Sakarya University Journal of Science, 22 (6), 1609-1616, 2018.

SAKARYA ÜNİVERSİTESİ FEN BİLİMLERİ ENSTİTÜSÜ DERGİSİ
SAKARYA UNIVERSITY JOURNAL OF SCIENCE

e-ISSN: 2147-835X
Dergi sayfası: http://www.saujs.sakarya.edu.tr

Received
10-08-2017
Accepted
20-03-2018

Doi
10.16984/saufenbilder.333936

Effects of different parameter estimators to error rate in discriminant analysis

Hayrinisa DEMİRCİ BİÇER*[1] , Cenker BİÇER[2]

ABSTRACT

Discriminant analysis is defined as a statistical technique that classifies a unit whose properties are measured, into one of the known finite numbers of populations. In this classifying process, an error occurs when the unit is classified to different population from its own population. This error is called the error rate or the probability of incorrect classification. It is desirable to minimize this error. This study focuses on determining the parameter estimation method that provides the minimum error rate, when the parameters of Weibull populations are not known. Maximum likelihood (ML), moments (MOM) and least squares (LS) methods are chosen from among parameter estimation methods. By a conducted simulation study, it is investigated that the error rate how is affected by the ML, LS and MOM estimates.

Keywords: optimal classification, Weibull distribution, error rate, discriminant analysis

1. INTRODUCTION

Discriminant analysis is a statistical method that classifies a unit to the one of the known and finite number of groups (populations) based on the measurements of unit. In this study, the number of groups is chosen as two for ease of operations.

Let Π_1 and Π_2 be two different populations each having a distribution from the same parametric class and $\underline{X}=\left(X_1,X_2,\cdots,X_p\right)'$ be a p-dimensional random vector. Also, let we denote the probability density function of the population Π_j $(j=1,2)$ by $f_j\left(\underline{x};\underline{\alpha}_j\right)$, where $\underline{\alpha}_j$ is the parameter vector of jth density.

Suppose that the measurements of the random vector $\underline{X}$ are in the p-dimensional sample space R^p. In the problem of classification, this sample space is divided into two regions such as B_1 and B_2, in which $B_1 \cup B_2 = R^p$ and $B_1 \cap B_2 = \varnothing$. Hence, If the observation vector $\underline{X}_0$ is in the region B_j, the unit $\underline{x}_0$ of $\underline{X}_0$ is assigned to B_j.

In classifying process, an error occurs when the unit is classified to different population from its own population.That is, the observation actually belongs to Π_j but it can be classified to Π_i. For the determined regions B_1 and B_2, the conditional probability of misclassifying of unit $\underline{x}_0$ from Π_j to Π_i is given by

$$P\left(\underline{x}_0 \text{ classify to } \Pi_i \mid \underline{x}_0 \in \Pi_j\right) = \int_{R_i} f_j\left(\underline{x},\underline{\alpha}_j\right) d\underline{x} = p_{ij} \qquad (1)$$

and the probability of misclassification of the observation $\underline{x}_0$ into Π_i is given by

$$P\left(\underline{x}_0 \text{ classify to } \Pi_i \text{with error}\right) = q_j p_{ij} \qquad (2)$$

* Corresponding Author
[1] Faculty of Arts and Sciences, Department of Statistics, Kırıkkale University, hdbicer@hotmail.com
[2] Faculty of Arts and Sciences, Department of Statistics, Kırıkkale University, cbicer@kku.edu.tr

Hence, the probability of total misclassification (TMC) for two groups is obtained as follows:

$$P(TMC) = \sum_{\substack{i=1 \\ i \neq j(j=1,2)}}^{2} q_j P\left(\underline{x}_0 \text{ classify to } \Pi_i \text{ when } \underline{x}_0 \in \Pi_j\right) \quad (3)$$

[1],[2].

Under the assumption that the equal cost of misclassification for each group, the classification problem can be defined as the determination of regions B_1 and B_2 with a minimum TMC. In the literature, this classification rule is known as the optimal classification rule (OCR). When the parameter vectors $\underline{\alpha}_1$ and $\underline{\alpha}_2$ are known, the OCR is given by

$$\xi : \begin{cases} \underline{x}_0 \text{ classify to } \Pi_1, \text{ if } U = \dfrac{f\left(\underline{x}_0;\underline{\alpha}_1\right)}{f\left(\underline{x}_0;\underline{\alpha}_2\right)} > k \\ \underline{x}_0 \text{ classify to } \Pi_2, \text{ otherwise} \end{cases} \quad (4)$$

where $k = \dfrac{q_2}{q_1}$.

In many real-life problems, the values of parameter vectors $\underline{\alpha}_1$ and $\underline{\alpha}_2$ are unknown. However, they can be estimated by using samples from Π_1 and Π_2 with sizes of n_1 and n_2, respectively. Let $\underline{\hat{\alpha}}_1$ and $\underline{\hat{\alpha}}_2$ be estimates of the parameter vectors $\underline{\alpha}_1$ and $\underline{\alpha}_2$, respectively. In this situation, the OCR based on the sample is given by

$$\hat{\xi} : \begin{cases} \underline{x}_0 \text{ classify to } \Pi_1, \text{ if } W = \dfrac{f\left(\underline{x}_0;\underline{\hat{\alpha}}_1\right)}{f\left(\underline{x}_0;\underline{\hat{\alpha}}_2\right)} > k \\ \underline{x}_0 \text{ classify to } \Pi_2, \text{ otherwise} \end{cases} \quad (5)$$

[3],[4].

