A Modified Likelihood Ratio Test for a Mean Vector with Monotone Missing Data

Ayaka Yagi*, Takashi Seo* and Muni Srivastava**

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

** *Department of Statistical Sciences*
University of Toronto
100 St. George Street, Toronto, ON, M5S 3G3, Canada

Abstract

In this study, we consider the likelihood ratio test (LRT) for a normal mean vector when the data have a monotone pattern of missing observations. We derive the modified likelihood ratio test (MLRT) statistics by using the decomposition of the likelihood ratio (LR). Further, we investigate the accuracy of the upper percentiles of these test statistics by Monte Carlo simulation.

Key Words and Phrases: Asymptotic expansion, Maximum likelihood estimator, Monte Carlo simulation.

1 Introduction

In statistical data analyses, testing a problem with missing data is an important problem. In this study we consider the one-sample test for a normal mean vector with monotone missing data. For the one-sample problem with k-step monotone missing data, the closed-form expressions for the MLEs of the mean vector and covariance matrix were given by Jinadasa and Tracy (1992). Kanda and Fujikoshi (1998) discussed the properties of the MLEs in the case of k-step monotone missing data using the conditional approach. The one-sample problem of the test for the mean vector with monotone missing data has been discussed by many authors. For discussions related to Hotelling's T^2-type statistic, see Krishnamoorthy and Pannala (1999); Chang and Richards (2009); Seko, Yamazaki and Seo (2012); and Yagi and Seo (2014), among others. For a discussion of the LRT statistic, see Krishnamoorthy and Pannala (1998) and Seko et al. (2012). For the two-sample problem, Seko, Kawasaki and Seo (2011) derived Hotelling's T^2-type statistic, the LRT statistic, and their approximate upper percentiles with two-step monotone missing data. In addition, Yu, Krishnamoorthy and Pannala (2006) derived an approximate distribution for Hotelling's T^2-type statistic using another approach. Recently, Yagi and Seo (2015a, 2015b) gave the approximate upper percentiles of the simplified Hotelling's T^2-type statistics for testing the equality of mean vectors and simultaneous confidence intervals with general k-step monotone missing data including the two- or three-step case.

In this paper, for the one-sample test, we give the LRT statistic for general monotone missing data and propose MLRT statistics by using the decomposition of the LR. This decomposition follows from Bhargava (1962) and Krishnamoorthy and Pannala (1998). This paper is organized in the following way. In Section 2, we first present the assumptions, notation and preliminaries. In Section 3, we derive the LRT statistic and MLRT statistics, which converges to the χ^2 distribution faster than the LRT statistic when the sample size is large. In Section 4, some simulation results for three- and five-step monotone missing data cases are presented to investigate the accuracy of the upper percentiles of the null distribution of the MLRT statistics.

2 Assumptions, notation and preliminaries

We consider the one-sample problem of testing for a mean vector with a k-step monotone missing data pattern. Let $\boldsymbol{x}_i$ be a $p_i \times 1$ normal random vector with the mean vector $\boldsymbol{\mu}_i$ and covariance matrix $\boldsymbol{\Sigma}_i$, where $\boldsymbol{\mu}_i = (\boldsymbol{\mu})_i = (\mu_1, \mu_2, \ldots, \mu_{p_i})'$, and $\boldsymbol{\Sigma}_i$ is the $p_i \times p_i$ principal submatrix of $\boldsymbol{\Sigma}(= \boldsymbol{\Sigma}_1)$ with $p = p_1 > p_2 > \cdots > p_k > 0$. Suppose that $\boldsymbol{x}_{i1}, \boldsymbol{x}_{i2}, \ldots, \boldsymbol{x}_{in_i}$ are independent and identically distributed samples from $\boldsymbol{x}_i$, $i = 1, 2, \ldots, k$, $n_1 > p$. Further, let $\boldsymbol{x}_i$, $i = 1, 2, \ldots, k$ be mutually independent. Note that k denotes the number of steps. That is, the above data set is called k-step monotone missing data (see Figure 1, where "$*$" indicates a missing observation).

