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

Graduate School of Pure and Applied Sciences
University of Tsukuba

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 \( \pi_i : N(\mu_i, \sigma_i^2) \), \( i = 1, \ldots, k \), where \( \mu_i \)'s and \( \sigma_i^2 \)'s are unknown. Let \( X_{i1}, X_{i2}, \ldots \) be a sequence of independent and identically distributed random variables from each \( \pi_i \). Having recorded \( X_{i1}, \ldots, X_{in_i} \) for each \( \pi_i \), let us write

\[
\overline{X}_{in_i} = \frac{1}{n_i} \sum_{j=1}^{n_i} X_{ij}
\]

and \( n=(n_1, \ldots, n_k) \). 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

\[
T_n = \sum_{i=1}^{k} b_i \overline{X}_{in_i}.
\]

We want to construct a fixed–width confidence interval such that

\[
P_\theta(|T_n - \mu| < d) \geq 1 - \alpha
\]

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

\[
P_\theta(|T_n - \mu| < d) = G \left( d^2 \left( \frac{\sum_{i=1}^{k} b_i^2 \sigma_i^2}{n_i} \right)^{-1} \right)
\]
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

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

(1.3)

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

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

(1.4)

for each $\pi_i$. However, since $\sigma_i$'s are unknown, the optimal fixed-sample sizes $C_i$'s should be estimated by using pilot samples from every $\pi_i$. 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 compared 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., $\lim \sup_{d \rightarrow 0} E_\theta (N_i - C_i) < \infty$. 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 $\sigma_{i*}$ for $\sigma_i$ such that

$$\sigma_i > \sigma_{i*}, \quad i = 1, \ldots, k.$$  

(1.5)

(T1) Having $m_0 (\geq 4)$ fixed, define

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

(1.6)

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

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

(1.7)

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

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

(1.8)

Let $N = (N_1, \ldots, N_k)$. 
Take an additional sample $X_{im+1}, \ldots, X_{iN_i}$ of size $N_i - m$ from each $\pi_i$. By combining the initial sample and the additional sample, calculate $\bar{X}_{iN_i} = N_i^{-1} \sum_{j=1}^{N_i} X_{ij}$ for each $\pi_i$. Finally, construct the fixed-width confidence interval with $T_N = \sum_{i=1}^{k} b_i \bar{X}_{iN_i}$.

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

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

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.,

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

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

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

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

$$\hat{s} = 1 + \frac{(a-1) \sum_{1\leq i \leq k} b_i^2 S_i^2 - k \tau_*}{2(\sum_{i=1}^{k} \sum_{j=1}^{k} b_i b_j S_i S_j)^2} \quad (2.1)$$

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

$$P_\theta(|T_N - \mu| < d) = E\theta \left\{ G \left( d^2 \left( \frac{\sum_{i=1}^{k} b_i^2 \sigma_i^2}{N_i} \right) \right) \right\}$$

$$= E\theta \left\{ G \left( a \left( \frac{\sum_{i=1}^{k} f_i C_i}{N_i} \right) \right) \right\}. \quad (2.2)$$

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

$$g(u_1, \ldots, u_k) = G(av^{-1}), \quad v = f_1 u_1^{-1} + \cdots + f_k u_k^{-1} \quad \text{for } u_i > 0, \ i = 1, \ldots, k.$$
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 $g(u_1, \ldots, u_k)$. For all $1 \leq i \neq j \leq k$, we have that
\[
\frac{\partial g}{\partial u_i} = aG'(a/v)f_i v^{-2}u_i^{-2},
\]
\[
\frac{\partial^2 g}{\partial u_i^2} = a\{aG''(a/v)f_i^2 v^{-4}u_i^{-4} + 2G'(a/v)f_i^2 v^{-3}u_i^{-4} - 2G'(a/v)f_i v^{-2}u_i^{-3}\},
\]
\[
\frac{\partial^2 g}{\partial u_i \partial u_j} = a\{aG''(a/v)f_if_j v^{-4}u_i^{-2}u_j^{-2} + 2G'(a/v)f_if_j v^{-3}u_i^{-2}u_j^{-2}\}.
\]

From (2.2), we use the Taylor expansion to claim that
\[
P_\theta(|T_N - \mu| < d) = E_\theta \left\{ g \left( \frac{N_1}{C_1}, \ldots, \frac{N_k}{C_k} \right) \right\}
= 1 - \alpha + aG'(a) \sum_{i=1}^{k} f_i E_\theta \left( \frac{N_i - C_i}{C_i} \right)
\]
\[
+ \frac{a}{2} \sum_{i=1}^{k} (aG''(a)f_i^2 v^{-4}u_i^{-4} - 2G'(a/v)f_i^2 v^{-3}u_i^{-4} - 2G'(a/v)f_i v^{-2}u_i^{-3})
\]
\[
+ E_\theta(\mathcal{R}),
\]
where
\[
E_\theta(\mathcal{R}) = \frac{1}{6} \sum_{1 \leq i \leq j \leq k} E_\theta \left\{ \frac{\partial^3 g}{\partial u_i \partial u_j \partial u_l} \bigg|_{u=\xi} \left( \frac{N_i - C_i}{C_i} \right) \left( \frac{N_j - C_j}{C_j} \right) \left( \frac{N_l - C_l}{C_l} \right) \right\}
\]
with suitable random variables $\xi_i$'s between $1$ and $N_i/C_i$, $i = 1, \ldots, k$, $u = (u_1, \ldots, u_k)$ and $\xi = (\xi_1, \ldots, \xi_k)$. With the help of Lemmas 5 and 6 in Appendix, we obtain the following expansion from (2.3):
\[
P_\theta(|T_N - \mu| < d) = 1 - \alpha
\]
\[
+ \frac{aG'(a)}{\nu} \left( s - 1 + \frac{1}{2} \sum_{i=1}^{k} f_i B_i + \sum_{i=1}^{k} f_i^2 + \frac{a}{G'(a)} \sum_{i=1}^{k} f_i^2 \right) + o(\nu^{-1}),
\]
where $B_i = C_i^{-1} \nu$ and $s$ is a constant such that $E_\theta(\tilde{s}) = s + o(1)$. Combining the results that $aG''(a)/G'(a) = (-a - 1)/2$ and $\sum_{i=1}^{k} f_i B_i = k\tau(\sum_{i=1}^{k} |b_i| \sigma_i)^{-2} + O(d^2)$ 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) up to 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
\[
P_\theta(|T_N - \mu| < d) = 1 - \alpha
\]
\[
+ \frac{aG'(a)}{2\nu} \left( a + 2k - 3 + \frac{k\tau(1-a)\sum_{i=1}^{k} b_i^2 \sigma_i^2}{(\sum_{i=1}^{k} |b_i| \sigma_i)^2} \right) + o(d^2) \text{ for all } \theta.
\]
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) $E_\theta(N_i - C_i) = (2\tau_\star)^{-1}\{[|b_i|\sigma_i \sum_{j=1}^{k} |b_j|\sigma_j + (a - 1)f_i \sum_{j=1}^{k} b_j^2 \sigma_j^2 + b_i^2 \sigma_i^2] + \frac{1}{2}(1 - k f_i) + o(1)\}$ for $i = 1, ..., k$,

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

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

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

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

\[
\frac{\mathcal{F}_{T_N}(\mu)}{\mathcal{F}_{T_C}(\mu)} = 1 + \frac{d^2(a + 1) \sum_{i=1}^{k} b_i^2 \sigma_i^2}{2a\tau_\star(\sum_{i=1}^{k} |b_i|\sigma_i)^2} + o(d^2),
\]

(2.7)

where $C = (C_1, ..., C_k)$ is defined by (1.4).

