7-5-2013

Inferences about Parameters of Trivariate Normal Distribution with Missing Data

Xing Wang
Florida International University, xwang019@fiu.edu

FLORIDA INTERNATIONAL UNIVERSITY

Miami, Florida

INFERENCES ABOUT PARAMETERS OF TRIVARIATE

NORMAL DISTRIBUTION WITH MISSING DATA

A thesis submitted in partial fulfillment of the

requirements for the degree of

MASTER OF SCIENCE

in

STATISTICS

by

Xing Wang

2013

To: Dean Kenneth Furton
    College of Arts and Sciences

This thesis, written by Xing Wang, and entitled Inferences about Parameters
of Trivariate Normal Distribution with Missing Data, having been approved
in respect to style and intellectual content, is referred to you for judgment.

We have read this thesis and recommend that it be approved.

_____
Florence George

_____
Kai Huang, Co-Major Professor

_____
Jie Mi, Co-Major Professor

Date of Defense: July 5, 2013

The thesis of Xing Wang is approved.

_____
Dean Kenneth Furton
College of Arts and Sciences

_____
Dean Lakshmi N. Reddi
University Graduate School

Florida International University, 2013

ABSTRACT OF THE THESIS

INFERENCES ABOUT PARAMETERS OF TRIVARIATE

NORMAL DISTRIBUTION WITH MISSING DATA

by

Xing Wang

Florida International University, 2013

Miami, Florida

Professor Jie Mi, Co-Major Professor

Professor Kai Huang, Co-Major Professor

Multivariate normal distribution is commonly encountered in any field, a frequent issue is the missing values in practice. The purpose of this research was to estimate the parameters in three-dimensional covariance permutation-symmetric normal distribution with complete data and all possible patterns of incomplete data. In this study, MLE with missing data were derived, and the properties of the MLE as well as the sampling distributions were obtained. A Monte Carlo simulation study was used to evaluate the performance of the considered estimators for both cases when ρ was known and unknown. All results indicated that, compared to estimators in the case of omitting observations with missing data, the estimators derived in this article led to better performance. Furthermore, when ρ was unknown, using the estimate of ρ would lead to the same conclusion.

Keywords: Trivariate Normal Distribution, Permutation-Symmetric Covariance, Missing Data, MLE.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

I. Introduction

In practice, the normal distribution is commonly encountered, since the sampling distribution of many multivariate statistics are proximately normal because of the a central limit effect. Early application of the multivariate normal distribution was with regard to biological studies. For instance, we often think that height and weight, when observed on the same individual, approximately follow bivariate normal distribution. We could extend this to the foot size, or any other variable of related physical characteristics, and these measurements together could follow the multivariate normal distribution.

Moreover, we often have to analyze data that contains missing values in practice. Incomplete normal data could arise in any number of scientific investigations: early detection of diseases, wildlife survey research, mental health research, and so on. Missing data are one of the most persuasive problems for analysis of data which can occur for a variety of reasons. For example, a participant may refuse to answer some of the questions; the use of new instruments results in incomplete historical data; some information may be purposely excised to protect confidentiality.

Estimation of parameters of a multivariate normal distribution when data are incomplete has been discussed by many authors. A systematic approach to missing values problem was derived using likelihoods of observed values. Wilks (1932) considered MLEs for a bivariate normal population with missing data in both variables. Srivastava and Zaatar (1973) used Monte Carlo simulation to compare four estimators of the covariance matrix of a bivariate normal distribution. Edgett (1956) gave maximum likelihood estimates of parameters of a trivariate normal distribution when observations on only one variable are missing. Lord (1955) and Matthai (1951) also found estimates of parameters

of a trivariate normal distribution in other special cases. Anderson (1957) indicated how one could obtain the Maximum Likelihood Estimates when the sample was monotone. Hocking and Smith (1968) developed a method of estimating parameters of a p-variate normal distribution with zero mean vector in which the missing observations are not required to follow certain patterns. Their estimation technique could be summarized as follows:

(a) The data were divided into groups according to which variates were missing.

(b) Initial estimates of the parameters were obtained from that group of observations with no missing variates.

(c) These initial estimators were modified by adjoining, optimally, the information in the remaining groups in a sequential manner until all data was used. Hocking and Marx (1979) used the same method as Hocking and Smith (1968) to derive estimates, but their use of matrices simplified the notation and gave the estimates in a form that was easily implemented on a computer. They also gave exact small sample moments of the estimators for the case of two data groups.

The purpose of my research is to estimate the parameters in three-dimensional covariance permutation-symmetric normal distribution with complete data and all possible patterns of incomplete data, and then to study the properties of the estimators. It is assumed that all correlation coefficients and variances are equal, which means that we focus on the covariance permutation-symmetric trivariate normal distribution. The special case of the covariance permutation-symmetric model are the exchangeable normal variables, which make their appearance in many statistical applications. For instance, in the Bayes theory concerning normal observations, it is generally assumed that the prior distribution

of θ, the population mean, has an $N(\mu, \omega^2)$ distribution for some μ and $\omega^2 > 0$. Thus, for given θ, if the random variables are (conditionally) $i.i.d$ normal variables, then the marginal distribution of the data variable $\boldsymbol{X} = (X_1, ..., X_n)'$ is a mixture. In this case $X_1, ..., X_n$ are exchangeable normal variables.

The thesis is organized as follows. The Maximum Likelihood Estimates of the parameters of trivariate normal distribution with complete data are reviewed in Section 2. Maximum Likelihood Estimates with missing data will be derived in Section 3. In Section 4, the properties of the MLEs will be obtained, followed by the sampling distributions in Section 5. The numerical studies based on Monte Carlo simulations are considered in Section 6, which include comparison of

(a) confidence regions of $\boldsymbol{\mu}$;

(b) the probability of Type I error, and power of testing $H_0 : \boldsymbol{\mu} = 0$ vs $H_a : \boldsymbol{\mu} \neq 0$;

(c) coverage probability and average width of confidence intervals with regard to μ_1;

(d) coverage probability and average width of confidence intervals with regard to $\mu_2 - \mu_1$ using i) only three-dimensional observations and ii) both three-dimensional observations and all possible incomplete observations. Finally, in Section 7, the simulation will follow the same procedure in Section 6 but based on unknown ρ.

II. Review of MLE with complete data

In order to obtain the trivariate normal likelihood function, let us assume that the 3×1 vectors $\boldsymbol{X_1}, \boldsymbol{X_2}, ..., \boldsymbol{X_n}$ represent a random sample from a three-dimensional multivariate normal population with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$, where $\boldsymbol{X_i} = (x_{i1}, x_{i2}, x_{i3})', i = 1, 2, ..., n$, $\boldsymbol{\mu} = (\mu_1, \mu_2, \mu_3)'$

and

$$\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_{22} & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_{33} \end{pmatrix}$$

Since $\boldsymbol{X_1}, \boldsymbol{X_2}, ..., \boldsymbol{X_n}$ are mutually independent and each has a distribution $N_3(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, the joint density function of all the observations is the product of the marginal normal densities:

$$
\begin{aligned}
f(\boldsymbol{X_1}, \boldsymbol{X_2}, ..., \boldsymbol{X_n}) &= \prod_{i=1}^{n} \{ \frac{1}{(2\pi)^{3/2}|\boldsymbol{\Sigma}|^{1/2}} exp[-\frac{(\boldsymbol{X_i} - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{X_i} - \boldsymbol{\mu})}{2}] \} \\
&= \frac{1}{(2\pi)^{3n/2}|\boldsymbol{\Sigma}|^{n/2}} exp[-\frac{\sum_{i=1}^{n}(\boldsymbol{X_i} - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{X_i} - \boldsymbol{\mu})}{2}]
\end{aligned}
$$

It has been derived that

$$\hat{\boldsymbol{\mu}} = \bar{\boldsymbol{X}} \tag{1}$$

and

$$\hat{\boldsymbol{\Sigma}} = \frac{\sum_{i=1}^{n}(\boldsymbol{X_i} - \bar{\boldsymbol{X}})(\boldsymbol{X_i} - \bar{\boldsymbol{X}})'}{n} = \frac{(n-1)S}{n} \tag{2}$$

are the maximum likelihood estimates of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ respectively, where

$$\bar{\boldsymbol{X}} = \frac{1}{n}\sum_{i=1}^{n} \bar{\boldsymbol{X_i}}, i = 1, 2, ..., n$$

4

and
$$S = \frac{\sum_{i=1}^{n}(\boldsymbol{X_i} - \bar{\boldsymbol{X}})(\boldsymbol{X_i} - \bar{\boldsymbol{X}})'}{n - 1}.$$