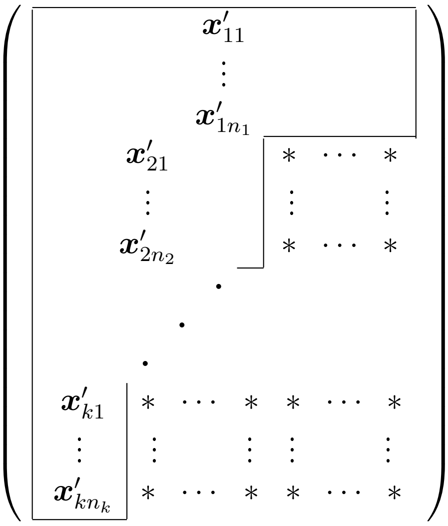

Figure 1. k-step monotone missing data

For a k-step monotone sample or k-step monotone missing data pattern, see Bhargava (1962), Srivastava and Carter (1983), Little and Rubin (2002), and Srivastava (2002), among others. In this paper, we adopt the notation from Yagi and Seo (2015b). As for the partitions of $\mathbf{\Sigma}$, for $1 \le i < j \le k$, let $(\mathbf{\Sigma}_i)_j$ be the principal submatrix of $\mathbf{\Sigma}_i$ of order $p_j \times p_j$; we define

$$\mathbf{\Sigma}_i = (\mathbf{\Sigma}_1)_i, \quad \mathbf{\Sigma}_1 = \mathbf{\Sigma} = \begin{pmatrix} \mathbf{\Sigma}_i & \mathbf{\Sigma}_{i2} \\ \mathbf{\Sigma}_{i2}' & \mathbf{\Sigma}_{i3} \end{pmatrix}, \quad \mathbf{\Sigma}_{i-1} = \begin{pmatrix} \mathbf{\Sigma}_i & \mathbf{\Sigma}_{(i-1,2)} \\ \mathbf{\Sigma}_{(i-1,2)}' & \mathbf{\Sigma}_{(i-1,3)} \end{pmatrix}, \quad i = 2, 3, \ldots, k.$$

For example, we can express $\mathbf{\Sigma}_1$ as

$$\mathbf{\Sigma}_1 = \left(\begin{array}{cc|c} \overbrace{\mathbf{\Sigma}_3}^{p_3} & \overbrace{\mathbf{\Sigma}_{(2,2)}}^{p_2-p_3} & \overbrace{\phantom{\mathbf{\Sigma}_{(1,2)}}}^{p_1-p_2} \\ \mathbf{\Sigma}_{(2,2)}' & \mathbf{\Sigma}_{(2,3)} & \mathbf{\Sigma}_{(1,2)} \\ \hline \multicolumn{2}{c|}{\mathbf{\Sigma}_{(1,2)}'} & \mathbf{\Sigma}_{(1,3)} \end{array}\right) \begin{array}{l} \}p_3 \\ \}p_2 - p_3 \\ \}p_1 - p_2 \end{array}.$$

Using this notation, we can express $\widehat{\mathbf{\Sigma}}$ concretely (see Jinadasa and Tracy (1992)).

Let

$$\boldsymbol{E}_i = \sum_{j=1}^{n_i} (\boldsymbol{x}_{ij} - \overline{\boldsymbol{x}}_i)(\boldsymbol{x}_{ij} - \overline{\boldsymbol{x}}_i)', \quad \overline{\boldsymbol{x}}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} \boldsymbol{x}_{ij}, \quad i = 1, 2, \ldots, k,$$

$$\boldsymbol{d}_1 = \overline{\boldsymbol{x}}_1, \quad \boldsymbol{d}_i = \frac{n_i}{N_{i+1}} \Big[\overline{\boldsymbol{x}}_i - \frac{1}{N_i} \sum_{j=1}^{i-1} n_j (\overline{\boldsymbol{x}}_j)_i \Big], \quad i = 2, 3, \ldots, k,$$

$$N_1 = 0, \quad N_{i+1} = N_i + n_i \ (= \sum_{j=1}^{i} n_j), \quad i = 1, 2, \ldots, k.$$

