Available online at www.sciencedirect.com
SCIENCE DIRECT
ELSEVIER
Journal of Multivariate Analysis 96 (2005) 1–19
Journal of Multivariate Analysis
www.elsevier.com/locate/jmva

L_1-norm error bounds for asymptotic expansions of multivariate scale mixtures and their applications to Hotelling's generalized T_0^2

Y. Fujikoshi[a,*], V.V. Ulyanov[b], R. Shimizu[c]

[a] *Department of Mathematics, Hiroshima University, 1-3-1, Kagamiyama, Higashi-Hiroshima 739-8526, Japan*
[b] *Moscow State University, Moscow, Russia*
[c] *The Institute of Statistical Mathematics, Japan*

Received 18 February 2003
Available online 11 November 2004

Abstract

This paper is concerned with the distribution of a multivariate scale mixture variate $\boldsymbol{X}=(X_1, \dots, X_p)'$ with $X_i = S_i Z_i$, where $Z_1, \dots, Z_p$ are $i.i.d.$ random variables, $S_i > 0 (i=1, \dots, p)$, and $\{S_1, \dots, S_p\}$ is independent of $\{Z_1, \dots, Z_p\}$. First we obtain L_1-norm error bounds for an asymptotic expansion of the density function of $\boldsymbol{X}$ in the multivariate case as well as in the univariate case. Then the results are applied in obtaining error bounds for asymptotic expansions of the null distribution of Hotelling's generalized T_0^2-statistic. The special features of our results are that our error bounds are given in explicit and computable forms. Further, their orders are the same as ones of the usual order estimates, and hence the paper provides a new proof for validity of the asymptotic expansions.

AMS 1991 subject classification: primary 62H10; secondary 62E20

Keywords: Asymptotic expansions; Error bounds; Hotelling's generalized T_0^2; L_1-norm error bound; Multivariate scale mixture

* Corresponding author. Fax: +81 0824 24 0710.
E-mail address: fuji@math.sci.hiroshima-u.ac.jp (Y. Fujikoshi).

0047-259X/$ - see front matter
doi:10.1016/j.jmva.2004.09.002

1. Introduction

A general multivariate scale mixture variate may be defined as

$$\boldsymbol{X} = (X_1, \ldots, X_p)' = \boldsymbol{S}\boldsymbol{Z}, \tag{1.1}$$

where $\boldsymbol{S}$ is a positive definite random matrix, and $\boldsymbol{S}$ is independent of $\boldsymbol{Z} = (Z_1, \ldots, Z_p)'$. Let $G_i (i = 1, \ldots, p)$ be the distribution of Z_i. In this paper we are concerned with a special case: $\boldsymbol{S} = \mathrm{diag}(S_1, \ldots, S_p)$ and $Z_1, \ldots, Z_p$ are independently and identically distributed, that is

$$X_i = S_i Z_i, \quad i = 1, \ldots, p, \quad Z_1, \ldots, Z_p \sim i.i.d. G, \tag{1.2}$$

where G is the distribution function of Z_1. Here it is tacitly assumed that all the scale factors S_i are close to 1 in some sense. Note that Z_i may be a general random variable with a smooth density function, though useful applications appear in two cases when Z_1 is distributed as the standard normal distribution and a gamma distribution. Having in mind statistical applications we consider a transformation from $\boldsymbol{S}$ to $\boldsymbol{Y} = \boldsymbol{S}^{\delta/\rho}$, where $\delta = \pm 1$ and ρ is a positive number. The transformation in case (1.2) is expressed as

$$Y_i = S_i^{\delta/\rho}, \quad i = 1, \ldots, p. \tag{1.3}$$

In practical applications the positive constant ρ is chosen as $\frac{1}{2}$ or 1 according to that Z_1 is distributed as the standard normal distribution or a gamma distribution. The constant δ may be chosen so that the first few moments of $S_i^{\delta/\rho}, i = 1, \ldots, p$ are computable or evaluated.

There is a considerable work on asymptotic expansions and their error bounds for the distribution function of $\boldsymbol{X}$ in the univariate case $p = 1$. For a summary, see, e.g., Hall [6], Fujikoshi and Shimizu [4], Fujikoshi [3], Shimizu and Fujikoshi [10], Ulyanov et al. [11]. For multivariate scale mixtures, some special cases have been studied. For the distribution function, Fujikoshi and Shimizu [4] studied the case $\boldsymbol{S} = s\boldsymbol{I}_p$. Fujikoshi and Shimizu [5] treated the case $\boldsymbol{S} - \boldsymbol{I}_p \geqslant \mathrm{O}$, $G_i = \Phi$, $\delta = 1$ and $\rho = \frac{1}{2}$, where Φ is the distribution function of $N(0, 1)$. The latter case has been applied to the distribution of the MLE in a general MANOVA model.

In this paper we consider L_1-norm error bounds for asymptotic expansions of the density function of $\boldsymbol{X}$. It may be noted that such results are useful for an asymptotic expansion for $\mathrm{P}(\boldsymbol{X} \in A)$ and its error bound. Fujikoshi and Shimizu [5] and Shimizu [9] obtained L_1-norm error bound when $G_i = \Phi$, $\delta = 1$ and $\rho = \frac{1}{2}$. We are interesting in the null distribution of Hotelling's generalized T_0^2-statistic defined by

$$T_0^2 = n \,\mathrm{tr}\, \boldsymbol{S}_h \boldsymbol{S}_e^{-1}, \tag{1.4}$$

where $\boldsymbol{S}_h$ and $\boldsymbol{S}_e$ are independently distributed as Wishart distributions $W_p(q, \boldsymbol{I}_p)$ and $W_p(n, \boldsymbol{I}_p)$, respectively. We note (see, Section 4) that T_0^2 can be expressed as $T_0^2 = X_1 + \cdots + X_p = (1, \ldots, 1)\boldsymbol{X}$, and

$$\mathrm{P}(T_0^2 \leqslant x) = \mathrm{P}(\boldsymbol{X} \in A_x), \tag{1.5}$$

where $A_x = \{(x_1, \ldots, x_p); x_1 + \cdots + x_p \leqslant x\}$. Here $\boldsymbol{X}$ is a multivariate scale mixture variate as in (1.2) such that $Z_1, \ldots, Z_p \sim \chi_q^2$ and $S_1^{-1} > \cdots > S_p^{-1} > 0$ are the characteristic roots of $\boldsymbol{W}$, where $n\boldsymbol{W} \sim W_p(n, \boldsymbol{I}_p)$. We derive L_1-norm error bounds for asymptotic expansions of the density function of $\boldsymbol{X}$ as in (1.2), especially in the case $\delta = -1$ and $Z_1 \sim \chi_q^2$. The result is used to get an error bound for an asymptotic expansion of the distribution of T_0^2. In Section 2 we discuss the univariate case $p = 1$. The main results in the multivariate case are given in Section 3. The proofs are given in Section 5. In Section 4 we give an error bound for an asymptotic expansion of the distribution of T_0^2. It may be noted that our error bounds are given in explicit and computable forms. Further, their orders are the same as ones of the usual order estimates, and hence by products, the paper provides a new proof for validity of the asymptotic expansions.

2. L_1-norm error bound in the univariate case

The multivariate scale mixture variate $\boldsymbol{X}$ in (1.1) or (1.2) is written for $p = 1$ as

$$X = SZ, \tag{2.1}$$

where S is a positive random variable, and Z and S are independent. Let f and g be the probability density functions of X and Z, respectively. Let $D = \{x \in \mathbf{R} : g(x) > 0\}$. We assume that for a given positive integer k,
A1. g is k times continuously differentiable on D.
Consider the transformation $Y = S^{\delta/\rho}$ as in (1.3). Then the conditional density of X given $Y = y$ is $h(x, y) = y^{-\delta\rho} g(xy^{-\delta\rho})$, and hence we have

$$f(x) = \mathrm{E}[Y^{-\delta\rho} g(xY^{-\delta\rho})]. \tag{2.2}$$

We consider an asymptotic expansion of $f(x)$ based on Taylor expansion of the conditional density function $h(x, y)$ around $y = 1$. Related to the expansion, define a function $b_{\delta,j}(x)$ for $j \geqslant 1$ and for $x \in D$, by formula

$$\frac{\partial^j}{\partial y^j}\left\{y^{-\delta\rho} g(xy^{-\delta\rho})\right\} = y^{-j} y^{-\delta\rho} b_{\delta,j}(xy^{-\delta\rho}) g(xy^{-\delta\rho}) \tag{2.3}$$

and $b_{\delta,j}(x) = 0$ for $x \notin D$. We put $b_{\delta,0}(x) = 1$ for $j = 0$. Equality (2.3) can be easily checked by mathematical induction. Then we can write Taylor expansion for $h(x, y)$ as follows: for $y > 0$ and $x \in \mathbf{R}^1$,

$$g_{\delta,k}(x, y) = g(x) + \sum_{j=1}^{k-1} \frac{1}{j!} b_{\delta,j}(x) g(x)(y-1)^j, \tag{2.4}$$

where $g_{\delta,1}(x, y) = g(x)$. This suggests an approximation $g_{\delta,k}(x)$ for $f(x)$:

$$\begin{aligned} g_{\delta,k}(x) &= \mathrm{E}[g_{\delta,k}(x, Y)] \\ &= g(x) + \sum_{j=1}^{k-1} \frac{1}{j!} b_{\delta,j}(x) g(x) \mathrm{E}[(Y-1)^j] \end{aligned} \tag{2.5}$$

under the assumption $\mathrm{E}[|Y-1|^k] < \infty$.

