Simultaneous testing of the mean vector and the covariance matrix with two-step monotone missing data

Miki Hosoya and Takashi Seo

Department of Mathematical Information Science
Tokyo University of Science
1-3, Kagurazaka, Shinjuku-ku, Tokyo 162-8601, Japan

Abstract. In this paper, we consider the problem of simultaneous testing of the mean vector and the covariance matrix when the data have a two-step monotone pattern that is missing observations. We give the likelihood ratio test (LRT) statistic and propose an approximate upper percentile of the null distribution using linear interpolation based on an asymptotic expansion of the modified LRT statistic in the case of a complete data set. As another approach, we give the modified LRT statistics with a two-step monotone missing data pattern using the coefficient of the modified LRT statistic with complete data. Finally, we investigate the asymptotic behavior of the upper percentiles of these test statistics by Monte Carlo simulation.

AMS 2010 *Mathematics Subject Classification.* 62E20, 62H10.

Key words and phrases. Asymptotic expansion, linear interpolation, modified likelihood ratio test statistic, two-step monotone missing data.

§1. Introduction

Let $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_{N_1}$ be distributed as the p-dimensional normal distribution $N_p(\boldsymbol{\mu}, \Sigma)$ and $\boldsymbol{x}_{1,N_1+1}, \boldsymbol{x}_{1,N_1+2}, \ldots, \boldsymbol{x}_{1N}$ be distributed as the p_1-dimensional normal distribution $N_{p_1}(\boldsymbol{\mu}_1, \Sigma_{11})$, where

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

We partition $\boldsymbol{x}_j$ into a $p_1 \times 1$ random vector and a $p_2 \times 1$ random vector as $\boldsymbol{x}_j = (\boldsymbol{x}'_{1j}, \boldsymbol{x}'_{2j})'$, where $\boldsymbol{x}_{ij} : p_i \times 1$, $i = 1, 2$, $j = 1, 2, \ldots, N_1$.

Such a data set has two-step monotone missing data:

$$\left(\begin{array}{c|cc|} \hline \boldsymbol{x}'_{11} & \boldsymbol{x}'_{21} & \\ \vdots & \vdots & \\ \boldsymbol{x}'_{1N_1} & \boldsymbol{x}'_{2N_1} & \\ \cline{2-3} \boldsymbol{x}'_{1,N_1+1} & * \cdots * & \\ \vdots & \vdots \quad \vdots & \\ \boldsymbol{x}'_{1N} & * \cdots * & \\ \cline{1-1} \end{array}\right),$$

where $p = p_1 + p_2, N_1 > p$, and "$*$" indicates a missing observation.

Missing data is an important problem in statistical data analyses. A variety of statistical procedures to deal with missing data have been developed by many authors, including Anderson (1957), Bhargava (1962), McLachlan and Krishnan (1997), and Little and Rubin (2002). For a general missing pattern, Srivastava (1985) discussed the LRT for mean vectors in one-sample and two-sample problems. Seo and Srivastava (2000) derived a test of equality of means and the simultaneous confidence intervals for the monotone missing data in a one-sample problem. Anderson (1957) developed an approach to derive the MLEs of the mean vector and the covariance matrix by solving the likelihood equations for monotone missing data with several missing patterns. Anderson and Olkin (1985) derived the MLEs for two-step monotone missing data in a one-sample problem. For the related discussion of the MLEs in cases of general k-step monotone missing data, see Jinadasa and Tracy (1992) and Kanda and Fujikoshi (1998).

Further, by the use of the MLEs of the mean vector and the covariance matrix, the LRT statistic and Hotelling's T^2-type statistic for tests of mean vectors with two or three-step monotone missing data has been discussed by Krishnamoorthy and Pannala (1999), Chang and Richards (2009), Seko, Yamazaki and Seo (2012), and Yagi and Seo (2014), among others. The problem of simultaneous testing of the mean and the variance under univariate and non-missing normality has been discussed by Choudhari, Kundu and Misra (2001) and Zhang, Xu and Chen (2012). For non-missing and multivariate normality, Davis (1971) gave the modified LRT statistic (see Muirhead (1982) and Srivastava (2002)). In this paper, the LRT and modified LRT statistics are given under multivariate normality with a two-step monotone missing data pattern.

The remainder of this paper is organized as follows. In Section 2, we consider the case in which the missing observations are of the two-step monotone type and provide an LRT statistic for the simultaneous testing of the mean vector and the covariance matrix. In Section 3, an approximation to the upper percentile of the LRT statistic and the modified LRT statistics are given. Finally, in Section 4, the accuracy of the approximation and the asymptotic behavior of modified statistics are investigated by Monte Carlo simulation.

§2. Likelihood ratio test statistic

In order to derive the LRT statistic of the simultaneous testing of the mean vector and the covariance matrix in the case of a two-step monotone missing data pattern, we present their MLEs, which are given by

$$\widehat{\boldsymbol{\mu}} = \begin{pmatrix} \widehat{\boldsymbol{\mu}}_1 \\ \widehat{\boldsymbol{\mu}}_2 \end{pmatrix} = \begin{pmatrix} \dfrac{1}{N}(N_1\overline{\boldsymbol{x}}_{(1)1} + N_2\overline{\boldsymbol{x}}_{(2)}) \\ \overline{\boldsymbol{x}}_{(1)2} - \widehat{\Sigma}_{21}\left\{\widehat{\Sigma}_{11}\right\}^{-1}(\overline{\boldsymbol{x}}_{(1)1} - \widehat{\boldsymbol{\mu}}_1) \end{pmatrix}, \tag{2.1}$$