Then, we note that $\widehat{\boldsymbol{\Sigma}}$ is given by

$$\widehat{\boldsymbol{\Sigma}} = \frac{1}{n_1} \boldsymbol{H}_1 + \sum_{i=2}^{k} \frac{1}{N_{i+1}} \boldsymbol{F}_i \left[\boldsymbol{H}_i - \frac{n_i}{N_i} \boldsymbol{L}_{i-1,1} \right] \boldsymbol{F}_i',$$

where

$$\begin{aligned}
&\boldsymbol{H}_1 = \boldsymbol{E}_1, \ \boldsymbol{H}_i = \boldsymbol{E}_i + \frac{N_i N_{i+1}}{n_i} \boldsymbol{d}_i \boldsymbol{d}_i', \ i = 2, 3, \ldots, k, \\
&\boldsymbol{L}_1 = \boldsymbol{H}_1, \ \boldsymbol{L}_i = (\boldsymbol{L}_{i-1})_i + \boldsymbol{H}_i, \ i = 2, 3, \ldots, k, \\
&\boldsymbol{L}_{i1} = (\boldsymbol{L}_i)_{i+1}, \ \boldsymbol{L}_i = \begin{pmatrix} \boldsymbol{L}_{i1} & \boldsymbol{L}_{i2} \\ \boldsymbol{L}_{i2}' & \boldsymbol{L}_{i3} \end{pmatrix}, \ i = 1, 2, \ldots, k-1, \\
&\boldsymbol{G}_1 = \boldsymbol{I}_{p_1}, \ \boldsymbol{G}_{i+1} = \begin{pmatrix} \boldsymbol{I}_{p_{i+1}} \\ \boldsymbol{L}_{i2}' \boldsymbol{L}_{i1}^{-1} \end{pmatrix}, \ i = 1, 2, \ldots, k-1, \\
&\boldsymbol{F}_1 = \boldsymbol{G}_1, \ \boldsymbol{F}_i = \boldsymbol{F}_{i-1} \boldsymbol{G}_i, \ i = 2, 3, \ldots, k.
\end{aligned}$$

When $k = 3$, we can see that the above result of $\widehat{\boldsymbol{\Sigma}}$ coincides with the result in the three-step case (see Yagi and Seo (2014)).

By the same derivation in Jinadasa and Tracy (1992), the MLE of $\boldsymbol{\Sigma}$ under H_0 can be written as

$$\widetilde{\boldsymbol{\Sigma}} = \frac{1}{n_1} \boldsymbol{V}_1 + \sum_{i=2}^{k} \frac{1}{N_{i+1}} \boldsymbol{P}_i \left[\boldsymbol{V}_i - \frac{n_i}{N_i} \boldsymbol{R}_{i-1,1} \right] \boldsymbol{P}_i',$$

where

$$\boldsymbol{V}_i = \sum_{j=1}^{n_i} \boldsymbol{x}_{ij} \boldsymbol{x}_{ij}', \quad i = 1, 2, \ldots, k,$$

and

$$\begin{aligned}
&\boldsymbol{R}_1 = \boldsymbol{V}_1,\ \boldsymbol{R}_i = (\boldsymbol{R}_{i-1})_i + \boldsymbol{V}_i,\ i = 2, 3, \ldots, k,\\
&\boldsymbol{R}_{i1} = (\boldsymbol{R}_i)_{i+1},\ \boldsymbol{R}_i = \begin{pmatrix} \boldsymbol{R}_{i1} & \boldsymbol{R}_{i2} \\ \boldsymbol{R}_{i2}' & \boldsymbol{R}_{i3} \end{pmatrix},\ i = 1, 2, \ldots, k-1,\\
&\boldsymbol{Q}_1 = \boldsymbol{I}_{p_1},\ \boldsymbol{Q}_{i+1} = \begin{pmatrix} \boldsymbol{I}_{p_{i+1}} \\ \boldsymbol{R}_{i2}'\boldsymbol{R}_{i1}^{-1} \end{pmatrix},\ i = 1, 2, \ldots, k-1,\\
&\boldsymbol{P}_1 = \boldsymbol{Q}_1,\ \boldsymbol{P}_i = \boldsymbol{P}_{i-1}\boldsymbol{Q}_i,\ i = 2, 3, \ldots, k.
\end{aligned}$$

