Hacettepe Journal of Mathematics & Statistics

Hacet. J. Math. Stat.
Volume 51 (4) (2022), 1211 – 1218
DOI : 10.15672/hujms.871588

Research Article

An exact test for equality of two normal mean vectors with monotone missing data

Jianqi Yu[1], Bin Wang*[1], Tao Zhang[2]

[1] College of Science, Guilin University of Technology, Guilin, Guangxi 541004, China
[2] College of Science, Guangxi University of Science and Technology, Liuzhou 545006, China

Abstract

The problem of testing equality of two normal mean vectors with incomplete data when the covariance matrices are equal is considered. For data matrices with monotone missing pattern, an exact test is proposed as an alternative one to the traditional likelihood ration test. Numerical power comparisons show that the powers of the proposed test and the likelihood ration test are comparable. However, the proposed test is an exact one. It is easy to use and useful to identify the component that caused the rejection of null hypothesis. It is illustrated using an example.

Mathematics Subject Classification (2020). 62H12, 62H15

Keywords. Monotone missing data, likelihood ration test, normal mean vector

1. Introduction

Missing data occurs very commonly in practical and has aroused an considerable amount of interest among statisticians. A variety of statistical methods have been developed to analyze data with missing values, see, e.g., [2, 9, 11, 13–15].

The reasons for data missing could be various which are not of our concern. To ignore the mechanics causing missing data, we assume that the data are missing at random (MAR). Lu and Copas [10] pointed out that inference from the likelihood method ignoring the missing data mechanism is valid if and only if the missing data mechanism is MAR. Refer to [8, 9] for the meaning of MAR.

There are a few missing patterns considered in the literature, but the incomplete data with monotone pattern (see Display (1.1)) not only occurs frequently in practice but also it is convenient for making inference. Moreover, if multivariate normality is assumed, then the monotone pattern allows the exact calculation of the maximum likelihood estimators (MLEs), the likelihood ratio statistics and relevant distributions. While it is an difficult job to achieve an exact test, many authors have considered the monotone missing pattern under the normality assumption, and provided asymptotic as well as approximate test procedures about the normal mean vector and covariance matrix. Anderson [1], one of the earliest papers in this area, put forward a simple approach to derive the MLEs of the mean

*Corresponding Author.
Email addresses: 498616729@qq.com (J. Yu), binwangglut@163.com (B. Wang), taozhangstat@gxust.edu.cn (T. Zhang)
Received: 01.02.2021; Accepted: 21.05.2022

and the covariance matrix by solving the likelihood equations for monotone missing data. Kanda and Fujikoshi [4] studied the distribution of the MLEs in the cases of 2-step, 3-step, and general k-step monotone missing data. Krishnamoorthy and Pannala [5, 6] provided an accurate and simple approach to construct the confidence region of the normal mean vector. Hao and Krishnamoorthy [3] developed an inferential procedure on a normal covariance matrix. For two populations, Yu et al. [17] considered the problem of testing equality of two normal mean vectors with the assumption that the two covariance matrices are equal, while Krishnamoorthy and Yu [7] considered the Behrens-Fisher Problem. Seko et al. [12] investigated the problem of testing equality of two normal mean vectors with 2-step monotone missing data by an other approximate approach, and Yagi and Seo [16] solved the same problem with 3-step monotone missing data. In this paper, we consider the problem of testing equality of two normal mean vectors with monotone missing data. Unlike the papers dealing with the same problem just mentioned, we present an exact test which is easy to use and useful to identify the component that caused the rejection of null hypothesis..

To formulate the problem, let $\boldsymbol{x}$ be a vector which follows a p-variate normal distribution with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$, i.e., $\boldsymbol{x} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Let $\boldsymbol{y} \sim N_p(\boldsymbol{\beta}, \boldsymbol{\Sigma})$ independently of $\boldsymbol{x}$. Suppose that we have a random sample of size N_1 on $\boldsymbol{x}$, and a random sample of size M_1 on $\boldsymbol{y}$. Assume that the samples have the following monotone pattern:

$$\begin{array}{ll} \boldsymbol{x}_{11},\ldots,\boldsymbol{x}_{1N_k},\ldots,\boldsymbol{x}_{1N_2},\ldots,\boldsymbol{x}_{1N_1} & \boldsymbol{y}_{11},\ldots,\boldsymbol{y}_{1M_k},\ldots,\boldsymbol{y}_{1M_2},\ldots,\boldsymbol{y}_{1M_1} \\ \boldsymbol{x}_{21},\ldots,\boldsymbol{x}_{2N_k},\ldots,\boldsymbol{x}_{2N_2} & \boldsymbol{y}_{21},\ldots,\boldsymbol{y}_{2M_k},\ldots,\boldsymbol{y}_{2M_2} \\ \cdots\cdots & \cdots\cdots \\ \boldsymbol{x}_{k1},\ldots,\boldsymbol{x}_{kN_k} & \boldsymbol{y}_{k1},\ldots,\boldsymbol{y}_{kM_k} \end{array} \tag{1.1}$$

where $\boldsymbol{x}_{ij}$ is a $p_i \times 1$ vector, $j = 1,\ldots N_i$, while $\boldsymbol{y}_{ij}$ is a $q_i \times 1$ vector, $j = 1,\ldots M_i$, $i = 1, ..., k$. In other words, in the $\boldsymbol{x}$-sample, there are N_1 observations available on the first p_1 components, N_2 observations available on the first $p_1 + p_2$ components, and so on. Notice that $N_1 \geq N_2 \geq ... \geq N_k$, $M_1 \geq M_2 \geq ... \geq M_k$, and $p_1 + ... + p_k = q_1 + ... + q_k = p$. We want to test

$$H_0 : \boldsymbol{\mu} = \boldsymbol{\beta} \quad \text{vs} \quad H_a : \boldsymbol{\mu} \neq \boldsymbol{\beta}. \tag{1.2}$$

