数理解析研究所講究録
第 1560 巻 2007 年 80-101

Asymptotic Second-Order Consistency For Fixed-Size Estimation and Its Applications

筑波大学大学院・数理物質科学研究科　矢田　和善 (Kazuyoshi Yata)
Graduate School of Pure and Applied Sciences
University of Tsukuba

筑波大学・数学系　青嶋　誠 (Makoto Aoshima)
Institute of Mathematics
University of Tsukuba

ABSTRACT

We consider fixed-size estimation for a linear function of means from independent and normally distributed populations having unknown and respective variances. We construct a fixed-width confidence interval with required accuracy about the magnitude of the length and the confidence coefficient. We propose a two-stage estimation methodology having the asymptotic second-order consistency with the required accuracy. The key is the asymptotic second-order analysis about the risk function. We give a variety of asymptotic characteristics about the estimation methodology, such as asymptotic sample size and asymptotic Fisher-information. With the help of the asymptotic second-order analysis, we also explore a number of generalizations and extensions of the two-stage methodology to such as bounded risk point estimation, multiple comparisons among components between the populations, and power analysis in equivalence tests to plan the appropriate sample size for a study.

Keywords: Bounded risk; Confidence interval; Efficiency; Equivalence tests; Fisher information; Multiple comparisons; Sample size determination; Second-order consistency; Two-stage estimation.

1. INTRODUCTION

Suppose that there exist k independent and normally distributed populations $\pi_i:\ N(\mu_i,\sigma_i^2)$, $i=1,...,k$, where μ_i's and σ_i^2's are unknown. Let $X_{i1},X_{i2},...$ be a sequence of independent and identically distributed random variables from each π_i. Having recorded $X_{i1},...,X_{in_i}$ for each π_i, let us write $\overline{X}_{in_i}=\sum_{j=1}^{n_i}X_{ij}/n_i$ and $\boldsymbol{n}=(n_1,...,n_k)$. We are interested in estimating the linear function $\mu=\sum_{i=1}^{k}b_i\mu_i$, where b_i's are known and nonzero scalars. Let $T_{\mathbf{n}}=\sum_{i=1}^{k}b_i\overline{X}_{in_i}$. We want to construct a fixed-width confidence interval such that

$$P_{\boldsymbol{\theta}}(|T_{\mathbf{n}}-\mu|<d)\geq 1-\alpha \tag{1.1}$$

for all $\boldsymbol{\theta}=(\mu_1,...,\mu_k,\sigma_1^2,...,\sigma_k^2)$, where $d\ (>0)$ and $\alpha\in(0,1)$ are both prespecified. Since

$$P_{\boldsymbol{\theta}}(|T_{\mathbf{n}}-\mu|<d)=G\left(d^2\left(\sum_{i=1}^{k}\frac{b_i^2\sigma_i^2}{n_i}\right)^{-1}\right) \tag{1.2}$$

with $G(\cdot)$ the cumulative distribution function (c.d.f.) of a chi–square random variable having one degree of freedom (d.f.), requirement (1.1) is satisfied if

$$d^2 \left(\sum_{i=1}^{k} \frac{b_i^2 \sigma_i^2}{n_i} \right)^{-1} \geq a, \tag{1.3}$$

where a is the constant such that $G(a) = 1 - \alpha$. It is easy to see that the sample sizes $\boldsymbol{n}$ which minimize the sum $\sum_{i=1}^{k} n_i$ subject to (1.3) are given as the smallest integer such that

$$n_i \geq \frac{a}{d^2} |b_i| \sigma_i \sum_{j=1}^{k} |b_j| \sigma_j \quad (= C_i, \text{ say}) \tag{1.4}$$

for each π_i. However, since σ_i's are unknown, the optimal fixed–sample–sizes C_i's should be estimated by using pilot samples from every π_i. It should be noted from Dantzig (1940) that any fixed–sample–size design cannot claim requirement (1.1).

Takada and Aoshima (1997) gave a two–stage estimation methodology in the spirit of Stein (1945) to satisfy requirement (1.1) for all the parameters. For the two–sample problem, see Banerjee (1967), Schwabe (1995) and Takada and Aoshima (1996). However, it tends to be oversampling especially when the pilot sample is fixed small campared to the size of C_i. Later, Takada (2004) gave a modification of the Takada–Aoshima procedure so as to make it *asymptotically second–order efficient*, i.e., $\limsup_{d \to 0} E_{\boldsymbol{\theta}}(N_i - C_i) < \infty$. Such a modification had been created and explored for the one–sample problem and the other problems by Mukhopadhyay and Duggan (1997, 1999), Aoshima and Takada (2000), and Aoshima and Mukhopadhyay (2002) among others. One may also refer to Aoshima (2005) for a review of this field.

Here, we summarize a modified two–stage procedure due to Takada (2004): Along the lines of Mukhopadhyay and Duggan (1997) and Takada (2004), we assume that there exists a known and positive lower bound $\sigma_{i\star}$ for σ_i such that

$$\sigma_i > \sigma_{i\star}, \quad i = 1, ..., k. \tag{1.5}$$

(T1) Having m_0 (≥ 4) fixed, define

$$m = \max \left\{ m_0, \left[\frac{a}{d^2} \min_{1 \leq i \leq k} |b_i| \sigma_{i\star} \sum_{j=1}^{k} |b_j| \sigma_{j\star} \right] + 1 \right\}, \tag{1.6}$$

where $[x]$ denotes the largest integer less than x. Take a pilot sample $X_{i1}, ..., X_{im}$ of size m and calculate $S_i^2 = \sum_{j=1}^{m} (X_{ij} - \overline{X}_{im})^2 / \nu$ for each π_i, where $\overline{X}_{im} = \sum_{j=1}^{m} X_{ij}/m$ and $\nu = m - 1$. Define the total sample size of each π_i by

$$N_i = \max \left\{ m, \left[\frac{u}{d^2} |b_i| S_i \sum_{j=1}^{k} |b_j| S_j \right] + 1 \right\}, \tag{1.7}$$

where the design constant u (> 0) is chosen as

$$u = a \left(1 + \frac{a + 2k - 1}{2\nu} \right). \tag{1.8}$$

Let $\boldsymbol{N} = (N_1, ..., N_k)$.

(T2) Take an additional sample $X_{im+1}, ..., X_{iN_i}$ of size $N_i - m$ from each π_i. By combining the initial sample and the additional sample, calculate $\overline{X}_{iN_i} = N_i^{-1}\sum_{j=1}^{N_i} X_{ij}$ for each π_i. Finally, construct the fixed-width confidence interval with $T_{\mathbf{N}} = \sum_{i=1}^k b_i \overline{X}_{iN_i}$.

Then, it holds as $d \to 0$ that

$$P_{\boldsymbol{\theta}}(|T_{\mathbf{N}} - \mu| < d) \geq 1 - \alpha + o(d^2) \quad \text{for all } \boldsymbol{\theta}. \tag{1.9}$$

In this paper, we give a different method to choose the constant u in (1.7). This method aims at making it *asymptotically second-order consistent* with the required accuracy as $d \to 0$, i.e.,

$$P_{\boldsymbol{\theta}}(|T_{\mathbf{N}} - \mu| < d) = 1 - \alpha + o(d^2) \quad \text{for all } \boldsymbol{\theta}. \tag{1.10}$$

With such the constant u, the required sample size is drastically reduced when compared with (1.8). The key is the asymptotic second-order analysis about the risk function. In Section 2, we show the asymptotic second-order consistency for such the modified two-stage procedure along with its asymptotic second-order characteristics. Also, we discuss asymptotic Fisher-information in the modified two-stage estimation methodology. In Section 3, with the help of the asymptotic second-order analysis, we explore a number of generalizations and extensions of the modified two-stage methodology to such as bounded risk point estimation, multiple comparisons among components between the populations, and power analysis in equivalence tests to plan the appropriate sample size for a study.

2. ASYMPTOTIC SECOND-ORDER CONSISTENCY

Throughout this section, we write that

$$\tau_\star = \min_{1 \leq i \leq k} |b_i|\sigma_{i\star} \sum_{j=1}^k |b_j|\sigma_{j\star}, \quad f_i = |b_i|\sigma_i \left(\sum_{j=1}^k |b_j|\sigma_j\right)^{-1} \quad (i = 1, ..., k).$$

Theorem 2.1. *Choose u in (1.7) as $u = a(1 + \nu^{-1}\hat{s})$ instead of (1.8), where*

$$\hat{s} = 1 + \frac{(a-1)\sum_{i=1}^k b_i^2 S_i^2 - k\tau_\star}{2(\sum_{i=1}^k |b_i| S_i)^2} \tag{2.1}$$

with S_i^2's calculated in (T1). Then, the modified two-stage procedure (1.6)-(1.7) is asymptotically second-order consistent as $d \to 0$ as stated in (1.10).

Proof of Theorem 2.1. We have from (1.2) that

$$\begin{aligned} P_{\boldsymbol{\theta}}(|T_{\mathbf{N}} - \mu| < d) &= E_{\boldsymbol{\theta}}\left\{G\left(d^2\left(\sum_{i=1}^k \frac{b_i^2\sigma_i^2}{N_i}\right)^{-1}\right)\right\} \\ &= E_{\boldsymbol{\theta}}\left\{G\left(a\left(\sum_{i=1}^k f_i\frac{C_i}{N_i}\right)^{-1}\right)\right\}. \end{aligned} \tag{2.2}$$

Now, let us define a new function as follows. We write

$$g(u_1, ..., u_k) = G(av^{-1}), \quad v = f_1 u_1^{-1} + \cdots + f_k u_k^{-1} \quad \text{for } u_i > 0, \ i = 1, ..., k.$$

Denoting $G'(w)$, $G''(w)$ for the first and second derivatives of $G(w)$ respectively, one can verify the following expressions of the partial derivatives of $g(u_1, ..., u_k)$. For all $1 \le i \ne j \le k$, we have that

$$\frac{\partial g}{\partial u_i} = aG'(a/v)f_i v^{-2} u_i^{-2},$$

$$\frac{\partial^2 g}{\partial u_i^2} = a\{aG''(a/v)f_i^2 v^{-4} u_i^{-4} + 2G'(a/v)f_i^2 v^{-3} u_i^{-4} - 2G'(a/v)f_i v^{-2} u_i^{-3}\},$$

$$\frac{\partial^2 g}{\partial u_i \partial u_j} = a\{aG''(a/v)f_i f_j v^{-4} u_i^{-2} u_j^{-2} + 2G'(a/v)f_i f_j v^{-3} u_i^{-2} u_j^{-2}\}.$$

