Jour. Ind. Soc. Ag. Statistics
54(2), 2001 :209–220

On the Performance of Two Sample Linear Discriminant Function

B. Singh
Indian Veterinary Research Insititute, Izatnagar 243122
(Received : June, 1999)

SUMMARY

Approximate expressions for moments and the probability of misclassification (PMC) are derived for two sample linear discriminant function (SLDF). The mean and variance of SLDF and PMC using both population linear discriminant function (PLDF) and SLDF are also obtained through simulated samples from two multivariate normal populations for examining the performance of SLDF and the validity of approximate theoretical results for practical applications. The numerical results reveal that PLDF under estimates the mean, variance and PMC for SLDF. The approximate expressions for SLDF provide good results for mean and PMC for all values of Δ^2 (Mahalanobis distance) and for variance for low and moderate values of Δ^2.

Key Words : Mahalanobis distance, Moments, Population LDF, Probability of misclassification, Sample LDF.

1. Introduction

Fisher's linear discriminant function is the popular technique in the field of discriminant analysis. An excellent account of this procedure can be found, for example, in Anderson [1] and McLachlan [6]. The linear discriminant function yields optimal results in the sense of smallest probability of misclassification (PMC) when parameters are known. The use of population linear discriminant function (PLDF) may not be justified in the same way when parameters are not known. Indeed except for asymptotic optimality and in special circumstances no finite sample optimality property has yet been found (Das Gupta [3] and Friedman [4]). To investigate the performance of two sample linear discriminant function (SLDF) one needs the sampling distribution of Anderson's classification statistic (W). The exact distribution of W was derived by Sitgreaves [9] but the expression was too complicated to be used, numerically. A method for computation of the cumulative distribution function of W by simulation was discussed by Teichroew and Sitgreaves [10] but actual simulation was not done due to the then low speed of computers.

In this paper, we derive the approximate moments and whence the sampling distribution of W and PMC for two group SLDF. We also obtain the numerical values through simulated samples from two multivariate normal populations for certain apriori values of parameters to study the performance of SLDF and the validity of theoretical results for practical applications.

2. *Population Linear Discriminant Function*

Let X be a random observation from a multivariate normal population. If population parameters are known the classification statistic is defined as

$$U = X'\Sigma^{-1}(\mu_1 - \mu_2) - \left(\frac{1}{2}\right)(\mu_1 + \mu_2)'\,\Sigma^{-1}(\mu_1 - \mu_2) \tag{2.1}$$

When X is distributed as N (μ_1, Σ), U is distributed normally with mean $\frac{\Delta^2}{2}$ and variance Δ^2. Similarly, when X is distributed as $N(\mu_2, \Sigma)$, U is distributed normally with mean $\left(-\frac{\Delta^2}{2}\right)$ and variance Δ^2. The Mahalanobis distance (Δ^2) between two multivariate populations is defined as

$$\Delta^2 = (\mu_1 - \mu_2)'\,\Sigma^{-1}(\mu_1 - \mu_2) \tag{2.2}$$

The probability of misclassification for PLDF is defined as $\Phi\,(-\Delta/2)$ for population π_1 and $[1 - \Phi\,(\Delta/2)]$ for population π_2.

3. *Sample Linear Discriminant Function*

In most applications the parameters are not known but are estimated from samples one from each population. Suppose that we have a sample $x_\alpha^{(1)}, (\alpha = 1, 2, \ldots, N_1)$, from population π_1 with distribution $N(\mu_1, \Sigma)$ and a sample $x_\alpha^{(2)}, (\alpha = 1, 2, \ldots, N_2)$, from population π_2 with distribution $N(\mu_2, \Sigma)$. These are taken as training samples to obtain estimates of μ_1, μ_2 and Σ. The estimates are

$$\bar{x}_1 = \left(\frac{1}{N_1}\right)\sum_{\alpha=1}^{N_1} x_\alpha^{(1)}, \quad \bar{x}_2 = \left(\frac{1}{N_2}\right)\sum_{\alpha=1}^{N_2} x_\alpha^{(2)}$$ for μ_1 and μ_2 respectively.

$$S=\left(\frac{1}{n}\right)\left[\sum_{\alpha=1}^{N_1}\left(x_\alpha^{(1)}-\bar{x}_1\right)\left(x_\alpha^{(1)}-\bar{x}_1\right)'+\sum_{\alpha=1}^{N_2}\left(x_\alpha^{(2)}-\bar{x}_2\right)\left(x_\alpha^{(2)}-\bar{x}_2\right)'\right] \text{ for } \Sigma$$