2. ERROR RATES IN DISCRIMINANT ANALYSIS

In the literature, there are introduced several types of error rates associated with discriminant rules, such as optimal error rate, conditional actual error rate and expected actual error rate [2], [3], [5], [6],[7]. Let we assume that $\alpha_i(\xi)$, $(i=1,2)$ is the probability of misclassification from Π_i and that the parameter vectors $\underline{\alpha}_1$ and $\underline{\alpha}_2$ are known. When the classification rule given by (4) is used, the rate of misclassification $\alpha_i(\xi)$ is defined as follows:

$$\alpha_1(\xi) = P\left(U \le k \middle| \underline{X} \in \Pi_1\right) \quad (6)$$

and

$$\alpha_2(\xi) = P\left(U > k \middle| \underline{X} \in \Pi_2\right). \quad (7)$$

These two probabilities of misclassification are known as the optimal error rate. Since the Π_1 and Π_2 are randomly tagged, it is enough to calculate one of $\alpha_1(\xi)$ and $\alpha_2(\xi)$ to obtain the optimal error rate. Hence, for a unit $\underline{X}_0$ which is known to come from Π_1, the probability of misclassification to Π_2 is given by

$$\alpha(\xi) = P(U \le k). \quad (8)$$

The conditional actual error rate is defined (like in the optimal error rate) as follows, based on the classification rule given by (5), and it depends on parameter estimates

$$\alpha\left(\hat{\xi}\right) = P(W \le k). \quad (9)$$

In the literature, $\alpha\left(\hat{\xi}\right)$ is also known as the conditional error rate. By taking the expectation of (9), we can immediately write

$$E\left(\alpha\left(\hat{\xi}\right)\right) = E\left(P(W \le k)\right). \quad (10)$$

Equation (10) is known as expected actual error rate or unconditional error rate [2], [3], [5], [7], [8], [9].

When the parameters of populations are unknown, theses error rates can be found by using the distribution of the sample discriminant function generated by the parameter estimates obtained from the samples. However, it is not always easy to obtain the distribution of the sample discriminant function. For this reason, there are many estimators for the error rates when the distribution of the sample discriminant function is unknown or cannot be obtained analytically.

In this study, the resubstitution estimator introduced by [10] is considered for the estimate of the error rate. From now on, resubstitution estimator will be indicated by $\hat{\alpha}^R$. The $\hat{\alpha}^R$ is calculated as follow. Assume that samples of sizes n_1 and n_2 from groups Π_1 and Π_2, respectively, are taken.These $n_1 + n_2$ observations are reclassified by using the classification rule given by (5). Also, let α_{n_1} and α_{n_2} indicate the numbers of the misclassified observations from Π_1 to Π_2 and Π_2 to Π_1, respectively. Under these assumptions, $\hat{\alpha}^R$ is defined as follows

$$\hat{\alpha}^R = \frac{\alpha_{n_1} + \alpha_{n_2}}{n_1 + n_2}. \quad (11)$$

[8], [9],[10]

3. DISCRIMINANT ANALYSIS FOR WEIBULL POPULATIONS

In this section, some basic information about the Weibull distribution and the optimal classification rule for the Weibull distributed populations proposed by [8] and [9] are discussed.

The Weibull distribution (WE) is an important distribution family for analyzing the nonsymmetric and the positive valued data. The WE has got two parameters, a scale parameter $\beta > 0$ and a shape parameter $\theta > 0$. The probability density function (pdf) of WE with the parameters θ and β is

$$f_{WE}(x;\theta,\beta) = \theta\beta x^{\beta-1} e^{-\theta(x)^{\beta}} \;, x > 0 \quad (12)$$

and the corresponding cumulative distribution function (cdf) is

$$F_{WE}(x;\theta,\beta) = 1 - e^{-\theta(x)^{\beta}} \;, x > 0 . \quad (13)$$

Also, the hazard function of WE is given by

$$h_{WE}(x;\theta,\beta) = \theta\beta^{\theta} x^{\theta-1} \quad (14)$$

The hazard function (or the failure rate) of this distribution is increasing when the parameter β is greater than 1, constant when β equal to 1 (the exponential case), and decreasing when β is less than 1. A detailed discussion on it has been provided by [11]. In the rest of this study, for brevity, the WE with parameters θ and β will be indicated as $WE(\theta,\beta)$. Suppose that $X_1, X_2, ..., X_n$ be a random sample from $WE(\theta,\beta)$. Several parameter estimators such as maximum likelihood, method of moment and least squares for the parameters of $WE(\theta,\beta)$ are given below.

The ML estimators of the parameters θ and β, say $\hat{\theta}_{ML}$ and $\hat{\beta}_{ML}$, respectively, are given by

$$\hat{\theta}_{ML} = \frac{n}{\sum_{i=1}^{n} x_i^{\hat{\beta}_{ML}}} \quad (15)$$

and

$$\hat{\beta}_{ML} = \left(\frac{\sum_{i=1}^{n} x_i^{\hat{\beta}_{ML}} \ln x_i}{\sum_{i=1}^{n} x_i^{\hat{\beta}_{ML}}} - \frac{1}{n}\sum_{i=1}^{n} \ln x_i \right)^{-1} . \quad (16)$$

The MOM estimators of the parameters θ and β, $\hat{\theta}_{MOM}$ and $\hat{\beta}_{MOM}$, respectively, are given by

$$\bar{X} - \left(\frac{1}{\hat{\theta}_{MOM}}\right)^{\frac{1}{\hat{\beta}_{MOM}}} \Gamma\left(1 + \frac{1}{\hat{\beta}_{MOM}}\right) = 0 \quad (17)$$

and

$$\sum \frac{x_i^2}{n} - \left(\frac{1}{\hat{\theta}_{MOM}}\right)^{\frac{2}{\hat{\beta}_{MOM}}} \Gamma\left(1 + \frac{2}{\hat{\beta}_{MOM}}\right) = 0 \quad (18)$$

Both ML and MOM estimators cannot be obtained analitycally but estimates of the parameters can be numerically computed from nonlinear systems (15)-(16) and (17)-(18), respectively.