III. The Maximum Likelihood Estimates with Missing Data

1. Notation and Assumption

In the present paper, we consider the covariance permutation-symmetric trivariate normal distribution, which means that $\sigma_1^2 = \sigma_2^2 = \sigma_3^2 \equiv \sigma^2$ and all correlation coefficients are equal: $\rho_{12} = \rho_{13} = \rho_{23} \equiv \rho$. By denoting $\tau = \sigma^2$, the marginal density function can be expressed as:

$$f(x_1, x_2, x_3) = \frac{e^{-w/[2\tau(2\rho^2 - \rho - 1)]}}{2\sqrt{2}\pi^{3/2}\tau^{3/2}\sqrt{1 - 3\rho^2 + 2\rho^3}} \tag{3}$$

where

$$
\begin{aligned}
w &= -(\rho + 1)\left[(x_1 - \mu_1)^2 + (x_2 - \mu_2)^2 + (x_3 - \mu_3)^2\right] \\
&\quad + 2\rho\left[(x_1 - \mu_1)(x_2 - \mu_2) + (x_1 - \mu_1)(x_3 - \mu_3) + (x_2 - \mu_2)(x_3 - \mu_3)\right]
\end{aligned}
$$

The data in our study are divided into groups according to which variables are missing, and initial estimates of the parameters are obtained from that group of observations with no missing variables. Specifically, it is assumed that a sample of size $n_0 + n_{12} + n_{13} + n_{23} + n_1 + n_2 + n_3$ is taken from a three-dimensional normal distribution but some of the observations have randomly occurring missing entries. In this case, the observations can be placed in 7 groups. Specifically, it means that the data consist of n_0 complete observations $\{(x_{i1}, x_{i2}, x_{i3}), i = 1, \cdots, n_0\}$ on $\boldsymbol{X} = (\boldsymbol{X_1}, \boldsymbol{X_2}, \boldsymbol{X_3})$, n_{12} observations $\{(x_{i1}^{(12)}, x_{i2}^{(12)}), i = 1, \cdots, n_{12}\}$ on $(\boldsymbol{X_1}, \boldsymbol{X_2})$, n_{13} observations $\{(x_{i1}^{(13)}, x_{i3}^{(13)}), i = 1, \cdots, n_{13}\}$ on $(\boldsymbol{X_1}, \boldsymbol{X_3})$, n_{23} observations $\{(x_{i2}^{(23)}, x_{i3}^{(23)}), i = 1, \cdots, n_{23}\}$ on $(\boldsymbol{X_2}, \boldsymbol{X_3})$, n_1 observations $\{x_{i1}^{(1)}, i = 1, \cdots, n_1\}$ on $\boldsymbol{X_1}$, n_2 observations $\{x_{i2}^{(2)}, i = 1, \cdots, n_2\}$ on $\boldsymbol{X_2}$, n_3 observations $\{x_{i3}^{(3)}, i = 1, \cdots, n_3\}$ on $\boldsymbol{X_3}$.

2. Likelihood Function

Following Hocking and Smith(1968), the likelihood function can be written as:

$$
\begin{aligned}
L(\boldsymbol{\mu}, \tau) &= L_0(\boldsymbol{\mu}, \tau) \cdot L_{12}(\mu_1, \mu_2, \tau) \cdot L_{13}(\mu_1, \mu_3, \tau) \cdot L_{23}(\mu_2, \mu_3, \tau) \cdot L_1(\mu_1, \tau) \cdot \\
& \quad L_2(\mu_2, \tau) \cdot L_3(\mu_3, \tau) \\
&= \prod_{i=1}^{n_0} f(\underline{x}_i; \boldsymbol{\mu}, \tau) \prod_{i=1}^{n_{12}} f(\underline{x}_i^{(12)}; \mu_1, \mu_2, \tau) \prod_{i=1}^{n_{13}} f(\underline{x}_i^{(13)}; \mu_1, \mu_3, \tau) \\
& \quad \prod_{i=1}^{n_{23}} f(\underline{x}_i^{(23)}; \mu_2, \mu_3, \tau) \prod_{i=1}^{n_1} f(x_i^{(1)}; \mu_1, \tau) \prod_{i=1}^{n_2} f(x_i^{(2)}; \mu_2, \tau) \prod_{i=1}^{n_3} f(x_i^{(3)}; \mu_3, \tau)
\end{aligned}
$$

$$(4)$$

Where $\underline{x}_i = (x_{i1}, x_{i2}, x_{i3})'$, $i = 1, \cdots, n_0$; $\underline{x}_i^{(12)} = (x_{i1}^{(12)}, x_{i2}^{(12)})'$, $i = 1, \cdots, n_{12}$;

$\underline{x}_i^{(13)} = (x_{i1}^{(13)}, x_{i3}^{(13)})'$, $i = 1, \cdots, n_{13}$; $\underline{x}_i^{(23)} = (x_{i2}^{(23)}, x_{i3}^{(23)})'$, $i = 1, \cdots, n_{23}$;

$x_i^{(1)} = x_{i1}^{(1)}$, $i = 1, \cdots, n_1$; $x_i^{(2)} = x_{i2}^{(2)}$, $i = 1, \cdots, n_2$; $x_i^{(3)} = x_{i3}^{(3)}$, $i = 1, \cdots, n_3$

L_0 is used to denote the likelihood function for the three-dimensional complete observations, where

$$
L_0(\boldsymbol{\mu}, \tau) = \frac{e^{-\sum_{i=1}^{n_0} w_{0i}/[2\tau(2\rho^2 - \rho - 1)]}}{(2\sqrt{2}\pi^{3/2}\tau^{3/2}\sqrt{1 - 3\rho^2 + 2\rho^3})^{n_0}}
$$

with

$$
\begin{aligned}
w_{0i} &= -(\rho + 1)\left[(x_{i1} - \mu_1)^2 + (x_{i2} - \mu_2)^2 + (x_{i3} - \mu_3)^2\right] \\
& \quad + 2\rho\left[(x_{i1} - \mu_1)(x_{i2} - \mu_2) + (x_{i1} - \mu_1)(x_{i3} - \mu_3) + (x_{i2} - \mu_2)(x_{i3} - \mu_3)\right]
\end{aligned}
$$

Similarly, other notations are defined as follows:

$$
L_{12}(\mu_1, \mu_2, \tau) = \frac{e^{-[\sum_{i=1}^{n_{12}} (x_{i1}^{(12)} - \mu_1)^2 - 2\rho \sum_{i=1}^{n_{12}} (x_{i1}^{(12)} - \mu_1)(x_{i2}^{(12)} - \mu_2) + \sum_{i=1}^{n_{12}} (x_{i2}^{(12)} - \mu_2)^2]/[2\tau(1 - \rho^2)]}}{(2\pi)^{n_{12}} \tau^{n_{12}} (1 - \rho^2)^{n_{12}/2}}
$$

$$L_{13}(\mu_1, \mu_3, \tau) = \frac{e^{-[\sum_{i=1}^{n_{13}} (x_{i1}^{(13)} - \mu_1)^2 - 2\rho \sum_{i=1}^{n_{13}} (x_{i1}^{(13)} - \mu_1)(x_{i3}^{(13)} - \mu_3) + \sum_{i=1}^{n_{13}} (x_{i3}^{(13)} - \mu_3)^2]/[2\tau(1-\rho^2)]}}{(2\pi)^{n_{13}} \tau^{n_{13}} (1 - \rho^2)^{n_{13}/2}}$$

$$L_{23}(\mu_2, \mu_3, \tau) = \frac{e^{-[\sum_{i=1}^{n_{23}} (x_{i2}^{(23)} - \mu_2)^2 - 2\rho \sum_{i=1}^{n_{23}} (x_{i2}^{(23)} - \mu_2)(x_{i3}^{(23)} - \mu_3) + \sum_{i=1}^{n_{23}} (x_{i3}^{(23)} - \mu_3)^2]/[2\tau(1-\rho^2)]}}{(2\pi)^{n_{23}} \tau^{n_{23}} (1 - \rho^2)^{n_{23}/2}}$$

$$L_1(\mu_1, \tau) = \frac{e^{-\sum_{i=1}^{n_1} (x_{i1}^{(1)} - \mu_1)^2/(2\tau)}}{(2\pi)^{n_1/2} \tau^{n_1/2}}$$

$$L_2(\mu_2, \tau) = \frac{e^{-\sum_{i=1}^{n_2} (x_{i2}^{(2)} - \mu_2)^2/(2\tau)}}{(2\pi)^{n_2/2} \tau^{n_2/2}}$$

$$L_3(\mu_3, \tau) = \frac{e^{-\sum_{i=1}^{n_3} (x_{i3}^{(3)} - \mu_3)^2/(2\tau)}}{(2\pi)^{n_3/2} \tau^{n_3/2}}$$