$n=(N_1+N_2-2)$

and classification statistic is defined as

$$W=X'S^{-1}(\bar{x}_1-\bar{x}_2)-\left(\frac{1}{2}\right)(\bar{x}_1+\bar{x}_2)'S^{-1}(\bar{x}_1-\bar{x}_2) \qquad (3.1)$$

3.1 Moments

We write W in (3.1) as

$$W=u'S^{-1}v \qquad (3.2)$$

where $u=(\bar{x}_1-\bar{x}_2)$, $v=X-\left(\frac{1}{2}\right)(\bar{x}_1+\bar{x}_2)$ and S is the pooled covariance matrix.

Suppose $X\in\pi_1$, then

$$u\sim N\left[\mu_1-\mu_2,\left(N_1^{-1}+N_2^{-1}\right)\Sigma\right] \text{ and } v\sim N\left[\frac{(\mu_1-\mu_2)}{2},\left(1+(4N_1)^{-1}+(4N_2)^{-1}\right)\Sigma\right]$$

Let $u_1=\sqrt{\left[\frac{N_1N_2}{(N_1+N_2)}\right]}\,u$ and $v_1=\sqrt{\left[\frac{4N_1N_2}{(N_1+N_2+4N_1N_2)}\right]}\,v$

Then

$$u_1\sim N\left[(\mu_1-\mu_2)\sqrt{\left\{\frac{N_1N_2}{N_1+N_2}\right\}},\Sigma\right],\ v_1\sim N\left[(\mu_1-\mu_2)\sqrt{\left\{\frac{N_1N_2}{(N_1+N_2+4N_1N_2)}\right\}},\Sigma\right]$$

$$W=k\left[(u_1+v_1)'S^{-1}(u_1+v_1)-(u_1-v_1)'S^{-1}(u_1-v_1)\right] \qquad (3.3)$$

where $k=\left(\frac{1}{8N_1N_2}\right)\sqrt{[(N_1+N_2)(N_1+N_2+4N_1N_2)]}$

Note that (u_1+v_1) and (u_1-v_1) are independently normally distributed (Moran [7]) as $(u_1+v_1)\sim N(\delta_1,k_1\Sigma)$ and $(u_1-v_1)\sim N(\delta_2,k_2\Sigma)$, where

$$\delta_1 = \left[(\mu_1 - \mu_2)\sqrt{N_1 N_2}\left\{(N_1 + N_2)^{-1/2} + (N_1 + N_2 + 4N_1 N_2)^{-1/2}\right\}\right]$$

$$k_1 = 2\left[1 + \frac{(N_1 - N_2)}{\{(N_1 + N_2)(N_1 + N_2 + 4N_1 N_2)\}^{1/2}}\right]$$

$$\delta_2 = \left[(\mu_1 - \mu_2)\sqrt{N_1 N_2}\left\{(N_1 + N_2)^{-1/2} - (N_1 + N_2 + 4N_1 N_2)^{-1/2}\right\}\right]$$

$$k_2 = 2\left[1 - \frac{(N_1 - N_2)}{\{(N_1 + N_2)(N_1 + N_2 + 4N_1 N_2)\}^{1/2}}\right]$$

Let $t_1 = (u_1 + v_1)k_1^{-1/2}$ and $t_2 = (u_1 - v_1)k_2^{-1/2}$

Then one writes

$$W = k\left[k_1 t_1' S^{-1} t_1 - k_2 t_2' S^{-1} t_2\right] \tag{3.4}$$

where t_1 and t_2 are independently distributed as

$$t_1 \sim N\left(\frac{\delta_1}{\sqrt{k_1}}, \Sigma\right) \text{and } t_2 \sim N\left(\frac{\delta_2}{\sqrt{k_2}}, \Sigma\right)$$

Now, by using the theorem (5.2.2) of Anderson [1], we write the classification statistic W as

$$W = kk_1 T_1^2 - kk_2 T_2^2 \tag{3.5}$$

where $T_1^2 \sim \left[\frac{np}{(n-p+1)}\right] F_{p,n-p+1}(\Delta_1^2)$ and $T_2^2 \sim \left[\frac{np}{(n-p+1)}\right] F_{p,n-p+1}(\Delta_2^2)$ with

