Hindawi
Journal of Mathematics
Volume 2021, Article ID 8837044, 9 pages
https://doi.org/10.1155/2021/8837044

Research Article

Hierarchical Missing Data and Multivariate Behrens–Fisher Problem

Jianqi Yu

Guilin University of Technology, Guilin 541004, Guangxi, China

Correspondence should be addressed to Jianqi Yu; 498616729@qq.com

Received 30 August 2020; Revised 1 January 2021; Accepted 5 January 2021; Published 27 January 2021

Academic Editor: Ljubisa Kocinac

This article firstly defines hierarchical data missing pattern, which is a generalization of monotone data missing pattern. Then multivariate Behrens–Fisher problem with hierarchical missing data is considered to illustrate that how ideas in dealing with monotone missing data can be extended to deal with hierarchical missing pattern. A pivotal quantity similar to the Hotelling T^2 is presented, and the moment matching method is used to derive its approximate distribution which is for testing and interval estimation. The precision of the approximation is illustrated through Monte Carlo data simulation. The results indicate that the approximate method is very satisfactory even for moderately small samples.

1. Introduction

Inferences with incomplete data have aroused lots of interest among statisticians in the past as well as present. The causes for missing data could be various which will not be discussed in this article. However, to ignore the process that causes missing data, it is usually assumed that the data are missing at random (MAR). For an exposition of such issues, we refer to Little and Rubin [1] or Little [2]. Lu and Copas [3] 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.

There are a few missing patterns considered in the literatures, but the incomplete data with monotone pattern (see display (1) and (2)) not only occur frequently in practice but also it allows the exact calculation of the maximum likelihood estimators (MLEs) and the likelihood ratio statistics and relevant distributions if multivariate normality is assumed. Anderson [4], one of the earliest authors in this area, gave a simple approach to derive the MLEs and present them for a special case of monotone pattern. Krishnamoorthy and Pannala [5, 6] provided an accurate, simple approach to construct a confidence region for a normal mean vector. Hao and Krishnamoorthy [7] developed an inferential procedure on a normal covariance matrix. Yu et al. [8] 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 [9] considered the Behrens–Fisher problem. Yu et al. [10] considered the problem of testing equality of two normal covariance matrices with monotone missing data.

Besides, Batsidis [11–13] extends the inferences on monotone missing data to the assumption of elliptically contoured distributions of which the multivariate normal is a special case. For theory and methods of multivariate analysis based on the elliptically contoured distributions, we refer to Fang and Zhang [14].

Most of the papers mentioned above use a similar strategy in dealing with the monotone missing data. To illustrate this, consider the data matrices with 2-block monotone pattern as shown below:

$$\begin{matrix} \mathbf{x}_1, \ldots, \mathbf{x}_n & \mathbf{x}_{n+1} \ldots, \mathbf{x}_{n+m} \\ \mathbf{y}_1, \ldots, \mathbf{y}_n & \end{matrix}. \quad (1)$$

The strategy is as follows: if we do not have the extradata on y, i.e., we have only the first n samples on (x, y), usually we already have a statistics, say Q, out of the complete data. Similarly, if we have only $(n + m)$ sample on x, we also have a similar statistics, say Q_1, for the lower-dimensional problem.

We then decompose Q into two parts $Q_1' + Q_2$, which correspond to the sample data on x and y, respectively. However, since we have extradata on y, Q_1' should be replaced with Q_1. Hence, we get the final statistics for inference $Q_1 + Q_2$.

In this article, we will define a new data missing pattern, the hierarchical data missing pattern, which is a generalization of monotone missing pattern. Moreover, the strategy just mentioned can also be used. To see this, we consider the multivariate Behrens–Fisher problem with hierarchical missing data. The approach that we will employ is based on the one due to Krishnamoorthy and Yu [9] for the monotone missing data.

The article is organized as follows: in the following section, we define the hierarchical data missing pattern. In Section 3, an approximate method for the multivariate Behrens–Fisher problem with hierarchical missing data is outlined. The accuracy of the approximation is investigated using the Monte Carlo simulation in Section 4. The methods are illustrated using an example in Section 5, and some concluding remarks are given in Section 6.

2. Hierarchical Data Missing Pattern

Suppose $(\mathbf{x_{1i}}, \mathbf{x_{2i}}, \ldots, \mathbf{x_{ki}})'$ is the i-th observation of the random vector $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x_k})'$, $i = 1, 2, \ldots, N_1$. The monotone pattern of missing data is like following data:

$$\begin{array}{r} \mathbf{x}_{11}, \ldots, \mathbf{x}_{1N_k}, \ldots, \mathbf{x}_{1N_2}, \ldots, \mathbf{x}_{1N_1}, \\ \mathbf{x}_{21}, \ldots, \mathbf{x}_{2N_k}, \ldots, \mathbf{x}_{2N_2}, \\ \ldots\ldots, \\ \mathbf{x}_{k1}, \ldots, \mathbf{x}_{kN_k}, \end{array} \tag{2}$$

where $\mathbf{x}_{ij}$ is a $p_i \times 1$ vector, $N_1 \geq N_2 \geq \cdots \geq N_k$, $i = 1, \ldots, k$. In other words, 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 $N_1 \geq N_2 \geq \cdots \geq N_k$ and $p_1 + \cdots + p_k = p$.

We define the hierarchical data missing pattern of as the following pattern:

$$\begin{array}{llll} \mathbf{x}_1, \ldots, \mathbf{x}_n & \mathbf{x}_{n+1}, \ldots, \mathbf{x}_{n+m} & \mathbf{x}_{n+m+1}, \ldots, \mathbf{x}_{n+m+l} & \mathbf{x}_{n+m+1+1}, \ldots, \mathbf{x}_N \\ \mathbf{y}_1, \ldots, \mathbf{y}_n & \mathbf{y}_{n+1}, \ldots, \mathbf{y}_{n+m} & & \\ & & \mathbf{z}_{n+m+1}, \ldots, \mathbf{z}_{n+m+l} & \\ & & & \mathbf{u}_{n+m+1+1}, \ldots, \mathbf{u}_N \\ \mathbf{v}_1, \ldots, \mathbf{v}_n & & & \\ & \mathbf{w}_{n+1}, \ldots, \mathbf{w}_{n+m} & & \end{array}, \tag{3}$$

where the index sets satisfy following conditions:

(1) The index set of the first row, i.e $(1, \ldots, n, n+1, \ldots, n+m, n+m+1, \ldots, n+m+1, n+m+1+1 \ldots, N)$, is the union of the index sets of all the other rows.

