Available online at www.sciencedirect.com
ScienceDirect
Journal of Statistical Planning and Inference 137 (2007) 2240–2249
journal of statistical planning and inference
www.elsevier.com/locate/jspi

Generalized inferences on the common mean vector of several multivariate normal populations

S.H. Lin[a,*], Jack C. Lee[b], R.S. Wang[c]

[a]*National Taichung Institute of Technology, 129 Sanmin Road Sec. 3, Taichung 404, Taiwan*
[b]*Institute of Statistics and Graduate Institute of Finance, National ChiaoTung University, Taiwan*
[c]*Institute of Statistics, National ChiaoTung University, Taiwan*

Received 24 May 2005; accepted 7 July 2006
Available online 2 October 2006

Abstract

The hypothesis testing and confidence region are considered for the common mean vector of several multivariate normal populations when the covariance matrices are unknown and possibly unequal. A generalized confidence region is derived using the concepts of generalized method based on the generalized p-value. The generalized confidence region is illustrated with two numerical examples. The merits of the proposed method are numerically compared with those of existing methods with respect to their expected area or expected d-dimensional volumes and coverage probabilities under different scenarios.

Keywords: Combined estimator; Coverage probability; Expected d-dimensional volumes; Generalized pivotal quantity; Generalized test variable

1. Introduction

Estimating the common mean vector of several multivariate normal populations with unknown and possibly unequal covariance matrices is one of the oldest and interesting problems in statistical literature. This problem arises, for example, when two or more independent measuring instruments or agencies are involved to measure like products, effects, or substances which are produced by the same production process to estimate the average quality in terms of several characteristics. If the samples collected by independent studies are assumed to come from multivariate normal populations with a common mean vector and unknown covariance matrices, then the problem of interest may be to estimate or construct a confidence region for the common mean vector $\boldsymbol{\mu}$ of these populations. If the unknown covariance matrices are assumed to be identical, then there are optimal methods available to make inferences on $\boldsymbol{\mu}$. However, when the covariance matrices are unknown and unequal, it is clear that the distribution of any combined estimators of $\boldsymbol{\mu}$ will involve nuisance parameters, and then the standard method has serious limitations for the purpose of finding an exact test and confidence region of $\boldsymbol{\mu}$. Therefore, constructing a generalized confidence region of $\boldsymbol{\mu}$ for models involving variance components deserves further attention.

* Corresponding author. Fax: +886 4 22196696.
E-mail address: suelin@ntit.edu.tw (S.H. Lin).

0378-3758/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.jspi.2006.07.005

Suppose there are $I\,(I \geqslant 2)$ d-variate normal populations with common mean vector $\boldsymbol{\mu}$ and unknown covariance matrices $\boldsymbol{\Sigma}_1, \ldots, \boldsymbol{\Sigma}_I$. Let $\mathbf{X}_{i1}, \ldots, \mathbf{X}_{in_i}$ be independent d-variate vector observations from the ith population, $i = 1, \ldots, I$, and $\mathbf{X}_{ij} \sim N_d(\boldsymbol{\mu}, \ \boldsymbol{\Sigma}_i)$, $j = 1, \ldots, n_i$. For the ith population, let

$$\bar{\mathbf{X}}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} \mathbf{X}_{ij} \quad \text{and} \quad \mathbf{S}_i = \frac{1}{n_i - 1} \sum_{j=1}^{n_i} (\mathbf{X}_{ij} - \bar{\mathbf{X}}_i)(\mathbf{X}_{ij} - \bar{\mathbf{X}}_i)' \tag{1.1}$$

be the sample mean vector and sample covariance matrix. We are interested in estimating the common mean vector $\boldsymbol{\mu}$, based on the minimal sufficient statistics $(\bar{\mathbf{X}}_1, \ldots, \bar{\mathbf{X}}_I, \mathbf{S}_1, \ldots, \mathbf{S}_I)$.

In the univariate case, the common mean problem has received considerable attention in the statistical literature; we refer the reader to Meier (1953), Maric and Graybill (1979), Pagurova and Gurskii (1979), Sinha (1985), Eberhardt et al. (1989), Fairweather (1972), Jordan and Krishnamoorthy (1996), Krishnamoorthy and Lu (2003), Lin and Lee (2005) and the references therein.

In the multivariate case, Chiou and Cohen (1985) showed that $\hat{\boldsymbol{\mu}}_{\mathrm{GD}d}$,

$$\hat{\boldsymbol{\mu}}_{\mathrm{GD}d} = \left(\sum_{i=1}^{I} n_i \mathbf{S}_i^{-1} \right)^{-1} \sum_{j=1}^{I} n_j \mathbf{S}_j^{-1} \bar{\mathbf{x}}_j, \tag{1.2}$$

dominates neither $\bar{\mathbf{X}}_1$ nor $\bar{\mathbf{X}}_2$, when $I = 2$ and $d \geqslant 2$, with respect to the covariance criterion, although Graybill and Deal (1959) got the opposite result in univariate two-sample case. Loh (1991) estimated the common mean vector from a symmetric loss function point of view as alternatives to $\hat{\boldsymbol{\mu}}_{\mathrm{GD}d}$. Zhou and Mathew (1994) proposed several combined tests for testing the common mean vector, but the problem of multiple comparisons had not been discussed when the null hypothesis was rejected. Jordan and Krishnamoorthy (1995) provided a confidence region of $\boldsymbol{\mu}$ centered at a weighted Graybill and Deal estimator $\hat{\boldsymbol{\mu}}_{\mathrm{JK}}$,

$$\hat{\boldsymbol{\mu}}_{\mathrm{JK}} = \left(\sum_{i=1}^{I} c_i n_i \mathbf{S}_i^{-1} \right)^{-1} \sum_{j=1}^{I} c_j n_j \mathbf{S}_j^{-1} \bar{\mathbf{x}}_j, \tag{1.3}$$

which does not always produce non-empty regions. Moreover, determination of the percentile points that are needed to construct the confidence region of $\boldsymbol{\mu}$ is quite difficult in practice, and thus approximation is necessary.

In this paper, we intend to provide a method that is readily applicable for both hypothesis testing and confidence region construction of the common mean vector $\boldsymbol{\mu}$. Our approach is based on the concepts of generalized p-values and generalized confidence intervals, introduced by Tsui and Weerahandi (1989) and Weerahandi (1993), respectively. These ideas have turned out to be very satisfactory for obtaining tests and confidence intervals for many complex problems; see Lin and Lee (2003), Lee and Lin (2004) and many others. Gamage et al. (2004) provided a generalized p-value and a generalized confidence region for the multivariate Behrens–Fisher problem and MANOVA. For a discussion of several applications, the readers are referred to the book by Weerahandi (1995). In terms of the expected area or d-dimensional volumes and coverage probability, our method is compared with the methods derived by the classical approach, Graybill and Deal (1959) and Jordan and Krishnamoorthy (1995), respectively. The numerical results in Sections 4 and 5 also show that our method performs better than these methods.