$$\begin{aligned} \widehat{\Sigma} &= \begin{pmatrix} \widehat{\Sigma}_{11} & \widehat{\Sigma}_{12} \\ \widehat{\Sigma}_{21} & \widehat{\Sigma}_{22} \end{pmatrix} \\ &= \begin{pmatrix} \dfrac{1}{N}(W_{(1)11} + W_{(2)}) & \widehat{\Sigma}_{11}\left\{W_{(1)11}\right\}^{-1}W_{(1)12} \\ W_{(1)21}\left\{W_{(1)11}\right\}^{-1}\widehat{\Sigma}_{11} & \dfrac{1}{N_1}W_{(1)22\cdot 1} + \widehat{\Sigma}_{21}\left\{\widehat{\Sigma}_{11}\right\}^{-1}\widehat{\Sigma}_{12} \end{pmatrix}, \end{aligned} \tag{2.2}$$

where

$$\overline{\boldsymbol{x}}_{(1)} = \begin{pmatrix} \overline{\boldsymbol{x}}_{(1)1} \\ \overline{\boldsymbol{x}}_{(1)2} \end{pmatrix}, \quad \overline{\boldsymbol{x}}_{(1)1} = \frac{1}{N_1}\sum_{j=1}^{N_1}\boldsymbol{x}_{1j}, \quad \overline{\boldsymbol{x}}_{(1)2} = \frac{1}{N_1}\sum_{j=1}^{N_1}\boldsymbol{x}_{2j},$$

$$\overline{\boldsymbol{x}}_{(2)} = \frac{1}{N_2}\sum_{j=N_1+1}^{N}\boldsymbol{x}_{1j},$$

and

$$W_{(1)} = \begin{pmatrix} W_{(1)11} & W_{(1)12} \\ W_{(1)21} & W_{(1)22} \end{pmatrix} = \sum_{j=1}^{N_1} (\boldsymbol{x}_j - \overline{\boldsymbol{x}}_{(1)})(\boldsymbol{x}_j - \overline{\boldsymbol{x}}_{(1)})',$$

$$W_{(2)} = \sum_{j=N_1+1}^{N} (\boldsymbol{x}_{1j} - \overline{\boldsymbol{x}}_{(2)})(\boldsymbol{x}_{1j} - \overline{\boldsymbol{x}}_{(2)})' + \frac{N_1 N_2}{N} (\overline{\boldsymbol{x}}_{(1)1} - \overline{\boldsymbol{x}}_{(2)})(\overline{\boldsymbol{x}}_{(1)1} - \overline{\boldsymbol{x}}_{(2)})',$$

$$W_{(1)22\cdot 1} = W_{(1)22} - W_{(1)21} \left\{ W_{(1)11} \right\}^{-1} W_{(1)12} \ .$$

These results follow from the results in Anderson and Olkin (1985) and Kanda and Fujikoshi (1998).

In the derivation, we use the following transformed parameters $(\boldsymbol{\eta}, \Delta)$:

$$\boldsymbol{\eta} = \begin{pmatrix} \boldsymbol{\eta}_1 \\ \boldsymbol{\eta}_2 \end{pmatrix} = \begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 - \Delta_{21}\boldsymbol{\mu}_1 \end{pmatrix},$$

$$\Delta = \begin{pmatrix} \Delta_{11} & \Delta_{12} \\ \Delta_{21} & \Delta_{22} \end{pmatrix} = \begin{pmatrix} \Sigma_{11} & \Sigma_{11}^{-1}\Sigma_{12} \\ \Sigma_{21}\Sigma_{11}^{-1} & \Sigma_{22\cdot 1} \end{pmatrix},$$

where $\Sigma_{22\cdot 1} = \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}$. We note that $(\boldsymbol{\eta}, \Delta)$ are in one-to-one correspondence to $(\boldsymbol{\mu}, \Sigma)$. After multiplying the observation vector $\boldsymbol{x}_j$ by the transformation matrix

$$A = \begin{pmatrix} I_{p_1} & O \\ -\Delta_{21} & I_{p_2} \end{pmatrix}$$

on the left side, the log likelihood function is derived, and the results can then be obtained by differentiation.

We consider the following hypothesis test when the data set is of a two-step monotone pattern.

$$H_0 : \boldsymbol{\mu} = \boldsymbol{\mu}_0, \ \Sigma = \Sigma_0 \text{ vs. } H_1 : \text{not } H_0. \tag{2.3}$$

Without loss of generality, we can assume that $\boldsymbol{\mu} = \mathbf{0}$ and $\Sigma = I_p$. Then, from the MLEs in (2.1) and (2.2), we obtain the following theorem.

Theorem 2.1. *Suppose that the data have a two-step monotone pattern that is missing observations and that* λ_{m} *is the likelihood ratio (LR) in the case of the two-step monotone* *missing data. Then, the LR of the hypothesis test* (2.3) *is given by*

$$\lambda_{\rm m} = |\widehat{\Sigma}_{11}|^{\frac{N}{2}} |\widehat{\Sigma}_{22\cdot 1}|^{\frac{N_1}{2}} \frac{\operatorname{etr}\left(-\frac{1}{2}\sum_{i=1}^{N}\boldsymbol{x}_{1i}\boldsymbol{x}_{1i}'\right) \operatorname{etr}\left(-\frac{1}{2}\sum_{i=1}^{N_1}\boldsymbol{x}_{2i}\boldsymbol{x}_{2i}'\right)}{\exp\left(-\frac{1}{2}Np_1\right)\exp\left(-\frac{1}{2}N_1p_2\right)}.$$