Our error bounds depend on

$$\xi_{\delta,j} = \frac{1}{j!}\left\|b_{\delta,j}(x)g(x)\right\|_1, \ (j = 0, 1, \ldots, k) \tag{2.6}$$

and more precisely they are expressed in terms of

$$\eta_{\delta,k} = \left\{\xi_{\delta,k}^{1/k} + \left(2 + \sum_{j=1}^{k-1} \xi_{\delta,j}\right)^{1/k}\right\}^k. \tag{2.7}$$

Here for any integrable function $h(\boldsymbol{x}) : \mathbf{R}^p \to \mathbf{R}^1$, we define its L_1-norm by

$$\|h(\boldsymbol{x})\|_{1;p} = \int_{\mathbf{R}^p} |h(\boldsymbol{x})|d\boldsymbol{x}$$

and in particular, we write $\|\cdot\|_1 = \|\cdot\|_{1;1}$.

Lemma 2.1. *For any $k \geqslant 1$ it holds that*

$$\left\|f(x) - g_{\delta,k}(x)\right\|_1 \leqslant \eta_{\delta,k} \mathrm{E}\left[|Y-1|^k\right]. \tag{2.8}$$

Proof. We use a Taylor formula (see, e.g., [7, p. 257]) for a function h with $k(\geqslant 1)$ continuous derivatives

$$\begin{aligned} h(y) &= h(1) + \sum_{j=1}^{k-1} \frac{1}{j!} h^{(j)}(1)(y-1)^j \\ &\quad + \frac{(y-1)^k}{(k-1)!} \mathrm{E}\left[(1-\tau)^{k-1} h^{(k)}(1+\tau(y-1))\right], \end{aligned} \tag{2.9}$$

where τ is a random variable with uniform distribution $(0, 1)$. For any $y > 0$ let

$$\Delta_{\delta,k}(x, y) \equiv y^{-\delta\rho} g(xy^{-\delta\rho}) - g_{\delta,k}(x, y). \tag{2.10}$$

Using (2.3), (2.4) and (2.9) we can write also for $k \geqslant 1$

$$\begin{aligned} \Delta_{\delta,k}(x, y) &= \frac{(y-1)^k}{(k-1)!} \mathrm{E}\left[(1-\tau)^{k-1} (1+\tau(y-1))^{-k-\delta\rho}\right. \\ &\quad \left. \times b_{\delta,k}\left(x(1+\tau(y-1))^{-\delta\rho}\right) g\left(x(1+\tau(y-1))^{-\delta\rho}\right)\right]. \end{aligned} \tag{2.11}$$

The idea of our proof is to use (2.10) or (2.11) depending on whether y is far from 1 or close to it. Let

$$\varphi = (\xi_{\delta,k}/\eta_{\delta,k})^{1/k}.$$

Note that $\varphi : 0 < \varphi < 1$. If $y : 0 < y < \varphi$, then it follows from (2.10) that

$$\begin{aligned} \left\|\Delta_{\delta,k}(x, y)\right\|_1 &\leqslant \left(1 + \sum_{j=0}^{k-1} \xi_{\delta,j}\right) \frac{(1-y)^k}{(1-\varphi)^k} \\ &= \eta_{\delta,k}|y-1|^k. \end{aligned} \tag{2.12}$$

If $y \geqslant \varphi$, then for any $\tau \in [0, 1]$ we have $1 + \tau(y - 1) \geqslant \varphi$. Therefore it follows from (2.11) and Fubini theorem that

$$\left\| \Delta_{\delta,k}(x, y) \right\|_1 \leqslant \xi_{\delta,k} \frac{|y-1|^k}{\varphi^k} = \eta_{\delta,k}|y-1|^k. \tag{2.13}$$

Combining (2.10), (2.12) and (2.13) we get (2.8). □

Remark 2.1. For a discussion of the selection of constant φ, see, e.g., [10].

It is well known (see, e.g., [8, Chapter 14]) that L_1-norm error bound for differences between two densities allows to find closeness of the corresponding distributions. More generally the result may be stated as follows.

Lemma 2.2. *Let* $\mathrm{Q}_i(A)$, $i = 1, 2$ *be two set functions on the Borel* σ*-field in* $\mathbf{R}^p$ *defined by*

$$\mathrm{Q}_i(A) = \int_A f_i(\boldsymbol{x})\, d\boldsymbol{x}, \quad i = 1, 2.$$

Suppose that $Q_1(\mathbf{R}^p) = Q_2(\mathbf{R}^p)$. *Then*

$$\begin{aligned} |\mathrm{Q}_1(A) - \mathrm{Q}_2(A)| &\leqslant \tfrac{1}{2} \int_{\mathbf{R}^p} |f_1(\boldsymbol{x}) - f_2(\boldsymbol{x})| d\boldsymbol{x} \\ &= \tfrac{1}{2} \| f_1(\boldsymbol{x}) - f_2(\boldsymbol{x}) \|_{1;p} . \end{aligned} \tag{2.14}$$

From Lemmas 2.1 and 2.2 we have the following theorem.

Theorem 2.1. *Le X be a scale mixture of Z defined by* (2.1), *and* $Y = S^{\delta/\rho}$, *where* $\delta = 1$ *or* -1 *and* $\rho > 0$. *Suppose that the density function g of* Z_1 *satisfies* A1 *and* $\mathrm{E}(Y^k) < \infty$, *for a given integer* $k (> 1)$. *Then we have for any Borel set* $A \subset \mathbf{R}^1$

$$|\mathrm{P}(X \in A) - \int_A g_{\delta,k}(x)\, dx| \leqslant \tfrac{1}{2}\, \eta_{\delta,k} \mathrm{E}[|Y-1|^k]. \tag{2.15}$$

For a special case $\delta = 1$, $\rho = \frac{1}{2}$, and $Z \sim N(0, 1)$, we have (see, e.g., [10])

$$\frac{\partial^j}{\partial y^j} \left(y^{-\frac{1}{2}} \phi(x y^{-\frac{1}{2}}) \right) \Bigg|_{y=1} = 2^{-j} H_{2j}(x) \phi(x), \tag{2.16}$$

where $\phi(x)$ is the density function of $N(0, 1)$, and $H_j(x)$ is Hermite polynomial defined by

$$H_j(x) = (-1)^j \{\phi(x)\}^{-1} \frac{d^j}{dx^j} \phi(x).$$

This implies that

$$\frac{\partial^j}{\partial y^j} \left(y^{-\frac{1}{2}} \phi(x y^{-\frac{1}{2}}) \right) \Bigg|_{y=1} = 2^{-j} \frac{d^{2j}}{dx^{2j}} \phi(x). \tag{2.17}$$

Therefore our approximation $g_{1,k}(x)$ can be written in the form

$$g_{1,k}(x) = \mathrm{E}\left[\phi(x) + \sum_{j=1}^{k-1} \frac{1}{2^j j!}(Y_1-1)^j \frac{d^{2j}}{dx^{2j}}\phi(x)\right]$$
$$= \phi(x) + \sum_{j=1}^{k-1} \frac{1}{2^j j!} H_{2j}(x)\phi(x)\mathrm{E}[(Y_1-1)^j].$$

