Far East Journal of Mathematical Sciences (FJMS)
Volume 34, Issue 3, 2009, Pages 393-404
Published Online: October 26, 2009
This paper is available online at http://www.pphmj.com

ASYMPTOTIC DISTRIBUTIONS OF PROFILE TEST STATISTICS UNDER ELLIPTICAL POPULATIONS

YOSIHITO MARUYAMA

Department of Mathematical Sciences
Osaka Prefecture University
1-1 Gakuen, Naka-ku Sakai-City
Osaka 599-8531, Japan

Abstract

In this paper, we obtain asymptotic expansions for the distributions of test statistics in profile analysis under elliptical populations. Those asymptotic expansions are given up to the order n^{-1}, where n is the sample size.

1. Introduction

Let $\mathbf{x}$ be a p-dimensional random vector according to an elliptical distribution with parameters $\boldsymbol{\mu}(p \times 1)$ and $\boldsymbol{\Lambda}(p \times p)$ whose probability density function is defined by $f(\mathbf{x}) = K_p |\boldsymbol{\Lambda}|^{-1/2} g((\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Lambda}^{-1}(\mathbf{x} - \boldsymbol{\mu}))$, for some non-negative function g, where K_p is the normalizing constant and $\boldsymbol{\Lambda}$ is positive definite. Notice that $\mathbf{x}^T$ is a transpose of $\mathbf{x}$. In Kelker [4], the characteristic function of the vector $\mathbf{x}$ is derived as

$$\phi(\vartheta) = \mathrm{E}[e^{i\vartheta^T \mathbf{x}}] = e^{i\vartheta^T \boldsymbol{\mu}} \Psi(\vartheta^T \boldsymbol{\Lambda} \vartheta),$$

for some function Ψ, where $i := \sqrt{-1}$, from which it follows that the mean vector and the covariance matrix, if they exist, are $\mathrm{E}[\mathbf{x}] = \boldsymbol{\mu}$ and $\mathrm{Cov}[\mathbf{x}] = -2\Psi'(0)\boldsymbol{\Lambda}$

2000 Mathematics Subject Classification: 62H10, 62E20.

Keywords and phrases: asymptotic expansion, elliptical distribution, perturbation method.

Received May 21, 2009

$=: \mathbf{\Sigma}$, respectively. We also define the kurtosis parameter by $\eta := \Psi^{(2)}(0)(\Psi'(0))^{-2} - 1$. When $\mathbf{x}$ is distributed as the elliptical distribution with parameters $\boldsymbol{\mu}$, $\mathbf{\Lambda}$, and η, we write $\mathbf{x} \sim \mathcal{E}_p(\boldsymbol{\mu}, \mathbf{\Lambda}, \eta)$.

Let $\mathbf{x}_1, \ldots, \mathbf{x}_{n_x}$ be independently and identically distributed (hereafter referred to as i.i.d.) as $\mathcal{E}_p(\boldsymbol{\mu}_x, \mathbf{\Lambda}, \eta_x)$. The sample mean vector and sample covariance matrix are, respectively,

$$\overline{\mathbf{x}} := \frac{1}{n_x}\sum_{j=1}^{n_x}\mathbf{x}_j,$$

$$\mathbf{S}_x := \frac{1}{n_x - 1}\sum_{j=1}^{n_x}(\mathbf{x}_j - \overline{\mathbf{x}})(\mathbf{x}_j - \overline{\mathbf{x}})^T.$$

Suppose that $\mathbf{y}_1, \ldots, \mathbf{y}_{n_y}$ are independently i.i.d. as $\mathcal{E}_p(\boldsymbol{\mu}_y, \mathbf{\Lambda}, \eta_y)$. The sufficient statistics are $\overline{\mathbf{y}}$ and $\mathbf{S}_y$, where

$$\overline{\mathbf{y}} := \frac{1}{n_y}\sum_{j=1}^{n_y}\mathbf{y}_j,$$

$$\mathbf{S}_y := \frac{1}{n_y - 1}\sum_{j=1}^{n_y}(\mathbf{y}_j - \overline{\mathbf{y}})(\mathbf{y}_j - \overline{\mathbf{y}})^T.$$

The pooled estimate $\mathbf{S}_{\mathrm{pl}}$ of $\mathbf{\Sigma}$ is given by $\mathbf{S}_{\mathrm{pl}} := (n_x + n_y - 2)^{-1}((n_x - 1)\mathbf{S}_x + (n_y - 1)\mathbf{S}_y)$.