(2) The index sets of two different rows are either disjoint, or inclusive. It is easy to see that the monotone pattern is a special case of the hierarchical pattern.

Now we consider the Behrens–Fisher problem with hierarchical missing data.

To formulate the problem, let $\mathbf{x}$ follows a p-variate normal distribution with mean vector μ and covariance matrix Σ, and we write this as $\mathbf{x} \sim N_p(\mu, \Sigma)$. Meanwhile, let $\mathbf{y} \sim N_p(\beta, \Delta)$, and $\mathbf{y}$ is independent of $\mathbf{x}$. It is assumed that Σ and Δ are unknown and arbitrary positive definite matrices. Let us consider the problem of testing:

$$H_0\colon\ \boldsymbol{\mu} = \boldsymbol{\beta} \text{ vs. } H_a\colon\ \boldsymbol{\mu} \neq \boldsymbol{\beta}. \tag{4}$$

Suppose that we have a sample of N_1 observations available on $\mathbf{x}$ and a sample of M_1 observations available on $\mathbf{y}$. We consider a simple 3-block hierarchical data as shown below (it is easy to extend the ideas and procedures for 3-block data to general case as in (3), but the notation will become very complicated):

$$\begin{array}{ll} \mathbf{x}_{11}, \ldots, \mathbf{x}_{1N_2}, & \mathbf{x}_{1,N_2+1}, \ldots, \mathbf{x}_{1N_1} \\ \mathbf{x}_{21}, \ldots, \mathbf{x}_{2N_2}, & \\ & \mathbf{x}_{3,N_2+1}, \ldots, \mathbf{x}_{3N_1} \end{array} \qquad \begin{array}{ll} \mathbf{y}_{11}, \ldots, \mathbf{y}_{1M_2}, & \mathbf{y}_{1,M_2+1}, \ldots, \mathbf{y}_{1M_1} \\ \mathbf{y}_{21}, \ldots, \mathbf{y}_{2M_2}, & \\ & \mathbf{y}_{3,M_2+1}, \ldots, \mathbf{y}_{3M_1} \end{array}, \tag{5}$$

where $\mathbf{x}_{ij}$ is a $p_i \times 1$ vector, $j = 1, \ldots, N_i$, while $\mathbf{y}_{ij}$ is a $q_i \times 1$ vector, $j = 1, \ldots, M_i$, $i = 1, 2, 3$. In other words, in the $\mathbf{x}$ 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 $N_1 - N_2$ observations available on the first p_1 and the last p_3 components. Notice that $N_1 \geq N_2$, $M_1 \geq M_2$, and $p_1 + p_2 + p_3 = q_1 + q_2 + q_3 = p$.

As pointed in Yu et al. [8], we do not need to consider the case of unequal pattern, i.e, $p_i \neq q_i$, for some $i = 1, 2, 3$, since any type of unequal patterns data can be rearranged to form an equal monotone pattern. For example, assume that $p_1 > q_1$, $p_2 > q_2$, $p_3 < q_3$, and $p_1 < q_1 + q_2$. Let $r_1 = q_1, r_2 = p_1 - q_1, r_3 = q_1 + q_2 - p_1, r_4 = p_1 + p_2 - q_1 - q_2, r_5 = p_3$, then it is obvious that we have a 5-block equal pattern.

Hence, without loss of generality, we assume that $p_i = q_i, i = 1, 2, 3$.

3. Inference on $\mu - \beta$

3.1. Preliminaries. Consider the data matrices in (5) and partition the data matrices as follows:

$$\mathbf{X}_1 = \left(\mathbf{x}_{11}, \ldots, \mathbf{x}_{1N_2}, \ldots, \mathbf{x}_{1N_1}\right)_{p_1 \times N_1},$$
$$\mathbf{X}_2 = \begin{pmatrix} \mathbf{x}_{11}, \ldots, \mathbf{x}_{1N_2} \\ \mathbf{x}_{21}, \ldots, \mathbf{x}_{2N_2} \end{pmatrix}_{(p_1+p_2)\times N_2}, \tag{6}$$
$$\mathbf{X}_3 = \begin{pmatrix} \mathbf{x}_{1,N_2+1}, \ldots, \mathbf{x}_{1N_1} \\ \mathbf{x}_{3,N_2+1}, \ldots, \mathbf{x}_{3N_1} \end{pmatrix}_{(p_1+p_3)\times N_3},$$

where $N_3 = N_1 - N_2$.

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

$$\mathbf{Y}_1 = \left(\mathbf{y}_{11}, \ldots, \mathbf{y}_{1M_2}, \ldots, \mathbf{y}_{1M_1}\right)_{p_1 \times M_1},$$
$$\mathbf{Y}_2 = \begin{pmatrix} \mathbf{y}_{11}, \ldots, \mathbf{y}_{1M_2} \\ \mathbf{y}_{21}, \ldots, \mathbf{y}_{2M_2} \end{pmatrix}_{(p_1+p_2)\times M_2}, \tag{7}$$
$$\mathbf{Y}_3 = \begin{pmatrix} \mathbf{y}_{1,M_2+1}, \ldots, \mathbf{y}_{1M_1} \\ \mathbf{y}_{3,M_2+1}, \ldots, \mathbf{y}_{3M_1} \end{pmatrix}_{(p_1+p_3)\times M_3},$$

where $M_3 = M_1 - M_2$.

Let $\overline{\mathbf{x}}_l$ and $\mathbf{S}_l$ denote, respectively, the sample mean vector and the sum of squares and sum of products matrix based on $\mathbf{X}_l$, $l = 1, 2, 3$. Similarly, let $\overline{\mathbf{y}}_l$ and $\mathbf{V}_l$ denote, respectively, the sample mean vector and the sums of squares and products matrix based on $\mathbf{Y}_l$, $l = 1, 2, 3$. We partition these means and matrices accordingly as follows:

$$\begin{aligned} \overline{\mathbf{x}}_1 &= \overline{\mathbf{x}}_1^{(1)}, \\ \overline{\mathbf{x}}_2 &= \begin{pmatrix} \overline{\mathbf{x}}_2^{(1)} \\ \overline{\mathbf{x}}_2^{(2)} \end{pmatrix}, \\ \overline{\mathbf{x}}_3 &= \begin{pmatrix} \overline{\mathbf{x}}_3^{(1)} \\ \overline{\mathbf{x}}_3^{(3)} \end{pmatrix}, \\ \mathbf{S}_1 &= \mathbf{S}_1^{(1,1)}, \\ \mathbf{S}_2 &= \begin{pmatrix} \mathbf{S}_2^{(1,1)} & \mathbf{S}_2^{(1,2)} \\ \mathbf{S}_2^{(2,1)} & \mathbf{S}_2^{(2,2)} \end{pmatrix}, \\ \mathbf{S}_3 &= \begin{pmatrix} \mathbf{S}_3^{(1,1)} & \mathbf{S}_3^{(1,3)} \\ \mathbf{S}_3^{(3,1)} & \mathbf{S}_3^{(3,3)} \end{pmatrix}. \end{aligned} \tag{8}$$