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

\[
\mathcal{F}_{T_N}(\mu) = E_\theta\left\{\left(\sum_{i=1}^{k} \frac{b_i^2 \sigma_i^2}{N_i} \right)^{-1}\right\}
\]

(2.8)

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

\[
\frac{\mathcal{F}_{T_N}(\mu)}{\mathcal{F}_{T_C}(\mu)} = E_\theta\left\{\left(\sum_{i=1}^{k} f_i C_i / N_i \right)^{-1}\right\}.
\]

(2.9)

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

\[
\frac{\mathcal{F}_{T_N}(\mu)}{\mathcal{F}_{T_C}(\mu)} = 1 + \sum_{i=1}^{k} f_i E_\theta\left(\frac{N_i - C_i}{C_i}\right) + \sum_{i=1}^{k} (f_i^2 - f_i) E_\theta\left(\frac{(N_i - C_i)^2}{C_i}\right)
\]

(2.10)

\[
+ \sum_{i \neq j} f_i f_j E_\theta\left(\frac{N_i - C_i}{C_i}\right)\left(\frac{N_j - C_j}{C_j}\right) + E_\theta(\Re),
\]

where

\[
E_\theta(\Re) = \frac{1}{6} \sum_{1 \leq i \leq j \leq k} E_\theta\left\{\frac{\partial^3 v^{-1}}{\partial u_i \partial u_j \partial u_l} \big|_{u=\xi}\left(\frac{N_i - C_i}{C_i}\right)\left(\frac{N_j - C_j}{C_j}\right)\left(\frac{N_l - C_l}{C_l}\right)\right\}
\]

(2.11)
with \( v = \sum_{i=1}^{k} f_i u_i^{-1} \) for \( u_i > 0, \ i = 1, \ldots, k \), suitable random variables \( \xi_i \)'s between 1 and \( N_i/C_i, \ i = 1, \ldots, k \), \( u = (u_1, \ldots, u_k) \) and \( \xi = (\xi_1, \ldots, \xi_k) \). With the help of Lemmas 5 and 6 in Appendix, we obtain the following expansion from (2.10):

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

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

**Remark 2.** For simplicity, we let \( k = 1 \) (\( b = 1 \)). Then, \( C = a \sigma^2/d^2 \). Under the assumption that \( \mathcal{F}_{X_N}(\mu) \) exceeds \( \mathcal{F}_{X_C}(\mu) \) for every fixed \( (\mu, \sigma^2) \), Mukhopadhyay (2005) proposed to determine the pilot sample size \( m \) for Stein’s (1945) two-stage estimation methodology as

\[
m = \text{smallest positive integer such that } \frac{\mathcal{F}_{X_N}(\mu)}{\mathcal{F}_{X_C}(\mu)} \leq 1 + \epsilon \tag{2.13}
\]

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

\[
m = \text{smallest positive integer such that } E_{\sigma^2}(N)/C \leq 1 + \epsilon + o(m^{-1}). \tag{2.14}
\]

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 \( \sigma^2 \), we should choose \( m \) in order \( O(d^c) \) with \( c \in (-1, 0) \) in order to specify quantity \( \epsilon \) free from \( \sigma^2 \). Then, we have that \( x = s \), so that \( m = s/\epsilon \) 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 \( \epsilon = s/m = d^{1/2} \). When \( \epsilon \) is specified as \( \epsilon = 0.1 \), we have that \( d = 10^{-4} \), 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

\[
\frac{\mathcal{F}_{T_N}(\mu)}{\mathcal{F}_{T_C}(\mu)} \leq 1 + \epsilon + o(m^{-1}),
\]

with \( \epsilon = (2a \tau_\star)^{-1}(a + 1)d^2 \). 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 \( T_N \) as \( d \to 0 \):

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

From (2.15), we have \( \epsilon = (2a \tau_\star)^{-1}(a + 3k - 1)d^2 \). It should be noted that the \( \epsilon \) 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

\[
\hat{s} = 1 - \frac{2 \sum_{i=1}^{k} b_i^2 \sigma_i^2 + k \tau_\star}{2(\sum_{i=1}^{k} |b_i| \sigma_i)^2}
\tag{2.16}
\]
instead of (2.1), the modified two-stage procedure (1.6)–(1.7) has the Fisher information in $T_N$ as $d \to 0$:
\[
\mathcal{F}_{T_N}(\mu)/\mathcal{F}_{T_C}(\mu) = 1 + o(m^{-1}).
\quad (2.17)
\]
Then, it holds as $d \to 0$:
\[
(i) \quad E_\theta(N_i - C_i) = (2\tau_\star)^{-1}\{ |b_i|\sigma_i \sum_{j=1}^{k} |b_j|\sigma_j - (2 \sum_{j=1}^{k} b_j^2\sigma^2_j + k\tau_\star)f_i + b_i^2\sigma^2_i \} + \frac{1}{2} + o(1) \quad \text{for } i = 1, \ldots, k,
\]
\[
(ii) \quad E_\theta(\sum_{i=1}^{k} N_i - \sum_{i=1}^{k} C_i) = (2\tau_\star)^{-1}\{ (\sum_{i=1}^{k} |b_i|\sigma_i)^2 - \sum_{i=1}^{k} b_i^2\sigma^2_i \} + o(1).
\]