$F_{a,b}(\Delta_i^2)$ as non central F variates and $\Delta_i^2 = \left(\frac{1}{k_i}\right)\delta_i' \Sigma^{-1} \delta_i, i = 1, 2,$ that is,

$$\Delta_1^2 = \left(\frac{N_1 N_2}{k_1}\right)\left[(N_1 + N_2)^{-1/2} + (N_1 + N_2 + 4N_1 N_2)^{-1/2}\right]^2 \Delta^2 \text{ and}$$

$$\Delta_2^2 = \left(\frac{N_1 N_2}{k_2}\right)\left[(N_1 + N_2)^{-1/2} - (N_1 + N_2 + 4N_1 N_2)^{-1/2}\right]^2 \Delta^2$$

A similar representation of W as a function of the elements of two 2 × 2 independent Wishart matrices has been provided by Bowker [2]. The exact distribution of W (3.5) is difficult to obtain since T_1^2 and T_2^2 are not independent variates. Their denominators are interrelated with identical distribution except when p = 1, in that case these Hotelling T^2 variates are independent with same denominator.

Here, we assume the same denominator for all values of p and derive the first two moments of W and whence its approximate sampling distribution to obtain PMC for SLDF. We examine the validity of these approximate results by comparing with corresponding results based on simulated samples from two multivariate normal populations. Although this comparison may not give exact answer but it frequently gives results that are sufficiently accurate for most practical purposes.

With the assumption of same denominator we write W as

$$W = \frac{(U_1 - U_2)}{V} \tag{3.6}$$

where $U_1 \sim g_1 \chi_p^2(\Delta_1^2)$, $U_2 \sim g_2 \chi_p^2(\Delta_2^2)$ and $V \sim \chi_{n-p+1}^2$ are independent chi-square variates. The constants $g_i = nkk_i$, i = 1, 2 are defined as

$$g_1 = \left(\frac{n}{4N_1N_2}\right)\left[N_1 - N_2 + \{(N_1+N_2)(N_1+N_2+4N_1N_2)\}^{1/2}\right] \text{ and}$$

$$g_2 = \left(\frac{n}{4N_1N_2}\right)\left[N_2 - N_1 + \{(N_1+N_2)(N_1+N_2+4N_1N_2)\}^{1/2}\right]$$

By using the expression for r-th raw moment of a non-central chi-square variate (Johnson and Kotz [5]) we obtain

$$E(U_1 - U_2) = nk\left[(k_1 - k_2)p + k_1\Delta_1^2 - k_2\Delta_2^2\right] \text{ and}$$

$$E(U_1 - U_2)^2 = n^2k^2\left[p(p+2)\left(k_1^2 + k_2^2\right) + 2(p+2)\left(k_1^2\Delta_1^2 + k_2^2\Delta_2^2\right) + \left(k_1^2\Delta_1^4 + k_2^2\Delta_2^4\right) - 2k_1k_2\left(p+\Delta_1^2\right)\left(p+\Delta_2^2\right)\right]$$

The r-th raw moment of W is expressed as

$$\mu_r'(W) = \mu_r'(U_1 - U_2)\mu_r'(V)$$

where $\mu_r'(V) = \left[\frac{\Gamma\{(q/2)-r\}}{\{2^r\Gamma(q/2)\}}\right]$ and $q = n - p + 1$

This gives $E\left(\frac{1}{V}\right)=(n-p-1)^{-1}$ and $E\left(\frac{1}{V^2}\right)=\left[(n-p-1)(n-p-3)\right]^{-1}$

The expressions for $X \in \pi_2$ can be obtained by interchanging δ_1^2 and δ_2^2. Finally we obtain

$$E(W)=\begin{bmatrix}\left[\left(\frac{n}{(n-p-1)}\right)\left[\left(\frac{(N_1-N_2)}{2N_1N_2}\right)p+\left(\frac{\Delta^2}{2}\right)\right], \text{ when } X\in\pi_1\right] \\ \left[\left(\frac{n}{(n-p-1)}\right)\left[\left(\frac{(N_1-N_2)}{2N_1N_2}\right)p-\left(\frac{\Delta^2}{2}\right)\right], \text{ when } X\in\pi_2\right]\end{bmatrix}$$

$$E\left(W^2\right)=E\left(U_1-U_2\right)^2 E\left(V^{-2}\right) \text{ and } Var(W)=E(W)^2-\left[E(W)\right]^2 \qquad (3.7)$$

3.2 Probability of Misclassification

By assuming that U_1 and U_2 are approximately distributed as $a\chi_b^2$ and $c\chi_d^2$ respectively, where the constants a,b,c and d are easily obtained by using the Patnaik's two moments approximation (Patnaik [8]). The moments of U_1 and U_2 are given in section (3.1).