In the following section, first we present the likelihood ratio test, then propose an exact test by combining two independent test using union-intersection principle. In Section 3, power comparisons are carried out using simulation. The simulation results indicate that the powers of the proposed test and the likelihood ration test are comparable. However, the proposed test is an exact one. It is easy to use and useful to identify the component that caused the rejection of null hypothesis. The method is illustrated using an example in Section 4.

2. Inference on $\boldsymbol{\mu} - \boldsymbol{\beta}$

Consider the data matrices in (1.1) with $k = 2$ and assume that $p_i = q_i$, $i = 1, 2$, and partition the data matrix $\mathbf{X}$ as follows:

$$\begin{aligned} \mathbf{X}_1 &= (\mathbf{x}_{11},\ldots,\mathbf{x}_{1N_2},\ldots,\mathbf{x}_{1N_1})_{p_1\times N_1}, \\ \mathbf{X}_2 &= \begin{pmatrix} \mathbf{x}_{11}, & \cdots & ,\mathbf{x}_{1N_2} \\ \mathbf{x}_{21}, & \cdots & ,\mathbf{x}_{2N_2} \end{pmatrix}_{p\times N_2}. \end{aligned} \tag{2.1}$$

Partition the matrix $\mathbf{Y}$ similarly. That is

$$\begin{aligned} \mathbf{Y}_1 &= (\mathbf{y}_{11},\ldots,\mathbf{y}_{1M_2},\ldots,\mathbf{y}_{1M_1})_{p_1\times M_1}, \\ \mathbf{Y}_2 &= \begin{pmatrix} \mathbf{y}_{11}, & \cdots & ,\mathbf{y}_{1M_2} \\ \mathbf{y}_{21}, & \cdots & ,\mathbf{y}_{2M_2} \end{pmatrix}_{p\times M_2}. \end{aligned} \tag{2.2}$$

Let $\bar{\boldsymbol{x}}_l$ and $\mathbf{S}_l$ denote respectively the sample mean vector and the sum of squares and sum of products matrix based on $\mathbf{X}_l$, $l = 1, 2$. We partition these means and matrices accordingly as follows:

$$\bar{\boldsymbol{x}}_1 = \bar{\boldsymbol{x}}_1^{(1)}, \quad \bar{\boldsymbol{x}}_2 = \begin{pmatrix} \bar{\boldsymbol{x}}_2^{(1)} \\ \bar{\boldsymbol{x}}_2^{(2)} \end{pmatrix}, \quad \mathbf{S}_1 = \mathbf{S}_1^{(1,1)} \text{ and } \mathbf{S}_2 = \begin{pmatrix} \mathbf{S}_2^{(1,1)} & \mathbf{S}_2^{(1,2)} \\ \mathbf{S}_2^{(2,1)} & \mathbf{S}_2^{(2,2)} \end{pmatrix}.$$

Notice that $\bar{\boldsymbol{x}}_l^{(i)}$ is the mean of the i-th block of the data matrix $\mathbf{X}_l$, $i = 1, ..., l$ and $l = 1, 2$. We also read $\mathbf{S}_l^{(i,j)}$ as the (i, j)-th block of $\mathbf{S}_l$ based on the data matrix $\mathbf{X}_l$, $l = 1, 2$. Similarly, let $\bar{\boldsymbol{y}}_l$ and $\mathbf{V}_l$ denote respectively the sample mean vector and the sums of squares and products matrix based on $\mathbf{Y}_l$, $l = 1, 2$ and they are also partitioned like $\bar{\boldsymbol{x}}_l$ and $\mathbf{S}_l$. That is, $\bar{\boldsymbol{y}}_l^{(i)}$ is the mean of the i-th block of data matrix $\mathbf{Y}_l$, $i = 1, ..., l$ and $l = 1, 2$, and $\mathbf{V}_l^{(i,j)}$ is the (i, j)-th block of $\mathbf{V}_l$, $i, j = 1, ..., l$ and $l = 1, 2$. Finally, we partition the parameters as follows:

$$\boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{pmatrix}, \quad \boldsymbol{\beta} = \begin{pmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \end{pmatrix}, \quad \boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{pmatrix}.$$

For likelihood ratio, it is simpler to use the following transformed parameters $\boldsymbol{\Delta}$:

$$\boldsymbol{\Delta} = \begin{pmatrix} \boldsymbol{\Delta}_{11} & \boldsymbol{\Delta}_{12} \\ \boldsymbol{\Delta}_{21} & \boldsymbol{\Delta}_{22} \end{pmatrix} = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1} & \boldsymbol{\Sigma}_{22.1} \end{pmatrix}$$

where $\boldsymbol{\Sigma}_{22.1} = \boldsymbol{\Sigma}_{22} - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12}$. Note that $\boldsymbol{\Delta}$ is in one-to-one correspondence with $\boldsymbol{\Sigma}$. We now give the MLEs of partitioned mean vectors and sub-matrices of $\boldsymbol{\Delta}$. Let

$$\mathbf{B}_{21} = \left(\mathbf{S}_2^{(2,1)} + \mathbf{V}_2^{(2,1)}\right)\left(\mathbf{S}_2^{(1,1)} + \mathbf{V}_2^{(1,1)}\right)^{-1}. \tag{2.3}$$