The rest of the article is organized as follows. The theory of generalized p-values and generalized confidence interval will be briefly introduced in Section 2. Our procedures for hypothesis testing and construction of the generalized confidence region about the common mean vector $\boldsymbol{\mu}$ are presented in Section 3. We apply these results to two sets of data and compare with five other methods numerically with respect to their simultaneous confidence intervals and the areas of 95% confidence regions in Section 4. One simulation study is presented in Section 5 to compare the expected areas or the expected d-dimensional volumes and the coverage probabilities of these methods in different combinations of sample sizes and covariance matrices.

2. Generalized p-values and generalized confidence intervals

A set-up where the generalized p-value may be defined is as follows. Let $\mathbf{X}$ be a random quantity having a density function $f(\mathbf{X}|\zeta)$, where $\zeta = (\theta, \ \boldsymbol{\eta})$ is a vector of unknown parameters, θ is the parameter of interest, and $\boldsymbol{\eta}$ is a vector

of nuisance parameters. Suppose we are interested in testing

$$\mathrm{H}_0 : \theta \leqslant \theta_0 \quad \text{vs.} \quad \mathrm{H}_1 : \theta > \theta_0, \tag{2.1}$$

where θ_0 is a pre-specified value.

Let $\mathbf{x}$ denote the observed value of $\mathbf{X}$ and consider the *generalized test variable* $T(\mathbf{X}; \mathbf{x}, \theta, \boldsymbol{\eta})$, which depends on the observed value $\mathbf{x}$ and the parameters ζ, and satisfies the following requirements:

(i) For fixed $\mathbf{x}$ and $\zeta = (\theta_0, \ \boldsymbol{\eta})$, the distribution of $T(\mathbf{X}; \mathbf{x}, \theta_0, \boldsymbol{\eta})$ is free of the nuisance parameters $\boldsymbol{\eta}$.
(ii) The observed value $T(\mathbf{x}; \mathbf{x}, \theta_0, \boldsymbol{\eta})$ of $T(\mathbf{X}; \mathbf{x}, \theta_0, \boldsymbol{\eta})$ does not depend on unknown parameters $\boldsymbol{\eta}$.
(iii) For fixed $\mathbf{x}$ and $\boldsymbol{\eta}$, $\Pr[T(\mathbf{X}; \mathbf{x}, \theta, \boldsymbol{\eta}) \geqslant t]$ is either increasing or decreasing in θ for any given t. (A)

Under the above conditions, if $T(\mathbf{X}; \mathbf{x}, \theta, \boldsymbol{\eta})$ is stochastically increasing in θ, then the generalized p-value for testing the hypothesis in (2.1) can be defined as

$$p = \sup_{\theta \leqslant \theta_0} \Pr[T(\mathbf{X}; \mathbf{x}, \theta, \boldsymbol{\eta}) \geqslant t] = \Pr[T(\mathbf{X}; \mathbf{x}, \theta_0, \boldsymbol{\eta}) \geqslant t], \tag{2.2}$$

where $t = T(\mathbf{x}; \mathbf{x}, \theta_0, \boldsymbol{\eta})$.

Under the same set-up, suppose $T_1(\mathbf{X}; \mathbf{x}, \theta, \boldsymbol{\eta})$ satisfies the following conditions:

(i) The distribution of $T_1(\mathbf{X}; \mathbf{x}, \theta, \boldsymbol{\eta})$ does not depend on any unknown parameters.
(ii) The observed value of $T_1(\mathbf{X}; \mathbf{x}, \theta, \boldsymbol{\eta})$ is free of nuisance parameters $\boldsymbol{\eta}$. (B)

Then we say $T_1(\mathbf{X}; \mathbf{x}, \theta, \boldsymbol{\eta})$ is a *generalized pivotal quantity* (GPQ). Let c_1 and c_2 be such that

$$\Pr[c_1 \leqslant T_1(\mathbf{X}; \mathbf{x}, \theta, \boldsymbol{\eta}) \leqslant c_2] = 1 - \alpha, \tag{2.3}$$

then $\{\theta : c_1 \leqslant T_1(\mathbf{x}; \mathbf{x}, \theta, \boldsymbol{\eta}) \leqslant c_2\}$ is a $100(1-\alpha)\%$ generalized confidence interval for θ. For example, if the value of $T_1(\mathbf{X}; \mathbf{x}, \theta, \boldsymbol{\eta})$ at $\mathbf{X} = \mathbf{x}$ is θ and $q_{\{T_1(\mathbf{X}); 1-\alpha\}}$ represents the $100(1-\alpha)$th percentile of $T_1(\mathbf{X}; \mathbf{x}, \theta, \boldsymbol{\eta})$, then $\{q_{\{T_1(\mathbf{x}); \alpha/2\}}, q_{\{T_1(\mathbf{x}); 1-\alpha/2\}}\}$ is a $100(1-\alpha)\%$ confidence interval for θ.

3. Confidence regions for $\boldsymbol{\mu}$

Suppose we have $I\,(I \geqslant 2)$ independent d-variate normal populations with common mean vector $\boldsymbol{\mu}$ and unequal covariance matrices $\boldsymbol{\Sigma}_1, \ldots, \boldsymbol{\Sigma}_I$. For the ith sample, $\bar{\mathbf{X}}_i$ and $\mathbf{S}_i$, defined in (1.1), are mutually independent with

$$\bar{\mathbf{X}}_i \sim N_d(\boldsymbol{\mu}, \ \boldsymbol{\Sigma}_i/n_i) \quad \text{and} \quad \mathbf{A}_i = (n_i - 1)\mathbf{S}_i \sim W_d(n_i - 1, \boldsymbol{\Sigma}_i), \quad i = 1, \ldots, I, \tag{3.1}$$

respectively, where $N_d(\boldsymbol{\mu}, \ \boldsymbol{\Sigma})$ is the d-variate normal distribution with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$ and $W_d(v, \ \boldsymbol{\Sigma})$ denotes the d-dimensional Wishart distribution with degrees of freedom v and scale matrix $\boldsymbol{\Sigma}$.

In this section, we will first provide a generalized confidence region of $\boldsymbol{\mu}$ based on a GPQ and then briefly review three other confidence regions of $\boldsymbol{\mu}$ proposed by Graybill and Deal (1959), Jordan and Krishnamoorthy (1995) and the classical method, respectively.