The region of classification for population π_1 is $W \geq 0$. The probability of misclassifying X to π_2 when it actually belongs to π_1, is given by

$$\begin{aligned} p(2|1) &= P(W \leq 0|\pi_1) \\ &= P(U_1 \leq U_2|\pi_1) \\ &= I_{w_0}(b/2, d/2) \end{aligned} \qquad (3.8)$$

where I_x (a, b) is the value of incomplete beta and $w_0=\frac{c}{(a+c)}$.

Similarly, the region of classification for π_2 is $W \leq 0$ and the probability of misclassifying X to π_1 when it actually belongs to π_2 given by

$$P(1|2)=I_w\left(\frac{d^*}{2},\frac{b^*}{2}\right) \qquad (3.9)$$

where $w=\frac{a^*}{\left(a^*+c^*\right)}$ and a^*, b^*, c^*, d^* are the corresponding constants when $X \in \pi_2$.

Table 1. Probability of Misclassification for PLDF

m_1	Σ_i	ρ	π_i	$N_1 = 20, N_2 = 20$				$N_1 = 25, N_2 = 15$				Δ^2	
m_2				$p = 3$		$p = 5$		$p = 3$		$p = 5$		$p = 3$	$p = 5$
m_3				T	S	T	S	T	S	T	S		
1	Σ_1	.3	π_1	.295	.293	.289	.285	.295	.291	.289	.286	1.16	1.23
0			π_2	.295	.311	.289	.304	.295	.315	.289	.311		
0		.5	π_1	.270	.262	.259	.248	.270	.267	.259	.258	1.50	1.67
			π_2	.270	.285	.259	.271	.270	.290	.259	.277		
		.7	π_1	.221	.219	.205	.203	.221	.228	.205	.203	2.36	2.72
			π_2	.221	.239	.205	.202	.221	.231	.205	.220		
1		.3	π_1	.138	.150	.123	.114	.138	.160	.123	.120	4.73	5.39
2			π_2	.138	.164	.123	.142	.138	.156	.123	.119		
0		.5	π_1	.121	.129	.093	.086	.121	.139	.093	.080	5.50	7.00
			π_2	.121	.146	.093	.109	.121	.144	.093	.088		
		.7	π_1	.081	.087	.047	.046	.081	.089	.047	.036	7.92	11.1
			π_2	.081	.098	.047	.054	.081	.090	.047	.041		
1		.3	π_1	.055	.057	.036	.031	.055	.059	.036	.034	10.4	12.9
2			π_2	.055	.045	.036	.051	.055	.043	.036	.044		
3		.5	π_1	.057	.060	.023	.018	.057	.063	.022	.019	10.0	16.0
			π_2	.057	.045	.023	.033	.057	.046	.022	.032		
		.7	π_1	.044	.043	.006	.004	.044	.046	.006	.002	11.7	24.6
			π_2	.044	.039	.006	.013	.044	.038	.006	.008		
1	Σ_2	.3	π_1	.301	.299	.301	.292	.301	.297	.301	.301	1.10	1.10
0			π_2	.301	.293	.301	.289	.301	.320	.301	.329		
0		.5	π_1	.282	.268	.282	.265	.282	.284	.282	.277	1.33	1.33
			π_2	.282	.281	.282	.274	.282	.299	.282	.300		
		.7	π_1	.241	.229	.241	.225	.241	.250	.241	.248	1.96	1.96
			π_2	.241	.250	.241	.228	.241	.251	.241	.260		
1		.3	π_1	.142	.148	.142	.137	.142	.164	.142	.155	4.57	4.57
2			π_2	.142	.171	.142	.159	.142	.160	.142	.145		
0		.5	π_1	.125	.131	.125	.118	.125	.143	.125	.141	5.33	5.33
			π_2	.125	.151	.125	.131	.125	.147	.125	.129		
		.7	π_1	.077	.082	.077	.062	.077	.085	.077	.073	8.16	8.16
			π_2	.077	.095	.077	.086	.077	.086	.077	.077		
1		.3	π_1	.052	.056	.046	.031	.052	.059	.046	.043	10.5	11.4
2			π_2	.052	.045	.046	.053	.052	.043	.046	.052		
3		.5	π_1	.063	.070	.039	.030	.063	.069	.039	.033	9.33	12.3
			π_2	.063	.055	.039	.048	.063	.049	.039	.047		
		.7	π_1	.063	.072	.017	.008	.063	.071	.017	.008	9.33	17.9
			π_2	.063	.058	.017	.026	.063	.049	.017	.016		