3. L_1-norm error bounds in the multivariate case

In this section we extend Theorem 2.1 to the multivariate scale mixture variate $\boldsymbol{X}$ as in (1.2). For its purpose, we derive an asymptotic expansion of the density function of $\boldsymbol{X}$ and its L_1-norm error bounds. Let $f_p(\boldsymbol{x})$ and $g_p(\boldsymbol{z})$ be the density functions of $\boldsymbol{X}$ and $\boldsymbol{Z}$, respectively. Then $g_p(\boldsymbol{z}) = g(z_1)\dots g(z_p)$, and the conditional density of $\boldsymbol{X}$ given $Y_i = y_i$, $i = 1, \dots, p$ is given by

$$h(\boldsymbol{x}, \boldsymbol{y}) = y_1^{-\delta\rho} g(x_1 y_1^{-\delta\rho}) \dots y_p^{-\delta\rho} g(x_p y_p^{-\delta\rho}), \tag{3.1}$$

where $\boldsymbol{x} = (x_1, \dots, x_p)'$ and $\boldsymbol{y} = (y_1, \dots, y_p)'$. We consider an approximation for $h(\boldsymbol{x}, \boldsymbol{y})$

$$\begin{aligned} g_{\delta,k,p}(\boldsymbol{x}, \boldsymbol{y}) &= g_p(\boldsymbol{x}) + \sum_{j=1}^{k-1} \frac{1}{j!}\left((y_1-1)\frac{\partial}{\partial z_1} + \cdots + (y_p-1)\frac{\partial}{\partial z_p}\right)^j \\ &\quad \times z_1^{-\delta\rho} g(x_1 z_1^{-\delta\rho}) \dots z_p^{-\delta\rho} g(x_p z_p^{-\delta\rho})\Big|_{z_1=\cdots=z_p=1} \\ &= g_p(\boldsymbol{x}) + \sum_{j=1}^{k-1} \sum_{(j)} \frac{1}{j_1! \dots j_p!} b_{\delta,j_1}(x_1) \dots b_{\delta,j_p}(x_p) g_p(\boldsymbol{x}) \\ &\quad \times (y_1-1)^{j_1} \dots (y_p-1)^{j_p}, \end{aligned} \tag{3.2}$$

which is an extension of $g_{\delta,k}(x, y)$, where the sum $\sum_{(j)}$ is taken over all non-negative integers such that $j_1 + \cdots + j_p = j$. This suggests an approximation for $f_p(\boldsymbol{x})$

$$\begin{aligned} g_{\delta,k,p}(\boldsymbol{x}) &= g_p(\boldsymbol{x}) + \sum_{j=1}^{k-1} \sum_{(j)} \frac{1}{j_1! \dots j_p!} b_{\delta,j_1}(x_1) \dots b_{\delta,j_p}(x_p) g_p(\boldsymbol{x}) \\ &\quad \times \mathrm{E}\left[(Y_1-1)^{j_1} \dots (Y_p-1)^{j_p}\right]. \end{aligned} \tag{3.3}$$

Our error bound depends on the quantity $\eta_{\delta,k,p}$ defined as follows. Let $\eta_{\delta,1,p} = 2 + v_{\delta,1,p}$ and for $k \geqslant 2$

$$\eta_{\delta,k,p} = \left\{ v_{\delta,k,p}^{1/k} + \left(2 + p\sum_{j=1}^{k-1} v_{\delta,j,p}\right)^{1/k} \right\}^k, \tag{3.4}$$

where

$$v_{\delta,j,p} = \sum_{[j]} \frac{(p-1)!}{i_1! \dots i_m!} \xi_{\delta,j_1} \dots \xi_{\delta,j_p}, \tag{3.5}$$

where the summation $[j]$ is taken over all non-negative integers $0 \leqslant j_1 \leqslant \cdots \leqslant j_p$ such that $j_1 + \cdots + j_p = j$, and the constants $m, i_1, \dots, i_m$ are positive integers such that

$$\begin{aligned} &0 \leqslant j_1 = \cdots = j_{i_1} < j_{i_1+1} = \cdots = j_{i_1+i_2} \\ &< \cdots < j_{i_1+\cdots+i_{m-1}+1} = \cdots = j_{i_1+\cdots+i_m} (= j_p) \leqslant j. \end{aligned}$$

For the motivation behind (3.5), see Lemma 5.3 and Remark 5.2 in Section 5. In particular, we have

$$\begin{aligned} v_{\delta,1,p} &= \xi_{\delta,1}, \\ v_{\delta,2,p} &= \xi_{\delta,2} + \tfrac{1}{2}(p-1)\xi_{\delta,1}^2, \\ v_{\delta,3,p} &= \xi_{\delta,3} + (p-1)\xi_{\delta,1}\xi_{\delta,2} + \tfrac{1}{6}(p-1)(p-2)\xi_{\delta,1}^3, \\ v_{\delta,4,p} &= \xi_{\delta,4} + \tfrac{1}{2}(p-1)\xi_{\delta,2}^2 + (p-1)\xi_{\delta,1}\xi_{\delta,3} \\ &\quad + \tfrac{1}{2}(p-1)(p-2)\xi_{\delta,1}^2\xi_{\delta,2} + \tfrac{1}{24}(p-1)(p-2)(p-3)\xi_{\delta,1}^4. \end{aligned} \tag{3.6}$$

In the following Theorems 3.1 and 3.2 we give two main results whose proofs are given in Section 5.

Theorem 3.1. *Let* $\boldsymbol{X} = \mathrm{diag}(S_1, \dots, S_p)\boldsymbol{Z}$ *be a multivariate scale mixture in* (1.2), *and* $Y_i = S_i^{\delta/\rho}$, $i = 1, \dots, p$, *where* $\delta = 1$ *or* -1 *and* $\rho > 0$. *Suppose that the density function* g *of* Z_1 *satisfies* A1 *and* $\mathrm{E}(Y_i^k) < \infty$, $i = 1, \dots, p$ *for a given integer k. Then we have for any Borel set* $A \subset \mathbf{R}^p$

$$|\mathrm{P}(\boldsymbol{X} \in A) - \int_A g_{\delta,k,p}(\boldsymbol{x}) d\boldsymbol{x}| \leqslant \tfrac{1}{2}\eta_{\delta,k,p} \sum_{i=1}^p \mathrm{E}[|Y_i - 1|^k]. \tag{3.7}$$

For an actual use of (3.7) we will take $k = 2$ or 4. If k is even, then the moment in the right-hand side can be expressed as

$$\sum_{i=1}^p \mathrm{E}[|Y_i - 1|^k] = \sum_{i=1}^p \mathrm{E}[|Y_i - 1|^k] = \mathrm{E}[\mathrm{tr}(\boldsymbol{W} - \boldsymbol{I}_p)^k]$$

which becomes more computable in applications, where $\boldsymbol{W} = \boldsymbol{H}\,\mathrm{diag}(Y_1, \dots, Y_p)\boldsymbol{H}'$ and $\boldsymbol{H}$ is any orthogonal matrix.

Theorem 3.2. *Under the same condition as in Theorem* 3.1 *we have for any Borel set* $A \subset \mathbf{R}^p$

$$|\mathrm{P}(\boldsymbol{X} \in A) - \int_A g_{\delta,k,p}(\boldsymbol{x})\, d\boldsymbol{x}| \leqslant \tfrac{1}{2} v_{\delta,k,p} \sum_{i=1}^p \mathrm{E}[|Y_i - 1|^k], \tag{3.8}$$

where $v_{\delta,k,p}$ are determined recursively by the relation

$$v_{\delta,k,p} = p^{-1}\left\{\eta_{\delta,k} + (p-1)\sum_{q=0}^{k-1} v_{\delta,k-q,p-1}\xi_{\delta,q}\right\}, \quad for\ k \geqslant 2, \tag{3.9}$$

with $v_{\delta,1,p} = \eta_{\delta,1}$, $v_{\delta,k,0} = 0$ and $v_{\delta,k,1} = \eta_{\delta,k}$ for all $k \geqslant 1$.