The statistics $\overline{\mathbf{y}_l}$ and $\mathbf{V}_l$ based on the data matrix $\mathbf{Y}$ in (7) are also partitioned like $\overline{\mathbf{x}}_l$ and $\mathbf{S}_l$:

$$\begin{aligned} \overline{\mathbf{y}}_1 &= \overline{\mathbf{y}}_1^{(1)}, \\ \overline{\mathbf{y}}_2 &= \begin{pmatrix} \overline{\mathbf{y}}_2^{(1)} \\ \overline{\mathbf{y}}_2^{(2)} \end{pmatrix}, \\ \overline{\mathbf{y}}_3 &= \begin{pmatrix} \overline{\mathbf{y}}_3^{(1)} \\ \overline{\mathbf{y}}_3^{(3)} \end{pmatrix}, \\ \mathbf{V}_1 &= \mathbf{V}_1^{(1,1)}, \\ \mathbf{V}_2 &= \begin{pmatrix} \mathbf{V}_2^{(1,1)} & \mathbf{V}_2^{(1,2)} \\ \mathbf{V}_2^{(2,1)} & \mathbf{V}_2^{(2,2)} \end{pmatrix}, \\ \mathbf{V}_3 &= \begin{pmatrix} \mathbf{V}_3^{(1,1)} & \mathbf{V}_3^{(1,3)} \\ \mathbf{V}_3^{(3,1)} & \mathbf{V}_3^{(3,3)} \end{pmatrix}. \end{aligned} \tag{9}$$

Finally, we partition the parameters as follows:

$$\begin{aligned} \boldsymbol{\mu} &= \begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \\ \boldsymbol{\mu}_3 \end{pmatrix}, \\ \boldsymbol{\Sigma} &= \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} & \boldsymbol{\Sigma}_{13} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} & \boldsymbol{\Sigma}_{23} \\ \boldsymbol{\Sigma}_{31} & \boldsymbol{\Sigma}_{32} & \boldsymbol{\Sigma}_{33} \end{pmatrix}, \\ \boldsymbol{\beta} &= \begin{pmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \\ \boldsymbol{\beta}_3 \end{pmatrix}, \\ \boldsymbol{\Delta} &= \begin{pmatrix} \boldsymbol{\Delta}_{11} & \boldsymbol{\Delta}_{12} & \boldsymbol{\Delta}_{13} \\ \boldsymbol{\Delta}_{21} & \boldsymbol{\Delta}_{22} & \boldsymbol{\Delta}_{23} \\ \boldsymbol{\Delta}_{31} & \boldsymbol{\Delta}_{32} & \boldsymbol{\Delta}_{33} \end{pmatrix}. \end{aligned} \tag{10}$$

where μ_i is p_i dimensional, $i = 1, 2, 3$.

Furthermore, define $\delta = \mu - \beta$ so that

$$\begin{aligned} \boldsymbol{\delta}_1 &= \boldsymbol{\mu}_1 - \boldsymbol{\beta}_1, \\ \boldsymbol{\delta}_2 &= \boldsymbol{\mu}_2 - \boldsymbol{\beta}_2, \\ \boldsymbol{\delta}_3 &= \boldsymbol{\mu}_3 - \boldsymbol{\beta}_3. \end{aligned} \tag{11}$$

Let $n_i = N_i - 1$ and $m_i = M_i - 1$, $i = 1, 2, 3$.

The following summary statistics are needed to define the pivotal quantity that we will use for hypothesis testing about δ. Let

$$\begin{aligned} \mathbf{C}_1 &= \frac{\mathbf{S}_1^{(1,1)}}{n_1 N_1} + \frac{\mathbf{V}_1^{(1,1)}}{m_1 M_1}, \\ \mathbf{C}_2 &= \frac{\mathbf{S}_2}{n_2 N_2} + \frac{\mathbf{V}_2}{m_2 M_2}, \\ \mathbf{C}_3 &= \frac{\mathbf{S}_3}{n_3 N_3} + \frac{\mathbf{V}_3}{m_3 M_3}, \\ \mathbf{B}_{2.1} &= \left[\frac{\mathbf{S}_2^{(2,1)}}{n_2 N_2} + \frac{\mathbf{V}_2^{(2,1)}}{m_2 M_2}\right]\left[\frac{\mathbf{S}_2^{(1,1)}}{n_2 N_2} + \frac{\mathbf{V}_2^{(1,1)}}{m_2 M_2}\right]^{-1}, \\ \mathbf{B}_{3.1} &= \left[\frac{\mathbf{S}_3^{(3,1)}}{n_3 N_3} + \frac{\mathbf{V}_3^{(3,1)}}{m_3 M_3}\right]\left[\frac{\mathbf{S}_3^{(1,1)}}{n_3 N_3} + \frac{\mathbf{V}_3^{(1,1)}}{m_3 M_3}\right]^{-1}, \\ \mathbf{C}_{2.1} &= \left[\frac{\mathbf{S}_2^{(2,2)}}{n_2 N_2} + \frac{\mathbf{V}_2^{(2,2)}}{m_2 M_2}\right] - \mathbf{B}_{2.1}\left[\frac{\mathbf{S}_2^{(1,2)}}{n_2 N_2} + \frac{\mathbf{V}_2^{(1,2)}}{m_2 M_2}\right], \\ \mathbf{C}_{3.1} &= \left[\frac{\mathbf{S}_3^{(3,3)}}{n_3 N_3} + \frac{\mathbf{V}_3^{(3,3)}}{m_3 M_3}\right] - \mathbf{B}_{3.1}\left[\frac{\mathbf{S}_3^{(1,3)}}{n_3 N_3} + \frac{\mathbf{V}_3^{(1,3)}}{m_3 M_3}\right]. \end{aligned} \tag{12}$$