The MLEs are given by

$$\hat{\boldsymbol{\mu}}_1 = \bar{\boldsymbol{x}}_1, \hat{\boldsymbol{\mu}}_2 = \bar{\boldsymbol{x}}_2^{(2)} - \mathbf{B}_{21}\left(\bar{\boldsymbol{x}}_2^{(1)} - \hat{\boldsymbol{\mu}}_1\right),$$

$$\hat{\boldsymbol{\beta}}_1 = \bar{\boldsymbol{y}}_1, \hat{\boldsymbol{\beta}}_2 = \bar{\boldsymbol{y}}_2^{(2)} - \mathbf{B}_{21}\left(\bar{\boldsymbol{y}}_2^{(1)} - \hat{\boldsymbol{\beta}}_1\right),$$

$$\hat{\boldsymbol{\Delta}} = \begin{pmatrix} \hat{\boldsymbol{\Delta}}_{11} & \hat{\boldsymbol{\Delta}}_{12} \\ \hat{\boldsymbol{\Delta}}_{21} & \hat{\boldsymbol{\Delta}}_{22} \end{pmatrix} = \begin{pmatrix} (\mathbf{S}_1^{(1,1)} + \mathbf{V}_1^{(1,1)})/(N_1 + M_1) & \mathbf{B}_{21}' \\ \mathbf{B}_{21} & \widehat{\boldsymbol{\Sigma}}_{2.1} \end{pmatrix},$$

with

$$\widehat{\boldsymbol{\Sigma}}_{2.1} = \frac{1}{N_2 + M_2}\left(\left(\mathbf{S}_2^{(2,2)} + \mathbf{V}_2^{(2,2)}\right) - \mathbf{B}_{21}\left(\mathbf{S}_2^{(1,2)} + \mathbf{V}_2^{(1,2)}\right)\right).$$

2.1. Likelihood ratio test

Under H_0, the two samples are actually from the same population. Denote the pooled samples by $\mathbf{Z}$ and denote the pooled mean vector and the sum of squares and sum of products matrix as

$$\bar{\mathbf{z}}_1 = \frac{1}{(N_1 + M_1)}(N_1\bar{\mathbf{x}}_1 + M_1\bar{\mathbf{y}}_1), \ \bar{\mathbf{z}}_2 = \begin{pmatrix} \bar{\mathbf{z}}_2^{(1)} \\ \bar{\mathbf{z}}_2^{(2)} \end{pmatrix} = \frac{1}{(N_2 + M_2)}(N_2\bar{\mathbf{x}}_2 + M_2\bar{\mathbf{y}}_2),$$

$$\mathbf{T}_1 = \mathbf{S}_1 + \mathbf{V}_1 + N_1(\bar{\mathbf{x}}_1 - \bar{\mathbf{z}}_1)(\bar{\mathbf{x}}_1 - \bar{\mathbf{z}}_1)' + M_1(\bar{\mathbf{y}}_1 - \bar{\mathbf{z}}_1)(\bar{\mathbf{y}}_1 - \bar{\mathbf{z}}_1)',$$

$$\mathbf{T}_2 = \begin{pmatrix} \mathbf{T}_2^{(1,1)} & \mathbf{T}_2^{(1,2)} \\ \mathbf{T}_2^{(2,1)} & \mathbf{T}_2^{(2,2)} \end{pmatrix} = \mathbf{S}_2 + \mathbf{V}_2 + N_2(\bar{\mathbf{x}}_2 - \bar{\mathbf{z}}_2)(\bar{\mathbf{x}}_2 - \bar{\mathbf{z}}_2)' + M_2(\bar{\mathbf{y}}_2 - \bar{\mathbf{z}}_2)(\bar{\mathbf{y}}_2 - \bar{\mathbf{z}}_2)'.$$

Moreover, let $\mathbf{C}_{21} = \mathbf{T}_2^{(2,1)}\left(\mathbf{T}_2^{(1,1)}\right)^{-1}$. The MLEs under H_0 are given by

$$\widetilde{\boldsymbol{\mu}}_1 = \widetilde{\boldsymbol{\beta}}_1 = \bar{\mathbf{z}}_1, \widetilde{\boldsymbol{\mu}}_2 = \widetilde{\boldsymbol{\beta}}_2 = \bar{\mathbf{z}}_2^{(2)} - \mathbf{C}_{21}\left(\bar{\mathbf{z}}_2^{(1)} - \bar{\mathbf{z}}_1\right),$$

$$\widetilde{\boldsymbol{\Delta}} = \begin{pmatrix} \widetilde{\boldsymbol{\Delta}}_{11} & \widetilde{\boldsymbol{\Delta}}_{12} \\ \widetilde{\boldsymbol{\Delta}}_{21} & \widetilde{\boldsymbol{\Delta}}_{22} \end{pmatrix} = \begin{pmatrix} \mathbf{T}_1/(N_1+M_1) & \mathbf{C}'_{21} \\ \mathbf{C}_{21} & \mathbf{T}_{2.1} \end{pmatrix},$$