3. Derivation of $\hat{\mu}$ and $\hat{\sigma}^2$

In the rest of this study it is assumed that $\rho \in (-1/2, 1)$ is known. In order to apply likelihood-based method to obtain the MLEs $\hat{\boldsymbol{\mu}}$ and $\hat{\sigma}^2$ of $\boldsymbol{\mu}$ and σ^2, the natural logarithms of the above likelihood functions are taken:

$$
\begin{aligned}
lnL_0(\boldsymbol{\mu}, \tau) &= -\frac{3n_0}{2} ln(2\pi) - \frac{3n_0}{2} ln\tau - \frac{n_0}{2} ln(1 - 3\rho^2 + 2\rho^3) \\
&\quad + \frac{\rho + 1}{2\tau(2\rho^2 - \rho - 1)} \left[\sum_{i=1}^{n_0} (x_{i1} - \mu_1)^2 + \sum_{i=1}^{n_0} (x_{i2} - \mu_2)^2 + \sum_{i=1}^{n_0} (x_{i3} - \mu_3)^2 \right] \\
&\quad - \frac{\rho}{\tau(2\rho^2 - \rho - 1)} \left[\sum_{i=1}^{n_0} (x_{i1} - \mu_1)(x_{i2} - \mu_2) + \sum_{i=1}^{n_0} (x_{i1} - \mu_1)(x_{i3} - \mu_3) \right. \\
&\quad \left. + \sum_{i=1}^{n_0} (x_{i2} - \mu_2)(x_{i3} - \mu_3) \right]
\end{aligned}
$$

$$lnL_{12}(\mu_1, \mu_2, \tau) = -n_{12} ln(2\pi) - n_{12} ln\tau - \frac{n_{12}}{2} ln(1 - \rho^2)$$

8

$$-\frac{1}{2\tau(1-\rho^2)}[\sum_{i=1}^{n_{12}}(x_{i1}^{(12)}-\mu_1)^2 - 2\rho\sum_{i=1}^{n_{12}}(x_{i1}^{(12)}-\mu_1)(x_{i2}^{(12)}-\mu_2)$$

$$+\sum_{i=1}^{n_{12}}(x_{i2}^{(12)}-\mu_2)^2]$$

$$lnL_{13}(\mu_1,\mu_3,\tau) = -n_{13}ln(2\pi) - n_{13}ln\tau - \frac{n_{13}}{2}ln(1-\rho^2)$$

$$-\frac{1}{2\tau(1-\rho^2)}[\sum_{i=1}^{n_{13}}(x_{i1}^{(13)}-\mu_1)^2 - 2\rho\sum_{i=1}^{n_{13}}(x_{i1}^{(13)}-\mu_1)(x_{i3}^{(13)}-\mu_3)$$

$$+\sum_{i=1}^{n_{13}}(x_{i3}^{(13)}-\mu_3)^2]$$

$$lnL_{23}(\mu_2,\mu_3,\tau) = -n_{23}ln(2\pi) - n_{23}ln\tau - \frac{n_{23}}{2}ln(1-\rho^2)$$

$$-\frac{1}{2\tau(1-\rho^2)}[\sum_{i=1}^{n_{23}}(x_{i2}^{(23)}-\mu_2)^2 - 2\rho\sum_{i=1}^{n_{23}}(x_{i2}^{(23)}-\mu_2)(x_{i3}^{(23)}-\mu_3)$$

$$+\sum_{i=1}^{n_{23}}(x_{i3}^{(23)}-\mu_3)^2]$$

$$lnL_1(\mu_1,\tau) = -\frac{n_1}{2}ln(2\pi) - \frac{n_1}{2}ln\tau - \frac{1}{2\tau}\sum_{i=1}^{n_1}(x_{i1}^{(1)}-\mu_1)^2$$

$$lnL_2(\mu_2,\tau) = -\frac{n_2}{2}ln(2\pi) - \frac{n_2}{2}ln\tau - \frac{1}{2\tau}\sum_{i=1}^{n_2}(x_{i2}^{(2)}-\mu_2)^2$$

$$lnL_3(\mu_3,\tau) = -\frac{n_3}{2}ln(2\pi) - \frac{n_3}{2}ln\tau - \frac{1}{2\tau}\sum_{i=1}^{n_3}(x_{i3}^{(3)}-\mu_3)^2$$

Combining all the above expressions together yields the log-likelihood function (5):

$$lnL(\boldsymbol{\mu},\tau) = -\frac{3n_0}{2}ln(2\pi) - \frac{3n_0}{2}ln\tau - \frac{n_0}{2}ln(1-3\rho^2+2\rho^3) - n_{12}ln(2\pi) - n_{12}ln\tau$$

$$-\frac{n_{12}}{2}ln(1-\rho^2) - n_{13}ln(2\pi) - n_{13}ln\tau - \frac{n_{13}}{2}ln(1-\rho^2) - n_{23}ln(2\pi)$$

$$-n_{23}ln\tau - \frac{n_{23}}{2}ln(1-\rho^2) - \frac{n_1}{2}ln(2\pi) - \frac{n_1}{2}ln\tau - \frac{n_2}{2}ln(2\pi)$$

$$-\frac{n_2}{2}ln\tau - \frac{n_3}{2}ln(2\pi) - \frac{n_3}{2}ln\tau$$

$$+\frac{\rho+1}{2\tau(2\rho^2-\rho-1)}[\sum_{i=1}^{n_0}(x_{i1}-\mu_1)^2+\sum_{i=1}^{n_0}(x_{i2}-\mu_2)^2+\sum_{i=1}^{n_0}(x_{i3}-\mu_3)^2]$$

$$-\frac{\rho}{\tau(2\rho^2-\rho-1)}[\sum_{i=1}^{n_0}(x_{i1}-\mu_1)(x_{i2}-\mu_2)+\sum_{i=1}^{n_0}(x_{i1}-\mu_1)(x_{i3}-\mu_3)$$

$$+\sum_{i=1}^{n_0}(x_{i2}-\mu_2)(x_{i3}-\mu_3)]-\frac{1}{2\tau(1-\rho^2)}[\sum_{i=1}^{n_{12}}(x_{i1}^{(12)}-\mu_1)^2$$

$$-2\rho\sum_{i=1}^{n_{12}}(x_{i1}^{(12)}-\mu_1)(x_{i2}^{(12)}-\mu_2)+\sum_{i=1}^{n_{12}}(x_{i2}^{(12)}-\mu_2)^2]$$

$$-\frac{1}{2\tau(1-\rho^2)}[\sum_{i=1}^{n_{13}}(x_{i1}^{(13)}-\mu_1)^2-2\rho\sum_{i=1}^{n_{13}}(x_{i1}^{(13)}-\mu_1)(x_{i3}^{(13)}-\mu_3)$$

$$+\sum_{i=1}^{n_{13}}(x_{i3}^{(13)}-\mu_3)^2]-\frac{1}{2\tau(1-\rho^2)}[\sum_{i=1}^{n_{23}}(x_{i2}^{(23)}-\mu_2)^2$$

$$-2\rho\sum_{i=1}^{n_{23}}(x_{i2}^{(23)}-\mu_2)(x_{i3}^{(23)}-\mu_3)+\sum_{i=1}^{n_{23}}(x_{i3}^{(23)}-\mu_3)^2]$$

$$-\frac{1}{2\tau}\sum_{i=1}^{n_1}(x_{i1}^{(1)}-\mu_1)^2-\frac{1}{2\tau}\sum_{i=1}^{n_2}(x_{i2}^{(2)}-\mu_2)^2-\frac{1}{2\tau}\sum_{i=1}^{n_3}(x_{i3}^{(3)}-\mu_3)^2$$

$$=\ c_0-\left(\frac{3n_0}{2}+n_{12}+n_{13}+n_{23}+\frac{n_1}{2}+\frac{n_2}{2}+\frac{n_3}{2}\right)ln\tau$$

$$+\frac{\rho+1}{2\tau(2\rho^2-\rho-1)}[\sum_{i=1}^{n_0}(x_{i1}-\mu_1)^2+\sum_{i=1}^{n_0}(x_{i2}-\mu_2)^2+\sum_{i=1}^{n_0}(x_{i3}-\mu_3)^2]$$

$$-\frac{1}{2\tau(1-\rho^2)}[\sum_{i=1}^{n_{12}}(x_{i1}^{(12)}-\mu_1)^2+\sum_{i=1}^{n_{12}}(x_{i2}^{(12)}-\mu_2)^2]$$

$$-\frac{1}{2\tau(1-\rho^2)}[\sum_{i=1}^{n_{13}}(x_{i1}^{(13)}-\mu_1)^2+\sum_{i=1}^{n_{13}}(x_{i3}^{(13)}-\mu_3)^2]$$

$$-\frac{1}{2\tau(1-\rho^2)}[\sum_{i=1}^{n_{23}}(x_{i2}^{(23)}-\mu_2)^2+\sum_{i=1}^{n_{23}}(x_{i3}^{(23)}-\mu_3)^2]$$

$$-\frac{1}{2\tau}\sum_{i=1}^{n_1}(x_{i1}^{(1)}-\mu_1)^2-\frac{1}{2\tau}\sum_{i=1}^{n_2}(x_{i2}^{(2)}-\mu_2)^2-\frac{1}{2\tau}\sum_{i=1}^{n_3}(x_{i3}^{(3)}-\mu_3)^2$$