From (2.2), we use the Taylor expansion to claim that

$$\begin{aligned} P_{\boldsymbol{\theta}}(|T_{\mathbf{N}} - \mu| < d) &= E_{\boldsymbol{\theta}}\left\{g\left(\frac{N_1}{C_1}, ..., \frac{N_k}{C_k}\right)\right\} \\ &= 1 - \alpha + aG'(a)\sum_{i=1}^{k} f_i E_{\boldsymbol{\theta}}\left(\frac{N_i - C_i}{C_i}\right) \\ &\quad + \frac{a}{2}\sum_{i=1}^{k}(aG''(a)f_i^2 + 2G'(a)f_i^2 - 2G'(a)f_i)E_{\boldsymbol{\theta}}\left\{\left(\frac{N_i - C_i}{C_i}\right)^2\right\} \\ &\quad + \frac{a}{2}\sum_{i \ne j}(aG''(a)f_i f_j + 2G'(a)f_i f_j)E_{\boldsymbol{\theta}}\left\{\left(\frac{N_i - C_i}{C_i}\right)\left(\frac{N_j - C_j}{C_j}\right)\right\} \\ &\quad + E_{\boldsymbol{\theta}}(\Re), \end{aligned} \tag{2.3}$$

where

$$E_{\boldsymbol{\theta}}(\Re) = \frac{1}{6}\sum_{1 \le i \le j \le \ell \le k} E_{\boldsymbol{\theta}}\left\{\frac{\partial^3 g}{\partial u_i \partial u_j \partial u_\ell}\Big|_{\mathbf{u}=\boldsymbol{\xi}}\left(\frac{N_i - C_i}{C_i}\right)\left(\frac{N_j - C_j}{C_j}\right)\left(\frac{N_\ell - C_\ell}{C_\ell}\right)\right\} \tag{2.4}$$

with suitable random variables ξ_i's between 1 and N_i/C_i, $i = 1, ..., k$, $\mathbf{u} = (u_1, ..., u_k)$ and $\boldsymbol{\xi} = (\xi_1, ..., \xi_k)$. With the help of Lemmas 5 and 6 in Appendix, we obtain the following expansion from (2.3):

$$\begin{aligned} P_{\boldsymbol{\theta}}(|T_{\mathbf{N}} - \mu| < d) &= 1 - \alpha \\ &+ \frac{aG'(a)}{\nu}\left(s - 1 + \frac{1}{2}\sum_{i=1}^{k} f_i B_i + \sum_{i=1}^{k} f_i^2 + a\frac{G''(a)}{G'(a)}\sum_{i=1}^{k} f_i^2\right) + o(\nu^{-1}), \end{aligned} \tag{2.5}$$

where $B_i = C_i^{-1}\nu$ and s is a constant such that $E_{\boldsymbol{\theta}}(\hat{s}) = s + o(1)$. Combining the results that $aG''(a)/G'(a) = (-a-1)/2$ and $\sum_{i=1}^{k} f_i B_i = k\tau_\star(\sum_{i=1}^{k}|b_i|\sigma_i)^{-2} + O(d^2)$ with (2.5), we claim assertion (1.10) as $d \to 0$. □

Remark 1. From Lemma 2 in Takada (2004), the constant u given by (1.8) is coincident with the one originally given by Takada and Aoshima (1997) upto the order $O(\nu^{-1})$. For the modified two-stage procedure (1.6)–(1.7) with (1.8), by putting $s = (a + 2k - 1)/2$ in (2.5), one has as $d \to 0$ that

$$\begin{aligned} P_{\boldsymbol{\theta}}(|T_{\mathbf{N}} - \mu| < d) &= 1 - \alpha \\ &+ \frac{aG'(a)}{2\nu}\left(a + 2k - 3 + \frac{k\tau_\star + (1-a)\sum_{i=1}^{k} b_i^2\sigma_i^2}{(\sum_{i=1}^{k}|b_i|\sigma_i)^2}\right) + o(d^2) \text{ for all } \boldsymbol{\theta}. \end{aligned} \tag{2.6}$$

Note that $\hat{s} < (a+2k-1)/2$ w.p.1. The use of (2.1) saves more samples when k is large.

Theorem 2.2. *The two-stage procedure (1.6)-(1.7) with (2.1) has as $d \to 0$:*

(i) $E_{\theta}(N_i - C_i) = (2\tau_\star)^{-1}\{\,|b_i|\sigma_i \sum_{j=1}^k |b_j|\sigma_j + (a-1)f_i \sum_{j=1}^k b_j^2\sigma_j^2 + b_i^2\sigma_i^2\,\}$
$+\frac{1}{2}(1-kf_i) + o(1)$ *for* $i = 1, ..., k,$

(ii) $E_{\theta}(\sum_{i=1}^k N_i - \sum_{i=1}^k C_i) = (2\tau_\star)^{-1}\{\,(\sum_{i=1}^k |b_i|\sigma_i)^2 + a\sum_{i=1}^k b_i^2\sigma_i^2\,\} + o(1).$

Proof of Theorem 2.2. The results are obtained by Lemma 5 in Appendix straightforwardly. □

Now, we evaluate the Fisher information in the statistic $T_{\mathbf{N}}$ that is calculated in (T2) with the constant u given by (2.1). We write the Fisher information in $T_{\mathbf{N}}$ about μ as $\mathcal{F}_{T_{\mathbf{N}}}(\mu)$.

Theorem 2.3. *The modified two-stage procedure (1.6)-(1.7) with (2.1) has the Fisher information in $T_{\mathbf{N}}$ as $d \to 0$:*

$$\frac{\mathcal{F}_{T_{\mathbf{N}}}(\mu)}{\mathcal{F}_{T_{\mathbf{C}}}(\mu)} = 1 + \frac{d^2(a+1)\sum_{i=1}^k b_i^2\sigma_i^2}{2a\tau_\star(\sum_{i=1}^k |b_i|\sigma_i)^2} + o(d^2), \tag{2.7}$$

where $\mathbf{C} = (C_1, ..., C_k)$ is defined by (1.4).

Proof of Theorem 2.3. In a way similar to Theorem 2.1 in Mukhopadhyay (2005), we have that

$$\begin{aligned}\mathcal{F}_{T_{\mathbf{N}}}(\mu) &= E_{\theta}\left\{\left(\sum_{i=1}^k \frac{b_i^2\sigma_i^2}{N_i}\right)^{-1}\right\} \\ &= E_{\theta}\left\{\frac{a}{d^2}\left(\sum_{i=1}^k f_i\frac{C_i}{N_i}\right)^{-1}\right\}.\end{aligned} \tag{2.8}$$

Then, one has that $\mathcal{F}_{T_{\mathbf{C}}}(\mu) = (\sum_{i=1}^k b_i^2\sigma_i^2/C_i)^{-1} = ad^{-2}$. So, we may write that

$$\frac{\mathcal{F}_{T_{\mathbf{N}}}(\mu)}{\mathcal{F}_{T_{\mathbf{C}}}(\mu)} = E_{\theta}\left\{\left(\sum_{i=1}^k f_i\frac{C_i}{N_i}\right)^{-1}\right\}. \tag{2.9}$$

From (2.9), we use the Taylor expansion to claim that

$$\begin{aligned}\frac{\mathcal{F}_{T_{\mathbf{N}}}(\mu)}{\mathcal{F}_{T_{\mathbf{C}}}(\mu)} &= 1 + \sum_{i=1}^k f_i E_{\theta}\left(\frac{N_i - C_i}{C_i}\right) + \sum_{i=1}^k (f_i^2 - f_i) E_{\theta}\left\{\left(\frac{N_i - C_i}{C_i}\right)^2\right\} \\ &\quad + \sum_{i \neq j} f_i f_j E_{\theta}\left\{\left(\frac{N_i - C_i}{C_i}\right)\left(\frac{N_j - C_j}{C_j}\right)\right\} + E_{\theta}(\Re),\end{aligned} \tag{2.10}$$

where

$$E_{\theta}(\Re) = \frac{1}{6}\sum_{1 \le i \le j \le \ell \le k} E_{\theta}\left\{\frac{\partial^3 v^{-1}}{\partial u_i \partial u_j \partial u_\ell}\Big|_{u=\xi}\left(\frac{N_i - C_i}{C_i}\right)\left(\frac{N_j - C_j}{C_j}\right)\left(\frac{N_\ell - C_\ell}{C_\ell}\right)\right\} \tag{2.11}$$

with $v = \sum_{i=1}^{k} f_i u_i^{-1}$ for $u_i > 0$, $i = 1, ..., k$, suitable random variables ξ_i's between 1 and N_i/C_i, $i = 1, ..., k$, $\boldsymbol{u} = (u_1, ..., u_k)$ and $\boldsymbol{\xi} = (\xi_1, ..., \xi_k)$. With the help of Lemmas 5 and 6 in Appendix, we obtain the following expansion from (2.10):

$$\frac{\mathcal{F}_{T_\mathbf{N}}(\mu)}{\mathcal{F}_{T_C}(\mu)} = 1 + \nu^{-1}\left(s - 1 + \sum_{i=1}^{k} f_i^2 + \frac{1}{2}\sum_{i=1}^{k} f_i B_i\right) + o(\nu^{-1}), \tag{2.12}$$

where $B_i = C_i^{-1}\nu$ and s is a constant such that $E_{\boldsymbol{\theta}}(\hat{s}) = s + o(1)$. Combining the result that $\sum_{i=1}^{k} f_i B_i = k\tau_\star(\sum_{i=1}^{k} |b_i|\sigma_i)^{-2} + O(d^2)$ with (2.12), we claim assertion (2.7) as $d \to 0$. □

Remark 2. For simplicity, we let $k = 1$ $(b = 1)$. Then, $C = a\sigma^2/d^2$. Under the assumption that $\mathcal{F}_{\overline{X}_N}(\mu)$ exceeds $\mathcal{F}_{\overline{X}_C}(\mu)$ for every fixed (μ, σ^2), Mukhopadhyay (2005) proposed to determine the pilot sample size m for Stein's (1945) two–stage estimation methodology as

$$m = \text{smallest positive integer such that } \mathcal{F}_{\overline{X}_N}(\mu)/\mathcal{F}_{\overline{X}_C}(\mu) \le 1 + \varepsilon \tag{2.13}$$

for a prespecified quantity ε (> 0) which is free from (μ, σ^2). Mukhopadhyay showed that $\mathcal{F}_{\overline{X}_N}(\mu) = \sigma^{-2}E_{\sigma^2}(N)$ and suggested that one may determine the pilot sample size m as

$$m = \text{smallest positive integer such that } E_{\sigma^2}(N)/C \le 1 + \varepsilon + o(m^{-1}). \tag{2.14}$$

Let us write that $E_{\sigma^2}(N)/C = 1 + x/m + o(m^{-1})$ with the design constant $u = a(1 + s/m) + O(m^{-2})$ where x is a constant free from m and $s = (a+1)/2$ for Stein's methodology. If m is completely free from σ^2, we should choose m in order $O(d^c)$ with $c \in (-1, 0)$ in order to specify quantity ε free from σ^2. Then, we have that $x = s$, so that $m = s/\varepsilon$ which is exactly the one given by (3.7) in Mukhopadhyay (2005). Now, let us say $c = -0.5$ and choose m in order $O(d^{-1/2})$. Let us simply write $m = sd^{-1/2}$. Then, we have that $\varepsilon = s/m = d^{1/2}$. When ε is specified as $\varepsilon = 0.1$ (0.01), we have that $d = 10^{-2}$ (10^{-4}), so that C should be very large. It would cause oversampling in the two–stage estimation methodology.

Remark 3. From (2.7), we have as $d \to 0$ that

$$\mathcal{F}_{T_\mathbf{N}}(\mu)/\mathcal{F}_{T_C}(\mu) \le 1 + \varepsilon + o(m^{-1}),$$

with $\varepsilon = (2a\tau_\star)^{-1}(a+1)d^2$. On the other hand, from (2.12) with $s = (a + 2k - 1)/2$, which is coincide with the one for Stein's (1945) methodology for $k = 1$, the modified two–stage procedure (1.6)–(1.7) with (1.8) has the Fisher information in $T_\mathbf{N}$ as $d \to 0$:

$$\frac{\mathcal{F}_{T_\mathbf{N}}(\mu)}{\mathcal{F}_{T_C}(\mu)} = 1 + \frac{d^2}{2a\tau_\star}\left(a + 2k - 3 + \frac{2\sum_{i=1}^{k} b_i^2\sigma_i^2 + k\tau_\star}{(\sum_{i=1}^{k} |b_i|\sigma_i)^2}\right) + o(d^2). \tag{2.15}$$

From (2.15), we have $\varepsilon = (2a\tau_\star)^{-1}(a+3k-1)d^2$. It should be noted that the ε part (redundancy) becomes small when we utilize (2.1) instead of (1.8).