3.1. Solutions based on the generalized method

As the focus of attention is the common mean vector of I d-variate normal distributions, the elements of the variance–covariance matrices $\boldsymbol{\Sigma}_i$, $i = 1, \ldots, I$, play the role of nuisance parameters. Thus, in order to form a GPQ for $\boldsymbol{\mu}$, one must replace $\boldsymbol{\Sigma}_i$ in $(\boldsymbol{\Sigma}_i/n_i)^{-1/2}(\bar{\mathbf{X}}_i - \boldsymbol{\mu}) = \mathbf{Z}_i \sim N_d(\mathbf{0}, \ \mathbf{I}_d)$ with its own GPQ. The result is a GPQ for $\boldsymbol{\mu}$, based on one sample, which is given by

$$\mathbf{T}_i^* = \bar{\mathbf{x}}_i - (\mathbf{u}_i^{1/2}\mathbf{R}_i^{-1}\mathbf{u}_i^{1/2})^{1/2}\mathbf{Z}_i, \tag{3.2}$$

where

$$\mathbf{R}_i = [\mathbf{u}_i^{-1/2}(\boldsymbol{\Sigma}_i/n_i)\mathbf{u}_i^{-1/2}]^{-1/2}(\mathbf{u}_i^{-1/2}\mathbf{U}_i\mathbf{u}_i^{-1/2})[\mathbf{u}_i^{-1/2}(\boldsymbol{\Sigma}_i/n_i)\mathbf{u}_i^{-1/2}]^{-1/2} \tag{3.3}$$

and $\mathbf{u}_i$ is the observed value of $\mathbf{U}_i$ with $\mathbf{U}_i = \mathbf{A}_i/n_i \sim W_d(n_i - 1,\ \Sigma_i/n_i)$. It is noted that $\mathbf{R}_i \sim W_d(n_i - 1,\ I_d)$ and the use of $\mathbf{R}_i^*$ can be found in Gamage et al. (2004) with $\mathbf{R}_i^* = \mathbf{R}_i/(n_i - 1)$. The individual GPQs for $\boldsymbol{\mu}$ are combined in a weighted manner, with weights $\mathbf{W}^{-1}\mathbf{W}_i$ are proportional to the sample sizes and the inverse of the covariance matrices with$\mathbf{W}^{-1}\mathbf{W}_i = \mathbf{W}^{-1}(\mathbf{u}_i^{1/2}\tilde{\mathbf{R}}_i^{-1}\mathbf{u}_i^{1/2})^{-1}$, where $\mathbf{W} = \sum_{i=1}^{I}\mathbf{W}_i$ and $\tilde{\mathbf{R}}_i$ are different Wishart random matrices with $\tilde{\mathbf{R}}_i \sim W_d(n_i - 1,\ I_d)$. Since the distribution of $\mathbf{Z}_i$'s and $\mathbf{R}_i$'s are free of any unknown parameters and these quantities are independent, it follows that the distributions of $\mathbf{T}_i^*$'s are free of any unknown parameters. Furthermore, using the definition of $\mathbf{Z}_i$'s and $\mathbf{R}_i$'s, we conclude that the observed value of $\mathbf{T}_i^*$ is $\boldsymbol{\mu}$ for all i, thus we can construct a weighted average of the GPQs based on individual samples. Let $\bar{\mathbf{X}} = (\bar{\mathbf{X}}_1, \ldots, \bar{\mathbf{X}}_I)$ and $\mathbf{U} = (\mathbf{U}_1, \ldots, \mathbf{U}_I)$, and $\bar{\mathbf{x}}$ and $\mathbf{u}$ be the corresponding observed values of $\bar{\mathbf{X}}$ and $\mathbf{U}$, respectively, then the GPQ can be expressed as

$$\mathbf{T}(\bar{\mathbf{X}}, \mathbf{U}; \bar{\mathbf{x}}, \mathbf{u}) = \mathbf{W}^{-1}\sum_{i=1}^{I}\mathbf{W}_i\mathbf{T}_i^*, \tag{3.4}$$

where $\mathbf{W} = \sum_{i=1}^{I}\mathbf{W}_i$ and $\mathbf{T}_i^*$ is given in (3.2).

To construct confidence region based on $\mathbf{T}$, we need to verify that $\mathbf{T}$ in (3.4) satisfies the two conditions in (B). The value of $\mathbf{T}_i^*$ in (3.2) at $(\bar{\mathbf{X}}_i, \mathbf{U}_i) = (\bar{\mathbf{x}}_i, \mathbf{u}_i)$ is $\boldsymbol{\mu}$ and $\mathbf{W}_i = n_i\boldsymbol{\Sigma}_i^{-1}$ at $\mathbf{U}_i = \mathbf{u}_i$, $i = 1, \ldots, I$. Therefore, $\mathbf{T} = \boldsymbol{\mu}$ at $(\bar{\mathbf{X}}, \mathbf{U}) = (\bar{\mathbf{x}}, \mathbf{u})$. It is also clear from (3.2) to (3.4) that, for a given$(\bar{\mathbf{x}}, \mathbf{u})$, the distribution of $\mathbf{T}$ is independent of any unknown parameters. Therefore, $\mathbf{T}$ in (3.4) satisfies the two conditions in (B) and is truly a GPQ, which can be used to construct a confidence region for $\boldsymbol{\mu}$.

As Krishnamoorthy and Lu (2003) pointed out it is better to use different Wishart random matrices $\mathbf{R}_i$ and $\tilde{\mathbf{R}}_i$ for constructing generalized variable $\mathbf{T}_i^*$and $\mathbf{W}_i$ in order to avoid producing too liberal results.

3.1.1. Hypothesis testing and confidence region

Given $(\bar{\mathbf{x}}, \mathbf{u})$, the distribution of $\mathbf{T}$ in (3.4) is free of unknown parameters and hence $\mathbf{T}$ can be used to construct a confidence region of $\boldsymbol{\mu}$ and test the hypothesis

$$\mathrm{H}_0 : \boldsymbol{\mu} = \boldsymbol{\mu}_0 \quad \text{vs.} \quad \mathrm{H}_1 : \boldsymbol{\mu} \neq \boldsymbol{\mu}_0, \tag{3.5}$$

where $\boldsymbol{\mu}_0$ is a pre-specified value.

Remark 1. If $\mathbf{a}$ is a $d \times 1$ column vector with elements $a_1, a_2, \ldots, a_d < \infty$, we write $\mathbf{a}' = (a_1, \ldots, a_d)$, and the *length* or *norm* of $\mathbf{a}$ is denoted by $\|\mathbf{a}\|$. Thus $\|\mathbf{a}\| = \sqrt{\mathbf{a}'\mathbf{a}} = (a_1^2 + a_2^2 + \cdots + a_d^2)^{1/2}$.

Remark 2. For a vector $\mathbf{y}$, $|\mathbf{b}'\mathbf{y}| \leqslant c(\mathbf{b}'\mathbf{b})^{1/2}$ if and only if $\mathbf{y}'\mathbf{y} \leqslant c^2$, for any nonzero vector $\mathbf{b}$, which is a consequence of the *Cauchy–Schwarz inequality*.