3 LRT and MLRT statistics

Consider the following hypothesis:

$$H_0 : \boldsymbol{\mu} = \boldsymbol{\mu}_0 \text{ vs. } H_1 : \boldsymbol{\mu} \neq \boldsymbol{\mu}_0,$$

where $\boldsymbol{\mu}_0$ is known. Without loss of generality, we can assume that $\boldsymbol{\mu}_0 = \mathbf{0}$. Then, the LR is given by

$$\lambda = \prod_{i=1}^{k} \left(\frac{|\widehat{\boldsymbol{\Sigma}}_i|}{|\widetilde{\boldsymbol{\Sigma}}_i|} \right)^{\frac{1}{2}n_i},$$

where $\widehat{\boldsymbol{\Sigma}}_i$ is the MLE of $\boldsymbol{\Sigma}_i$ under H_1, and $\widetilde{\boldsymbol{\Sigma}}_i$ is the MLE of $\boldsymbol{\Sigma}_i$ under H_0, as in Section 2.

We note that the null distribution of the LRT statistic $Q(= -2\log\lambda)$ is asymptotically a χ^2 distribution with p degrees of freedom. However, it may be noted that the upper percentiles of the χ^2 distribution are not a good approximation to those of the LRT statistic when the sample size is not large. For example, Table 1 gives the simulated values of the upper percentiles of Q and the actual type I error rates for the three- and five-step monotone missing data cases. It may be seen that the upper percentiles of the χ^2 distribution are useful as an approximation to the upper percentiles of Q for cases in which the sample size is considerably large.

Therefore, we consider the MLRT statistic, which converges to the χ^2 distribution faster than the LRT statistic even when the sample size is small.

Table 1 : The upper percentiles of $Q(=-2\log\lambda)$ and the actual type I error rates

	$(p_1,p_2,p_3)=(8,4,2)$		$(p_1,p_2,p_3,p_4,p_5)=(15,12,9,6,3)$	
n_1	$q(\alpha)$	α_q	$q(\alpha)$	α_q
10	41.98	0.567	–	–
20	20.74	0.167	46.86	0.478
30	18.53	0.112	33.86	0.230
40	17.70	0.092	30.91	0.161
50	17.18	0.082	29.52	0.131
100	16.29	0.064	27.15	0.084
200	15.88	0.057	26.03	0.065
400	15.68	0.053	25.53	0.057
∞	15.51	0.050	25.00	0.050

Note. $q(\alpha)$: the upper 100α percentiles of Q, $\alpha_q=\Pr\{Q>\chi^2_{p,\alpha}\}$,
$n_2=n_3=\cdots=n_5=10,\ \alpha=0.05$.

Using the notation in Section 2, we define

$$\mathbf{\Sigma}_{(i,3)\cdot i+1}=\mathbf{\Sigma}_{(i,3)}-\mathbf{\Sigma}'_{(i,2)}\mathbf{\Sigma}^{-1}_{i+1}\mathbf{\Sigma}_{(i,2)},\quad i=1,2,\ldots,k-1.$$

Then we decompose λ as $\lambda=\prod_{i=1}^{k}\lambda_i$, where

$$\lambda_i=\left(\frac{|\widehat{\mathbf{\Sigma}}_{(i,3)\cdot i+1}|}{|\widetilde{\mathbf{\Sigma}}_{(i,3)\cdot i+1}|}\right)^{\frac{N_{i+1}}{2}},\quad i=1,2,\ldots,k-1,\quad \lambda_k=\left(\frac{|\widehat{\mathbf{\Sigma}}_k|}{|\widetilde{\mathbf{\Sigma}}_k|}\right)^{\frac{N_{k+1}}{2}}.$$