3. APPLICATIONS

3.1. Bounded risk estimation
Suppose that there exist $k$ independent and normally distributed populations $\pi_i : N_p(\mu_i, \Sigma_i)$, $i = 1, \ldots, k$, where $\mu_i$'s $\in \mathbb{R}^p$ and $\Sigma_i$'s are both unknown, but $\Sigma_i$'s are $p \times p$ p.d. matrices. Let $X_{i1}, X_{i2}, \ldots$ be a sequence of independent and identically distributed random vectors from each $\pi_i$. Having recorded $X_{i1}, \ldots, X_{in_i}$ for each $\pi_i$, let us write $\overline{X}_{in_i} = \sum_{j=1}^{n_i} X_{ij}/n_i$ and $n = (n_1, \ldots, n_k)$. 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 $T_n = \sum_{i=1}^{k} b_i \overline{X}_{in_i}$. For a prespecified constant $W > 0$, we want to construct $T_n$ such that
\[
E_\theta(||T_n - \mu||^2) \leq W \quad (3.1)
\]
for all $\theta = (\mu_1, \ldots, \mu_k, \Sigma_1, \ldots, \Sigma_k)$, where $|| \cdot ||$ is the Euclidean norm. Since
\[
E_\theta(||T_n - \mu||^2) = \sum_{i=1}^{k} b_i^2 \text{tr}(\Sigma_i)/n_i, \quad (3.2)
\]
it is easy to see that the sample sizes $n$ which minimize the sum $\sum_{i=1}^{k} n_i$ subject to (3.1) are given as the smallest integer such that
\[
n_i \geq \frac{1}{W} |b_i|\sqrt{\text{tr}(\Sigma_i)} \sum_{j=1}^{k} |b_j|\sqrt{\text{tr}(\Sigma_j)} \quad (= C_i, \text{ say}) \quad (3.3)
\]
for each $\pi_i$.

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 $\sigma_{\star}$ for $\text{tr}(\Sigma_i)$ such that
\[
\text{tr}(\Sigma_i) > \sigma_{\star}, \quad i = 1, \ldots, k. \quad (3.4)
\]
Having $m_0 (\geq 4)$ fixed, define

$$m = \max \left\{ m_0, \left[ 1 \min_{1 \leq i \leq k} |b_i| \sigma_i \sum_{j=1}^k |b_j| \sigma_{j*} \right] + 1 \right\}. \quad (3.5)$$

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

$$N_i = \max \left\{ m, \left[ \frac{u}{W} |b_i| \sqrt{\text{tr}(S_i)} \sigma_{i*} \sum_{j=1}^k |b_j| \sqrt{\text{tr}(S_j)} \right] + 1 \right\}, \quad (3.6)$$

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

(T2) Take an additional sample $X_{i_{m+1}}, \ldots, X_{i_{N_i}}$ of size $Ni - m$ from each $\pi_i$. By combining the initial sample and the additional sample, calculate $\overline{X}_{iN} = N_i^{-1} \sum_{j=1}^{N_i} X_{ij}$ for each $\pi_i$. Finally, estimate $\mu$ by $T_N = \sum_{i=1}^k b_i \overline{X}_{iN}$. Theorem 3.1. Let $\tau_* = \min_{1 \leq i \leq k} |b_i| \sigma_i \sum_{j=1}^k |b_j| \sigma_{j*}$, where $\sigma_i$ is given by (3.4). Choose $u$ in (3.6) as $u = 1 + \nu^{-1} \hat{s}$, where

$$\hat{s} = \frac{2 \sum_{i=1}^k (\text{tr}(S_i^2) / (\text{tr}(S_i))^2) \left( b_i^2 \text{tr}(S_i) + |b_i| \sqrt{\text{tr}(S_i)} \sum_{j=1}^k |b_j| \sqrt{\text{tr}(S_j)} \right) - k \tau_*}{2 \left( \sum_{i=1}^k |b_i| \sqrt{\text{tr}(S_i)} \right)^2} \quad (3.7)$$

with $S_i$'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.,

$$E_\theta(||T_N - \mu||^2) = W + o(W^2) \quad \text{for all } \theta. \quad (3.8)$$

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

$$E_\theta(||T_N - \mu||^2) = E_\theta \left( \sum_{i=1}^k b_i^2 \text{tr}(\Sigma_i) / N_i \right)
= W E_\theta \left( \sum_{i=1}^k f_i C_i / N_i \right),$$

where $f_i = |b_i| \sqrt{\text{tr}(\Sigma_i)} / \sum_{j=1}^k |b_j| \sqrt{\text{tr}(\Sigma_j)}$. Use the Taylor expansion to claim that

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

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

$$E_\theta \left( \sum_{i=1}^k f_i C_i / N_i \right) = 1 + \frac{1}{2\nu} \sum_{i=1}^k f_i \left( -2s - B_i + A_i \left( f_i + \frac{3}{2} \right) + \sum_{j=1}^k f_j A_j \left( f_j + \frac{1}{2} \right) \right) + o(\nu^{-1}), \quad (3.10)$$
where $A_i = \text{tr}(\Sigma_i^2)/(\text{tr}(\Sigma_i))^2$, $B_i = \nu C_i^{-1}$, and $s$ is a constant such that $E_\theta(s) = s + o(1)$. 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 $\pi_i : N_p(\mu_i, \Sigma_i)$, $i = 1, \ldots, k$, where $p \geq 2$, and $\mu_i$'s $\in R^p$ and $\Sigma_i$'s are both unknown, but $\Sigma_i = (\sigma(i)_{rs}) (\geq 0)$ has a spherical structure such that

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

with $\delta_i (> 0)$ unknown parameter for each $\pi_i$. A special case of such the model is the intraclass correlation model, that is, $\Sigma_i = \sigma_i^2 \{1 - \rho_i I_p + \rho_i J\}$ for some $\rho_i$, where $J$ denotes a $p \times p$ matrix of all 1's. We consider multiple comparisons experiments for correlated components of $\mu = \Sigma_i = \sigma_i \{1 - \rho_i I_p + \rho_i J\}$. Let us write $\mu = (\xi_1, \ldots, \xi_p)$. Similarly to Section 3.1, we use $T_n = \Sigma_i = \sigma_i \{1 - \rho_i I_p + \rho_i J\}$ as an estimate of $\mu$. Let us write $T_n = (T_{1n}, \ldots, T_{pn})$. For a prespecified constant $d (> 0)$, we define three types of simultaneous confidence intervals for $(\xi_1, \ldots, \xi_p)$:

(MCA) $R_n = \{\mu | \xi_r - \xi_s \in [T_{rn} - T_{sn} - d, T_{rn} - T_{sn} + d], 1 \leq r < s \leq p\};$

(MCB) $R_n = \{\mu | \xi_r - \max_{\epsilon \neq r} \xi_{\epsilon} \in [(T_{rn} - \max_{s \neq r} T_{sn} - d)^-, +(T_{rn} - \max_{s \neq r} T_{sn} + d)^+],$

$\quad r = 1, \ldots, p\},$

where $+x^+ = \max\{0, x\}$ and $-x^- = \min\{0, x\};$

(MCC) $R_n = \{\mu | \xi_r - \xi_p \in [T_{rn} - T_{pn} - d, T_{rn} - T_{pn} + d], r = 1, \ldots, p - 1\}.$

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 $R_n$ such that

$$P_\theta(\mu \in R_n) \geq 1 - \alpha \quad \text{for all } \theta = (\mu_1, \ldots, \mu_k, \Sigma_1, \ldots, \Sigma_k)$$

(3.12)

with $\Sigma_i$'s defined by (3.11).