Finally, to obtain the LS estimators of parameters θ and β, say $\hat{\theta}_{LS}$ and $\hat{\beta}_{LS}$, respectively, first, the random sample $X_1, X_2, ..., X_n$ is sorted ascending order. On the other hand, By using the transformation $\eta = \theta^{1/\beta}$, the following equation can be easily written from (13).

$$\ln x_i = -\ln\eta + \frac{1}{\beta}\left(\ln\left(-\ln\left(1 - F(x_i)\right)\right)\right). \quad (19)$$

Hence, the following simple linear regression equation can be written by taking

$Y_i = \ln x_i$, $\beta_0 = -\ln\eta$, $\beta_1 = \frac{1}{\beta}$ and

$X_i = \left(\ln\left(-\ln\left(1 - F(x_i)\right)\right)\right)$

$$Y_i = \beta_0 + \beta_1 X_i. \quad (20)$$

Considering the linear regression model given by (20), $\hat{\beta}_0$ and $\hat{\beta}_1$ can be easily estimated by using

$$\hat{\beta}_1 = \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n\sum_{i=1}^{n} x_i^2 - \left(\sum x_i\right)^2} \quad (21)$$

and

$$\hat{\beta}_0 = \bar{y}_i - \hat{\beta}_1 \sum_{i=1}^{n} \frac{x_i}{n}. \quad (22)$$

Therefore, LS estimates of the parameters are obtained as $\hat{\beta}_{LS} = \frac{1}{\hat{\beta}_1}$, $\hat{\eta} = e^{-\hat{\beta}_0}$ and $\hat{\theta}_{LS} = \hat{\eta}^{\hat{\beta}_{LS}}$

Where $F(x_i)$ is unknown but it can be estimated with formula $\frac{i}{n+1}$ or $\frac{i-0.3}{n+0.4}$, see [12].

Now, we consider the classification problem for the Weibull distributed populations.

Suppose that $X_{11}, X_{12}, \ldots, X_{1n_1}$ and $X_{21}, X_{22}, \ldots, X_{2n_2}$ are two independent random samples from Π_1 and Π_2, respectively, in which Π_1 and Π_2 follow $WE(\theta_i, \beta_i)$ $i = 1,2$. Then, under the assumptions equal classification cost and equal prior probability, the OCR is given by

$$R_{WE}(s): \begin{cases} \underline{x}_0 \text{ classify into } \Pi_1, \text{if } x_0^{\hat{\beta}_1 - \hat{\beta}_2} e^{-\left[\hat{\theta}_1 x_0^{\hat{\beta}_1} - \hat{\theta}_2 x_0^{\hat{\beta}_2}\right]} > \frac{\hat{\beta}_2 \hat{\theta}_2}{\hat{\beta}_1 \hat{\theta}_1} \\ \underline{x}_0 \text{ classify into } \Pi_2, \text{otherwise} \end{cases} \quad (23)$$

where $\hat{\beta}_1$, $\hat{\beta}_2$, $\hat{\theta}_1$ and $\hat{\theta}_2$ are estimates of the parameters β_1, β_2, θ_1 and θ_2, respectively. See [8] and [9] for further information on this OCR and its derivation.

In literature, unknown values of the parameters are generally estimated by using the maximum likelihood method, because of the fact that they are easy to obtain and to calculate. In this study, the classification problem for WE populations is considered in case of using the ML, the LS and the MOM estimators instead of unknown values of the parameters and it is investigated that the error rate how is affected by these estimates.

When the parameter values are unknown, the OCR is given by two cases $(\beta_1 = \beta_2 = \beta)$ and $(\theta_1 = \theta_2 = \theta)$.

Case 1: The OCR $R_{WE}(s)$ for the case of $\beta_1 = \beta_2 = \beta$ is

$$R_{WE}(s): \begin{cases} \underline{x}_0 \text{ classify into } \Pi_1, \text{if } \bar{x}_1 < \bar{x}_2 \text{ and } \underline{x}_0 < c_1 \\ \underline{x}_0 \text{ classify into } \Pi_2, \text{if } \bar{x}_1 < \bar{x}_2 \text{ and } \underline{x}_0 \geq c_1 \end{cases} \quad (24)$$

or

$$R_{WE}(s): \begin{cases} \underline{x}_0 \text{ classify into } \Pi_1, \text{if } \bar{x}_1 \geq \bar{x}_2 \text{ and } \underline{x}_0 \geq c_1 \\ \underline{x}_0 \text{ classify into } \Pi_2, \text{if } \bar{x}_1 \geq \bar{x}_2 \text{ and } \underline{x}_0 < c_1 \end{cases}. \quad (25)$$

Where $c_1 = \left[\frac{\log\hat{\theta}_2 - \log\hat{\theta}_1}{\hat{\theta}_2 - \hat{\theta}_1}\right]^{1/\hat{\beta}}$ and $\hat{\beta} = \left(\hat{\beta}_1 + \hat{\beta}_2\right)/2$.