We note that the values of $\lambda_i,\ i=1,2,\ldots,k$ are mutually all independent (see Hao and Krishnamoorthy (2001)). Using an asymptotic expansion of the null distribution of λ_i, we can derive a modified LR as

$$\lambda^*=\prod_{i=1}^{k}\lambda_i^{\rho_i},$$

where

$$\rho_i=1-\frac{1}{2N_{i+1}}(p_i+p_{i+1}+2),\quad i=1,2,\ldots,k-1,\quad \rho_k=1-\frac{1}{2N_{k+1}}(p_k+2).$$

For simplicity, in order to show the derivation of the value of ρ_i, we first derive an asymptotic expansion of the null distribution of λ_i in the case of $k=2$. For $k=2$, we have $\lambda = \lambda_1 \lambda_2$, where

$$\lambda_1 = \left(\frac{|\widehat{\boldsymbol{\Sigma}}_{(1,3)\cdot 2}|}{|\widetilde{\boldsymbol{\Sigma}}_{(1,3)\cdot 2}|} \right)^{\frac{n_1}{2}}, \quad \lambda_2 = \left(\frac{|\widehat{\boldsymbol{\Sigma}}_2|}{|\widetilde{\boldsymbol{\Sigma}}_2|} \right)^{\frac{N_3}{2}}.$$

We note that $\boldsymbol{W}_1 (= n_1 \widehat{\boldsymbol{\Sigma}}_{(1,3)\cdot 2})$ follows a Wishart distribution, $W_{p_1 - p_2}(\boldsymbol{\Sigma}_{(1,3)\cdot 2}, n_1 - p_2 - 1)$. Let $\overline{\boldsymbol{w}}_1$ be a $(p_1 - p_2) \times 1$ vector such that

$$n_1 \widetilde{\boldsymbol{\Sigma}}_{(1,3)\cdot 2} = \boldsymbol{W}_1 + (n_1 - p_2) \overline{\boldsymbol{w}}_1 \overline{\boldsymbol{w}}_1'.$$

Then λ_1 can be written as

$$\lambda_1 = \frac{|\boldsymbol{W}_1|^{\frac{n_1}{2}}}{|\boldsymbol{W}_1 + (n_1 - p_2) \overline{\boldsymbol{w}}_1 \overline{\boldsymbol{w}}_1'|^{\frac{n_1}{2}}}.$$

Further, letting

$$\frac{1}{n_1 - p_2 - 1} \boldsymbol{W}_1 = \boldsymbol{I}_{p_1 - p_2} + \frac{1}{\sqrt{n_1 - p_2}} \boldsymbol{V}_1,$$
$$\overline{\boldsymbol{w}}_1 = \frac{1}{\sqrt{n_1 - p_2}} \boldsymbol{z}_1,$$

then the characteristic function of $Q_1 (= -2 \log \lambda_1)$ can be expanded as

$$\begin{aligned} &E[\exp\{it(Q_1)\}] \\ &= E\Big[\exp\{(it)\boldsymbol{z}_1'\boldsymbol{z}_1\}\Big(1 - \frac{1}{\sqrt{n_1}}(it)\boldsymbol{z}_1'\boldsymbol{V}_1\boldsymbol{z}_1 \\ &\quad + \frac{1}{n_1}\Big\{(it)(p_2+1)\boldsymbol{z}_1'\boldsymbol{z}_1 - \frac{1}{2}(it)(\boldsymbol{z}_1'\boldsymbol{z}_1)^2 + (it)\boldsymbol{z}_1'\boldsymbol{V}_1^2\boldsymbol{z}_1 + \frac{1}{2}(it)^2(\boldsymbol{z}_1'\boldsymbol{V}_1\boldsymbol{z}_1)^2\Big\}\Big)\Big] \\ &\quad + O(n_1^{-\frac{3}{2}}). \end{aligned}$$

Therefore, inverting the characteristic function, we have

$$\Pr(Q_1 \le x) = G_{p_1 - p_2}(x) + \frac{1}{n_1}\left[\beta_1 G_{p_1 - p_2}(x) - \beta_1 G_{p_1 - p_2 + 2}(x)\right] + O(n_1^{-2}),$$

where

$$\beta_1 = -\frac{1}{4}(p_1 - p_2)(p_1 + p_2 + 2),$$

and $G_p(x)$ is the distribution function of a χ^2-variate with p degrees of freedom. Therefore, if $\rho_1 = 1 - (p_1 + p_2 + 2)/(2n_1)$, then the cumulative distribution function of $Q_1^* (=$