Further, the LR can be expressed as

$$\begin{aligned}\lambda_{\rm m} &= \left(\frac{e}{N}\right)^{\frac{1}{2}Np_1} |W_{(1)11} + W_{(2)}|^{\frac{1}{2}N} \\ &\times \operatorname{etr}\left[-\frac{1}{2}\left\{W_{(1)11}+W_{(2)} + \frac{1}{N}(N_1\overline{\boldsymbol{x}}_{(1)1}+N_2\overline{\boldsymbol{x}}_{(2)})(N_1\overline{\boldsymbol{x}}_{(1)1}+N_2\overline{\boldsymbol{x}}_{(2)})'\right\}\right] \\ &\times \left(\frac{e}{N_1}\right)^{\frac{1}{2}N_1p_2} |W_{(1)22\cdot 1}|^{\frac{1}{2}N_1}\operatorname{etr}\left\{-\frac{1}{2}(W_{(1)22} + N_1\overline{\boldsymbol{x}}_{(1)2}\overline{\boldsymbol{x}}_{(1)2}')\right\}.\end{aligned}$$

The result in Theorem 2.1 coincides with the result in Hao and Krishnamoorthy (2001). We note that under H_0, $-2\log\lambda_{\rm m}$ is asymptotically distributed as a χ^2 distribution with $g = p(p+3)/2$ degrees of freedom when $N_1, N \to \infty$ with $N_1/N \to \delta \in (0,1]$. However, when the sample size is not large, the χ^2 distribution is not a good approximation to the upper percentile of $-2\log\lambda_{\rm m}$. Further, it is not easy to find the exact distribution of the LRT statistic $-2\log\lambda_{\rm m}$. In the next section, we give an approximate upper percentile of $-2\log\lambda_{\rm m}$ and propose modified LRT statistics whose upper percentile is close to that of the χ^2 distribution even for small samples.

§3. The modified LRT statistics and an approximate upper percentile of the LRT statistic

In this section, we propose an approximate upper percentile of the null distribution of $-2\log\lambda_{\rm m}$ using linear interpolation based on an asymptotic expansion of the modified LRT statistic in the case of a complete data set. Further, as another approach, we give the modified LRT statistics using the coefficient of the modified LRT statistic for the complete data.

3.1. Modified coefficient approximation procedure

We first consider the LR in the case of a complete data set. Let $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_{N,} \sim N_p(\boldsymbol{\mu}, \Sigma)$, and let λ_{c} be the LR for the complete data set. Then, the LR is given by

$$\lambda_{\text{c}} = \left(\frac{\text{e}}{N}\right)^{\frac{Np}{2}} |V|^{\frac{N}{2}} \text{etr}\left\{-\frac{1}{2}(V + N\overline{\boldsymbol{x}}\,\overline{\boldsymbol{x}}')\right\},$$

where

$$\overline{\boldsymbol{x}} = \frac{1}{N}\sum_{i=1}^{N} \boldsymbol{x}_i, \ V = \sum_{i=1}^{N}(\boldsymbol{x}_i - \overline{\boldsymbol{x}})(\boldsymbol{x}_i - \overline{\boldsymbol{x}})'.$$

Further, the modified LRT statistic is given by $-2\rho_{N,p}\log\lambda_{\text{c}}$, where $\rho_{N,p} = 1 - (2p^2 + 9p + 11)/\{6N(p+3)\}$, and its cumulative distribution function can be expanded as

$$\Pr(-2\rho_{N,p}\log\lambda_{\text{c}} \le x) = G_g(x) + \frac{\gamma}{M^2}\{G_{g+4}(x) - G_g(x)\} + O(M^{-3}), \tag{3.1}$$

where

$$M = \rho_{N,p}N, \quad \gamma = \frac{p}{288(p+3)}(2p^4 + 18p^3 + 49p^2 + 36p - 13),$$

and $G_g(x)$ and $G_{g+4}(x)$ are the cumulative distribution functions of the χ^2 distribution with $g(= p(p+3)/2)$ and $g+4$ degrees of freedoms, respectively.

This result was derived by Davis (1971) (see Muirhead (1982) and Srivastava (2002)). This means that if the χ^2 distribution is used as an approximation to the distribution of $-2\rho_{N,p}\log\lambda_{\text{c}}$, the error involved is not of order M^{-1} but of order M^{-2}.

If we denote the coefficients of the modified LRT statistics in the case of complete data sets N and N_1 by $\rho_{N,p}$ and $\rho_{N_1,p}$, respectively, then it may be noted that ρ_{miss} is between $\rho_{N,p}$ and $\rho_{N_1,p}$, where ρ_{miss} is the coefficient of the modified LRT statistic $-2\rho_{\text{miss}}\log\lambda_{\text{m}}$. From the linear interpolation, we propose an approximation to the modified LRT statistic in the case of two-step monotone missing data. Calculating the approximate coefficient $\rho^*_{\text{miss}} = (p_1\rho_{N,p} + p_2\rho_{N_1,p})/p$, we can obtain an approximate modified LRT statistic $-2\rho^*_{\text{miss}}\log\lambda_{\text{m}}$, where

$$\rho^*_{\text{miss}} = 1 - \frac{1}{N}\left(1 + \frac{N_2p_2}{N_1p}\right)\frac{2p^2 + 9p + 11}{6(p+3)}.$$

3.2. Asymptotic expansion approximation procedure

In this subsection, we give an approximate upper percentile of $-2\log\lambda_{\mathrm{m}}$ when the data have a two-step monotone pattern that is missing observations. First, in the case of a complete data set, we obtain the following lemma.

Lemma 3.1. *Suppose that $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_N$ are distributed as $N_p(\boldsymbol{\mu}, \Sigma)$. Then, under the null hypothesis H_0 in* (2.3)*, the upper percentile of the modified LRT statistic,$-2\rho_{N,p}\log\lambda_{\mathrm{c}}$, can be expanded as*

$$q_{\mathrm{MLR\cdot c}}(\alpha) = \chi^2_g(\alpha) + \frac{1}{M^2}\frac{2\gamma}{g(g+2)}\chi^2_g(\alpha)\left\{\chi^2_g(\alpha) + g + 2\right\} + \mathrm{o}(M^{-2}),$$

where

$$M = \rho_{N,p}N,\ \ \rho_{N,p} = 1 - \frac{2p^2+9p+11}{6N(p+3)},\ \ g = \frac{1}{2}p(p+3),$$

and $\chi^2_g(\alpha)$ is the upper percentile of the χ^2 distribution with g degrees of freedom.