$$-\frac{\rho}{\tau(2\rho^2-\rho-1)}[\sum_{i=1}^{n_0}(x_{i1}-\mu_1)(x_{i2}-\mu_2)+\sum_{i=1}^{n_0}(x_{i1}-\mu_1)(x_{i3}-\mu_3)$$

$$+\sum_{i=1}^{n_0}(x_{i2}-\mu_2)(x_{i3}-\mu_3)]+\frac{\rho}{\tau(1-\rho^2)}[\sum_{i=1}^{n_{12}}(x_{i1}^{(12)}-\mu_1)(x_{i2}^{(12)}-\mu_2)$$

$$+\sum_{i=1}^{n_{13}}(x_{i1}^{(13)}-\mu_1)(x_{i3}^{(13)}-\mu_3)+\sum_{i=1}^{n_{23}}(x_{i2}^{(23)}-\mu_2)(x_{i3}^{(23)}-\mu_3)]$$

$$(5)$$

where

$$
\begin{aligned}
c_0 \;=\; & -\frac{3n_0}{2}ln(2\pi) - \frac{n_0}{2}ln(1 - 3\rho^2 + 2\rho^3) - n_{12}ln(2\pi) - \frac{n_{12}}{2}ln(1 - \rho^2) - n_{13}ln(2\pi) \\
& -\frac{n_{13}}{2}ln(1 - \rho^2) - n_{23}ln(2\pi) - \frac{n_{23}}{2}ln(1 - \rho^2) - \frac{n_1}{2}ln(2\pi) - \frac{n_2}{2}ln(2\pi) \\
& -\frac{n_3}{2}ln(2\pi)
\end{aligned}
$$

In order to find the Maximum Likelihood Estimates of $\boldsymbol{\mu} = (\mu_1, \mu_2, \mu_3)'$ and τ, the partial derivatives of lnL with respect to $\boldsymbol{\mu}$ and τ are taken, and are set to be zero:

$$
\frac{\partial lnL}{\partial \boldsymbol{\mu}} = \mathbf{0}
$$

$$
\frac{\partial lnL}{\partial \tau} = 0
$$

which are equivalent to the following expressions:

$$
\begin{aligned}
\frac{\partial lnL}{\partial \mu_1} \;=\; & -\frac{\rho + 1}{\tau(2\rho^2 - \rho - 1)}\sum_{i=1}^{n_0}(x_{i1} - \mu_1) + \frac{1}{\tau(1 - \rho^2)}[\sum_{i=1}^{n_{12}}(x_{i1}^{(12)} - \mu_1) \\
& +\sum_{i=1}^{n_{13}}(x_{i1}^{(13)} - \mu_1)] + \frac{1}{\tau}\sum_{i=1}^{n_1}(x_{i1}^{(1)} - \mu_1) + \frac{\rho}{\tau(2\rho^2 - \rho - 1)}[\sum_{i=1}^{n_0}(x_{i2} - \mu_2) \\
& +\sum_{i=1}^{n_0}(x_{i3} - \mu_3)] - \frac{\rho}{\tau(1 - \rho^2)}[\sum_{i=1}^{n_{12}}(x_{i2}^{(12)} - \mu_2) + \sum_{i=1}^{n_{13}}(x_{i3}^{(13)} - \mu_3)] = 0 \quad (6)
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial lnL}{\partial \mu_2} \;=\; & -\frac{\rho + 1}{\tau(2\rho^2 - \rho - 1)}\sum_{i=1}^{n_0}(x_{i2} - \mu_2) + \frac{1}{\tau(1 - \rho^2)}[\sum_{i=1}^{n_{12}}(x_{i2}^{(12)} - \mu_2) \\
& +\sum_{i=1}^{n_{23}}(x_{i2}^{(23)} - \mu_2)] + \frac{1}{\tau}\sum_{i=1}^{n_2}(x_{i2}^{(2)} - \mu_2) + \frac{\rho}{\tau(2\rho^2 - \rho - 1)}[\sum_{i=1}^{n_0}(x_{i1} - \mu_1) \\
& +\sum_{i=1}^{n_0}(x_{i3} - \mu_3)] - \frac{\rho}{\tau(1 - \rho^2)}[\sum_{i=1}^{n_{12}}(x_{i1}^{(12)} - \mu_1) + \sum_{i=1}^{n_{23}}(x_{i3}^{(23)} - \mu_3)] = 0 \quad (7)
\end{aligned}
$$

11

$$\frac{\partial lnL}{\partial \mu_3} = -\frac{\rho+1}{\tau(2\rho^2-\rho-1)}\sum_{i=1}^{n_0}(x_{i3}-\mu_3) + \frac{1}{\tau(1-\rho^2)}[\sum_{i=1}^{n_{13}}(x_{i3}^{(13)}-\mu_3)$$

$$+\sum_{i=1}^{n_{23}}(x_{i3}^{(23)}-\mu_3)] + \frac{1}{\tau}\sum_{i=1}^{n_3}(x_{i3}^{(3)}-\mu_3) + \frac{\rho}{\tau(2\rho^2-\rho-1)}[\sum_{i=1}^{n_0}(x_{i1}-\mu_1)$$

$$+\sum_{i=1}^{n_0}(x_{i2}-\mu_2)] - \frac{\rho}{\tau(1-\rho^2)}[\sum_{i=1}^{n_{13}}(x_{i1}^{(13)}-\mu_1) + \sum_{i=1}^{n_{23}}(x_{i2}^{(23)}-\mu_2)] = 0 \quad (8)$$

$$\frac{\partial lnL}{\partial \tau} = -\frac{1}{\tau}\left(\frac{3n_0}{2} + n_{12} + n_{13} + n_{23} + \frac{n_1}{2} + \frac{n_2}{2} + \frac{n_3}{2}\right)$$

$$-\frac{\rho+1}{2\tau^2(2\rho^2-\rho-1)}[\sum_{i=1}^{n_0}(x_{i1}-\mu_1)^2 + \sum_{i=1}^{n_0}(x_{i2}-\mu_2)^2 + \sum_{i=1}^{n_0}(x_{i3}-\mu_3)^2]$$

$$+\frac{\rho}{\tau^2(2\rho^2-\rho-1)}[\sum_{i=1}^{n_0}(x_{i1}-\mu_1)(x_{i2}-\mu_2) + \sum_{i=1}^{n_0}(x_{i1}-\mu_1)(x_{i3}-\mu_3)$$

$$+\sum_{i=1}^{n_0}(x_{i2}-\mu_2)(x_{i3}-\mu_3)] + \frac{1}{2\tau^2(1-\rho^2)}[\sum_{i=1}^{n_{12}}(x_{i1}^{(12)}-\mu_1)^2$$

$$-2\rho\sum_{i=1}^{n_{12}}(x_{i1}^{(12)}-\mu_1)(x_{i2}^{(12)}-\mu_2) + \sum_{i=1}^{n_{12}}(x_{i2}^{(12)}-\mu_2)^2 + \sum_{i=1}^{n_{13}}(x_{i1}^{(13)}-\mu_1)^2$$

$$-2\rho\sum_{i=1}^{n_{13}}(x_{i1}^{(13)}-\mu_1)(x_{i3}^{(13)}-\mu_3) + \sum_{i=1}^{n_{13}}(x_{i3}^{(13)}-\mu_3)^2 + \sum_{i=1}^{n_{23}}(x_{i2}^{(23)}-\mu_2)^2$$

$$-2\rho\sum_{i=1}^{n_{23}}(x_{i2}^{(23)}-\mu_2)(x_{i3}^{(23)}-\mu_3) + \sum_{i=1}^{n_{23}}(x_{i3}^{(23)}-\mu_3)^2]$$

$$+\frac{1}{2\tau^2}[\sum_{i=1}^{n_1}(x_{i1}^{(1)}-\mu_1)^2 + \sum_{i=1}^{n_2}(x_{i2}^{(2)}-\mu_2)^2 + \sum_{i=1}^{n_3}(x_{i3}^{(3)}-\mu_3)^2] = 0$$

$$(9)$$

The system of linear equations (6) (7) (8) can be rewritten as:

$$-(\rho+1)(1-\rho^2)\sum_{i=1}^{n_0}(x_{i1}-\mu_1) + (2\rho^2-\rho-1)[\sum_{i=1}^{n_{12}}(x_{i1}^{(12)}-\mu_1) + \sum_{i=1}^{n_{13}}(x_{i1}^{(13)}-\mu_1)]$$

$$+(2\rho^2-\rho-1)(1-\rho^2)\sum_{i=1}^{n_1}(x_{i1}^{(1)}-\mu_1) + \rho(1-\rho^2)[\sum_{i=1}^{n_0}(x_{i2}-\mu_2) + \sum_{i=1}^{n_0}(x_{i3}-\mu_3)]$$

$$-\rho(2\rho^2-\rho-1)[\sum_{i=1}^{n_{12}}(x_{i2}^{(12)}-\mu_2) + \sum_{i=1}^{n_{13}}(x_{i3}^{(13)}-\mu_3)] = 0 \quad (10)$$