with $\mathbf{T}_{2.1} = \left(\mathbf{T}_2^{(2,2)} - \mathbf{T}_2^{(2,1)}\left(\mathbf{T}_2^{(1,1)}\right)^{-1}\mathbf{T}_2^{(1,2)}\right)/(N_2+M_2)$. Define

$$W_1^{-1} = N_1^{-1} + M_1^{-1}, W_2^{-1} = N_2^{-1} + M_2^{-1},$$

$$\hat{\boldsymbol{\mu}}_{2.1} = \hat{\boldsymbol{\mu}}_2 - B_{21}\hat{\boldsymbol{\mu}}_1, \hat{\boldsymbol{\beta}}_{2.1} = \hat{\boldsymbol{\beta}}_2 - \mathbf{B}_{21}\hat{\boldsymbol{\beta}}_1.$$

The likelihood ratio for testing $H_0 : \boldsymbol{\mu} = \boldsymbol{\beta}$ vs $H_a : \boldsymbol{\mu} \neq \boldsymbol{\beta}$ (LRT for short) is given by

$$\begin{aligned} \Lambda &= \left(\frac{\left|\hat{\boldsymbol{\Delta}}_{11}\right|}{\left|\widetilde{\boldsymbol{\Delta}}_{11}\right|}\right)^{\frac{1}{2}(N_1+M_1)} \times \left(\frac{\left|\hat{\boldsymbol{\Delta}}_{22}\right|}{\left|\widetilde{\boldsymbol{\Delta}}_{22}\right|}\right)^{\frac{1}{2}(N_2+M_2)} \\ &= \left(1 + \tfrac{Q_1}{N_1+M_1}\right)^{-\frac{1}{2}(N_1+M_1)} \times \left(1 + \tfrac{R}{N_2+M_2}\right)^{-\frac{1}{2}(N_2+M_2)} \\ &:= \Lambda_1 \times \Lambda_2. \end{aligned}$$

Here

$$\begin{aligned} Q_1 &= W_1\left(\hat{\boldsymbol{\mu}}_1 - \hat{\boldsymbol{\beta}}_1 - (\boldsymbol{\mu}_1 - \boldsymbol{\beta}_1)\right)'\left[\hat{\boldsymbol{\Sigma}}_{11}\right]^{-1}\left(\hat{\boldsymbol{\mu}}_1 - \hat{\boldsymbol{\beta}}_1 - (\boldsymbol{\mu}_1 - \boldsymbol{\beta}_1)\right), \\ R &= \frac{Q_2}{1+Q_{2d}}, \end{aligned}$$

with

$$\begin{aligned} Q_2 &= W_2\left(\hat{\boldsymbol{\mu}}_{2.1} - \hat{\boldsymbol{\beta}}_{2.1} - (\boldsymbol{\mu}_2 - \boldsymbol{\beta}_2 - \mathbf{B}_{21}(\boldsymbol{\mu}_1 - \boldsymbol{\beta}_1))\right)'\left[\hat{\boldsymbol{\Sigma}}_{2.1}\right]^{-1} \\ &\quad \times \left(\hat{\boldsymbol{\mu}}_{2.1} - \hat{\boldsymbol{\beta}}_{2.1} - (\boldsymbol{\mu}_2 - \boldsymbol{\beta}_2 - \mathbf{B}_{21}(\boldsymbol{\mu}_1 - \boldsymbol{\beta}_1))\right), \\ \hat{\boldsymbol{\mu}}_{2.1} &= \hat{\boldsymbol{\mu}}_2 - \mathbf{B}_{21}\hat{\boldsymbol{\mu}}_1, \quad \hat{\boldsymbol{\beta}}_{2.1} = \hat{\boldsymbol{\beta}}_2 - \mathbf{B}_{21}\hat{\boldsymbol{\beta}}_1, \\ Q_{2d} &= W_2\left(\bar{\boldsymbol{x}}_2^{(1)} - \bar{\boldsymbol{y}}_2^{(1)} - (\boldsymbol{\mu}_1 - \boldsymbol{\beta}_1)\right)'\left[S_2^{(1,1)} + V_2^{(1,1)}\right]^{-1} \\ &\quad \times \left(\bar{\boldsymbol{x}}_2^{(1)} - \bar{\boldsymbol{y}}_2^{(1)} - (\boldsymbol{\mu}_1 - \boldsymbol{\beta}_1)\right). \end{aligned}$$

The LRT statistics, $-2\log(\Lambda)$, is asymptotically distributed as χ^2 with p degrees of freedom. Thus, for a given level of significance level α and an observed value Λ_0 of Λ, the LRT rejects $H_0 : \boldsymbol{\mu} = \boldsymbol{\beta}$ when

$$P\left(-2\log(\Lambda) > -2\log(\Lambda)\right) = P\left(\chi^2(p) > -2\log(\Lambda)\right) < \alpha.$$