Case 2: The OCR $R_{WE}(s)$ for the case of $\theta_1 = \theta_2 = \theta$ is

$$R_{WE}(s): \begin{cases} \underline{x}_0 \text{ classify into } \Pi_1, \text{if } \bar{x}_1 < \bar{x}_2 \text{ and } \underline{x}_0 < c_2 \\ \underline{x}_0 \text{ classify into } \Pi_2, \text{if } \bar{x}_1 < \bar{x}_2 \text{ and } \underline{x}_0 \geq c_2 \end{cases} \quad (26)$$

or

$$R_{WE}(s): \begin{cases} \underline{x}_0 \text{ classify into } \Pi_1, \text{if } \bar{x}_1 \geq \bar{x}_2 \text{ and } \underline{x}_0 \geq c_2 \\ \underline{x}_0 \text{ classify into } \Pi_2, \text{if } \bar{x}_1 \geq \bar{x}_2 \text{ and } \underline{x}_0 < c_2 \end{cases}. \quad (27)$$

Where $c_2 = \left(\log\hat{\beta}_2 - \log\hat{\beta}_1\right)$ and $\hat{\theta} = \left(\hat{\theta}_1 + \hat{\theta}_2\right)/2$, see [9].

4. SIMULATION STUDY

In this section, some simulation studies are performed to comparatively present the error rate performance of OCR discussed in the previous section using the estimates of unknown parameters obtained by ML, LS and MOM. Two cases $\beta_1 = \beta_2 = \beta$ and $\theta_1 = \theta_2 = \theta$ given in the previous section are considered in the simulation studies. In the case 1, the parameter β was choosen as $0.5,1,3.6$. When the estimates ML, LS and MOM used instead of unknown parameters, the error rates calculated for $n_1 = n_2 = 20,40,60$ and 100 and for $\theta_1 = 0.5,2$ and 4 and at different values of θ_2. Simulated results based on 1000 Monte-Carlo simulations are tabulated in Table 1-3.

Table 1. Error rates for Case 1, $\beta = 0.5$

θ_1	θ_2	Method	n=20	n=40	n=60	n=100	θ_1	θ_2	Method	n=20	n=40	n=60	n=100
0.5	1	ML	0.36320	0.36946	0.37047	0.37245	2	4	ML	0.36305	0.36915	0.37511	0.37198
		LS	0.36500	0.37175	0.37198	0.37349			LS	0.36483	0.37084	0.37566	0.37279
		MOM	0.42080	0.42529	0.42381	0.42377			MOM	0.38313	0.38455	0.38493	0.38011
	2	ML	0.26035	0.25818	0.26142	0.26311		6	ML	0.30103	0.30259	0.30658	0.30715
		LS	0.25998	0.25760	0.26128	0.26331			LS	0.30040	0.30321	0.30655	0.30694
		MOM	0.35333	0.35549	0.36044	0.34680			MOM	0.34558	0.35696	0.36025	0.37241
	4	ML	0.17035	0.17334	0.17376	0.17391		8	ML	0.25938	0.26268	0.26378	0.26411
		LS	0.16895	0.17298	0.17358	0.17358			LS	0.25870	0.26218	0.26365	0.26390
		MOM	0.29765	0.30720	0.30095	0.29920			MOM	0.29653	0.30425	0.31174	0.31169
	8	ML	0.10683	0.10730	0.10981	0.11085		10	ML	0.22853	0.23454	0.23093	0.23106
		LS	0.10720	0.10743	0.11033	0.11068			LS	0.22760	0.23404	0.23062	0.23112
		MOM	0.28473	0.31843	0.30274	0.31232			MOM	0.24280	0.24449	0.23662	0.23260
	10	ML	0.08933	0.09465	0.09358	0.09384		20	ML	0.14803	0.14704	0.15143	0.15050
		LS	0.09070	0.09499	0.09374	0.09416			LS	0.14840	0.14709	0.15113	0.15064
		MOM	0.25093	0.25784	0.24228	0.25135			MOM	0.14918	0.14814	0.15166	0.15087
	12	ML	0.08023	0.08169	0.08122	0.08157	4	6	ML	0.40125	0.41518	0.41754	0.42201
		LS	0.08185	0.08229	0.08141	0.08196			LS	0.40590	0.41979	0.42076	0.42386
		MOM	0.22253	0.23663	0.22443	0.22755			MOM	0.43178	0.44839	0.45113	0.45667
	14	ML	0.06912	0.07256	0.07204	0.07157		8	ML	0.36078	0.37015	0.37089	0.37484
		LS	0.07100	0.07310	0.07262	0.07192			LS	0.36493	0.37300	0.37190	0.37501
		MOM	0.21878	0.23491	0.23272	0.22964			MOM	0.39268	0.40130	0.39933	0.40386
	16	ML	0.06555	0.06529	0.06703	0.06699		10	ML	0.32708	0.32991	0.33304	0.33668
		LS	0.06740	0.06630	0.06759	0.06734			LS	0.32765	0.33044	0.33357	0.33736
		MOM	0.21845	0.23196	0.23379	0.23567			MOM	0.34678	0.34266	0.34213	0.34368
	20	ML	0.05430	0.05429	0.05360	0.05571		20	ML	0.22553	0.23076	0.23112	0.23212
		LS	0.05695	0.05537	0.05463	0.05649			LS	0.22423	0.23061	0.23058	0.23190
		MOM	0.22515	0.20900	0.22004	0.20418			MOM	0.23830	0.23734	0.23740	0.23713