It is shown for MCA and MCC that

$$P_\theta(\mu \in R_n) = G_p \left( d^2 \left( \sum_{i=1}^{k} \frac{b_i^2 \delta_i^2}{n_i} \right)^{-1} \right),$$

(3.13)

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

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

(3.14)

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

(3.15)

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

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

(3.16)
where

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

(3.17)

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

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

(3.18)

for each \( \pi_i \), 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 \( \sigma_{i*} \) for \( \delta_i \) such that

\[ \delta_i > \sigma_{i*}, \quad i = 1, \ldots, k. \]  

(3.19)

(T1) Having \( m_0 \ (\geq 4) \) fixed, define

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

(3.20)

Take a pilot sample \( X_{ij} = (X_{ij1}, \ldots, X_{ijp}), \ j = 1, \ldots, m, \) and calculate \( S_{ip}^2 = \nu_p^{-1} \sum_{r=1}^{p} \sum_{j=1}^{m} (X_{1jr} - \overline{X}_{ijr} - \overline{X}_{ir} + \overline{X}_{i.})^2 \) with \( \nu_p = (p - 1)(m - 1) \) for each \( \pi_i \). Here, \( \overline{X}_{ij} = m^{-1} \sum_{j=1}^{m} X_{ijr}, \overline{X}_{i.} = m^{-1} \sum_{j=1}^{m} X_{ijr} \) 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 \( \nu_p \) d.f. Define the total sample size of each \( \pi_i \) by

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

(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 \( \mathcal{N} = (N_1, \ldots, N_k) \).

(T2) Take an additional sample \( X_{im+1}, \ldots, X_{in_i} \) of size \( N_i - m \) from each \( \pi_i \). 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 \( \pi_i \). Finally, for each method, construct \( R_N \) with the components \((T_{1N}, \ldots, T_{pN})\) of \( T_N = \sum_{i=1}^{k} b_i \overline{X}_{iN_i} \).

The following theorem can be obtained similarly to Theorem 2.1.

**Theorem 3.2.** Let \( \tau_* = \min_{1 \leq i \leq k} |b_i| \sigma_{i*} \sum_{j=1}^{k} |b_j| \sigma_{j*} \), where \( \sigma_{i*} \) 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

\[ \hat{s} = 1 - \frac{2(a \frac{\partial^2 \nu_p}{\partial \sigma_{i*}^2} + 1) \sum_{i=1}^{k} b_i^2 S_{ip}^2 + k(p - 1) \tau_*}{2(\sum_{i=1}^{k} |b_i| S_{ip}^2)^2} \]  

(3.22)

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

\[ P_{\theta}(\mu \in R_N) = 1 - \alpha + o(d^2) \quad \text{for all} \ \theta \quad (MCA, MCC); \]  

(3.23)

\[ P_{\theta}(\mu \in R_N) \geq 1 - \alpha + o(d^2) \quad \text{for all} \ \theta \quad (MCB). \]  

(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 \( \alpha \), note that \( aG'_p(a)/G_p(a) \leq -1 \). Then, from (3.22), we have that \( \hat s < s \) w.p.l. 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 \( \pi_i : N(\mu_i, \sigma_i^2), i = 1, 2 \), with \( \mu_i \)'s and \( \sigma_i^2 \)'s both unknown. We want to design a test of

\[
H_0 : |\mu| = |\mu_1 - \mu_2| \geq d \quad \text{against} \quad H_a : |\mu| < d
\]  (3.25)

which has size \( \alpha \) and power no less than \( 1 - \beta \) at \( |\mu| \leq \gamma d \) for all \( \theta = (\mu_1, \mu_2, \sigma_1^2, \sigma_2^2) \), where \( \alpha, \beta \in (0, 1), \gamma \in [0, 1) \), 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 \( \sigma_i^2 \)'s had been known, we would take a sample from each \( \pi_i \) of size

\[
n_i \geq \frac{\delta^2}{d^2} \sigma_i \sum_{j=1}^{2} \sigma_j \quad (= C_i, \text{ say})
\]  (3.26)

and test the hypothesis by

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

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

\[
P(|N(0,1) + x| < R(x)) = \alpha
\]  (3.28)

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

\[
P(|N(0,1) + \gamma \delta| < R(\delta)) = 1 - \beta.
\]  (3.29)

When \( \sigma_i^2 \)'s are unknown but common \( (\sigma_1^2 = \sigma_2^2) \), Liu (2003) proposed \( k \) (\( k \geq 3 \))-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 \( \sigma_{\star} \) for \( \sigma_i \) such that

\[
\sigma_i > \sigma_{\star}, \quad i = 1, 2.
\]  (3.30)

(T1) Having \( m_0 \) (\( \geq 4 \)) fixed, define

\[
m = \max \left\{ m_0, \left[ \frac{\delta^2}{d^2} \min_{1 \leq i \leq 2} \sigma_{\star} \sum_{j=1}^{2} \sigma_{j\star} \right] + 1 \right\}.
\]  (3.31)
Take a pilot sample $X_{11}, \ldots, X_{1m}$ of size $m$ and calculate $S_{i1}^2 = \sum_{j=1}^{m} (X_{ij} - \overline{X}_{im})^2/\nu$ with $\nu = m-1$ for each $\pi_i$. Define the total sample size of each $\pi_i$ by

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

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 $X_{im+1}, \ldots, X_{iN}$ of size $N_i - m$ from each $\pi_i$. By combining the initial sample and the additional sample, calculate $\overline{X}_{iN_i} = \frac{1}{N_i} \sum_{j=1}^{N_i} X_{ij}$ for each $\pi_i$. Then, test the hypothesis by

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

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

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

$$\hat{s} = 1 - \frac{\sum_{i=1}^{2} S_i^2}{2(\sum_{i=1}^{2} S_i)^2} - \frac{\tau_*}{(\sum_{i=1}^{2} S_i)^2} + \frac{(\phi(\epsilon_1) + \phi(\epsilon_2))(\eta_1^2 \phi(\eta_1) + \eta_2^2 \phi(\eta_2)) - (\phi(\eta_1) + \phi(\eta_2)) (\epsilon_1^2 \phi(\epsilon_1) + \epsilon_2^2 \phi(\epsilon_2))}{(\phi(\epsilon_1) + \phi(\epsilon_2))(\eta_1 \phi(\eta_1) + \eta_2 \phi(\eta_2)) - (\phi(\eta_1) + \phi(\eta_2)) (\epsilon_1 \phi(\epsilon_1) + \epsilon_2 \phi(\epsilon_2))} \sum_{i=1}^{2} S_i^2,$$  

where $\phi(\cdot)$ is the p.d.f. of $N(0, 1)$, $\epsilon_1 = R(\delta) - \delta$, $\epsilon_2 = R(\delta) + \delta$, $\eta_1 = R(\delta) - \gamma \delta$, $\eta_2 = R(\delta) + \gamma \delta$, and $S_i^2$'s are calculated in (T1).