Note that the LRT for

$$H_{01} : \boldsymbol{\mu}_1 = \boldsymbol{\beta}_1 \text{ vs } \mathrm{H_{a1}} : \boldsymbol{\mu}_1 \neq \boldsymbol{\beta}_1$$

rejects H_{01} if Λ_1 is too small, and the LRT for

$$H_{02} : \boldsymbol{\mu}_2 = \boldsymbol{\beta}_2, \boldsymbol{\mu}_1 = \boldsymbol{\beta}_1 \text{ vs } \mathrm{H_{a2}} : \boldsymbol{\mu}_2 \neq \boldsymbol{\beta}_2, \boldsymbol{\mu}_1 = \boldsymbol{\beta}_1$$

rejects H_{02} for small values of Λ_2. Recall that

$$Q_1 \sim \frac{p_1(N_1+M_1)}{N_1+M_1-p_1-1}F_{p_1,N_1+M_1-p_1-1}$$

independently of

$$R \sim \frac{p_2(N_2+M_2)}{N_2+M_2-p-1}F_{p_2,N_2+M_2-p-1}.$$

The testing problem in (1.2) can be decomposed into two independent testing problems and they can be combined using union-intersection principle to get a single exact test.

2.2. Union-intersection test

The test based on union-intersection principle rejects $H_0 : \boldsymbol{\mu} = \boldsymbol{\beta}$ for large values of $\max\{Q_1, R\}$. Instead of $\max\{Q_1, R\}$, we use $M_0 := \max\{Q_1^*, R^*\}$, where

$$Q_1^* = \frac{N_1 + M_1 - 2}{N_1 + M_1} Q_1, \quad R^* = \frac{N_2 + M_2 - p_1 - 2}{N_2 + M_2} R.$$

That is, $\hat{\boldsymbol{\Sigma}}_{2.1}$ in R and $\hat{\boldsymbol{\Sigma}}_{11}$ in Q_1 are replaced by unbiased estimators of $\boldsymbol{\Sigma}_{2.1}$ and $\boldsymbol{\Sigma}_{11}$ respectively. This type of modification was suggested by [2]. Although the modification does not change the tests, it is found that the test based on $\max\{Q_1^*, R^*\}$ is better than the one based on $\max\{Q_1, R\}$. For a given level of significance level α and an observed value M_0, the union-intersection test (UIT for short) rejects H_0 if

$$P(\max\{Q_1^*, R^*\} > M_0) < \alpha$$

or equivalently

$$P\left(F_{p_1, N_1+M_1-p_1-1} \le \frac{M_0(N_1 + M_1 - p_1 - 1)}{p_1(N_1 + M_1 - 2)}\right)$$
$$\times P\left(F_{p_2, N_2+M_2-p-1} \le \frac{M_0(N_2 + M_2 - p - 1)}{p_2(N_2 + M_2 - p_1 - 2)}\right) > 1 - \alpha.$$

3. Power comparison

Since it is difficult to derive power functions of the LRT and the proposed test in Sections 2.1 and 2.2, we estimate the powers of two tests using simulations (1000000 runs). It is obvious that two tests are lower triangular invariant, hence we take $\boldsymbol{\Sigma}$ to be an identity matrix. The powers are estimated for different values of

$$\boldsymbol{\delta}_1 = (\boldsymbol{\mu}_1 - \boldsymbol{\beta}_1)' \boldsymbol{\Sigma}_{11}^{-1} (\boldsymbol{\mu}_1 - \boldsymbol{\beta}_1),$$

$$\boldsymbol{\delta}_2 = (\boldsymbol{\mu}_{2.1} - \boldsymbol{\beta}_{2.1})' \boldsymbol{\Sigma}_{2.1}^{-1} (\boldsymbol{\mu}_{2.1} - \boldsymbol{\beta}_{2.1}),$$

$$\boldsymbol{\delta}_3 = (\boldsymbol{\mu}_{3.21} - \boldsymbol{\beta}_{3.21})' \boldsymbol{\Sigma}_{3.21}^{-1} (\boldsymbol{\mu}_{3.21} - \boldsymbol{\beta}_{3.21}),$$

$$\boldsymbol{\delta}_4 = (\boldsymbol{\mu}_{4.321} - \boldsymbol{\beta}_{4.321})' \boldsymbol{\Sigma}_{4.321}^{-1} (\boldsymbol{\mu}_{4.321} - \boldsymbol{\beta}_{4.321}),$$

and presented in Tables 1-4 for monotone pattern.

We see from the tables that in most of cases, the powers of the LRT are higher than the powers of UIT. For $p = 2$, the differences between powers are small, and meanwhile the differences between sizes are also small. For higher dimension like $p = 8$, the differences between powers become bigger. For example, the biggest difference is $0.49 - 0.28 = 0.21$, However, the differences between sizes also become larger with the biggest value $0.16 - 0.05 = 0.09$. We also observe that when the true difference of the mean become larger, or the sample sizes becomes larger, the powers of the two tests become closer and closer. In addition, the percentage of missing data seems not very related to this kind change of differences between powers of two tests.

Hence, none of LRT and UIT is expected to dominate the other since they are different funtions of the same set of pivots obtained from likelihood ratio test statistics. However, if the true difference of the two mean vectors or the sample sizes are big enough, UIT is preferable since it is almost as powerful as LRT and controls the size concisely.

Table 1. Simulated powers of the LRT and UIT when $p_1 = p_2 = 1$.

