by the random sample is given by

$$L_1(\theta_1,\ \theta_2,\ \lambda,\ p\,|\underline{x}) = \prod_{j=1}^{n} H(x_j)\ S(x_j),$$

(2.2)

where H(x_j) = $p\ g_1(x_j) + q\ g_1(x_j)$, j =1,2,…,n

The natural logarithm of equation (2.2) is

$$L^* = \sum_{j=1}^{n} \left[\ln\ H\ (x_j) + \ln\ \ S(x_j)\right]. \tag{2.3}$$

Differentiating (2.3) with respect to the parameters p, θ_1, θ_2 and λ partially and equating to zero gives

$$\frac{\partial L1}{\partial p} = \ \sum_{j=1}^{n} \frac{g_1(x) - g_2(x)}{H(x_j)} = 0 \tag{2.4}$$

$$\frac{L1}{\partial \theta_i} = \ \sum_{j=1}^{n} \text{W1j}\ \frac{(g_1)}{(\theta_i)}\,((\frac{x_j}{\theta_i})^l - 2) = 0 \tag{2.5}$$

$$\frac{\partial L1}{\partial \lambda} = \frac{n}{l} + 2\sum_{j=1}^{n} Ln\big(X_j\big) - \sum_{i=j}^{2}\sum_{j=1}^{n} \text{W1j}(2Ln(\theta_i) + (\frac{x_j}{\theta_i})^l\ (Ln\frac{x_j}{\theta_i}))) = 0 \tag{2.6}$$

The amount w_{ij} represents the chance that the observation leaves the module and is able to be conveyed. . $w_{1j} = \dfrac{pg_1(x_j)}{H(x_j)}$, $w_{2j} = 1 - w_{1j}$.

Therefore

$$\text{W1j} = \frac{1}{1+\exp[\ a+b(\ x_j\ )^l\ ]} \tag{2.7}$$

Where a=${\theta_1}^{-l} - {\theta_2}^{-l}$, b= Ln($\frac{q}{p}$) – 2lLn$\frac{\theta_2}{\theta_1}$

we solve the nonlinear system of (2.4), (2.5), (2.6), we contract

$$\hat{p} = \frac{1}{n}\ \sum_{j=1}^{n} w1(\ x_j) \tag{2.8}$$

$$\hat{\theta_i} = (\frac{\sum_{j=1}^{n} 2\text{W}i(x_{j)}}{\sum_{j=1}^{n} \text{W}i(x_j){x_j}^l})^{\frac{-1}{\lambda}} \quad ,\ i=1,2 \tag{2.9}$$

$$\hat{\lambda} = \frac{n}{\sum_{i=j}^{2}\sum_{j=1}^{n} \text{W1j}(2Ln(\theta_i)+(\frac{x_j}{\theta_i})^l\ (Ln\frac{x_j}{\theta_i}))\ -2LnX_j} \tag{2.10}$$

Where $\hat{w}_{ij}$ be an estimated value of w_{ij} and be gotten by substituting a, b in (2.7) with MLEs, $\hat{a}, \hat{b}$. The equations that are nonlinear (2.4), (2.5), (2.6) you can solve the nonlinear equations (2. 4), (2. 5), and (2.6) in an iterative manner, using methods like the expectation-maximization algorithm introduced by Dempster et al. (1977). In this study, my research involves using the Quasi-Newton method to directly address the nonlinear system. This approach expands on the secant method for groups of equations that are nonlinear, especially through a method called Broyden's method, (see, Broyden (1965)). In this approach, the Jacobian matrix used in Newton's method is replaced by an estimated version that gets updated with every iteration [15].

3. Optimal Discriminant Function

Let two WPMD populations Π_i , i=1,2 .The function for nonlinear discrimination

$$NL_{1o}(x) = a + b\, x^{\lambda} \tag{3.1}$$

where a , b were definite in (2.5). The likelihood of an individual, as determined by the posterior probability x of unidentified source has originated from Π_1 is given by

$$\Pr(x \in \Pi_1) = \{1 + \exp[NL_{1O}(x)]\}^{-1} .$$

we can categorize x in Π_1 if $NL_{1O}(x) < 0$, and in Π_2 if $NL_{1O}(x) \geq 0$. At the parameters of the populations Π_1 and Π_2 were all recognized, we achieve optimal $NL_{1O}(x)$ given by (3.1).

4. Estimation Discriminant Function

Overall, the characteristics of the populations are still not known. We utilize the existing data to make estimates for both the parameters and the discriminant functions. Our goal is to focus on these kinds of data:

(i) Classified sample "c": This occurs at information is got by taking samples from a combination of two populations, with the source of each observation recognized post-sampling.