Let $\tilde{\mathbf{T}}$ denote the standardized expression of $\mathbf{T}$ with $\tilde{\mathbf{T}} = \boldsymbol{\Sigma}_{\mathbf{T}}^{-1/2}(\mathbf{T} - \boldsymbol{\mu}_{\mathbf{T}})$, where $\boldsymbol{\mu}_{\mathbf{T}}$ and $\boldsymbol{\Sigma}_{\mathbf{T}}$ represent the mean and covariance matrix of $\mathbf{T}$. Define $\tilde{\boldsymbol{\mu}}_0 = \boldsymbol{\Sigma}_{\mathbf{T}}^{-1/2}(\boldsymbol{\mu}_0 - \boldsymbol{\mu}_{\mathbf{T}})$, and then the generalized p-value for testing (3.5) can be given by

$$p = P\{\|\tilde{\mathbf{T}}\| > \|\tilde{\boldsymbol{\mu}}_0\| | \bar{\mathbf{x}}, \mathbf{u}\}, \tag{3.6}$$

and H_0 will be rejected whenever p is less than α. Furthermore, let $q_{\{\|\tilde{\mathbf{T}}\|;\ \gamma\}}$ be the 100γth percentile of $\|\tilde{\mathbf{T}}\|$, so we have

$$P\{\tilde{\mathbf{T}}'\tilde{\mathbf{T}} = (\mathbf{T} - \boldsymbol{\mu}_{\mathbf{T}})'\boldsymbol{\Sigma}_{\mathbf{T}}^{-1}(\mathbf{T} - \boldsymbol{\mu}_{\mathbf{T}}) \leqslant q^2_{\{\|\tilde{\mathbf{T}}\|;\ \gamma\}}\} = \gamma. \tag{3.7}$$

Since the observed value of $\mathbf{T}$ is $\boldsymbol{\mu}$, the $100(1 - \alpha)\%$ confidence region of $\boldsymbol{\mu}$ can be solved by the inequality

$$\{\boldsymbol{\mu} : (\boldsymbol{\mu} - \boldsymbol{\mu}_{\mathbf{T}})'\boldsymbol{\Sigma}_{\mathbf{T}}^{-1}(\boldsymbol{\mu} - \boldsymbol{\mu}_{\mathbf{T}}) \leqslant q^2_{\{\|\tilde{\mathbf{T}}\|;1-\alpha\}}\}, \tag{3.8}$$

which is equivalent to solving the inequality $\{\boldsymbol{\mu} : \|\boldsymbol{\Sigma}_{\mathbf{T}}^{-1/2}(\boldsymbol{\mu} - \boldsymbol{\mu}_{\mathbf{T}})\| \leqslant q_{\{\|\tilde{\mathbf{T}}\|;\ 1-\alpha\}}\}$.

Simultaneous confidence intervals for the d-components of $\boldsymbol{\mu}$ can be developed from consideration of confidence intervals for $\mathbf{a}'\mathbf{T}$, where $\mathbf{a}$ is any nonzero d-variate vector.

According to Remark 2, let $\mathbf{y}=\boldsymbol{\Sigma}_{\mathbf{T}}^{-1/2}(\mathbf{T}-\boldsymbol{\mu}_{\mathbf{T}})$, $\mathbf{b}=\boldsymbol{\Sigma}_{\mathbf{T}}^{1/2}\mathbf{a}$ and $c=q_{\{\|\tilde{\mathbf{T}}\|;\ 1-\alpha\}}$, then from (3.7), we have the following:

$$(\mathbf{T}-\boldsymbol{\mu}_{\mathbf{T}})'\boldsymbol{\Sigma}_{\mathbf{T}}^{-1}(\mathbf{T}-\boldsymbol{\mu}_{\mathbf{T}})\leqslant q^2_{\{\|\tilde{\mathbf{T}}\|;\ 1-\alpha\}} \quad \text{if and only if } |\mathbf{a}'(\mathbf{T}-\boldsymbol{\mu}_{\mathbf{T}})|\leqslant q_{\{\|\tilde{\mathbf{T}}\|;\ 1-\alpha\}}\sqrt{\mathbf{a}'\boldsymbol{\Sigma}_{\mathbf{T}}\mathbf{a}}. \tag{3.9}$$

The inequality (3.9) implies that

$$P\{\mathbf{a}'\boldsymbol{\mu}_{\mathbf{T}}-q_{\{\|\tilde{\mathbf{T}}\|;\ 1-\alpha\}}\sqrt{\mathbf{a}'\boldsymbol{\Sigma}_{\mathbf{T}}\mathbf{a}}\leqslant\mathbf{a}'\mathbf{T}\leqslant\mathbf{a}'\boldsymbol{\mu}_{\mathbf{T}}+q_{\{\|\tilde{\mathbf{T}}\|;\ 1-\alpha\}}\sqrt{\mathbf{a}'\boldsymbol{\Sigma}_{\mathbf{T}}\mathbf{a}}\}=1-\alpha, \tag{3.10}$$

for all nonzero vector $\mathbf{a}$. If $\mathbf{a}$ is the vector with 1 for the lth element and 0 elsewhere, the simultaneous $100(1-\alpha)\%$ confidence interval for the lth component of common mean vector $\boldsymbol{\mu}$, μ_l, is

$$(\mu_{\mathbf{T}(l)}-q_{\{\|\tilde{\mathbf{T}}\|;1-\alpha\}}\sqrt{\Sigma_{\mathbf{T}}^{(l,l)}},\ \mu_{\mathbf{T}(l)}+q_{\{\|\tilde{\mathbf{T}}\|;\ 1-\alpha\}}\sqrt{\Sigma_{\mathbf{T}}^{(l,l)}}),\quad l=1,\ldots,d, \tag{3.11}$$

where $\mu_{\mathbf{T}(l)}$ is the lth component of $\boldsymbol{\mu}_{\mathbf{T}}$ and $\Sigma_{\mathbf{T}}^{(l,l)}$ is the (l,l)th component of $\boldsymbol{\Sigma}_{\mathbf{T}}$. In practice, we need the following algorithm.

Algorithm 1. For a given $(n_1,\ldots,n_I)$, $(\bar{\mathbf{x}}_1,\ldots,\bar{\mathbf{x}}_I)$ and $(\mathbf{u}_1,\ldots,\mathbf{u}_I)$:

For $j=1,\ldots,m$:
Generate $\mathbf{Z}_1,\ldots,\mathbf{Z}_I$ from $N_d(\mathbf{0},\ \mathbf{I}_d)$.
Generate $\mathbf{R}_i$ and $\tilde{\mathbf{R}}_i$ from $W_d(n_i-1,\ \mathbf{I}_d)$, $i=1,\ldots,I$.
Compute $\mathbf{W}_1,\ldots,\mathbf{W}_I$ and $\mathbf{W}$.
Compute $\mathbf{T}_j=\mathbf{W}^{-1}\sum_{i=1}^{I}\mathbf{W}_i\,\mathbf{T}_i^*$.
(End j loop)
Compute $\hat{\boldsymbol{\mu}}_{\mathbf{T}}=1/m\sum_{j=1}^{m}\mathbf{T}_j$ and $\hat{\boldsymbol{\Sigma}}_{\mathbf{T}}=1/(m-1)\sum_{j=1}^{m}(\mathbf{T}_j-\hat{\boldsymbol{\mu}}_{\mathbf{T}})(\mathbf{T}_j-\hat{\boldsymbol{\mu}}_{\mathbf{T}})'$.
Compute $\|\hat{\tilde{\mathbf{T}}}_j\|$ and $\|\hat{\tilde{\boldsymbol{\mu}}}_0\|$,where $\hat{\tilde{\mathbf{T}}}_j=\hat{\boldsymbol{\Sigma}}_{\mathbf{T}}^{-1/2}(\mathbf{T}_j-\hat{\boldsymbol{\mu}}_{\mathbf{T}})$, $j=1,\ldots,m$, and $\hat{\tilde{\boldsymbol{\mu}}}_0=\hat{\boldsymbol{\Sigma}}_{\mathbf{T}}^{-1/2}(\boldsymbol{\mu}_0-\hat{\boldsymbol{\mu}}_{\mathbf{T}})$.
Let $\tau_j=1$ if $\|\hat{\tilde{\mathbf{T}}}_j\|\geqslant\|\hat{\tilde{\boldsymbol{\mu}}}_0\|$; else $\tau_j=0$.
$1/m\sum_{j=1}^{m}\tau_j$ is a Monte Carlo estimate of the generalized p-value for testing (3.6).

Let $q_{\{\|\hat{\tilde{\mathbf{T}}}\|;\ 1-\alpha\}}$ be the $100(1-\alpha)$th percentile of $\|\hat{\tilde{\mathbf{T}}}_j\|$, $j=1,\ldots,m$, then the confidence region of $\boldsymbol{\mu}$ and the simultaneous confidence interval of μ_l, $l=1,\ldots,d$, can be obtained through (3.8) and (3.11), respectively.

3.1.2. The expected area or the expected d-dimensional volume and coverage probability of the confidence region

We will compute the coverage probabilities and the expected surface areas or the expected d-dimensional volumes of the generalized confidence regions under $d\geqslant 2$.

Remark 3. Suppose we have a confidence region of $\boldsymbol{\mu}$ which satisfies the following inequality: $(\boldsymbol{\mu}-\hat{\boldsymbol{\mu}})'\mathbf{V}^{-1}(\boldsymbol{\mu}-\hat{\boldsymbol{\mu}})\leqslant c^2$, where $\mathbf{V}$ is a $d\times d$ positive definite matrix. The ellipsoid center is $\hat{\boldsymbol{\mu}}$, and the axes of the ellipsoid are $\pm|c|\sqrt{e_l}$ in the direction of $\boldsymbol{\xi}_l$, where e_l's are the eigenvalues of $\mathbf{V}$ and $\boldsymbol{\xi}_l$'s are the corresponding eigenvectors, $l=1,\ldots,d$. Thus the d-dimensional volume of $\boldsymbol{\mu}$ can be computed by $\pi^{d/2}c^d/\Gamma(1+d/2)\sqrt{|\mathbf{V}|}$, where $|\mathbf{V}|$ is the determinant of $\mathbf{V}$ and $\Gamma(\cdot)$is gamma function. Specifically, for $d=2$ and 3, the area and volume can be reduced to $\pi c^2\times\sqrt{|\mathbf{V}|}$ and $4\pi c^3/3\cdot\sqrt{|\mathbf{V}|}$, respectively.

According to Remark 3, the d-dimensional volume of the confidence region in (3.8) derived by generalized method are

$$\frac{\pi^{d/2}\cdot q^d_{\{\|\tilde{\mathbf{T}}\|;\ 1-\alpha\}}}{\Gamma(1+d/2)}\cdot\sqrt{|\boldsymbol{\Sigma}_{\mathbf{T}}|}. \tag{3.12}$$

The algorithm for computing the expected d-dimensional volume as well as coverage probability is given as follows.

Algorithm 2. For a given $(n_1, \ldots, n_I)$, $(\boldsymbol{\Sigma}_1, \ldots, \boldsymbol{\Sigma}_I)$ and $\boldsymbol{\mu}$:

For $k = 1, \ldots, K$:
Generate $\bar{\mathbf{X}}_i^{(k)} \sim N_d(\boldsymbol{\mu},\ \boldsymbol{\Sigma}_i/n_i),\ i = 1, \ldots, I$.
Generate $\mathbf{U}_i^{(k)} \sim W_d(n_i - 1,\ \boldsymbol{\Sigma}_i/n_i), i = 1, \ldots, I$.
Use Algorithm 1 to compute d-dimensional volume H_k of the confidence region in the kth iteration,
$H_k = \pi^{d/2} \cdot q^{d(k)}_{\{\|\tilde{\mathbf{T}}\|;1-\alpha\}}/\Gamma(1 + d/2)\sqrt{|\hat{\boldsymbol{\Sigma}}_{\mathbf{T}}^{(k)}|}$.
Use Algorithm 1 to compute coverage probability, set $\delta_k = 1$ if
$\|\hat{\boldsymbol{\Sigma}}_{\mathbf{T}}^{-1/2(k)}(\boldsymbol{\mu} - \hat{\boldsymbol{\mu}}_{\mathbf{T}}^{(k)})\| \leqslant q^{(k)}_{\{\|\hat{\tilde{\mathbf{T}}}\|;1-\alpha\}}$;else $\delta_k = 0$.
(End k loop)

$1/K \sum_{k=1}^{K} H_k$ and $1/K \sum_{k=1}^{K} \delta_k$ are Monte Carlo estimates of the d-dimensional volume and coverage probability of the generalized confidence region, respectively.

3.2. Solutions based on other combined tests

We will briefly review three other confidence regions that are utilized to compare with our procedure in numerical examples.

3.2.1. Solution based on Graybill and Deal (1959)

Graybill and Deal (1959) suggested a combined estimator $\hat{\mu}_{\mathrm{GD1}}$, $\hat{\mu}_{\mathrm{GD1}} = \sum_{i=1}^{I} n_i \bar{x}_i/s_i^2 / \sum_{i=1}^{I} n_i/s_i^2$, for the common mean μ and we further extend the idea to construct a confidence region of $\boldsymbol{\mu}$ centered at the estimator $\hat{\boldsymbol{\mu}}_{\mathrm{GD}d}$, $\hat{\boldsymbol{\mu}}_{\mathrm{GD}d} = (\sum_{i=1}^{I} n_i \mathbf{S}_i^{-1})^{-1} \sum_{j=1}^{I} n_j \mathbf{S}_j^{-1} \bar{\mathbf{x}}_j$. If we can find a cut-off point a_α such that

$$P\left\{\sum_{i=1}^{I} n_i(\bar{\mathbf{X}}_i - \boldsymbol{\mu})'\mathbf{S}_i^{-1}(\bar{\mathbf{X}}_i - \boldsymbol{\mu}) \leqslant a_\alpha\right\} = 1 - \alpha, \tag{3.13}$$

then the values of $\boldsymbol{\mu}$ that satisfy (3.13) form the $100(1 - \alpha)\%$ confidence region of $\boldsymbol{\mu}$.