Remark 3.1. In the case $\delta = 1$, $\rho = \frac{1}{2}$ and $Z_1 \sim N(0,1)$ a similar result as in Theorem 3.2 has been obtained in Theorem 2 in Shimizu [9] with the same recurrence relation as (3.8) but with another initial value for $v_{\delta,k,1}$. Using (2.17) we can write the function $g_{1,k,p}$ as

$$\begin{aligned} g_{1,k,p}(\boldsymbol{x}) &= \mathrm{E}\left[\phi_p(\boldsymbol{x}) + \sum_{j=1}^{k-1} \frac{1}{2^j j!}\left(\partial'_{\boldsymbol{x}}(\boldsymbol{Y} - \boldsymbol{I}_p)\partial_{\boldsymbol{x}}\right)^j \phi_p(\boldsymbol{x})\right] \\ &= \phi_p(\boldsymbol{x}) + \sum_{j=1}^{k-1}\sum_{(j)}\prod_{i=1}^{p} \frac{1}{2^{j_i} j_i!} H_{2j_i}(x_i)\cdot\phi_p(\boldsymbol{x})\mathrm{E}\left[\sum_{i=1}^{p} |Y_i - 1|^k\right], \end{aligned}$$

where $\phi_p(\boldsymbol{x}) = \phi(x_1)\ldots\phi(x_p)$.

From relation (3.9) the constants $v_{\delta,k,p}$ for $k = 1, \ldots, 4$ are determined recursively as follows.

$$\begin{aligned} v_{\delta,1,p} &= \eta_{\delta,1}, \\ v_{\delta,2,p} &= \eta_{\delta,2} + \tfrac{1}{2}(p-1)\xi_{\delta,1}\eta_{\delta,1}, \\ v_{\delta,3,p} &= \eta_{\delta,3} + \tfrac{1}{2}(p-1)\{\xi_{\delta,1}\eta_{\delta,2} + \eta_{\delta,2}\eta_{\delta,1}\} \\ &\quad + \tfrac{1}{6}(p-1)(p-2)\xi_{\delta,1}\eta_{\delta,1}, \\ v_{\delta,4,p} &= \eta_{\delta,4} + \tfrac{1}{2}(p-1)\{\xi_{\delta,1}\eta_{\delta,3} + \xi_{\delta,2}\eta_{\delta,2} + \xi_{\delta,3}\eta_{\delta,1}\} \\ &\quad + \tfrac{1}{6}(p-1)(p-2)\{\xi_{\delta,1}^2\eta_{\delta,2} + 2\xi_{\delta,1}\xi_{\delta,2}\eta_{\delta,1}\} \\ &\quad + \tfrac{1}{24}(p-1)(p-2)(p-3)\xi_{\delta,1}^3\eta_{\delta,1}. \end{aligned} \tag{3.10}$$

Combining Theorems 3.1 and 3.2 we have

$$|\mathrm{P}(\boldsymbol{X} \in A) - \int_A g_{\delta,k,p}(\boldsymbol{x})\,d\boldsymbol{x}| \leqslant \tfrac{1}{2}\min(\eta_{\delta,k,p}, v_{\delta,k,p}) \sum_{i=1}^{p} \mathrm{E}[|Y_i - 1|^k]. \tag{3.11}$$

Note that

$$\eta_{\delta,1,p} = v_{\delta,1,p}, \quad \eta_{\delta,2,p} \geqslant v_{\delta,2,p}. \tag{3.12}$$

Now in special cases we show how to simplify the approximation

$$P_{\delta,k,p}(A) = \int_A g_{\delta,k,p}(\boldsymbol{x})\,d\boldsymbol{x} = \sum_{j=0}^{k-1} P^{(j)}_{\delta,k,p}(A), \tag{3.13}$$

where $P^{(0)}_{\delta,k,p}(A) = \int_A g_p(\boldsymbol{x})\,d\boldsymbol{x} = I_\delta(A)$ and for $j \geqslant 1$

$$P^{(j)}_{\delta,k,p}(A) = \sum_{(j)} \frac{1}{j_1! \dots j_p!} \int_A b_{\delta,j_1}(x_1) \dots b_{\delta,j_p}(x_p) g_p(\boldsymbol{x})\,d\boldsymbol{x} \times \mathrm{E}\left[(Y_1 - 1)^{j_1} \dots (Y_p - 1)^{j_p}\right]. \tag{3.14}$$

Note that in applications to the distribution of T_0^2 the set A is invariant, that is, A stays the same for any permutation of coordinates $x_1, \dots, x_p$. In the following, assume that A is invariant and let

$$\begin{aligned} I_{\delta,i}(A) &= \int_A b_{\delta,i}(x_1) g_p(\boldsymbol{x}) d\boldsymbol{x}, \\ I_{\delta,ij}(A) &= \int_A b_{\delta,i}(x_1) b_{\delta,j}(x_2) g_p(\boldsymbol{x}) d\boldsymbol{x}, \\ I_{\delta,ijk}(A) &= \int_A b_{\delta,i}(x_1) b_{\delta,j}(x_2) b_{\delta,j}(x_3) g_p(\boldsymbol{x}) d\boldsymbol{x}, \text{ so on.} \end{aligned} \tag{3.15}$$

Then we have

$$\begin{aligned} P^{(1)}_{\delta,k,p}(A) &= I_{\delta,1}(A) M_{1;p}, \\ P^{(2)}_{\delta,k,p}(A) &= \tfrac{1}{2} I_{\delta,2}(A) M_{2;p} + \tfrac{1}{2} I_{\delta,11}(A) M_{11;p}, \\ P^{(3)}_{\delta,k,p}(A) &= \tfrac{1}{6} I_{\delta,3}(A) M_{3;p} + \tfrac{1}{2} I_{\delta,21}(A) M_{12;p} + \tfrac{1}{6} I_{\delta,111}(A) M_{11,1;p}, \end{aligned} \tag{3.16}$$

where

$$\begin{aligned} M_{\alpha;p} &= \mathrm{E}[\sum_{i=1}^{p} (Y_i - 1)^\alpha], \\ M_{\alpha\beta;p} &= \mathrm{E}[\sum_{i \neq j}^{p} (Y_i - 1)^\alpha (Y_j - 1)^\beta], \\ M_{\alpha\beta\gamma;p} &= \mathrm{E}[\sum_{i \neq j \neq k}^{p} (Y_i - 1)^\alpha (Y_j - 1)^\beta (Y_j - 1)^\gamma], \quad \text{etc.} \end{aligned} \tag{3.17}$$

4. Hotelling's generalized T_0^2-statistic

In this section we consider error bounds for asymptotic expansion of the null distribution of Hotelling's generalized T_0^2-statistic defined by (1.5). The statistic is used as one of the test statistics in multivariate linear model. The limiting of T_0^2 is a chi-square distribution χ_r^2 with $r = pq$ degrees of freedom. Further, it is known (see, e.g., [1]) that T_0^2 has an asymptotic expansion

$$P(T_0^2 \leqslant x) = G_r(x) + \frac{r}{4n}\{(q-p-1)G_r(x) - 2qG_{r+2}(x) + (q+p+1)G_{r+4}(x)\} + \mathrm{O}(n^{-2}),$$

where G_r is the distribution function of χ_r^2 variate.

Lemma 4.1. *We can write* T_0^2 *in terms of a multivariate scale mixture* $\boldsymbol{X} = (X_1, \ldots, X_p)' = \mathrm{diag}(S_1, \ldots, S_p)\boldsymbol{Z}$ *as*

$$T_0^2 = X_1 + \cdots + X_p, \tag{4.1}$$

where $\boldsymbol{Z} = (Z_1, \ldots, Z_p)'$, $Z_1, \ldots, Z_p$ *are i.i.d. random variables,* $Z_1 \sim \chi_q^2$, $S_i = Y_i^{-1} (i = 1, \ldots, p)$, *and* $Y_1 > \cdots > Y_p > 0$ *are the characteristic roots of* $\boldsymbol{W}$ *such that* $n\boldsymbol{W} \sim W_p(n, I_p)$.