$-2\rho_1 \log \lambda_1$) is given by

$$\Pr(Q_1^* \leq x) = G_{p_1-p_2}(x) + O(n_1^{-2}).$$

As with the case of λ_1, by the perturbation method, the cumulative distribution function of $Q_2(= -2\log \lambda_2)$ can be written as

$$\Pr(Q_2 \leq x) = G_{p_2}(x) + \frac{1}{N_3}\left[\beta_2 G_{p_2}(x) - \beta_2 G_{p_2+2}(x)\right] + O(N_3^{-2}),$$

where

$$\beta_2 = -\frac{1}{4}p_2(p_2+2).$$

Therefore, if $\rho_2 = 1 - (p_2+2)/(2N_3)$, then the cumulative distribution function of $Q_2^*(= -2\rho_2 \log \lambda_2)$ is given by

$$\Pr(Q_2^* \leq x) = G_{p_2}(x) + O(N_3^{-2}).$$

As a remark, for k-step monotone missing data case, the MLRT statistic $Q^\dagger(= -2\rho \log \lambda)$ can be obtained by the above result, where

$$\rho = 1 - \frac{1}{2p}\left\{\sum_{i=1}^{k-1}\frac{1}{N_{i+1}}(p_i - p_{i+1})(p_i + p_{i+1} + 2) + \frac{1}{N_{k+1}}p_k(p_k+2)\right\}.$$

It holds that the cumulative distribution function of $Q^\dagger$ is given by

$$\Pr(Q^\dagger \leq x) = G_p(x) + O(n_1^{-2}).$$

We note that the value of ρ coincides with that of Krishnamoorthy and Pannala (1998) when $k = 2$.

4 Simulation studies

In this section, we will study the numerical accuracy of the upper percentiles of the MLRT statistics using the actual test sizes. In order to investigate the accuracy of the approximation for the one-sample case, we compute the upper percentiles of Q, Q^*, and

$Q^\dagger$ with monotone missing data by Monte Carlo simulation. For each parameter, the simulation was executed 10^6 times using normal random vectors generated from $N_{p_i}(\mathbf{0}, \boldsymbol{I}_{p_i})$, $i = 1, 2, \ldots, k$.

In Tables 2 and 3, we provide the simulated upper 100α percentiles of Q, Q^* and $Q^\dagger$ for the three-step and five-step cases. Further, we provide the actual test sizes, α_q, α_{q^*} and $\alpha_{q^\dagger}$, given by

$$\alpha_q = \Pr\{Q > \chi^2_{p,\alpha}\}, \quad \alpha_{q^*} = \Pr\{Q^* > \chi^2_{p,\alpha}\}, \quad \text{and} \quad \alpha_{q^\dagger} = \Pr\{Q^\dagger > \chi^2_{p,\alpha}\},$$

respectively, where $\chi^2_{p,\alpha}$ is the upper percentile of the χ^2 distribution with p degrees of freedom. It may be noted from Tables 2 and 3 that each value of q, q^* and $q^\dagger$ is closer to the upper percentiles of the χ^2 distribution with p degrees of freedom, $\chi^2_{p,\alpha}$, when the n_1 becomes large. It is seen from Table 2 that $q^*(\alpha)$ for $(p_1, p_2, p_3) = (8, 4, 2)$ and $(15, 12, 9)$ is a considerably good approximate value when n_1 is greater than 20. Similarly, it is seen from Table 3 that $q^*(\alpha)$ for $(p_1, p_2, p_3.p_4, p_5) = (15, 12, 9, 6, 3)$ is a considerably good approximate value when n_1 is greater than 20 without regard to the sample size of n_i, $i \geq 2$. It may be noted from the simulation results that the MLRT statistic Q^* converges to the χ^2 distribution faster than the MLRT statistic $Q^\dagger$ in almost all cases.