Choose $\lambda$ in (3.33) as $\lambda = (1 + \nu^{-1}\hat{t})$ with

$$\hat{t} = \frac{(\epsilon_1^2 \phi(\epsilon_1) + \epsilon_2^2 \phi(\epsilon_2))(\eta_1 \phi(\eta_1) + \eta_2 \phi(\eta_2)) - (\epsilon_1 \phi(\epsilon_1) + \epsilon_2 \phi(\epsilon_2))(\eta_1^2 \phi(\eta_1) + \eta_2^2 \phi(\eta_2))}{2R(\delta) \left\{ (\phi(\epsilon_1) + \phi(\epsilon_2))(\eta_1 \phi(\eta_1) + \eta_2 \phi(\eta_2)) - (\epsilon_1 \phi(\epsilon_1) + \epsilon_2 \phi(\epsilon_2))(\phi(\eta_1) + \phi(\eta_2)) \right\} (\sum_{i=1}^{2} S_i)^2},$$

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.,

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

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

$$E_\theta \left\{ \Phi \left( (\sqrt{\lambda} R(\delta) - \delta) \left( \sum_{i=1}^{2} f_i \frac{C_i}{N_i} \right)^{-1/2} \right) \right\} = E_\theta \left\{ \Phi \left( (\sqrt{\lambda} R(\delta) + \delta) \left( \sum_{i=1}^{2} f_i \frac{C_i}{N_i} \right)^{-1/2} \right) \right\}$$

$$= \Phi(\epsilon_1) - \Phi(-\epsilon_2) + \frac{1}{4\nu} \epsilon_1 \phi(\epsilon_1) \left( 2s + \frac{2R(\delta)}{\epsilon_1} t - 2 + \sum_{i=1}^{2} f_i B_i + \sum_{i=1}^{2} f_i^2 + \epsilon_1 \frac{\phi'(\epsilon_1)}{\phi(\epsilon_1)} \right)$$

$$+ \frac{1}{4\nu} \epsilon_2 \phi(\epsilon_2) \left( 2s + \frac{2R(\delta)}{\epsilon_2} t - 2 + \sum_{i=1}^{2} f_i B_i + \sum_{i=1}^{2} f_i^2 + \epsilon_2 \frac{\phi'(\epsilon_2)}{\phi(\epsilon_2)} \right) + E_\theta(\Re),$$

(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

$$
E_{\theta} \left\{ \Phi \left( \left( \sqrt{\lambda} R(\delta) - \gamma \delta \right) \left( \sum_{i=1}^{2} f_i \frac{C_i}{N_i} \right)^{-1/2} \right) \right\} - E_{\theta} \left\{ \Phi \left( \left( \sqrt{\lambda} R(\delta) + \gamma \delta \right) \left( \sum_{i=1}^{2} f_i \frac{C_i}{N_i} \right)^{-1/2} \right) \right\}
$$

$$
= \Phi(\eta_1) - \Phi(-\eta_2) + \frac{1}{4\nu} \eta_1 \phi(\eta_1) \left( 2s + \frac{2R(\delta)}{\eta_1} t - 2 + \sum_{i=1}^{2} f_i B_i + \sum_{i=1}^{2} f_i^2 + \eta_1 \frac{\phi'(\eta_1)}{\phi(\eta_1)} \right)
$$

$$
+ \frac{1}{4\nu} \eta_2 \phi(\eta_2) \left( 2s + \frac{2R(\delta)}{\eta_2} t - 2 + \sum_{i=1}^{2} f_i B_i + \sum_{i=1}^{2} f_i^2 + \eta_2 \frac{\phi'(\eta_2)}{\phi(\eta_2)} \right) + E_{\theta}(\Re). \quad (3.38)
$$

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 $d \rightarrow 0$ in (3.37)-(3.38).

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

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

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

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

and test the hypothesis by

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

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

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

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

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

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

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

$$
H_0 : \mu = \mu_1 - \mu_2 \leq -d \quad \text{against} \quad H_a : \mu > -d \quad (3.44)
$$
which has size $\alpha$ and power no less than $1 - \beta$ at $\mu \geq -\gamma d$ for all $\theta$. 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 $\sigma_i^2$'s had been known, we would take a sample from each $\pi_i$ of size

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

and test the hypothesis by

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

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

$$\hat{\delta} = 1 + (z_\alpha^2 + z_{1-\beta}^2 + z_\alpha z_{1-\beta} - 1) \frac{\sum_{i=1}^{2} S_i^2}{2(\sum_{i=1}^{2} S_i)^2} - \frac{\tau_*}{(\sum_{i=1}^{2} S_i)^2};$$

$$\hat{\iota} = z_\alpha z_{1-\beta}(z_\alpha + z_{1-\beta}) \frac{1 - r}{\gamma z_\alpha - z_{1-\beta}} \frac{\sum_{i=1}^{2} S_i^2}{2(\sum_{i=1}^{2} S_i)^2};$$

Then, the test of (3.44), given by

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

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

4. APPENDIX

Throughout this section, we write that

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

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 $a\tau_i/d^2$ always remains an integer. Then, from (1.6), we may write that $m = a\tau_i/d^2$. We note that $\nu S_i^2/\sigma_i^2$, $i = 1, ..., k$, are independently distributed as a chi-square distribution with $\nu$ d.f. Let $W_i$, $i = 1, ..., k$, denote random variables such that $\nu W_i$, $i = 1, ..., k$, are independently distributed as the chi-square distribution with $\nu$ 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 $E(\omega_i^{2t}) = O(\nu^{-t})$, $t=1,2,\ldots$

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

$$E_\theta(|Y_i - \tau_i|^t) = O(\nu^{-t/2}) \quad (t \geq 2).$$
Proof. We write that
\[ S_i S_j - \sigma_i \sigma_j = \sigma_i \sigma_j \{(\sqrt{1 + \omega_i} - 1)(\sqrt{1 + \omega_j} - 1) + (\sqrt{1 + \omega_i} - 1) + (\sqrt{1 + \omega_j} - 1)\}. \]

By noting that \( E_\theta((1 + \omega_i)^{1/2} - 1)^t = O(\nu^{-t/2}) \) \((t \geq 2)\), we have that \( E_\theta(|S_i S_j - \sigma_i \sigma_j |^t) = O(\nu^{-t/2}) \) \((t \geq 2)\). Hence, it holds that
\[ E_\theta(|Y_i - \tau_i|^t) = E_\theta\left(\left|\sum_{j=1}^{k}|b_i||b_j|(S_{ij} - \sigma_i \sigma_j)\right|^t\right) = O(\nu^{-t/2}) \) \((t \geq 2)\).