Proof. It is well known that the distribution of T_0^2 can be expressed as

$$\begin{aligned} T_0^2 &= n\,\mathrm{tr}(\boldsymbol{U}'\boldsymbol{U})\boldsymbol{S}_e^{-1} \\ &= n\,\mathrm{tr}(\boldsymbol{H}'\boldsymbol{U}'\boldsymbol{U}\boldsymbol{H})(\boldsymbol{H}'\boldsymbol{S}_e\boldsymbol{H})^{-1}, \end{aligned}$$

where $\boldsymbol{U}$ is a $q \times p$ random matrix whose elements are independent identically distributed as $N(0, 1)$, and $\boldsymbol{H}$ is an orthogonal matrix. Note that the distributions of $\boldsymbol{U}\boldsymbol{H}$ and $\boldsymbol{H}'\boldsymbol{S}_e\boldsymbol{H}$ are the same as $\boldsymbol{U}$ and $\boldsymbol{S}_e$, respectively. The result is obtained by choosing $\boldsymbol{H}$ such that $\boldsymbol{H}'\boldsymbol{S}_e\boldsymbol{H} = \mathrm{diag}(Y_1, \ldots, Y_p)$. □

Note that from Lemma 4.1

$$\mathrm{P}(T_0^2 \leqslant x) = \mathrm{P}(\boldsymbol{X} \in A_x) \simeq \int_{A_x} g_{\delta,k,p}(\boldsymbol{x})d\boldsymbol{x}$$

as in (1.5). Now we use Theorems 3.1 and 3.2 with $\delta = -1$, $\rho = 1$, $Z_1 \sim \chi_q^2$, and $k = 2$ or 4. Let $g_q(x)$ be a density function of χ_q^2, i.e.

$$g_q(x) = \frac{1}{2^{q/2}\Gamma(q/2)}\, x^{q/2-1} \exp(-x/2).$$

Then the functions $b_{-1,j}(x)$ defined by (2.3) are given by

$$\begin{aligned} b_{-1,1}(x) &= -\tfrac{1}{2}\,(x - q), \\ b_{-1,2}(x) &= \tfrac{1}{4}\,\{x^2 - 2qx + q(q-2)\}, \\ b_{-1,3}(x) &= -\tfrac{1}{8}\,\{x^3 - 3qx^2 + 3q(q-2)x - q(q-2)(q-4)\}. \end{aligned} \tag{4.2}$$

It is easy to see that

$$\begin{aligned} b_{-1,1}(x)g_q(x) &= \tfrac{1}{2}\,q\left\{g_q(x) - g_{q+2}(x)\right\}, \\ b_{-1,2}(x)g_q(x) &= \tfrac{1}{4}\,q\left\{(q-2)g_q(x) - 2qg_{q+2}(x) + (q+2)g_{q+4}(x)\right\}, \\ b_{-1,3}(x)g_q(x) &= \tfrac{1}{8}\,q\left\{(q-2)(q-4)g_q(x) - 3q(q-2)g_{q+2}(x)\right. \\ &\quad \left. + 3q(q+2)g_{q+4}(x) - (q+2)(q+4)g_{q+6}(x)\right\}. \end{aligned} \tag{4.3}$$

Using expression (4.3) we have

$$\begin{aligned} I_{-1,1}(A_x) &= \tfrac{1}{2}\,q\left[G_r(x) - G_{r+2}(x)\right], \\ I_{-1,2}(A_x) &= \tfrac{1}{4}\,q\,\{(q-2)G_r(x) - 2qG_{r+2}(x) + (q+2)G_{r+4}(x)\}, \\ I_{-1,11}(A_x) &= \tfrac{1}{4}\,q^2\,\{G_r(x) - 2G_{r+2}(x) + G_{r+4}(x)\}, \end{aligned}$$

$$
\begin{aligned}
I_{-1,3}(A_x) &= \tfrac{1}{8}\,q\,\{(q-2)(q-4)G_r(x) - 3q(q-2)G_{r+2}(x) \\
&\quad +3q(q+2)G_{r+4}(x) - (q+2)(q+4)G_{r+6}(x)\}, \\
I_{-1,21}(A_x) &= \tfrac{1}{8}\,q^2\,\{(q-2)G_r(x) - (3q-2)G_{r+2}(x) \\
&\quad +(3q+2)G_{r+4}(x) - (q+2)G_{r+6}(x)\}, \\
I_{-1,111}(A_x) &= \tfrac{1}{8}\,q^3\,\{G_r(x) - 3G_{r+2}(x) + 3G_{r+4}(x) - G_{r+6}(x)\}.
\end{aligned}
\tag{4.4}
$$

Let $\boldsymbol{V} = \boldsymbol{W} - \boldsymbol{I}_p$. Then it is easily seen that

$$
\begin{aligned}
M_{j;p} &= \sum_{i=1}^{p} (Y_i - 1)^j = \mathrm{tr}\,\boldsymbol{V}^j, \quad j = 1, 2, \ldots, \\
M_{11;p} &= \sum_{i\neq j}^{p} (Y_i - 1)(Y_j - 1) = (\mathrm{tr}\,\boldsymbol{V})^2 - \mathrm{tr}\,\boldsymbol{V}^2, \\
M_{21;p} &= \sum_{i\neq j}^{p} (Y_i - 1)^2(Y_j - 1) = \mathrm{tr}\,\boldsymbol{V}\,\mathrm{tr}\,\boldsymbol{V}^2 - \mathrm{tr}\,\boldsymbol{V}^3, \\
M_{111;p} &= \sum_{i\neq j\neq k}^{p} (Y_i - 1)(Y_j - 1)(Y_k - 1) = (\mathrm{tr}\,\boldsymbol{V})^3 - 3\,\mathrm{tr}\,\boldsymbol{V}\,\mathrm{tr}\,\boldsymbol{V}^2 + 2\,\mathrm{tr}\,\boldsymbol{V}^3.
\end{aligned}
\tag{4.5}
$$

These imply the following expressions:

$$
\begin{aligned}
P_{\delta,1,p}(A_x) &= I_{\delta,1}(A_x), \\
P_{\delta,2,p}(A_x) &= P_{\delta,1,p}(A_x) + I_{\delta,1}(A_x)\mathrm{E}[\mathrm{tr}\,\boldsymbol{V}], \\
P_{\delta,3,p}(A_x) &= P_{\delta,2,p}(A_x) + \tfrac{1}{2}\,I_{\delta,2,p}(A_x)\mathrm{E}[\mathrm{tr}\,\boldsymbol{V}^2] \\
&\quad +\tfrac{1}{2}\,I_{\delta,11}(A_x)\mathrm{E}[\{(\mathrm{tr}\,\boldsymbol{V})^2 - \mathrm{tr}\,\boldsymbol{V}^2\}], \\
P_{\delta,4,p}(A_x) &= P_{\delta,3,p}(A_x) + \tfrac{1}{6}\,I_{\delta,3}(A_x)\mathrm{E}[\mathrm{tr}\,\boldsymbol{V}^3] \\
&\quad +\tfrac{1}{2}\,I_{\delta,21}(A_x)\mathrm{E}[\{\mathrm{tr}\,\boldsymbol{V}\,\mathrm{tr}\,\boldsymbol{V}^2 - \mathrm{tr}\,\boldsymbol{V}^3\}] \\
&\quad +\tfrac{1}{6}\,I_{\delta,111}(A_x)\mathrm{E}[\{(\mathrm{tr}\,\boldsymbol{V})^3 - 3\,\mathrm{tr}\,\boldsymbol{V}\,\mathrm{tr}\,\boldsymbol{V}^2 + 2\,\mathrm{tr}\,\boldsymbol{V}^3\}].
\end{aligned}
\tag{4.6}
$$

Further, from moment formulas (see, e.g., [2]) on Whishart matrix we have the following results:

$$
\begin{aligned}
&\mathrm{E}[\mathrm{tr}\,\boldsymbol{V}] = 0, \quad \mathrm{E}[\mathrm{tr}\,\boldsymbol{V}^2] = \frac{1}{n}\,p(p+1), \\
&\mathrm{E}[(\mathrm{tr}\,\boldsymbol{V})^2] = \frac{1}{n}\,2p, \quad \mathrm{E}[\mathrm{tr}\,\boldsymbol{V}^3] = \frac{1}{n^2}\,p(p^2+3p+4), \\
&\mathrm{E}[\mathrm{tr}\,\boldsymbol{V}\,\mathrm{tr}\,\boldsymbol{V}^2] = \frac{4}{n^2}\,p(p+1), \quad \mathrm{E}[(\mathrm{tr}\,\boldsymbol{V})^3] = \frac{8}{n^2}\,p, \\
&\mathrm{E}[\mathrm{tr}\,\boldsymbol{V}^4] = \frac{1}{n^2}\,p(2p^2+5p+5).
\end{aligned}
\tag{4.7}
$$

