著者 | 竹田和義, 青島真実
--- | ---
タイトル | 軽度の収縮方法を用いた試料数を決定するための交差構造の推定
氏名 | 田代学
種類 | 種類
日付 | 日付
ページ | ページ
年 | 年
(C) | (C) 2008 Elsevier B.V.
URL | URL

DOI | 10.1016/j.jspi.2008.03.032
Double shrink methodologies to determine the sample size via covariance structures

Kazuyoshi Yata\textsuperscript{a}, Makoto Aoshima \textsuperscript{b,*}

\textsuperscript{a}Graduate School of Pure and Applied Sciences, University of Tsukuba, Ibaraki, Japan

\textsuperscript{b}Institute of Mathematics, University of Tsukuba, Ibaraki, Japan

Abstract

We consider fixed–size estimation for a linear function of mean vectors from $\pi_i : N_p(\mu_i, \Sigma_i), i = 1, ..., k$, when every $\Sigma_i$ has some structure. The goal of inference is to construct a fixed–span confidence region with required accuracy. We find a sample size for each $\pi_i$ with the help of the ‘double shrink methodology’, that is introduced by this paper, via covariance structures of $\Sigma_i, i = 1, ..., k$. We estimate the sample size in a two–stage sampling and give a fixed–span confidence region that has the coverage probability approximately second–order consistent with the required accuracy. Some simulations are carried out to see moderate sample size performances of the proposed methodologies.

Key words: Asymptotic uniformity; Intraclass correlation model; Optimal sample size; Two–stage sampling

1 Introduction

Suppose that we have $\pi_i : N_p(\mu_i, \Sigma_i), i = 1, ..., k$, independent, normally distributed populations, having unknown mean vector $\mu_i$ and unknown covariance matrix $\Sigma_i$. We assume that $\Sigma_i$ for every $i$ has the following structure

\textsuperscript{*} 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.

\textsuperscript{*} Corresponding author. Institute of Mathematics, University of Tsukuba, Ibaraki 305-8571, Japan; Fax: +81-298-53-6501.

Email address: aoshima@math.tsukuba.ac.jp (Makoto Aoshima).

Preprint submitted to Elsevier 18 April 2008
with unknown positive scalars \( \sigma_{ij}, \ j = 1, \ldots, \ell \):

\[
\Sigma_i = \sigma_{i1}^2 A_1 + \cdots + \sigma_{i\ell}^2 A_\ell
\]  

(1)

with a fixed \( \ell \) \((1 \leq \ell \leq p)\), where \( A_j \) is a \( p \times p \) known symmetric matrix with rank \( r_j \) having \( 1 \leq r_1 \leq \cdots \leq r_\ell \), \( \sum_{j=1}^\ell r_j = p \) and \( \Sigma_{ij} = I_p \).

A special case of the structure is the intraclass correlation model defined by \( \Sigma_i = \sigma_{i1}^2 \{ (1 - \rho_i) I_p + \rho_i 11^T \}, \ 1 = (1, \ldots, 1)^T: \ A_1 = p^{-1} 11^T (r_1 = 1), \ A_2 = I_p - A_1 (r_2 = p - 1), \ \sigma_{11}^2 = \sigma_{i1}^2 (p \rho_i + 1 - \rho_i), \ \sigma_{i2}^2 = \sigma_{i1}^2 (1 - \rho_i) \) and \( \ell = 2 \).

Let \( X_{i1}, X_{i2}, \ldots \) be a sequence of independent and identically distributed (i.i.d.) random vectors from each \( \pi_i \). Having recorded \( X_{i1}, X_{i2}, \ldots, X_{in_i} \) for each \( \pi_i \), let us write \( \bar{X}_{i n_i} = \sum_{n_i=1} X_{i n_i} / n_i \) and \( n = (n_1, \ldots, n_k) \). We are interested in estimating the linear function \( \theta = \sum_{i=1}^k b_i \mu_i \), where \( b_i \)'s are known and nonzero scalars. Let \( T_n = \sum_{i=1}^k b_i \bar{X}_{in_i} \). Given \( d > 0 \), we define

\[
R_n = \{ \mu \in \mathbb{R}^p : ||T_n - \mu|| \leq d \},
\]

(2)

where \( || \cdot || \) is the Euclidean norm. Now, with some given \( \alpha \in (0, 1) \), our goal is to construct a fixed–span confidence region \( R_n \) such that

\[
P_\theta(\mu \in R_n) \geq 1 - \alpha \quad \text{for all} \ \theta,
\]

(3)

where \( \theta = (\mu_1, \ldots, \mu_k, \sigma_{11}, \ldots, \sigma_{1\ell}, \ldots, \sigma_{k1}, \ldots, \sigma_{k\ell}) \).

From (1), there exists a \( p \times p \) orthogonal matrix \( H \) such that

\[
H^T \Sigma_i H = \begin{pmatrix} \sigma_{i1}^2 I_{r_1} & \quad O \\ & \ddots \\ O & \quad \sigma_{i\ell}^2 I_{r_\ell} \end{pmatrix}
\]

for all \( \pi_i, \ i = 1, \ldots, k \). Then, \( H^T (T_n - \mu) \) is distributed as \( N_p(0, \sum_{i=1}^k n_i^{-1} \sigma_{i1}^2 H^T \Sigma_i H) \). Let \( V_j \ (j = 1, \ldots, \ell) \) be a mutually independent chi–square random variable with \( r_j \ (j = 1, \ldots, \ell) \) degrees of freedom (d.f.). We note that

\[
P_\theta(\mu \in R_n) = P_\theta \left( \sum_{j=1}^\ell \sum_{i=1}^k b_i^2 n_i \sigma_{ij}^2 V_j \leq d^2 \right).
\]

(4)

The purpose of this paper is to find the sample size for each \( \pi_i \) in order to have the coverage probability, given by (4), satisfying requirement (3).

There are many literatures related to this inference problem. For instance, see Aoshima (2005) and Ghosh et al. (1997) for a review. We, however, emphasize on a previous work given by Aoshima and Takada (2006) for the present problem. They gave a method to find the sample size for each \( \pi_i \) and constructed a region \( R_n \) satisfying requirement (3) in a Stein’s (1945) two–stage sampling scheme. They proposed an approximation to the sample size in order to overcome the complexity of its calculation and showed that the approximation
gives a region $R_n$ with accuracy approximately equal to $1 - \alpha$. In this paper, we propose a different method, the ‘double shrink methodology’, to find the sample size for each $\pi_i$. It is shown that our findings in this paper improve both the sample size and the approximation to the coverage probability.

In Section 2, we introduce a new methodology, the ‘double shrink methodology’, to find the sample size for each $\pi_i$ via covariance structures. In Section 3, we give an approximation to the coverage probability along with an expansion formula of the sample size to satisfy requirement (3). In Section 4, we numerically show that the double shrink methodology, proposed in this paper, improves both the sample size and the approximation to the coverage probability in several simulation studies.

2 Double shrink methodology

In this section, we introduce a new methodology to find the sample sizes satisfying requirement (3) via covariance structures. We name it the double shrink methodology. Let us explain how it finds the sample sizes to satisfy requirement (3) via covariance structures. We name it the double shrink methodology. Let us explain how it finds the sample sizes to satisfy requirement (3) via covariance structures.

Let $F_p(\cdot)$ denote the cumulative distribution function of a chi–square random variable with $p$ d.f. Let $a$ be the constant such that $F_p(a) = 1 - \alpha$. We tentatively consider $n_i = (a/d^2)|b_i|\sigma_{i1} \sum_{\ell=1}^{k-1} |b_{i\ell}|\sigma_{i\ell+1}$ that yields the smallest sum $\sum_{k=1}^{k-1} n_i$ to hold that $\sum_{k=1}^{k-1} b_{i\ell}\sigma_{i\ell}^2/n_i = d^2/a$. Let $\delta_j = (\sum_{\ell=1}^{k-1} |b_{i\ell}|\sigma_{i\ell+1})^{-1} \sum_{k=1}^{k-1} |b_{i\ell}|\sigma_{ij}^2/\sigma_{i1}$, $j = 2, ..., \ell$. Let $\delta = \max_{j=2,...,\ell} \delta_j$. If $\delta \leq 1$, we have from (4) that

$$P_{\theta}(\mu \in R_n) = P_{\theta}\left(V_1 + \sum_{j=2}^{\ell} \delta_j V_j \leq a\right) \geq P_{\theta}\left(\sum_{j=1}^{\ell} V_j \leq a\right) = F_p(a) = 1 - \alpha.$$ 

If $\delta > 1$, we modify the sample size as $n_i = (a\delta/d^2)|b_i|\sigma_{i1} \sum_{\ell=1}^{k-1} |b_{i\ell}|\sigma_{i\ell+1}$ in order to have that

$$P_{\theta}(\mu \in R_n) = P_{\theta}\left(\frac{1}{\delta} V_1 + \sum_{j=2}^{\ell} \delta_j V_j \leq a\right) \geq P_{\theta}\left(\sum_{j=1}^{\ell} V_j \leq a\right) = F_p(a) = 1 - \alpha.$$ 

When we consider both the cases of $\delta$ simultaneously, the sample size required to satisfy requirement (3) is given by $n_i = (a/d^2)|b_i|\sigma_{i1} \max_{1 \leq j \leq \ell} \sum_{\ell=1}^{k-1} |b_{i\ell}|\sigma_{i\ell+1}^2/\sigma_{i1}$. 

Figures 1–3 summarize the idea of the double shrink methodology in the case that $p = 2$, $\ell = 2$ and $\delta = \delta_2 > 1$. In Figure 1, the ellipse indicates a $1 - \alpha$ confidence region for $\mu$ and the shaded circle indicates the fixed region with the radius of $d$ from the centre $\mu$. When the sample is taken up to size $n_i = (a/d^2)|b_i|\sigma_{i1} \sum_{\ell=1}^{k-1} |b_{i\ell}|\sigma_{i\ell+1}$ for each $\pi_i$, the confidence region is shrunk up to the circle along the eigenvector having axis number $l = 1$ as seen in Figure 2. Next, when the sample is additionally taken until the total sample size
for each \( \pi_i \) becomes of size \( n_i = (a/d^2)|b_i|\sigma_{i1}^{\max}1\leq j \leq \ell \sum_{i'=1}^k |b_{i'}|\sigma_{i'j}^2/\sigma_{i'1} \), the confidence region is shrunk up to the circle along the eigenvector having axis number \( l = 2 \) as seen in Figure 3.

On the other hand, Aoshima and Takada (2006) considered the sample size given by

\[
n_i \geq \frac{a}{d^2} \max_{1 \leq j \leq \ell} (|b_i|\sigma_{ij} \sum_{i'=1}^k |b_{i'}|\sigma_{i'j}) \quad (= \tilde{n}_i, \text{ say}) \tag{5}
\]

for each \( \pi_i \). As seen in Figure 4, the confidence region of such a size is included in the circle and it may not come in contact with the circle.

Comparing Figure 4 with Figure 3, it should be noted that the double shrink methodology improves the approximation to the coverage probability in (4). One can expect that the double shrink methodology successfully prevents the confidence region from both meeting requirement (3) excessively and oversampling too much.

Generally, for a fixed \( l (1 \leq l \leq \ell) \), the double shrink methodology finds the sample size for each \( \pi_i \) as

\[
n_i \geq \frac{a}{d^2} |b_i|\sigma_{il}^{\max}1\leq j \leq \ell \sum_{i'=1}^k |b_{i'}|\sigma_{i'lj}/\sigma_{i'l} \quad (= C_i, \text{ say}) \tag{6}
\]

Then, the region \( R_n \) given by (2) satisfies requirement (3). Since \( \sigma_{ij} \)'s are unknown, it is necessary to estimate \( C_i \)'s in (6) with some pilot samples. We consider a two–stage estimation methodology to determine the sample sizes \( n \) under the following assumptions: For fixed \( l (1 \leq l \leq \ell) \),

(A1) There exists some \( j_l \) such that \( \sum_{i=1}^k |b_i|\sigma_{ij_l}^2/\sigma_{ij} > \sum_{i=1}^k |b_i|\sigma_{ij}^2/\sigma_{ij} \) for all \( j (1 \leq j \neq j_l \leq \ell) \);

(A2) There exists a known and positive lower bound \( \sigma_{il}^{*} \) for \( \sigma_{il} \) such that \( \sigma_{il} > \sigma_{il}^{*} \), \( i = 1, \ldots, k \)

(cf. Mukhopadhyay and Duggan, 1999).