Table 2. Probability of Misclassification for SLDF

m_1	Σ_i	ρ	π_i	$N_1 = 20, N_2 = 20$				$N_1 = 25, N_2 = 15$				Δ^2	
m_2				p = 3		p = 5		p = 3		p = 5		p = 3	p = 5
m_3				T	S	T	S	T	S	T	S		
1	Σ_1	.3	π_1	.364	.302	.334	.314	.301	.313	.305	.324	1.16	1.23
0			π_2	.364	.345	.334	.347	.321	.343	.335	.374		
0		.5	π_1	.303	.264	.286	.270	.282	.285	.277	.284	1.50	1.67
			π_2	.303	.312	.286	.296	.295	.315	.300	.327		
		.7	π_1	.232	.241	.226	.211	.228	.239	.216	.223	2.36	2.72
			π_2	.232	.260	.226	.229	.236	.259	.234	.256		
1		.3	π_1	.144	.160	.133	.145	.144	.181	.130	.126	4.73	5.39
2			π_2	.144	.174	.133	.161	.148	.182	.140	.151		
0		.5	π_1	.126	.132	.102	.109	.126	.157	.099	.093	5.50	7.00
			π_2	.126	.156	.102	.134	.131	.163	.107	.115		
		.7	π_1	.118	.092	.053	.058	.083	.102	.052	.046	7.92	11.1
			π_2	.118	.105	.053	.070	.086	.108	.055	.070		
1		.3	π_1	.059	.068	.041	.048	.056	.065	.039	.046	10.4	12.9
2			π_2	.059	.051	.041	.065	.059	.048	.042	.060		
3		.5	π_1	.061	.071	.026	.034	.061	.069	.025	.037	10.0	16.0
			π_2	.061	.053	.026	.056	.063	.050	.025	.049		
		.7	π_1	.047	.050	.008	.008	.046	.051	.008	.008	11.7	24.6
			π_2	.047	.045	.008	.022	.048	.036	.008	.020		
1	Σ_2	.3	π_1	.314	.299	.331	.292	.309	.319	.319	.327	1.10	1.10
0			π_2	.314	.348	.331	.289	.324	.345	.350	.380		
0		.5	π_1	.301	.291	.310	.306	.291	.306	.295	.302	1.33	1.33
			π_2	.301	.319	.310	.307	.309	.338	.327	.356		
		.7	π_1	.256	.251	.267	.263	.250	.264	.254	.277	1.96	1.96
			π_2	.256	.273	.267	.261	.257	.284	.280	.321		
1		.3	π_1	.145	.163	.155	.156	.145	.187	.150	.171	4.57	4.57
2			π_2	.145	.178	.155	.189	.153	.185	.163	.171		
0		.5	π_1	.129	.137	.135	.138	.128	.163	.131	.146	5.33	5.33
			π_2	.129	.158	.135	.167	.132	.170	.141	.159		
		.7	π_1	.082	.089	.085	.092	.081	.099	.081	.091	8.16	8.16
			π_2	.082	.101	.085	.101	.085	.104	.082	.093		
1		.3	π_1	.057	.065	.050	.051	.055	.065	.049	.061	10.5	11.4
2			π_2	.057	.050	.050	.067	.058	.048	.052	.071		
3		.5	π_1	.068	.078	.017	.050	.066	.071	.043	.045	9.33	12.3
			π_2	.068	.062	.017	.067	.069	.055	.044	.070		
		.7	π_1	.068	.078	.020	.024	.066	.073	.019	.026	9.33	17.9
			π_2	.068	.063	.020	.037	.069	.059	.020	.026		