Remark 4. If we choose u in (1.7) as $u = a(1 + \nu^{-1}\hat{s})$ with

$$\hat{s} = 1 - \frac{2\sum_{i=1}^{k} b_i^2 S_i^2 + k\tau_\star}{2(\sum_{i=1}^{k} |b_i|S_i)^2} \tag{2.16}$$

instead of (2.1), the modified two–stage procedure (1.6)–(1.7) has the Fisher information in $T_{\mathbf{N}}$ as $d \to 0$:

$$\mathcal{F}_{T_{\mathbf{N}}}(\mu)/\mathcal{F}_{T_C}(\mu) = 1 + o(m^{-1}). \tag{2.17}$$

Then, it holds as $d \to 0$:

(i) $E_{\boldsymbol{\theta}}(N_i - C_i) = (2\tau_\star)^{-1}\{\, |b_i|\sigma_i \sum_{j=1}^k |b_j|\sigma_j - (2\sum_{j=1}^k b_j^2\sigma_j^2 + k\tau_\star)f_i + b_i^2\sigma_i^2\,\}$
$+\frac{1}{2} + o(1)$ for $i = 1, ..., k$,

(ii) $E_{\boldsymbol{\theta}}(\sum_{i=1}^k N_i - \sum_{i=1}^k C_i) = (2\tau_\star)^{-1}\{\, (\sum_{i=1}^k |b_i|\sigma_i)^2 - \sum_{i=1}^k b_i^2\sigma_i^2\} + o(1)$.

3. APPLICATIONS

3.1. Bounded risk estimation

Suppose that there exist k independent and normally distributed populations $\pi_i : N_p(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i)$, $i = 1, ..., k$, where $\boldsymbol{\mu}_i$'s $\in R^p$ and $\boldsymbol{\Sigma}_i$'s are both unknown, but $\boldsymbol{\Sigma}_i$'s are $p \times p$ p.d. matrices. Let $\boldsymbol{X}_{i1}, \boldsymbol{X}_{i2}, ...$ be a sequence of independent and identically distributed random vectors from each π_i. Having recorded $\boldsymbol{X}_{i1}, ..., \boldsymbol{X}_{in_i}$ for each π_i, let us write $\overline{\boldsymbol{X}}_{in_i} = \sum_{j=1}^{n_i} \boldsymbol{X}_{ij}/n_i$ and $\boldsymbol{n} = (n_1, ..., n_k)$. We are interested in estimating the linear function $\boldsymbol{\mu} = \sum_{i=1}^k b_i\boldsymbol{\mu}_i$, where b_i's are known and nonzero scalars. Let $\boldsymbol{T}_{\mathbf{n}} = \sum_{i=1}^k b_i\overline{\boldsymbol{X}}_{in_i}$. For a prespecified constant W (> 0), we want to construct $\boldsymbol{T}_{\mathbf{n}}$ such that

$$E_{\boldsymbol{\theta}}(||\boldsymbol{T}_{\mathbf{n}} - \boldsymbol{\mu}||^2) \le W \tag{3.1}$$

for all $\boldsymbol{\theta} = (\boldsymbol{\mu}_1, ..., \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_1, ..., \boldsymbol{\Sigma}_k)$, where $||\cdot||$ is the Euclidean norm. Since

$$E_{\boldsymbol{\theta}}(||\boldsymbol{T}_{\mathbf{n}} - \boldsymbol{\mu}||^2) = \sum_{i=1}^k b_i^2 \mathrm{tr}(\boldsymbol{\Sigma}_i)/n_i, \tag{3.2}$$

it is easy to see that the sample sizes $\boldsymbol{n}$ which minimize the sum $\sum_{i=1}^k n_i$ subject to (3.1) are given as the smallest integer such that

$$n_i \ge \frac{1}{W}|b_i|\sqrt{\mathrm{tr}(\boldsymbol{\Sigma}_i)}\sum_{j=1}^k |b_j|\sqrt{\mathrm{tr}(\boldsymbol{\Sigma}_j)} \quad (= C_i, \text{ say}) \tag{3.3}$$

for each π_i.

When $p = 1$, Ghosh et al. (1997, Chap. 6) considered a two–stage estimation methodology to satisfy requirement (3.1). Later, Aoshima and Takada (2002) considered the present problem and gave a different two–stage estimation methodology. Aoshima and Takada showed that their procedure satisfies requirement (3.1) with fewer samples than those in Ghosh et al. When applying the asymptotic second–order analysis to the present problem, we give a modified two–stage estimation methodology to hold the asymptotic second–order consistency as $W \to 0$ as stated in (3.8): We assume that there exists a known and positive lower bound $\sigma_{i\star}$ for $\mathrm{tr}(\boldsymbol{\Sigma}_i)$ such that

$$\mathrm{tr}(\boldsymbol{\Sigma}_i) > \sigma_{i\star}, \quad i = 1, ..., k. \tag{3.4}$$

(T1) Having m_0 (≥ 4) fixed, define

$$m = \max\left\{m_0, \left[\frac{1}{W}\min_{1\leq i\leq k}|b_i|\sigma_{i\star}\sum_{j=1}^{k}|b_j|\sigma_{j\star}\right]+1\right\}. \tag{3.5}$$

Take a pilot sample $\boldsymbol{X}_{i1}, ..., \boldsymbol{X}_{im}$ of size m and calculate $\boldsymbol{S}_i = \sum_{j=1}^{m}(\boldsymbol{X}_{ij}-\overline{\boldsymbol{X}}_{im})(\boldsymbol{X}_{ij}-\overline{\boldsymbol{X}}_{im})'/\nu$ for each π_i, where $\overline{\boldsymbol{X}}_{im} = \sum_{j=1}^{m}\boldsymbol{X}_{ij}/m$ and $\nu = m-1$. Define the total sample size of each π_i by

$$N_i = \max\left\{m, \left[\frac{u}{W}|b_i|\sqrt{\mathrm{tr}(\boldsymbol{S}_i)}\sum_{j=1}^{k}|b_j|\sqrt{\mathrm{tr}(\boldsymbol{S}_j)}\right]+1\right\}, \tag{3.6}$$

where u is chosen as $u = 1+\nu^{-1}\hat{s}$ with $\hat{s}$ given by (3.7). Let $\boldsymbol{N} = (N_1, ..., N_k)$.

(T2) Take an additional sample $\boldsymbol{X}_{im+1}, ..., \boldsymbol{X}_{iN_i}$ of size $N_i - m$ from each π_i. By combining the initial sample and the additional sample, calculate $\overline{\boldsymbol{X}}_{iN_i} = N_i^{-1}\sum_{j=1}^{N_i}\boldsymbol{X}_{ij}$ for each π_i. Finally, estimate $\boldsymbol{\mu}$ by $\boldsymbol{T}_{\mathbf{N}} = \sum_{i=1}^{k} b_i\overline{\boldsymbol{X}}_{iN_i}$.

Theorem 3.1. *Let $\tau_\star = \min_{1\leq i\leq k}|b_i|\sigma_{i\star}\sum_{j=1}^{k}|b_j|\sigma_{j\star}$, where $\sigma_{i\star}$ is given by (3.4). Choose u in (3.6) as $u = 1+\nu^{-1}\hat{s}$, where*

$$\hat{s} = \frac{2\sum_{i=1}^{k}(\mathrm{tr}(\boldsymbol{S}_i^2)/(\mathrm{tr}(\boldsymbol{S}_i))^2)\left(b_i^2\mathrm{tr}(\boldsymbol{S}_i)+|b_i|\sqrt{\mathrm{tr}(\boldsymbol{S}_i)}\sum_{j=1}^{k}|b_j|\sqrt{\mathrm{tr}(\boldsymbol{S}_j)}\right)-k\tau_\star}{2\left(\sum_{i=1}^{k}|b_i|\sqrt{\mathrm{tr}(\boldsymbol{S}_i)}\right)^2} \tag{3.7}$$

with $\boldsymbol{S}_i$'s calculated in (T1). Then, the modified two-stage procedure (3.5)–(3.6) is asymptotically second-order consistent as $W\to 0$, i.e.,

$$E_{\boldsymbol{\theta}}(||\boldsymbol{T}_{\mathbf{N}}-\boldsymbol{\mu}||^2) = W + o(W^2) \quad \textit{for all } \boldsymbol{\theta}. \tag{3.8}$$

Proof of Theorem 3.1. We have from (3.2) that

$$\begin{aligned} E_{\boldsymbol{\theta}}(\|\boldsymbol{T}_{\mathbf{N}}-\boldsymbol{\mu}\|)^2 =& E_{\boldsymbol{\theta}}\left(\sum_{i=1}^{k}b_i^2\mathrm{tr}(\boldsymbol{\Sigma}_i)/N_i\right) \\ =& WE_{\boldsymbol{\theta}}\left(\sum_{i=1}^{k}f_i\frac{C_i}{N_i}\right), \end{aligned}$$

where $f_i = |b_i|\sqrt{\mathrm{tr}(\boldsymbol{\Sigma}_i)}/\sum_{j=1}^{k}|b_j|\sqrt{\mathrm{tr}(\boldsymbol{\Sigma}_j)}$. Use the Taylor expansion to claim that

$$E_{\boldsymbol{\theta}}\left(\sum_{i=1}^{k}f_i\frac{C_i}{N_i}\right) = 1-\sum_{i=1}^{k}f_iE_{\boldsymbol{\theta}}\left(\frac{N_i-C_i}{C_i}\right)+\sum_{i=1}^{k}f_iE_{\boldsymbol{\theta}}\left\{\left(\frac{N_i-C_i}{C_i}\right)^2\right\}+E_{\boldsymbol{\theta}}(\Re), \tag{3.9}$$

where $E_{\boldsymbol{\theta}}(\Re) = -\sum_{i=1}^{k}f_iE_{\boldsymbol{\theta}}\left\{\xi_i^{-4}C_i^{-3}(N_i-C_i)^3\right\}$ with suitable random variables ξ_i's between 1 and N_i/C_i, $i = 1, ..., k$. One may apply Lemma 6 in Appendix to claim that $E_{\boldsymbol{\theta}}(\Re) = o(\nu^{-1})$ as $W\to 0$. With the help of Remark 12 in Appendix, we obtain the following expansion from (3.9):

$$E_{\boldsymbol{\theta}}\left(\sum_{i=1}^{k}f_i\frac{C_i}{N_i}\right) = 1+\frac{1}{2\nu}\sum_{i=1}^{k}f_i\left(-2s-B_i+A_i\left(f_i+\frac{3}{2}\right)+\sum_{j=1}^{k}f_jA_j\left(f_j+\frac{1}{2}\right)\right)+o(\nu^{-1}), \tag{3.10}$$

where $A_i = \mathrm{tr}(\Sigma_i^2)/(\mathrm{tr}(\Sigma_i))^2$, $B_i = \nu C_i^{-1}$, and s is a constant such that $E_{\theta}(\hat{s}) = s + o(1)$. From (3.10), we obtain (3.8) straightforwardly. □

Remark 5. Aoshima and Takada (2002) gave a two-stage estimation methodology to satisfy requirement (3.1) without assumption (3.4). In their methodology, the constant u is given by $u = \nu/(\nu-2) = 1 + 2/\nu + O(\nu^{-2})$. From (3.7), note that $\hat{s} < 2$ w.p.1. The use of (3.7) saves more samples when k is large.