In conclusion, we have developed the MLRT statistics Q^* and $Q^\dagger$ with general monotone missing data in one-sample problem. Our MLRT statistics are considerably more accurate than the χ^2 approximation, even for small samples.

Table 2: The upper percentiles of $Q(=-2\log\lambda)$, $Q^*(=-2\log\lambda^*)$, $Q^\dagger(=-2\rho\log\lambda)$ and the actual type I error rates

Sample Size			Upper Percentile			Type I Error Rate		
n_1	n_2	n_3	$q(\alpha)$	$q^*(\alpha)$	$q^\dagger(\alpha)$	α_q	α_{q^*}	$\alpha_{q^\dagger}$
				$(p_1,p_2,p_3)=(8,4,2)$				
				$\alpha=0.05$				
10	10	10	41.98	16.77	24.49	.567	.072	.221
20	10	10	20.74	15.58	16.16	.167	.051	.061
30	10	10	18.53	15.52	15.72	.112	.050	.054
40	10	10	17.70	15.55	15.65	.092	.051	.052
50	10	10	17.18	15.52	15.57	.082	.050	.051
100	10	10	16.29	15.50	15.51	.064	.050	.050
200	10	10	15.88	15.48	15.49	.057	.050	.050
				$\alpha=0.01$				
10	10	10	58.22	22.08	33.96	.393	.018	.105
20	10	10	27.09	20.20	21.11	.059	.010	.014
30	10	10	24.09	20.12	20.44	.032	.010	.011
40	10	10	22.94	20.12	20.28	.024	.010	.011
50	10	10	22.28	20.11	20.19	.020	.010	.010
100	10	10	21.14	20.09	20.12	.014	.010	.010
200	10	10	20.56	20.05	20.05	.012	.010	.010
				$(p_1,p_2,p_3)=(15,12,9)$				
				$\alpha=0.05$				
20	10	10	47.04	25.17	32.73	.485	.052	.174
30	10	10	34.05	25.08	26.56	.235	.051	.072
40	10	10	31.07	25.06	25.68	.165	.051	.060
50	10	10	29.58	25.00	25.34	.133	.050	.055
100	10	10	27.17	25.00	25.06	.084	.050	.051
200	10	10	26.09	25.02	25.03	.066	.050	.050
400	10	10	25.51	24.98	24.98	.057	.050	.050
				$\alpha=0.01$				
20	10	10	60.58	30.84	42.15	.297	.011	.072
30	10	10	42.08	30.69	32.81	.094	.010	.018
40	10	10	38.13	30.66	31.52	.056	.010	.013
50	10	10	36.29	30.61	31.08	.040	.010	.012
100	10	10	33.28	30.64	30.71	.021	.010	.010
200	10	10	31.91	30.61	30.62	.015	.010	.010
400	10	10	31.25	30.59	30.60	.012	.010	.010

Note. $\alpha_q=\Pr\{Q>\chi^2_{p,\alpha}\}$, $\alpha_{q^*}=\Pr\{Q^*>\chi^2_{p,\alpha}\}$, $\alpha_{q^\dagger}=\Pr\{Q^\dagger>\chi^2_{p,\alpha}\}$, $\chi^2_{8,0.05}=15.51$, $\chi^2_{8,0.01}=20.09$, $\chi^2_{15,0.05}=25.00$, $\chi^2_{15,0.01}=30.58$.

Table 3: The upper percentiles of $Q(= -2\log\lambda)$, $Q^*(= -2\log\lambda^*)$, $Q^\dagger(= -2\rho\log\lambda)$ and the actual type I error rates