12

$$-(\rho+1)(1-\rho^2)\sum_{i=1}^{n_0}(x_{i2}-\mu_2)+(2\rho^2-\rho-1)[\sum_{i=1}^{n_{12}}(x_{i2}^{(12)}-\mu_2)+\sum_{i=1}^{n_{23}}(x_{i2}^{(23)}-\mu_2)]$$

$$+(2\rho^2-\rho-1)(1-\rho^2)\sum_{i=1}^{n_2}(x_{i2}^{(2)}-\mu_2)+\rho(1-\rho^2)[\sum_{i=1}^{n_0}(x_{i1}-\mu_1)+\sum_{i=1}^{n_0}(x_{i3}-\mu_3)]$$

$$-\rho(2\rho^2-\rho-1)[\sum_{i=1}^{n_{12}}(x_{i1}^{(12)}-\mu_1)+\sum_{i=1}^{n_{23}}(x_{i3}^{(23)}-\mu_3)]=0 \tag{11}$$

$$-(\rho+1)(1-\rho^2)\sum_{i=1}^{n_0}(x_{i3}-\mu_3)+(2\rho^2-\rho-1)[\sum_{i=1}^{n_{13}}(x_{i3}^{(13)}-\mu_3)+\sum_{i=1}^{n_{23}}(x_{i3}^{(23)}-\mu_3)]$$

$$+(2\rho^2-\rho-1)(1-\rho^2)\sum_{i=1}^{n_3}(x_{i3}^{(3)}-\mu_3)+\rho(1-\rho^2)[\sum_{i=1}^{n_0}(x_{i1}-\mu_1)+\sum_{i=1}^{n_0}(x_{i2}-\mu_2)]$$

$$-\rho(2\rho^2-\rho-1)[\sum_{i=1}^{n_{13}}(x_{i1}^{(13)}-\mu_1)+\sum_{i=1}^{n_{23}}(x_{i2}^{(23)}-\mu_2)]=0 \tag{12}$$

They could be further expressed as follows:

$$[n_0(\rho+1)^2+(n_{12}+n_{13})(2\rho+1)+n_1(2\rho+1)(1+\rho)(1-\rho)]\mu_1$$

$$-\rho[n_{12}(2\rho+1)+n_0(1+\rho)]\mu_2-\rho[n_{13}(2\rho+1)+n_0(1+\rho)]\mu_3$$

$$-(\rho+1)^2\sum_{i=1}^{n_0}x_{i1}-(2\rho+1)(\sum_{i=1}^{n_{12}}x_{i1}^{(12)}+\sum_{i=1}^{n_{13}}x_{i1}^{(13)})$$

$$-(2\rho+1)(1+\rho)(1-\rho)\sum_{i=1}^{n_1}x_{i1}^{(1)}+\rho(1+\rho)\sum_{i=1}^{n_0}(x_{i2}+x_{i3})$$

$$+\rho(2\rho+1)(\sum_{i=1}^{n_{12}}x_{i2}^{(12)}+\sum_{i=1}^{n_{13}}x_{i3}^{(13)})=0 \tag{13}$$

$$-\rho[n_{12}(2\rho+1)+n_0(1+\rho)]\mu_1+[n_0(\rho+1)^2+(n_{12}+n_{23})(2\rho+1)$$

$$+n_2(2\rho+1)(1+\rho)(1-\rho)]\mu_2-\rho[n_{23}(2\rho+1)+n_0(1+\rho)]\mu_3$$

$$-(\rho+1)^2\sum_{i=1}^{n_0}x_{i2}-(2\rho+1)(\sum_{i=1}^{n_{12}}x_{i2}^{(12)}+\sum_{i=1}^{n_{23}}x_{i2}^{(23)})$$

$$- (2\rho + 1)(1 + \rho)(1 - \rho) \sum_{i=1}^{n_2} x_{i2}^{(2)} + \rho(1 + \rho) \sum_{i=1}^{n_0} (x_{i1} + x_{i3})$$

$$+ \rho(2\rho + 1)(\sum_{i=1}^{n_{12}} x_{i1}^{(12)} + \sum_{i=1}^{n_{23}} x_{i3}^{(23)}) = 0 \tag{14}$$

$$- \rho[n_{13}(2\rho + 1) + n_0(1 + \rho)]\mu_1 - \rho[n_{23}(2\rho + 1) + n_0(1 + \rho)]\mu_2$$

$$+ [n_0(\rho + 1)^2 + (n_{13} + n_{23})(2\rho + 1) + n_3(2\rho + 1)(1 + \rho)(1 - \rho)]\mu_3$$

$$- (\rho + 1)^2 \sum_{i=1}^{n_0} x_{i3} - (2\rho + 1)(\sum_{i=1}^{n_{13}} x_{i3}^{(13)} + \sum_{i=1}^{n_{23}} x_{i3}^{(23)})$$

$$- (2\rho + 1)(1 + \rho)(1 - \rho) \sum_{i=1}^{n_3} x_{i3}^{(3)} + \rho(1 + \rho) \sum_{i=1}^{n_0} (x_{i1} + x_{i2})$$

$$+ \rho(2\rho + 1)(\sum_{i=1}^{n_{13}} x_{i1}^{(13)} + \sum_{i=1}^{n_{23}} x_{i2}^{(23)}) = 0 \tag{15}$$

For the sake of convenience, the above system of equations (13) (14) (15) could be denoted by matrix and vectors:

$$\boldsymbol{A\mu} = \boldsymbol{b} \tag{16}$$

where Matrix $\boldsymbol{A}=$

$$\begin{pmatrix} a_{11} & -\rho(n_{12}\gamma_3 + n_0\gamma_1) & -\rho(n_{13}\gamma_3 + n_0\gamma_1) \\ -\rho(n_{12}\gamma_3 + n_0\gamma_1) & a_{22} & -\rho(n_{23}\gamma_3 + n_0\gamma_1) \\ -\rho(n_{13}\gamma_3 + n_0\gamma_1) & -\rho(n_{23}\gamma_3 + n_0\gamma_1) & a_{33} \end{pmatrix} \tag{17}$$

with

$$a_{11} = n_0\gamma_1^2 + (n_{12} + n_{13})\gamma_3 + n_1\gamma_3\gamma_2\gamma_1,$$

$$a_{22} = n_0\gamma_1^2 + (n_{12} + n_{23})\gamma_3 + n_2\gamma_3\gamma_2\gamma_1,$$

$$a_{33} = n_0\gamma_1^2 + (n_{13} + n_{23})\gamma_3 + n_3\gamma_3\gamma_2\gamma_1,$$

$$\gamma_1 = \rho + 1, \gamma_2 = 1 - \rho, \gamma_3 = 2\rho + 1$$

$$\boldsymbol{\mu} = (\mu_1, \mu_2, \mu_3)', \boldsymbol{b} = (b_1, b_2, b_3)' \tag{18}$$

$$
\begin{aligned}
b_1 &= \gamma_1^2 n_0 \bar{x}_{.1} + \gamma_3 (n_{12} \bar{x}_{.1}^{(12)} + n_{13} \bar{x}_{.1}^{(13)}) + \gamma_3 \gamma_2 \gamma_1 n_1 \bar{x}_{.1}^{(1)} - \rho \gamma_1 n_0 (\bar{x}_{.2} + \bar{x}_{.3}) \\
&\quad - \rho \gamma_3 (n_{12} \bar{x}_{.2}^{(12)} + n_{13} \bar{x}_{.3}^{(13)})
\end{aligned}
$$

$$
\begin{aligned}
b_2 &= \gamma_1^2 n_0 \bar{x}_{.2} + \gamma_3 (n_{12} \bar{x}_{.2}^{(12)} + n_{23} \bar{x}_{.2}^{(23)}) + \gamma_3 \gamma_2 \gamma_1 n_2 \bar{x}_{.2}^{(2)} - \rho \gamma_1 n_0 (\bar{x}_{.1} + \bar{x}_{.3}) \\
&\quad - \rho \gamma_3 (n_{12} \bar{x}_{.1}^{(12)} + n_{23} \bar{x}_{.3}^{(23)})
\end{aligned}
$$

$$
\begin{aligned}
b_3 &= \gamma_1^2 n_0 \bar{x}_{.3} + \gamma_3 (n_{13} \bar{x}_{.3}^{(13)} + n_{23} \bar{x}_{.3}^{(23)}) + \gamma_3 \gamma_2 \gamma_1 n_3 \bar{x}_{.3}^{(3)} - \rho \gamma_1 n_0 (\bar{x}_{.1} + \bar{x}_{.2}) \\
&\quad - \rho \gamma_3 (n_{13} \bar{x}_{.1}^{(13)} + n_{23} \bar{x}_{.2}^{(23)})
\end{aligned}
$$

where $\bar{x}_{.j} = \frac{\sum_{i=1}^{n_0} x_{ij}}{n_0}$, $\bar{x}_{.j}^{(12)} = \frac{\sum_{i=1}^{n_{12}} x_{ij}^{(12)}}{n_{12}}$, $\bar{x}_{.j}^{(13)} = \frac{\sum_{i=1}^{n_{13}} x_{ij}^{(13)}}{n_{13}}$, $\bar{x}_{.j}^{(23)} = \frac{\sum_{i=1}^{n_{23}} x_{ij}^{(23)}}{n_{23}}$,

$\bar{x}_{.j}^{(1)} = \frac{\sum_{i=1}^{n_1} x_{ij}^{(1)}}{n_1}$, $\bar{x}_{.j}^{(2)} = \frac{\sum_{i=1}^{n_2} x_{ij}^{(2)}}{n_2}$, $\bar{x}_{.j}^{(3)} = \frac{\sum_{i=1}^{n_3} x_{ij}^{(3)}}{n_3}$, $j = 1, 2, 3$.