3.2. Multiple comparisons among components

Suppose that there exist k independent and normally distributed populations $\pi_i : N_p(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i)$, $i = 1, ..., k$, where $p \geq 2$, and $\boldsymbol{\mu}_i$'s $\in R^p$ and $\boldsymbol{\Sigma}_i$'s are both unknown, but $\boldsymbol{\Sigma}_i = (\sigma_{(i)rs})$ $(> \mathbf{0})$ has a spherical structure such that

$$\sigma_{(i)rr} + \sigma_{(i)ss} - 2\sigma_{(i)rs} = 2\delta_i^2 \quad (1 \leq r < s \leq p) \tag{3.11}$$

with δ_i (> 0) unknown parameter for each π_i. A special case of such the model is the intraclass correlation model, that is, $\boldsymbol{\Sigma}_i = \sigma_i^2\{(1-\rho_i)\boldsymbol{I}_p + \rho_i \boldsymbol{J}\}$ for some ρ_i, where $\boldsymbol{J}$ denotes a $p \times p$ matrix of all 1's. We consider multiple comparisons experiments for correlated components of $\boldsymbol{\mu} = \sum_{i=1}^k b_i \boldsymbol{\mu}_i$. Let us write $\boldsymbol{\mu} = (\xi_1, ..., \xi_p)$. Similarly to Section 3.1, we use $\boldsymbol{T}_{\mathbf{n}} = \sum_{i=1}^k b_i \overline{\boldsymbol{X}}_{in_i}$ as an estimate of $\boldsymbol{\mu}$. Let us write $\boldsymbol{T}_{\mathbf{n}} = (T_{1\mathbf{n}}, ..., T_{p\mathbf{n}})$. For a prespecified constant d (> 0), we define three types of simultaneous confidence intervals for $(\xi_1, ..., \xi_p)$:

$$\text{(MCA)}\ R_{\mathbf{n}} = \{\boldsymbol{\mu} |\ \xi_r - \xi_s \in [T_{r\mathbf{n}} - T_{s\mathbf{n}} - d,\ T_{r\mathbf{n}} - T_{s\mathbf{n}} + d],\ 1 \leq r < s \leq p\};$$

$$\text{(MCB)}\ R_{\mathbf{n}} = \{\boldsymbol{\mu} |\ \xi_r - \max_{s \neq r} \xi_s \in [-(T_{r\mathbf{n}} - \max_{s\neq r} T_{s\mathbf{n}} - d)^-,\ +(T_{r\mathbf{n}} - \max_{s \neq r} T_{s\mathbf{n}} + d)^+],$$
$$r = 1, ..., p\},$$
$$\text{where } +x^+ = \max\{0, x\} \text{ and } -x^- = \min\{0, x\};$$

$$\text{(MCC)}\ R_{\mathbf{n}} = \{\boldsymbol{\mu} |\ \xi_r - \xi_p \in [T_{r\mathbf{n}} - T_{p\mathbf{n}} - d,\ T_{r\mathbf{n}} - T_{p\mathbf{n}} + d],\ r = 1, ..., p-1\}.$$

For the details of these multiple comparisons methods, see Aoshima and Kushida (2005) and its references. For each of them, for d (> 0) and $\alpha \in (0, 1)$ both specified, we want to construct $R_{\mathbf{n}}$ such that

$$P_{\boldsymbol{\theta}}(\boldsymbol{\mu} \in R_{\mathbf{n}}) \geq 1 - \alpha \quad \text{for all } \boldsymbol{\theta} = (\boldsymbol{\mu}_1, ..., \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_1, ..., \boldsymbol{\Sigma}_k) \tag{3.12}$$

with $\boldsymbol{\Sigma}_i$'s defined by (3.11).

It is shown for MCA and MCC that

$$P_{\boldsymbol{\theta}}(\boldsymbol{\mu} \in R_{\mathbf{n}}) = G_p\left(d^2\left(\sum_{i=1}^k \frac{b_i^2\delta_i^2}{n_i}\right)^{-1}\right), \tag{3.13}$$

where $G_p(y)$ for $y > 0$ is defined by

$$G_p(y) = p\int_{-\infty}^{\infty}\{\Phi(x) - \Phi(x - \sqrt{y})\}^{p-1} d\Phi(x) \quad \text{(for MCA)}, \tag{3.14}$$

$$G_p(y) = \int_{-\infty}^{\infty}\{\Phi(x + \sqrt{y}) - \Phi(x - \sqrt{y})\}^{p-1} d\Phi(x) \quad \text{(for MCC)} \tag{3.15}$$

with $\Phi(\cdot)$ the c.d.f. of a $N(0, 1)$ random variable. It is shown for MCB that

$$P_{\boldsymbol{\theta}}(\boldsymbol{\mu} \in R_{\mathbf{n}}) \geq G_p\left(d^2\left(\sum_{i=1}^k \frac{b_i^2\delta_i^2}{n_i}\right)^{-1}\right), \tag{3.16}$$

where

$$G_p(y) = \int_{-\infty}^{\infty} \{\Phi(x+\sqrt{y})\}^{p-1} d\Phi(x). \tag{3.17}$$

So, the sample sizes $\boldsymbol{n}$ that minimize the sum $\sum_{i=1}^{k} n_i$ while satisfying requirement (3.12) are given as the smallest integer such that

$$n_i \geq \frac{a}{d^2}|b_i|\delta_i \sum_{j=1}^{k} |b_j|\delta_j \quad (= C_i, \text{ say}) \tag{3.18}$$

for each π_i, where a (> 0) is a constant such that $G_p(a) = 1-\alpha$ with $G_p(\cdot)$ defined for each method by (3.14), (3.15) or (3.17), respectively.

When applying the asymptotic second-order analysis to this problem, we give a modified two-stage estimation methodology to hold the asymptotic second-order consistency as $d \to 0$ as stated in (3.23)–(3.24): We assume that there exists a known and positive lower bound $\sigma_{i\star}$ for δ_i such that

$$\delta_i > \sigma_{i\star}, \quad i = 1, ..., k. \tag{3.19}$$

(T1) Having m_0 (≥ 4) fixed, define

$$m = \max\left\{m_0, \left[\frac{a}{d^2} \min_{1\leq i\leq k} |b_i|\sigma_{i\star} \sum_{j=1}^{k} |b_j|\sigma_{j\star}\right] + 1\right\}. \tag{3.20}$$

Take a pilot sample $\boldsymbol{X}_{ij} = (X_{ij1}, ..., X_{ijp})$, $j = 1, ..., m$, and calculate $S_{ip}^2 = \nu_p^{-1}\sum_{r=1}^{p}\sum_{j=1}^{m}(X_{ijr} - \overline{X}_{ij.} - \overline{X}_{i.r} + \overline{X}_{i..})^2$ with $\nu_p = (p-1)(m-1)$ for each π_i. Here, $\overline{X}_{ij.} = p^{-1}\sum_{r=1}^{p} X_{ijr}$, $\overline{X}_{i.r} = m^{-1}\sum_{j=1}^{m} X_{ijr}$ and $\overline{X}_{i..} = (pm)^{-1}\sum_{r=1}^{p}\sum_{j=1}^{m} X_{ijr}$. Note that $\nu_p S_{ip}^2/\delta_i^2$ is distributed as a chi-square distribution with ν_p d.f. Define the total sample size of each π_i by

$$N_i = \max\left\{m, \left[\frac{u}{d^2}|b_i|S_{ip} \sum_{j=1}^{k} |b_j|S_{jp}\right] + 1\right\}, \tag{3.21}$$

where u is chosen as $u = a(1+\nu_p^{-1}\hat{s})$ with a given for each method and $\hat{s}$ given by (3.22). Let $\boldsymbol{N} = (N_1, ..., N_k)$.

(T2) Take an additional sample $\boldsymbol{X}_{im+1}, ..., \boldsymbol{X}_{iN_i}$ of size $N_i - m$ from each π_i. By combining the initial sample and the additional sample, calculate $\overline{\boldsymbol{X}}_{iN_i} = N_i^{-1}\sum_{j=1}^{N_i} \boldsymbol{X}_{ij}$ for each π_i. Finally, for each method, construct $R_{\mathbf{N}}$ with the components $(T_{1\mathbf{N}}, ..., T_{p\mathbf{N}})$ of $\boldsymbol{T}_{\mathbf{N}} = \sum_{i=1}^{k} b_i \overline{\boldsymbol{X}}_{iN_i}$.

The following theorem can be obtained similarly to Theorem 2.1.

Theorem 3.2. *Let $\tau_\star = \min_{1\leq i\leq k} |b_i|\sigma_{i\star}\sum_{j=1}^{k}|b_j|\sigma_{j\star}$, where $\sigma_{i\star}$ is given by (3.19). Choose u in (3.21) as $u = a(1+\nu_p^{-1}\hat{s})$ with a given for each method, where*

$$\hat{s} = 1 - \frac{2(a\frac{G_p''(a)}{G_p'(a)} + 1)\sum_{i=1}^{k} b_i^2 S_{ip}^2 + k(p-1)\tau_\star}{2(\sum_{i=1}^{k}|b_i|S_{ip})^2} \tag{3.22}$$

with S_{ip}^2's calculated in (T1). Then, the modified two-stage procedure (3.20)–(3.21) is asymptotically second-order consistent as $d \to 0$, i.e.,

$$P_{\boldsymbol{\theta}}(\boldsymbol{\mu} \in R_{\mathbf{N}}) = 1-\alpha+o(d^2) \quad \textit{for all } \boldsymbol{\theta} \quad \textit{(MCA, MCC)}; \tag{3.23}$$

$$P_{\boldsymbol{\theta}}(\boldsymbol{\mu} \in R_{\mathbf{N}}) \geq 1-\alpha+o(d^2) \quad \textit{for all } \boldsymbol{\theta} \quad \textit{(MCB)}. \tag{3.24}$$

Remark 6. The two–stage estimation methodology (3.20)–(3.21) was given by Aoshima and Kushida (2005), but they chose the constant u in (3.21) as $u = a(1+\nu_p^{-1}s)$ with $s = k-1-aG_p''(a)/G_p'(a)$. For a nominal value of α, note that $aG_p'(a)/G_p(a) \le -1$. Then, from (3.22), we have that $\hat{s} < s$ w.p.1. The use of (3.22) saves more samples when k is large.