The proof is completed. \(\square\)

Remark 9. As for (3.6), let \( \tau_i = |b_i|\sqrt{\text{tr}(\Sigma_i)}\sum_{j=1}^{k}|b_j|\sqrt{\text{tr}(\Sigma_j)} \) and \( Y_i = |b_i|\sqrt{\text{tr}(S_i)}\sum_{j=1}^{k}|b_j|\sqrt{\text{tr}(S_j)} \).

Let \( W_{ij}, i = 1, \ldots, k; j = 1, \ldots, p \), denote random variables such that \( \nu W_{ij} \) are independently distributed as a chi-square distribution with \( \nu \) d.f. One may write that \( \text{tr}(S_i) = \text{tr}(\Sigma_i) + \sum_{j=1}^{p}\lambda_{ij}(W_{ij} - 1) \), where \( \lambda_{ij} \)'s are latent roots of \( \Sigma_i \). 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
\[ E_\theta\left(N_i - \left[\frac{u}{d^2} Y_i\right] - 1\right) = E_\theta\left\{I_{\{N_i=m\}}\left(m - \left[\frac{u}{d^2} Y_i\right] - 1\right)\right\} \leq \sqrt{P_\theta(N_i=m)E_\theta\left\{(m - \left[\frac{u}{d^2} Y_i\right] - 1)^2\right\}}. \]

Then, it follows that
\[ P_\theta(N_i = m) = P_\theta\left(\frac{uY_i}{d^2} \leq m\right) = P_\theta\left(\frac{uY_i}{d^2C_i} - \frac{C_i+1}{C_i} \leq \frac{m - (C_i+1)}{C_i}\right) \leq P_\theta\left(\frac{uY_i}{d^2C_i} - 1 - \frac{1}{C_i} \leq \frac{\tau_s - \tau_i}{\tau_i}\right) \leq P_\theta\left(\frac{|uY_i/a\tau_i| - 1}{C_i} \geq \frac{\tau_i - \tau_s}{\tau_i}\right) \leq \left(\frac{\tau_i - \tau_s}{\tau_i}\right)^6 E_\theta\left\{\left(\frac{|uY_i/a\tau_i| - 1}{C_i} + C_i^{-1}\right)^6\right\}. \]

Now, one can yield that
\[ E_\theta\left\{|uY_i/a\tau_i - 1|^t\right\} \leq E_\theta\left\{\left(\frac{1}{\tau_i} \left(|Y_i - \tau_i| + \frac{\delta Y_i}{\nu}\right)\right)^t\right\} = O(\nu^{-t/2}) \) \((t \geq 2)\).
Here, (4.3) follows from the result that for any $x \geq 0$ and $y \geq 0$ such that $x + y = t \geq 2$, we have from Lamma 1 that

$$E_{\theta}(|Y_i - \tau_i|^2 | \nu^{-1} \hat{s} Y_i|^p) \leq \sqrt{E_{\theta}(|Y_i - \tau_i|^{2p}) E_{\theta}(|\nu^{-1} \hat{s} Y_i|^{2p})}$$

$$= O \left( \nu^{-(x/2+y)} \right) = O \left( \nu^{-(t/2+y/2)} \right).$$

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

$$P_{\theta}(N_i = m) = O(\emptyset).$$

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

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

$$P(U \leq x) = \sum_{i=0}^{\infty} P(U \leq x, i \leq qX_{b\nu} - h < i + 1)$$

$$= \sum_{i=0}^{\infty} P(i \leq qX_{b\nu} - h < i + x)$$

$$= \sum_{i=0}^{\infty} \left( F_{b\nu} \left( \frac{i + h + x}{q} \right) - F_{b\nu} \left( \frac{i + h}{q} \right) \right)$$

$$= \frac{x}{q} \sum_{i=0}^{\infty} F_{b\nu}' \left( \frac{i + h + x_i}{q} \right),$$

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 $i_*$ such that $(i_* + h + x_i)/q \leq b\nu - 2 < (i_* + 1 + h + x_i)/q$. Then, we have that

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

Hence, it follows that

$$\sum_{i=0}^{\infty} F_{b\nu}' \left( \frac{i + h + x_i}{q} \right) \leq \int_{h + x_i}^{x_i} F_{b\nu}' \left( \frac{z}{q} \right) dz + F_{b\nu}' \left( \frac{i_* + h + x_i}{q} \right)$$

$$\leq \int_{h + x_i}^{x_i} F_{b\nu}' \left( \frac{z}{q} \right) dz + \sup_{x_i} F_{b\nu}'(z).$$

(4.6)
If \((h + x_i)/q > b
u - 2\), we have that
\[
\int_i^{i+1} F'_{b
u} \left( \frac{z + h + x_i}{q} \right) dz \leq \begin{cases} 
F'_{b
u} \left( \frac{i+1+h+x_i}{q} \right) & (i < i_*) \\
F'_{b
u} \left( \frac{i+h+x_i}{q} \right) & (i \geq i_* + 1).
\end{cases}
\]
Hence, it follows that
\[
\int_{h+x_i}^{\infty} F'_{b
u} \left( \frac{z}{q} \right) dz - \sup_z F'_{b
u}(z) \leq \sum_{i=0}^{\infty} F'_{b
u} \left( \frac{i+h+x_i}{q} \right). 
\tag{4.7}
\]
Combining (4.6) and (4.7) with (4.5), we have that
\[
x - x F_{b
u} \left( \frac{h + x_i}{q} \right) - \frac{x}{q} \sup_z F_{b
u}'(z) \leq P(U \leq x)
\leq x - x F_{b
u} \left( \frac{h + x_i}{q} \right) + \frac{x}{q} \sup_z F_{b
u}'(z). 
\tag{4.8}
\]
For the second term in (4.8), it is expanded as
\[
F_{b
u} \left( \frac{h + x_i}{q} \right) = \frac{h + x_i}{q} F_{b
u}' \left( \frac{h_i'}{q} \right), 
\tag{4.9}
\]
where \(h_i' \in (0, h + x_i)\). For the third term in (4.8), it is evaluated by Stirling’s formula that
\[
\sup_z F_{b
u}'(z) = F_{b
u}'(b
u - 2) = O(\nu^{-1/2}) \quad \text{as} \quad \nu \to \infty. 
\tag{4.10}
\]
By combining (4.9) and (4.10) with (4.8), we have that
\[
P(U \leq x) = x - x F_{b
u} \left( \frac{h + x_i}{q} \right) + O(\nu^{-1/2}) \quad \text{as} \quad \nu \to \infty.
\tag{4.11}
\]
It completes the proof.
\[\square\]