Furthermore, let

$$\widehat{\boldsymbol{\delta}}_1 = \overline{\mathbf{x}}_1^{(1)} - \overline{\mathbf{y}}_1^{(1)}, \widehat{\boldsymbol{\delta}}_2 = \overline{\mathbf{x}}_2^{(2)} - \overline{\mathbf{y}}_2^{(2)}, \widehat{\boldsymbol{\delta}}_3 = \overline{\mathbf{x}}_3^{(3)} - \overline{\mathbf{y}}_3^{(3)}, \widehat{\boldsymbol{\delta}} = (\widehat{\boldsymbol{\delta}}_1', \widehat{\boldsymbol{\delta}}_2', \widehat{\boldsymbol{\delta}}_3')'. \tag{13}$$

The pivotal quantity that we propose for hypothesis testing and confidence estimation of δ which is given by

$$\begin{aligned} Q &= (\hat{\boldsymbol{\delta}}_1 - \boldsymbol{\delta}_1)' \mathbf{C}_1^{-1} (\hat{\boldsymbol{\delta}}_1 - \boldsymbol{\delta}_1) \\ &\quad + ((\hat{\boldsymbol{\delta}}_2 - \boldsymbol{\delta}_2) - \mathbf{B}_{2.1} (\overline{\mathbf{x}}_2^{(1)} - \overline{\mathbf{y}}_2^{(1)} - \boldsymbol{\delta}_1))' (\mathbf{C}_{2.1})^{-1} ((\hat{\boldsymbol{\delta}}_2 - \boldsymbol{\delta}_2) - \mathbf{B}_{2.1} (\overline{\mathbf{x}}_2^{(1)} - \overline{\mathbf{y}}_2^{(1)} - \boldsymbol{\delta}_1)) \\ &\quad + ((\hat{\boldsymbol{\delta}}_3 - \boldsymbol{\delta}_3) - \mathbf{B}_{3.1} (\overline{\mathbf{x}}_3^{(1)} - \overline{\mathbf{y}}_3^{(1)} - \boldsymbol{\delta}_1))' (\mathbf{C}_{3.1})^{-1} ((\hat{\boldsymbol{\delta}}_3 - \boldsymbol{\delta}_3) - \mathbf{B}_{3.1} (\overline{\mathbf{x}}_3^{(1)} - \overline{\mathbf{y}}_3^{(1)} - \boldsymbol{\delta}_1)) \\ &= Q_1 + Q_2 + Q_3, \text{ say.} \end{aligned} \tag{14}$$

The idea behind Q is as follows: if there are only $N_2\,(M_2)$ observations on the first p_1 components of $\mathbf{X}\,(\mathbf{Y})$, the appropriate statistic for hypothesis testing and confidence estimation of $(\delta_1', \delta_2')' = ((\mu_1 - \beta_1)', (\mu_1 - \beta_2)')'$ can be decomposed as the sum of two parts after some algebra:

$$\left[\left(\overline{\mathbf{x}}_2^{(1)} - \overline{\mathbf{y}}_2^{(1)}\right) - \boldsymbol{\delta}_1\right]' \left[\frac{\mathbf{S}_2^{(1,1)}}{n_2 N_2} + \frac{\mathbf{V}_2^{(1,1)}}{m_2 M_2}\right]^{-1} \cdot \left[\left(\overline{\mathbf{x}}_2^{(1)} - \overline{\mathbf{y}}_2^{(1)}\right) - \boldsymbol{\delta}_1\right] + Q_2. \tag{15}$$

Since there are additional observations on the first p_1 components, the first part above should be replaced by Q_1.

Similarly, If there are only the last $N_3\,(M_3)$ observations on the first p_1 components of $X\,(Y)$, the appropriate statistic for hypothesis testing and confidence estimation of $(\delta_1', \delta_3')'$ can be decomposed as the sum of two parts after some algebra:

$$\left[\left(\overline{\mathbf{x}}_3^{(1)} - \overline{\mathbf{y}}_3^{(1)}\right) - \boldsymbol{\delta}_1\right]' \left[\frac{\mathbf{S}_3^{(1,1)}}{n_3 N_3} + \frac{\mathbf{V}_3^{(1,1)}}{m_3 M_3}\right]^{-1} \cdot \left[\left(\overline{\mathbf{x}}_3^{(1)} - \overline{\mathbf{y}}_3^{(1)}\right) - \boldsymbol{\delta}_1\right] + Q_3. \tag{16}$$

Again, the first part should also be replaced by Q_1.

3.2. Hypothesis Test and Confidence Region for $\mu - \beta$. Because Q is resembling the Hotelling-T^2 statistic, and its distribution is free of any parameters, it is reasonable to approximate its distribution by the distribution of $dF_{p,\nu}$, where d is a positive constant and $F(a,b)$ denotes the F random variable with numerator degrees of freedom a and the denominator degrees of freedom b.

To find an approximation to the distribution of Q, we evaluated its first two approximate moments in the Appendix. Then, using the "moment matching" method, the distribution of Q is approximated by $dF_{p,\nu}$, where d is a positive constant, and $F_{a,b}$ denotes the F random variable with numerator degrees of freedom a and the denominator degrees of freedom b. The unknown constants d and ν can be determined so that the first two moments of Q are equal to those of $dF_{p,\nu}$. Using the modified Wishart approximation (see Lemma A.1 in Appendix) and following the lines of Krishnamoorthy and Pannala [6], we evaluated an approximation G_1 for $E(Q)$ and an approximation G_2 for $E(Q^2)$ in Appendix. To express G_1 and G_2, we need the following terms.

Let $\tilde{\mathbf{S}}_1 = (\mathbf{S}_1^{(1,1)}/(n_1 N_1)), \tilde{\mathbf{V}}_1 = (\mathbf{V}_1^{(1,1)}/(m_1 M_1)), \mathbf{C}_1 = \tilde{\mathbf{S}}_1 + \tilde{\mathbf{V}}_1$, and

$$f_1 = \frac{p_1 + p_1^2}{(1/n_1)\left\{\operatorname{tr}\left[\left(\tilde{\mathbf{S}}_1 \mathbf{C}_1^{-1}\right)^2\right] + \left[\operatorname{tr}\left(\tilde{\mathbf{S}}_1 \mathbf{C}_1^{-1}\right)\right]^2\right\} + (1/m_1)\left\{\operatorname{tr}\left[\left(\tilde{\mathbf{V}}_1 \mathbf{C}_1^{-1}\right)^2\right] + \left[\operatorname{tr}\left(\tilde{\mathbf{V}}_1 \mathbf{C}_1^{-1}\right)\right]^2\right\}}. \tag{17}$$