Proof. Putting the upper percentile of $-2\rho_{N,p}\log\lambda_{\mathrm{c}}$ with

$$q_{\mathrm{MLR\cdot c}}(\alpha) = \chi^2_g(\alpha) + \frac{1}{M^2}h + \mathrm{o}(M^{-2}),$$

where h is a constant, we have

$$1-\alpha = G_g(q_{\mathrm{MLR\cdot c}}(\alpha)) - g_g(\chi^2_g(\alpha))\frac{1}{M^2}h + \mathrm{o}(M^{-2}), \tag{3.2}$$

where $G_g(x)$ and $g_g(x)$ are, respectively, the cumulative distribution function and the density function of the χ^2 distribution with g degrees of freedom. On the other hand, from (3.1), we can write

$$\begin{aligned} 1-\alpha &= \Pr\left\{-2\rho_{N,p}\log\lambda_{\mathrm{c}} \le q_{\mathrm{MLR\cdot c}}(\alpha)\right\} \\ &= G_g(q_{\mathrm{MLR\cdot c}}(\alpha)) + \frac{\gamma}{M^2}\left\{G_{g+4}(q_{\mathrm{MLR\cdot c}}(\alpha)) - G_g(q_{\mathrm{MLR\cdot c}}(\alpha))\right\} \\ &\quad + \mathrm{o}(M^{-2}). \end{aligned} \tag{3.3}$$

Therefore, using $G_{g+2j}(x) = -2g_{g+2j}(x) + G_{g+2(j-1)}(x)$, $j = 0, 1, 2$ and comparing (3.2) with (3.3), we obtain

$$h = \frac{2\gamma}{g(g+2)}\chi^2_g(\alpha)\left\{\chi^2_g(\alpha) + g + 2\right\} + \mathrm{o}(M^{-2}).$$

□

From Lemma 3.1 and $M^{-2} = N^{-2} + \mathrm{O}(N^{-3})$, we can expand the upper percentile of $-2\log\lambda_{\mathrm{c}}$ as

$$q_{\mathrm{LR\cdot c}}(\alpha) = \chi_g^2(\alpha) + \frac{\nu}{N}\chi_g^2(\alpha) + \frac{1}{N^2}\chi_g^2(\alpha)\left\{\nu^2 + \frac{2\gamma}{g} + \frac{2\gamma}{g(g+2)}\chi_g^2(\alpha)\right\} + \mathrm{o}(N^{-2}),$$

where

$$\nu = \frac{2p^2 + 9p + 11}{6(p+3)}.$$

From the linear interpolation, letting $q_{\mathrm{LR\cdot m}}(\alpha)$ be the upper percentile of $-2\log\lambda_{\mathrm{m}}$, an approximate upper percentile of $-2\log\lambda_{\mathrm{m}}$ can be obtained as

$$\begin{aligned} q^*_{\mathrm{LR\cdot m}}(\alpha) &= \chi_g^2(\alpha) + \frac{1}{N}\left(p_1 + \frac{1}{c_1}p_2\right)\frac{\nu}{p}\chi_g^2(\alpha) \\ &\quad + \frac{1}{N^2}\left(p_1 + \frac{1}{c_1^2}p_2\right)\frac{\chi_g^2(\alpha)}{p}\left\{\nu^2 + \frac{2\gamma}{g} + \frac{2\gamma}{g(g+2)}\chi_g^2(\alpha)\right\} + \mathrm{o}(N^{-2}), \end{aligned}$$

where $c_1 = N_1/N$.

3.3. The LRT statistic's decomposition procedure

In this section, we give other modified LRT statistics by the decomposition of λ_{m}. We first consider the following test problem for Σ.

$$H_{01} : \Sigma = \Sigma_0 = I \text{ vs. } H_{11} : \Sigma \neq \Sigma_{01}.$$

Hao and Krishnamoorthy (2001) derived the modified LRT statistic $\lambda_{\mathrm{m\cdot\Sigma}}$ in the case of two-step monotone missing data, which is given by

$$\begin{aligned} \lambda_{\mathrm{m\cdot\Sigma}} = &\left(\frac{\mathrm{e}}{n}\right)^{\frac{1}{2}np_1} \left|W_{(1)11} + W_{(2)}\right|^{\frac{1}{2}n} \exp\left\{-\frac{1}{2}\mathrm{tr}(W_{(1)11} + W_{(2)})\right\} \\ &\times \left(\frac{\mathrm{e}}{n_1}\right)^{\frac{1}{2}n_1p_2} \left|W_{(1)22\cdot1}\right|^{\frac{1}{2}n_1} \exp\left\{-\frac{1}{2}\mathrm{tr}W_{(1)22\cdot1}\right\} \\ &\times \exp\left\{-\frac{1}{2}\mathrm{tr}(W_{(1)22} - W_{(1)22\cdot1})\right\}, \end{aligned}$$

where $n = N-1$, $n_1 = N_1-p_1-1$. We note that the modified LRT statistic $-2\log\lambda_{\mathrm{m}\cdot\Sigma}$ is an unbiased test statistic (see Hao and Krishnamoorthy (2001) and Chang and Richards (2010)). Further, after modifying and rearranging some terms, they expressed the modified LR for H_0 in (2.3) as $\lambda_{\mathrm{m}\cdot\Sigma}\lambda_1\lambda_2$, where

$$\lambda_1 = \exp\left\{-\frac{1}{2N}(N_1\overline{\boldsymbol{x}}_{(1)1} + N_2\overline{\boldsymbol{x}}_{(2)})'(N_1\overline{\boldsymbol{x}}_{(1)1} + N_2\overline{\boldsymbol{x}}_{(2)})\right\},$$
$$\lambda_2 = \exp\left\{-\frac{1}{2}N_1\overline{\boldsymbol{x}}'_{(1)2}\overline{\boldsymbol{x}}_{(1)2}\right\}.$$