1. Having \( m_0 (\geq 4) \) fixed, define

\[
m = \max \left\{ m_0, \left[ \frac{a}{d^2} \min_{1 \leq i \leq k} |b_i|\sigma_{il}^{*} \sum_{i'=1}^k |b_{i'}|\sigma_{i'l} \right] + 1 \right\}, \tag{7}
\]

where \( [x] \) denotes the largest integer less than \( x \). According to (7), take a pilot sample \( \mathbf{X}_{is}, \ s = 1, \ldots, m \) of size \( m \) and calculate \( \mathbf{S}_i = \nu^{-1} \sum_{s=1}^m (\mathbf{X}_{is} - \mathbf{X}_{im})(\mathbf{X}_{is} - \mathbf{X}_{im})^T \) for all \( \pi_i, \ i = 1, \ldots, k \), where \( \mathbf{X}_{im} = \sum_{s=1}^m \mathbf{X}_{is}/m \) and \( \nu = m - 1 \). Then, define

\[
S_{ij}^2 = r_j^{-1} \text{tr} (\mathbf{A}_j \mathbf{S}_i) \tag{8}
\]

as an unbiased estimate of \( \sigma_{ij}^2 \) for all \( i = 1, \ldots, k; \ j = 1, \ldots, \ell \). Define the total
sample size for all \( \pi_i, \ i = 1, \ldots, k \), by

\[
N_i = \max \left\{ m, \left[ \frac{u_l}{d^2} |b_i| S_{it} \max_{1 \leq j \leq \ell} \sum_{i' = 1}^k \left| b_{i'} \right| \frac{S_{ij}^2}{S_{ij}} \right] + 1 \right\},
\]

(9)

where \( u_l (>0) \) is determined later. Let \( \mathbf{N} = (N_1, \ldots, N_k) \).

2. Take an additional sample \( \mathbf{X}_{is}, \ s = m + 1, \ldots, N_i \) for each \( \pi_i \). By combining the initial sample and the additional sample, calculate \( \tilde{X}_{iN_i} = \left( \sum_{i' = 1}^k b_{i'} \overline{X}_{iN_i} \right) \) for each \( \pi_i \). Then, define the region \( R_{\mathbf{N}} \) as in (2) with \( T_{\mathbf{N}} = \sum_{i=1}^k b_i \tilde{X}_{iN_i} \).

3 Second–order approximations

For fixed \( l (1 \leq l \leq \ell) \), let \( j_l \) and \( j_{lm} \) be the indices such that

\[
\max_{1 \leq j \leq \ell} \sum_{i = 1}^k |b_i| \frac{\sigma_{ij}^2}{\sigma_{il}} = \sum_{i = 1}^k |b_i| \frac{\sigma_{ij_l}^2}{\sigma_{il}} \quad \text{in (6)},
\]

(10)

\[
\max_{1 \leq j \leq \ell} \sum_{i = 1}^k \left| b_i \right| \frac{S_{ij}^2}{S_{il}} = \sum_{i = 1}^k \left| b_i \right| \frac{S_{ij_m}^2}{S_{il}} \quad \text{in (9)},
\]

respectively. Throughout, we write that

\[
\tau_{ls} = \min_{1 \leq i \leq k} \left| b_i \right| \frac{\sigma_{il}^*}{\sigma_{il}} \sum_{i' = 1}^k \left| b_{i'} \right| \sigma_{i' l*},
\]

\[
f_i = \left| b_i \right| \frac{\sigma_{ij_l}^2 / \sigma_{il}}{\left( \sum_{i' = 1}^k \left| b_{i'} \right| \sigma_{i' j_l}^2 / \sigma_{i' l} \right)^{-1}} \quad (i = 1, \ldots, k).
\]

**Theorem 3.1.** Assume that \( a \geq p - 2 \). Choose \( u_l \) in (9) as \( u_l = a(1 + \nu^{-1} \bar{s}_l) \) where

\[
\bar{s}_l = \begin{cases} \frac{1}{r_l} + \frac{(a-p) \sum_{i=1}^k b_i^2 S_{il}^2 - r_l k r_{ls}}{2 r_l (\sum_{i=1}^k \left| b_i \right| S_{il})^2} & (j_{lm} = l), \\ \left( \sum_{i=1}^k \left| b_i \right| \frac{S_{ij_m}^2}{S_{il}} \right)^{-2} & (j_{lm} \neq l) \end{cases}
\]

(11)

with \( S_{ij}^2 \)’s given by (8). Then, the region \( R_{\mathbf{N}} \) is asymptotically second–order consistent as \( d \to 0 \) in the sense that

\[
P_{\theta}(\mu \in R_{\mathbf{N}}) \geq E_{\theta} \left\{ F_p \left( d^2 \left( \max_{1 \leq j \leq \ell} \sum_{i = 1}^k \left| b_i \right| \frac{S_{ij}^2}{S_{ij}} \right)^{-1} \right) \right\} = 1 - \alpha + o(d^2) \quad \text{for all } \theta.
\]

(12)
For the first term in (13), we use the Taylor expansion to claim that

$$\max_{1 \leq j \leq \ell} \sum_{i=1}^{k} \frac{b_i^2 \sigma_{ij}^2}{N_i} = \sum_{i=1}^{k} \frac{b_i^2 \sigma_{ijn}^2}{N_i}.$$ 

We have from (4) that

$$P_{\theta}(\mu \in R_N) \geq E_{\theta} \left\{ F_p \left( d^2 \left( \sum_{i=1}^{k} \frac{b_i^2 \sigma_{ijn}^2}{N_i} \right)^{-1} \right) \right\}$$

$$= E_{\theta} \left\{ F_p \left( a \left( \sum_{i=1}^{k} f_i C_i \frac{C_i}{N_i} \right)^{-1} \right) \right\} + E_{\theta} \left\{ I_{\{i \neq j, N\}} F_p \left( d^2 \left( \sum_{i=1}^{k} \frac{b_i^2 \sigma_{ijn}^2}{N_i} \right)^{-1} \right) \right\},$$

where $I_{\{i \neq j, N\}}$ is the indicator function. Now, let us define a new function as follows.

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

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

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

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

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

For the first term in (13), we use the Taylor expansion to claim that

$$E_{\theta} \left\{ F_p \left( a \left( \sum_{i=1}^{k} f_i C_i \frac{C_i}{N_i} \right)^{-1} \right) \right\} = E_{\theta} \left\{ g \left( \frac{N_1}{C_1}, ..., \frac{N_k}{C_k} \right) \right\}$$

$$= 1 - \alpha + a F_p'(a) \sum_{i=1}^{k} f_i E_{\theta} \left( \frac{N_i - C_i}{C_i} \right)$$

$$+ \frac{a}{2} \sum_{i=1}^{k} (a F_p''(a) f_i^2 + 2 F_p'(a) f_i^2 - 2 F_p'(a) f_i) E_{\theta} \left\{ \left( \frac{N_i - C_i}{C_i} \right)^2 \right\}$$

$$+ \frac{a}{2} \sum_{i \neq j} (a F_p''(a) f_i f_j + 2 F_p'(a) f_i f_j) E_{\theta} \left\{ \left( \frac{N_i - C_i}{C_i} \right) \left( \frac{N_j - C_j}{C_j} \right) \right\}$$

$$+ E_{\theta} (\mathbb{R}),$$

(14)
Combining (18) and (19) with (13), the result can be obtained.

For the second term in (13), with the help of Lemma 6, we evaluate as

\[
\sum_{i,j \neq j_N} E_\theta \left\{ I_{(j_i \neq j_N)} F_p \left( d^2 \left( \sum_{i=1}^k b_i^2 \sigma_{i,j_i}^2 \right) N_i \right)^{-1} \right\} \leq P_\theta(j_i \neq j_N) = O(d^6).
\]

Combining (18) and (19) with (13), the result can be obtained.

\[
E_\theta(\mathcal{R}) = \frac{1}{6} \sum_{i,j} E_\theta \left\{ \frac{\partial^3 g}{\partial u_i \partial u_j \partial u_t} \right\} \mathbf{u} \cdot \mathbf{\xi} \left( \frac{N_i - C_i}{C_i} \right) \left( \frac{N_j - C_j}{C_j} \right) \left( \frac{N_t - C_t}{C_t} \right) \right \}
\]
\textbf{Theorem 3.2.} The two–stage procedure (7)–(9) with (11) has as \(d \to 0:\)

\[ E_\Theta(N_i - C_i) = \begin{cases}
(2\tau_i r_i)^{-1} \{ |b_i| \sigma_{d} \sum_{i'=1}^{k} |b_{i'}| \sigma_{d^*} + (a - p) f_i \sum_{i'=1}^{k} \sigma_{i'\ell}^2 + b_i^2 \sigma_{i\ell}^2 \} \\
+ \frac{1}{2} \left( 1 - k f_i \right) + o(1) \quad \text{for } i = 1, \ldots, k \quad (j_i = \ell),
\end{cases} \]

\[ \begin{cases}
(2\tau_i r_{ji})^{-1} \left( \left( a + 2 - p \right) \frac{\sigma_{i'j}^2}{\sigma_{i'\ell}} f_i \sum_{i'=1}^{k} \sigma_{i'\ell} \frac{p_i}{\tau_i} \left( |b_i| \sigma_{d} \sum_{i'=1}^{k} |b_{i'}| \sigma_{d^*} \right) \right) \\
+ \frac{1}{2} \left( 1 - \frac{\sigma_{ij}^2}{\sigma_{ij^*}} \right) + o(1) \quad \text{for } i = 1, \ldots, k \quad (j_i \neq \ell).
\end{cases} \]

\textbf{Proof.} The results are obtained by Lemma 5 in Appendix straightforwardly. \(\square\)

Now, in the two–stage procedure (7)–(9) with (11), one may determine the axis number \(l\) \((1 \leq l \leq \ell)\) so as to minimize the sum \(\sum_{i=1}^{k} N_i\) as follows: Let \(l_o\) denote the axis number \(l\) that minimizes the sum \(\sum_{i=1}^{k} C_i\) in (6). We assume that \(l_o\) is determined uniquely. For fixed \(l_o\), let \(j_o\) be the index such that

\[ \frac{\max_{1 \leq j \leq \ell} \sum_{i=1}^{k} |b_i| \sigma_{ij}^2}{\sigma_{ij}} = \sum_{i=1}^{k} \frac{|b_i| \sigma_{ij^*}^2}{\sigma_{ij^*}}. \]

Let \(\tau_{ijl} = |b_i| \sigma_{d} \sum_{i'=1}^{k} |b_{i'}| \sigma_{i'\ell}^2 / \sigma_{d^*} \) \((i = 1, \ldots, k; j = 1, \ldots, \ell; l = 1, \ldots, \ell)\). We assume for (A1)–(A2) that

(A1') There exists some \(j_i\) for each \(l\) such that \(\tau_{ijl} > \tau_{ijj}\) for all \(j\) \((1 \leq j \neq j_i \leq \ell)\);

(A2') There exists a known and positive lower bound \(\sigma_{ij^*}\) for \(\sigma_{ij}\) such that

\[ \sigma_{ij} > \sigma_{ij^*}, \quad i = 1, \ldots, k; \ j = 1, \ldots, \ell. \]

Having \(m_0\) \((\geq 4)\) fixed, we start with a pilot sample of size \(m\) as

\[ m = \max \left\{ m_0, \frac{a}{d^2} \min_{1 \leq i \leq \ell} |b_i| \sigma_{ij^*} \sum_{i'=1}^{k} |b_{i'}| \sigma_{i'j^*} \right\} + 1 \quad (20) \]

on behalf of (7). We carry out the first–stage sampling of the two–stage procedure and calculate the sum of the required total sample sizes, \(\sum_{i=1}^{k} N_i\) for each \(l\) \((1 \leq l \leq \ell)\), using (9) after replacing \(\tau_{i*}\) with

\[ \tau_{o*} = \min_{1 \leq j \leq \ell} |b_i| \sigma_{ij^*} \sum_{i'=1}^{k} |b_{i'}| \sigma_{i'j^*} \]

in (11). Now, find the number \(l\) \(\left(= l_{om}, \text{ say}\right)\) that gives the minimum sum \(\sum_{i=1}^{k} N_i\) and fix \(l_{om}\) as the axis number hereafter. One may utilize the two–stage estimation methodology stated in Section 2 after replacing both (7) and
respectively.