**Lemma 4.** For the two-stage procedure (1.6)–(1.7) with (2.1), we have as \(d \to 0\) that
\[
E( \hat{s} ) \leq \frac{u |b_i| S_i \sum_{j=1}^{k} |b_j|S_j}{d^2 |b_i|^2 S_i \sum_{j=1}^{k} |b_j|S_j} \leq \frac{1}{2} + O(d).
\]

**Proof.** Let \(X_{k
u} = \nu \sum_{i=1}^{k} W_i\) and \(V_i = \nu W_i / X_{k
u}, i = 1, \ldots, k\). Then, \(X_{k
u}\) 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 \(X_{k
u}\) and \(\tilde{V} = (V_1, \ldots, V_k)\) are independent. We write \(\hat{s}\) as
\[
\hat{s} = 1 + \frac{(a - 1)b_i^2 \sigma_i^2 V_i \sum_{j=1}^{k} b_j^2 \sigma_j^2 V_j}{2 Z_i^2} - \nu b_i^2 \sigma_i^2 V_i k \tau_* \frac{2 X_{k
u} Z_i^2}{2X_{k
u} Z_i^2},
\]
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 \((u/d^2)|b_i|S_i \sum_{j=1}^{k} |b_j|S_j = (u/d^2)\) \( X_{k\nu}Z_i \) where

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

Let us denote

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

(4.12)

Then, we have that \((u/d^2)X_{k\nu}Z_i = QX_{k\nu} - H\). Let \( U = (QX_{k\nu} - H) - [QX_{k\nu} - H] \).

Let \( U = (QX_{k\nu} - H) - [QX_{k\nu} - H] \).

Then, \( P_\theta(U \leq x|\tilde{V} = \tilde{v}) \geq x - \frac{x(h + x_i)}{q} F'_{X_{k\nu}}(\frac{h_i'}{q}) - \frac{x}{q} \sup_z F'_{X_{k\nu}}(z), \)

\[
P_\theta(U \leq x|\tilde{V} = \tilde{v}) \leq x - \frac{x(h + x_i)}{q} F'_{X_{k\nu}}(\frac{h_i'}{q}) + \frac{x}{q} \sup_z F'_{X_{k\nu}}(z),
\]

where \( x_i \in (0, x) \) and \( h_i' \in (0, h + x_i) \). We evaluate that \( H/Q \leq k\tau_*/(2\sum_{j=1}^{k} b_j^2\sigma_j^2V_j) \leq k\tau_*/(2\min_{1 \leq i \leq k} b_i^2\sigma_i^2)(= \gamma) \), and \( 1/Q \leq \tau_*/Z_i \leq \tau_*/(b_i^2\sigma_i^2V_i) \). Then, we have \( E_\theta(x_i/Q) \leq (\tau_*/b_i^2\sigma_i^2)(k\nu-2)/\nu \).

From (4.10), one can yield that

\[
E_\theta \left\{ \frac{H + x_i}{Q} F'_{X_{k\nu}} \left( \frac{h_i'}{Q} \right) \right\} \leq E_\theta \left\{ \frac{H + x_i}{Q} \sup_z F'_{X_{k\nu}}(z) \right\} = O(d),
\]

\[
E_\theta \left\{ \frac{x}{Q} \sup_z F'_{X_{k\nu}}(z) \right\} = O(d),
\]

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

\[
P_\theta(U \leq x) = x + O(d) \quad \text{as} \quad d \to 0.
\]

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

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

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

(4.13)

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

\[
Q = \frac{\delta^2Z_i}{d^2\nu} \left( 1 + \frac{1}{\nu} \left( 1 + \frac{(s_i - 1)b_i^2\sigma_i^2V_i \sum_{j=1}^{k} b_j^2\sigma_j^2V_j}{2Z_i^2} \right) \right), \quad H = \frac{\delta^2b_i^2\sigma_i^2V_i \tau_*}{d^2Z_i\nu}.
\]

(4.14)
with
\[
\phi (\epsilon_1) + \phi (\epsilon_2)) \left( \frac{\eta_1^2 \phi (\eta_1) + \eta_2^2 \phi (\eta_2)}{(\phi (\epsilon_1) + \phi (\epsilon_2)) (\eta_1 \phi (\eta_1) + \eta_2 \phi (\eta_2))} \right)
\frac{\eta_1^2 \phi (\eta_1) + \eta_2^2 \phi (\eta_2)}{\phi (\epsilon_1) + \phi (\epsilon_2)) (\eta_1 \phi (\eta_1) + \eta_2 \phi (\eta_2))}
\frac{(\eta_1 \phi (\eta_1) + \eta_2 \phi (\eta_2)) - (\phi (\eta_1) + \phi (\eta_2)) (\epsilon_1 \phi (\epsilon_1) + \epsilon_2 \phi (\epsilon_2))}{(\phi (\epsilon_1) + \phi (\epsilon_2)) (\eta_1 \phi (\eta_1) + \eta_2 \phi (\eta_2)) - (\phi (\eta_1) + \phi (\eta_2)) (\epsilon_1 \phi (\epsilon_1) + \epsilon_2 \phi (\epsilon_2))}.
\]
(4.16)

When \( \hat{s} \) is given by (3.47), one may write that
\[
Q = \left( \frac{z_\alpha - z_{1-\beta}}{1-\gamma} \right)^2 \frac{Z_i}{d^2 \nu} \left\{ 1 + \frac{1}{\nu} \left( 1 + \frac{(s_0 - 1)b_1^2 \sigma_i^2 V_i \sum_{j=1}^{k} b_j^2 \sigma_j^2 V_j}{Z_i^2} \right) \right\},
\]
\[
H = \left( \frac{z_\alpha - z_{1-\beta}}{1-\gamma} \right)^2 \frac{b_1^2 \sigma_i^2 V_i \tau_s}{d^2 Z_i \nu},
\]
(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
\[
Q = \frac{Z_{Li}}{W \nu} \left\{ 1 + \frac{1}{\nu} \left( \frac{L_i^2 \sum_{j=1}^{k} \Lambda_{ij} (b_i^2 L_i^2 + Z_{L_{ij}})}{Z_{Li}^2} \right) \right\},
\]
\[
H = \frac{ab_1^2 \delta_{i}^2 V_{\mu} k \tau_s}{2d^2 Z_{Li} \nu},
\]
(4.18)