If we denote

$$\lambda_3 = \left(\frac{\mathrm{e}}{N}\right)^{\frac{1}{2}Np_1}|W_{(1)11}+W_{(2)}|^{\frac{1}{2}N}\exp\left\{-\frac{1}{2}\mathrm{tr}(W_{(1)11}+W_{(2)})\right\},$$
$$\lambda_4 = \left(\frac{\mathrm{e}}{N_1}\right)^{\frac{1}{2}N_1p_2}|W_{(1)22\cdot1}|^{\frac{1}{2}N_1}\exp\left\{-\frac{1}{2}\mathrm{tr}(W_{(1)22\cdot1})\right\},$$
$$\lambda_5 = \exp\left\{-\frac{1}{2}\mathrm{tr}(W_{(1)21}W_{(1)11}^{-1}W_{(1)12})\right\},$$

we can express $\lambda_{\mathrm{m}} = \prod_{i=1}^{5}\lambda_i$. Since $\lambda_1\lambda_3$ and $\lambda_2\lambda_4$ are of the form of LR for H_0 under non-missing normality, we can give the modified LRT statistics, $-2\rho_{N,p_1}\log\lambda_1\lambda_3$ and $-2\rho_{N_1,p_2}\log\lambda_2\lambda_4$, respectively, where

$$\rho_{N,p_1} = 1-\frac{2p_1^2+9p_1+11}{6N(p_1+3)}\ ,\quad \rho_{N_1,p_2} = 1-\frac{2p_2^2+9p_2+11}{6N_1(p_2+3)}\ .$$

Thus, we propose a new modified LRT statistic given by $-2\log\tau$, where

$$\tau = (\lambda_1\lambda_3)^{\rho_{N,p_1}}(\lambda_2\lambda_4)^{\rho_{N_1,p_2}}\lambda_5\ .$$

In addition, we denote

$$\lambda_3^* = \left(\frac{\mathrm{e}}{n}\right)^{\frac{1}{2}np_1}|W_{(1)11}+W_{(2)}|^{\frac{1}{2}n}\exp\left\{-\frac{1}{2}\mathrm{tr}(W_{(1)11}+W_{(2)})\right\},$$
$$\lambda_4^* = \left(\frac{\mathrm{e}}{n_1}\right)^{\frac{1}{2}n_1p_2}|W_{(1)22\cdot1}|^{\frac{1}{2}n_1}\exp\left\{-\frac{1}{2}\mathrm{tr}(W_{(1)22\cdot1})\right\}.$$

Then, we can propose the modified LRT statistic $-2\log\tau^*$, where

$$\tau^* = \lambda_1\lambda_2(\lambda_3^*)^{\rho_{n,p_1}}(\lambda_4^*)^{\rho_{n_1,p_2}}\lambda_5$$

and

$$\rho_{n,p_1} = 1 - \frac{2p_1^2 + 3p_1 - 1}{6n(p_1+1)}, \quad \rho_{n_1,p_2} = 1 - \frac{2p_2^2 + 3p_2 - 1}{6n_1(p_2+1)}.$$

§4. Simulation studies

We evaluate the accuracy and the asymptotic behaviors of the χ^2 approximations by Monte Carlo simulation (10^6 runs).

In Table 1, we provide the simulated upper 100α percentiles of $-2\log\lambda_{\mathrm{m}}$ and $-2\rho^*_{\mathrm{miss}}\log\lambda_{\mathrm{m}}$ and the approximate upper percentiles of $-2\log\lambda_{\mathrm{m}}$, that is, $q^*_{\mathrm{LR\cdot m}}(\alpha)$ for $(p_1, p_2) = (8, 4)$; $\alpha = 0.05,\ 0.01$; and for the following three cases of (N_1, N_2),

$$(N_1, N_2) = \begin{cases} (m, m), & m = 20, 40, 80, 160, 320, \\ (2m, m), & m = 10, 20, 40, 80, 160, \\ (m, 2m), & m = 20, 40, 80, 160. \end{cases}$$

In Table 2, we provide the same upper percentiles as those given in Table 1 for $(p_1, p_2) = (8, 4)$; $\alpha = 0.05,\ 0.01$; $(N_1, N_2) = (m_1, m_2), m_1 = 40, 80, 160, 320,\ m_2 = 10, 30, 60, 120,$ where the sets of (N_1, N_2) are combinations of m_1 and m_2.