Profile analysis pertains to situations where a battery of p treatments (tests, questions, and so forth) is administered to two or more groups of subjects. It is assumed that the responses for the different groups are independent of one another, but all responses must be expressed in similar units. Ordinarily we might pose the question: Are the population mean vectors the same? In profile analysis, the question of equality of mean vectors is divided into several specific possibilities.

When we consider the population means $\boldsymbol{\mu}_x = (\mu_{x1}, \ldots, \mu_{xp})^T$ representing the average responses to p treatments for the group, a plot of these means connected by straight lines (broken-line graph) is the profile for the population. Profiles can be constructed for each population (group). We shall concentrate on two groups.

First, the test statistic for the parallelism hypothesis $\mathbb{H}_1 : \mu_{xi-1} - \mu_{xi} = \mu_{yi-1} - \mu_{yi}$ for $i = 2, \ldots, p$ is defined by

$$T_1 := \frac{n_x + n_y - p}{(n_x + n_y - 2)(p-1)} \left(\frac{|\mathbf{S}_{\mathrm{pl}} + \mathbf{W}|}{|\mathbf{S}_{\mathrm{pl}}|} \frac{\mathbf{1}^T (\mathbf{S}_{\mathrm{pl}} + \mathbf{W})^{-1} \mathbf{1}}{\mathbf{1}^T \mathbf{S}_{\mathrm{pl}}^{-1} \mathbf{1}} - 1 \right), \tag{1}$$

where $\mathbf{W} := n_x n_y (n_x + n_y)^{-1} (\bar{\mathbf{x}} - \bar{\mathbf{y}})(\bar{\mathbf{x}} - \bar{\mathbf{y}})^T$ and $\mathbf{1}$ is a $p \times 1$ vector with all elements equal to 1. For two-sample problem, if $\mathbb{H}_1$ is true, we test the level hypothesis $\mathbb{H}_2 : \sum_{i=1}^{p} \mu_{xi} = \sum_{i=1}^{p} \mu_{yi}$ or the flatness hypothesis $\mathbb{H}_3 : \mu_{x1} = \cdots = \mu_{xp}$ and $\mu_{y1} = \cdots = \mu_{yp}$. The test statistic for testing hypothesis $\mathbb{H}_2$ is

$$T_2 := \frac{n_x n_y}{n_x + n_y} \mathbf{1}^T (\bar{\mathbf{x}} - \bar{\mathbf{y}}) (\mathbf{1}^T \mathbf{S}_{\mathrm{pl}} \mathbf{1})^{-1} (\bar{\mathbf{x}} - \bar{\mathbf{y}})^T \mathbf{1}. \tag{2}$$

When $\mathbb{H}_1$ and $\mathbb{H}_2$ are tenable, the common mean vector is estimated using all $n_x + n_y$ observations by $(n_x + n_y)^{-1}(n_x \bar{\mathbf{x}} + n_y \bar{\mathbf{y}})$, and the test criterion for $\mathbb{H}_3$ is given by substitution of

$$\mathbf{W} = \frac{(n_x \bar{\mathbf{x}} + n_y \bar{\mathbf{y}})(n_x \bar{\mathbf{x}} + n_y \bar{\mathbf{y}})^T}{n_x + n_y},$$

into (1), which is denoted by T_3. These test statistics are essentially Hotelling type statistic in normal populations. The distribution of Hotelling statistic has been studied under some other underlying distributions including a mixture of two normal populations and a general population by Siotani [7, 8], Srivastava and Awan [9], Kabe and Gupta [3], Fujikoshi [2], etc.

It is easily seen that the limiting distribution of them as the size of sample tends to infinity is a chi-square distribution. The purpose of this paper is to obtain asymptotic expansions for distributions of test statistics on above three hypotheses tests under elliptical populations.