Table 2. Error rates for Case 1, $\beta = 1$

θ_1	θ_2	Method	n=20	n=40	n=60	n=100	θ_1	θ_2	Method	n=20	n=40	n=60	n=100
0.5	1	ML	0.36005	0.36778	0.37417	0.37389	2	4	ML	0.36368	0.37061	0.37329	0.37384
		LS	0.36325	0.37108	0.37557	0.37490			LS	0.36680	0.37130	0.37368	0.37469
		MOM	0.36613	0.37186	0.37622	0.37557			MOM	0.36828	0.37506	0.37483	0.37542
	2	ML	0.25610	0.26056	0.26231	0.26279		6	ML	0.30013	0.30330	0.30597	0.30515
		LS	0.25685	0.26075	0.26241	0.26312			LS	0.30070	0.30419	0.30571	0.30495
		MOM	0.25753	0.26185	0.26298	0.26310			MOM	0.30268	0.30519	0.30699	0.30547
	4	ML	0.17598	0.17390	0.17540	0.17350		8	ML	0.25700	0.26073	0.26174	0.26410
		LS	0.17515	0.17388	0.17525	0.17347			LS	0.25693	0.26010	0.26160	0.26412
		MOM	0.17598	0.17388	0.17568	0.17355			MOM	0.25853	0.26105	0.26197	0.26398
	8	ML	0.10955	0.11035	0.10886	0.10974		10	ML	0.22820	0.23034	0.23266	0.23206
		LS	0.10963	0.11104	0.10903	0.11018			LS	0.22855	0.22971	0.23232	0.23179
		MOM	0.10913	0.11040	0.10866	0.10982			MOM	0.22933	0.23034	0.23255	0.23208
	10	ML	0.09225	0.09505	0.09379	0.09405		20	ML	0.14900	0.15179	0.15013	0.15078
		LS	0.09368	0.09560	0.09435	0.09426			LS	0.14890	0.15188	0.15006	0.15069
		MOM	0.09230	0.09500	0.09392	0.09417			MOM	0.19008	0.18458	0.17601	0.16585
	12	ML	0.07840	0.08123	0.08131	0.08250	4	6	ML	0.40180	0.41466	0.42130	0.42241
		LS	0.07988	0.08171	0.08165	0.08322			LS	0.40753	0.41805	0.42492	0.42463
		MOM	0.07915	0.08140	0.08134	0.08259			MOM	0.40610	0.41989	0.42572	0.42562
	14	ML	0.06635	0.07264	0.07183	0.07349		8	ML	0.36188	0.36844	0.37336	0.37319
		LS	0.06962	0.07414	0.07253	0.07380			LS	0.36455	0.36984	0.37485	0.37383
		MOM	0.07190	0.07345	0.07233	0.07361			MOM	0.36488	0.37097	0.37531	0.37504
	16	ML	0.06130	0.06594	0.06590	0.06602		10	ML	0.33110	0.33101	0.33573	0.33575
		LS	0.06300	0.06736	0.06656	0.06631			LS	0.33180	0.33134	0.33597	0.33574
		MOM	0.07020	0.07055	0.06697	0.06699			MOM	0.33450	0.33240	0.33782	0.33639
	20	ML	0.05155	0.05381	0.05645	0.05589		20	ML	0.22815	0.22843	0.23363	0.23316
		LS	0.05395	0.05536	0.05726	0.05631			LS	0.22740	0.22830	0.23366	0.23301
		MOM	0.09735	0.09934	0.09378	0.07454			MOM	0.25820	0.25749	0.25083	0.24563

Table 3. Errror rates for Case 1, $\beta = 3.6$

θ_1	θ_2	Method	n=20	n=40	n=60	n=100
0.5	1	ML	0.36265	0.36725	0.36924	0.37327
		LS	0.36460	0.36830	0.37106	0.37396
		MOM	0.36285	0.36664	0.36947	0.37263
	2	ML	0.25948	0.26269	0.26338	0.26244
		LS	0.25773	0.26219	0.26316	0.26248
		MOM	0.25898	0.26249	0.26320	0.26245
	4	ML	0.17435	0.17114	0.17331	0.17464
		LS	0.17390	0.17093	0.17316	0.17470
		MOM	0.17428	0.17090	0.17334	0.17467
	8	ML	0.10598	0.10831	0.10924	0.10965
		LS	0.10658	0.10886	0.10928	0.10963
		MOM	0.10573	0.10845	0.10922	0.10963
	10	ML	0.09033	0.09209	0.09384	0.09349
		LS	0.09118	0.09376	0.09422	0.09381
		MOM	0.09060	0.09234	0.09392	0.09355
	12	ML	0.07740	0.08074	0.08270	0.08210
		LS	0.07918	0.08119	0.08308	0.08218
		MOM	0.07798	0.08080	0.08284	0.08218
	14	ML	0.06850	0.07198	0.07258	0.07365
		LS	0.07000	0.07185	0.07334	0.07416
		MOM	0.06897	0.07153	0.07265	0.07363
	16	ML	0.06472	0.06595	0.06492	0.06576
		LS	0.06637	0.06676	0.06569	0.06643
		MOM	0.06477	0.06608	0.06501	0.06595
	20	ML	0.05217	0.05373	0.05541	0.05561
		LS	0.05460	0.05484	0.05619	0.05633
		MOM	0.05292	0.05409	0.05539	0.05575