It may be noted from Tables 1 and 2 that the simulated values are closer to the upper percentile of the χ^2 distribution when the sample size becomes large. In addition, it can be seen from both tables that the upper percentile of $-2\rho^*_{\mathrm{miss}}\log\lambda_{\mathrm{m}}$ is considerably better than that of $-2\log\lambda_{\mathrm{m}}$ even for small sample sizes. Further, Tables 1 and 2 list the simulated coverage probabilities for the upper percentiles of $-2\log\lambda_{\mathrm{m}}$ and $-2\rho^*_{\mathrm{miss}}\log\lambda_{\mathrm{m}}$ as well as $q^*_{\mathrm{LR\cdot m}}(\alpha)$, which are given by

$$\mathrm{CP}_{\mathrm{m}}(\chi^2) = 1 - \Pr\left\{-2\log\lambda_{\mathrm{m}} > \chi^2_g(\alpha)\right\},$$
$$\mathrm{CP}^*_{\mathrm{m}}(\chi^2) = 1 - \Pr\left\{-2\rho^*_{\mathrm{miss}}\log\lambda_{\mathrm{m}} > \chi^2_g(\alpha)\right\},$$

and

$$\mathrm{CP}(q^*_{\mathrm{LR\cdot m}}) = 1 - \Pr\left\{-2\log\lambda_{\mathrm{m}} > q^*_{\mathrm{LR\cdot m}}(\alpha)\right\},$$

respectively. It appears from the simulated results that the approximate value $q^*_{\text{LR}\cdot\text{m}}(\alpha)$ based on the asymptotic expansion is good for all cases, even when $N_1 < N_2$. Therefore, it can be concluded that our approximation procedures are very accurate for most of the cases.

In Tables 3 and 4, we provide the simulated upper percentiles of $-2\log\tau$ and $-2\log\tau^*$ for the same cases as those in Tables 1 and 2. It may also be noted that the upper percentiles of $-2\log\tau^*$ are considerably good even for small sample sizes. Tables 3 and 4 list the simulated coverage probabilities for the upper percentiles of $-2\log\tau$ and $-2\log\tau^*$, which are given by

$$\text{CP}_\tau(\chi^2) = 1 - \Pr\left\{-2\log\tau > \chi^2_g(\alpha)\right\}$$

and

$$\text{CP}_{\tau^*}(\chi^2) = 1 - \Pr\left\{-2\log\tau^* > \chi^2_g(\alpha)\right\},$$

respectively. The results for coverage probabilities also show that our modified LRT statistic $-2\log\tau^*$ yields considerably good χ^2 approximations for cases in which the sample size is small.

In conclusion, we have developed the approximate upper percentiles of the LRT statistic $-2\log\lambda_{\text{m}}$ and some modified LRT statistics for simultaneous testing of the mean vector and the covariance matrix for the case of two-step monotone missing data. The null distribution of the modified LRT statistic $-2\log\tau^*$ proposed in this paper has considerably good approximation to the χ^2 distribution even when the sample size is small.

Table 1: The simulated values for $-2\log\lambda_{\mathrm{m}}$ and $-2\rho^*_{\mathrm{miss}}\log\lambda_{\mathrm{m}}$, and the approximate value for $-2\log\lambda_{\mathrm{m}}$, and the simulated coverage probabilities when $(p_1, p_2) = (8, 4)$

Sample Size		Upper Percentile			Coverage Probability		
N_1	N_2	$-2\log\lambda_{\mathrm{m}}$	$-2\rho^*_{\mathrm{miss}}\log\lambda_{\mathrm{m}}$	$q^*_{\mathrm{LR\cdot m}}(\alpha)$	$\mathrm{CP_m}(\chi^2)$	$\mathrm{CP^*_m}(\chi^2)$	$\mathrm{CP}(q^*_{\mathrm{LR\cdot m}})$
			$\alpha = 0.05$				
20	20	149.97	127.36	134.65	0.406	0.798	**0.818**
40	40	126.88	117.32	122.79	0.798	0.917	**0.920**
80	80	119.32	114.82	117.69	0.897	0.938	**0.939**
160	160	116.12	113.94	115.35	0.928	**0.945**	**0.945**
320	320	114.59	113.51	114.23	0.940	0.947	**0.948**
20	10	152.74	125.88	138.65	0.359	0.819	**0.836**
40	20	128.09	116.83	124.50	0.779	0.921	**0.925**
80	40	119.79	114.52	118.47	0.891	0.940	**0.941**
160	80	116.39	113.83	115.72	0.925	**0.945**	**0.945**
320	160	114.73	113.47	114.41	0.939	**0.948**	**0.948**
20	40	147.56	129.02	130.99	0.449	0.773	**0.797**
40	80	125.81	117.91	121.16	0.815	0.912	**0.916**
80	160	118.85	115.11	116.93	0.902	**0.936**	**0.936**
160	320	115.87	114.05	114.98	0.930	**0.944**	**0.944**
			$\alpha = 0.01$				
20	20	165.26	140.35	147.80	0.642	0.929	**0.939**
40	40	139.32	128.82	134.71	0.931	0.980	**0.981**
80	80	130.99	126.05	129.11	0.973	**0.987**	**0.987**
160	160	127.35	124.95	126.53	0.983	**0.989**	**0.989**
320	320	125.71	124.52	125.31	0.987	**0.989**	**0.989**
20	10	168.20	138.62	152.21	0.594	0.939	**0.947**
40	20	140.52	128.16	136.60	0.922	0.982	**0.983**
80	40	131.41	125.63	129.97	0.971	0.987	**0.988**
160	80	127.74	124.93	126.94	0.983	**0.989**	**0.989**
320	160	125.95	124.56	125.51	0.987	**0.989**	**0.989**
20	40	162.72	142.28	143.76	0.682	0.916	**0.928**
40	80	138.16	129.48	132.93	0.939	0.978	**0.979**
80	160	130.39	126.30	128.27	0.975	**0.986**	**0.986**
160	320	127.10	125.10	126.13	0.984	**0.988**	**0.988**

Note. The closest to $1-\alpha$ in coverage probabilities $\mathrm{CP_m}(\chi^2)$, $\mathrm{CP^*_m}(\chi^2)$, and $\mathrm{CP}(q^*_{\mathrm{LR\cdot m}})$ of each low is in bold. $\chi^2_g(0.05) = 113.145$, $\chi^2_g(0.01) = 124.116$.