(N_1, N_2, M_1, M_2)		(18,15,17,14)		(25,16,25,18)		(30,15,30,15)		(40,12,40,13)	
percentage of missing data		8.6%		16.0%		25.0%		34.4%	
δ_1	δ_2	LRT	UIT	LRT	UIT	LRT	UIT	LRT	UIT
0.0	0.0	0.07	0.05	0.07	0.05	0.07	0.05	0.07	0.05
0.0	0.2	0.20	0.17	0.21	0.19	0.20	0.17	0.18	0.15
0.0	0.5	0.40	0.36	0.44	0.42	0.40	0.37	0.35	0.33
0.0	1.0	0.66	0.63	0.74	0.72	0.69	0.67	0.61	0.58
0.0	1.3	0.78	0.76	0.84	0.82	0.79	0.78	0.71	0.70
0.0	1.5	0.84	0.82	0.89	0.88	0.85	0.84	0.78	0.76
0.0	2.0	0.93	0.92	0.96	0.96	0.94	0.93	0.88	0.88
0.1	0.0	0.15	0.11	0.17	0.14	0.19	0.16	0.25	0.20
0.3	0.0	0.30	0.26	0.41	0.37	0.48	0.43	0.59	0.53
0.5	0.0	0.46	0.41	0.61	0.58	0.69	0.66	0.82	0.78
1.1	0.0	0.79	0.76	0.92	0.92	0.96	0.95	0.99	0.99
1.4	0.0	0.88	0.86	0.97	0.97	1.00	1.00	1.00	1.00
2.1	0.0	0.97	0.97	1.00	1.00	1.00	1.00	1.00	1.00
3.1	0.0	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
4.0	0.0	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.1	0.1	0.21	0.16	0.25	0.20	0.27	0.21	0.31	0.23
0.1	0.2	0.28	0.22	0.33	0.27	0.34	0.27	0.36	0.28
0.1	0.3	0.35	0.28	0.41	0.33	0.40	0.32	0.42	0.33
0.1	0.5	0.47	0.39	0.55	0.47	0.54	0.45	0.53	0.42
0.1	0.7	0.59	0.51	0.67	0.60	0.64	0.55	0.61	0.51
0.2	0.1	0.29	0.22	0.37	0.30	0.40	0.33	0.48	0.40
0.3	0.1	0.30	0.30	0.48	0.41	0.53	0.46	0.62	0.55
0.5	0.1	0.53	0.45	0.66	0.59	0.74	0.67	0.84	0.79
0.2	0.2	0.36	0.27	0.45	0.36	0.47	0.38	0.55	0.45
0.2	0.3	0.42	0.33	0.52	0.43	0.54	0.43	0.58	0.48
0.2	0.5	0.54	0.45	0.65	0.54	0.64	0.52	0.67	0.55
0.3	0.2	0.44	0.34	0.55	0.46	0.59	0.49	0.68	0.59
0.5	0.2	0.58	0.48	0.72	0.63	0.77	0.70	0.86	0.80
0.4	0.4	0.62	0.49	0.73	0.63	0.78	0.68	0.83	0.74
0.8	0.7	0.88	0.78	0.95	0.90	0.97	0.93	0.99	0.97
1.0	1.0	0.95	0.88	0.99	0.96	0.99	0.98	1.00	1.00
1.5	1.5	0.99	0.97	1.00	1.00	1.00	1.00	1.00	1.00
2.0	2.0	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
3.5	3.4	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

Table 2. Simulated powers of the LRT and UIT when $p_1 = p_2 = 2$.

(N_1, N_2, M_1, M_2)				(18,15,17,14)		(25,16,25,18)		(30,15,30,15)		(40,12,40,13)	
percentage of missing data				8.6%		16.0%		25.0%		34.4%	
δ_1	δ_2	δ_3	δ_4	LRT	UIT	LRT	UIT	LRT	UIT	LRT	UIT
0.0	0.0	0.0	0.0	0.08	0.05	0.08	0.05	0.08	0.05	0.08	0.05
0.0	0.0	0.0	0.1	0.12	0.08	0.12	0.09	0.12	0.08	0.11	0.08
0.0	0.0	0.0	0.3	0.23	0.16	0.25	0.20	0.23	0.18	0.20	0.15
0.0	0.0	0.0	0.5	0.34	0.26	0.37	0.31	0.33	0.27	0.29	0.23
0.0	0.0	0.0	0.7	0.43	0.35	0.49	0.42	0.44	0.37	0.38	0.31
0.0	0.0	0.0	1.0	0.57	0.48	0.63	0.58	0.58	0.52	0.49	0.41
0.0	0.0	0.0	2.0	0.86	0.81	0.92	0.89	0.88	0.85	0.79	0.74
0.0	0.0	0.0	3.0	0.97	0.96	0.99	0.98	0.97	0.96	0.92	0.90
0.0	0.0	0.0	4.0	0.99	0.99	1.00	1.00	0.99	0.99	0.98	0.98
0.0	0.0	0.2	0.2	0.28	0.21	0.3	0.24	0.28	0.22	0.24	0.18
0.0	0.0	0.4	0.4	0.48	0.39	0.54	0.47	0.49	0.42	0.43	0.35
0.0	0.0	0.6	0.6	0.66	0.59	0.73	0.68	0.66	0.60	0.57	0.50
0.0	0.0	0.8	0.8	0.78	0.72	0.84	0.80	0.79	0.75	0.70	0.64
0.0	0.0	1.0	1.0	0.87	0.82	0.91	0.90	0.87	0.84	0.80	0.75
0.0	0.0	2.0	2.0	0.99	0.99	1.00	1.00	0.99	0.99	0.98	0.97
0.15	0.1	0.1	0.2	0.38	0.25	0.46	0.34	0.49	0.35	0.55	0.41
0.25	0.35	0.3	0.5	0.77	0.68	0.86	0.74	0.89	0.79	0.92	0.84
0.45	0.55	0.6	0.7	0.93	0.82	0.98	0.94	1.00	0.99	0.99	0.98
0.85	0.8	0.75	0.8	0.99	0.94	1.00	1.00	1.00	1.00	1.00	1.00
1.0	1.1	1.1	1.0	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