3.3. Testing for equivalence

We consider the problem to test the equivalence of two independent normal populations π_i : $N(\mu_i, \sigma_i^2)$, $i = 1, 2$, with μ_i's and σ_i^2's both unknown. We want to design a test of

$$H_0 : |\mu| = |\mu_1 - \mu_2| \ge d \quad \text{against} \quad H_a : |\mu| < d \tag{3.25}$$

which has size α and power no less than $1-\beta$ at $|\mu| \le \gamma d$ for all $\boldsymbol{\theta} = (\mu_1, \mu_2, \sigma_1^2, \sigma_2^2)$, where α, $\beta \in (0,1)$, $\gamma \in [0,1)$, and $d > 0$ (the limit of equivalence) are four prescribed constants. So, the two populations are deemed to be equivalent if the mean difference between the two populations is smaller than d. Let us write $\overline{X}_{in_i} = \sum_{j=1}^{n_i} X_{ij}/n_i$, $i = 1, 2$, similarly to Section 1. If σ_i^2's had been known, we would take a sample from each π_i of size

$$n_i \ge \frac{\delta^2}{d^2}\sigma_i \sum_{j=1}^{2} \sigma_j \quad (= C_i, \text{ say}) \tag{3.26}$$

and test the hypothesis by

$$\text{rejecting } H_0 \iff |\overline{X}_{1n_1} - \overline{X}_{2n_2}| < \left(\sum_{i=1}^{2} \frac{\sigma_i^2}{C_i}\right)^{1/2} R\left(d\left(\sum_{i=1}^{2}\frac{\sigma_i^2}{C_i}\right)^{-1/2}\right) = \frac{dR(\delta)}{\delta}. \tag{3.27}$$

Here, the function $R(\cdot)$ is determined uniquely by the equation

$$P(|N(0,1) + x| < R(x)) = \alpha \tag{3.28}$$

with $N(0,1)$ a standard normal random variable, and $\delta = \delta(\alpha, \beta, \gamma)$ is the unique solution of the equation

$$P(|N(0,1) + \gamma\delta| < R(\delta)) = 1 - \beta. \tag{3.29}$$

When σ_i^2's are unknown but common ($\sigma_1^2 = \sigma_2^2$), Liu (2003) proposed k (≥ 3)–stage procedure having the size $\alpha + o(n^{-1})$ and the minimum power $1 - \beta + o(n^{-1})$. When applying the asymptotic second–order analysis to the present problem, we give a modified two–stage procedure to hold the asymptotic second–order consistency, which has the accuracy of the same degree as in Liu, as stated in (3.36): We assume that there exists a known and positive lower bound $\sigma_{i\star}$ for σ_i such that

$$\sigma_i > \sigma_{i\star}, \quad i = 1, 2. \tag{3.30}$$

(T1) Having m_0 (≥ 4) fixed, define

$$m = \max\left\{m_0, \left[\frac{\delta^2}{d^2} \min_{1\le i\le 2} \sigma_{i\star} \sum_{j=1}^{2} \sigma_{j\star}\right] + 1\right\}. \tag{3.31}$$

Take a pilot sample $X_{i1},...,X_{im}$ of size m and calculate $S_i^2=\sum_{j=1}^m(X_{ij}-\overline{X}_{im})^2/\nu$ with $\nu=m-1$ for each π_i. Define the total sample size of each π_i by

$$N_i=\max\left\{m,\left[\frac{u}{d^2}S_i\sum_{j=1}^2 S_j\right]+1\right\}, \tag{3.32}$$

where u is chosen as $u=\delta^2(1+\nu^{-1}\hat{s})$ with $\hat{s}$ given by (3.34).

(T2) Take an additional sample $X_{im+1},...,X_{iN_i}$ of size N_i-m from each π_i. By combining the initial sample and the additional sample, calculate $\overline{X}_{iN_i}=N_i^{-1}\sum_{j=1}^{N_i}X_{ij}$ for each π_i. Then, test the hypothesis by

$$\text{rejecting } H_0 \iff |\overline{X}_{1N_1}-\overline{X}_{2N_2}|<\sqrt{\lambda}\frac{dR(\delta)}{\delta}, \tag{3.33}$$

where λ is chosen as $\lambda=1+\nu^{-1}\hat{t}$ with $\hat{t}$ given by (3.35).

Theorem 3.3. *Let $\tau_\star=\min_{1\le i\le 2}|b_i|\sigma_{i\star}\sum_{j=1}^2|b_j|\sigma_{j\star}$, where $\sigma_{i\star}$ is given by (3.30). Choose u in (3.32) as $u=\delta^2(1+\nu^{-1}\hat{s})$ with*

$$\begin{aligned}\hat{s}=1&-\frac{\sum_{i=1}^2 S_i^2}{2(\sum_{i=1}^2 S_i)^2}-\frac{\tau_\star}{(\sum_{i=1}^2 S_i)^2}\\&+\frac{(\phi(\varepsilon_1)+\phi(\varepsilon_2))(\eta_1^3\phi(\eta_1)+\eta_2^3\phi(\eta_2))-(\phi(\eta_1)+\phi(\eta_2))(\varepsilon_1^3\phi(\varepsilon_1)+\varepsilon_2^3\phi(\varepsilon_2))}{(\phi(\varepsilon_1)+\phi(\varepsilon_2))(\eta_1\phi(\eta_1)+\eta_2\phi(\eta_2))-(\phi(\eta_1)+\phi(\eta_2))(\varepsilon_1\phi(\varepsilon_1)+\varepsilon_2\phi(\varepsilon_2))}\frac{\sum_{i=1}^2 S_i^2}{2(\sum_{i=1}^2 S_i)^2},\end{aligned} \tag{3.34}$$

where $\phi(\cdot)$ is the p.d.f. of $N(0,1)$, $\varepsilon_1=R(\delta)-\delta$, $\varepsilon_2=R(\delta)+\delta$, $\eta_1=R(\delta)-\gamma\delta$, $\eta_2=R(\delta)+\gamma\delta$, and S_i^2's are calculated in (T1). Choose λ in (3.33) as $\lambda=(1+\nu^{-1}\hat{t})$ with

$$\hat{t}=\frac{(\varepsilon_1^3\phi(\varepsilon_1)+\varepsilon_2^3\phi(\varepsilon_2))(\eta_1\phi(\eta_1)+\eta_2\phi(\eta_2))-(\varepsilon_1\phi(\varepsilon_1)+\varepsilon_2\phi(\varepsilon_2))(\eta_1^3\phi(\eta_1)+\eta_2^3\phi(\eta_2))}{2R(\delta)\left\{(\phi(\varepsilon_1)+\phi(\varepsilon_2))(\eta_1\phi(\eta_1)+\eta_2\phi(\eta_2))-(\varepsilon_1\phi(\varepsilon_1)+\varepsilon_2\phi(\varepsilon_2))(\phi(\eta_1)+\phi(\eta_2))\right\}}\frac{\sum_{i=1}^2 S_i^2}{(\sum_{i=1}^2 S_i)^2}, \tag{3.35}$$

where S_i^2's are calculated in (T1). Then, the test (3.33) of (3.25), with (3.31)-(3.32), is asymptotically second-order consistent as $d\to 0$, i.e.,

$$\textit{size}=\alpha+o(d^2)\quad\textit{and}\quad\textit{minimum power}=1-\beta+o(d^2)\quad\textit{for all }\boldsymbol{\theta}. \tag{3.36}$$

Proof of Theorem 3.3. From (3.33), we have the size at $|\mu_1-\mu_2|=d$ that

$$\begin{aligned}&E_{\boldsymbol{\theta}}\left\{\Phi\left((\sqrt{\lambda}R(\delta)-\delta)\left(\sum_{i=1}^2 f_i\frac{C_i}{N_i}\right)^{-1/2}\right)\right\}-E_{\boldsymbol{\theta}}\left\{\Phi\left(-(\sqrt{\lambda}R(\delta)+\delta)\left(\sum_{i=1}^2 f_i\frac{C_i}{N_i}\right)^{-1/2}\right)\right\}\\&=\Phi(\varepsilon_1)-\Phi(-\varepsilon_2)+\frac{1}{4\nu}\varepsilon_1\phi(\varepsilon_1)\left(2s+\frac{2R(\delta)}{\varepsilon_1}t-2+\sum_{i=1}^2 f_iB_i+\sum_{i=1}^2 f_i^2+\varepsilon_1\frac{\phi'(\varepsilon_1)}{\phi(\varepsilon_1)}\right)\\&\quad+\frac{1}{4\nu}\varepsilon_2\phi(\varepsilon_2)\left(2s+\frac{2R(\delta)}{\varepsilon_2}t-2+\sum_{i=1}^2 f_iB_i+\sum_{i=1}^2 f_i^2+\varepsilon_2\frac{\phi'(\varepsilon_2)}{\phi(\varepsilon_2)}\right)+E_{\boldsymbol{\theta}}(\Re),\end{aligned} \tag{3.37}$$

where $\phi'(\cdot)$ denotes the first derivative of $\phi(\cdot)$. Similarly, we have the minimum power at $|\mu_1 - \mu_2| = \gamma d$ that

$$E_{\theta}\left\{\Phi\left((\sqrt{\lambda}R(\delta)-\gamma\delta)\left(\sum_{i=1}^{2} f_i\frac{C_i}{N_i}\right)^{-1/2}\right)\right\} - E_{\theta}\left\{\Phi\left(-(\sqrt{\lambda}R(\delta)+\gamma\delta)\left(\sum_{i=1}^{2} f_i\frac{C_i}{N_i}\right)^{-1/2}\right)\right\}$$
$$= \Phi(\eta_1) - \Phi(-\eta_2) + \frac{1}{4\nu}\eta_1\phi(\eta_1)\left(2s + \frac{2R(\delta)}{\eta_1}t - 2 + \sum_{i=1}^{2} f_iB_i + \sum_{i=1}^{2} f_i^2 + \eta_1\frac{\phi'(\eta_1)}{\phi(\eta_1)}\right)$$
$$+ \frac{1}{4\nu}\eta_2\phi(\eta_2)\left(2s + \frac{2R(\delta)}{\eta_2}t - 2 + \sum_{i=1}^{2} f_iB_i + \sum_{i=1}^{2} f_i^2 + \eta_2\frac{\phi'(\eta_2)}{\phi(\eta_2)}\right) + E_{\theta}(\Re). \tag{3.38}$$

Here, in both (3.37)–(3.38), s and t are constants such that $E_{\theta}(\hat{s}) = s + o(1)$ and $E_{\theta}(\hat{t}) = t + o(1)$. One may apply Lemma 6 in Appendix to claim that $E_{\theta}(\Re) = o(\nu^{-1})$ as $d \to 0$ in (3.37)–(3.38). Note that $\Phi(\varepsilon_1) - \Phi(-\varepsilon_2) = \alpha$ and $\Phi(\eta_1) - \Phi(-\eta_2) = 1 - \beta$. The assertion (3.36) can be shown straightforwardly. □

Remark 7. Let us consider the case that our goal is to design a two–sided test of

$$H_0 : \mu = \mu_1 - \mu_2 = 0 \quad \text{against} \quad H_a : \mu \neq 0 \tag{3.39}$$

which has size α and power $1 - \beta$ at $|\mu| = d$ for all $\boldsymbol{\theta}$, where α, $\beta \in (0, 1)$ and $d > 0$ are three prescribed constants. If σ_i^2's had been known, we would take a sample from each π_i of size

$$n_i \geq \frac{c^2(\alpha,\beta)}{d^2}\sigma_i\sum_{j=1}^{2}\sigma_j \tag{3.40}$$

and test the hypothesis by

$$\text{rejecting } H_0 \iff |\overline{X}_{1n_1} - \overline{X}_{2n_2}| > \frac{dz_{\alpha/2}}{c(\alpha,\beta)}, \tag{3.41}$$

where z_x is the upper x point of $N(0, 1)$, and $c(\alpha, \beta)$ (> 0) is the unique solution of the equation

$$P(|N(0,1) + c(\alpha,\beta)| > z_{\alpha/2}) = 1 - \beta. \tag{3.42}$$

One may utilize the two–stage procedure described above for this goal as well after replacing $(\delta,\ R(\delta),\ \gamma)$ with $(c(\alpha,\beta),\ z_{\alpha/2},\ 0)$, respectively, in (3.31)–(3.32) and (3.34)–(3.35). Then, the test of (3.39), given by

$$\text{rejecting } H_0 \iff |\overline{X}_{1n_1} - \overline{X}_{2n_2}| > \sqrt{\lambda}\frac{dz_{\alpha/2}}{c(\alpha,\beta)}, \tag{3.43}$$

is asymptotically second–order consistent as $d \to 0$ as stated in (3.36).