**Theorem 3.3.** Assume that \( a \geq p - 2 \). For the axis number \( l_{om} \), let \( j_{om} \) be the index such that

\[
\max_{1 \leq j \leq t} \sum_{i=1}^{k} |b_i| \frac{S_{ij}^2}{S_{i\ell_{om}}} = \sum_{i=1}^{k} |b_i| \frac{S_{ijom}^2}{S_{i\ell_{om}}}. 
\]

Choose \( u_{l_{om}} \) in (21) as \( u_{l_{om}} = a(1 + \nu^{-1} \hat{s}_{l_{om}}) \) where

\[
\hat{s}_{l_{om}} = \begin{cases} 
\frac{1}{r_{l_{om}}} + \frac{(a-p) \sum_{i=1}^{k} S_{ijam}^2 - r_{l_{om}} k r_{o}}{2 r_{l_{om}} (\sum_{i=1}^{k} |b_i| S_{i\ell_{om}})^2} & (j_{om} = l_{om}), \\
\left( \sum_{i=1}^{k} |b_i| \frac{S_{ijom}^2}{S_{i\ell_{om}}} \right)^{-2} \left( \frac{a+2-p}{2} \sum_{i=1}^{k} b_i^2 \frac{S_{ijom}^2}{S_{i\ell_{om}}} - \frac{r_{o}}{2} \sum_{i=1}^{k} \frac{S_{ijom}^2}{S_{i\ell_{om}}} \right) & (j_{om} \neq l_{om})
\end{cases}
\]

with \( S_{ij}^2 \)'s given by (8). For the two-stage procedure (20)–(21), the region \( R_\mathbf{N} \) is asymptotically second-order consistent as \( d \to 0 \) as stated in (12).

**Proof.** With the help of Lemma 8 in Appendix, the result can be obtained similarly to Theorem 3.1. \( \square \)

**Theorem 3.4.** The two-stage procedure (20)–(21) with (22) has as \( d \to 0 \):

\[
E \theta(N_i - C_i) = \begin{cases} 
(2 r_{o} r_{l_{o}})^{-1} \left\{ |b_i| \sigma_{il_{o}} \sum_{i' = 1}^{k} |b_{i'}| \sigma_{i'l_{o}} + (a - p) f_{oi} \sum_{i' = 1}^{k} b_{i'}^2 \sigma_{i'l_{o}}^2 + b_{i}^2 \sigma_{i{l_{o}}}^2 \right\} \\
+ \frac{1}{2} (1 - k f_{o}) + o(1) & \text{for } i = 1, \ldots, k \\
(2 r_{o} r_{j_{o}})^{-1} \left\{ (a + 2 - p) \frac{\sigma_{ij_{o}}^2}{\sigma_{i'l_{o}}^2} f_{oi} \sum_{i' = 1}^{k} b_{i'}^2 \sigma_{i'l_{o}}^2 + r_{o} \frac{\sigma_{ij_{o}}^2}{\sigma_{i'l_{o}}^2} \right\} \\
+ \frac{1}{2} \left( 1 - \frac{\sigma_{ij_{o}}^2}{\sigma_{j_{o}'l_{o}}^2} \right) f_{oi} \sum_{i' = 1}^{k} \frac{\sigma_{ij_{o}}^2}{\sigma_{i'l_{o}}^2} + o(1) & \text{for } i = 1, \ldots, k
\end{cases} 
\]

where

\[
f_{oi} = \left| b_i \right| \left( \frac{\sigma_{ij_{o}}^2}{\sigma_{i'l_{o}}^2} \right) \left( \sum_{i' = 1}^{k} |b_{i'}| \sigma_{i'l_{o}}^2 \right)^{-1}
\]

**Proof.** The results are obtained by Lemmas 5 and 8 in Appendix straightforwardly. \( \square \)
Remark 1. If there exists a number \( l \) such that \( j_1 = l \) in (10), we have that \( \tilde{n}_i \geq C_i \), where \( \tilde{n}_i \) is given by (5) and \( C_i \) is given by (6) with such \( l \). It follows for \( \tilde{n}_i \)'s that

\[
F_p \left( d^2 \left( \max_{1 \leq j \leq l} \sum_{i=1}^{k} \frac{b_{ij}^2 \sigma_{ij}^2}{\tilde{n}_i} \right)^{-1} \right) \geq 1 - \alpha.
\]

Therefore, a region \( R_N \) with a sample size estimating \( \tilde{n}_i \) for each \( \pi_i \) can no longer satisfy (12). If \( \tilde{n}_i \) is modified by

\[
\tilde{n}_i^* = \frac{a}{d^2} \left( \max_{1 \leq j \leq l} \frac{\sum_{i'=1}^{k} \left| b_{ij} \right| \sigma_{ij} \sum_{i'=1}^{k} \left| b_{ij'} \right| \sigma_{ij'}}{\tilde{n}_i} \right)^{-1} \max_{1 \leq j \leq l} \left( \left| b_{ij} \right| \sigma_{ij} \sum_{i'=1}^{k} \left| b_{ij'} \right| \sigma_{ij'} \right),
\]

one has that

\[
F_p \left( d^2 \left( \max_{1 \leq j \leq l} \sum_{i=1}^{k} \frac{b_{ij}^2 \sigma_{ij}^2}{\tilde{n}_i^*} \right)^{-1} \right) = 1 - \alpha.
\]

Therefore, a region \( R_N \) with a sample size estimating \( \tilde{n}_i^* \) for each \( \pi_i \) should satisfy (12). However, it would be much complicated to estimate \( \tilde{n}_i^* \).

Remark 2. Let us consider the case that there exists a common and known number \( j_1 \) for all \( \pi_i \) such that \( \sigma_{ij} > \sigma_{ij} \) for all \( j \) \((1 \leq j \neq j_1 \leq l)\) in (1). One can reduce the two–stage procedure (7)–(9) with (11) as follows: We have for any fixed \( j_2 \) \((1 \leq j_2 \neq j_1 \leq l)\) that

\[
\sum_{i=1}^{k} \left| b_{ij} \sigma_{ij_2} \right| \max_{1 \leq j \leq l} \left( \sum_{i'=1}^{k} \left| b_{ij'} \sigma_{ij_2} \right| \sigma_{ij_2} \right) = \sum_{i=1}^{k} \left| b_{ij} \sigma_{ij_2} \right| \sum_{i'=1}^{k} \left| b_{ij'} \sigma_{ij_2} \right| \sigma_{ij_2} \sigma_{ij_2} = \sum_{i=1}^{k} \left| b_{ij} \sigma_{ij_2} \right| \sigma_{ij_2} \sigma_{ij_2} \geq 2 \sum_{i=1}^{k} \left| b_{ij} \sigma_{ij_2} \right| \sigma_{ij_2} \sigma_{ij_2} \geq 2 \sum_{i=1}^{k} \left| b_{ij} \sigma_{ij_2} \right| \sigma_{ij_2} \sigma_{ij_2} = \sum_{i=1}^{k} \left| b_{ij} \sigma_{ij_2} \right| \sigma_{ij_2} \sigma_{ij_2}.
\]

So, one naturally chooses \( j_1 \) as the axis number \( l \) in (6) in order to have the minimum sum \( \sum_{i=1}^{k} C_i \). Under (A1)–(A2), with \( j_1 \) for \( l \), we carry out the first–stage sampling with (7). Now, define the total sample size of each \( \pi_i \) as

\[
N_i = \max \left\{ m, \left[ \frac{n_1}{d^2} \left| b_{ij_1} \right| S_{ij_1} \sum_{i'=1}^{k} \left| b_{ij'} \right| S_{ij'} \right] + 1 \right\}, \tag{23}
\]

10
where \(u_1\) is given by

\[
    u_1 = a \left( 1 + \frac{1}{\nu} \left( \frac{1}{r_{j_1}} + \frac{(a - p)\sum_{i=1}^{k} b_i^2 s_{i1}^2 - r_{j_1} k_{j_1}}{2r_{j_1} \left( \sum_{i=1}^{k} b_i^2 s_{i1}^2 \right)^2} \right) \right)
\]

with \(\tau_{j_1} = \min_{1 \leq i \leq k} |b_i| \sigma_{j_1, *}, \sum_{i=1}^{k} |b_i| \sigma_{j_1, *}. \) We carry out the second–stage sampling with (23)–(24). Then, this reduced procedure holds the asymptotic characteristic stated in Theorem 3.1. On the other hand, the total sample size due to Aoshima and Takada (2006) is naturally reduced to (23) with

\[
    u_1 = a + \frac{a}{2p\nu} \left( \frac{p + 2\ell}{p + 2} + 2k\ell - p \right).
\]

**Remark 3.** When the structure (1) specifies an intraclass correlation model as \(\Sigma_i = \sigma_i^2 \{ (1 - \rho_i)I_p + \rho_i I_1^T \}, \) the parameters in (1) are given by \(\sigma_i^2 = \sigma_i^2 (p\rho_i + 1 - \rho_i) \) and \(\sigma_i^2 = \sigma_i^2 (1 - \rho_i) \) with \(r_1 = 1, r_2 = p - 1 \) and \(\ell = 2. \) Let us consider the case where \(\rho_i > 0 \) for all \(\pi_i, \) such as a large dimensional case. Since \(\sigma_i > \sigma_i^2 \) for all \(\pi_i, \) one may follow the two–stage procedure (23)–(24) in Remark 2. Then, a candidate for \(\sigma_{i1,*} \) appearing in (24) is given by \(\sigma_{i1,*} = \sigma_{i,*} \) with \(\sigma_{i,*} \) a known and positive lower bound for \(\pi_i, \)’s standard deviation.

**Remark 4.** One may apply Theorems 3.1 and 3.2 to constructing a fixed–size ellipsoidal confidence region for regression parameters. Let us consider a linear regression model as follows:

\[
    Y_i = x_i^T \beta + \varepsilon_i, \quad i = 1, 2, \ldots,
\]

where \(x_i, \)’s are known \(p–vectors, \) \(\beta \) is an unknown \(p–vector \) of regression parameters, and \(\varepsilon_i, \)’s are i.i.d. \(N(0, \sigma^2) \) with unknown \(\sigma (> 0). \) Let us denote \(y_n = (Y_1, \ldots, Y_n) \) and \(X_n = (x_1, \ldots, x_n), \) and assume that rank \((X_n) = p (\leq n). \) Having recorded \((x_i, Y_i), \) \(i = 1, \ldots, n \) with \(n > p, \) we estimate \(\beta \) by the least square estimator \(\hat{\beta}_n = (X_n^T X_n)^{-1} X_n^T y_n. \) Along the lines of Ghosh et al. (pp. 198–200, 1997), a fixed–size ellipsoidal confidence region for \(\beta \) is defined by

\[
    R_{\beta_n} = \{ \beta \in \mathbb{R}^p : (\hat{\beta}_n - \beta)^T (n^{-1} X_n^T X_n)(\hat{\beta}_n - \beta) \leq d^2 \}
\]

for given \(d (> 0) \), and our goal is to hold \(P_{\beta, \sigma^2}(\beta \in R_{\beta_n}) \geq 1 - \alpha \) for all \(\beta, \sigma^2 \) for given \(\alpha \in (0, 1). \) It is easy to see that \(P_{\beta, \sigma^2}(\beta \in R_{\beta_n}) = F_p(nd^2/\sigma^2), \) so that \(n \) has to be \(n \geq a \sigma^2/d^2 \) (\(= C, \) say). We assume that there exists a known and positive lower bound \(\sigma_* \) for \(\sigma \) such that \(\sigma > \sigma_* ^2. \) One may estimate \(\sigma^2 \) by \(S^2 = (m - p)^{-1} (y_m - X_m \beta_m)^T (y_m - X_m \beta_m) \) with a pilot sample of size \(m = \max \{ m_0, \lfloor a \sigma_*^2/d^2 \rfloor + 1 \} \) where \(m_0 = \max \{ 4, p + 1 \}. \) Then, the total sample size of the two–stage estimation methodology is defined by \(N = \max \{ m, \lfloor u S^2/d^2 \rfloor + 1 \} \), where

\[
    u = a(1 + \nu^{-1} \hat{s}) \quad \text{with} \quad \nu = m - p \quad \text{and} \quad \hat{s} = (a - p + 2)/2 - 0.5(\sigma_*^2/S^2).
\]
Then, from Theorems 3.1 and 3.2 (with \( k = 1, \ell = 1 \) and \( r_1 = 1 \)), it holds as \( d \to 0 \) that

\[
\begin{align*}
P_{\beta, \sigma^2}(\beta \in R_{\beta_N}) &= 1 - \alpha + o(d^2) \quad \text{for all } (\beta, \sigma^2), \\
E_{\beta, \sigma^2}(N - C) &= (2\sigma^2_i)^{-1}\sigma^2(a - p + 2) + o(1).
\end{align*}
\]

4 Moderate sample performances

In order to study the performance of the two-stage procedure, we take resort to computer simulations. We fix \( p = 4 \) and \( b_i = 1, i = 1, \ldots, k \). Our goal is to construct 90\% fixed-span confidence regions for \( \mu = \mu_1 + \cdots + \mu_k \). In other words, we have \( \alpha = 0.1 \) (that is, \( a = 7.779 \)). Independent pseudorandom normal observations from \( \pi_i : N_4(\mu_i, \Sigma_i), i = 1, \ldots, k \), were generated where \( \Sigma_i \)'s were fixed as \( \Sigma_i = \sigma_i^2\{1 - \rho_i\mathbf{I}_4 + \rho_i\mathbf{1}\mathbf{1}^T\} \). Then, \( \ell = 2, (r_1, r_2) = (1, 3), \sigma_1^2 = \sigma_i^2, \rho_1 = 1 - \rho_i \) and \( \sigma_2^2 = \sigma_i^2(1 - \rho_i) \), \( i = 1, \ldots, k \). We choose \( \sigma_i = 1, \ i = 1, \ldots, k \). We set \( \sigma_{ij*}^2 = 0.7\sigma_{ij}^2, \ i = 1, \ldots, k; \ j = 1, 2 \), with \( \rho_i \) fixed in each table.