Let $\tilde{\mathbf{S}}_2 = (\mathbf{S}_2/(n_2 N_2))$, $\tilde{\mathbf{V}}_2 = (\mathbf{V}_2/(m_2 M_2))$, $\mathbf{C}_2 = \tilde{\mathbf{S}}_2 + \tilde{\mathbf{V}}_2$, and

$$f_2 = \frac{(p_1 + p_2) + (p_1 + p_2)^2}{(1/n_2)\left\{\operatorname{tr}\left[\left(\tilde{\mathbf{S}}_2 \mathbf{C}_2^{-1}\right)^2\right] + \left[\operatorname{tr}\left(\tilde{\mathbf{S}}_2 \mathbf{C}_2^{-1}\right)\right]^2\right\} + (1/m_2)\left\{\operatorname{tr}\left[\left(\tilde{\mathbf{V}}_2 \mathbf{C}_2^{-1}\right)^2\right] + \left[\operatorname{tr}\left(\tilde{\mathbf{V}}_2 \mathbf{C}_2^{-1}\right)\right]^2\right\}}. \tag{18}$$

Let $\tilde{\mathbf{S}}_3 = (\mathbf{S}_3/(n_3 N_3))$, $\tilde{\mathbf{V}}_3 = (\mathbf{V}_3/m_3 M_3)$, $\mathbf{C}_3 = \tilde{\mathbf{S}}_3 + \tilde{\mathbf{V}}_3$, and

$$f_3 = \frac{(p_1 + p_3) + (p_1 + p_3)^2}{(1/n_3)\left\{\operatorname{tr}\left[\left(\tilde{\mathbf{S}}_3 \mathbf{C}_3^{-1}\right)^2\right] + \left[\operatorname{tr}\left(\tilde{\mathbf{S}}_3 \mathbf{C}_3^{-1}\right)\right]^2\right\} + (1/m_3)\left\{\operatorname{tr}\left[\left(\tilde{\mathbf{V}}_3 \mathbf{C}_3^{-1}\right)^2\right] + \left[\operatorname{tr}\left(\tilde{\mathbf{V}}_3 \mathbf{C}_3^{-1}\right)\right]^2\right\}}. \tag{19}$$

In terms of the above quantities, we have

$$G_1 = \frac{f_1 p_1}{f_1 - p_1 - 1} + \frac{p_2 f_2 (f_2 - 1)}{(f_2 - p_1 - p_2 - 1)(f_2 - p_1 - 1)} + \frac{p_3 f_3 (f_3 - 1)}{(f_3 - p_1 - p_3 - 1)(f_3 - p_1 - 1)}$$

$$\begin{aligned} G_2 = {} & \frac{p_1 (p_1 + 2) f_1^2}{(f_1 - p_1 - 1)(f_1 - p_1 - 3)} + 2\left(\frac{f_1 p_1}{f_1 - p_1 - 1}\right)\left(\frac{p_2 f_2 (f_2 - 1)}{(f_2 - p_1 - p_2 - 1)(f_2 - p_1 - 1)}\right) \\ & + 2\left(\frac{f_1 p_1}{f_1 - p_1 - 1}\right)\left(\frac{p_3 f_3 (f_3 - 1)}{(f_3 - p_1 - p_3 - 1)(f_3 - p_1 - 1)}\right) \\ & + 2\left(\frac{p_2 f_2 (f_2 - 1)}{(f_2 - p_1 - p_2 - 1)(f_2 - p_1 - 1)}\right)\left(\frac{p_3 f_3 (f_3 - 1)}{(f_3 - p_1 - p_3 - 1)(f_3 - p_1 - 1)}\right) \\ & + \frac{p_2 (p_2 + 2) f_2^2 (f_2 - 1)(f_2 - 3)}{(f_2 - p_1 - p_2 - 1)(f_2 - p_1 - p_2 - 3)(f_2 - p_1 - 1)(f_2 - p_1 - 3)} \\ & + \frac{p_3 (p_3 + 2) f_3^2 (f_3 - 1)(f_3 - 3)}{(f_3 - p_1 - p_3 - 1)(f_3 - p_1 - p_3 - 3)(f_3 - p_1 - 1)(f_3 - p_1 - 3)}, \end{aligned} \tag{20}$$

and an approximation to the distribution of the pivotal quantity Q in (10) is given by $Q \sim \mathrm{d}F_{p,\nu}$ approximately, where

$$\nu = \frac{4pG_2 - 2(p+2)G_1^2}{pG_2 - (p+2)G_1^2}, \qquad d = G_1 \frac{\nu - 2}{\nu}. \tag{21}$$

Thus, for a given level α and an observed value Q_0 of Q, the null hypothesis that $\delta = \mu - \beta = 0$ will be rejected whenever the p value

$$P\left(Q_0 > \mathrm{d}F_{p,\nu} | H_0\right) < \alpha. \tag{22}$$

Furthermore, an approximate $1 - \alpha$ confidence set for $\mu - \beta$ is the set of values of δ that satisfy

$$Q \le \mathrm{d}F_{p,\nu}(1 - \alpha), \tag{23}$$

where Q is given in (9) and $F_{p,\nu}(1-\alpha)$ is the $(1-\alpha)$th quantile of the $F_{p,\nu}$ distribution.

4. Accuracy of the Approximations

We have used two approximations, one for approximating the sum of two Wishart matrices with different scale matrices and another for approximating the moments of Q to derive the distribution of Q. So, to understand the accuracy of the approximation, we estimated the sizes of the test for hypotheses in (4) when the nominal level is 0.05 using the Monte Carlo simulation.

To select the parameter configurations for Monte Carlo simulation, we note that the distribution of Q is location invariant, and so without loss of generality, we can assume that $\mu = \beta = 0$ to estimate the sizes. As pointed out in the study of Krishnamoorthy and Yu [9], we can also take Σ as a diagonal matrix with positive elements and Δ as a correlation matrix.

The estimated sizes are presented in Table 1 for the case of $p_1 = 2$, $p_2 = 1$, $p_3 = 1$, and a few selected sample sizes. The sample sizes are chosen so that the number of data missing is relatively small in some cases and large in other cases. It is clear from Table 1 that the coverage probabilities are very close to 0.95 for all the cases considered. In the worst situations, the coverage probabilities are around 0.93.

5. An Illustrative Example

We shall now illustrate the methods using "Fisher's Iris Data" which represent measurements of the sepal length and width and pedal 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 downloaded them from http://javeeh.net/sasintro/intro151.html. For illustration purpose, we use the data on virginica (x) and setosa (y). Since the sample size is large enough, we simply assume that the data are following approximately a multivariate normal distribution.