2. Theoretical Results

We consider asymptotic expansions for distributions of three test statistics T_1, T_2 and T_3. It is well known that these are reduced to Hotelling type statistic (F-statistic) when populations have multivariate normal distributions. However, under the class of elliptical distributions, those statistics are no longer as F-statistic. In this section, therefore, we derive asymptotic expansions for T_1, T_2 and T_3 by perturbation method. Hereafter write $n := n_x + n_y$ and suppose $n_x \geq n_y = \varepsilon n_x$ $(0 < \varepsilon \leq 1)$.

Theorem 1. *The distribution of* T_1 *is expanded as*

$$\Pr(T_1 \leq t) = G_{p-1}(t) + \frac{1}{n}\sum_{k=0}^{2} a_k G_{p+2k-1}(t) + o(n^{-1}),$$

where $G_\nu(t)$ *is the distribution function of the central chi-square distribution with* ν *degrees of freedom*, *and each coefficient is given by*

$$a_0 = -\frac{(p-1)^2}{4} + \frac{p^2-1}{8}\left(\frac{(\varepsilon^2 - 2\varepsilon - 1)\eta_x}{1+\varepsilon} - \frac{(\varepsilon^2 + 2\varepsilon - 1)\eta_y}{(1+\varepsilon)\varepsilon}\right),$$

$$a_1 = -\frac{p-1}{2} - \frac{p^2-1}{4}\left(\frac{(\varepsilon^2 - 4\varepsilon + 1)\eta_x}{1+\varepsilon} + \frac{(\varepsilon^2 - 4\varepsilon + 1)\eta_y}{(1+\varepsilon)\varepsilon}\right),$$

$$a_2 = \frac{p^2-1}{4} + \frac{p^2-1}{8}\left(\frac{(\varepsilon^2 - 6\varepsilon + 3)\eta_x}{1+\varepsilon} + \frac{(3\varepsilon^2 - 6\varepsilon + 1)\eta_y}{(1+\varepsilon)\varepsilon}\right).$$

Corollary 2. *The distribution of* T_2 *is expanded as*

$$\Pr(T_2 \leq t) = G_1(t) + \frac{1}{n}\sum_{k=0}^{2} b_k G_{1+2k}(t) + o(n^{-1}),$$

where each coefficient is given by substituting $p = 2$ *into* a_k $(k = 0, 1, 2)$, *that is*,

$$b_0 = -\frac{1}{4} + \frac{3}{8}\left(\frac{(\varepsilon^2 - 2\varepsilon - 1)\eta_x}{1+\varepsilon} - \frac{(\varepsilon^2 + 2\varepsilon - 1)\eta_y}{(1+\varepsilon)\varepsilon}\right),$$

$$b_1 = -\frac{1}{2} - \frac{3}{4}\left(\frac{(\varepsilon^2 - 4\varepsilon + 1)\eta_x}{1+\varepsilon} + \frac{(\varepsilon^2 - 4\varepsilon + 1)\eta_y}{(1+\varepsilon)\varepsilon}\right),$$

$$b_2 = \frac{3}{4} + \frac{3}{8}\left(\frac{(\varepsilon^2 - 6\varepsilon + 3)\eta_x}{1+\varepsilon} + \frac{(3\varepsilon^2 - 6\varepsilon + 1)\eta_y}{(1+\varepsilon)\varepsilon}\right).$$

Theorem 3. *The distribution of* T_3 *is expanded as*

$$\Pr(T_3 \le t) = G_{p-1}(t) + \frac{1}{n}\sum_{k=0}^{2} c_k G_{p+2k-1}(t) + o(n^{-1}),$$

where

$$c_0 = -\frac{(p-1)^2}{4} - \frac{(p^2-1)(\eta_x + \varepsilon\eta_y)}{4(1+\varepsilon)},$$

$$c_1 = -\frac{p-1}{2} + \frac{(p^2-1)(\eta_x + \varepsilon\eta_y)}{2(1+\varepsilon)},$$

$$c_2 = \frac{p^2-1}{4} - \frac{(p^2-1)(\eta_x + \varepsilon\eta_y)}{4(1+\varepsilon)}.$$