If $\boldsymbol{A}$ is a positive definite matrix, then the equation (16) can be solved and thus allow us to obtain the matrix expression of Maximum Likelihood Estimates of $\boldsymbol{\mu}$. To prove the positive definiteness of $\boldsymbol{A}$, matrix $\boldsymbol{A}$ is firstly denoted as the sum of two matrices: $\boldsymbol{B_1}$ and $\boldsymbol{B_2}$. The next step is to show that $\boldsymbol{B_1}$ is positive semi-definite and $\boldsymbol{B_2}$ is positive definite, where

$$\boldsymbol{B_1} = \gamma_3 \begin{pmatrix} n_{12} + n_{13} & -\rho n_{12} & -\rho n_{13} \\ -\rho n_{12} & n_{12} + n_{23} & -\rho n_{23} \\ -\rho n_{13} & -\rho n_{23} & n_{13} + n_{23} \end{pmatrix} \tag{19}$$

$$\boldsymbol{B_2} = \gamma_1 \begin{pmatrix} n_0 \gamma_1 + n_1 \gamma_2 \gamma_3 & -\rho n_0 & -\rho n_0 \\ -\rho n_0 & n_0 \gamma_1 + n_2 \gamma_2 \gamma_3 & -\rho n_0 \\ -\rho n_0 & -\rho n_0 & n_0 \gamma_1 + n_3 \gamma_2 \gamma_3 \end{pmatrix}. \tag{20}$$

It can be proved that matrix $\boldsymbol{B_1}$ is positive semi-definite by rewriting $\boldsymbol{B_1}$ as the sum of $\boldsymbol{C_1}, \boldsymbol{C_2}$ and $\boldsymbol{C_3}$, where

$$
\boldsymbol{C_1} = \gamma_3 \begin{pmatrix} n_{12} & -\rho n_{12} & 0 \\ -\rho n_{12} & n_{12} & 0 \\ 0 & 0 & 0 \end{pmatrix}
\tag{21}
$$

$$
\boldsymbol{C_2} = \gamma_3 \begin{pmatrix} n_{13} & 0 & -\rho n_{13} \\ 0 & 0 & 0 \\ -\rho n_{13} & 0 & n_{13} \end{pmatrix}
\tag{22}
$$

$$
\boldsymbol{C_3} = \gamma_3 \begin{pmatrix} 0 & 0 & 0 \\ 0 & n_{23} & -\rho n_{23} \\ 0 & -\rho n_{23} & n_{23} \end{pmatrix}.
\tag{23}
$$

Proof. A matrix is positive semi-definite if and only if every principal minor of the matrix is nonnegative. Since $n_{12} \geq 0, \rho \in (-\frac{1}{2}, 1), \gamma_3 = 2\rho + 1 > 0$, the first-order principal minors of matrix $\boldsymbol{C_1}$ are nonnegative; the second-order principal minors of matrix $\boldsymbol{C_1}$ are $\gamma_3 n_{12}^2 (1 - \rho^2) \geq 0$ and 0; the third-order principal minor of matrix $\boldsymbol{C_1}$ is 0. Therefore, $\boldsymbol{C_1}$ is a positive semi-definite matrix. Similarly, since $n_{13} \geq 0$ and $n_{23} \geq 0$, all the principal minors of $\boldsymbol{C_2}$ and $\boldsymbol{C_3}$ are nonnegative, so that $\boldsymbol{C_2}$ and $\boldsymbol{C_3}$ are also positive semi-definite matrices.

Based on the results above, matrix $\boldsymbol{B_1} = \boldsymbol{C_1} + \boldsymbol{C_2} + \boldsymbol{C_3}$ is a positive semi-definite matrix. ■

Applying the similar argument to matrix $\boldsymbol{B_2}$, we could write it as the sum of $\boldsymbol{C_4}$ and $\boldsymbol{C_5}$ given below. By showing that $\boldsymbol{C_4}$ is positive definite and $\boldsymbol{C_5}$ is positive semi-definite, $\boldsymbol{B_2}$ can be proved as a positive definite matrix.

$$C_4 = \gamma_1 \begin{pmatrix} n_0\gamma_1 & -n_0\rho & -n_0\rho \\ -n_0\rho & n_0\gamma_1 & -n_0\rho \\ -n_0\rho & -n_0\rho & n_0\gamma_1 \end{pmatrix} \tag{24}$$

$$C_5 = \gamma_1 \begin{pmatrix} n_1\gamma_2\gamma_3 & 0 & 0 \\ 0 & n_2\gamma_2\gamma_3 & 0 \\ 0 & 0 & n_3\gamma_2\gamma_3 \end{pmatrix} \tag{25}$$

Proof. Since $n_0 > 0, \rho \in (-\frac{1}{2}, 1), \gamma_1 = \rho + 1 > 0, \gamma_2 = 1 - \rho > 0, \gamma_3 = 2\rho + 1 > 0$, the first-order principal minor of matrix C_4 is $n_0\gamma_1^2 > 0$; the second-order principal minor is $n_0^2\gamma_1(\gamma_1^2 - \rho^2) > 0$; the third-order principal minor is $n_0^3(\rho + 1)^3(1 - \rho)(2\rho + 1)^2 > 0$. Therefore, C_4 is a positive definite matrix.

Since $n_1 \geq 0, n_2 \geq 0, n_3 \geq 0$, the first-order principal minors of matrix C_5 are $n_1\gamma_1\gamma_2\gamma_3 \geq 0$, $n_2\gamma_1\gamma_2\gamma_3 \geq 0$ and $n_3\gamma_1\gamma_2\gamma_3 \geq 0$. Similarly, it can be checked that all the second and third-order principal minors are nonnegative. Therefore C_5 is a positive semi-definite matrix. As a result, matrix $B_2 = C_4 + C_5$ is a positive definite matrix.

B_1 is shown to be positive semi-definite and B_2 is positive definite, therefore $A = B_1 + B_2$ is a positive definite matrix. ∎

Note that A is positive definite, thus it is a nonsingular matrix and its inverse matrix exists. It could be derived from expression (16) that

$$\hat{\mu} = A^{-1}b$$

where

$$\hat{\mu} = (\hat{\mu}_1, \hat{\mu}_2, \hat{\mu}_3)'. \tag{26}$$

Moreover, the Maximum Likelihood Estimator of τ could be obtained by plugging $\hat{\mu}$ into equation (9):

$$\hat{\tau} = \frac{1}{\frac{3n_0}{2} + n_{12} + n_{13} + n_{23} + \frac{n_1}{2} + \frac{n_2}{2} + \frac{n_3}{2}} \cdot \left\{ \frac{\rho+1}{2(2\rho+1)(1-\rho)} \left[\sum_{i=1}^{n_0}(x_{i1} - \hat{\mu}_1)^2 \right.\right.$$

$$+ \sum_{i=1}^{n_0}(x_{i2} - \hat{\mu}_2)^2 + \sum_{i=1}^{n_0}(x_{i3} - \hat{\mu}_3)^2] - \frac{\rho}{(2\rho+1)(1-\rho)}\left[\sum_{i=1}^{n_0}(x_{i1} - \hat{\mu}_1)(x_{i2} - \hat{\mu}_2)\right.$$

$$+ \sum_{i=1}^{n_0}(x_{i1} - \hat{\mu}_1)(x_{i3} - \hat{\mu}_3) + \sum_{i=1}^{n_0}(x_{i2} - \hat{\mu}_2)(x_{i3} - \hat{\mu}_3)]$$

$$+ \frac{1}{2(1+\rho)(1-\rho)}\left[\sum_{i=1}^{n_{12}}(x_{i1}^{(12)} - \hat{\mu}_1)^2 - 2\rho\sum_{i=1}^{n_{12}}(x_{i1}^{(12)} - \hat{\mu}_1)(x_{i2}^{(12)} - \hat{\mu}_2)\right.$$

$$+ \sum_{i=1}^{n_{12}}(x_{i2}^{(12)} - \hat{\mu}_2)^2 + \sum_{i=1}^{n_{13}}(x_{i1}^{(13)} - \hat{\mu}_1)^2 - 2\rho\sum_{i=1}^{n_{13}}(x_{i1}^{(13)} - \hat{\mu}_1)(x_{i3}^{(13)} - \hat{\mu}_3)$$