We created hierarchical patterns by discarding the last 15 measurements on x_3 (pedal length of virginica) and the first 35 measurements on x_4 (pedal width of virginica), the last 30 measurements on y_3 (pedal length of setosa) and the first 20 measurements on y_4 (pedal width of setosa). That is, we have $p_1 = 2, p_2 = 1, p_3 = 1$, and $(N_1, N_2, N_3) = (50, 15, 35)$, $(M_1, M_2, M_3) = (50, 20, 30)$. Let $\mu' = (\mu_1, \mu_2, \mu_3) =$ (average of the sepal length and width, pedal length, and pedal width) of virginica and $\beta' = (\beta_1, \beta_2, \beta_3) =$ (average of the sepal length and width, pedal length, and pedal width) of setosa. We want to test

Table 1: Monte Carlo estimates of the coverage probabilities of the confidence region; $\Sigma = \mathrm{diag}(\lambda_1, \lambda_2, \lambda_3)$, $\Delta = \begin{pmatrix} 1 & \rho_{12} & \rho_{13} \\ \rho_{12} & 1 & \rho_{23} \\ \rho_{13} & \rho_{23} & 1 \end{pmatrix}(p_1, p_2, p_3) = (1, 1, 1)$; $\alpha = 0.05$.

	$(N_1, N_2, N_3, M_1, M_2, M_3)$			
$(\lambda_1, \lambda_2, \lambda_3, \rho_{12}, \rho_{13}, \rho_{23})$	(13, 6, 14, 7)	(14, 6, 20, 12)	(18, 10, 20, 10)	(25, 13, 15, 8)
(1, 2, 3, 0.3, 0.2, −0.4)	0.954	0.949	0.952	0.956
(8, 2, 6, −0.4, −0.5, 0.3)	0.948	0.967	0.954	0.953
(0.4, 3, 9, 0.9, 0.3, −0.1)	0.948	0.941	0.947	0.962
(1, 1, 1, 0.2, 0.2, 0.2)	0.957	0.952	0.954	0.953
(3, 3, 3, 0.9, 0.9, 0.9)	0.946	0.942	0.949	0.951
(0.4, 0.3, 0.1, −0.4, −0.2, −0.1)	0.955	0.951	0.950	0.954
(2, 0.1, 12, 0.1, −0.9, −0.5)	0.952	0.954	0.949	0.944
(0.5, 0.9, 0.1, 0.9, 0.4, 0.5)	0.952	0.944	0.950	0.955
(0.6, 0.5, 0.1, −0.7, 0.5, 0.2)	0.954	0.950	0.952	0.954
(1, 25, 50, −0.5, 0.5, 0.1)	0.944	0.939	0.946	0.956
(23, 33, 55, 0.1, −0.1, 0.2)	0.943	0.936	0.954	0.951
(1, 1, 40, −0.4, 0.4, 0.2)	0.954	0.949	0.948	0.955

$$H_0\colon \boldsymbol{\mu} - \boldsymbol{\beta} = 0 \text{ vs. } H_0\colon \boldsymbol{\mu} - \boldsymbol{\beta} \neq 0. \tag{24}$$

After careful calculation, we get $Q_1 = 536.925$, $Q_2 = 1299.851$, and $Q_3 = 1121.327$, so $Q = 2958.103$. The required values to compute the critical value are $G_1 = E(Q) = 4.307$, $G_2 = E(Q^2) = 31.341$, $d = 4.250$, and $\nu = 150.759$. The critical value $dF_{p,\nu}(0.95) = 11.325$.

Since Q is much larger than the critical value, we have sufficient evidence to reject H_0 at 95% confidence level.

6. Concluding Remarks

In this article, we define hierarchical data missing pattern and point out that the strategy in many papers dealing with monotone missing data can be extended to deal with hierarchical missing data. To illustrate this, the multivariate Behrens–Fisher problem is considered. Based on the procedures due to Krishnamoorthy and Yu [9] dealing with the monotone missing data, we proposed a Hotelling T^2 type test for Behrens–Fisher problem. The test is simple to use, and the hierarchical patterns of the two samples are not necessarily the same.

As pointed out by two reviewers, this paper is based on multivariate normal population. Like what did in Batsidis [11–13] for monotone missing data, an extension of the results given in this paper for hierarchical missing data from elliptic distribution is an interesting open problem. Moreover, the proposed study can be extended for the neutrosophic statistics as future research. For details of neutrosophic statistics, see Aslam [15, 16] and Kashif et al. [17].

Appendix

The following two lemmas are needed to find approximate moments of Q in (14). In Lemma A.1, we propose the modified version of the Nel and van der Merwe [18] Wishart approximation given in Krishnamoorthy and Yu [19]. For a proof of Lemma A.2, see Seber [20]; p. 52.

Lemma A.1. *Let* $\mathbf{A}_1 \sim W_p(m_1 - 1, \Delta_1)$ *independently of* $\mathbf{A}_2 \sim W_p(m_2 - 1, \Delta_2)$. *Define*

$$\widetilde{\mathbf{A}}_i = \frac{\mathbf{A}_i}{m_i(m_i - 1)}, \qquad \widetilde{\boldsymbol{\Delta}}_i = \frac{\boldsymbol{\Delta}_i}{m_i}, \quad i = 1, 2. \tag{A.1}$$

Then

$$\widetilde{\mathbf{A}} = \widetilde{\mathbf{A}}_1 + \widetilde{\mathbf{A}}_2 \sim W_p\left(f, \frac{1}{f}\left(\widetilde{\boldsymbol{\Delta}}_1 + \widetilde{\boldsymbol{\Delta}}_2\right)\right) \text{approximately}, \tag{A.2}$$

where

$$f = \frac{p + p^2}{(1/(m_1 - 1))\left\{\mathrm{tr}\left[\left(\widetilde{\mathbf{A}}_1\widetilde{\mathbf{A}}^{-1}\right)^2\right] + \left[\mathrm{tr}\left(\widetilde{\mathbf{A}}_1\widetilde{\mathbf{A}}^{-1}\right)\right]^2\right\} + (1/(m_2 - 1))\left\{\mathrm{tr}\left[\left(\widetilde{\mathbf{A}}_2\widetilde{\mathbf{A}}^{-1}\right)^2\right] + \left[\mathrm{tr}\left(\widetilde{\mathbf{A}}_2\widetilde{\mathbf{A}}^{-1}\right)\right]^2\right\}}. \tag{A.3}$$