(ii) Mixed sample "m": At this point, all observations in the available data go unnoticed.

4.1. Classified sample case

Let n_i original explanations existing from Π_i , where Π_i is $WPMD(\theta_i, \lambda)$, i=1,2.

The estimated nonlinear discriminant function can be expressed as

$$NL_{1c}(x) = \tilde{a} + \tilde{b}\ x^{\lambda}, \tag{4.1}$$

where $\tilde{a}$ and $\tilde{b}$ are obtained from (2.7) by replacing the parameters p, θ_1, θ_2 and λ by their MLEs based on classified sample $\tilde{p}, \breve{\theta}_1, \breve{\theta}_2$ and $\tilde{\lambda}$

$$\tilde{p} = \frac{n_1}{n} \tag{4.2}$$

$$\breve{\theta} i = \left(\frac{\sum_{j=1}^{n} n_2}{\sum_{j=1}^{n}(1/x_j{}^{l})}\right)^{\frac{-1}{\lambda}} \tag{4.3}$$

$$\tilde{\lambda} = \frac{n}{\sum_{i=j}^{2}\sum_{j=1}^{n}(2Ln(\theta_i)+\frac{x_j}{\theta_i}(Ln\frac{x_j}{\theta_i}))) - 2LnX_j} \tag{4.4}$$

where $n = n_1 + n_2$, $i = 1,2$.

4.2. Mixed sample case

The initial explanations should all originate from the combination itself. (2.1) Π_i , $i = 1,2$ The nonlinear discriminant function derived from this mixed sample is expressed as follows:

$$NL_{1c}(x) = \tilde{a} + \tilde{b}\ x^{\lambda} \tag{٤.٥}$$

since $\hat{a}$, $\hat{b}$ are MLE's achieved through substituting the parameters p, θ_1, θ_2 and λ with the solution of the nonlinear system (2.4), (2.5) , (2.6), $\hat{p}, \breve{\theta}_1, \breve{\theta}_2$ and $\hat{\lambda}$ in (2.7).

5. Definition of Errors of Misclassification

Assume we have an observation x that belongs to one of two WPMD populations $\Pi_i, i = 1,2$. The conditional probabilities of misclassifying x by the nonlinear discriminant function $NL_{1j}(x)$, $j = o,c,m$ can be expressed as follows

$$E_{1j} = \Pr[NL_{1j}(x) > 0 \,|x \in \Pi_1] = \Pr[a + \frac{b}{x} > 0 \,|x \in \Pi_1], \quad \text{if} \quad \theta_1 > \theta_2 .$$

Hence

$$E_{1j} = 1 - F(\gamma_j, \theta_1, \lambda) \text{ and } E_{2j} = F(\gamma_j, \theta_2, \lambda),$$

where γ_j has the values $\frac{-b}{a}, \frac{-\tilde{b}}{\tilde{a}}, \frac{-\hat{b}}{\hat{a}},$ for $j = o,c,m,$ respectively, and $F(\gamma_j, \theta_i, \lambda)$ be cumulative distribution function (CDF) of WPMD. The corresponding total probability of misclassification is given by

$$E_j = p\,E_{1j} + q\,E_{2j} . \tag{5.1}$$

6. Simulation Experiments and Results

We performed a series of simulation experiments to investigate the performance of $NL_m(x)$ relative to $NL_C(x)$ and $NL_O(x)$ for samples. We used the acceptance-rejection method to generate random variables from WPMD.

Theorem: (see, Rubinstein(1981)) Suppose X be a random variable distributed with the PDF $f_X(x), x \in I$, which is represented as

$$f_X(x) = C\, g(x)\, h(x),$$

where $C \geq 1,\ 0 < g(x) \leq 1$ and $h(x)$ is also a PDF let U and Y be distributed as $u(0,1)$ and $h(y)$ respectively. Then

$$f_Y(x \mid U \leq g(Y)) = f_X(x).$$

Now consider,

$$WPMD \equiv f(x;\theta,\lambda),\ f_X(x;\theta,\lambda) = \left(\frac{\lambda}{\theta}\right)\left(\frac{x}{\theta}\right)^{2\lambda-1} e^{-\left(\frac{x}{\theta}\right)^{\lambda}} = C\ g(x)\ h(x)$$

where, $C = \lambda$, g(x) = $\theta^{-3\lambda} x^{2\lambda-1} e^{x^{\lambda-1}}$ h(x) = $\frac{1}{\theta^{\lambda}} e^{-\frac{x}{\theta^{\lambda}}}$.