$$+ \sum_{i=1}^{n_{13}}(x_{i3}^{(13)} - \hat{\mu}_3)^2 + \sum_{i=1}^{n_{23}}(x_{i2}^{(23)} - \hat{\mu}_2)^2 - 2\rho\sum_{i=1}^{n_{23}}(x_{i2}^{(23)} - \hat{\mu}_2)(x_{i3}^{(23)} - \hat{\mu}_3)$$

$$\left.+ \sum_{i=1}^{n_{23}}(x_{i3}^{(23)} - \hat{\mu}_3)^2] + \frac{1}{2}\left[\sum_{i=1}^{n_1}(x_{i1}^{(1)} - \hat{\mu}_1)^2 + \sum_{i=1}^{n_2}(x_{i2}^{(2)} - \hat{\mu}_2)^2 + \sum_{i=1}^{n_3}(x_{i3}^{(3)} - \hat{\mu}_3)^2\right]\right\}$$

$$(27)$$

By denoting $c_1 = \frac{3n_0}{2} + n_{12} + n_{13} + n_{23} + \frac{n_1}{2} + \frac{n_2}{2} + \frac{n_3}{2}$, $\gamma_1 = \rho + 1$, $\gamma_2 = 1 - \rho$, $\gamma_3 = 2\rho + 1$, the equation above can be further simplified as:

$$\hat{\tau} = \frac{1}{c_1} \cdot \left\{ \frac{\gamma_1}{2\gamma_3\gamma_2}\left[\sum_{i=1}^{n_0}(x_{i1} - \hat{\mu}_1)^2 + \sum_{i=1}^{n_0}(x_{i2} - \hat{\mu}_2)^2 + \sum_{i=1}^{n_0}(x_{i3} - \hat{\mu}_3)^2\right]\right.$$

$$- \frac{\rho}{\gamma_3\gamma_2}\left[\sum_{i=1}^{n_0}(x_{i1} - \hat{\mu}_1)(x_{i2} - \hat{\mu}_2) + \sum_{i=1}^{n_0}(x_{i1} - \hat{\mu}_1)(x_{i3} - \hat{\mu}_3)\right.$$

$$\left.+ \sum_{i=1}^{n_0}(x_{i2} - \hat{\mu}_2)(x_{i3} - \hat{\mu}_3)\right] + \frac{1}{2\gamma_1\gamma_2}\left[\sum_{i=1}^{n_{12}}(x_{i1}^{(12)} - \hat{\mu}_1)^2\right.$$

$$- 2\rho\sum_{i=1}^{n_{12}}(x_{i1}^{(12)} - \hat{\mu}_1)(x_{i2}^{(12)} - \hat{\mu}_2) + \sum_{i=1}^{n_{12}}(x_{i2}^{(12)} - \hat{\mu}_2)^2 + \sum_{i=1}^{n_{13}}(x_{i1}^{(13)} - \hat{\mu}_1)^2$$

$$- 2\rho\sum_{i=1}^{n_{13}}(x_{i1}^{(13)} - \hat{\mu}_1)(x_{i3}^{(13)} - \hat{\mu}_3) + \sum_{i=1}^{n_{13}}(x_{i3}^{(13)} - \hat{\mu}_3)^2 + \sum_{i=1}^{n_{23}}(x_{i2}^{(23)} - \hat{\mu}_2)^2$$

$$- 2\rho\sum_{i=1}^{n_{23}}(x_{i2}^{(23)} - \hat{\mu}_2)(x_{i3}^{(23)} - \hat{\mu}_3) + \sum_{i=1}^{n_{23}}(x_{i3}^{(23)} - \hat{\mu}_3)^2]$$

$$\left.+ \frac{1}{2}\left[\sum_{i=1}^{n_1}(x_{i1}^{(1)} - \hat{\mu}_1)^2 + \sum_{i=1}^{n_2}(x_{i2}^{(2)} - \hat{\mu}_2)^2 + \sum_{i=1}^{n_3}(x_{i3}^{(3)} - \hat{\mu}_3)^2\right]\right\}$$

IV. Properties of MLEs

1. Mean of $\hat{\boldsymbol{\mu}}$

The estimate of $\boldsymbol{\mu}$ in our model with missing data is an unbiased estimator. This property could be easily proved as follows.

Proof. Since $E(\bar{x}_{\cdot j}) = E(\bar{x}_{\cdot j}^{(12)}) = E(\bar{x}_{\cdot j}^{(13)}) = E(\bar{x}_{\cdot j}^{(23)}) = E(\bar{x}_{\cdot j}^{(1)}) = E(\bar{x}_{\cdot j}^{(2)}) = E(\bar{x}_{\cdot j}^{(3)}) = \mu_j, \ \ j = 1, 2, 3.$

$$E(\boldsymbol{b}) = \begin{pmatrix} E(b_1) \\ E(b_2) \\ E(b_3) \end{pmatrix}$$

$$=$$

$$\begin{pmatrix} \gamma_1^2 n_0 \mu_1 + \gamma_3(n_{12}\mu_1 + n_{13}\mu_1) + \gamma_3\gamma_2\gamma_1 n_1 \mu_1 - \rho\gamma_1 n_0(\mu_2 + \mu_3) - \rho\gamma_3(n_{12}\mu_2 + n_{13}\mu_3) \\ \gamma_1^2 n_0 \mu_2 + \gamma_3(n_{12}\mu_2 + n_{23}\mu_2) + \gamma_3\gamma_2\gamma_1 n_2 \mu_2 - \rho\gamma_1 n_0(\mu_1 + \mu_3) - \rho\gamma_3(n_{12}\mu_1 + n_{23}\mu_3) \\ \gamma_1^2 n_0 \mu_3 + \gamma_3(n_{13}\mu_3 + n_{23}\mu_3) + \gamma_3\gamma_2\gamma_1 n_3 \mu_3 - \rho\gamma_1 n_0(\mu_1 + \mu_2) - \rho\gamma_3(n_{13}\mu_1 + n_{23}\mu_2) \end{pmatrix}$$

$$=$$

$$\begin{pmatrix} [n_0\gamma_1^2 + (n_{12} + n_{13})\gamma_3 + n_1\gamma_3\gamma_2\gamma_1]\mu_1 - \rho(n_{12}\gamma_3 + n_0\gamma_1)\mu_2 - \rho(n_{13}\gamma_3 + n_0\gamma_1)\mu_3 \\ -\rho(n_{12}\gamma_3 + n_0\gamma_1)\mu_1 + [n_0\gamma_1^2 + (n_{12} + n_{23})\gamma_3 + n_2\gamma_3\gamma_2\gamma_1]\mu_2 - \rho(n_{23}\gamma_3 + n_0\gamma_1)\mu_3 \\ -\rho(n_{13}\gamma_3 + n_0\gamma_1)\mu_1 - \rho(n_{23}\gamma_3 + n_0\gamma_1)\mu_2 + [n_0\gamma_1^2 + (n_{13} + n_{23})\gamma_3 + n_3\gamma_3\gamma_2\gamma_1]\mu_3 \end{pmatrix}$$

$$= \boldsymbol{A}\boldsymbol{\mu}$$

Consequently

$$E(\hat{\boldsymbol{\mu}}) = E(\boldsymbol{A}^{-1}\boldsymbol{b}) = \boldsymbol{A}^{-1}E(\boldsymbol{b}) = \boldsymbol{A}^{-1}\boldsymbol{A}\boldsymbol{\mu} = \boldsymbol{\mu}$$

∎

2. Variance of $\hat{\boldsymbol{\mu}}$

With matrix notation, the variance of $\hat{\boldsymbol{\mu}}$ can be easily denoted as

$$Var(\hat{\boldsymbol{\mu}}) = Var(\boldsymbol{A}^{-1}\boldsymbol{b}) = \boldsymbol{A}^{-1}Var(\boldsymbol{b})(\boldsymbol{A}^{-1})' \tag{29}$$

where $\boldsymbol{b}$ is the same vector as in (18). For the sake of convenience, $\boldsymbol{b}$ could be denoted as

$$\boldsymbol{b} = \boldsymbol{DX}^*$$

where

$$\boldsymbol{D}' = \begin{pmatrix} n_0\gamma_1^2 & -\rho n_0\gamma_1 & -\rho n_0\gamma_1 \\ -\rho n_0\gamma_1 & n_0\gamma_1^2 & -\rho n_0\gamma_1 \\ -\rho n_0\gamma_1 & -\rho n_0\gamma_1 & n_0\gamma_1^2 \\ n_{12}\gamma_3 & -\rho n_{12}\gamma_3 & 0 \\ -\rho n_{12}\gamma_3 & n_{12}\gamma_3 & 0 \\ n_{13}\gamma_3 & 0 & -\rho n_{13}\gamma_3 \\ -\rho n_{13}\gamma_3 & 0 & n_{13}\gamma_3 \\ 0 & n_{23}\gamma_3 & -\rho n_{23}\gamma_3 \\ 0 & -\rho n_{23}\gamma_3 & n_{23}\gamma_3 \\ n_1\gamma_3\gamma_2\gamma_1 & 0 & 0 \\ 0 & n_2\gamma_3\gamma_2\gamma_1 & 0 \\ 0 & 0 & n_3\gamma_3\gamma_2\gamma_1 \end{pmatrix}$$

and

$$\boldsymbol{X}^* = (\bar{x}_{.1}, \bar{x}_{.2}, \bar{x}_{.3}, \bar{x}_{.1}^{(12)}, \bar{x}_{.2}^{(12)}, \bar{x}_{.1}^{(13)}, \bar{x}_{.3}^{(13)}, \bar{x}_{.2}^{(23)}, \bar{x}_{.3}^{(23)}, \bar{x}_{.1}^{(1)}, \bar{x}_{.2}^{(2)}, \bar{x}_{.3}^{(3)})'.$$