The Cornish-Fisher type expansion gives the approximate percentile of the distribution of T_1 as a simple function of the chi-square percentile. The upper α percentile of T_1 is

$$\chi^2_{p-1,\alpha} - \frac{2\chi^2_{p-1,\alpha}}{(p-1)n}\left(a_0 - \frac{a_2\chi^2_{p-1,\alpha}}{p+1}\right), \tag{3}$$

where $\chi^2_{v,\alpha}$ is the upper α percentile of the central chi-square distribution with v degrees of freedom. Furthermore, the upper 100α percentage point of T_2 is

$$\chi^2_{1,\alpha} - \frac{2\chi^2_{1,\alpha}}{n}\left(b_0 - \frac{b_2\chi^2_{1,\alpha}}{3}\right), \tag{4}$$

and that of T_3 is given by

$$\chi^2_{p-1,\alpha} - \frac{2\chi^2_{p-1,\alpha}}{(p-1)n}\left(c_0 - \frac{c_2\chi^2_{p-1,\alpha}}{p+1}\right). \tag{5}$$

Remark 1. The approximate upper percentile of T_2 given by (4) is not related to p. When $n_x = n_y$, the coefficients a_k coincide with c_k for $k = 0, 1, 2$ and hence the critical point given by (3) is equal to that by (5).

Remark 2. If two populations have elliptical distributions with the same kurtosis parameter, that is $\eta_x = \eta_y =: \eta$, then coefficients in Theorem 1 are the same as the result of Maruyama [6], that is

$$a_0 = -\frac{(p-1)^2}{4} + \frac{(p^2-1)(\varepsilon^2 - 4\varepsilon + 1)\eta}{8\varepsilon},$$

$$a_1 = -\frac{p-1}{2} - \frac{(p^2-1)(\varepsilon^2 - 4\varepsilon + 1)\eta}{4\varepsilon},$$

$$a_2 = \frac{p^2-1}{4} + \frac{(p^2-1)(\varepsilon^2 - 4\varepsilon + 1)\eta}{8\varepsilon}.$$

The fact mentioned above is also true for Corollary 2 and Theorem 3. Especially, the coefficients c_k $(k = 0, 1, 2)$ reduce to

$$c_0 = -\frac{(p-1)^2 + (p^2-1)\eta}{4},$$

$$c_1 = -\frac{p-1-(p^2-1)\eta}{2},$$

$$c_2 = \frac{(p^2-1)(1-\eta)}{4}.$$

These are essentially equivalent to the asymptotic expansion of Hotelling T^2 statistic.

3. Real Data Analysis

We briefly state the numerical result of theorems derived in the previous section through empirical studies. In this section, all computations were made using the statistical software SAS release 9.1.3 (SAS Institute Inc., Cary, NC).

The 55 urine specimens, given in Table 1, were analyzed in an effort to determine if certain physical characteristics of the urine might be related to the formation of calcium oxalate crystals. The data in Table 1 are a part of them in Andrews and Herzberg [1]. The five physical characteristics of the urine are:

x_1, y_1 : *specific gravity*, the density of the urine relative to water;

x_2, y_2 : *pH*, the negative logarithm of the hydrogen ion;

x_3, y_3 : *osmolarity* (mOsm), a unit used in biology and medicine but not in physical chemistry. Osmolarity is proportional to the concentration of molecules in solution;

x_4, y_4 : *conductivity* (mMho milliMho), one Mho is one reciprocal Ohm. Conductivity is proportional to the concentration of charged ions in solution;

x_5, y_5 : *urea concentration* in millimoles per liter.