In Table 1, pretending that \( \rho_i > 0, \ i = 1, \ldots, k \), we compare the performance of the two-stage procedure in Remark 2 with (24), which is given in the first block, with that with (25) due to Aoshima and Takada (2006), which is given in the second block. Let \( C = \sum_{i=1}^k C_i \). We consider three cases that (i) \( k = 2, (\rho_1, \rho_2) = (1/10, 1/5) \) and \( C = 90 \); (ii) \( k = 3, (\rho_1, \rho_2, \rho_3) = (1/10, 3/20, 1/5) \) and \( C = 135 \); (iii) \( k = 4, (\rho_1, \rho_2, \rho_3, \rho_4) = (1/10, 2/15, 1/6, 1/5) \) and \( C = 180 \), whereas with a fixed value of \( C \) one easily obtains from (6) that \( d = 0.707, 0.866 \) and 1.001, respectively. Then, note that \( m = 30 \) obtained from (7) is same as the one given by Aoshima and Takada (2006).

In Tables 2 and 3, we examine the two-stage procedure (20)–(21) with (22) in the first block, the two-stage procedure (7)–(9) with (11) having axis number \( l = 2 \) in the second block, and the two-stage procedure due to Aoshima and Takada (2006) in the third block. In Table 2, we consider two cases that (i) \( k = 2, (\rho_1, \rho_2) = (-0.2, 0.1) \) and \( d = 0.458 \); (ii) \( k = 4, (\rho_1, \rho_2, \rho_3, \rho_4) = (-0.2, -0.2, 0.1, 0.1) \) and \( d = 0.648 \). In Table 3, we consider two cases that (i) \( k = 2, (\rho_1, \rho_2) = (-0.1, 0.2) \) and \( d = 0.567 \); (ii) \( k = 4, (\rho_1, \rho_2, \rho_3, \rho_4) = (-0.1, -0.1, 0.2, 0.2) \) and \( d = 0.801 \). We calculated the value of \( m \) from (20) in the first block, from (7) with \( l = 2 \) in the second block, and from (2.6) given in Aoshima and Takada (2006) in the third block, respectively, in each table.

In Tables 1–3, the findings obtained by averaging the outcomes from 10,000 (= \( R \), say) replications are summarized in each situation. Under a fixed scenario, suppose that the \( r \)th replication ends with \( N_i = n_{ir} \) \( (i = 1, \ldots, k) \) observations and the corresponding fixed-span confidence region \( R_n \), based on \( n_r = (n_{1r}, \ldots, n_{kr}) \) for \( r = 1, \ldots, R \). Now, \( \pi_i = R^{-1}\sum_{r=1}^R n_{ir} \) which estimates \( C_i \).
with its estimated standard error \( s(\bar{\pi}_i) \), where \( s^2(\bar{\pi}_i) = (R^2 - R)^{-1} \sum_{r=1}^{R} (n_{ir} - \bar{n})^2 \), \( i = 1, ..., k \). Then, \( \bar{\pi} (= \bar{\pi}_1 + \cdots + \bar{\pi}_k) \) estimates the total fixed sample size \( C \) with its estimated standard error \( s(\bar{\pi}) \), computed analogously. In the end of the \( r \)th replication, we also check whether \( \mu \) belongs to the constructed confidence region \( R_{n_r} \) and define \( p_r = 1 \) (or 0) accordingly as \( \mu \) does (or does not) belong to \( R_{n_r} \), \( r = 1, ..., R \). Let \( \bar{p} = R^{-1} \sum_{r=1}^{R} p_r \), which estimates the target coverage probability, having its estimated standard error \( s(\bar{p}) \) where \( s^2(\bar{p}) = R^{-1}(1 - \bar{p}) \). For the two–stage procedure (7)–(9) with (11) (or (20)–(21) with (22) or Remark 2 with (24)), the value of \( u \) is given as the average number of the outcomes from 10,000 replications. For the two–stage procedure due to Aoshima and Takada (2006), that is given by (25). At the last column, we gave the approximate value of \( E_\theta(N_i - C_i) \), which was obtained from Theorem 3.2 (or Theorem 3.4) in Section 3 and from Theorem 2.2 in Aoshima and Takada (2006).

Let us explain, for example, the entries from the first block for the case when \( k = 2 \) in Table 1. We consider \( C = 90 \), and hence \( d = 0.707 \), \( C_1 = 42.67 \), \( C_2 = 47.33 \) from (6) with \( l = 1 \). One obtains \( m = 30 \) using (7) (having \( m_0 = 4 \), for example). From 10,000 independent simulations, we observed \( u = 8.211 \), \( \pi_1 = 45.18 \), \( s(\pi_1) = 0.091 \), \( \pi_2 = 50.03 \), \( s(\pi_2) = 0.104 \), and \( \bar{n} = 95.21 \), \( s(\bar{n}) = 0.175 \). Also, we had \( \bar{p} = 0.9712 \), \( s(\bar{p}) = 0.00167 \), and \( \pi_1 - C_1 = 2.51 \), \( \pi_2 - C_2 = 2.70 \), \( \bar{n} - C = 5.21 \). At the last column, we had \( E(N_1 - C_1) = 2.43 \), \( E(N_2 - C_2) = 2.68 \), \( E(N - C) = 5.12 \) where \( N = \sum_{i=1}^{k} N_i \). Theorem 3.2 with \( j_l = l = 1 \) indicates that one may expect \( \pi_i - C_i \) to fall in the vicinity of the value of \( E(N_i - C_i) \), \( i = 1, 2 \). One will observe that the values of \( E_\theta(N_i - C_i) \) are approximated fairly well by these asymptotic values for small \( d \).

Throughout, the two–stage procedure (7)–(9) with (11) or (20)–(21) with (22) or Remark 2 with (24) reduces the sample size required in the two–stage procedure due to Aoshima and Takada (2006). It is obvious specially when the number of populations, \( k \), becomes large. In Table 2, one will observe that the two–stage procedure (7)–(9) with (11) having axis number \( l = 2 \) reduces the sample size required in the modified procedure (20)–(21) with (22). It is because the former can start with preferable \( m \) for the specified axis number, \( l = 2 \), which is the same number as \( l_o \) to minimize the sum \( \sum_{i=1}^{k} C_i \) for the parameter configuration in Table 2. In contrast, in Table 3, one will observe the converse. It is because the axis number \( l = 2 \) does not coincide with \( l_o \) for the parameter configuration in Table 3.

Acknowledgements

The authors would like to thank an anonymous referee and the Associate Editor for their valuable comments.
A Appendix

Throughout this section, for fixed $l$ ($1 \leq l \leq \ell$), we write that

$$\tau_{ijl} = |b_i| \sigma_{il} \sum_{i'=1}^k |b_{i'}| \frac{\sigma^2_{i'i}}{\sigma_{il}}$$

$$Y_{ijl} = |b_i| S_{il} \sum_{i'=1}^k |b_{i'}| \frac{\sigma^2_{i'i}}{\sigma_{il}}$$

$$\tilde{Y}_d = \max(Y_{i1}, ..., Y_{d1}), \quad Y_l = \max(Y_{1l}, ..., Y_{ll})$$

for $i = 1, ..., k$; $j = 1, ..., \ell$. Then, note that $\tilde{Y}_d = |b_i| S_{il} \tilde{Y}_l$. From (6), we write that $C_i = a \tau_{ijl}/d^2$. Let $d (> 0)$ go to zero through a sequence such that $a \tau_{ij}/d^2$ always remains an integer. Then, from (7), we may write that $m = a \tau_{ij}/d^2$.

Let $\nu_j = \nu_{rj}$, $j = 1, ..., \ell$. For $S^2_{ij}$, $j = 1, ..., \ell$ ($i = 1, ..., k$), defined by (8), we note that $
u_j S^2_{ij}/\sigma^2_{ij}$, $j = 1, ..., \ell$, are independently distributed as the chi–square distribution with $\nu_j$ d.f., $j = 1, ..., \ell$, respectively. For each $i$, let $W_{ij}$, $j = 1, ..., \ell$, denote the random variables such that $\nu_j W_{ij}$. $j = 1, ..., \ell$, are independently distributed as the chi–square distribution with $\nu_j$ d.f., $j = 1, ..., \ell$, respectively. Let $\delta_{ij} = W_{ij} - 1$. Then, we have that $S^2_{ij} = \sigma^2_{ij}(1 + \delta_{ij})$, and $E(\delta_{ij}) = 0$, $E(\delta^2_{ij}) = 2\nu_j^{-1}$, $E(\delta_{ij}^{2t-1}) = O(\nu^{-t})$, $E(\delta_{ij}^{2t}) = O(\nu^{-t})$ and $E((1 + \delta_{ij})^t) = O(1)$, $t = 1, 2, ...$

**Lemma 1** For fixed $l$ ($1 \leq l \leq \ell$), if there exists some $j$ such that $\tau_{ijl} > \tau_{ijl}'$ for all $j'$ ($1 \leq j' \neq j \leq \ell$), we have as $\nu \to \infty$ that

$$E_{\theta}(\tilde{Y}_d) = E_{\theta}(Y_{ijd}) + O(\nu^{-3/2}) \quad \text{and} \quad E_{\theta}(|Y_{ijl} - \tau_{ijl}|^t) = O(\nu^{-t/2}) \quad (t \geq 2).$$

**Proof.** We write that

$$\frac{S^2_{ij}}{S_{il}} - \frac{\sigma^2_{ij}}{\sigma_{il}} = \frac{\sigma^2_{il}}{\sigma_{il}} \left\{ (1 + \delta_{ij}) \left( \frac{1}{\sqrt{1 + \delta_{il}}} - 1 \right) + \delta_{ij} \right\}$$

and

$$\frac{S_{il} S^2_{ij}}{S_{il} S_{il}} - \frac{\sigma_{il} \sigma^2_{ij}}{\sigma_{il} \sigma_{il}} = \frac{\sigma_{il} \sigma^2_{il}}{\sigma_{il} \sigma_{il}} \left\{ (1 + \delta_{ij}) \left( \frac{1}{\sqrt{1 + \delta_{il}}} - 1 \right) \left( \frac{1}{\sqrt{1 + \delta_{il}}} - 1 \right) + \delta_{ij} \right\}.$$  

By noting that $E_{\theta}(|(1 + \delta_{il})^{-1/2} - 1|^t) = O(\nu^{-t/2})$ and $E_{\theta}(|(1 + \delta_{il})^{-1/2} - 1|^t) = O(\nu^{-t/2})$ ($t \geq 2$), we have for $t \geq 2$ that

$$E_{\theta}\left( \frac{S^2_{ij}}{S_{il}} - \frac{\sigma^2_{ij}}{\sigma_{il}} \right)^t = O(\nu^{-t/2}) \quad \text{and} \quad E_{\theta}\left( \left| \frac{S_{il} S^2_{ij}}{S_{il} S_{il}} - \frac{\sigma_{il} \sigma^2_{ij}}{\sigma_{il} \sigma_{il}} \right|^t \right) = O(\nu^{-t/2})$$
and hence that

\[
E_\theta(|Y_{jl} - \tau_{jl}|^t) = E_\theta \left( \sum_{i=1}^{k} |b_i| \left( \frac{S_{ij}^2}{S_{il}} - \frac{\sigma_i^2}{\sigma_l^2} \right) \right)^t = O(\nu^{-t/2}), \tag{A.1}
\]

\[
E_\theta(|Y_{ijl} - \tau_{ijl}|^t) = E_\theta \left( \sum_{i'=1}^k |b_{i'}||b_{i'}| \left( \frac{S_{il}^2S_{ij}^2}{S_{i'l'}} - \frac{\sigma_{il}^2\sigma_{ij}^2}{\sigma_{i'l'}^2} \right) \right)^t = O(\nu^{-t/2}).
\]