Generate U_1 from $u(0,1)$ using RNUN routine from IMSL .

Generate Y from h(x) from IMSL

If $U_1 \leq g(Y)$, deliver Y as the variable generated from $WPMD(\theta,\lambda)$.

Go to Step 1.

The simulation experiments are used to examine how MLEs perform for the mixed samples:

$\theta_1 = 1,\quad \theta_2 = 2.0$, $\lambda = 1.0, 3.0, 5.0$, $p = 0.25,\ 0.5$ and different samples $n = 30$, $n = 50$ and $n = 100$.

The samples are created using the following process:

-Generate U_2 from $u(0,1)$ using RNUN routine from IMSL.

- If $U_2 \leq p$, then $X \equiv WPMD(\theta_1,\lambda)$ otherwise $X \equiv WPMD(\theta_2,\lambda)$

This process is repeated n times, which leads to n_1 findings from the first part and n_2 from the another part. This produces a combined sample of size $n = n_1 + n_2$.

In Table1 the average values from the mixed samples and the parameters from the classified samples are estimated, corresponding square errors, and biases are contrasted. In general, the MLEs obtained from classified samples were found to perform better than those obtained from mixture , particularly when λ is substantial and n was on the rise.

In Table2 the likelihood of misclassifying cases related to three different discrimination methods is assessed.

In Table3 display the differences in misclassification errors $\bar{e}_m$, $\bar{e}_C$ with $\bar{e}_O$ and the standard deviations for $\bar{e}_m$, $\bar{e}_C$ (in parentheses) and their relative percentage biases for $p = 0.25,\ 0.5$. The initial entry located in every cell below $B(\bar{e}_j)$, $j = m, c$ be the value of the absolute bias from $\bar{e}_O$ standardized by the standard deviation (SD_j) of $\bar{e}_j$,

given by $B(\bar{e}_j) = \dfrac{|\bar{e}_j - \bar{e}_O|}{SD_j}$, $j = m, c$.

The second entry is the value absolute bias from $\bar{e}_O$ standardized by $\bar{e}_O$,

given by $B(\bar{e}_j) = \dfrac{|\bar{e}_j - \bar{e}_O|}{\bar{e}_O}$. On the other hand, $|B|$ represents the percentage bias to $\bar{e}_m$ from $\bar{e}_C$ standardized by $\bar{e}_C$, given by $|B| = \dfrac{|\bar{e}_m - \bar{e}_C|}{\bar{e}_C}$.

From Table 2 (a, b) We observe that the misclassification errors are not good. At λ , $d = |\mu_1 - \mu_2|$ is small. We discovered the impact of the mistakes on the performance of the classified samples $\bar{e}_C$ standardized by $\bar{e}_O$ and SD are better than the presentation of the mistakes in the mixture samples $\bar{e}_m$ .Usually when $p = 0.25$, The pair of categories processes' combined performance is better than when $p = 0.5$. As the size of samples goes up from. $n = 30$ to $n = 100$, each estimation for the two sets of parameters being analyzed improves.

The following tables summarizes the key results.

Table١ : Estimated means and mean square errors and biases , $p = 0.25$, $\theta_2 = 0.5$.