Table 1. Physical characteristics of urines with and without crystals

No crystals **x**					Crystals **y**				
x_1	x_2	x_3	x_4	x_5	y_1	y_2	y_3	y_4	y_5
1.017	5.74	577	20.0	296	1.021	5.94	774	27.9	325
1.011	5.51	408	12.6	224	1.024	5.60	866	29.5	360
1.020	5.27	668	25.3	252	1.021	5.53	775	31.2	302
1.012	5.62	461	17.4	195	1.024	5.36	853	27.6	364
1.015	5.41	543	21.9	170	1.013	5.86	531	21.4	197
1.021	6.13	779	25.7	382	1.010	6.27	371	11.2	188
1.025	5.53	907	28.4	448	1.011	7.01	443	21.4	124
1.011	5.21	450	17.9	161	1.011	6.13	364	10.9	159
1.018	4.90	684	26.1	284	1.031	5.73	874	17.4	516
1.025	6.81	947	32.6	395	1.020	7.94	567	19.7	212
1.014	6.14	565	23.6	214	1.040	6.28	838	14.3	486
1.024	6.30	874	29.9	380	1.025	5.71	854	27.0	385
1.019	5.47	760	33.8	199	1.026	6.19	956	27.6	473
1.014	7.38	577	30.1	87	1.015	5.98	487	14.8	198
1.020	5.96	631	11.2	422	1.013	5.58	516	20.8	184
1.023	5.68	749	29.0	239	1.014	5.90	456	17.8	164
1.017	6.76	455	8.8	270	1.012	6.75	251	5.10	141
1.017	7.61	527	25.8	75	1.025	6.90	945	33.6	396
1.020	5.44	781	29.0	349	1.026	6.29	833	22.2	457
1.017	7.92	680	25.3	282	1.028	4.76	312	12.4	10
1.019	5.98	579	15.5	297	1.027	5.40	840	24.5	395
1.017	6.56	559	15.8	317	1.018	5.14	703	29.0	272
1.020	5.66	702	23.6	330	1.022	5.09	736	19.8	418
1.020	6.35	704	24.5	260	1.025	7.90	721	23.6	301
1.018	6.18	694	23.3	311	1.017	4.81	410	13.3	195
1.021	5.33	815	26.0	385	1.024	5.40	803	21.8	394
1.015	6.79	541	20.9	187	1.015	6.03	416	12.8	178
1.020	5.68	876	35.8	308					

The 55 patients are divided into two groups, with and without crystals, with 27 and 28 patients, respectively. It may be noted that there are missing observations in original data by Andrews and Herzberg [1]. But each group has only one missing observation. Hence the comparison may be carried out without them.

We conduct the profile analysis tests based on two methods for approximating the exact null distributions of the test statistics. One of the methods is based on the quantiles of the exact distribution of the test statistic is assuming multivariate normality. For testing $\mathbb{H}_1$, the parallelism hypothesis, we see that $n_x = 28$, $n_y = 27$ and $p = 5$. The sample mean vectors and pooled sample covariance matrix are

$$\bar{\mathbf{x}} = \begin{pmatrix} 1.018 \\ 6.047 \\ 660.464 \\ 23.564 \\ 275.679 \end{pmatrix}, \qquad \bar{\mathbf{y}} = \begin{pmatrix} 1.021 \\ 5.981 \\ 647.963 \\ 20.689 \\ 288.667 \end{pmatrix},$$

$$\mathbf{S}_{\text{pl}} = \begin{pmatrix} 0.328 \times 10^{-4} & -0.321 \times 10^{-3} & 0.807 & 0.148 \times 10^{-1} & 0.467 \\ -0.321 \times 10^{-3} & 0.599 & -8.688 & -0.633 \times 10^{-1} & -10.295 \\ 0.807 & -8.688 & 34225.206 & 993.680 & 17586.016 \\ 0.148 \times 10^{-1} & -0.633 \times 10^{-1} & 993.680 & 48.850 & 255.522 \\ 0.467 & -10.295 & 17586.016 & 255.522 & 13037.436 \end{pmatrix},$$

and

$$\mathbf{W} = \begin{pmatrix} 0.827 \times 10^{-4} & -0.224 \times 10^{-2} & -0.421 & -0.969 \times 10^{-1} & 0.438 \\ -0.224 \times 10^{-2} & 0.606 \times 10^{-1} & 11.410 & 2.624 & -11.855 \\ -0.421 & 11.410 & 2148.182 & 494.098 & -2231.827 \\ -0.969 \times 10^{-1} & 2.624 & 494.098 & 113.646 & -513.337 \\ 0.438 & -11.855 & -2231.827 & -513.337 & 2318.729 \end{pmatrix}.$$

Hence the test statistic given by (1)

$$\frac{55 - 5}{53 \times 4} \times 7.595 = 1.791$$

is not significant compared to the F-value for 4 and 50 degrees of freedom of 2.557 at the 0.05 level of significance (denote $F_{4, 50, 0.05} = 2.557$). We do not reject $\mathbb{H}_1$ and conclude that the profiles are parallel.