Table 3. Mean Values of SLDF

m_1	Σ_i	ρ	π_i	$N_1 = 20, N_2 = 20$				$N_1 = 25, N_2 = 15$				Δ^2	
m_2				p = 3		p = 5		p = 3		p = 5		p = 3	p = 5
m_3				T	S	T	S	T	S	T	S		
1	Σ_1	.3	π_1	0.649	0.672	0.733	0.798	0.694	0.721	0.821	0.805	1.161	1.234
0			π_2	−0.649	−0.616	−0.733	−0.733	−0.604	−0.551	−0.653	−0.580		
0		.5	π_1	0.838	0.869	0.990	1.061	0.883	0.915	1.069	1.079	1.500	1.667
			π_2	−0.838	−0.797	−0.990	−0.981	−0.794	−0.739	−0.911	−0.821		
		.7	π_1	1.319	1.363	1.614	1.694	1.364	1.404	1.694	1.730	2.361	2.719
			π_2	−1.319	−1.267	−1.614	−1.596	−1.275	−1.222	−1.535	−1.423		
1		.3	π_1	2.644	2.556	3.200	3.414	2.689	2.697	3.279	3.369	4.732	5.390
2			π_2	−2.644	−2.568	−3.200	−3.090	−2.600	−2.465	−3.121	−3.132		
0		.5	π_1	3.074	2.972	4.156	4.431	3.118	3.123	4.235	4.388	5.500	7.000
			π_2	−3.074	−2.966	−4.156	−4.014	−3.029	−2.864	−4.077	−4.077		
		.7	π_1	4.424	4.299	6.614	6.998	4.469	4.469	6.694	6.942	7.917	11.140
			π_2	−4.424	−4.274	−6.614	−6.441	−4.380	−4.171	−6.535	−6.533		
1		.3	π_1	5.788	5.865	7.711	7.816	5.832	5.885	7.790	7.868	10.357	12.987
2			π_2	−5.788	−5.984	−7.711	−7.584	−5.743	−5.838	−7.632	−7.715		
3		.5	π_1	5.588	5.664	9.500	9.646	5.633	5.684	9.579	9.742	10.000	16.000
			π_2	−5.588	−5.780	−9.500	−9.286	−5.543	−5.635	−9.421	−9.493		
		.7	π_1	6.520	6.611	14.583	14.788	6.564	6.622	14.662	14.953	11.667	24.561
			π_2	−6.520	−6.739	−14.580	−14.260	−6.475	−6.586	−14.500	−14.590		
1	Σ_2	.3	π_1	0.614	0.645	0.653	0.694	0.659	0.687	0.732	0.687	1.099	1.099
0			π_2	−0.614	−0.597	−0.653	−0.687	−0.569	−0.531	−0.573	−0.507		
0		.5	π_1	0.745	0.787	0.791	0.823	0.790	0.822	0.871	0.822	1.333	1.333
			π_2	−0.745	−0.728	−0.791	−0.834	−0.700	−0.669	−0.712	−0.632		
		.7	π_1	1.096	1.153	1.164	1.182	1.141	1.179	1.244	1.188	1.961	1.961
			π_2	−1.096	−1.078	−1.164	−1.220	−1.051	−1.033	−1.085	−0.983		
1		.3	π_1	2.554	2.471	2.714	2.880	2.599	2.609	2.793	2.812	4.571	4.571
2			π_2	−2.554	−2.486	−2.714	−2.670	−2.510	−2.383	−2.635	−2.674		
0		.5	π_1	2.980	2.881	3.166	3.366	3.025	3.030	3.246	3.295	5.333	5.333
			π_2	−2.980	−2.877	−3.166	−3.144	−2.935	−2.776	−3.087	−3.127		
		.7	π_1	4.558	4.432	4.843	5.115	4.603	4.603	4.922	5.030	8.157	8.157
			π_2	−4.558	−4.405	−4.843	−4.833	−4.514	−4.301	−4.764	−4.821		
1		.3	π_1	5.871	5.950	6.766	6.876	5.916	5.969	6.846	6.843	10.506	11.396
2			π_2	−5.871	−6.070	−6.766	−6.774	−5.826	−5.923	−6.687	−6.822		
3		.5	π_1	5.216	5.283	7.323	7.480	5.260	5.309	7.402	7.452	9.333	12.333
			π_2	−5.216	−5.394	−7.323	−7.264	−5.171	−5.252	−7.244	−7.371		
		.7	π_1	5.216	5.276	10.676	10.907	5.260	5.309	10.755	10.880	9.333	17.980
			π_2	−5.216	−5.388	−10.68	−10.52	−5.171	−5.247	−10.60	−10.73		