Table 2: The simulated values for $-2\log\lambda_{\mathrm{m}}$ and $-2\rho^*_{\mathrm{miss}}\log\lambda_{\mathrm{m}}$, and the approximate value for $-2\log\lambda_{\mathrm{m}}$, and the simulated coverage probabilities when $(p_1, p_2) = (8, 4)$

Sample Size		Upper Percentile			Coverage Probability		
N_1	N_2	$-2\log\lambda_{\mathrm{m}}$	$-2\rho^*_{\mathrm{miss}}\log\lambda_{\mathrm{m}}$	$q^*_{\mathrm{LR\cdot m}}(\alpha)$	$\mathrm{CP_m}(\chi^2)$	$\mathrm{CP^*_m}(\chi^2)$	$\mathrm{CP}(q^*_{\mathrm{LR\cdot m}})$
				$\alpha = 0.05$			
40	10	129.12	116.46	125.92	0.763	0.925	**0.929**
80	10	120.61	114.30	119.55	0.883	0.942	**0.943**
160	10	116.78	113.61	116.35	0.922	**0.947**	**0.947**
320	10	115.01	113.41	114.75	0.937	**0.948**	**0.948**
40	30	127.40	117.11	123.51	0.790	0.919	**0.923**
80	30	120.07	114.52	118.76	0.889	0.940	**0.941**
160	30	116.65	113.70	116.12	0.923	0.946	**0.947**
320	30	114.88	113.35	114.68	0.937	**0.949**	**0.949**
40	60	126.21	117.65	121.80	0.808	0.914	**0.918**
80	60	119.57	114.74	118.02	0.894	0.938	**0.939**
160	60	116.48	113.79	115.86	0.924	**0.946**	**0.946**
320	60	114.86	113.41	114.60	0.938	**0.948**	**0.948**
40	120	125.31	118.22	120.38	0.822	0.908	**0.913**
80	120	119.07	115.03	117.23	0.900	0.936	**0.937**
160	120	116.28	113.94	115.51	0.926	**0.945**	**0.945**
320	120	114.69	113.37	114.48	0.939	0.948	**0.949**
				$\alpha = 0.01$			
40	10	141.67	127.79	138.17	0.914	0.982	**0.984**
80	10	132.28	125.36	131.15	0.968	**0.988**	**0.988**
160	10	128.08	124.60	127.63	0.982	**0.989**	**0.989**
320	10	126.09	124.34	125.87	0.986	**0.990**	**0.990**
40	30	139.93	128.63	135.51	0.928	0.980	**0.982**
80	30	131.67	125.58	130.28	0.970	0.987	**0.988**
160	30	127.85	124.62	127.38	0.982	**0.989**	**0.989**
320	30	126.12	124.44	125.81	0.986	**0.989**	**0.989**
40	60	138.51	129.11	133.63	0.936	0.979	**0.980**
80	60	131.18	125.88	129.47	0.972	**0.987**	**0.987**
160	60	127.63	124.68	127.09	0.983	**0.989**	**0.989**
320	60	126.04	124.45	125.72	0.986	**0.989**	**0.989**
40	120	137.51	129.74	132.07	0.943	0.977	**0.979**
80	120	130.65	126.22	128.60	0.974	**0.986**	**0.986**
160	120	127.40	124.83	126.71	0.983	**0.989**	**0.989**
320	120	125.74	124.29	125.58	0.987	**0.990**	**0.990**

Note. The closest to $1-\alpha$ in coverage probabilities $\mathrm{CP_m}(\chi^2)$, $\mathrm{CP^*_m}(\chi^2)$, and $\mathrm{CP}(q^*_{\mathrm{LR\cdot m}})$ of each low is in bold. $\chi^2_g(0.05) = 113.145$, $\chi^2_g(0.01) = 124.116$.

Table 3: The simulated values for $-2\log\tau$ and $-2\log\tau^*$, and the simulated coverage probabilities when $(p_1, p_2) = (8, 4)$

Sample Size		Upper Percentile		Coverage Probability	
N_1	N_2	$-2\log\tau$	$-2\log\tau^*$	$\mathrm{CP}_\tau(\chi^2)$	$\mathrm{CP}_{\tau^*}(\chi^2)$
		$\alpha = 0.05$			
20	20	140.32	114.86	0.572	**0.938**
40	40	123.20	113.82	0.850	**0.945**
80	80	117.59	113.43	0.914	**0.948**
160	160	115.34	113.31	0.934	**0.949**
320	320	114.17	113.20	0.943	**0.950**
20	10	141.20	115.61	0.556	**0.932**
40	20	123.55	114.11	0.846	**0.943**
80	40	117.81	113.60	0.912	**0.947**
160	80	115.37	113.37	0.934	**0.949**
320	160	114.29	113.30	0.942	**0.949**
20	40	139.63	114.28	0.585	**0.942**
40	80	122.95	113.64	0.854	**0.947**
80	160	117.53	113.35	0.915	**0.949**
160	320	115.19	113.22	0.935	**0.950**
		$\alpha = 0.01$			
20	20	154.54	126.08	0.788	**0.987**
40	40	135.06	124.78	0.955	**0.989**
80	80	129.00	124.41	0.979	**0.990**
160	160	126.44	124.28	0.986	**0.990**
320	320	125.23	124.20	0.988	**0.990**
20	10	155.31	126.80	0.776	**0.985**
40	20	135.48	125.19	0.953	**0.988**
80	40	129.25	124.64	0.978	**0.989**
160	80	126.51	124.33	0.985	**0.990**
320	160	125.38	124.32	0.988	**0.990**
20	40	153.67	125.34	0.798	**0.988**
40	80	135.00	124.69	0.956	**0.989**
80	160	128.91	124.36	0.979	**0.990**
160	320	126.50	124.30	0.986	**0.990**

Note. The closer to $1-\alpha$ in coverage probabilities $\mathrm{CP}_\tau(\chi^2)$ and $\mathrm{CP}_{\tau^*}(\chi^2)$ of each low is in bold.
$\chi^2_g(0.05) = 113.145$, $\chi^2_g(0.01) = 124.116$.