3.2.2. Solution based on Jordan and Krishnamoorthy (1995)

Jordan and Krishnamoorthy (1995) constructed a confidence region of $\boldsymbol{\mu}$ centered at a weighted Graybill and Deal estimator $\hat{\boldsymbol{\mu}}_{\mathrm{JK}}$,

$$\hat{\boldsymbol{\mu}}_{\mathrm{JK}} = \left(\sum_{i=1}^{I} c_i n_i \mathbf{S}_i^{-1}\right)^{-1} \sum_{j=1}^{I} c_j n_j \mathbf{S}_j^{-1} \bar{\mathbf{x}}_j,$$

where the c_i's are some positive constants such that $\sum_{i=1}^{I} c_i = 1$. Thus Graybill and Deal's method can be treated as a special case of Jordan and Krishnamoorthy's method with $c_1 = \cdots = c_I$. Since $\sum_{i=1}^{I} c_i n_i (\bar{\mathbf{X}}_i - \boldsymbol{\mu})'\mathbf{S}_i^{-1}(\bar{\mathbf{X}}_i - \boldsymbol{\mu})$ is distributed as $\sum_{i=1}^{I} c_i T_i^2$ and with $(n_i - 1)T_i^2 \sim (d/(n_i - d))F_{d,\ n_i - d}$, thus the weights c_i can be chosen as

$$c_i = [Var(T_i^2)]^{-1} \Big/ \sum_{i=1}^{I} [Var(T_i^2)]^{-1} \quad \text{and} \quad Var(T_i^2) = \frac{2d(n_i - 1)^2(n_i - 2)}{(n_i - d - 2)^2(n_i - d - 4)}, \quad i = 1, \ldots, I.$$

Therefore, when the sample sizes are equal, the method proposed by Jordan and Krishnamoorthy (1995) is the same as Graybill and Deal (1959). Finding a cut-off point b_α which satisfies

$$P\left\{\sum_{i=1}^{I} c_i n_i (\bar{\mathbf{X}}_i - \boldsymbol{\mu})'\mathbf{S}_i^{-1}(\bar{\mathbf{X}}_i - \boldsymbol{\mu}) \leqslant b_\alpha\right\} = 1 - \alpha \tag{3.14}$$

Table 1
Summary data in Example 1 and Example 2

	n_1	n_2	$\bar{\mathbf{x}}_1$	$\bar{\mathbf{x}}_2$	$\mathbf{s}_1$	$\mathbf{s}_2$
Example 1	12	8	$\begin{pmatrix}4.73\\7.93\end{pmatrix}$	$\begin{pmatrix}5.29\\8.57\end{pmatrix}$	$\begin{pmatrix}2.71 & 4.47\\4.47 & 10.92\end{pmatrix}$	$\begin{pmatrix}4.23 & 1.58\\1.58 & 2.23\end{pmatrix}$
Example 2	8	12	$\begin{pmatrix}4.94\\8.04\end{pmatrix}$	$\begin{pmatrix}5.21\\8.89\end{pmatrix}$	$\begin{pmatrix}2.08 & 4.28\\4.28 & 12.99\end{pmatrix}$	$\begin{pmatrix}5.38 & 1.37\\1.37 & 2.83\end{pmatrix}$

is extremely complicated in practice, so Jordan and Krishnamoorthy (1995) suggested approximating b_α by $\lambda F_{Id,\gamma}$ with an additional requirement of $n_i > d+4$, for all i, where

$$\gamma = \frac{4IdM_2 - 2(Id+2)M_1^2}{IdM_2 - (Id+2)M_1^2}, \quad \lambda = \frac{\gamma-2}{\gamma}M_1, \quad M_1 = d\sum_{i=1}^{I}\frac{c_i(n_i-1)}{n_i-d-2} \quad \text{and}$$
$$M_2 = d(d+2)\sum_{i=1}^{I}\frac{c_i^2(n_i-1)^2}{(n_i-d-2)(n_i-d-4)} + 2d^2\sum_{i>j}\frac{c_ic_j(n_i-1)(n_j-1)}{(n_i-d-2)(n_j-d-2)}. \tag{3.15}$$

3.2.3. Solutions based on the classical method

In the classical procedure, practitioners usually ignore the non-homogeneity for mathematical tractability and simply apply the Hotelling's T^2 test. That is, we will assume that $\boldsymbol{\Sigma}_1 = \cdots = \boldsymbol{\Sigma}_I = \boldsymbol{\Sigma}$ and then the point estimator of $\boldsymbol{\mu}$ and the pool covariance matrix are

$$\hat{\boldsymbol{\mu}}_{\text{cla.}} = \sum_{j=1}^{I} n_j\bar{\mathbf{x}}_j/N \quad \text{and} \quad \mathbf{S} = \sum_{j=1}^{I}(n_j-1)\mathbf{S}_j/(N-I), \tag{3.16}$$

respectively, where $N = \sum_{i=1}^{I} n_i$. The $100(1-\alpha)\%$ confidence region of $\boldsymbol{\mu}$ is

$$\left\{\boldsymbol{\mu} : N(\hat{\boldsymbol{\mu}}_{\text{cla.}} - \boldsymbol{\mu})'\mathbf{S}^{-1}(\hat{\boldsymbol{\mu}}_{\text{cla.}} - \boldsymbol{\mu}) \leqslant \frac{d(N-I)}{N-I-d+1}F_{1-\alpha}(d,\ N-I-d+1)\right\}, \tag{3.17}$$

where $F_{1-\alpha}(d,\ N-I-d+1)$ is the $100(1-\alpha)$th percentile of the $F_{d,\ N-I-d+1}$ distribution.

4. Illustrative examples

Two numerical examples are given to illustrate the advantages of our proposed method for setting confidence limits and calculating confidence region for the common mean vector. For comparison purposes, these data are taken from Jordan and Krishnamoorthy (1995). Furthermore, for the reasons that the results of Jordan and Krishnamoorthy (1995) and Graybill and Deal (1959) are the same when the sample sizes are identical with the restriction $n_i > d+4$, for all i, which is needed for the Jordan and Krishnamoorthy (1995)'s method, we choose $n_1 = 12,\ n_2 = 8$ in the first example and $n_1 = 8,\ n_2 = 12$ in the second example. The summary data are given in Table 1.

In order to construct 95% simultaneous confidence intervals of μ_1 and μ_2 and calculate the area of the confidence region of $\boldsymbol{\mu}$, we use Algorithm 1 with $m = 10,000$ runs. For demonstration purposes, we will provide the results of six procedures to make a comparison. The results appear in Table 2.