Table 4. Variance of SLDF

m_1	Σ_i	ρ	π_i	$N_1 = 20, N_2 = 20$				$N_1 = 25, N_2 = 15$				Δ^2	
m_2				p = 3		p = 5		p = 3		p = 5		p = 3	p = 5
m_3				T	S	T	S	T	S	T	S		
1	Σ_1	.3	π_1	2.359	2.173	3.355	3.218	2.410	2.341	3.442	3.127	1.161	1.234
0			π_2	2.359	2.240	3.355	3.221	2.366	2.234	3.385	3.418		
0		.5	π_1	2.977	2.735	4.319	4.033	3.301	2.954	4.410	3.934	1.500	1.667
			π_2	2.977	2.784	4.319	4.052	2.973	2.761	4.330	4.134		
		.7	π_1	4.664	4.182	6.878	6.022	4.718	4.485	6.965	6.032	2.361	2.719
			π_2	4.664	4.210	6.878	6.094	4.627	4.092	6.839	5.984		
1		.3	π_1	10.19	7.408	14.75	11.42	12.20	9.368	14.75	10.66	4.732	5.390
2			π_2	10.19	8.040	14.75	11.60	10.02	8.349	14.50	11.95		
0		.5	π_1	12.26	8.495	20.44	15.36	12.24	11.06	20.35	13.37	5.500	7.000
			π_2	12.26	9.333	20.44	15.26	12.02	9.848	20.03	15.22		
		.7	π_1	19.64	12.39	38.38	25.60	19.48	16.16	37.85	21.57	7.917	11.14
			π_2	19.64	13.64	38.38	25.06	19.18	14.40	37.34	24.48		
1		.3	π_1	28.46	19.82	47.91	28.08	28.08	17.88	47.11	28.59	10.36	12.99
2			π_2	28.46	19.58	47.91	30.93	27.69	17.06	46.51	31.65		
3		.5	π_1	27.08	19.08	65.48	35.62	26.74	17.19	64.12	35.93	10.00	16.00
			π_2	27.08	18.85	65.48	39.81	26.36	16.42	63.38	39.03		
		.7	π_1	33.75	22.62	129.1	59.46	33.24	20.38	125.4	59.55	11.67	24.56
			π_2	33.75	22.34	129.1	66.07	32.79	19.45	124.3	62.58		
1	Σ_2	.3	π_1	2.249	2.135	3.065	2.859	2.300	2.169	3.151	2.904	1.099	1.099
0			π_2	2.249	2.147	3.065	2.947	2.257	2.121	3.100	3.228		
0		.5	π_1	2.670	2.556	3.571	3.218	2.722	2.579	3.659	3.297	1.333	1.333
			π_2	2.670	2.511	3.571	3.406	2.671	2.488	3.597	3.631		
		.7	π_1	3.859	3.652	5.004	4.227	3.914	3.691	5.095	4.418	1.961	1.961
			π_2	3.859	3.511	5.004	4.638	3.839	3.470	5.004	4.753		
1		.3	π_1	9.773	7.197	12.13	9.706	9.787	8.997	12.17	9.814	4.571	4.571
2			π_2	9.773	7.779	12.13	9.739	9.611	8.028	11.96	9.921		
0		.5	π_1	11.79	8.243	14.56	11.84	11.78	10.71	14.57	11.18	5.333	5.333
			π_2	11.79	9.047	14.56	11.49	11.58	9.533	14.32	11.15		
		.7	π_1	20.45	12.79	24.98	18.97	20.27	16.67	24.79	17.07	8.157	8.157
			π_2	20.45	14.08	24.98	17.95	19.96	14.86	24.41	16.88		
1		.3	π_1	29.04	20.14	39.64	23.93	28.65	18.15	39.08	24.62	10.51	11.40
2			π_2	29.04	19.89	39.64	25.86	28.25	17.32	38.56	27.64		
3		.5	π_1	24.59	17.68	44.43	25.80	24.32	15.96	43.73	26.32	9.333	12.33
			π_2	24.59	17.48	44.43	28.56	23.96	15.23	43.16	29.81		
		.7	π_1	24.59	17.66	78.39	39.46	24.32	16.01	76.60	39.24	9.333	17.98
			π_2	24.59	17.46	78.39	43.38	23.96	15.22	75.76	44.09		