Lemma A.2. *Let* $\mathbf{A} \sim W_p(m, \Sigma)$ *independently of* $\mathbf{X} \sim N_p(\mu, \Sigma)$. *Write*

$$\mathbf{X} = \begin{pmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \end{pmatrix}, \quad \boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{pmatrix}, \quad \boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{21} \\ \boldsymbol{\Sigma}_{12} & \boldsymbol{\Sigma}_{22} \end{pmatrix}, \tag{A.4}$$

so that $\mathbf{X}_1$ *and* μ_1 *are of order* $p_1 \times 1$ *and* Σ_{11} *is of order* $p_1 \times p_1$. *Define* $\mathbf{X}_{2.1} = \mathbf{X}_2 - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{X}_1$, $\mu_{2.1} = \mu_2 - \Sigma_{21}\Sigma_{11}^{-1}\mu_1$, $\mathbf{A}_{2.1} = \mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12}$, *and* $\Sigma_{2.1} = \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}$. *Then,*

$$m\mathbf{X}'\mathbf{A}^{-1}\mathbf{X} = m\mathbf{X}_1'\mathbf{A}_{11}^{-1}\mathbf{X}_1 + m\mathbf{X}_{2.1}\mathbf{A}_{2.1}^{-1}\mathbf{X}_{2.1}, \tag{A.5}$$

and when $\mu_{2.1} = 0$,

$$\frac{\mathbf{X}_{2.1}'\mathbf{A}_{2.1}^{-1}\mathbf{X}_{2.1}}{1 + \mathbf{X}_1'\mathbf{A}_{11}^{-1}\mathbf{X}_1} \sim \frac{p_2}{m - p + 1} F_{p_2, m-p+1}. \tag{A.6}$$

The above statistic is independent of $\mathbf{X}_1'\mathbf{A}_{11}^{-1}\mathbf{X}_1$ whether $\mu_{2.1} = 0$ or not.

Let

$$\widetilde{\mathbf{S}}_1 = \frac{\mathbf{S}_1^{(1,1)}}{n_1 N_1}, \quad \widetilde{\mathbf{V}}_1 = \frac{\mathbf{V}_1^{(1,1)}}{m_1 M_1}. \tag{A.7}$$

Note that $\widetilde{\mathbf{S}}_1$ and $\widetilde{\mathbf{V}}_1$ are independent with

$$\widetilde{\mathbf{S}}_1 \sim W_{p_1}\left(n_1, \frac{\boldsymbol{\Sigma}_{11}}{n_1 N_1}\right), \quad \widetilde{\mathbf{V}}_1 \sim W_{p_1}\left(m_1, \frac{\boldsymbol{\Delta}_{11}}{m_1 M_1}\right), \tag{A.8}$$

and so using Lemma A.1, we have

$$\mathbf{C}_1 \sim W_{p_1}\left(f_1, \frac{1}{f_1}\left(\frac{\boldsymbol{\Sigma}_{11}}{N_1} + \frac{\boldsymbol{\Delta}_{11}}{M_1}\right)\right) \text{approximately}, \tag{A.9}$$

where f_1 is given in (17).

Since $\mathbf{C}_1$ and $\widehat{\delta}_1$ are independent, using (A.1), we have

$$Q_1 \sim \frac{p_1 f_1}{f_1 - p_1 + 1} F_{p_1, f_1 - p_1 + 1} \text{approximately}. \tag{A.10}$$

Define

$$Q_{2d} = \frac{1}{f_2}\left[\left(\overline{\mathbf{x}}_2^{(1)} - \overline{\mathbf{y}}_2^{(1)}\right) - \boldsymbol{\delta}_1\right]' \left[\frac{\mathbf{S}_2^{(1,1)}}{n_2 N_2} + \frac{\mathbf{V}_2^{(1,1)}}{m_2 M_2}\right]^{-1} \cdot \left[\left(\overline{\mathbf{x}}_2^{(1)} - \overline{\mathbf{y}}_2^{(1)}\right) - \boldsymbol{\delta}_1\right], \tag{A.11}$$

where f_2 is given in (18),

$$Q_{3d} = \frac{1}{f_3}\left[\left(\overline{\mathbf{x}}_3^{(1)} - \overline{\mathbf{y}}_3^{(1)}\right) - \boldsymbol{\delta}_1\right]' \left[\frac{\mathbf{S}_3^{(1,1)}}{n_3 N_3} + \frac{\mathbf{V}_3^{(1,1)}}{m_3 M_3}\right]^{-1} \cdot \left[\left[\overline{\mathbf{x}}_3^{(1)} - \overline{\mathbf{y}}_3^{(1)}\right] - \boldsymbol{\delta}_1\right], \tag{A.12}$$

where f_3 is given in (19), and

$$R_2 = \frac{Q_2}{(1 + Q_{2d})}, \quad R_3 = \frac{Q_3}{(1 + Q_{3d})}. \tag{A.13}$$

Using Lemma A.1, we have

$$\left[\frac{\mathbf{S}_2^{(1,1)}}{n_2 N_2} + \frac{\mathbf{V}_2^{(1,1)}}{m_2 M_2}\right] \sim W_{p_1}\left(f_2, \frac{1}{f_2}\left(\frac{\boldsymbol{\Sigma}_{11}}{N_2} + \frac{\boldsymbol{\Delta}_{11}}{M_2}\right)\right) \text{approximately}, \tag{A.14}$$

and independently of $(\overline{\mathbf{x}}_2^{(1)} - \overline{\mathbf{y}}_2^{(1)}) \sim N_p(\delta_1, ((\Sigma_{11}/N_2) + (\Delta_{11}/M_2)))$,

$$\left[\frac{\mathbf{S}_3^{(1,1)}}{n_3 N_3} + \frac{\mathbf{V}_3^{(1,1)}}{m_3 M_3}\right] \sim W_{p_1}\left(f_3, \frac{1}{f_3}\left(\frac{\boldsymbol{\Sigma}_{11}}{N_3} + \frac{\boldsymbol{\Delta}_{11}}{M_3}\right)\right) \text{approximately}, \tag{A.15}$$

and independently of $(\overline{\mathbf{x}}_3^{(1)} - \overline{\mathbf{y}}_3^{(1)}) \sim N_p(\delta_1, ((\Sigma_{11}/N_3) + (\Delta_{11}/M_3)))$.

Hence,

$$Q_{2d} \sim \frac{p_1}{f_2 - p_1 + 1} F_{p_1, f_2 - p_1 + 1}, \quad Q_{3d} \sim \frac{p_1}{f_3 - p_1 + 1} F_{p_1, f_3 - p_1 + 1} \text{approximately}. \tag{A.16}$$

Using Lemma A.2, we have

$$R_2 \sim \frac{f_2 p_2}{f_2 - p_1 - p_2 + 1} F_{p_2, f_2 - p_1 - p_2 + 1} \text{approximately}, \tag{A.17}$$

And it is independent of Q_2 and Q_{2d}.

$$R_3 \sim \frac{f_3 p_3}{f_3 - p_1 - p_3 + 1} F_{p_3, f_2 - p_1 - p_3 + 1} \text{approximately}, \tag{A.18}$$

and it is independent of Q_2 and Q_{2d}.