where \( L_i = \sqrt{\sum_{j=1}^{p} \lambda_{ij} V_{ij}} \), \( \Lambda_{ij} = \sum_{j=1}^{p} \lambda_{ij}^2 V_{ij}^2 / \left( \sum_{j=1}^{p} \lambda_{ij} V_{ij} \right)^2 \), \( Z_{\delta} = \sum_{j=1}^{k} |b_j| \delta_j \sqrt{V_{pj}} \) and \( V_{\mu} \) denotes a beta random variable having parameters \( \nu_p/2 \) and \( (k-1)\nu_p/2 \). Note that, for nominal values of \( \alpha \) and \( \beta \), it holds that \( s \geq -1 \) in (4.16) and \( G''_p(a)/G'_p(a) < 0 \) in (4.19). We can evaluate that \( E_\theta (1/Q) = O(1) \) and \( E_\theta (H/Q) = O(1) \) 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) \( E_\theta \{ C_i^{-1} (N_i - C_i) \} = (2\nu)^{-1} (2s - 1 + f_i + B_i) + O(d^3) \),

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

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

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

**Proof.** Let us write that
\[
N_i = rC_i T_i + (1 + [rC_i T_i] - rC_i T_i) + (N_i - [rC_i T_i] - 1),
\]
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 \( c \geq 1 \) is fixed. Then, we have that
\[
C_i^{-1}(N_i - C_i) = (rT_i - 1) + \nu^{-1}B_i U_i + C_i^{-1}D_i. \quad (4.20)
\]
By noting that \( E_{\theta}(\hat{s}) = s + o(1) \), we obtain the following results:
\[
\begin{align*}
E_{\theta}(rT_i - 1) &= (2\nu)^{-1}(2s - 1 + f_i) + O(d^3), \\
E_{\theta}\{(rT_i - 1)^2\} &= (2\nu)^{-1}(1 + 2f_i + \sum_{i'=1}^{k} f_{i'}^2) + O(d^3), \\
E_{\theta}\{(rT_i - 1)(rT_j - 1)\} &= (2\nu)^{-1}(f_i + f_j + \sum_{i'=1}^{k} f_{i'}^2) + O(d^3) \quad (i \neq j).
\end{align*}
\]
Let us combine these results with the expectations of \( (4.20) \). Let \( U_i = U_0 + \epsilon_i \), where \( U_0 \) is a \( U(0,1) \) random variable and \( \epsilon_i \) is the remainder term. Then, note that \( E\{(rT_i - 1)^{-1}\epsilon_j\}^{2} \leq E\{(rT_i - 1)^{-1}\} E(\nu^{-2}\epsilon_j^2) = o(\nu^{-3}) \) so that \( E\{(rT_i - 1)^{-1}\epsilon_j\} = o(\nu^{-3/2}) \). The results are obtained straightforwardly.

\[ \square \]

**Remark 12.** For the two-stage procedure \((3.5)-(3.6)\) with \((3.7)\), we have as \( W \to 0 \) that
\[
(i) \quad E_{\theta}(C_i^{-1}(N_i - C_i)) = (2\nu)^{-1}\{2s + B_i + A_i(f_i - 0.5) - 0.5 \sum_{j=1}^{k} f_j A_j\} + O(W^{3/2}),
\]
\[
(ii) \quad E_{\theta}(C_i^{-2}(N_i - C_i)^2) = (2\nu)^{-1}\{A_i(1 + 2f_i) + \sum_{j=1}^{k} f_j^2 A_j\} + O(W^{3/2}),
\]
where \( A_i = \text{tr}(\Sigma_i^2)/(\text{tr}(\Sigma_i))^2 \), \( B_i = \nu/C_i \), and \( s \) is a constant such that \( E_{\theta}(\hat{s}) = s + o(1) \).

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

**Proof.** In order to verify this lemma, we have to deal with the terms such as \( E_{\theta}(I_i) \), \( E_{\theta}(I_{ij}) \) and \( E_{\theta}(I_{ij\ell}) \), where
\[
I_i = \frac{\partial^3 g}{\partial u_i^3}\bigg|_{u=\epsilon} \left( \frac{N_i - C_i}{C_i} \right)^3, \quad I_{ij} = \frac{\partial^3 g}{\partial u_i^2 u_j}\bigg|_{u=\epsilon} \left( \frac{N_i - C_i}{C_i} \right)^2 \left( \frac{N_j - C_j}{C_j} \right), \quad I_{ij\ell} = \frac{\partial^3 g}{\partial u_i u_j u_\ell}\bigg|_{u=\epsilon} \left( \frac{N_i - C_i}{C_i} \right) \left( \frac{N_j - C_j}{C_j} \right) \left( \frac{N_\ell - C_\ell}{C_\ell} \right)
\]
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
\[
A_{i_1}^{-p_1} \xi_{i_2}^{-p_2} \cdots \xi_{i_k}^{-p_k}
\]
with \( A \geq 0, \ p_i \geq 0, \ r = 1, \ldots, k \), which are independent of \( d \). Let \( A \) also denote a generic positive constant, independent of \( d \). Let us write \( N_i^* = C_i^{-1}(N_i - C_i) \) for \( i = 1, \ldots, k \). Then, we obtain that
\[
|E_{\theta}(I_i)| \leq AE_{\theta}(\xi_{i_1}^{-p_1} \xi_{i_2}^{-p_2} \cdots \xi_{i_k}^{-p_k} |N_i^*|^{3}). \quad (4.22)
\]
We observe that \( \xi_{i} > C_i^{-1}m = \tau_i^{-1}T \) w.p.1 for all \( i = 1, \ldots, k \). Also, we observe that \( E_{\theta}(|N_i^*|^{3}) = O(\nu^{-3/2}) \) since \( E_{\theta}(|N_i|^{4}) = O(\nu^{-2}) \) from the facts that \( E_{\theta}\{(rT_i - 1)^3\} = O(\nu^{-2}) \), \( E_{\theta}\{(rT_i - 1)^4\} = O(\nu^{-1}) \), and
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 $E_\theta(|N_i^*|^2|N_j^*|) = O(\nu^{-3/2})$ and $E_\theta(|N_i^*||N_j^*||N_\ell^*|) = O(\nu^{-3/2})$ to show that $|E_\theta(I_{ij})| = O(\nu^{-3/2})$ and $|E_\theta(I_{ijp})| = O(\nu^{-3/2})$ for $1 \leq i < j < \ell \leq k$. Therefore, we conclude that $E_\theta(\mathcal{R}) = O(\nu^{-3/2}) = o(\nu^{-1})$. □

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