Sample Size		Upper Percentile			Type I Error Rate		
n_1	$n_2 = \cdots = n_5$	$q(\alpha)$	$q^*(\alpha)$	$q^\dagger(\alpha)$	α_q	α_{q^*}	$\alpha_{q^\dagger}$
		$(p_1, p_2, p_3, p_4, p_5) = (15, 12, 9, 6, 3)$					
		$\alpha = 0.05$					
20	10	46.86	25.08	33.06	.478	.051	.179
30	10	33.86	24.99	26.63	.230	.050	.074
40	10	30.91	24.98	25.69	.161	.050	.060
50	10	29.52	24.99	25.39	.131	.050	.055
100	10	27.15	25.02	25.08	.084	.050	.051
200	10	26.03	24.98	24.99	.065	.050	.050
400	10	25.53	25.00	25.01	.057	.050	.050
20	5	48.88	25.14	32.38	.530	.052	.169
40	10	30.91	24.97	25.69	.162	.050	.060
80	20	27.47	25.01	25.15	.089	.050	.052
160	40	26.14	25.01	25.04	.067	.050	.051
320	80	25.56	25.02	25.02	.058	.050	.050
640	160	25.28	25.01	25.01	.054	.050	.050
		$\alpha = 0.01$					
20	10	60.36	30.67	42.59	.292	.010	.075
30	10	41.82	30.55	32.89	.092	.010	.018
40	10	37.97	30.63	31.56	.054	.010	.013
50	10	36.16	30.54	31.10	.039	.010	.012
100	10	33.23	30.58	30.70	.021	.010	.010
200	10	31.89	30.60	30.62	.015	.010	.010
400	10	31.26	30.61	30.61	.012	.010	.010
20	5	62.54	30.76	41.42	.337	.011	.068
40	10	37.97	30.56	31.56	.054	.010	.013
80	20	33.65	30.61	30.81	.023	.010	.011
160	40	31.98	30.61	30.63	.015	.010	.010
320	80	31.22	30.54	30.56	.012	.010	.010
640	160	30.91	30.58	30.59	.011	.010	.010

Note. $\alpha_q = \Pr\{Q > \chi^2_{p,\alpha}\}$, $\alpha_{q^*} = \Pr\{Q^* > \chi^2_{p,\alpha}\}$, $\alpha_{q^\dagger} = \Pr\{Q^\dagger > \chi^2_{p,\alpha}\}$, $\chi^2_{15,0.05} = 25.00$, $\chi^2_{15,0.01} = 30.58$.

Acknowledgment

The first and second authors' research was partly supported by a Grant-in-Aid for JSPS Fellows (15J00414) and a Grant-in-Aid for Scientific Research (C) (26330050), respectively.

References

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

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

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

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

[5] Kanda, T. and Fujikoshi, Y. (1998). Some basic properties of the MLE's for a multivariate normal distribution with monotone missing data. *American Journal of Mathematical and Management Sciences*, **18**, 161–190.

[6] Krishnamoorthy, K. and Pannala, M. K. (1998). Some simple test procedures for normal mean vector with incomplete data. *Annals of the Institute of Statistical Mathematics*, **50**, 531–542.

[7] Krishnamoorthy, K. and Pannala, M. K. (1999). Confidence estimation of a normal mean vector with incomplete data. *The Canadian Journal of Statistics*, **27**, 395–407.

[8] Little, R. J. A. and Rubin, D. B. (2002). *Statistical Analysis with Missing Data, 2nd ed.* Wiley, New York.

[9] Seko, N., Kawasaki, T. and Seo, T. (2011). Testing equality of two mean vectors with two-step monotone missing data. *American Journal of Mathematical and Management Sciences*, **31**, 117–135.

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

[11] Srivastava, M. S. (2002). *Methods of Multivariate Statistics.* Wiley, New York.

[12] Srivastava, M. S. and Carter, E. M. (1983). *An Introduction to Applied Multivariate Statistics.* Elsevier North-Holland, New York.

[13] Yagi, A. and Seo, T. (2014). A test for mean vector and simultaneous confidence intervals with three-step monotone missing data. *American Journal of Mathematical and Management Sciences*, **33**, 161–175.

[14] Yagi, A. and Seo, T. (2015a) Tests for equality of mean vectors and simultaneous confidence intervals with two-step or three-step monotone missing data patterns. *American Journal of Mathematical and Management Sciences*, **34**, 213–233.

[15] Yagi, A. and Seo, T. (2015b). Tests for normal mean vectors with monotone incomplete data. *Technical Report No.15–07, Statistical Research Group, Hiroshima University, Hiroshima, Japan.*

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