Next, there would exist some \( \varepsilon_{A_l} (> 0) \) such that \( \tau_{jl} > \tau_{j'tl} + \varepsilon_{A_l} \) for all \( j' \neq jl \). Define \( A_{jl} = \{ |Y_{jl} - \tau_{jl}| < \varepsilon_{A_l}/2 \}, j = 1, \ldots, \ell, \) and \( A_l = \cap_{j=1}^\ell A_{jl} \). Let \( \tilde{Y}_l = (Y_{1l}, \ldots, Y_{d_l}) \). Note that

\[
\tilde{Y}_l = Y_{j'l} \quad \text{and} \quad \tilde{Y}_d = Y_{ijl} \quad \text{when} \quad Y_l \in A_l. \tag{A.2}
\]

Otherwise, from (A.1), we evaluate that

\[
P_\theta(A_{jl}^c) \leq \sum_{j=1}^\ell P_\theta(A_{jl}^c) = \sum_{j=1}^\ell P_\theta(|Y_{jl} - \tau_{jl}| \geq \varepsilon_{A_l}/2)
\]

\[
\leq (2/\varepsilon_{A_l})^6 \sum_{j=1}^\ell E_\theta(|Y_{jl} - \tau_{jl}|^6) = O(\nu^{-3}). \tag{A.3}
\]

Let \( I_{A_l} \) be the indicator function. Since \( |E_\theta(Y_{ijl}I_{A_l})| \leq \sqrt{E_\theta(Y_{ijl}^2)P_\theta(A_{jl})} = O(\nu^{-3/2}) \), we have that

\[
E_\theta(\tilde{Y}_d I_{A_l}) = O(\nu^{-3/2}). \tag{A.4}
\]

From (A.2) and (A.4), we obtain that

\[
E_\theta(\tilde{Y}_d) = E_\theta(\tilde{Y}_d I_{A_l}) + E_\theta(\tilde{Y}_d I_{A_l}^c) = E_\theta(Y_{ijl}) + O(\nu^{-3/2}).
\]

The proof is completed. \qed

**Lemma 2.** For the two–stage procedure (7)–(9) with (11), we have as \( d \to 0 \) that

\[
E_\theta \left( N_i - \left[ \frac{u_d \tilde{Y}_d}{d^2} \right] - 1 \right) = O(d).
\]

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

\[
E_\theta \left( N_i - \left[ \frac{u_d \tilde{Y}_d}{d^2} \right] - 1 \right) = E_\theta \left( I_{\{N_i = m\}} \left( m - \left[ \frac{u_d \tilde{Y}_d}{d^2} \right] - 1 \right) \right)
\]

\[
\leq \sqrt{P_\theta(N_i = m)E_\theta \left( \left( m - \left[ \frac{u_d \tilde{Y}_d}{d^2} \right] - 1 \right)^2 \right)}. \tag{A.5}
\]

15
From (A.2), we note that $j_{il} = j_l$ when $(Y_{il}, \ldots, Y_{il}) \in A_l$. From assumptions (A1)–(A2), we have that $\tau_{ijl} > \tau_i$. Then, it follows that

$$P_\theta(N_i = m) = P_\theta \left( \frac{u_{ijl}}{\bar{d}^2} \leq m \right) = P_\theta \left( \frac{u_{ijl}}{d^2C_i} - \frac{C_i + 1}{C_i} \leq \frac{m - (C_i + 1)}{C_i} \right)$$

$$\leq P_\theta \left( \frac{u_{ijl}}{\alpha \tau_{ijl}} - 1 - \frac{1}{C_i} \leq \frac{\tau^* - \tau_{ijl}}{\tau_{ijl}} \right)$$

$$\leq P_\theta \left( \frac{u_{ijl}}{\alpha \tau_{ijl}} - 1 - \frac{1}{C_i} \leq \frac{\tau^* - \tau_{ijl}}{\tau_{ijl}} \right) \cap A_l \right) + P_\theta(A^c_l)$$

$$\leq P_\theta \left( \left| \frac{u_{ijl}}{\alpha \tau_{ijl}} - 1 \right| + C_i^{-1} \geq \frac{\tau^* - \tau_{ijl}}{\tau_{ijl}} \right) + P_\theta(A^c_l) \quad (A.6)$$

$$\leq \left( \frac{\tau^* - \tau_{ijl}}{\tau_{ijl}} \right)^{-6} E_\theta \left\{ \left( \left| \frac{u_{ijl}}{\alpha \tau_{ijl}} - 1 \right| + C_i^{-1} \right)^6 \right\} + P_\theta(A^c_l), \quad (A.7)$$

where $u_{ijl} (= a(1 + \nu^{-1}\hat{s}_{ijl})$, say) in (A.6) denotes the value of $u_l$ when $(Y_{3l}, \ldots, Y_{dl}) \in A_l$. Now, one yields that

$$E_\theta \left\{ \left| \frac{u_{ijl}}{\alpha \tau_{ijl}} - 1 \right| \right\} \leq E_\theta \left\{ \left( \frac{1}{\tau_{ijl}} \left| Y_{ijl} - \tau_{ijl} \right| + \left| \hat{s}_{ijl} Y_{ijl} \right| \right) \right\}$$

$$= O(\nu^{-t/2}) \quad (t \geq 2). \quad (A.8)$$

Here, (A.8) follows from the result that for any $x \geq 0$ and $y \geq 0$ such that $x + y = t \geq 2$, we have from Lemma 1 that

$$E_\theta \left( \left| Y_{ijl} - \tau_{ijl} \right|^x \left| \nu^{-1}\hat{s}_{ijl} Y_{ijl} \right|^y \right) \leq \sqrt{E_\theta \left( \left| Y_{ijl} - \tau_{ijl} \right|^{2x} \right) E_\theta \left( \left| \nu^{-1}\hat{s}_{ijl} Y_{ijl} \right|^{2y} \right)$$

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

By combining (A.8) and (A.3) with (A.3), we evaluate that

$$P_\theta(N_i = m) = O(d^6). \quad (A.9)$$

We obtain the result in view of (A.5) and (A.9). \hfill \Box

The following result was given by Aoshima and Yata (2007).

**Lemma 3.** Let $q \geq 0$ and $x \geq 0$ be constants. For a fixed $b \geq 1$, let $X_{bw}$ denote a chi-square random variable with $bv$ d.f. Let $U = qX_{bw} - h - [qX_{bw} - h]$, where $[x]$ denotes the largest integer less than $x$. Then, we have that

$$x - \frac{x(1 + h + x)}{q} \sup_z F'_{bw}(z) \leq P(U \leq x) \leq x + \frac{q}{x} \sup_z F'_{bw}(z),$$

$$16$$
where \( x \in (0, 1) \) and \( x_i \in (0, x) \), and \( F_{b'}(\cdot) \) denotes the first derivative of \( F_{b}(\cdot) \).

**Lemma 4.** For the two-stage procedure (7)–(9) with (11), we have as \( d \to 0 \) that

\[
E_\theta \left\{ \frac{u_l}{d^2} | b_i | S_d \sum_{l'=1}^{k} | b_{l'} | \frac{S_{l'} \tilde{m}}{S_{l' l}} - \left[ \frac{u_l}{d^2} | b_i | S_d \sum_{l'=1}^{k} | b_{l'} | \frac{S_{l'} \tilde{m}}{S_{l' l}} \right] \right\} = \frac{1}{2} + O(d).
\]

**Proof.** Let \( U_{(ijl)} = (u_{(ij)}/d^2)|b_i|S_d \sum_{l'=1}^{k} | b_{l'} | S_{l' j} / S_{l' l} \) with \( u_{(ij)} = a(1 + \nu^{-1} \tilde{s}_{l(i)}) \). Then, from (A.2)–(A.3), we have as \( d \to 0 \) that

\[
E_\theta \{ U_{(ijm)_l} - [U_{(ijm)_l}] \} = E_\theta \{ (U_{(ijm)_l} - [U_{(ijm)_l}])I_{A_l} \} + E_\theta \{ (U_{(ijm)_l} - [U_{(ijm)_l}])I_{A_l} \} = E_\theta \{ U_{(ijl)} - [U_{(ijl)}] \} + O(d). \tag{A.10}
\]

Let \( W_{ij}, \ j = 1, \ldots, \ell \) \((i = 1, \ldots, k)\), be independently distributed as the chi–square distribution with \( \nu_j \) d.f., \( j = 1, \ldots, \ell \), respectively. Let \( X_j = \sum_{l=1}^{k} W_{ij} \) and \( V_{ij} = W_{ij}/X_j, \ i = 1, \ldots, k; \ j = 1, \ldots, \ell \). Then, \( X_j \) is distributed as the chi–square distribution with \( k \nu_j \) d.f., \( V_{ij} \) is distributed as the beta distribution with parameters \( \nu_j / 2 \) and \((k-1)\nu_j / 2\), and \( X_j \) and \( V_i = (V_{i\ell_1}, \ldots, V_{i\ell_l}) \) are independent.

When \( ji = l \), we write \( \tilde{s}_{l(i)} \) as

\[
\tilde{s}_{l(i)} = \frac{1}{\nu_l} + \frac{(a-p)b_i^2 \sigma_i^2 V_{il} \sum_{l'=1}^{k} | b_{l'} | \sigma_{l'l} \sqrt{V_{l'l}}}{2\nu_l Z_i^2} - \frac{b_i^2 \sigma_i^2 V_{il} k \tau_{l*}}{2X_i Z_i^2},
\]

where \( Z_i = |b_i| \sigma_i \sqrt{V_{il}} \sum_{l'=1}^{k} | b_{l'} | \sqrt{V_{l'l}} \). Then, we have that \( (u_{(ij)/d^2})|b_i|S_d \sum_{l'=1}^{k} | b_{l'} | S_{l' i} = (u_{(ij)/d^2})X_i Z_i \) where

\[
\frac{u_{(ij)/d^2}X_i Z_i}{d^2 \nu_l} = X_i \frac{aZ_i}{d^2 \nu_l} \left( 1 + \frac{1}{\nu_l} \left( 1 + \frac{(a-p)b_i^2 \sigma_i^2 V_{il} \sum_{l'=1}^{k} | b_{l'} | \sigma_{l' l} \sqrt{V_{l' l}}}{2Z_i^2} \right) \right)
\]

\[
- \frac{ab_i^2 \sigma_i^2 V_{il} k \tau_{l*}}{2d^2 Z_i \nu_l}.
\]

Let us denote

\[
Q = \frac{aZ_i}{d^2 \nu_l} \left( 1 + \frac{1}{\nu_l} \left( 1 + \frac{(a-p)b_i^2 \sigma_i^2 V_{il} \sum_{l'=1}^{k} | b_{l'} | \sigma_{l' l} \sqrt{V_{l' l}}}{2Z_i^2} \right) \right),
\]

\[
H = \frac{ab_i^2 \sigma_i^2 V_{il} k \tau_{l*}}{2d^2 Z_i \nu_l}.
\]

Then, we have that \( (u_{(ij)/d^2})X_i Z_i = QX_l - H \). Let \( U = (QX_l - H) - [QX_l - H] \). From Lemma 3, the conditional distribution of \( U \), given \( \tilde{V}_i = \ldots \),
\( \tilde{v}_i \) \((H = h, \ Q = q)\), is given for \( x \in (0, 1) \) that

\[
x - x(1 + h + x) \sup_x F'_{\kappa v}(z) \leq P_\theta(U \leq x | \tilde{V} = \tilde{v}) \leq x + x \sup_z F'_{\kappa v}(z),
\]

where \( x_i \in (0, x) \). If \( a \geq p - 2 \), we evaluate that \( 1/Q \leq r_i \tau_{l*}/Z_i \leq r_i \tau_{l*}/(b_i^i \sigma_i^2 V_{il}) \) and \( H/Q \leq r_i k \tau_{l*}/(2 \sum_{i=1}^k b_i^i \sigma_i^2 V_{il}) \leq r_i k \tau_{l*}/(2 \min_{1 \leq i \leq k} b_i^i \sigma_i^2) \) \((= \gamma)\). Then, we have \( E_\theta(1/Q) \leq (r_i \tau_{l*}/b_i^i \sigma_i^2)(k \nu_{l*} - 2)/(\nu_{l*} - 2) \). Here, \( H/Q \) is uniformly integrable since \( |H/Q| \leq \gamma \), and \( 1/Q \) is uniformly integrable since \( 1/Q \leq r_i \tau_{l*}/(b_i^i \sigma_i^2 V_{il}) \) with \( r_i \tau_{l*}/(b_i^i \sigma_i^2 V_{il}) \) being uniformly integrable. From the result that \( \sup_z F'_{\kappa v}(z) = O(\nu^{-1/2}) \) as \( \nu \to \infty \), one can yield that

\[
E_\theta \left( \frac{x}{Q} \sup_z F'_{\kappa v}(z) \right) = E_\theta \left( \frac{x(1 + H + x)}{Q} \sup_z F'_{\kappa v}(z) \right) = O(d).
\]