Actual values of the parameters			Estimated means, mean square errors, biases.							
			Classified				Mixture			
n	θ_1	λ	$\tilde{P}$	$\breve{\theta}_1$	$\breve{\theta}_2$	$\tilde{\lambda}$	$\hat{P}$	$\breve{\theta}_1$	$\breve{\theta}_2$	$\hat{\lambda}$
30	2.0	1	.2402	.5398	2.4051	1.0019	.3501	.7240	2.4026	.9760
			.0069	.1290	.7601	.0291	.0297	.1405	1.0306	.0193
			.0114	.0431	.3513	.0019	.0921	.2069	.3634	.0151
		3	.2440	.5902	2.1421	3.3918	.3831	.7701	2.3018	2.9379
			.0047	.1075	.5572	3.3100	.0447	.1505	.4140	1.4321
			.0120	.0701	.1293	.5402	.1311	.2481	.2828	.0441
		5	.2321	.5543	2.1748	5.1621	.3211	.6203	2.0393	5.2071
			.0050	.1072	.6245	1.7381	.0346	.0641	.5222	5.5611
			.0064	.0631	.2746	.2247	.0581	.1242	.0198	.3081
٥٠	2.0	1	.2439	.5517	2.1045	1.0254	.3491	.5404	2.4113	.8761
			.0017	.0694	.2114	.0091	.0245	.0472	1.7434	.0031
			.0049	.0562	.1062	.0167	.1081	.1305	.5123	.0124
		3	.2461	.5291	2.1415	2.8716	.3217	.6156	2.3385	2.9061
			.0041	.0801	.1752	.2441	.0349	.0588	1.0788	.2312
			.0077	.0373	.1327	.1171	.0828	.1306	.3366	.0935
		5	.2354	.4851	2.1030	5.0657	.3078	.6602	1.8149	4.7285
			.0023	.0451	.1501	2.2763	.0244	.0623	.3061	2.0171
			.0035	.0044	.1044	.0765	.0579	.1504	.0870	.1827
100	2.0	1	.2421	.5354	2.0318	1.0166	.2676	.5413	2.3081	1.0164
			.0015	.0387	.1271	.0047	.0017	.0776	.2107	.0069
			.0011	.0456	.0528	.0186	.0047	.0512	.1057	.0175
		3	.2604	.5343	2.1480	3.0216	.2578	.5382	2.1337	2.8827
			.0015	.0362	.1304	.1280	.0035	.0813	.1862	.2542
			.0019	.0311	.1470	.0148	.0078	.0382	.1337	.1172
		5	.2361	.5031	2.0710	5.2643	.2374	.4757	2.1030	5.0667
			.0016	.0262	.1291	2.5417	.0024	.0465	.1546	2.2863
			.0012	.0006	.0621	.3361	.0029	.0042	.1020	.0668

Note: In Table 1, the first row shows the estimated means, the second row represents the estimated mean square errors, and the third row displays the estimated biases.

Table2 :- Individual probabilities of misclassification., $\theta_2 = 0.5$

Actual values of the parameters				classification procedures					
				mixture		classified		optimal	
n	p	θ_1	λ	$\bar{e}_{1m}$	$\bar{e}_{2m}$	$\bar{e}_{1c}$	$\bar{e}_{2c}$	$\bar{e}_{1o}$	$\bar{e}_{2o}$
30	.25	2.0	1.0	.3104	.6652	.1021	.7998	.1232	.7968
	.5			.3217	.6172	.5911	.2899	.3354	.5624
	.25		3.0	.3513	.6263	.0629	.9182	.1203	.7968
	.5			.٢٩٣٨	.6579	.1934	.8001	.3209	.5624
	.25		5.0	.3016	.6947	.0487	.9125	.1142	.٨٧٧٨
	.5			.3615	.6256	.2997	.8021	.3351	.٦٦٣٦
٥٠	.25	2.0	1.0	.2782	.6423	.1239	.7861	.1232	.7968
	.5			.2971	.5918	.3806	.4982	.3354	.5624
	.25		3.0	.2301	.6758	.1135	.7962	.1203	.7968
	.5			.2622	.6973	.3218	.4931	.3209	.5624
	.25		5.0	.1868	.7256	.1153	.7983	.1142	.7968
	.5			.4921	.3984	.2807	.5807	.3351	.5624
100	.25	2.0	1.0	.2209	.6582	.1241	.7854	.1232	.7968
	.5			.3752	.4497	.3116	.4879	.3354	.5624
	.25		3.0	.1983	.6904	.0761	.8209	.1203	.7968
	.5			.3828	.4829	.2967	.5771	.3209	.5624
	.25		5.0	.2017	.6658	.0792	.8098	.1142	.7968
	.5			.3291	.4709	.2814	.5337	.3351	.5624

Table 3:- Total probabilities of misclassification and percentage standardized biases, $p = 0.25$, $\theta_2 = 0.5$.

Actual values of the parameters			classification procedures			relative bias to		
			mixture	completely classified	optimal	optimal		completely classified
n	θ_1	λ	$\bar{e}_m$	$\bar{e}_C$	$\bar{e}_o$	$B(\bar{e}_m)$	$B(\bar{e}_c)$	$\lvert \boldsymbol{B} \rvert$
30	2.0	1.0	.5322 (.1011)	.7071 (.0779)	.5763	5.5115 13.1889	2.1088 6.5231	7.7310
		3.0	.5207 (.1209)	.7261 (.0631)	.5746	6.1224 16.6344	4.9216 18.2257	11.5151
		5.0	.6382 (.1207)	.6261 (.0696)	.5758	3.8851 10.5621	3.7152 13.6917	8.3678
٥٠	2.0	1.0	.5295 (.1102)	.6544 (.0287)	.5763	6.4521 17.8612	2.5773 11.9972	5.5466
		3.0	.5519 (.1117)	.6816 (.0457)	.5746	2.7650 6.6231	.2751 1.1231	4.4016
		5.0	.5930 (.1242)	.7008 (.0403)	.5758	.7414 1.0123	1.4532 4.9982	3.2336
100	2.0	1.0	.5522 (.1077)	.5949 (.0213)	.5763	3.3762 9.99151	.6253 2.3987	2.9877
		3.0	.5400 (.1034)	.6059 (.0258)	.5746	4.9837 14.0126	1.1023 8.1207	6.8154
		5.0	.5455 (.1122)	.6087 (.0279)	.5758	4.1215 11.7101	1.8501 8.7219	6.2617