In keeping with our assumption that profiles are similar at the 5% level, we test $\mathbb{H}_2$ for coincident profiles and find that $\mathbf{1} = (1, 1, 1, 1, 1)^T$. Using (2), the test statistic $T_2 = 0.973 \times 10^{-3}$ is much smaller than $F_{1, 53, 0.05} = 4.023$. We cannot reject $\mathbb{H}_2$ and conclude that two groups are similar. To test $\mathbb{H}_3$ we find that

$$\frac{n_x \overline{\mathbf{x}} + n_y \overline{\mathbf{y}}}{n_x + n_y} = \begin{pmatrix} 1.019 \\ 6.015 \\ 654.327 \\ 22.153 \\ 282.055 \end{pmatrix},$$

and compare

$$T_3 = \frac{55 - 5}{53 \times 4} \times 3153.352 = 743.715,$$

with the critical value $F_{4, 50, 0.05} = 2.557$. Since $743.715 > 2.557$, $\mathbb{H}_3$ is rejected. We conclude that there is a difference in response in the type of request.

We will continue the other method based on quantiles given by (3), (4) and (5) when two samples have elliptical distributions and both covariance matrices are equal; however, actual tests are not available for these assumptions. Let $\varepsilon = n_y / n_x = 27/28 = 0.964$ and $\alpha = 0.05$. Note here that an estimate of kurtosis parameter, which is a proper consistent estimate, is known as the estimation by multivariate sample kurtosis. Let $\mathbf{S}_x = \mathbf{HLH}^T$, where $\mathbf{H}$ is an orthogonal matrix, $\mathbf{L} = \text{diag}(l_1, \ldots, l_p)$, $\boldsymbol{\xi}_j = (\xi_{1j}, \ldots, \xi_{pj})^T := \mathbf{H}^T \mathbf{x}_j$ and $\overline{\boldsymbol{\xi}} = (\overline{\xi}_1, \ldots, \overline{\xi}_p)^T := \mathbf{H}^T \overline{\mathbf{x}}$. Then a consistent estimator of η_x is given by

$$\hat{\eta}_x := \frac{1}{3pn_x} \sum_{i=1}^{p} l_i^{-2} \sum_{j=1}^{n_x} (\xi_{ij} - \overline{\xi}_i)^4 - 1.$$

In a similar way the consistent estimator for η_y is also given and denoted by $\hat{\eta}_y$. It may be noted that more efficient estimate could be found. In this paper, however, we present the approximations based on $\hat{\eta}_x$ and $\hat{\eta}_y$ instead of η_x and η_y. The estimation of the kurtosis parameter in elliptical distributions is discussed by Maruyama [5].

From Table 1, we obtain $\hat{\eta}_x = -0.195$, $\hat{\eta}_y = -0.158$ and approximate upper 5 percentile 10.703 by (3). Since the test statistic is not significant $(T_1 = 1.791 < 10.703)$, ... the parallelism hypothesis is not rejected. Next, we consider the level hypothesis $\mathbb{H}_2$ and have the approximate critical value 4.016 by (4). Then the level hypothesis is not rejected $(T_2 = 0.973 \times 10^{-3} < 4.016)$. That is, the responses of crystals and no crystals to the five characteristics appear to be the same. Finally, from (5) the approximation of critical point is given by 10.704. Then the flatness hypothesis is rejected $(T_3 = 743.715 > 10.704)$. Summing up, it is found that the two profiles are parallel and same level; however these are not flatness at $\alpha = 0.05$ significant level. Notice that consideration for multiplicity of tests is needed while practicing the actual tests.