θ_1	θ_2	Method	n=20	n=40	n=60	n=100
2	4	ML	0.36025	0.37093	0.36977	0.37434
		LS	0.36670	0.37279	0.37086	0.37489
		MOM	0.36055	0.37011	0.36914	0.37421
	6	ML	0.29755	0.30448	0.30619	0.30640
		LS	0.29670	0.30470	0.30676	0.30638
		MOM	0.29723	0.30449	0.30599	0.30624
	8	ML	0.25760	0.26056	0.26177	0.26046
		LS	0.25745	0.26099	0.26138	0.26066
		MOM	0.25715	0.26061	0.26148	0.26045
	10	ML	0.23245	0.23229	0.22940	0.23156
		LS	0.23120	0.23168	0.22878	0.23183
		MOM	0.23170	0.23220	0.22934	0.23162
	20	ML	0.14783	0.14993	0.15041	0.15033
		LS	0.14780	0.14964	0.15027	0.15027
		MOM	0.14813	0.14973	0.15030	0.15031
4	6	ML	0.40007	0.41291	0.41840	0.42318
		LS	0.40540	0.41670	0.42074	0.42525
		MOM	0.39850	0.41239	0.41765	0.42278
	8	ML	0.35970	0.36914	0.37276	0.37305
		LS	0.36328	0.37018	0.37411	0.37421
		MOM	0.35933	0.36881	0.37265	0.37313
	10	ML	0.32913	0.33314	0.33415	0.33644
		LS	0.33013	0.33355	0.33423	0.33664
		MOM	0.32923	0.33256	0.33428	0.33661
	20	ML	0.22788	0.22754	0.23349	0.23208
		LS	0.22755	0.22726	0.23326	0.23180
		MOM	0.22773	0.22750	0.23345	0.23204

In the case of 2, as similar to case of 1, $\theta_1 = \theta_2 = \theta$ was choosen as 0.5, 1 and 3.6. The error rates calculated for $n_1 = n_2 = 20, 40, 60$ and 100 and for $\beta_1 = 0.5, 2$ and 4 and at different values of β_2. Simulated results based on 1000 Monte-Carlo simulation are compatible with Table 4-6.

Table 4: Error rates for Case 2, $\theta = 0.5$

β_1	β_2	Method	n=20	n=40	n=60	n=100
.5	1	ML	0.31513	0.31666	0.32199	0.32292
		LS	0.31675	0.31841	0.32364	0.32337
		MOM	0.39233	0.39506	0.39439	0.38608
	2	ML	0.18550	0.19228	0.19281	0.19339
		LS	0.18890	0.19359	0.19407	0.19438
		MOM	0.29990	0.30461	0.30807	0.31167
	4	ML	0.10715	0.10979	0.11151	0.11100
		LS	0.10813	0.11171	0.11222	0.11188
		MOM	0.23680	0.25583	0.25870	0.26139
	8	ML	0.05987	0.06340	0.06282	0.06326
		LS	0.06070	0.06451	0.06333	0.06369
		MOM	0.21680	0.23880	0.23730	0.22728
	10	ML	0.04842	0.05128	0.05292	0.05245
		LS	0.04940	0.05174	0.05353	0.05295
		MOM	0.20930	0.21186	0.22623	0.21329
	12	ML	0.04387	0.04274	0.04548	0.04560
		LS	0.04405	0.04293	0.04565	0.04599
		MOM	0.20003	0.21520	0.22078	0.21472
	14	ML	0.03997	0.04007	0.03954	0.03964
		LS	0.03980	0.04021	0.03986	0.04000
		MOM	0.21265	0.20933	0.21515	0.22244
	16	ML	0.03160	0.03418	0.03475	0.03532
		LS	0.03172	0.03469	0.03498	0.03555
		MOM	0.19830	0.21331	0.20146	0.21852
	20	ML	0.02762	0.02761	0.02878	0.02896
		LS	0.02755	0.02776	0.02866	0.02901
		MOM	0.19865	0.21020	0.19699	0.20729

β_1	β_2	Method	n=20	n=40	n=60	n=100
2	4	ML	0.31083	0.32019	0.31993	0.32190
		LS	0.31463	0.32234	0.32118	0.32311
		MOM	0.31093	0.32013	0.31922	0.32190
	6	ML	0.23550	0.23848	0.23616	0.24133
		LS	0.23828	0.24019	0.23733	0.24267
		MOM	0.23555	0.23873	0.23630	0.24176
	8	ML	0.18575	0.19384	0.19351	0.19391
		LS	0.18883	0.19531	0.19485	0.19443
		MOM	0.18708	0.19444	0.19407	0.19414
	10	ML	0.15595	0.15919	0.16153	0.16345
		LS	0.15920	0.16125	0.16229	0.16413
		MOM	0.15785	0.16028	0.16175	0.16364
	20	ML	0.08695	0.09173	0.09400	0.09369
		LS	0.08905	0.09328	0.09457	0.09441
		MOM	0.08890	0.09285	0.09445	0.09411
4	6	ML	0.37785	0.38453	0.38832	0.38838
		LS	0.38130	0.38810	0.38996	0.38929
		MOM	0.37825	0.38534	0.38817	0.38851
	8	ML	0.31313	0.32059	0.32139	0.32177
		LS	0.31518	0.32129	0.32289	0.32197
		MOM	0.31268	0.32045	0.32204	0.32210
	10	ML	0.26450	0.26965	0.27476	0.27513
		LS	0.26703	0.27171	0.27563	0.27616
		MOM	0.26528	0.27066	0.27480	0.27551
	20	ML	0.15583	0.16078	0.16113	0.16313
		LS	0.15880	0.16226	0.16247	0.16364
		MOM	0.15778	0.16199	0.16215	0.16336