Table 3. Simulated powers of the LRT and UIT when $p_1 = 4, p_2 = 4$.

(N_1, N_2, M_1, M_2)	(18,15,17,14)		(25,16,25,18)		(30,15,30,15)		(40,12,40,13)	
percentage of missing data	8.6%		16.0%		25.0%		34.4%	
$(\delta_1, \delta_2, \delta_3, \delta_4, \delta_5, \delta_6, \delta_7, \delta_8)$	LRT	UIT	LRT	UIT	LRT	UIT	LRT	UIT
(0,0,0,0,0,0,0,0)	0.14	0.05	0.11	0.05	0.13	0.05	0.14	0.05
(0,0,0,0,0,0,0,0.5)	0.33	0.17	0.34	0.20	0.33	0.19	0.31	0.16
(0,0,0,0,0,0,0,1)	0.53	0.33	0.34	0.21	0.52	0.36	0.48	0.30
(0,0,0,0,0,0,0.5,0.5)	0.53	0.32	0.57	0.41	0.52	0.37	0.47	0.29
(0,0,0,0,0,0,1,1)	0.80	0.64	0.85	0.76	0.80	0.69	0.73	0.56
(0,0,0,0,0.35,0.5,0.6,0.8)	0.85	0.70	0.88	0.81	0.84	0.74	0.77	0.61
(0,0,0,0,1,1.5,1,1.2)	0.99	0.96	1.00	0.99	0.99	0.97	0.97	0.92
(0,0,0,0.5,0,0,0,0)	0.38	0.19	0.46	0.31	0.53	0.36	0.68	0.45
(0,0,0,1.0,0,0,0,0)	0.60	0.43	0.74	0.64	0.82	0.71	0.93	0.84
(1,1,1,1,0,0,0,0)	0.99	0.99	1.00	1.00	1.00	1.00	1.00	1.00
(0.45,0.6,0.35,0.7,0.3,0.3,0.4,0.6)	0.99	0.90	1.00	0.98	1.00	0.99	1.00	1.00
(0.75,1,0.68,1.2,0.7,1,0.8,1)	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
(1,1, 2,2,2,2,3,3)	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

Table 4. Simulated powers of the LRT and UIT when $p_1 = 2, p_2 = 6$.

(N_1, N_2, M_1, M_2)	(18,15,17,14)		(25,16,25,18)		(30,15,30,15)		(40,12,40,13)	
percentage of missing data	12.8%		24.0%		37.5%		51.0%	
$(\delta_1, \delta_2, \delta_3, \delta_4, \delta_5, \delta_6, \delta_7, \delta_8)$	LRT	UIT	LRT	UIT	LRT	UIT	LRT	UIT
(0,0,0,0,0,0,0,0)	0.14	0.05	0.12	0.05	0.14	0.05	0.16	0.05
(0,0,0,0,0,0,0,0.5)	0.33	0.18	0.35	0.22	0.33	0.18	0.32	0.15
(0,0,0,0,0,0,0,1)	0.33	0.20	0.58	0.42	0.54	0.36	0.49	0.27
(0,0,0,0,0,0,0.5,0.5)	0.53	0.34	0.57	0.43	0.54	0.36	0.49	0.27
(0,0,0,0,0,0,1,1)	0.80	0.64	0.85	0.75	0.80	0.68	0.74	0.53
(0,0,0,0,0.35,0.5,0.6,0.8)	0.85	0.71	0.90	0.82	0.86	0.73	0.79	0.59
(0,0,0,0,1,1.5,1,1.2)	0.99	0.97	1.00	1.00	0.99	0.98	0.97	0.91
(0,0,0,0.5,0,0,0,0)	0.34	0.18	0.35	0.21	0.33	0.18	0.33	0.15
(0,0,0,1.0,0,0,0,0)	0.53	0.33	0.58	0.42	0.54	0.37	0.49	0.28
(1,1,1,1,0,0,0,0)	1.00	0.98	0.99	0.94	0.99	0.95	1.00	0.98
(0.45,0.6,0.35,0.7,0.3,0.3,0.4,0.6)	0.99	0.82	0.99	0.88	0.99	0.85	0.99	0.82
(0.75,1,0.68,1.2,0.7,1,0.8,1)	1.00	0.99	1.00	1.00	1.00	0.99	0.99	0.99
(1,1, 2,2,2,2,3,3)	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

4. An illustrative example

Now, we use "Fishers Iris data" to illustrate the method. These data represent measurements of the sepal length and width, and petal length and width in centimeters of fifty plants for each of three types of iris: Iris setosa, Iris versicolor and Iris virginica. The data sets are posted in many websites, and we use data set IRIS in R language. For illustrative purposes, we use the data on versicolor ($\boldsymbol{x}$) and virginica($\boldsymbol{y}$). Also, we only use sepal length and width as two components.