Appendix: Proofs of Theorems

We first derive the asymptotic expansion for test statistic T_1 under the null hypothesis $\mathbb{H}_1$. The following result is discussed by Williams [10]:

$$(\bar{\mathbf{x}} - \bar{\mathbf{y}})^T \mathbf{S}_{\text{pl}}^{-1} (\bar{\mathbf{x}} - \bar{\mathbf{y}}) - \frac{((\bar{\mathbf{x}} - \bar{\mathbf{y}})^T \mathbf{S}_{\text{pl}}^{-1} \mathbf{1})^2}{\mathbf{1}^T \mathbf{S}_{\text{pl}}^{-1} \mathbf{1}} = (\mathbf{C}(\bar{\mathbf{x}} - \bar{\mathbf{y}}))^T (\mathbf{C}\mathbf{S}_{\text{pl}}\mathbf{C}^T)^{-1} \mathbf{C}(\bar{\mathbf{x}} - \bar{\mathbf{y}}), \quad (6)$$

where $\mathbf{C}$ is any $(p-1) \times p$ contrast matrix. It can be seen from (6) that, $T_{\text{H}} := n_x n_y (n_x + n_y)^{-1} (\mathbf{C}(\bar{\mathbf{x}} - \bar{\mathbf{y}}))^T (\mathbf{C}\mathbf{S}_{\text{pl}}\mathbf{C}^T)^{-1} \mathbf{C}(\bar{\mathbf{x}} - \bar{\mathbf{y}})$ is regarded as Hotelling type statistic by constructing the transformed observations $\mathbf{C}\mathbf{x}_1, \ldots, \mathbf{C}\mathbf{x}_{n_x}$ and $\mathbf{C}\mathbf{y}_1, \ldots, \mathbf{C}\mathbf{y}_{n_y}$ for independent samples of sizes n_x and n_y from the two populations. These have sample mean vectors $\mathbf{C}\bar{\mathbf{x}}$ and $\mathbf{C}\bar{\mathbf{y}}$, respectively, and pooled covariance matrix $\mathbf{C}\mathbf{S}_{\text{pl}}\mathbf{C}^T$. Then T_1 is rewritten as

$$\frac{n_x + n_y - p}{(n_x + n_y - 2)(p-1)} T_{\text{H}}.$$

Without loss of generality we assume $\mathbf{\Sigma} = \mathbf{I}_p$. Let

$$\bar{\mathbf{x}} = \boldsymbol{\mu}_x + \frac{\mathbf{v}_x}{\sqrt{n_x}},$$

$$\left(1 - \frac{1}{n_x}\right)\mathbf{S}_x + (\bar{\mathbf{x}} - \boldsymbol{\mu}_x)(\bar{\mathbf{x}} - \boldsymbol{\mu}_x)^T = \mathbf{I}_p + \frac{\mathbf{Z}_x}{\sqrt{n_x}}.$$

Similarly, define

$$\overline{\mathbf{y}} = \boldsymbol{\mu}_y + \frac{\mathbf{v}_y}{\sqrt{n_y}},$$

$$\left(1 - \frac{1}{n_y}\right)\mathbf{S}_y + (\overline{\mathbf{y}} - \boldsymbol{\mu}_y)(\overline{\mathbf{y}} - \boldsymbol{\mu}_y)^T = \mathbf{I}_p + \frac{\mathbf{Z}_y}{\sqrt{n_y}}.$$

Then $\mathbf{C}\mathbf{S}_{\mathrm{pl}}\mathbf{C}^T$ can be expressed by $\mathbf{u}_x := \mathbf{C}\mathbf{v}_x$, $\mathbf{u}_y := \mathbf{C}\mathbf{v}_y$, $\mathbf{M}_x := \mathbf{C}\mathbf{Z}_x\mathbf{C}^T$ and $\mathbf{M}_y := \mathbf{C}\mathbf{Z}_y\mathbf{C}^T$ as follows:

$$\mathbf{C}\mathbf{S}_{\mathrm{pl}}\mathbf{C}^T = \frac{n_x}{n_x - \frac{2}{1+\varepsilon}}\left(\mathbf{I}_{p-1} + \frac{\mathbf{M}_x + \sqrt{\varepsilon}\mathbf{M}_y}{\sqrt{n_x}(1+\varepsilon)} - \frac{\mathbf{u}_x\mathbf{u}_x^T + \mathbf{u}_y\mathbf{u}_y^T}{n_x(1+\varepsilon)}\right),$$

hence $(\mathbf{C}\mathbf{S}_{\mathrm{pl}}\mathbf{C}^T)^{-1} = \mathbf{I}_{p-1} + n_x^{-1/2}\mathbf{A}_1 + n_x^{-1}\mathbf{A}_2 + o(n_x^{-1})$, where