Theorem 4.1. *Let T_0^2 be the generalized Hotelling statistic defined by* (1.5). *Under the assumption $n \geqslant p$, it holds that*

$$\text{(i) } |\mathrm{P}(T_0^2 \leqslant x) - G_r(x)| \leqslant \frac{1}{2n} p(p+1) v_{-1,2,p}(q),$$

$$\begin{aligned}\text{(ii) } |\mathrm{P}(T_0^2 \leqslant x) &- G_r(x) - \frac{r}{4n}\{(q-p-1)G_r(x) \\ &-2qG_{r+2}(x) + (q+p+1)G_{r+4}(x)\}| \\ \leqslant& \frac{r}{48n^2}(|h| + 24q|p-q+1| + 24q(p+q+1) \\ &+\frac{1}{2n^2} p(2p^2+5p+5)\min\{\eta_{-1,4,p}(q), v_{-1,4,p}(q)\},\end{aligned}$$

where $r = pq$ and $h = 8\{p^2 - 3(q-1)p + q^2 - 3q + 4\}$. Here $\eta_{-1,k,p}(q)$ and $v_{-1,k,p}(q)$ are the $\eta_{\delta,k,p}$ in (3.4) *and $v_{\delta,k,p}$ in* (3.9), *respectively, for the case $Z_1 \sim \chi_q^2$.*

Proof. Substituting (4.4) and (4.7) into (4.6) we have

$$\begin{aligned} P_{-1,1,p}(A_x) &= \mathrm{P}(\mathbf{Z} \in A_x) = \mathrm{P}(\chi_r^2 \leqslant x) = G_r(x), \\ P_{-1,2,p}(A_x) &= G_r(x), \\ P_{-1,4,p}(A_x) &= G_r(x) + \frac{r}{4n}\{(q-p-1)G_r(x) \\ &\quad -2qG_{r+2}(x) + (q+p+1)G_{r+4}(x)\} \\ &\quad +\frac{r}{48n^2}\{h(G_r(x) - G_{r+6}(x)) - 24q(p-q+1)(G_{r+2}(x) - G_{r+4}(x)) \\ &\quad +24(p+q+1)q(G_{r+4}(x) - G_{r+6}(x))\}. \end{aligned} \tag{4.8}$$

Therefore, it follows from (3.11) that

$$|\mathrm{P}(T_0^2 \leqslant x) - P_{-1,2,p}(A_x)| \leqslant \frac{1}{2n} p(p+1) v_{-1,2,p}(q), \tag{4.9}$$

since $\eta_{-1,2,p}(q) \geqslant v_{-1,2,p}(q)$ (see (3.12)), and

$$\begin{aligned} &|\mathrm{P}(T_0^2 \leqslant x) - P_{-1,4,p}(A_x)| \\ &\quad \leqslant \frac{1}{2n^2} p(2p^2+5p+5)\min\{\eta_{-1,4,p}(q), v_{-1,4,p}(q)\}. \end{aligned} \tag{4.10}$$

These imply the theorem. □

The quantities of $\eta_{-1,k,p}(q)$ and $v_{-1,k,p}(q)$ are numerically obtained. For $k = 2, 4$, $q =$ 1–5 and $p =$ 1–4, see Tables 1 and 2.

From Tables 1 and 2 it seems that the bound $v_{-1,k,p}(q)$ is better than the bound $\eta_{-1,k,p}(q)$, at least, for some region of q and p.

5. Proofs of Theorems 3.1 and 3.2

Lemma 5.1. *For a positive integer j, let*

$$0 \leqslant j_1 \leqslant j_2 \leqslant \cdots \leqslant j_p$$

Table 1
Values of $\eta_{-1,2,p}(q)$ and $\eta_{-1,4,p}(q)$

q	$\eta_{-1,2,p}(q)$				$\eta_{-1,4,p}(q)$			
	p				p			
	1	2	3	4	1	2	3	4
1.0	4.57	5.67	6.76	7.85	15.68	25.43	39.21	57.77
2.0	5.71	7.64	9.56	11.47	22.58	44.24	78.44	128.19
3.0	6.69	9.34	11.97	14.59	30.36	66.82	127.69	219.52
4.0	7.61	10.91	14.19	17.46	38.88	93.15	186.92	331.36
5.0	8.45	12.37	16.26	20.15	47.97	122.61	254.72	461.11

Table 2
Values of $v_{-1,2,p}(q)$ and $v_{-1,4,p}(q)$

q	$v_{-1,2,p}(q)$				$v_{-1,4,p}(q)$			
	p				p			
	1	2	3	4	1	2	3	4
1.0	4.57	5.18	5.78	6.38	15.68	18.74	22.41	26.76
2.0	5.71	6.72	7.72	8.73	22.58	28.93	37.04	47.18
3.0	6.69	8.04	9.40	10.75	30.36	40.61	54.16	71.59
4.0	7.61	9.28	10.95	12.62	38.88	53.69	73.72	99.93
5.0	8.45	10.43	12.40	14.36	47.97	67.91	95.31	131.65

be integers such that $j_1 + \cdots + j_p = j$ *and* $a_1, \ldots, a_p$ *be non-negative real numbers. Then*

$$\sum a_1^{\ell_1} \cdots a_p^{\ell_p} \leqslant (p-1)!(a_1^j + \cdots + a_p^j), \tag{5.1}$$

where summation on the left-hand side is taken over all $p!$ *permutations* $(\ell_1, \ldots, \ell_p)$ of $(j_1, \ldots, j_p)$.

Proof. We prove (5.1) by mathematical induction on p. If $p = 1$, then (5.1) is obvious. We assume that (5.1) is valid for $p - 1$ ($\geqslant 1$). We write the left-hand side of (5.1) in the form

$$\sum a_1^{\ell_1} \ldots a_p^{\ell_p} = a_1^{j_1} P_{j-j_1}(a_2, \ldots, a_p) + \cdots + a_1^{j_p} P_{j-j_p}(a_2, \ldots, a_p),$$

where $P_{j-j_1}(a_2, \ldots, a_p) = \sum a_2^{\ell_2} \ldots a_p^{\ell_p}$ and the summation here is taken over all $(p-1)!$ permutations $(\ell_2, \ldots, \ell_p)$ of $(j_2, \ldots, j_p)$. Polynomials $P_{j-j_2}, \ldots, P_{j-j_p}$ are defined similarly. The hypothesis of the induction asserts that for all $k = 1, 2, \ldots, p$ we have

$$P_{j-j_k}(a_2, \ldots, a_p) \leqslant (p-2)!(a_2^{j-j_k} + \cdots + a_p^{j-j_k}).$$

Therefore we get

$$\sum a_1^{\ell_1} \ldots a_p^{\ell_p} \leqslant (p-2)! \Big[a_1^{j_1} \left(a_2^{j-j_1} + \cdots + a_p^{j-j_1} \right) + \cdots + a_1^{j_p} \left(a_2^{j-j_p} + \cdots + a_p^{j-j_p} \right) \Big]. \tag{5.2}$$

It is clear that on the right-hand side of (5.2) we can replace a_1 by any other a_i, that is, for $i = 1, 2, \ldots, p$ we have

$$\sum a_1^{\ell_1} \cdots a_p^{\ell_p} \leqslant (p-2)! \left[a_i^{j_1} \sum_{k=1, k \neq i}^{p} a_k^{j-j_1} + \cdots + a_i^{j_p} \sum_{k=1, k \neq i}^{p} a_k^{j-j_p} \right]. \tag{5.3}$$

Note that for any positive b_1 and b_2 a function $b_1^{j-x} b_2^x + b_1^x b_2^{j-x}$ of x is convex on $[0, j]$ and is equal to $b_1^j + b_2^j$ for $x = 0$ and $x = j$ (cf. Lemma 2 in [9]). Therefore for all integers $i = 0, 1, \ldots, j$ we have

$$b_1^{j-i} b_2^i + b_1^i b_2^{j-i} \leqslant b_1^j + b_2^j. \tag{5.4}$$

Thus summing up inequalities (5.3) for $i = 1, 2, \ldots, p$ we get from (5.4) that

$$p \sum a_1^{\ell_1} \cdots a_p^{\ell_p} \leqslant (p-2)!(p-1)p(a_1^j + \cdots + a_p^j).$$

Hence we obtain (5.1). □

Remark 5.1. Note when $a_1 = \cdots = a_p = 1$ inequality (5.1) is written in the form $p! \leqslant p!$. Therefore, (5.1) is sharp.