These comparisons presented in Table 2 correspond to

(1) GD: Graybill and Deal (1959).
(2) JK: Jordan and Krishnamoorthy (1995).
(3) Classical: the classical procedure with assumption of identical covariance.
(4) Generalized method: the generalized method proposed in this article.
(5) Sample 1: the procedure based on sample 1 alone.
(6) Sample 2: the procedure based on sample 2 alone.

Table 2
Areas of 95% confidence region of $\boldsymbol{\mu}$ and the simultaneous 95% confidence intervals of μ_1 and μ_2 for Example 1 and Example 2

Methods	Point estimate $\hat{\boldsymbol{\mu}}$	95% Confidence interval of μ_1	95% confidence interval of μ_2	Area of 95% confidence region of $\boldsymbol{\mu}$
Example 1				
(1)	$\begin{pmatrix}4.92\\8.39\end{pmatrix}$	(3.61, 6.29)	(6.52, 10.26)	6.07
(2)	$\begin{pmatrix}4.87\\8.25\end{pmatrix}$	(3.61, 6.12)	(6.12, 10.37)	5.79
(3)	$\begin{pmatrix}4.95\\8.19\end{pmatrix}$	(3.83, 6.07)	(6.50, 9.88)	4.43
(4)	$\begin{pmatrix}4.95\\8.38\end{pmatrix}$	(4.03, 5.87)	(7.07, 9.68)	2.88
(5)	$\begin{pmatrix}4.73\\7.93\end{pmatrix}$	(3.30, 6.15)	(5.07, 10.80)	7.35
(6)	$\begin{pmatrix}5.29\\8.57\end{pmatrix}$	(2.77, 7.81)	(6.75, 10.40)	12.40
Example 2				
(1)	$\begin{pmatrix}5.18\\8.78\end{pmatrix}$	(3.90, 6.46)	(6.91, 10.66)	6.42
(2)	$\begin{pmatrix}5.20\\8.85\end{pmatrix}$	(3.78, 6.62)	(7.28, 10.41)	6.33
(3)	$\begin{pmatrix}5.10\\8.55\end{pmatrix}$	(3.85, 6.35)	(6.94, 10.16)	5.55
(4)	$\begin{pmatrix}5.18\\8.77\end{pmatrix}$	(4.29, 6.07)	(7.50, 10.05)	3.06
(5)	$\begin{pmatrix}4.94\\8.04\end{pmatrix}$	(3.17, 6.71)	(3.63, 12.46)	13.88
(6)	$\begin{pmatrix}5.21\\8.89\end{pmatrix}$	(3.20, 7.22)	(7.43, 10.35)	8.65

The results in Table 2 suggest that the classical method performs quite well in these two examples. The method proposed by Jordan and Krishnamoorthy (1995) is slightly better than that of Graybill and Deal (1959) when the sample sizes are not identical and the two-covariance matrices are not seriously heteroscedastic. The areas constructed by methods (1)–(4), whose procedures are based on all samples, are smaller than the areas obtained by (5) and (6) which are based only on one individual sample. The numerical results indicate that our proposed procedures based on the generalized p-value and generalized confidence intervals are more satisfactory compared to the other procedures considered.

5. Simulation studies

In simulation studies, we used Algorithm 2 with $K = 1000$ iterations and Algorithm 1 with $m = 5000$ runs to calculate the coverage probabilities of the confidence regions of $\boldsymbol{\mu}$ and the expected areas or expected d-dimensional volumes. Jordan and Krishnamoorthy (1995) have mentioned that the confidence regions of (5) and (6) are larger than (2) under the situation that the "generalized variances" between the sample mean vectors are not too large. Furthermore, we think it is not reasonable to construct a confidence region of the common mean vector based only on one particular sample. Therefore, we compared methods (1)–(4) in simulation studies with the various combinations of samples sizes and covariance matrices. Since the sampling distributions of $\bar{\mathbf{X}}$'s are location invariant, without loss of generality, these normal random vectors are generated with zero mean vector. Furthermore, according to *eigen decomposition theorem*, for any positive definite matrix $\sum$, there exists an orthogonal matrix P such that $P' \sum P$ is diagonal (see also Rao, 1973, p. 39). Thus we chose $\Sigma_1 = I_d$ and Σ_i to be diagonal in simulation studies. The estimated expected surface areas or the expected hyper-volumes of 95% confidence regions of $\boldsymbol{\mu}$ as well as the corresponding coverage probabilities under different scenarios are given in Tables 3–6.

Table 3
Expected areas and corresponding coverage probabilities of 95% confidence regions of $\boldsymbol{\mu}$ under $\boldsymbol{\Sigma}_1 = \mathbf{I}_2$ and $\boldsymbol{\Sigma}_2 = \frac{n_2}{n_1} a\mathbf{I}_2$

n = (20, 10)	(1)	(2)	(3)	(4)
a				
3	1.425 (0.953)	1.133 (0.952)	1.141 (0.943)	0.789 (0.946)
9	1.582 (0.949)	1.197 (0.948)	2.414 (0.922)	0.981 (0.943)
15	1.630 (0.948)	1.218 (0.948)	3.683 (0.913)	1.048 (0.954)
25	1.639 (0.947)	1.222 (0.948)	5.763 (0.903)	1.106 (0.947)
50	1.666 (0.951)	1.234 (0.950)	10.950 (0.898)	1.109 (0.945)
100	1.669 (0.952)	1.233 (0.951)	21.407 (0.897)	1.114 (0.950)
500	1.700 (0.958)	1.250 (0.957)	104.620 (0.898)	1.127 (0.948)

Table 4
Expected areas and corresponding coverage probabilities of 95% confidence regions of $\boldsymbol{\mu}$ under $\boldsymbol{\Sigma}_1 = \mathbf{I}_2$ and $\boldsymbol{\Sigma}_2 = \text{diag}(1,\ a)$

n = (20, 10)	(1)	(2)	(3)	(4)
a				
4	1.259 (0.949)	1.053 (0.954)	0.976 (0.947)	0.825 (0.947)
9	1.304 (0.951)	1.072 (0.946)	1.311 (0.938)	0.853 (0.952)
16	1.341 (0.953)	1.091 (0.954)	1.678 (0.841)	0.881 (0.946)
25	1.330 (0.952)	1.083 (0.951)	2.063 (0.934)	0.869 (0.950)
50	1.333 (0.948)	1.083 (0.946)	2.856 (0.936)	0.887 (0.951)
100	1.334 (0.942)	1.084 (0.945)	3.968 (0.925)	0.874 (0.945)
500	1.342 (0.944)	1.088 (0.949)	8.676 (0.930)	0.876 (0.953)