$$\mathbf{A}_1 = -\frac{\mathbf{M}_x + \sqrt{\varepsilon}\mathbf{M}_y}{1+\varepsilon},$$

$$\mathbf{A}_2 = \frac{\mathbf{M}_x^2 + \varepsilon\mathbf{M}_y^2 + \sqrt{\varepsilon}(\mathbf{M}_x\mathbf{M}_y + \mathbf{M}_y\mathbf{M}_x)}{(1+\varepsilon)^2} + \frac{\mathbf{u}_x\mathbf{u}_x^T - \mathbf{u}_y\mathbf{u}_y^T - 2\mathbf{I}_{p-1}}{1+\varepsilon}.$$

The characteristic function of T_1 is expressed with $\boldsymbol{\tau} := (\varepsilon/(1+\varepsilon))^{1/2}\mathbf{u}_x - (1+\varepsilon)^{-1/2}\mathbf{u}_y$ as

$$\mathrm{E}[e^{i\vartheta T_1}] = \mathrm{E}\left[e^{i\vartheta\boldsymbol{\tau}^T\boldsymbol{\tau}}\left\{1 - \frac{i\vartheta\boldsymbol{\tau}^T\mathbf{A}_1\boldsymbol{\tau}}{\sqrt{n_x}} + \frac{1}{n_x}\left(i\vartheta\boldsymbol{\tau}^T\mathbf{A}_2\boldsymbol{\tau} + \frac{(i\vartheta\boldsymbol{\tau}^T\mathbf{A}_1\boldsymbol{\tau})^2}{2}\right)\right\}\right] + o(n_x^{-1}).$$

Calculating expectations with respect to $\mathbf{u}_x$, $\mathbf{u}_y$, $\mathbf{M}_x$ and $\mathbf{M}_y$ by using the joint density of sample mean vector and sample covariance matrix, we obtain

$$\mathrm{E}[e^{i\vartheta T_1}] = (1 - 2i\vartheta)^{-(p-1)/2}\left(1 + \frac{a_0 + a_1(1-2i\vartheta)^{-1} + a_2(1-2i\vartheta)^{-2}}{n}\right) + o(n^{-1}).$$

Therefore, inverting the characteristic function, we have Theorem 1. By using the same way as mentioned above, we obtain Corollary 2 and Theorem 3. □

Acknowledgements

The author would like to thank the editor and referees for their helpful comments.

References

[1] D. F. Andrews and A. M. Herzberg, Data: A Collection of Problems from Many Fields for the Student and Research Worker, Springer-Verlag, New York, 1985.

[2] Y. Fujikoshi, An asymptotic expansion for the distribution of Hotelling's T^2-statistic under nonnormality, J. Multivariate Anal. 61(2) (1997), 187-193.

[3] D. G. Kabe and A. K. Gupta, Hotelling's T^2-distribution for a mixture of two normal populations, South African Statist. J. 24(2) (1990), 87-92.

[4] D. Kelker, Distribution theory of spherical distributions and a location-scale parameter generalization, Sankhyā Ser. A 32 (1970), 419-438.

[5] Y. Maruyama, Asymptotic properties for measures of multivariate kurtosis in elliptical distributions, Int. J. Pure Appl. Math. 25(3) (2005), 407-421.

[6] Y. Maruyama, Asymptotic expansions of the null distributions of some test statistics for profile analysis in general distributions, J. Statist. Plann. Inference 137(2) (2007), 506-526.

[7] M. Siotani, On the distributions of the Hotelling's T^2-statistics, Ann. Inst. Statist. Math. 8 (1956), 1-14.

[8] M. Siotani, An asymptotic expansion of the non-null distribution of Hotelling's generalized T_0^2-statistic, Ann. Math. Statist. 42 (1971), 560-571.

[9] M. S. Srivastava and H. M. Awan, On the robustness of Hotelling's T^2-test and distribution of linear and quadratic forms in sampling from a mixture of two multivariate normal populations, Comm. Statist. Theory Methods 11(1) (1982), 81-107.

[10] E. J. Williams, Comparing means of correlated variates, Biometrika 57(2) (1970), 459-461.