From the fact that \( E_\theta \{P_\theta(U \leq x | \tilde{V} = \tilde{v})\} = P_\theta(U \leq x) \), we obtain that

\[
P_\theta(U \leq x) = x + O(d) \quad \text{as} \quad d \to 0. \quad (A.11)
\]

Next, when \( j_l \neq l \), let us denote

\[
Q = \frac{aZ_i}{d^2 \nu_{j_l}} \left( 1 + \frac{(a + 2 - p)b_i^i \sigma_i^2 V_{il}}{2 \nu_{j_l} Z_i} \sum_{i=1}^k b_i^i \sigma_i^2 \nu_{j_l} V_{i^2} \right),
\]

\[
H = \frac{a\tau_{l*} b_i^i \sigma_i^2 V_{il}}{2 \nu_{j_l} Z_i} \sum_{i=1}^k \sigma_i^2 \nu_{j_l} V_{i^2},
\]

where \( Z_i = |b_i| \sigma_i \sqrt{V_{il} \sum_{i=1}^k |b_{i^2}| (\sigma_i^2 \nu_{j_l} V_{i^2} / \sigma_i \sqrt{V_{i^2}})} \). Then, \( (u_{i^2}/d^2)|b_i| S_{i^2} \sum_{i=1}^k |b_{i^2}| S_{i^2} / S_{j_l} = QX_{j_l} - H \). Since \( a \geq p - 2 \), we evaluate that

\[
E_\theta \left( \frac{H}{Q} \right) \leq \frac{r_{j_l} \tau_{l*} \max_{1 \leq i \leq k} \sigma_i^2 / \sigma_i^2}{2 \min_{1 \leq i \leq k} b_i^i \sigma_i^2 / \sigma_i^2} E_\theta \left( \sum_{i=1}^k \frac{1}{V_{i^2}} \right) = \frac{kr_{j_l} \tau_{l*} \max_{1 \leq i \leq k} \sigma_i^2 / \sigma_i^2}{2 \min_{1 \leq i \leq k} b_i^i \sigma_i^2 / \sigma_i^2} \left( \frac{k \nu_{l*} - 2}{\nu_{l*} - 2} \right),
\]

\[
E_\theta \left( \frac{1}{Q} \right) \leq \frac{r_{j_l} \tau_{l*} b_i^i \sigma_i^2 / \sigma_i^2}{b_i^i \sigma_i^2} \left( \frac{k \nu_{l*} - 2}{\nu_{l*} - 2} \right).
\]

Then, we claim (A.11) similarly to the case when \( j_l = l \). Hence, \( U \) is asymptotically uniform on \((0, 1)\) as \( d \to 0 \). In view of (A.10), the proof is completed.

\( \square \)

**Remark 5.** When the design value 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 \( k \geq 2 \). When \( u_l \) is not a constant as in this paper, it is much more complicated.
Lemma 5. When \( j_l = 1 \), the two-stage procedure (7)–(9) with (11) has as \( d \to 0 \):

(i) \( E_{\theta}(C_i^{-1}(N_i - C_i)) = (2\nu_1)^{-1}(2s_ir_l - 1 + f_i + B_i) + O(d^3) \),

(ii) \( E_{\theta}(C_i^{-2}(N_i - C_i)^2) = (2\nu_1)^{-1}(1 + 2f_i + \sum_{i' = 1}^k f_{i'}^2) + O(d^3) \),

(iii) \( E_{\theta}(C_i^{-1}(N_i - C_i)C_j^{-1}(N_j - C_j)) = (2\nu_1)^{-1}(f_i + f_j + \sum_{i' = 1}^k f_{i'}^2) + O(d^3) \) \((i \neq j)\);

When \( j_l \neq 1 \), it has as \( d \to 0 \):

(i) \( E_{\theta}(C_i^{-1}(N_i - C_i)) = (2\nu_1)^{-1}(2s_ir_l + 1 - f_i + B_i) + O(d^3) \),

(ii) \( E_{\theta}(C_i^{-2}(N_i - C_i)^2) = (2\nu_1)^{-1}(1 - 2f_i + (4r_{j_l}^{-1}r_l + 1) \sum_{i' = 1}^k f_{i'}^2) + O(d^3) \),

(iii) \( E_{\theta}(C_i^{-1}(N_i - C_i)C_j^{-1}(N_j - C_j)) = (2\nu_1)^{-1}((4r_{j_l}^{-1}r_l + 1) \sum_{i' = 1}^k f_{i'}^2 - f_i - f_j) + O(d^3) \) \((i \neq j)\),

where \( B_i = \nu_1/C_i \) and \( s_l \) is given by (17).

Proof. Let us write that
\[
N_i = rC_iT_i + (1 + [rC_iT_i] - rC_iT_i) + (N_i - [rC_iT_i] - 1),
\]
where \( r = \nu_1/a = 1 + \nu^{-1}s_l \) and \( T_i = \tau_{ijl}^{-1}Y_i \). 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
\[
C_i^{-1}(N_i - C_i) = (rT_i - 1) + \nu^{-1}B_iU_i + C_i^{-1}D_i. \tag{A.12}
\]

We have from Lemma 1 that \( E_{\theta}(s_l) = s_l + o(1) \). When \( j_l = l \), we obtain the following results:

\[
E_{\theta}(rT_i - 1) = (2\nu_1)^{-1}(2s_ir_l - 1 + f_i) + O(d^3),
\]
\[
E_{\theta}((rT_i - 1)^2) = (2\nu_1)^{-1}(1 + 2f_i + \sum_{i' = 1}^k f_{i'}^2) + O(d^3), \tag{A.13}
\]
\[
E_{\theta}((rT_i - 1)(rT_j - 1)) = (2\nu_1)^{-1}(f_i + f_j + \sum_{i' = 1}^k f_{i'}^2) + O(d^3) \ (i \neq j).
\]

When \( j_l \neq l \), we obtain the following results:

\[
E_{\theta}(rT_i - 1) = (2\nu_1)^{-1}(2s_ir_l + 1 - f_i) + O(d^3),
\]
\[
E_{\theta}((rT_i - 1)^2) = (2\nu_1)^{-1}(1 - 2f_i + (4r_{j_l}^{-1}r_l + 1) \sum_{i' = 1}^k f_{i'}^2) + O(d^3), \tag{A.14}
\]
\[
E_{\theta}((rT_i - 1)(rT_j - 1)) = (2\nu_1)^{-1}(-f_i - f_j + (4r_{j_l}^{-1}r_l + 1) \sum_{i' = 1}^k f_{i'}^2) + O(d^3) \ (i \neq j).
\]
Let us combine these results with the expectations of (A.12). Let \( U_i = U_0 + \varepsilon_i \), where \( U_0 \) is a \( U(0,1) \) random variable and \( \varepsilon_i \) is the remainder term. Then, note that \( E\{(rT_i - 1)^2\} \leq E\{(rT_i - 1)^2\}E(\nu^{-2}\varepsilon_j^2) = o(\nu^{-3}) \) so that \( E\{(rT_i - 1)^2\} = o(\nu^{-3/2}) \). The results are obtained straightforwardly.

**Lemma 6.** For the two–stage procedure (7)–(9) with (11), let \( j_i \) be the index such that \( \max_{1 \leq j \leq t} \left( \frac{b_i^2 \sigma_{ij}}{N_i} \right) = \frac{b_i^2 \sigma_{ij_i}}{N_i} \). Then, we have as \( d \to 0 \) that

\[
P_{\theta}(j_i \neq j) = O(d^6).
\]

**Proof.** There would exist some \( \varepsilon_{j_i} > 0 \) such that \( \sum_{i=1}^{k} \frac{b_i^2 \sigma_{ij_i}^2}{\sigma_{ij_i}} > \sum_{i=1}^{k} \frac{b_i^2 \sigma_{ij_i}^2}{\tau_{ij_i}} + \varepsilon_{j_i} \) for all \( j' \neq j \). Define \( J_i = \{ | \sum_{i=1}^{k} \frac{b_i^2 \sigma_{ij_i}^2}{N_i} - \sum_{i=1}^{k} \frac{b_i^2 \sigma_{ij_i}^2}{C_i} | < \varepsilon_{j_i}(d^2/2a), \ j = 1, \ldots, t, \) and \( J = \bigcap_{i=1}^{t} J_i \). Note that

\[
\hat{j}_i = j \text{ when } \left( \sum_{i=1}^{k} \frac{b_i^2 \sigma_{ij_i}^2}{N_i}, \ldots, \sum_{i=1}^{k} \frac{b_i^2 \sigma_{ij_i}^2}{N_i} \right) \in J.
\]

Now, from (A.8), we evaluate that

\[
E_{\theta} \left\{ \sum_{i=1}^{k} \left( \frac{b_i^2 \sigma_{ij_i}^2}{u_i Y_{ij_i}} - \frac{b_i^2 \sigma_{ij_i}^2}{a \tau_{ij_i}} \right) \right\} \leq E_{\theta} \left\{ \left( \frac{1}{u_i Y_{ij_i}} \sum_{i=1}^{k} \frac{b_i^2 \sigma_{ij_i}^2}{a \tau_{ij_i}} - 1 \right)^t \right\} = O(\nu^{-t/2}) \quad (t \geq 2). \tag{A.15}
\]

Next, define \( D_i = \{ 1 + |u_i Y_{ij_i}/d^2| > m \}, \ i = 1, \ldots, k, \) and \( D_i = \bigcap_{i=1}^{t} D_i \). Then, from Lemmas 1-2, we evaluate that

\[
P_{\theta}(J_i) \leq \sum_{j=1}^{\ell} P_{\theta}(J_i) = \sum_{j=1}^{\ell} P_{\theta} \left( \left| \sum_{i=1}^{k} \frac{b_i^2 \sigma_{ij_i}^2}{N_i} - \sum_{i=1}^{k} \frac{b_i^2 \sigma_{ij_i}^2}{C_i} \right| > \varepsilon_{j_i}(d^2/2a) \right) \]

\[
\leq \sum_{j=1}^{\ell} P_{\theta} \left( \left| \sum_{i=1}^{k} \frac{b_i^2 \sigma_{ij_i}^2}{N_i} - \sum_{i=1}^{k} \frac{b_i^2 \sigma_{ij_i}^2}{C_i} \right| > \varepsilon_{j_i}(d^2/2a) \right) \cap A_i \cap \cap D_i \right) + P_{\theta}(A_i) + P_{\theta}(D_i) \]

\[
= \sum_{j=1}^{\ell} P_{\theta} \left( \left| \sum_{i=1}^{k} \frac{b_i^2 \sigma_{ij_i}^2}{N_i} - \sum_{i=1}^{k} \frac{b_i^2 \sigma_{ij_i}^2}{C_i} \right| > \varepsilon_{j_i}(d^2/2a) \right) \cap A_i \right) + P_{\theta}(A_i) + P_{\theta}(D_i) \]

\[
= \sum_{j=1}^{\ell} P_{\theta} \left( \sum_{i=1}^{k} \frac{b_i^2 \sigma_{ij_i}^2}{u_i Y_{ij_i} + d^2 U_i} - \frac{b_i^2 \sigma_{ij_i}^2}{a \tau_{ij_i}} \right) > \varepsilon_{j_i}/2a \right) + O(d^6). \tag{A.16}
\]

20
where \( U_i = 1 + [u_{ij}Y_{ijkl}/d^2] - u_{ij}Y_{ijkl}/d^2 \), \( i = 1, \ldots, k \). Here, using the Taylor expansion and (A.15), we evaluate in (A.16) that

\[
P_\theta \left( k \sum_{i=1}^{k} \left| \frac{b_i^2 \sigma_{ij}^2}{u_{ij}(Y_{ijkl}) + d^2 U_i} - \frac{b_i^2 \sigma_{ij}^2}{a r_{ij}} \right| > \varepsilon J_i / 2a \right) \\
\leq P_\theta \left( \sum_{i=1}^{k} \left| \frac{b_i^2 \sigma_{ij}^2}{u_{ij}(Y_{ijkl})} - \frac{b_i^2 \sigma_{ij}^2}{a r_{ij}} \right| + d^2 \sum_{i=1}^{k} U_i \left( \frac{b_i^2 \sigma_{ij}^2}{u_{ij}(Y_{ijkl}) + d^2 U_i} \right)^2 > \varepsilon J_i / 2a \right) \\
\leq P_\theta \left( \sum_{i=1}^{k} \left| \frac{b_i^2 \sigma_{ij}^2}{u_{ij}(Y_{ijkl})} - \frac{b_i^2 \sigma_{ij}^2}{a r_{ij}} \right| + d^2 \sum_{i=1}^{k} \left( \frac{b_i^2 \sigma_{ij}^2}{u_{ij}(Y_{ijkl})} \right)^2 > \varepsilon J_i / 2a \right) \\
\leq \left( \frac{2a}{\varepsilon J_i} \right)^6 E_\theta \left\{ \left( \sum_{i=1}^{k} \left| \frac{b_i^2 \sigma_{ij}^2}{u_{ij}(Y_{ijkl})} - \frac{b_i^2 \sigma_{ij}^2}{a r_{ij}} \right| + d^2 \sum_{i=1}^{k} \left( \frac{b_i^2 \sigma_{ij}^2}{u_{ij}(Y_{ijkl})} \right)^2 \right)^6 \right\} = O(d^6),
\]

(A.17)

where \( U'_i \in (0, U_i), \ i = 1, \ldots, k \). From (A.16) and (A.17), we conclude the result.