Table 5: Error rates for Case 2, $\theta = 1$

β_1	β_2	Method	n=20	n=40	n=60	n=100	β_1	β_2	Method	n=20	n=40	n=60	n=100
0.5	1	ML	0.33605	0.33758	0.33938	0.34378	2	4	ML	0.33005	0.34254	0.34242	0.34378
		LS	0.33670	0.33785	0.34132	0.34485			LS	0.33413	0.34441	0.34311	0.34439
		MOM	0.34603	0.34434	0.34256	0.34618			MOM	0.32970	0.34265	0.34224	0.34387
	2	ML	0.21085	0.21756	0.21621	0.21768		6	ML	0.25795	0.26440	0.26342	0.26367
		LS	0.21423	0.21965	0.21770	0.21867			LS	0.26308	0.26628	0.26486	0.26422
		MOM	0.22018	0.22448	0.22228	0.22176			MOM	0.25970	0.26508	0.26356	0.26354
	4	ML	0.12478	0.12813	0.12917	0.12865		8	ML	0.20830	0.21391	0.21772	0.21792
		LS	0.12815	0.13086	0.13028	0.12934			LS	0.21210	0.21608	0.21943	0.21869
		MOM	0.13365	0.13524	0.13334	0.13231			MOM	0.21050	0.21485	0.21791	0.21794
	8	ML	0.06920	0.07358	0.07423	0.07466		10	ML	0.17713	0.18346	0.18536	0.18555
		LS	0.07060	0.07473	0.07513	0.07520			LS	0.18158	0.18645	0.18686	0.18670
		MOM	0.07818	0.08306	0.08066	0.07829			MOM	0.18045	0.18485	0.18620	0.18615
	10	ML	0.06072	0.06109	0.06316	0.06193		20	ML	0.10388	0.10846	0.10940	0.11031
		LS	0.06260	0.06173	0.06377	0.06239			LS	0.10610	0.10990	0.11097	0.11132
		MOM	0.07295	0.06740	0.06926	0.06672			MOM	0.10563	0.10950	0.11049	0.11084
	12	ML	0.05105	0.05184	0.05362	0.05342	4	6	ML	0.38528	0.39671	0.40001	0.40371
		LS	0.05222	0.05307	0.05447	0.05378			LS	0.39148	0.39999	0.40104	0.40494
		MOM	0.06105	0.06115	0.06015	0.05690			MOM	0.38645	0.39766	0.40016	0.40366
	14	ML	0.04582	0.04671	0.04800	0.04750		8	ML	0.33335	0.34303	0.34218	0.34409
		LS	0.04607	0.04728	0.04817	0.04799			LS	0.33610	0.34573	0.34337	0.34387
		MOM	0.05945	0.05508	0.05232	0.05389			MOM	0.33463	0.34311	0.34219	0.34403
	16	ML	0.03905	0.04020	0.04131	0.04211		10	ML	0.29163	0.29715	0.29779	0.29746
		LS	0.03917	0.04074	0.04152	0.04250			LS	0.29583	0.29913	0.29910	0.29868
		MOM	0.05130	0.04994	0.04881	0.04691			MOM	0.29420	0.29790	0.29881	0.29800
	20	ML	0.03315	0.03381	0.03487	0.03485		20	ML	0.17750	0.18320	0.18428	0.18764
		LS	0.03320	0.03451	0.03526	0.03509			LS	0.18133	0.18593	0.18573	0.18846
		MOM	0.04737	0.04594	0.04369	0.04018			MOM	0.18045	0.18540	0.18494	0.18818

Table 6: Error rates for Case 2, $\theta = 3.6$

β_1	β_2	Method	n=20	n=40	n=60	n=100	β_1	β_2	Method	n=20	n=40	n=60	n=100
0.5	1	ML	0.28888	0.29654	0.29532	0.29628	2	4	ML	0.28943	0.29419	0.29736	0.29617
		LS	0.29298	0.30040	0.29825	0.29773			LS	0.29448	0.29774	0.30001	0.29773
		MOM	0.29770	0.30234	0.29828	0.30007			MOM	0.29115	0.29469	0.29739	0.29653
	2	ML	0.14678	0.15398	0.15441	0.15445		6	ML	0.20075	0.20168	0.20514	0.20497
		LS	0.15083	0.15666	0.15619	0.15553			LS	0.20538	0.20439	0.20690	0.20635
		MOM	0.15388	0.15676	0.15880	0.15661			MOM	0.20213	0.20271	0.20558	0.20541
	4	ML	0.06970	0.07290	0.07157	0.07333		8	ML	0.15020	0.15273	0.15353	0.15521
		LS	0.07060	0.07409	0.07232	0.07370			LS	0.15403	0.15541	0.15494	0.15612
		MOM	0.08065	0.07700	0.07259	0.07403			MOM	0.15180	0.15320	0.15428	0.15534
	8	ML	0.03302	0.03344	0.03411	0.03394		10	ML	0.11775	0.12011	0.12128	0.12082
		LS	0.03320	0.03354	0.03446	0.03428			LS	0.12093	0.12244	0.12265	0.12159
		MOM	0.03832	0.03730	0.03558	0.03559			MOM	0.11935	0.12163	0.12192	0.12121
	10	ML	0.02590	0.02520	0.02585	0.02687		20	ML	0.05387	0.05690	0.05731	0.05698
		LS	0.02550	0.02536	0.02598	0.02696			LS	0.05440	0.05735	0.05798	0.05757
		MOM	0.03382	0.02917	0.02871	0.02676			MOM	0.05402	0.05726	0.05785	0.05735
	12	ML	0.02075	0.02165	0.02196	0.02253	4	6	ML	0.36143	0.37055	0.37341	0.37540
		LS	0.02090	0.02140	0.02213	0.02269			LS	0.36865	0.37433	0.37653	0.37764
		MOM	0.03152	0.02416	0.02437	0.02409			MOM	0.36328	0.37141	0.37488	0.37548
	14	ML	0.01700	0.01824	0.01928	0.01996		8	ML	0.28970	0.29195	0.29794	0.29815
		LS	0.01715	0.01837	0.01950	0.02000			LS	0.29498	0.29558	0.30023	0.29991
		MOM	0.02615	0.02145	0.02233	0.01992			MOM	0.29280	0.29378	0.29913	0.29892
	16	ML	0.01520	0.01655	0.01652	0.01643		10	ML	0.24300	0.24120	0.24231	0.24542
		LS	0.01543	0.01679	0.01653	0.01640			LS	0.24760	0.24421	0.24452	0.24658
		MOM	0.02425	0.01892	0.01952	0.01632			MOM	0.24565	0.24286	0.24346	0.24627
	20	ML	0.01225	0.01285	0.01309	0.01341		20	ML	0.11785	0.11869	0.12319	0.12211
		LS	0.01233	0.01284	0.01298	0.01335			LS	0.12038	0.12074	0.12485	0.12281
		MOM	0.01915	0.01581	0.01538	0.01478			MOM	0.11965	0.12014	0.12455	0.12266