Lemma 5.2. *In Lemma* 5.1 *we assume*

$$0 \leqslant j_1 = \cdots = j_{i_1} < j_{i_1+1} = \cdots = j_{i_1+i_2} < \cdots < j_{i_1+\cdots+i_{m-1}+1} = \cdots = j_{i_1+\cdots+i_m} (= j_p) \leqslant j.$$

Then

$$\sum_1 a_1^{\ell_1} \cdots a_p^{\ell_p} \leqslant \frac{(p-1)!}{i_1! \cdots i_m!} (a_1^j + \cdots + a_p^j), \tag{5.5}$$

where summation on the left-hand side is taken over all different permutations $(\ell_1, \ldots, \ell_p)$ *of* $(j_1, \ldots, j_p)$.

Proof. The result follows from

$$\sum a_1^{\ell_1} \cdots a_p^{\ell_p} = i_1! \cdots i_m! \sum_1 a_1^{\ell_1} \cdots a_p^{\ell_p}$$

and by applying Lemma 5.1 to the left-hand side of the above equality. □

Lemma 5.3. *Assume that* $g(x)$ *satisfies* A1 *in Section* 2. *Let* $j\,(0 < j \leqslant k)$ *be a positive integer and*

$$I(\boldsymbol{x}) = \sum_{(j)} \prod_{i=1}^{p} a_i^{j_i} \frac{1}{j_i!} y_{ij_i}^{-j_i} b_{\delta,j_i}(x_i y_{ij_i}^{-\delta\rho}) g(x_i y_{ij_i}^{-\delta\rho}),$$

where $a_1, \ldots, a_p$ *are positive numbers. If all* $y_{ij_i} \geqslant \varphi > 0$, $i = 1, 2, \ldots, p$, *then*

$$\|I(x)\|_{1,p} \leqslant \varphi^{-j} v_{\delta,j,p}(a_1^j + \cdots + a_p^j), \tag{5.6}$$

where $v_{\delta,j,p}$ *is given by* (3.5).

Proof. Since all $y_{ij_i} \geqslant \varphi$ and for any permutation of $(j_1, \ldots, j_p)$ a product $\xi_{\delta,j_1} \ldots \xi_{\delta,j_p}$ does not change, we get

$$\|I(\boldsymbol{x})\|_{1,p} \leqslant \varphi^{-j} \sum_{[j]} \left(\xi_{j_1} \ldots \xi_{j_p} \sum_{1} a_1^{\ell_1} \ldots a_p^{\ell_p} \right), \tag{5.7}$$

where $\sum_{[j]}$ denotes summation over all non-negative integers $0 \leqslant j_1 \leqslant \cdots \leqslant j_p$ such that $j_1 + \cdots + j_p = j$, and $\sum_1$ means summation over all different permutations $\{\ell_1, \ell_2, \ldots, \ell_p\}$ of a fixed set $\{j_1, j_2, \ldots, j_p\}$. We get (5.6) from (5.7) and Lemma 5.2. □

Remark 5.2. Let $h(\boldsymbol{x}, \boldsymbol{y})$ be the conditional density function defined by (3.1). Assume that $g(x)$ satisfies A1. Then for any positive $a_1, \ldots, a_p$ and a positive integer $j : 0 < j \leqslant k$ we have

$$\left\| \left(a_1 \frac{\partial}{\partial y_1} + \cdots + a_p \frac{\partial}{\partial y_p} \right)^j h(\boldsymbol{x}, \boldsymbol{y}) \Big|_{y_i = y_{i0}, i=1,2,\ldots,p} \right\|_{1;p} \leqslant j! \varphi^{-j} v_{\delta,j,p}(a_1^j + \cdots + a_p^j), \tag{5.8}$$

provided $y_{i0} \geqslant \varphi > 0$, $i = 1, 2, \ldots, p$, where $v_{\delta,j,p}$ is defined by (3.5). The result follows immediately from Lemma 5.3 and the fact that the left-hand side equals

$$\sum_{(j)} j! \prod_{i=1}^{p} a_i^{j_i} \frac{1}{j_i!} y_{i0}^{-j_i - \delta\rho} b_{\delta,j_i}(x_i y_{i0}^{-\delta\rho}) g(x_i y_{i0}^{-\delta\rho}).$$

Proof of Theorem 3.1. Note that

$$f_p(\boldsymbol{x}) = \mathrm{E}\,[h(\boldsymbol{x}, \boldsymbol{Y})]\,,$$

where $\boldsymbol{Y} = (Y_1, \ldots, Y_p)'$ and a function h is defined in (3.1). We construct an expansion for using (2.9) sequentially. Namely, at first we apply (2.9) to $y_1^{-\delta\rho} g(x_1 y_1^{-\delta\rho})$.

We get

$$h(\boldsymbol{x},\boldsymbol{y}) = \left\{ g(x_1) + \sum_{j=1}^{k-1} \frac{1}{j!} b_{\delta,j}(x_1) g(x_1)(y_1-1)^j + R_{\delta,k;1}(y_1-1)^k \right\} \times h_2(\boldsymbol{x},\boldsymbol{y}), \tag{5.9}$$

where

$$R_{\delta,k;1} = \frac{1}{(k-1)!} \mathrm{E}\left[(1-\tau)^{k-1} \frac{\partial}{\partial y^k} \left(y^{-\delta\rho} g(x_1 y^{-\delta\rho}) \right) \Big|_{y=1+\tau(y_1-1)} \right],$$

$$h_2(\boldsymbol{x},\boldsymbol{y}) = \prod_{i=2}^{p} y_i^{-\delta\rho} g(x_i y_i^{-\delta\rho}).$$

Now we apply (2.9) for a function $y_2^{-\delta\rho} g(x_2 y_2^{-\delta\rho})$ so that for a summand

$$\frac{1}{j!} b_{\delta,j}(x_1) g(x_1)(y_1-1)^j h_2(\boldsymbol{x},\boldsymbol{y}),$$

we apply (2.9) with k replaced by $k-j$. At last we obtain the following expansion:

$$h(\boldsymbol{x},\boldsymbol{y}) = g(x_1)\ldots g(x_p) + \sum_{j=1}^{k-1} \sum_{(j)} \prod_{i=1}^{p} \frac{1}{j_i!} b_{\delta,j_i}(x_i) g(x_i)(y_1-1)^j + \Delta_{\delta,k,p}(\boldsymbol{x},\boldsymbol{y}), \tag{5.10}$$

where $R_{\delta,k,p}$ is a sum of terms each of which can be written in the form

$$(y_1-1)^{k_1} \cdots (y_p-1)^{k_p} I_{k_1}(y_1) \cdots I_{k_p}(y_p) \tag{5.11}$$

with $k_i \geqslant 0$ for $i = 1, 2, \ldots, p$ and $k_1 + \cdots + k_p = k$. Each factor I_j in (5.11) has one of the following form:

$$I_k(y) = \frac{1}{(k-1)!} \int_0^1 (1-\tau)^{k-1} \frac{\partial^k}{\partial y_1^k} \left(y_1^{-\delta\rho} g(x y_1^{-\delta\rho}) \right) \Big|_{y_1=1+\tau(y-1)} d\tau, \tag{5.12}$$

$I_0(y) = g(x)$ or $I_0(y) = y^{-\delta\rho} g(x y^{-\delta\rho})$ and when $j : 1 \leqslant j \leqslant k-1$, we have for $I_j(y)$ one of the two representations:

$$\frac{1}{j!} b_{\delta,j}(x) g(x) \quad \text{or}$$

$$\frac{1}{(j-1)!} \int_0^1 (1-\tau)^{j-1} \frac{\partial^j}{\partial y_1^j} \left(y_1^{-\delta\rho} g(x y_1^{-\delta\rho}) \right) \Big|_{y_1=1+\tau(y-1)} d\tau.$$

Let

$$\varphi = \left(v_{\delta,k,p} / \eta_{\delta,k,p} \right)^{1/k}. \tag{5.13}$$

At first we consider the case when $0 < \min(y_1, \ldots, y_p) \leqslant \varphi$. Assume that y_1 is such that $0 < y_1 \leqslant \varphi$. We have for any $j : 1 \leqslant j \leqslant k$,

$$\begin{aligned} |1-y_1|^j + \cdots + |1-y_p|^j &\leqslant \frac{1}{(1-\varphi)^{k-j}} \Big(|1-y_1|^k \\ &\quad + |1-y_1|^{k-j} |1-y_2|^j + \cdots + |1-y_1|^{k-j} |1-y_p|^j \Big) \\ &\leqslant \frac{p}{(1-\varphi)^{k-j}} \Big(|1-y_1|^k + \cdots + |1-y_p|^k \Big). \end{aligned} \tag{5.14}$$