We applied the LRT to check the equality of covariance matrices. The test produced a p-value of 0.398, and so the assumption of equality of covariance matrices are tenable.

We created monotone patterns by discarding the last 20 measurements on $\boldsymbol{x}_2$ (sepal width of virginica), the last 25 measurements on $\boldsymbol{y}_2$ (sepal width of setosa). That is, we have $p_1 = 2, p_2 = 2$, $(N_1, N_2) = (50, 30)$, and $(M_1, M_2) = (50, 25)$.

Let μ_1 and μ_2 be the average sepal length and width of versicolor respectively, $\boldsymbol{\mu}' := (\mu_1, \mu_2)$. And let β_1 and β_2 be the average sepal length and width of virginica respectively, $\boldsymbol{\beta}' := (\beta_1, \beta_2)$. We want to test

$$H_0 : \boldsymbol{\mu} = \boldsymbol{\beta} \quad \text{vs.} \quad H_a : \boldsymbol{\mu} \neq \boldsymbol{\beta}$$

After careful calculation, we get $\log T_1 = -71.602$, $\log T_2 = -25.582$, and $\Lambda = 187.532$. The critical value $Q_1^* = 31.688, R^* = 0.00581$. So, $M_0 = 31.688$. Since p-value=9.09e-7, we have sufficient evidence to reject H_0 at 95% confidence level.

5. Conclusion

In this article, we have developed an exact test for equality of two normal mean vectors based on monotone missing data with the assumption that the population covariance matrices are equal. It is easy to use and as powerful as the LRTs. Furthermore, the proposed test is useful to identify the components that caused the rejection of H_0.

Acknowledgment. The authors would like to thank the referees very much for their detailed and valuable comments improving the paper and thank NSFC (Nos. 11961015, 11861014), GXNSF (Nos. 2018GXNSFAA050031, 2018GXNSFAA281145), and a start-up grant from Guilin University of Technology for partial support.

References

[1] T.W. Anderson, *Maximum likelihood estimates for a multivariate normal distribution when some observations are missing*, J. Amer. Statist. Assoc. **52** (278), 200-203, 1957.

[2] R. Bhargava, *Multivariate tests of hypotheses with incomplete data*, PhD Thesis, Stanford University, 1962.

[3] J. Hao and K. Krishnamoorthy, *Inferences on normal covariance matrix and generalized variance with incomplete data*, J. Multivariate Anal. **78** (1), 62-82, 2001.

[4] T. Kanda and Y. Fujikoshi, *Some basic properties ofthe MLE's for a multivariate normal distribution with monotone missing data*, Amer. J. Math. Management Sci. **18** (1–2), 161-190, 1998.

[5] K. Krishnamoorthy and M. Pannala, *Some simple test procedures for normal mean vector with incomplete data*, Ann. Inst. Statist. Math. **50** (3), 531-542, 1998.

[6] K. Krishnamoorthy and M. Pannala, *Confidence estimation of normal mean vector with incomplete data*, Canad. J. Statist. **27** (2), 395-407, 1999.

[7] K. Krishnamoorthy and J. Yu, *Multivariate Behrens-Fisher problem with missing data*, J. Multivariate Anal. **105** (1), 141-150, 2012.

[8] R.J.A. Little, *A test of missing completely at random for multivariate data with missing values*, J. Amer. Statist. Assoc. **83** (404), 1198-1202, 1988.

[9] R.J.A. Little and D.B. Rubin, *Statistical Analysis with Missing Data*, Wiley, New York, 1987.

[10] G.B. Lu and J.B. Copas, *Missing at random, likelihood ignorability and model completeness*, Ann. Statist. **32** (2), 754-765, 2004.

[11] G.J. McLachlan and T. Krishnan, *The EM Alxorithm and Extensions*, 4th ed., Wiley, New York, 1997.

[12] N. Seko, T. Kawasaki and T. Seo, *Testing equality of two mean vectors with two-step monotone missing data*, Amer. J. Math. Management Sci. **31** (1), 117-135, 2011.

[13] N. Shutoh, M. Kusumi, W. Morinaga, S. Yamada and T. Seo, *Testing equality of mean vectors in two sample problems with missing data*, Comm. Statist. Simulation Comput. **39** (3), 487-500, 2010.

[14] M.S. Srivastava, *Multivariate data with missing observations*, Comm. Statist. Theory Methods **14** (4), 775-792, 1985.

[15] M.S. Srivastava and E.M. Carter, *The maximum likelihood method for non-response in sample survey*, Surv. Methodol. **12**, 61-72, 1986.

[16] A. Yagi and T. Seo, *Tests for equality of mean vectors and simultaneous confidence intervals with two-step or three-step monotone missing data patterns*, Amer. J. Math. and Management Sci. **34** (3), 213-233, 2015.

[17] J. Yu, K. Krishnamoorthy and M. Pannala, *Two-sample inference for normal mean vectors based on monotone missing data*, J. Multivariate Anal. **97** (10), 2162-2176, 2006.