Remark 8. Let us consider the case that our goal is to design a one–sided equivalence test of

$$H_0 : \mu = \mu_1 - \mu_2 \leq -d \quad \text{against} \quad H_a : \mu > -d \tag{3.44}$$

which has size α and power no less than $1-\beta$ at $\mu \geq -\gamma d$ for all $\boldsymbol{\theta}$. So, one wants to demonstrate that a treatment is no worse than a standard or one treatment is no worse than another treatment in paired comparison by amount d. If σ_i^2's had been known, we would take a sample from each π_i of size

$$n_i \geq \left(\frac{z_\alpha - z_{1-\beta}}{(1-\gamma)d}\right)^2 \sigma_i \sum_{j=1}^{2} \sigma_j, \tag{3.45}$$

and test the hypothesis by

$$\text{rejecting } H_0 \iff \overline{X}_{1n_1} - \overline{X}_{2n_2} > -d\left(\frac{\gamma z_\alpha - z_{1-\beta}}{z_\alpha - z_{1-\beta}}\right). \tag{3.46}$$

One may utilize the two–stage procedure for this goal as well. Replace δ^2 with $(z_\alpha - z_{1-\beta})^2/(1-\gamma)^2$ in (3.31) and in the choice of u of (3.32). Choose

$$\hat{s} = 1 + (z_\alpha^2 + z_{1-\beta}^2 + z_\alpha z_{1-\beta} - 1)\frac{\sum_{i=1}^2 S_i^2}{2(\sum_{i=1}^2 S_i)^2} - \frac{\tau_\star}{(\sum_{i=1}^2 S_i)^2}, \tag{3.47}$$

$$\hat{t} = z_\alpha z_{1-\beta}(z_\alpha + z_{1-\beta})\frac{1-r}{\gamma z_\alpha - z_{1-\beta}}\frac{\sum_{i=1}^2 S_i^2}{2(\sum_{i=1}^2 S_i)^2}. \tag{3.48}$$

Then, the test of (3.44), given by

$$\text{rejecting } H_0 \iff \overline{X}_{1N_1} - \overline{X}_{2N_2} > -\sqrt{\lambda} d\left(\frac{\gamma z_\alpha - z_{1-\beta}}{z_\alpha - z_{1-\beta}}\right) \tag{3.49}$$

with $\lambda = 1 + \nu^{-1}\hat{t}$, is asymptotically second–order consistent as $d \to 0$ as stated in (3.36).

4. APPENDIX

Throughout this section, we write that

$$\tau_i = |b_i|\sigma_i \textstyle\sum_{j=1}^k |b_j|\sigma_j, \quad Y_i = |b_i| S_i \sum_{j=1}^k |b_j| S_j$$

for $i = 1, ..., k$. From (1.4), we write that $C_i = a\tau_i/d^2$. Let d (> 0) go to zero thorough a sequence such that $a\tau_\star/d^2$ always remains an integer. Then, from (1.6), we may write that $m = a\tau_\star/d^2$. We note that $\nu S_i^2/\sigma_i^2$, $i = 1, ..., k$, are independently distributed as a chi–square distribution with ν d.f. Let W_i, $i = 1, ..., k$, denote random variables such that νW_i, $i = 1, ..., k$, are independently distributed as the chi–square distribution with ν d.f. Let $\omega_i = W_i - 1$. Then, we have that $S_i^2 = \sigma_i^2(1+\omega_i)$, and $E(\omega_i) = 0$, $E(\omega_i^2) = 2\nu^{-1}$, $E(\omega_i^{2t-1}) = O(\nu^{-t})$ and $E(\omega_i^{2t}) = O(\nu^{-t})$, $t = 1, 2, ...$

Lemma 1 *For each i, we have as $\nu \to \infty$ that*

$$E_{\boldsymbol{\theta}}(|Y_i - \tau_i|^t) = O(\nu^{-t/2}) \quad (t \geq 2).$$

Proof. We write that

$$\begin{aligned}&S_iS_j - \sigma_i\sigma_j \\ &= \sigma_i\sigma_j\{(\sqrt{1+\omega_i}-1)(\sqrt{1+\omega_j}-1) + (\sqrt{1+\omega_i}-1) + (\sqrt{1+\omega_j}-1)\}.\end{aligned}$$

By noting that $E_{\theta}(|(1+\omega_i)^{1/2}-1|^t) = O(\nu^{-t/2})$ $(t \geq 2)$, we have that $E_{\theta}(|S_iS_j - \sigma_i\sigma_j|^t) = O(\nu^{-t/2})$ $(t \geq 2)$. Hence, it holds that

$$E_{\theta}(|Y_i - \tau_i|^t) = E_{\theta}\left(\left|\sum_{j=1}^{k} |b_i||b_j|(S_iS_j - \sigma_i\sigma_j)\right|^t\right) = O(\nu^{-t/2}) \quad (t \geq 2).$$

The proof is completed. □

Remark 9. As for (3.6), let $\tau_i = |b_i|\sqrt{\mathrm{tr}(\boldsymbol{\Sigma}_i)}\sum_{j=1}^{k}|b_j|\sqrt{\mathrm{tr}(\boldsymbol{\Sigma}_j)}$ and $Y_i = |b_i|\sqrt{\mathrm{tr}(\boldsymbol{S}_i)}\sum_{j=1}^{k}|b_j|$ $\sqrt{\mathrm{tr}(\boldsymbol{S}_j)}$. Let W_{ij}, $i = 1,...,k$; $j = 1,...,p$, denote random variables such that νW_{ij}, $i = 1,...,k$; $j = 1,...,p$, are independently distributed as a chi-square distribution with ν d.f. One may write that $\mathrm{tr}(\boldsymbol{S}_i) = \mathrm{tr}(\boldsymbol{\Sigma}_i) + \sum_{j=1}^{p}\lambda_{ij}(W_{ij}-1)$, where λ_{ij}'s are latent roots of $\boldsymbol{\Sigma}_i$. Then, we can obtain the same result as in Lemma 1 for (3.6) as well.

Lemma 2. *For the two-stage procedure (1.6)-(1.7) with (2.1), we have as $d \to 0$ that*

$$E_{\theta}\left(N_i - \left[\frac{u}{d^2}Y_i\right] - 1\right) = O(d).$$

Proof. Let $I_{\{N_i=m\}}$ be the indicator function. Then, we have that

$$\begin{aligned} E_{\theta}\left(N_i - \left[\tfrac{u}{d^2}Y_i\right] - 1\right) &= E_{\theta}\left\{I_{\{N_i=m\}}\left(m - \left[\tfrac{u}{d^2}Y_i\right] - 1\right)\right\} \\ &\leq \sqrt{P_{\theta}(N_i = m)E_{\theta}\left\{\left(m - \left[\tfrac{u}{d^2}Y_i\right] - 1\right)^2\right\}}. \end{aligned} \tag{4.1}$$

Then, it follows that

$$\begin{aligned} P_{\theta}(N_i = m) &= P_{\theta}\left(\frac{uY_i}{d^2} \leq m\right) \\ &= P_{\theta}\left(\frac{uY_i}{d^2C_i} - \frac{C_i+1}{C_i} \leq \frac{m-(C_i+1)}{C_i}\right) \\ &\leq P_{\theta}\left(\frac{uY_i}{a\tau_i} - 1 - \frac{1}{C_i} \leq \frac{\tau_\star - \tau_i}{\tau_i}\right) \\ &\leq P_{\theta}\left(\left|\frac{uY_i}{a\tau_i} - 1\right| + C_i^{-1} \geq \frac{\tau_i - \tau_\star}{\tau_i}\right) \\ &\leq \left(\frac{\tau_i - \tau_\star}{\tau_i}\right)^{-6} E_{\theta}\left\{\left(\left|\frac{uY_i}{a\tau_i} - 1\right| + C_i^{-1}\right)^6\right\}. \end{aligned} \tag{4.2}$$

Now, one can yield that

$$E_{\theta}\left\{\left|\frac{uY_i}{a\tau_i} - 1\right|^t\right\} \leq E_{\theta}\left\{\left(\frac{1}{\tau_i}\left(|Y_i - \tau_i| + \left|\frac{\hat{s}Y_i}{\nu}\right|\right)\right)^t\right\} = O(\nu^{-t/2}) \quad (t \geq 2). \tag{4.3}$$

Here, (4.3) follows from the result that for any x (≥ 0) and y (≥ 0) such that $x+y=t$ (≥ 2), we have from Lamma 1 that

$$\begin{aligned}E_{\theta}(|Y_i-\tau_i|^x|\nu^{-1}\hat{s}Y_i|^y) \leq& \sqrt{E_{\theta}(|Y_i-\tau_i|^{2x})E_{\theta}(|\nu^{-1}\hat{s}Y_i|^{2y})}\\ =& O\left(\nu^{-(x/2+y)}\right) = O\left(\nu^{-(t/2+y/2)}\right).\end{aligned}$$

By combining (4.3) with (4.2), we have that

$$P_{\theta}(N_i=m)=O(d^6). \tag{4.4}$$

The result can be obtained in view of (4.1) and (4.4). □

Lemma 3. *Let q (>0) and h (≥ 0) be constants. For a fixed b (≥ 1), let $X_{b\nu}$ denote a chi-square random variable with $b\nu$ d.f. Then, we have as $\nu \to \infty$ that*

$$E(qX_{b\nu}-h-[qX_{b\nu}-h])=\frac{1}{2}+O(\nu^{-1/2}).$$

Proof. Let $U=qX_{b\nu}-h-[qX_{b\nu}-h]$. Then, we have for $x \in (0,1)$ and $x_i \in (0,x)$ that

$$\begin{aligned}P(U\leq x) &= \sum_{i=0}^{\infty} P(U\leq x,\ i\leq qX_{b\nu}-h<i+1)\\ &= \sum_{i=0}^{\infty} P(i\leq qX_{b\nu}-h<i+x)\\ &= \sum_{i=0}^{\infty}\left(F_{b\nu}\left(\frac{i+h+x}{q}\right)-F_{b\nu}\left(\frac{i+h}{q}\right)\right)\\ &= \frac{x}{q}\sum_{i=0}^{\infty}F'_{b\nu}\left(\frac{i+h+x_i}{q}\right),\end{aligned} \tag{4.5}$$

where $F_{b\nu}(\cdot)$ is the c.d.f. of a chi-square random variable with $b\nu$ d.f., and $F'_{b\nu}(\cdot)$ denotes the first derivative of $F_{b\nu}(\cdot)$. Since $m \geq 4$ and $b \geq 1$, we have that $b\nu \geq 3$. Here, there is at most one constant c $(=b\nu-2)$ satisfying $\sup_z F'_{b\nu}(z)=F'_{b\nu}(c)$, $z>0$. If $(h+x_i)/q \leq b\nu-2$, there exists integer $i_\star$ such that $(i_\star+h+x_i)/q \leq b\nu-2 < (i_\star+1+h+x_i)/q$. Then, we have that

$$\int_i^{i+1} F'_{b\nu}\left(\frac{z+h+x_i}{q}\right)dz \geq \begin{cases} F'_{b\nu}\left(\frac{i+h+x_i}{q}\right) & (i<i_\star),\\ F'_{b\nu}\left(\frac{i+1+h+x_i}{q}\right) & (i\geq i_\star+1).\end{cases}$$

Hence, it follows that

$$\begin{aligned}\sum_{i=0}^{\infty}F'_{b\nu}\left(\frac{i+h+x_i}{q}\right) &\leq \int_{h+x_i}^{\infty}F'_{b\nu}\left(\frac{z}{q}\right)dz+F'_{b\nu}\left(\frac{i_\star+h+x_i}{q}\right)\\ &\leq \int_{h+x_i}^{\infty}F'_{b\nu}\left(\frac{z}{q}\right)dz+\sup_z F'_{b\nu}(z).\end{aligned} \tag{4.6}$$

If $(h+x_i)/q > b\nu - 2$, we have that

$$\int_i^{i+1} F'_{b\nu}\left(\frac{z+h+x_i}{q}\right) dz \le \begin{cases} F'_{b\nu}\left(\frac{i+1+h+x_i}{q}\right) & (i < i_\star), \\ F'_{b\nu}\left(\frac{i+h+x_i}{q}\right) & (i \ge i_\star + 1). \end{cases}$$