According to simulation results given by Tables 1-6, for both cases, we can say that the classification performance of the method increases when the unknown parameter values draw away from each other. Namely, the error rate decreases. In another saying, if one of the populations is overlapping the other, the error rate increases, otherwise, the error rate decreases. This is an expected case because the shapes of the pdf's of the overlapping distributions are almost the same. This situation can be seen in Figures 1-2.

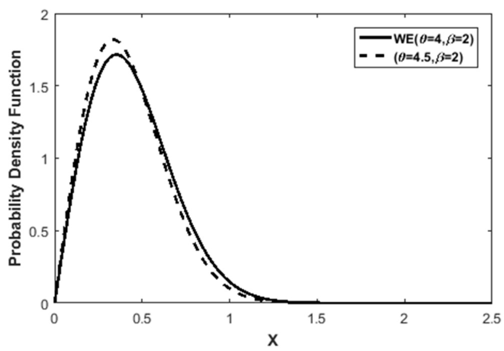

Figure 1. Overlapping distributions

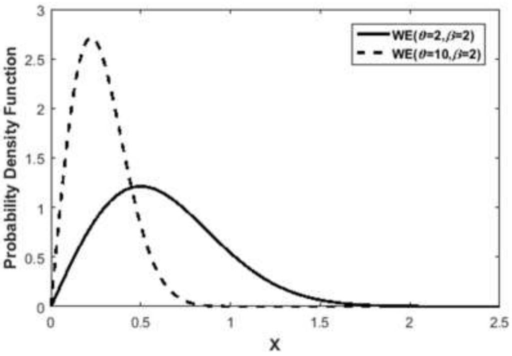

Figure 2. Non-overlapping distributions

However, even if the population parameters are close to each other, it can be said that the OCR works quite well. Furthermore, it can be seen from Tables 1-6 that the ML estimators provide the best classification performance in all cases. Besides, it is concluded from Table 1-6 that the LS estimators provide nearly same classification performance with the ML estimators. Since the LS estimators can be easily calculated for the Weibull parameters, their use instead of the unknown parameter values in the classification process may provide an advantage for researchers.

5. CONCLUSION

In this study, the problem of discriminant analysis for Weibull distributed populations is investigated considering the optimal classifier proposed by [8] and [9]. Also, when the values of the population parameters are unknown, a series of simulation studies are carried out with the aim to determine the parameter estimator which makes the error rate to be minimum. In accordance with this purpose, we employe the three estimators frequently used in the literature such as ML, LS, and MOM. In the numerical study, two basic cases are considered, such as $\beta_1 = \beta_2 = \beta$ and $\theta_1 = \theta_2 = \theta$. Each case is evaluated at different values of the parameters and at different sample sizes. The results of the numerical studies show that the OCR works well in all cases, even in the overlapping distributions. The findings show that the minimum error rate for the OCR is obtained when the ML estimates are used instead of the unknown parameter values. Therefore, based on the simulation study results, to obtain the minimum error rate in the discriminant analysis between Weibull distributed populations, It can be said that the use of ML estimates is appropriate instead of the unknown values of the distribution parameters.

REFERENCES

[1] P. A. Lachenburch, "Discriminant Analysis", *New York, Hafner*, 1975.

[2] G. A. F. Seber, "Multivariate Observations", *New York, John Wiley & Sons. Inc*, 1984.

[3] T. W. Anderson, "Introduction to Multivariate Statistical Analysis", *Wiley, New York*, 1984.

[4] B. L. Welch, "Note on Discriminant Functions", *Biometrika*, 31, 218–220, 1939.

[5] P. A. Lachenburch and R. Mickey, "Estimation of Error Rates in Discriminant Analysis", *Technometrics*, 10, 1–11, 1968.

[6] S. M. Snapin," An Evaluation of Smooted Error Rate Estimators in Discriminant Analysis", *Institute of Statistics Miemeo Series*, Chapel Hill, 1983.

[7] M. Hills, "Allocation Rules and Their Error Rates", *J. Roy. Stat. Soc. B*, 28, 1-31, 1966.

[8] H. D. Biçer, "Discriminant analiysis in censored data: Weibull distribution case", *Doctoral dissertation, University of Ankara*, 2011.

[9] H. D. Biçer, C. Atakan and C. Biçer, "İki parametreli weibull dağılımına sahip kitlelerde diskriminant analizi", *NWSA Physical sciences*, *6*(4), 124-133, 2011.

[10] C. A. B. Smith, "Some Examples of Discrimination", *Annals of Eugenics*, 18, 272-282, 1978.

[11] N. L. Johnson, S. Kotz and N. Balakrishnan, "Continuous univariate distributions", *New York, Wiley*, 1995.

[12] M. Tiryakioğlu and D. Hudak, "Unbiased estimates of the Weibull parameters by the linear regression method", *Journal of Materials Science*, 43, 1914-1919, 2008.