7. Conclusion

This study utilizes the maximum likelihood method to derive MLEs of WPMD parameters from both mixed and classified datasets. These MLEs are subsequently applied to estimate the corresponding nonlinear discriminant function for both types of samples. The approximate nonlinear discriminant function's performance is assessed and benchmarked against the optimal discriminant function through a series of Monte Carlo simulations, with evaluation metrics based on mean squared error (MSE) and total likelihood. The results indicate that the hybrid discrimination scheme outperforms the fully segmented method. Overall, leveraging total probability, we find that the mixed discriminant way demonstrates more favorable performance compared to the fully classified approach.

References

[1] K.S. Sultan, A.S. Al-Moisheer, Estimation of a discriminant function from a mixture of two inverse Weibull distributions, J. Stat. Comput. Simul., 83 (2013) 888–905.

[2] K.E. Ahmad, Z.F. Jaheen, H.S. Mohamed, Finite mixture of Burr type XII distribution and its reciprocal: properties and applications, Stat. Pap., 52 (2011) 835–845.

[3] T.J. O'Neill, Normal discriminant with unclassified observations, J. Amer. Stat. Assoc., 73 (1978) 821–826.

[4] S. Ganesalingam, G.J. McLachlan, The efficiency of a linear discriminant function based on unclassified initial samples, Biometrika, 65 (1978) 658–662.

[5] S. Ganesalingam, G.J. McLachlan, Small sample results for a linear discriminant function estimated from a mixture of normal populations, J. Stat. Comput. Simul., 9 (1979) 151–158.

[6] K. Amoh, Estimation of a discriminant function from a mixture of two inverse Gaussian distributions when the sample size is small, J. Stat. Comput. Simul., 20 (1985) 275–286.

[7] M.A.W. Mahmoud, H.M. Moustafa, Estimation of a discriminant function from a mixture of two gamma distributions when the sample size is small, *Math. Comput. Modelling*, **18** (1993) 87–95.

[8] K.E. Ahmad, Small sample results for a nonlinear discriminant function estimated from a mixture of two Burr type XII distributions, *Comput. Math. Appl.*, **28** (1994) 13–20.

[9] K.E. Ahmad, A.M. Abd-Elrahman, Updating a nonlinear discriminant function estimated from a mixture of two Weibull distributions, *Math. Comput. Modelling*, **19** (1994) 41–51.

[10] M.A.W. Mahmoud, H.M. Moustafa, Errors of misclassification associated with gamma distribution, *Math. Comput. Modelling*, **22** (1995) 105–119.

[11] K.E. Ahmad, The efficiency of a nonlinear discriminant function based on unclassified initial samples from a mixture of two Burr type XII populations, *Comput. Math. Appl.*, **30** (1995) 1–7.

[12] H.M. Moustafa, S.G. Ramadan, On MLE of a nonlinear discriminant function from a mixture of two Gompertz distributions based on small sample sizes, *J. Stat. Comput. Simul.*, **73** (2003) 867–885.

[13] K.E. Ahmad, Z.F. Jaheen, A.A. Modhesh, Estimation of a discriminant function based on small sample size from a mixture of two Gumbel distributions, *Comm. Statist. Simulation Comput.*, **39** (2010) 713–725.

[14] A.M. Abdelrahman, A better alternative to the generalized Bilal distribution, *Int. J. Reliab. Qual. Saf. Eng.*, **30** (2023) Article 350027. https://doi.org/10.1142/S0218539323500274

[15] H.M. Moustafa, S.G. Ramadan, S.A. Mohammed, Nonlinear discriminant functions for mixed random walk models, *Commun. Stat. Simul. Comput.*, **39** (2010) 735–748.

[1٦] C.G. Broyden, A class of methods for solving nonlinear simultaneous equations, *Math. Comput.*, **19** (1965) 577–593.

[1٧] R. Y. Rubinstein, *Random Variate Generation*, Wiley Series in Probability and Statistics, Wiley, 1981.

[١٨] **IMSL**, *Reference Manual*, IMSL, Houston, Texas, 1995.