The following result can be obtained similarly to Lemma 6 given in Aoshima and Yata (2007).

**Lemma 7.** For the two-stage procedure (7)–(9) with (11), one has from (A.13)–(A.14) as \( d \to 0 \) that \( E_\theta(\mathcal{R}) = o(d^2) \) in (15).

**Lemma 8.** For the two-stage procedure (20)–(21) with (22), let \( \hat{u} \tilde{Y}_i, \ i = 1, \ldots, k \), satisfy that

\[
\sum_{i=1}^{k} \max \left\{ m, \left[ \frac{\hat{u} \tilde{Y}_i}{d^2} \right] + 1 \right\} = \min \left( \sum_{i=1}^{k} \max \left\{ m, \left[ \frac{u_1 \tilde{Y}_i}{d^2} \right] + 1 \right\}, \ldots, \sum_{i=1}^{k} \max \left\{ m, \left[ \frac{u_l \tilde{Y}_i}{d^2} \right] + 1 \right\} \right),
\]

where \( m = \lfloor a d^{-2} \tau \rfloor + 1 \), and \( u_j \) is the design constant with axis number \( j (= 1, \ldots, \ell) \). Then, we have as \( d \to 0 \) that

\[
E_\theta(\hat{Y}_i) = E_\theta(Y_{ijkl}) + O(d^2).
\]

**Proof.** Let \( \hat{u} \tilde{Y}_i, \ i = 1, \ldots, k \), satisfy that

\[
\sum_{i=1}^{k} \left( 1 + \left[ \frac{\hat{u} \tilde{Y}_i}{d^2} \right] \right) = \min \left( \sum_{i=1}^{k} \left( 1 + \left[ \frac{u_{ij} \tilde{Y}_{ijkl}}{d^2} \right] \right), \ldots, \sum_{i=1}^{k} \left( 1 + \left[ \frac{u_{ijkl} \tilde{Y}_{ijl}}{d^2} \right] \right) \right).
\]

From (A.2), we note that \( u_{ij}(Y_{ijkl}) = u_\hat{Y}_i \) when \( (Y_{il}, Y_{ij}) \in A_l, \ i = 1, \ldots, k; \ l = 1, \ldots, \ell \). We note that \( \max(m, \left[ u_d d^{-2} \tilde{Y}_d \right] + 1) = \left[ u_d d^{-2} \tilde{Y}_d \right] + 1 \) when
for all \( l = 1, \ldots, \ell \). Let \( A = \cap_{l=1}^{\ell} A_l \) and \( D = \cap_{l=1}^{\ell} D_l \). Let \( Y_l = (Y_{il}, \ldots, Y_{il}) \), \( l = 1, \ldots, \ell \), and \( \hat{u} \hat{Y} = (\sum_{i=1}^{k} u_{i1} \hat{Y}_{i1}, \ldots, \sum_{i=1}^{k} u_{i\ell} \hat{Y}_{i\ell}) \). Note that 

\[
\hat{Y}_i = \hat{Y}_i \quad \text{when} \quad (Y_1, \ldots, Y_\ell) \in A \quad \text{and} \quad \hat{u} \hat{Y} \in D \quad \text{(A.18)}
\]

Next, let us write \( \sum_{i=1}^{k} \tau_{ij}, o = \min(\sum_{i=1}^{k} \tau_{ij1}, \ldots, \sum_{i=1}^{k} \tau_{ij, l}) \). Recall the definition of \((l_o, j_o)\), which was given right after Theorem 3.2, to claim that 

\[
\sum_{i=1}^{k} \tau_{ij, l} < \sum_{i=1}^{k} \tau_{ij, l'} \quad \text{for all} \quad l' \neq l_o.
\]

Then, there would exist some \( \varepsilon_B > 0 \) such that \( \sum_{i=1}^{k} \tau_{ij, l} > \varepsilon_B \sum_{i=1}^{k} \tau_{ij, l'} \) for all \( l' \neq l_o \). Define \( B_l = \{ \frac{1}{2} \sum_{i=1}^{k} (u_{ij}d - 2Y_{ij}) \} \), \( i = 1, \ldots, \ell \), and \( B = \cap_{l=1}^{\ell} B_l \). Let \( uY = (\sum_{i=1}^{k} u_{ij1}d - 2Y_{ij1}, \ldots, \sum_{i=1}^{k} u_{ij, l}d - 2Y_{ij, l}) \). Note that 

\[
\hat{Y}_i = Y_{ij, l_o} \quad \text{when} \quad uY \in B. \quad \text{(A.19)}
\]

From (A.8), we have that 

\[
E_\theta \left\{ \left| \sum_{i=1}^{k} d^2 \left( 1 + \frac{u_{ij}}{d^2} Y_{ij} \right) - \frac{a}{d^2} \tau_{ij} \right|^t \right\} 
= E_\theta \left\{ \sum_{i=1}^{k} \left( u_{ij}Y_{ij} - a\tau_{ij} \right) + d^2 \sum_{i=1}^{k} \left( 1 + \frac{u_{ij}}{d^2} Y_{ij} \right) - \frac{u_{ij}}{d^2} \tau_{ij} \right|^t \right\} 
\leq E_\theta \left\{ \left( \sum_{i=1}^{k} \left( u_{ij}Y_{ij} - a\tau_{ij} \right) + d^2 \sum_{i=1}^{k} \left( 1 + \frac{u_{ij}}{d^2} Y_{ij} \right) - \frac{u_{ij}}{d^2} \tau_{ij} \right) \right|^t \right\} 
\leq E_\theta \left\{ \left( \sum_{i=1}^{k} a\tau_{ij} + \frac{u_{ij}}{d^2} Y_{ij} - 1 \right) + d^2 k \right\} 
= O(d^t) \quad (t \geq 2).
\]

Then, we evaluate that 

\[
P_\theta(B^c) \leq \sum_{l=1}^{\ell} P_\theta(B_l^c) = \sum_{l=1}^{\ell} P_\theta \left\{ \left| \sum_{i=1}^{k} \left( u_{ij}Y_{ij} - a\tau_{ij} \right) \right| > \varepsilon_B(a/2d^2) \right\} 
\leq \left( \frac{2}{a\varepsilon_B} \right)^6 \sum_{l=1}^{\ell} E_\theta \left\{ \left| \sum_{i=1}^{k} d^2 \left( 1 + \frac{u_{ij}}{d^2} Y_{ij} \right) - \frac{a}{d^2} \tau_{ij} \right|^6 \right\} \leq O(d^6).
\]

Let \( I_B \) be the indicator function. Note that 

\[
|E_\theta(Y_{ij}I_{B^c})| \leq \sqrt{E_\theta(Y_{ij}^2)P_\theta(B^c)} = O(d^3).
\]

22
From (A.3), we evaluate that

\[ P_{\theta}(A^c) \leq \sum_{j=1}^{\ell} \sum_{l=1}^{\ell} P_{\theta}(A_{jl}^c) = O(\nu^{-3}), \]

From (A.7), we evaluate that

\[ P_{\theta}(D^c) \leq \sum_{i=1}^{k} \sum_{l=1}^{\ell} P_{\theta}(D_{iil}^c) = O(d^6). \]

Note that \( |E_{\theta}(\tilde{Y}_{ijl}I_{A^c})| = O(\nu^{-3/2}) \) and \( |E_{\theta}(\tilde{Y}_{ijl}I_{D^c})| = O(\nu^{-3/2}) \). Let \( I_{(A \cap B \cap D)} \) be the indicator function. We have that

\[ E_{\theta}(\tilde{Y}_{i}I_{(A \cap B \cap D)^c}) \leq E_{\theta}(\tilde{Y}_{i}(I_{A^c} + I_{B^c} + I_{D^c})) = O(d^3). \] (A.20)

From (A.18), (A.19) and (A.20), we obtain that

\[ E_{\theta}(\tilde{Y}_{i}) = E_{\theta}(\tilde{Y}_{i}(I_{A^c} + I_{B^c} + I_{D^c})), \]

\[ = E_{\theta}(Y_{ijl}I_{o}) + O(d^3). \]

The proof is completed. \( \square \)

References

Fig. 1. Fixed region with radius $d$ from the centre $\mu$.

Fig. 2. When the sample is taken up to size $n_i = \left(\frac{a}{d^2}\right) |b_i| \sigma_{i1} \sum_{v=1}^{k} |b_{iv}| \sigma_{iv}$ for each $\pi_i$.

Fig. 3. When the sample is additionally taken until the total sample for each $\pi_i$ becomes of size $n_i = \left(\frac{a\delta}{d^2}\right) |b_i| \sigma_{i1} \sum_{v=1}^{k} |b_{iv}| \sigma_{iv}$.

Fig. 4. When the sample is taken up to size $n_i = \left(\frac{a/d^2}{\max_{1\leq j\leq \ell} |b_{ij}| \sigma_{ij} \sum_{v=1}^{k} |b_{ijv}| \sigma_{ijv}}\right)$ for each $\pi_i$.
Table 1. Simulated results

<table>
<thead>
<tr>
<th>(k = 2), ((\rho_1, \rho_2) = (1/10, 1/5)), (\tau_{1*} = 1.920)</th>
<th>(m = 30) ((d = 0.707))</th>
</tr>
</thead>
<tbody>
<tr>
<td>(C)</td>
<td>90</td>
</tr>
<tr>
<td>(C_1)</td>
<td>42.67</td>
</tr>
<tr>
<td>(C_2)</td>
<td>47.33</td>
</tr>
</tbody>
</table>

Remark 2 with (24)

| \(C\) | 90 | 8.261 | 95.75 | 0.174 | 0.9720 | 0.00165 | 5.75 | 5.66 |
| \(C_1\) | 42.67 | 45.44 | 0.091 | 2.77 | 2.69 |
| \(C_2\) | 47.33 | 50.31 | 0.104 | 2.98 | 2.97 |

Remark 2 with (25)

<table>
<thead>
<tr>
<th>(k = 3), ((\rho_1, \rho_2, \rho_3) = (1/10, 3/20, 1/5)), (\tau_{1*} = 2.881)</th>
<th>(m = 30) ((d = 0.866))</th>
</tr>
</thead>
<tbody>
<tr>
<td>(C)</td>
<td>135</td>
</tr>
<tr>
<td>(C_1)</td>
<td>42.65</td>
</tr>
<tr>
<td>(C_2)</td>
<td>45.04</td>
</tr>
<tr>
<td>(C_3)</td>
<td>47.31</td>
</tr>
</tbody>
</table>

Remark 2 with (24)

| \(C\) | 135 | 8.396 | 146.06 | 0.217 | 0.9746 | 0.00157 | 11.06 | 10.38 |
| \(C_1\) | 42.65 | 46.10 | 0.083 | 3.45 | 3.29 |
| \(C_2\) | 45.04 | 48.76 | 0.089 | 3.72 | 3.46 |
| \(C_3\) | 47.31 | 51.20 | 0.095 | 3.88 | 3.62 |

Remark 2 with (25)

<table>
<thead>
<tr>
<th>(k = 4), ((\rho_1, \rho_2, \rho_3, \rho_4) = (1/10, 2/15, 1/6, 1/5)), (\tau_{1*} = 3.841)</th>
<th>(m = 30) ((d = 1.001))</th>
</tr>
</thead>
<tbody>
<tr>
<td>(C)</td>
<td>180</td>
</tr>
<tr>
<td>(C_1)</td>
<td>42.64</td>
</tr>
<tr>
<td>(C_2)</td>
<td>44.25</td>
</tr>
<tr>
<td>(C_3)</td>
<td>45.80</td>
</tr>
<tr>
<td>(C_4)</td>
<td>47.31</td>
</tr>
</tbody>
</table>