Table 5
Expected five-dimensional volumes and corresponding coverage probabilities of 95% confidence region of $\boldsymbol{\mu}$ under $\boldsymbol{\Sigma}_1 = \mathbf{I}_5$, $\boldsymbol{\Sigma}_2 = 3 \times \mathbf{I}_5$, and $\boldsymbol{\Sigma}_3 = a \cdot \mathbf{I}_5$

n = (10, 20, 15)	(1)	(2)	(3)	(4)
a				
3	13.836 (0.940)	16.735 (0.946)	2.367 (0.952)	8.672 (0.955)
9	20.483 (0.931)	28.597 (0.941)	9.357 (0.937)	15.164 (0.960)
15	24.746 (0.952)	35.098 (0.962)	21.992 (0.926)	16.959 (0.954)
25	26.038 (0.944)	38.793 (0.962)	57.681 (0.916)	17.986 (0.959)
50	27.629 (0.951)	40.912 (0.958)	250.797 (0.860)	17.851 (0.950)
100	28.117 (0.937)	42.170 (0.956)	1189.15 (0.850)	18.674 (0.953)
500	28.883 (0.937)	43.674 (0.956)	59807.72 (0.830)	19.565 (0.955)

We also considered the cases with $n = (10, 20)$ and $n = (10, 10)$ in Tables 3 and 4, $n = (10, 15, 20)$ and $n = (10, 10, 10)$ in Table 5 and $n = (15, 12, 20)$ and $n = (15, 15, 15)$ in Table 6. The numerical results showed the same pattern and hence are not reported in the paper. From these tables, we find that the classical method performs well only when the populations are almost homogeneous, but, as expected, its performance grows worse as the degree of non-homogeneity increases. For overall comparisons, the three methods allowing heteroscedasticity are much better than the classical approach when non-homogeneity is present. Between methods (1) and (2), there is no clear-cut winner. Graybill and Deal's method is preferable in the situation in which smaller sample sizes are associated with smaller variances, whereas Jordan and Krishnamoorthy's method is better when smaller sample sizes are associated with larger variances. Hence, some knowledge regarding the relation between the sample size and population variance is necessary in order to choose between these two methods. Furthermore, these two methods are under the restrictions of $n_i > d + 4$, thus they cannot be applied to populations with small sample sizes and high dimensions. However, it is clear that our proposed method, derived through the concepts of generalized p-value and generalized confidence region, is not constrained by this restriction and the results are mostly better than any of the existing methods in the senses of almost

Table 6
Expected seven-dimensional volumes and corresponding coverage probabilities of 95% confidence region of $\boldsymbol{\mu}$ under $\boldsymbol{\Sigma}_1 = \mathbf{I}_7$, $\boldsymbol{\Sigma}_2 = 3 \cdot \mathbf{I}_7$, and $\boldsymbol{\Sigma}_3 = a \cdot \mathbf{I}_7$

n = (12, 15, 20)	(1)	(2)	(3)	(4)
a				
3	60.071 (0.941)	93.157 (0.941)	3.175 (0.950)	27.749 (0.952)
9	128.119 (0.950)	680.555 (0.949)	34.596 (0.930)	75.758 (0.957)
15	146.985 (0.948)	1231.954 (0.957)	132.133 (0.901)	104.231 (0.948)
25	167.940 (0.955)	1953.954 (0.958)	588.661 (0.896)	128.694 (0.951)
50	183.259 (0.941)	2837.132 (0.945)	5127.91 (0.879)	149.833 (0.953)
100	193.690 (0.943)	3689.861 (0.956)	51495.91 (0.837)	163.402 (0.949)
500	208.448 (0.942)	4579.084 (0.956)	129905.94 (0.842)	169.208 (0.956)

always, with few exceptions, having the smallest expected areas or *d*-dimensional volumes as well as having good coverage probabilities. Thus, we conclude that the generalized method is very efficient and readily applicable for practical use.

Acknowledgement

We are grateful to the associate editor and two anonymous referees for their helpful comments and suggestions which improved the final version of the paper.

References

Chiou, W., Cohen, A., 1985. On estimating a common multivariate normal mean vector. Ann. Inst. Statist. Math. 37, 499–506.
Eberhardt, K.R., Reeve, C.P., Spiegelman, C.H., 1989. A minimax approach to combining means, with practical examples. Chemometrics Intell. Lab. Systems 5, 129–148.
Fairweather, W.R., 1972. A method of obtaining an exact confidence interval for the common mean of several normal populations. Appl. Statist. 21, 229–233.
Gamage, J., Mathew, T., Weerahandi, S., 2004. Generalized *p*-values and generalized confidence regions for the multivariate Behrens–Fisher problem and MANOVA. J. Multivariate Anal. 88, 177–189.
Graybill, F.A., Deal, R.B., 1959. Combining unbiased estimators. Biometrics 15, 543–550.
Jordan, S.M., Krishnamoorthy, K., 1995. Confidence regions for the common mean vector of several multivariate normal populations. Canad. J. Statist. 23, 283–297.
Jordan, S.M., Krishnamoorthy, K., 1996. Exact confidence intervals for the common mean of several normal populations. Biometrics 52, 77–86.
Krishnamoorthy, K., Lu, Y., 2003. Inferences on the common mean of several normal populations based on the generalized variable method. Biometrics 59, 237–247.
Lee, J.C., Lin, S.H., 2004. Generalized confidence intervals for the ratio of means of two normal populations. J. Statist. Plann. Inference 123, 49–60.
Lin, S.H., Lee, J.C., 2003. Exact tests in simple growth curve models and one-way ANOVA with equicorrelation error structure. J. Multivariate Anal. 84, 351–368.
Lin, S.H., Lee, J.C., 2005. Generalized confidence interval for the common mean of several normal populations. J. Statist. Plann. Inference 134, 568–582.
Loh, W.L., 1991. Estimating the common mean of two multivariate normal distributions. Ann. Statist. 19, 283–296.
Maric, N., Graybill, F.A., 1979. Small samples confidence intervals on common mean of two normal distributions with unequal variances. Comm. Statist. Theory Methods A 8, 1255–1269.
Meier, P., 1953. Variance of a weighted mean. Biometrics 9, 59–73.
Pagurova, V.I., Gurskii, V.V., 1979. A confidence interval for the common mean of several normal distributions. Theory Probab. Appl. 88, 882–888.
Rao, C.R., 1973. Linear Statistical Inference and its Applications. Wiley, New York.
Sinha, B.K., 1985. Unbiased estimation of the variance of the Graybill–Deal estimator of the common mean of several normal populations. Canad. J. Statist. 13, 243–247.
Tsui, K., Weerahandi, S., 1989. Generalized *p*-values in significance testing of hypotheses in the presence of nuisance parameters. J. Amer. Statist. Assoc. 84, 602–607.
Weerahandi, S., 1993. Generalized confidence intervals. J. Amer. Statist. Assoc. 88, 899–905.
Weerahandi, S., 1995. Exact Statistical Methods for Data Analysis. Springer, New York.
Zhou, L., Mathew, T., 1994. Combining independent tests in multivariate linear models. J. Multivariate Anal. 51, 265–276.