Using the above approximate distributional results and treating f_i's as constants, we evaluate the following moments:

$$\begin{aligned}
E(Q_1) &\simeq \frac{f_1 p_1}{f_1 - p_1 - 1},\\
E(Q_2) &= E(R_2(1+Q_{2d})) = E(R_2)E(1+Q_{2d}) \simeq \frac{p_2 f_2 (f_2-1)}{(f_2-p_1-p_2-1)(f_2-p_1-1)},\\
E(Q_3) &= E(R_3(1+Q_{3d})) = E(R_3)E(1+Q_{3d}) \simeq \frac{p_3 f_3 (f_3-1)}{(f_3-p_1-p_3-1)(f_3-p_1-1)}\\
E\left(Q_1^2\right) &\simeq \frac{p_1(p_1+2)f_1^2}{(f_1-p_1-1)(f_1-p_1-3)},\\
E\left(Q_2^2\right) &\simeq \frac{p_2(p_2+2)f_2^2(f_2-1)(f_2-3)}{(f_2-p_1-p_2-1)(f_2-p_1-p_2-3)(f_2-p_1-1)(f_2-p_1-3)},\\
E\left(Q_3^2\right) &\simeq \frac{p_3(p_3+2)f_3^2(f_3-1)(f_3-3)}{(f_3-p_1-p_3-1)(f_3-p_1-p_3-3)(f_3-p_1-1)(f_3-p_1-3)}.
\end{aligned} \tag{A.19}$$

Using the arguments of Krishnamoorthy and Pannala [6], it can be shown that $E(Q_1Q_2) \simeq E(Q_1)E(Q_2), E(Q_1Q_3) \simeq E(Q_1)E(Q_3), E(Q_3Q_2) \simeq E(Q_3)E(Q_2)$. Thus, we have

$$\begin{aligned}
E(Q) &= E(Q_1) + E(Q_2) + E(Q_3) = G_1,\\
E\left(Q^2\right) &\simeq E\left(Q_1^2\right) + E\left(Q_2^2\right) + E\left(Q_3^2\right) + 2E(Q_1)E(Q_2) + 2E(Q_1)E(Q_3) + 2E(Q_3)E(Q_2) = G_2.
\end{aligned} \tag{A.20}$$

Data Availability

The data sets can be downloaded freely from http://javeeh.net/sasintro/intro151.html.

Conflicts of Interest

The author declares no conflicts of interest.

References

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

[2] R. J. A. Little, “A test of missing completely at random for multivariate data with missing values,” *Journal of the American Statistical Association*, vol. 83, no. 404, pp. 1198–1202, 1988.

[3] G. B. Lu and J. B. Copas, “Missing at random, likelihood ignorability and model completeness,” *Annals of Statistics*, vol. 32, pp. 754–765, 2004.

[4] T. W. Anderson, “Maximum likelihood estimates for a multivariate normal distribution when some observations are missing,” *Journal of the American Statistical Association*, vol. 52, no. 278, pp. 200–203, 1957.

[5] K. Krishnamoorthy and M. K. Pannala, “Some simple test procedures for normal mean vector with incomplete data,” *Annals of the Institute of Statistical Mathematics*, vol. 50, no. 3, pp. 531–542, 1998.

[6] K. Krishnamoorthy and M. K. Pannala, “Confidence estimation of a normal mean vector with incomplete data,” *Canadian Journal of Statistics*, vol. 27, no. 2, pp. 395–407, 1999.

[7] J. Hao and K. Krishnamoorthy, “Inferences on a normal covariance matrix and generalized variance with monotone missing data,” *Journal of Multivariate Analysis*, vol. 78, no. 1, pp. 62–82, 2001.

[8] J. Yu, K. Krishnamoorthy, and M. K. Pannala, “Two-sample inference for normal mean vectors based on monotone missing data,” *Journal of Multivariate Analysis*, vol. 97, no. 10, pp. 2162–2176, 2006.

[9] K. Krishnamoorthy and J. Yu, “Multivariate behrens-fisher problem with missing data,” *Journal of Multivariate Analysis*, vol. 105, no. 1, pp. 141–150, 2012.

[10] J. Yu, K. Krishnamoorthy, and Y. He, “Testing equality of two normal covariance matrices with monotone missing data,” *Communications in Statistics - Theory and Methods*, vol. 49, no. 16, pp. 3911–3918, 2019.

[11] A. Batsidis and K. Zografos, “Statistical inference for location and scale of elliptically contoured models with monotone missing data,” *Journal of Statistical Planning and Inference*, vol. 136, no. 8, pp. 2606–2629, 2006.

[12] A. Batsidis and K. Zografos, “Multivariate linear regression model with elliptically contoured distributed errors and monotone missing dependent variables,” *Communications in Statistics - Theory and Methods*, vol. 37, no. 3, pp. 349–372, 2008.

[13] A. Batsidis, “Robustness of the likelihood ratio test for detection and estimation of a mean change point in a sequence of elliptically contoured observations,” *Statistics*, vol. 44, no. 1, pp. 17–24, 2010.

[14] K. T. Fang and Y. T. Zhang, *Generalized Multivariate Analysis*, Science Press Beijing and Springer, Berlin, Germany, 1990.
[15] M. Aslam, "Neutrosophic analysis of variance: application to university students," *Complex & Intelligent Systems*, vol. 5, no. 4, p. 403, 2019.
[16] M. Aslam, "A new goodness of fit test in the presence of uncertain parameters," *Complex & Intelligent Systems*, 2020.
[17] M. Kashif, H. Nida, M. Imran Khan, and M. Aslam, "Decomposition of matrix under neutrosophic environment," *Neutrosophic Sets and Systems*, vol. 30, pp. 143–148, 2019.
[18] D. G. Nel and C. A. Van Der Merwe, "A solution to the multivariate behrens-fisher problem," *Communications in Statistics - Theory and Methods*, vol. 15, no. 12, pp. 3719–3735, 1986.
[19] K. Krishnamoorthy and J. Yu, "Modified nel and van der merwe test for the multivariate Behrens-Fisher problem," *Statistics & Probability Letters*, vol. 66, no. 2, pp. 161–169, 2004.
[20] G. A. F. Seber, *Multivariate Observations*, Wiley, New York, NY, USA, 1984.