Remark 2 with (24)

| \(C\) | 180 | 8.530 | 196.62 | 0.256 | 0.9742 | 0.00159 | 16.62 | 16.60 |
| \(C_1\) | 42.64 | 46.61 | 0.079 | 3.97 | 3.95 |
| \(C_2\) | 44.25 | 48.32 | 0.084 | 4.07 | 4.09 |
| \(C_3\) | 45.80 | 50.00 | 0.086 | 4.19 | 4.22 |
| \(C_4\) | 47.31 | 51.70 | 0.090 | 4.39 | 4.35 |
Table 2. Simulated results

<table>
<thead>
<tr>
<th>$u$</th>
<th>$\pi$</th>
<th>$s(\pi)$</th>
<th>$\bar{p}$</th>
<th>$s(\bar{p})$</th>
<th>$\pi - C$</th>
<th>$E(N - C)$</th>
</tr>
</thead>
</table>

$k = 2$, $d = 0.458$
$$(\rho_1, \rho_2) = (-0.2, 0.1)$, $(\tau_{1*}, \tau_{2*}) = (0.785, 1.357)$

(20)–(21) with (22): $m = 30$

<table>
<thead>
<tr>
<th>$C$</th>
<th>155</th>
<th>7.948</th>
<th>163.22</th>
<th>0.176</th>
<th>0.9153</th>
<th>0.00278</th>
<th>8.22</th>
<th>3.02</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C_1$</td>
<td>83.06</td>
<td>84.34</td>
<td>0.131</td>
<td>1.28</td>
<td>1.60</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$C_2$</td>
<td>71.94</td>
<td>78.87</td>
<td>0.149</td>
<td>6.94</td>
<td>1.42</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(7)–(9) with (11) having $l = 2$: $m = 51$

<table>
<thead>
<tr>
<th>$C$</th>
<th>158.04</th>
<th>8.261</th>
<th>174.23</th>
<th>0.182</th>
<th>0.9368</th>
<th>0.00243</th>
<th>16.20</th>
<th>10.07</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C_1$</td>
<td>83.06</td>
<td>85.92</td>
<td>0.084</td>
<td>2.86</td>
<td>0.91</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$C_2$</td>
<td>71.94</td>
<td>74.38</td>
<td>0.068</td>
<td>2.44</td>
<td>0.84</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Aoshima and Takada (2006): $m = 30$

<table>
<thead>
<tr>
<th>$C$</th>
<th>155</th>
<th>7.867</th>
<th>160.30</th>
<th>0.138</th>
<th>0.9171</th>
<th>0.00276</th>
<th>5.30</th>
<th>1.75</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C_1$</td>
<td>83.06</td>
<td>88.46</td>
<td>0.106</td>
<td>5.39</td>
<td>5.41</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$C_2$</td>
<td>71.94</td>
<td>74.38</td>
<td>0.068</td>
<td>2.44</td>
<td>0.84</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$k = 4$, $d = 0.648$
$$(\rho_1, \rho_2, \rho_3, \rho_4) = (-0.2, -0.2, 0.1, 0.1)$, $(\tau_{1*}, \tau_{2*}) = (1.570, 3.135)$

(20)–(21) with (22): $m = 30$

<table>
<thead>
<tr>
<th>$C$</th>
<th>310</th>
<th>7.873</th>
<th>318.27</th>
<th>0.246</th>
<th>0.9091</th>
<th>0.00287</th>
<th>8.27</th>
<th>3.91</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C_1$</td>
<td>83.06</td>
<td>84.18</td>
<td>0.099</td>
<td>1.12</td>
<td>1.02</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$C_2$</td>
<td>83.06</td>
<td>84.20</td>
<td>0.099</td>
<td>1.13</td>
<td>1.02</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$C_3$</td>
<td>71.94</td>
<td>74.98</td>
<td>0.103</td>
<td>3.04</td>
<td>0.93</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$C_4$</td>
<td>71.94</td>
<td>74.91</td>
<td>0.100</td>
<td>2.98</td>
<td>0.93</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(7)–(9) with (11) having $l = 2$: $m = 59$

<table>
<thead>
<tr>
<th>$C$</th>
<th>310</th>
<th>7.801</th>
<th>313.18</th>
<th>0.169</th>
<th>0.9100</th>
<th>0.00286</th>
<th>3.17</th>
<th>1.96</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C_1$</td>
<td>83.06</td>
<td>83.87</td>
<td>0.061</td>
<td>0.80</td>
<td>0.49</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$C_2$</td>
<td>83.06</td>
<td>83.86</td>
<td>0.061</td>
<td>0.80</td>
<td>0.49</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$C_3$</td>
<td>71.94</td>
<td>72.72</td>
<td>0.051</td>
<td>0.78</td>
<td>0.49</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$C_4$</td>
<td>71.94</td>
<td>72.72</td>
<td>0.051</td>
<td>0.79</td>
<td>0.49</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Aoshima and Takada (2006): $m = 30$

<table>
<thead>
<tr>
<th>$C$</th>
<th>316.08</th>
<th>8.530</th>
<th>355.89</th>
<th>0.266</th>
<th>0.9426</th>
<th>0.00233</th>
<th>39.82</th>
<th>29.92</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C_1$</td>
<td>83.06</td>
<td>90.92</td>
<td>0.093</td>
<td>7.86</td>
<td>8.13</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$C_2$</td>
<td>83.06</td>
<td>90.93</td>
<td>0.093</td>
<td>7.87</td>
<td>8.13</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$C_3$</td>
<td>74.97</td>
<td>87.02</td>
<td>0.112</td>
<td>12.04</td>
<td>6.83</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$C_4$</td>
<td>74.97</td>
<td>87.02</td>
<td>0.111</td>
<td>12.05</td>
<td>6.83</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 3. Simulated results

<table>
<thead>
<tr>
<th></th>
<th>u</th>
<th>π</th>
<th>s(π)</th>
<th>p</th>
<th>s(p)</th>
<th>π − C</th>
<th>E(N − C)</th>
</tr>
</thead>
</table>

\( k = 2, \ d = 0.567 \)
\( (\rho_1, \rho_2) = (-0.1, 0.2), \ (\tau_{1*}, \tau_{2*}) = (1.231, 1.217) \)

(20)–(21) with (22): \( m = 30 \)

<table>
<thead>
<tr>
<th></th>
<th>( C )</th>
<th>( \pi )</th>
<th>( s(\pi) )</th>
<th>( p )</th>
<th>( s(p) )</th>
<th>( \pi − C )</th>
<th>( E(N − C) )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>107</td>
<td>8.084</td>
<td>116.41</td>
<td>0.188</td>
<td>0.9275</td>
<td>0.00259</td>
<td>9.41</td>
</tr>
<tr>
<td></td>
<td>42.60</td>
<td>50.12</td>
<td>0.087</td>
<td>7.52</td>
<td>2.53</td>
<td>3.80</td>
<td></td>
</tr>
<tr>
<td></td>
<td>64.40</td>
<td>66.29</td>
<td>0.167</td>
<td>1.88</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(7)–(9) with (11) having \( l = 2 \): \( m = 30 \)

<table>
<thead>
<tr>
<th></th>
<th>( C )</th>
<th>( \pi )</th>
<th>( s(\pi) )</th>
<th>( p )</th>
<th>( s(p) )</th>
<th>( \pi − C )</th>
<th>( E(N − C) )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>115.64</td>
<td>8.134</td>
<td>124.24</td>
<td>0.241</td>
<td>0.9429</td>
<td>0.00232</td>
<td>8.60</td>
</tr>
<tr>
<td></td>
<td>62.41</td>
<td>67.09</td>
<td>0.140</td>
<td>4.67</td>
<td>3.89</td>
<td>3.25</td>
<td></td>
</tr>
<tr>
<td></td>
<td>53.23</td>
<td>57.15</td>
<td>0.110</td>
<td>3.93</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Aoshima and Takada (2006): \( m = 30 \)

<table>
<thead>
<tr>
<th></th>
<th>( C )</th>
<th>( \pi )</th>
<th>( s(\pi) )</th>
<th>( p )</th>
<th>( s(p) )</th>
<th>( \pi − C )</th>
<th>( E(N − C) )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>113.78</td>
<td>8.261</td>
<td>122.98</td>
<td>0.171</td>
<td>0.9428</td>
<td>0.00232</td>
<td>9.20</td>
</tr>
<tr>
<td></td>
<td>49.38</td>
<td>54.42</td>
<td>0.064</td>
<td>5.05</td>
<td>3.38</td>
<td>3.08</td>
<td></td>
</tr>
<tr>
<td></td>
<td>64.40</td>
<td>68.55</td>
<td>0.143</td>
<td>4.15</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\( k = 4, \ d = 0.801 \)
\( (\rho_1, \rho_2, \rho_3, \rho_4) = (-0.1, -0.1, 0.2, 0.2), \ (\tau_{1*}, \tau_{2*}) = (2.462, 2.853) \)

(20)–(21) with (22): \( m = 30 \)

<table>
<thead>
<tr>
<th></th>
<th>( C )</th>
<th>( \pi )</th>
<th>( s(\pi) )</th>
<th>( p )</th>
<th>( s(p) )</th>
<th>( \pi − C )</th>
<th>( E(N − C) )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>214</td>
<td>7.983</td>
<td>225.60</td>
<td>0.271</td>
<td>0.9233</td>
<td>0.00266</td>
<td>11.59</td>
</tr>
<tr>
<td></td>
<td>42.60</td>
<td>47.16</td>
<td>0.076</td>
<td>4.56</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>42.60</td>
<td>47.12</td>
<td>0.077</td>
<td>4.52</td>
<td>1.66</td>
<td>1.66</td>
<td></td>
</tr>
<tr>
<td></td>
<td>64.40</td>
<td>65.66</td>
<td>0.137</td>
<td>1.26</td>
<td>2.37</td>
<td>2.37</td>
<td></td>
</tr>
<tr>
<td></td>
<td>64.40</td>
<td>65.66</td>
<td>0.137</td>
<td>1.26</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(7)–(9) with (11) having \( l = 2 \): \( m = 35 \)

<table>
<thead>
<tr>
<th></th>
<th>( C )</th>
<th>( \pi )</th>
<th>( s(\pi) )</th>
<th>( p )</th>
<th>( s(p) )</th>
<th>( \pi − C )</th>
<th>( E(N − C) )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>231.28</td>
<td>7.909</td>
<td>239.04</td>
<td>0.169</td>
<td>0.942</td>
<td>0.00234</td>
<td>7.76</td>
</tr>
<tr>
<td></td>
<td>62.41</td>
<td>64.49</td>
<td>0.061</td>
<td>2.08</td>
<td>1.77</td>
<td>1.77</td>
<td></td>
</tr>
<tr>
<td></td>
<td>62.41</td>
<td>64.50</td>
<td>0.061</td>
<td>2.10</td>
<td>1.77</td>
<td>1.77</td>
<td></td>
</tr>
<tr>
<td></td>
<td>53.23</td>
<td>55.01</td>
<td>0.051</td>
<td>1.78</td>
<td>1.53</td>
<td>1.53</td>
<td></td>
</tr>
<tr>
<td></td>
<td>53.23</td>
<td>55.03</td>
<td>0.051</td>
<td>1.80</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Aoshima and Takada (2006): \( m = 30 \)

<table>
<thead>
<tr>
<th></th>
<th>( C )</th>
<th>( \pi )</th>
<th>( s(\pi) )</th>
<th>( p )</th>
<th>( s(p) )</th>
<th>( \pi − C )</th>
<th>( E(N − C) )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>227.56</td>
<td>8.530</td>
<td>251.82</td>
<td>0.250</td>
<td>0.952</td>
<td>0.00214</td>
<td>24.27</td>
</tr>
<tr>
<td></td>
<td>49.38</td>
<td>55.50</td>
<td>0.054</td>
<td>6.13</td>
<td>4.93</td>
<td>4.93</td>
<td></td>
</tr>
<tr>
<td></td>
<td>49.38</td>
<td>55.51</td>
<td>0.054</td>
<td>6.13</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>64.40</td>
<td>70.43</td>
<td>0.125</td>
<td>6.03</td>
<td>5.79</td>
<td>5.79</td>
<td></td>
</tr>
<tr>
<td></td>
<td>64.40</td>
<td>70.38</td>
<td>0.125</td>
<td>5.98</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>