4. Simulation

Here, we generate $N_1 + N_2 + 2$ observations from two p-variate normal populations, $N_1 + 1$ from π_1 and $N_2 + 1$ from π_2, with certain apriori values of parameters. The first $N_1 + N_2$ p-variate observations are used to obtain SLDF. The PLDF is obtained by using the population mean vectors and the dispersion matrix. The remaining two observations, one from each population were used to get numerical value for PLDF and SLDF for each group, separately. This process was repeated 1000 times to get one value for each of PMC for PLDF, PMC for SLDF, mean for SLDF and variance for SLDF for each group, separately, for one fixed set of parameters p, N_1 and N_2 . The corresponding theoretical values are also computed from the formulae given in previous sections for comparison with simulated results. The numerical results presented in Tables 1-4 are for the following apriori values :

$$\Sigma_1 = (\sigma_{ij}),\ \sigma_{ii} = 1 \text{ and } \sigma_{ij} = \rho,\ i \neq j \text{ and } \Sigma_2 = (\sigma_{ij}),\ \sigma_{ij} = \rho^{|i-j|},\ \forall\ i \text{ and } j$$

$$p = 3, 5, (1)\ N_1 = N_2 = 20, (2)\ N_1 = 25, N_2 = 15, \rho = 0.3, 0.5, 0.7$$

$$\mu_1 = (m_1, m_2, m_3, 0, \ldots, 0) \text{ and } \mu_2 = (0, 0, 0, \ldots, 0),\ m_1 = 1,\ m_2 = 0, 2,\ m_3 = 0, 3$$

5. Numerical Results

The numerical results in Table 1 reveal that the simulated (S) values of PMC for PLDF agree with the corresponding theoretical (T) values for all values of Mahalanobis Distance (Δ^2) between the two multivariate normal populations. This agreement supports the simulated results for the study. The results in Tables 2-4 indicate that the values for mean, variance and PMC are more for SLDF than PLDF. The mean and PMC of SLDF obtained from the approximate expressions in section 3 are close to simulated values presented in Tables 2 & 3 for all values of Δ^2 and overestimate the variance of SLDF in comparison to the simulated values in Table-4 particularly for the highly distant populations. This implies that the theoretical expressions give good approximation for mean and PMC of SLDF for all values of Δ^2 and the variance of SLDF for low and moderate values of Δ^2 . Thus, the sampling distribution of W considered here is good approximation for all practical purposes for low and moderate values of Δ^2.

REFERENCES

[1] Anderson, T.W. (1984). An *Introduction to Multivariate Statistical Analysis* (second edition). John Wiley & Sons, New York.

[2] Bowker, A.H. (1961). A representation of Hotelling's and Anderson's classification statistic W in terms of sample statistics. *Studies in Item Analysis and Prediction* (H. Solomon, Ed), Stanford University Press.

[3] Das Gupta, S. (1965). Optimum classification rules for classification into two multivariate normal population. *Ann. Math. Statist.*, **36,** 1174-1184.

[4] Friedman, J.H. (1989). Regularized discriminant analysis. *J. Amer. Statist. Assoc.*, **84**, 165-175.

[5] Johnson, N.L. and Kotz, S. (1970). *Continuous Univariate Distributions.* John Wiley & Sons, New York.

[6] McLachlan, G.J. (1992). *Discriminant Analysis and Statistical Pattern Recognition.* John Wiley & Sons, New York.

[7] Moran, M.A. (1975). On the expectation of errors of allocation associated with linear discriminant function. *Biometrika,* **62**, 141-148.

[8] Patnaik, P.B. (1949). The non-central χ^2 and F distributions and their applications. *Biometrika,* **36**, 202-232.

[9] Sitgreaves, R. (1961). Some results on the distribution of the W classification statistic. *Studies in Item Analysis and Prediction* (H. Solomon, Ed), Stanford University Press.

[10] Teichroew, T. and Sitgreaves, R. (1961). Computation of an empirical sampling distribution for the W-classification statistic. *Studies in Item Analysis and Prediction* (H. Solomon, Ed), Stanford University Press.