Therefore, using Lemma 5.3, (5.10) and (5.13) we get

$$\begin{aligned} \|R_{\delta,k,p}\|_{1;p} &\leqslant 2 + \sum_{j=1}^{k-1} \Big(|y_1-1|^j + \cdots + |y_p-1|^j \Big) v_{\delta,j,p} \\ &\leqslant \frac{1}{(1-\varphi)^k} \Big(|1-y_1|^k + \cdots + |1-y_p|^k \Big) \left(2 + p \sum_{j=1}^{k-1} v_{\delta,j,p} \right) \\ &\leqslant \eta_{\delta,k,p} \Big(|y_1-1|^k + \cdots + |y_p-1|^k \Big). \end{aligned} \tag{5.15}$$

If $\min(y_1, \ldots, y_p) > \varphi$ then using Lemma 5.3, (5.12) and representations for summands contained in $R_{\delta,k,p}$ we get

$$\begin{aligned} \|R_{\delta,k,p}\|_{1;p} &\leqslant \varphi^{-k} v_{\delta,k,p} \Big(|y_1-1|^k + \cdots + |y_p-1|^k \Big) \\ &= \eta_{\delta,k,p} \Big(|y_1-1|^k + \cdots + |y_p-1|^k \Big). \end{aligned} \tag{5.16}$$

According to Remark 5.2 and combining (5.15) and (5.16) we finish the proof of Theorem 3.1. □

Proof of Theorem 3.2. Let $h(\boldsymbol{x}, \boldsymbol{y}) = \prod_{i=1}^p y_i^{-\delta\rho} g(x_i y_i^{-\delta\rho})$, and

$$\Delta_{\delta,k,p}(\boldsymbol{x}, \boldsymbol{y}) = h(\boldsymbol{x}, \boldsymbol{y}) - g_{\delta,k,p}(\boldsymbol{x}, \boldsymbol{y}), \tag{5.17}$$

where $g_{\delta,k,p}(\boldsymbol{x}, \boldsymbol{y})$ is defined by (3.3). In order to prove (3.8) it is enough as usual to show that

$$\|\Delta_{\delta,k,p}(\boldsymbol{x}, \boldsymbol{y})\|_{1;p} \leqslant v_{\delta,k,p} \sum_{i=1}^p |y_i - 1|^k, \tag{5.18}$$

where y_i, $i = 1, \ldots, p$, are considered as positive real numbers. In the following we show that the result can be proved by using arguments similar to the proof of Lemma 2 in [9]. We prove (5.18) by mathematical induction with respect to p. For $p = 1$, $v_{\delta,k,1} = \eta_{\delta,k}$, and hence inequality (5.18) was proved in Lemma 2.1 or Theorem 2.1. For $p \geqslant 2$ we use

$$h(\boldsymbol{x}, \boldsymbol{y}) = \left\{ g(x_p) + \sum_{j=1}^{k-1} \frac{1}{j!}(y_p - 1)^j b_{\delta,j}(x_p) g(x_p) + \Delta_{\delta,k,1}(x_p, y_p) \right\} \times \prod_{i=1}^{p-1} y_i^{-\delta\rho} g(x_i y_i^{-\delta\rho}), \tag{5.19}$$

where $\|\Delta_{\delta,k,1}(x_p, y_p)\|_1 \leqslant \eta_{\delta,k}|y_p - 1|^k$. We apply equality (5.17) to $\prod_{i=1}^{p-1} y_i^{-\delta\rho} g(x_i y_i^{-\delta\rho})$ with p replaced by $p-1$ and k replaced by $k-j$ when $\prod_{i=1}^{p-1} y_i^{-\delta\rho} g(x_i y_i^{-\delta\rho})$ is a factor by $(y_p - 1)^j$ in (5.19). Thus we get

$$\begin{aligned} \Delta_{\delta,k,p}(\boldsymbol{x}, \boldsymbol{y}) = {} & g(x_p)\Delta_{\delta,k,p-1}(\boldsymbol{x}_{(-p)}, \boldsymbol{y}_{(-p)}) \\ & + \sum_{j=1}^{k-1} \frac{1}{j!}(y_p - 1)^j\, b_{\delta,j}(x_p) g(x_p) \Delta_{\delta,k-j,p-1}(\boldsymbol{x}_{(-p)}, \boldsymbol{y}_{(-p)}) \\ & + \Delta_{\delta,k,1}(x_p, y_p) \prod_{i=1}^{p-1} y_i^{-\delta\rho} g(x_i y_i^{-\delta\rho}), \end{aligned} \tag{5.20}$$

where $\boldsymbol{x}_{(-p)} = (x_1, \ldots, x_{p-1})'$ and $\boldsymbol{y}_{(-p)} = (y_1, \ldots, y_{p-1})'$. Assume that (5.18) holds for $p-1$. Then from (5.20) we get

$$\begin{aligned} \|\Delta_{\delta,k,p}(\boldsymbol{x}, \boldsymbol{y})\|_{1;p} \leqslant {} & \eta_{\delta,k}|y_p - 1|^k \\ & + \sum_{q=0}^{k-1} \xi_{\delta,q}|y_p - 1|^q v_{\delta,k-q,p-1} \sum_{i=1}^{p-1} |y_i - 1|^{k-q}. \end{aligned} \tag{5.21}$$

It is clear we could use the same arguments to the function $\prod_{i=1, i\neq j}^{p} y_i^{-\delta\rho} g(x_i y_i^{-\delta\rho})$ with any $j = 1, 2, \ldots, p$. Then we could get (5.21) with $|y_p - 1|$ replaced by $|y_j - 1|$. Since in all these inequalities the left-hand sides will coincide, summing up the inequalities for $j = 1, 2, \ldots, p$ and using

$$\sum_{i \neq j}^{p} |y_i - 1|^{k-q}|y_j - 1|^q \leqslant (p-1) \sum_{i=1}^{p} |y_i - 1|^k, \tag{5.22}$$

we obtain

$$p\|\Delta_{\delta,k,p}(\boldsymbol{x}, \boldsymbol{y})\|_{1;p} \leqslant \left\{ \eta_{\delta,k} + (p-1) \sum_{q=0}^{k-1} v_{\delta,k-q,p-1}\xi_{\delta,q} \right\} \sum_{i=1}^{p} |y_i - 1|^k. \tag{5.23}$$

Note that (5.22) follows from (5.4). Therefore, we come to (5.18) and recurrence formula for $v_{\delta,k,p}$ stated in Theorem 3.2. □

Acknowledgments

The authors would like to thank a reviewer and editors for their careful readings and useful comments.

References

[1] T.W. Anderson, An Introduction to Multivariate Analysis, second ed., Wiley, New York, 1984.
[2] Y. Fujikoshi, Asymptotic expansions of the distributions of test statistics in multivariate analysis, J. Sci. Hiroshima Univ. Ser. A-I 34 (1970) 73–144.
[3] Y. Fujikoshi, Error bounds for asymptotic approximations of some distribution functions, in: C.R. Rao (Ed.), Multivariate Analysis: Future Directions, North-Holland Publishing Company, Amsterdam, 1993, pp. 181–208.
[4] Y. Fujikoshi, R. Shimizu, Asymptotic expansions of some mixtures of univariate and multivariate distributions, J. Multivariate Anal. 30 (1989) 279–291.
[5] Y. Fujikoshi, R. Shimizu, Asymptotic expansions of some mixtures of the multivariate normal distribution and their error bounds, Ann. Statist. 17 (1989) 1124–1132.
[6] P. Hall, On measures of the distance of a mixture from its parent distribution, Stochastic Process. Appl. 8 (1979) 357–365.
[7] A.I. Khuri, Advanced Calculus with Applications in Statistics, second ed., Wiley, New York, 2003.
[8] S.T. Rachev, Probability Metrics and Stability of Stochastic Models, Wiley, New York, 1991.
[9] R. Shimizu, Expansion of the scale mixture of the multivariate normal distribution, J. Multi. Anal. 53 (1) (1995) 126–138.
[10] R. Shimizu, Y. Fujikoshi, Sharp error bounds for asymptotic expansions of the distribution functions of scale mixtures, Ann. Inst. Statist. Math. 49 (1997) 285–297.
[11] V.V. Ulyanov, Y. Fujikoshi, R. Shimizu, Nonuniform error bounds in asymptotic expansions for scale mixtures under mild moment conditions, J. Math. Sci. 93 (1999) 600–608.