Hence, it follows that

$$\int_{h+x_i}^{\infty} F'_{b\nu}\left(\frac{z}{q}\right) dz - \sup_z F'_{b\nu}(z) \le \sum_{i=0}^{\infty} F'_{b\nu}\left(\frac{i+h+x_i}{q}\right). \tag{4.7}$$

Combining (4.6) and (4.7) with (4.5), we have that

$$\begin{aligned} x - xF_{b\nu}\left(\frac{h+x_i}{q}\right) - \frac{x}{q}\sup_z F'_{b\nu}(z) &\le P(U \le x) \\ &\le x - xF_{b\nu}\left(\frac{h+x_i}{q}\right) + \frac{x}{q}\sup_z F'_{b\nu}(z). \end{aligned} \tag{4.8}$$

For the second term in (4.8), it is expanded as

$$F_{b\nu}\left(\frac{h+x_i}{q}\right) = \frac{h+x_i}{q} F'_{b\nu}\left(\frac{h'_i}{q}\right), \tag{4.9}$$

where $h'_i \in (0, h+x_i)$. For the third term in (4.8), it is evaluated by Stirling's formula that

$$\sup_z F'_{b\nu}(z) = F'_{b\nu}(b\nu - 2) = O(\nu^{-1/2}) \quad \text{as} \quad \nu \to \infty. \tag{4.10}$$

By combining (4.9) and (4.10) with (4.8), we have that

$$P(U \le x) = x - \frac{x(h+x_i)}{q} F'_{b\nu}\left(\frac{h'_i}{q}\right) + O(\nu^{-1/2}) \quad \text{as} \quad \nu \to \infty.$$

Then, from the fact that $F'_{b\nu}(h'_i/q) \le \sup_z F'_{b\nu}(z) = O(\nu^{-1/2})$, we obtain that

$$P(U \le x) = x + O(\nu^{-1/2}) \quad \text{as} \quad \nu \to \infty. \tag{4.11}$$

It completes the proof. □

Lemma 4. *For the two-stage procedure (1.6)-(1.7) with (2.1), we have as* $d \to 0$ *that*

$$E_\theta\left\{\frac{u}{d^2}|b_i|S_i\sum_{j=1}^k |b_j|S_j - \left[\frac{u}{d^2}|b_i|S_i\sum_{j=1}^k |b_j|S_j\right]\right\} = \frac{1}{2} + O(d).$$

Proof. Let $X_{k\nu} = \nu\sum_{i=1}^k W_i$ and $V_i = \nu W_i/X_{k\nu}$, $i = 1, ..., k$. Then, $X_{k\nu}$ is distributed as the chi-square distribution with $k\nu$ d.f., V_i is distributed as the beta distribution with parameters $\nu/2$ and $(k-1)\nu/2$, and $X_{k\nu}$ and $\widetilde{V} = (V_1, ..., V_k)$ are independent. We write $\hat{s}$ as

$$\hat{s} = 1 + \frac{(a-1)b_i^2\sigma_i^2 V_i \sum_{j=1}^k b_j^2\sigma_j^2 V_j}{2Z_i^2} - \nu\frac{b_i^2\sigma_i^2 V_i k\tau_\star}{2X_{k\nu}Z_i^2},$$

where $Z_i = |b_i|\sigma_i\sqrt{V_i}\sum_{j=1}^k |b_j|\sigma_j\sqrt{V_j}$. Then, we have that $(u/d^2)|b_i|S_i\sum_{j=1}^k |b_j|S_j = (u/d^2\nu)$ $X_{k\nu}Z_i$ where

$$\frac{uX_{k\nu}Z_i}{d^2\nu} = X_{k\nu}\frac{aZ_i}{d^2\nu}\left\{1+\frac{1}{\nu}\left(1+\frac{(a-1)b_i^2\sigma_i^2V_i\sum_{j=1}^k b_j^2\sigma_j^2V_j}{2Z_i^2}\right)\right\} - \frac{ab_i^2\sigma_i^2V_ik\tau_\star}{2d^2Z_i\nu}.$$

Let us denote

$$Q = \frac{aZ_i}{d^2\nu}\left\{1+\frac{1}{\nu}\left(1+\frac{(a-1)b_i^2\sigma_i^2V_i\sum_{j=1}^k b_j^2\sigma_j^2V_j}{2Z_i^2}\right)\right\}, \quad H = \frac{ab_i^2\sigma_i^2V_ik\tau_\star}{2d^2Z_i\nu}. \tag{4.12}$$

Then, we have that $(u/d^2\nu)X_{k\nu}Z_i = QX_{k\nu} - H$. Let $U = (QX_{k\nu} - H) - [QX_{k\nu} - H]$. From Lemma 3, the conditional distribution of U, given $\widetilde{V} = \widetilde{v}$ $(H = h,\ Q = q)$, is given for $x \in (0,1)$ that

$$P_\theta(U \le x|\widetilde{V} = \widetilde{v}) \ge x - \frac{x(h+x_i)}{q}F'_{k\nu}\left(\frac{h'_i}{q}\right) - \frac{x}{q}\sup_z F'_{k\nu}(z),$$
$$P_\theta(U \le x|\widetilde{V} = \widetilde{v}) \le x - \frac{x(h+x_i)}{q}F'_{k\nu}\left(\frac{h'_i}{q}\right) + \frac{x}{q}\sup_z F'_{k\nu}(z),$$

where $x_i \in (0,x)$ and $h'_i \in (0, h+x_i)$. We evaluate that $H/Q \le k\tau_\star/(2\sum_{j=1}^k b_j^2\sigma_j^2V_j) \le k\tau_\star/(2\min_{1\le i\le k} b_i^2\sigma_i^2)$ $(=\gamma)$, and $1/Q \le \tau_\star/Z_i \le \tau_\star/(b_i^2\sigma_i^2V_i)$. Then, we have $E_\theta(x_i/Q) \le (\tau_\star/b_i^2\sigma_i^2)(k\nu-2)/(\nu-2)$. Here, H/Q is uniformly integrable since $|H/Q| \le \gamma$, and $1/Q$ is uniformly integrable since $|1/Q| \le \tau_\star/(b_i^2\sigma_i^2V_i)$ with $\tau_\star/(b_i^2\sigma_i^2V_i)$ being uniformly integrable. From (4.10), one can yield that

$$E_\theta\left\{\frac{H+x_i}{Q}F'_{k\nu}\left(\frac{H'_i}{Q}\right)\right\} \le E_\theta\left\{\frac{H+x_i}{Q}\sup_z F'_{k\nu}(z)\right\} = O(d),$$
$$E_\theta\left\{\frac{x}{Q}\sup_z F'_{k\nu}(z)\right\} = O(d),$$

where $H'_i \in (0, H+x_i)$. From the fact that $E_\theta\{P_\theta(U \le x|\widetilde{V} = \widetilde{v})\} = P_\theta(U \le x)$, we obtain that

$$P_\theta(U \le x) = x + O(d) \quad \text{as} \quad d \to 0. \tag{4.13}$$

Hence, U is asymptotically uniform on $(0,1)$ as $d \to 0$. The proof is completed. □

Remark 10. When $\hat{s}$ is given by (2.16), one may write that

$$Q = \frac{aZ_i}{d^2\nu}\left\{1+\frac{1}{\nu}\left(1-\frac{b_i^2\sigma_i^2V_i\sum_{j=1}^k b_j^2\sigma_j^2V_j}{Z_i^2}\right)\right\}, \quad H = \frac{ab_i^2\sigma_i^2V_ik\tau_\star}{2d^2Z_i\nu}. \tag{4.14}$$

When $\hat{s}$ is given by (3.34), one may write that

$$Q = \frac{\delta^2Z_i}{d^2\nu}\left\{1+\frac{1}{\nu}\left(1+\frac{(s_t-1)b_i^2\sigma_i^2V_i\sum_{j=1}^2 b_j^2\sigma_j^2V_j}{2Z_i^2}\right)\right\}, \quad H = \frac{\delta^2b_i^2\sigma_i^2V_i\tau_\star}{d^2Z_i\nu} \tag{4.15}$$

with

$$s_t = \frac{(\phi(\varepsilon_1)+\phi(\varepsilon_2))(\eta_1^3\phi(\eta_1)+\eta_2^3\phi(\eta_2)) - (\phi(\eta_1)+\phi(\eta_2))(\varepsilon_1^3\phi(\varepsilon_1)+\varepsilon_2^3\phi(\varepsilon_2))}{(\phi(\varepsilon_1)+\phi(\varepsilon_2))(\eta_1\phi(\eta_1)+\eta_2\phi(\eta_2)) - (\phi(\eta_1)+\phi(\eta_2))(\varepsilon_1\phi(\varepsilon_1)+\varepsilon_2\phi(\varepsilon_2))}. \tag{4.16}$$

When $\hat{s}$ is given by (3.47), one may write that

$$\begin{aligned} Q &= \left(\frac{z_\alpha - z_{1-\beta}}{1-\gamma}\right)^2 \frac{Z_i}{d^2\nu}\left\{1+\frac{1}{\nu}\left(1+\frac{(s_0-1)b_i^2\sigma_i^2V_i\sum_{j=1}^2 b_j^2\sigma_j^2V_j}{2Z_i^2}\right)\right\}, \\ H &= \left(\frac{z_\alpha - z_{1-\beta}}{1-\gamma}\right)^2 \frac{b_i^2\sigma_i^2V_i\tau_*}{d^2Z_i\nu} \end{aligned} \tag{4.17}$$

with $s_0 = z_\alpha^2 + z_{1-\beta}^2 + z_\alpha z_{1-\beta}$. When $\hat{s}$ is given by (3.7), one may write that

$$Q = \frac{Z_{L_i}}{W\nu}\left\{1+\frac{1}{\nu}\left(\frac{L_i^2\sum_{j=1}^k \Lambda_j(b_j^2L_j^2+Z_{L_j})}{Z_{L_i}^2}\right)\right\}, \quad H = \frac{b_i^2L_i^2k\tau_*}{2WZ_{L_i}\nu}, \tag{4.18}$$

where $L_i = \sqrt{\sum_{j=1}^p \lambda_{ij}V_{ij}}$, $\Lambda_i = \sum_{j=1}^p \lambda_{ij}^2V_{ij}^2/(\sum_{j=1}^p \lambda_{ij}V_{ij})^2$, $Z_{L_i} = |b_i|L_i\sum_{j=1}^k |b_j|L_j$, and V_{ij} denotes a beta random variable having parameters $\nu/2$ and $(k-1)\nu/2$. When $\hat{s}$ is given by (3.22), one may write that

$$Q = \frac{aZ_{\delta i}}{d^2\nu_p}\left\{1+\frac{1}{\nu_p}\left(1-\frac{\left(a\frac{G_p''(a)}{G_p'(a)}+1\right)b_i^2\delta_i^2V_{pi}\sum_{j=1}^k b_j^2\delta_j^2V_{pj}}{Z_{\delta i}^2}\right)\right\}, \quad H = \frac{ab_i^2\delta_i^2V_{pi}k\tau_*}{2d^2Z_{\delta i}\nu}, \tag{4.19}$$

where $Z_{\delta i} = |b_i|\delta_i\sqrt{V_{pi}}\sum_{j=1}^k |b_j|\delta_j\sqrt{V_{pj}}$ and V_{pi} denotes a beta random variable having parameters $\nu_p/2$ and $(k-1)\nu_p/2$. Note that, for nominal values of α and β, it holds that $s_t \geq -1$ in (4.16) and $G_p''(a)/G_p'(a) < 0$ in (4.19). We can evaluate that $E_{\boldsymbol\theta}(1/Q) = O(1)$ and $E_{\boldsymbol\theta}(H/Q) = O(1)$ for Q's and H's described above. Hence, the result similar to Lemma 4 is obtained for those cases as well.