Table 4: The simulated values for $-2\log\tau$ and $-2\log\tau^*$, and the simulated coverage probabilities when $(p_1, p_2) = (8, 4)$

Sample Size		Upper Percentile		Coverage Probability	
N_1	N_2	$-2\log\tau$	$-2\log\tau^*$	$\mathrm{CP}_\tau(\chi^2)$	$\mathrm{CP}_{\tau^*}(\chi^2)$
		$\alpha = 0.05$			
40	10	123.78	114.33	0.843	**0.942**
80	10	118.00	113.71	0.911	**0.946**
160	10	115.51	113.49	0.933	**0.948**
320	10	114.31	113.28	0.942	**0.949**
40	30	123.39	113.97	0.848	**0.944**
80	30	117.90	113.68	0.912	**0.946**
160	30	115.38	113.39	0.934	**0.948**
320	30	114.33	113.29	0.942	**0.949**
40	60	123.05	113.74	0.853	**0.946**
80	60	117.73	113.54	0.913	**0.947**
160	60	115.41	113.39	0.933	**0.948**
320	60	114.29	113.29	0.942	**0.949**
40	120	122.78	113.46	0.856	**0.948**
80	120	117.62	113.39	0.914	**0.948**
160	120	115.24	113.25	0.935	**0.949**
320	120	114.17	113.18	0.943	**0.950**
		$\alpha = 0.01$			
40	10	135.92	125.47	0.952	**0.988**
80	10	129.33	124.76	0.978	**0.989**
160	10	126.68	124.45	0.985	**0.989**
320	10	125.34	124.18	0.988	**0.990**
40	30	135.30	125.05	0.954	**0.988**
80	30	129.35	124.62	0.978	**0.989**
160	30	126.71	124.42	0.985	**0.990**
320	30	125.33	124.28	0.988	**0.990**
40	60	135.00	124.70	0.956	**0.989**
80	60	129.15	124.55	0.979	**0.989**
160	60	126.61	124.43	0.985	**0.989**
320	60	125.42	124.35	0.988	**0.990**
40	120	134.60	124.49	0.958	**0.989**
80	120	129.06	124.45	0.979	**0.989**
160	120	126.52	124.24	0.986	**0.990**
320	120	125.27	124.16	0.988	**0.990**

Note. The closer to $1-\alpha$ in coverage probabilities $\mathrm{CP}_\tau(\chi^2)$ and $\mathrm{CP}_{\tau^*}(\chi^2)$ of each low is in bold.
$\chi^2_g(0.05) = 113.145$, $\chi^2_g(0.01) = 124.116$.

References

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

[2] Anderson, T. W. and Olkin, I. (1985). Maximum-likelihood estimation of the parameters of a multivariate normal distribution, *Linear Algebra Appl.*, **70**, 147–171.

[3] Bhargava, R. (1962). Multivariate tests of hypotheses with incomplete data, *Technical Report No.3, Applied Mathematics and Statistics Laboratories, Stanford University*, Stanford, California.

[4] Chang, W. -Y. and Richards, D. St. P. (2009). Finite-sample inference with monotone incomplete multivariate normal data, I, *J. Multivariate Anal.*, **100**, 1883–1899.

[5] Chang, W. -Y. and Richards, D. St. P. (2010). Finite-sample inference with monotone incomplete multivariate normal data, II, *J. Multivariate Anal.*, **101**, 603–620.

[6] Choudhari, P., Kundu, D. and Misra, N. (2001). Likelihood ratio test for simultaneous testing of the mean and the variance of a normal distribution, *J. Statist. Comuput. Simul.*, **71**, 313–333.

[7] Davis, A. W. (1971). Percentile approximations for a class of likelihood ratio criteria, *Biometrika*, **58**, 349–356.

[8] Hao, J. and Krishnamoorthy, K. (2001). Inferences on a normal covariance matrix and generalized variance with monotone missing data, *J. Multivariate Anal.*, **78**, 62–82.

[9] Jinadasa, K. G. and Tracy, D. S. (1992). Maximum likelihood estimation for multivariate normal distribution with monotone sample, *Comm. Statist. Theory Methods*, **21**, 41–50.

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

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

[12] Little, R. J. and Rubin, D. B. (2002). *Statistical Analysis with Missing Data, 2nd ed.*, Wiley-Interscience (John Wiley & Sons), Hoboken, NJ.

[13] McLachlan, J. G. and Krishnan, T. (1997). *The EM Algorithm and Extensions*, John Wiley & Sons, Inc., New York.

[14] Muirhead, R. J. (1982). *Aspects of Multivariate Statistical Theory*, John Wiley & Sons, Inc., New York.

[15] Seo, T. and Srivastava, M. S. (2000). Testing equality of means and simultaneous confidence intervals in repeated measures with missing data, *Biometrical J.*, **42**, 981–993.

[16] Seko, N., Yamazaki, A. and Seo, T. (2012). Tests for mean vector with two-step monotone missing data, *SUT J. Math.*, **48**, 13–36.

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

[18] Srivastava, M. S. (2002). *Methods of Multivariate Statistics*, Wiley-Interscience (John Wiley & Sons), New York.

[19] Yagi, A. and Seo, T. (2014). A test for mean vector and simultaneous confidence intervals with three-step monotone missing data, to appear in *Amer. J. Math. Management Sci.*, **33**.

[20] Zhang, L., Xu, X. and Chen, G. (2012). The exact likelihood ratio test for equality of two normal populations, *Amer. Statist.*, **66**, 180–184.