Remark 11. When the design constant is defined as a constant, the asymptotic uniformity of $P(U \leq x)$ was studied by several authors. See Hall (1981) for $k=1$ and Takada (2004) for $k \geq 2$.

Lemma 5. *The two-stage procedure (1.6)-(1.7) with (2.1) has as $d \to 0$:*

(i) $E_{\boldsymbol\theta}\{C_i^{-1}(N_i - C_i)\} = (2\nu)^{-1}(2s - 1 + f_i + B_i) + O(d^3)$,

(ii) $E_{\boldsymbol\theta}\{C_i^{-2}(N_i - C_i)^2\} = (2\nu)^{-1}(1 + 2f_i + \sum_{i'=1}^k f_{i'}^2) + O(d^3)$,

(iii) $E_{\boldsymbol\theta}\{C_i^{-1}(N_i - C_i)C_j^{-1}(N_j - C_j)\} = (2\nu)^{-1}(f_i + f_j + \sum_{i'=1}^k f_{i'}^2) + O(d^3) \quad (i \neq j)$;

where $B_i = \nu/C_i$ and s is a constant such that $E_{\boldsymbol\theta}(\hat{s}) = s + o(1)$.

Proof. Let us write that

$$N_i = rC_iT_i + (1 + [rC_iT_i] - rC_iT_i) + (N_i - [rC_iT_i] - 1),$$

where $r = u/a = 1 + \nu^{-1}\hat{s}$ and $T_i = \tau_i^{-1}Y_i$. Here, from Lemma 4, $U_i = 1 + [rC_iT_i] - rC_iT_i$ is asymptotically distributed as $U(0,1)$. Let $D_i = N_i - [rC_iT_i] - 1$. From Lemma 2, it follows that $E\{(D_i/\nu)^c\} = O(\nu^{-3/2})$ as $d \to 0$, where c (≥ 1) is fixed. Then, we have that

$$C_i^{-1}(N_i - C_i) = (rT_i - 1) + \nu^{-1}B_iU_i + C_i^{-1}D_i. \tag{4.20}$$

By noting that $E_{\boldsymbol{\theta}}(\hat{s}) = s + o(1)$, we obtain the following results:

$$\begin{aligned} &E_{\boldsymbol{\theta}}(rT_i - 1) = (2\nu)^{-1}(2s - 1 + f_i) + O(d^3), \\ &E_{\boldsymbol{\theta}}\{(rT_i - 1)^2\} = (2\nu)^{-1}(1 + 2f_i + \sum_{i'=1}^{k} f_{i'}^2) + O(d^3), \\ &E_{\boldsymbol{\theta}}\{(rT_i - 1)(rT_j - 1)\} = (2\nu)^{-1}(f_i + f_j + \sum_{i'=1}^{k} f_{i'}^2) + O(d^3) \quad (i \neq j). \end{aligned} \tag{4.21}$$

Let us combine these results with the expectations of (4.20). Let $U_i = U_0 + \varepsilon_i$, where U_0 is a $U(0,1)$ random variable and ε_i is the remainder term. Then, note that $(E\{(rT_i - 1)\nu^{-1}\varepsilon_j\})^2 \leq E\{(rT_i - 1)^2\}E(\nu^{-2}\varepsilon_j^2) = o(\nu^{-3})$ so that $E\{(rT_i - 1)\nu^{-1}\varepsilon_j\} = o(\nu^{-3/2})$. The results are obtained straightforwardly. □

Remark 12. For the two-stage procedure (3.5)–(3.6) with (3.7), we have as $W \to 0$ that

(i) $E_{\boldsymbol{\theta}}\{C_i^{-1}(N_i - C_i)\} = (2\nu)^{-1}\{2s + B_i + A_i(f_i - 0.5) - 0.5\sum_{j=1}^{k} f_jA_j\} + O(W^{3/2})$,

(ii) $E_{\boldsymbol{\theta}}\{C_i^{-2}(N_i - C_i)^2\} = (2\nu)^{-1}\{A_i(1 + 2f_i) + \sum_{j=1}^{k} f_j^2A_j\} + O(W^{3/2})$,

where $A_i = \mathrm{tr}(\Sigma_i^2)/(\mathrm{tr}(\Sigma_i))^2$, $B_i = \nu/C_i$, and s is a constant such that $E_{\boldsymbol{\theta}}(\hat{s}) = s + o(1)$.

Lemma 6. *For the two-stage procedure (1.6)–(1.7) with (2.1), one has as $d \to 0$ that $E_{\boldsymbol{\theta}}(\Re) = o(\nu^{-1})$ in (2.4).*

Proof. In order to verify this lemma, we have to deal with the terms such as $E_{\boldsymbol{\theta}}(I_i)$, $E_{\boldsymbol{\theta}}(I_{ij})$ and $E_{\boldsymbol{\theta}}(I_{ij\ell})$, where

$$\begin{aligned} I_i &= \frac{\partial^3 g}{\partial u_i^3}\Big|_{\boldsymbol{u}=\boldsymbol{\xi}} \left(\frac{N_i - C_i}{C_i}\right)^3, \quad I_{ij} = \frac{\partial^3 g}{\partial u_i^2 \partial u_j}\Big|_{\boldsymbol{u}=\boldsymbol{\xi}} \left(\frac{N_i - C_i}{C_i}\right)^2 \left(\frac{N_j - C_j}{C_j}\right), \\ I_{ij\ell} &= \frac{\partial^3 g}{\partial u_i \partial u_j \partial u_\ell}\Big|_{\boldsymbol{u}=\boldsymbol{\xi}} \left(\frac{N_i - C_i}{C_i}\right)\left(\frac{N_j - C_j}{C_j}\right)\left(\frac{N_\ell - C_\ell}{C_\ell}\right) \end{aligned}$$

for all $1 \leq i < j < \ell \leq k$. Note that each third–order partial derivative's magnitude can be bounded from above by a finite sum of terms of the type

$$A\xi_1^{-p_1}\xi_2^{-p_2}\cdots\xi_k^{-p_k}$$

with $A \geq 0$, $p_r \geq 0$, $r = 1, ..., k$, which are independent of d. Let A also denote a generic positive constant, independent of d. Let us write $N_i^\star = C_i^{-1}(N_i - C_i)$ for $i = 1, ..., k$. Then, we obtain that

$$|E_{\boldsymbol{\theta}}(I_i)| \leq AE_{\boldsymbol{\theta}}(\xi_1^{-p_1}\xi_2^{-p_2}\cdots\xi_k^{-p_k}|N_i^\star|^3). \tag{4.22}$$

We observe that $\xi_i > C_i^{-1}m = \tau_i^{-1}\tau_\star$ w.p.1 for all $i = 1, ..., k$. Also, we observe that $E_{\boldsymbol{\theta}}(|N_i^\star|^3) = O(\nu^{-3/2})$ since $E_{\boldsymbol{\theta}}(|N_i^\star|^4) = O(\nu^{-2})$ from the facts that $E_{\boldsymbol{\theta}}\{(rT_i - 1)^3\} = O(\nu^{-2})$, $E_{\boldsymbol{\theta}}\{(rT_i -$

$1)^4\} = O(\nu^{-2})$ and so on together with (4.21). Hence, from (4.22), it follows that $|E_\theta(I_i)| = O(\nu^{-3/2})$.

Similarly, one may use the facts that $E_\theta(|N_i^\star|^2|N_j^\star|) = O(\nu^{-3/2})$ and $E_\theta(|N_i^\star||N_j^\star||N_\ell^\star|) = O(\nu^{-3/2})$ to show that $|E_\theta(I_{ij})| = O(\nu^{-3/2})$ and $|E_\theta(I_{ij\ell})| = O(\nu^{-3/2})$ for $1 \leq i < j < \ell \leq k$. Therefore, we conclude that $E_\theta(\Re) = O(\nu^{-3/2}) = o(\nu^{-1})$. □

ACKNOWLEDGMENTS

Research of the second author was partially supported by Grant-in-Aid for Scientific Research (B), Japan Society for the Promotion of Science (JSPS), under Contract Number 18300092.

REFERENCES

Aoshima, M. (2005). Statistical inference in two–stage sampling, *Selected Papers on Differential Equations and Analysis, Amer. Math. Soc. Transl.*, 215: 125–145.

Aoshima, M. and Kushida, T. (2005). Asymptotic second–order efficiency for two–stage multiple comparisons with components of a linear function of mean vectors, *Advances in Ranking and Selection, Multiple Comparisons, and Reliability: Methodology and Applications*, N. Balakrishnan, N. Kannan, and H.N. Nagaraja, eds., Birkhäuser, 191–211.

Aoshima, M. and Mukhopadhyay, N. (2002). Two–stage estimation of a linear function of normal means with second–order approximations, *Sequential Anal.*, 21: 109–144.

Aoshima, M. and Takada, Y. (2000). Second–order properties of a two–stage procedure for comparing several treatments with a control, *J. Japan Statist. Soc.*, 30: 27–41.

Aoshima, M. and Takada, Y. (2002). Bounded risk point estimation of a linear function of k multinormal mean vectors when covariance matrices are unknown, *Advances on Theoretical and Methodological Aspects of Probability and Statistics*, N. Balakrishnan, ed., Taylor & Francis, New York, 279–287.

Banerjee, S. (1967). Confidence interval of preassigned length for the Behrens–Fisher problem, *Ann. Math. Statist.*, 38: 1175–1179.

Dantzig, G.B. (1940). On the non–existence of tests of "Student's" hypothesis having power functions independent of σ, *Ann. Math. Statist.*, 11: 186–192.

Ghosh, M., Mukhopadhyay, N. and Sen, P.K. (1997). *Sequential Estimation,* John Wiley & Sons, Inc., New York.

Hall, P. (1981). Asymptotic theory of triple sampling for sequential estimation of a mean. *Ann. Statist.*, 9: 1229–1238.

Liu, W. (2003). Some sequential equivalence tests that control both the size and power, *South African Statist. J.*, 37: 105–125.

Mukhopadhyay, N. (2005). A new approach to determine the pilot sample size in two–stage sampling, *Commun. Statist.-Theory Meth* 34: 1275–1295.

Mukhopadhyay, N. and Duggan, W.T. (1997). Can a two–stage procedure enjoy second–order properties ?, *Sankhya, Series A*, 59: 435–448.

Mukhopadhyay, N. and Duggan, W.T. (1999). On a two–stage procedure having second–order properties with applications, *Ann. Inst. Statist. Math.,* 51: 621–636.

Schwabe, R. (1995). Some two–stage procedures for treating the Behrens–Fisher problem, Model-Oriented Data Analysis, W.G. Müller, H. P. Wynn and A. A. Zhigljavsky eds, Physica, *Heidelberg* 81–89.

Stein, C. (1945). A two–sample test for a linear hypothesis whose power is independent of the variance, *Annals of Mathematical Statistics* 16: 243–258.

Takada, Y. (2004). Asymptotic second–order efficiency of a two–stage procedure for estimating a linear function of normal means, *Sequential Analysis* 23: 103–120.

Takada, Y. and Aoshima, M. (1996). Two–stage procedures for the difference of two multinormal means with covariance matrices different only by unknown scalar multipliers, *Commun. Statist.-Theory Meth* 25: 2371–2379.

Takada, Y. and Aoshima, M. (1997). Two–stage procedures for estimating a linear function of multinormal mean vectors, *Sequential Analysis* 16: 353–362.