Since the $Var(\boldsymbol{b})$ could be further expressed as:

$$Var(\boldsymbol{b}) = Var(\boldsymbol{DX}^*) = \boldsymbol{D}Var(\boldsymbol{X}^*)\boldsymbol{D}', \tag{30}$$

combining (29) and (30), we get:

$$Var(\hat{\boldsymbol{\mu}}) = (\boldsymbol{A}^{-1}\boldsymbol{D})Var(\boldsymbol{X}^*)(\boldsymbol{A}^{-1}\boldsymbol{D})' \tag{31}$$

20

where the variance of $\boldsymbol{X}^*$ can be obtained by

$$Var(\boldsymbol{X}^*) = \begin{pmatrix}
\frac{\sigma^2}{n_0} & \frac{\rho\sigma^2}{n_0} & \frac{\rho\sigma^2}{n_0} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\frac{\rho\sigma^2}{n_0} & \frac{\sigma^2}{n_0} & \frac{\rho\sigma^2}{n_0} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\frac{\rho\sigma^2}{n_0} & \frac{\rho\sigma^2}{n_0} & \frac{\sigma^2}{n_0} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & \frac{\sigma^2}{n_{12}} & \frac{\rho\sigma^2}{n_{12}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & \frac{\rho\sigma^2}{n_{12}} & \frac{\sigma^2}{n_{12}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & \frac{\sigma^2}{n_{13}} & \frac{\rho\sigma^2}{n_{13}} & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & \frac{\rho\sigma^2}{n_{13}} & \frac{\sigma^2}{n_{13}} & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{\sigma^2}{n_{23}} & \frac{\rho\sigma^2}{n_{23}} & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{\rho\sigma^2}{n_{23}} & \frac{\sigma^2}{n_{23}} & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{\sigma^2}{n_1} & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{\sigma^2}{n_2} & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{\sigma^2}{n_3}
\end{pmatrix}.$$

3. Variance of $\hat{\mu}_1 - \hat{\mu}_2$

The variance of $\hat{\boldsymbol{\mu}}$ is a 3×3 matrix which has the format as follows:

$$Var(\hat{\boldsymbol{\mu}}) = \begin{pmatrix}
Var(\hat{\mu}_1) & Cov(\hat{\mu}_1, \hat{\mu}_2) & Cov(\hat{\mu}_1, \hat{\mu}_3) \\
Cov(\hat{\mu}_1, \hat{\mu}_2) & Var(\hat{\mu}_2) & Cov(\hat{\mu}_2, \hat{\mu}_3) \\
Cov(\hat{\mu}_1, \hat{\mu}_3) & Cov(\hat{\mu}_2, \hat{\mu}_3) & Var(\hat{\mu}_3)
\end{pmatrix}. \tag{32}$$

Consequently, the variance of $\hat{\mu}_1 - \hat{\mu}_2$ can be obtained as:

$$Var(\hat{\mu}_1 - \hat{\mu}_2) = Var(\hat{\mu}_1) + Var(\hat{\mu}_2) - 2Cov(\hat{\mu}_1, \hat{\mu}_2)$$

V. Sampling distributions

1. The distribution of $\hat{\boldsymbol{\mu}} - \boldsymbol{\mu}$

Since vector $\boldsymbol{X} = (\boldsymbol{X_1}, \boldsymbol{X_2}, \boldsymbol{X_3})$ has a joint normal distribution, $\hat{\boldsymbol{\mu}}$ is a linear combination of $\boldsymbol{b}$, which is also the linear combination of the components of $\boldsymbol{X}$. As a

21

result, $\hat{\boldsymbol{\mu}}$ follows the normal distribution: $N_3(E(\hat{\boldsymbol{\mu}} - \boldsymbol{\mu}), Var(\hat{\boldsymbol{\mu}} - \boldsymbol{\mu}))$, where $E(\hat{\boldsymbol{\mu}} - \boldsymbol{\mu})$

and $Var(\hat{\boldsymbol{\mu}} - \boldsymbol{\mu})$ could be simply calculated as:

$$E(\hat{\boldsymbol{\mu}} - \boldsymbol{\mu}) = E(\hat{\boldsymbol{\mu}}) - E(\boldsymbol{\mu}) = \boldsymbol{\mu} - \boldsymbol{\mu} = \mathbf{0}$$

$$Var(\hat{\boldsymbol{\mu}} - \boldsymbol{\mu}) = Var(\hat{\boldsymbol{\mu}}) = (\boldsymbol{A}^{-1}\boldsymbol{D})Var(\boldsymbol{X}^*)(\boldsymbol{A}^{-1}\boldsymbol{D})'.$$

2. The distribution of $\hat{\mu}_1 - \hat{\mu}_2$

Since $\hat{\boldsymbol{\mu}} = (\hat{\mu}_1, \hat{\mu}_2, \hat{\mu}_3)'$ has a joint normal distribution, $\hat{\mu}_1 - \hat{\mu}_2$ follows normal

distribution with mean

$$E(\hat{\mu}_1 - \hat{\mu}_2) = E(\hat{\mu}_1) - E(\hat{\mu}_2) = \mu_1 - \mu_2$$

and variance

$$Var(\hat{\mu}_1 - \hat{\mu}_2) = Var(\hat{\mu}_1) + Var(\hat{\mu}_2) - 2Cov(\hat{\mu}_1, \hat{\mu}_2).$$

VI. Simulation Study with Known ρ

In this section, a simulation study is conducted to compare the performance of

estimates in case (a) only three-dimensional observations and (b) both three-dimensional

observations and all possible incomplete observations. The efficiency of this estimator is

evaluated by means of coverage probabilities and average width of the confidence intervals,

type I error and power of testing $H_0 : \boldsymbol{\mu} = \mathbf{0} \ vs. H_a : \boldsymbol{\mu} \neq \mathbf{0}$. The sample size is chosen as

$n_0 = 10, n_{12} = n_{13} = n_{23} = 15, n_1 = n_2 = n_3 = 20$, and the following five levels of stan-

dard deviations are considered: $\sigma = 0.25, 0.5, 1, 2, 4$. It is assumed that ρ is known in this

section, so the real value of ρ is used, which ranges from 0.1 to 0.9. The population mean

$\boldsymbol{\mu}$ is chosen to be $\boldsymbol{\mu} = (1, 2, 3)'$ when calculating the power of testing, coverage probabili-

ties and average width of the confidence intervals. The number of replication is $r = 10000$.

1. Compare the confidence regions of $\boldsymbol{\mu}$

Table 1: Coverage Probabilities of 95% confidence regions of $\boldsymbol{\mu} = (1, 2, 3)'$

	ρ	σ				
		0.25	0.5	1	2	4
(a)	0.1	0.9518	0.9511	0.9447	0.9520	0.9492
	0.2	0.9517	0.9492	0.9488	0.9495	0.9497
	0.3	0.9503	0.9506	0.9527	0.9525	0.9486
	0.4	0.9518	0.9488	0.9495	0.9507	0.9463
	0.5	0.9520	0.9492	0.9498	0.9516	0.9469
	0.6	0.9495	0.9497	0.9509	0.9511	0.9447
	0.7	0.9525	0.9486	0.9517	0.9492	0.9488
	0.8	0.9507	0.9463	0.9503	0.9506	0.9527
	0.9	0.9516	0.9469	0.9518	0.9488	0.9495
	ρ	σ				
		0.25	0.5	1	2	4
(b)	0.1	0.9501	0.9511	0.9496	0.9503	0.9452
	0.2	0.9468	0.9503	0.9520	0.9499	0.9469
	0.3	0.9525	0.9525	0.9501	0.9512	0.9478
	0.4	0.9510	0.9505	0.9539	0.9495	0.9495
	0.5	0.9523	0.9439	0.9496	0.9507	0.9516
	0.6	0.9503	0.9462	0.9525	0.9520	0.9519
	0.7	0.9522	0.9489	0.9503	0.9494	0.9526
	0.8	0.9516	0.9495	0.9538	0.9535	0.9507
	0.9	0.9514	0.9513	0.9542	0.9514	0.9521
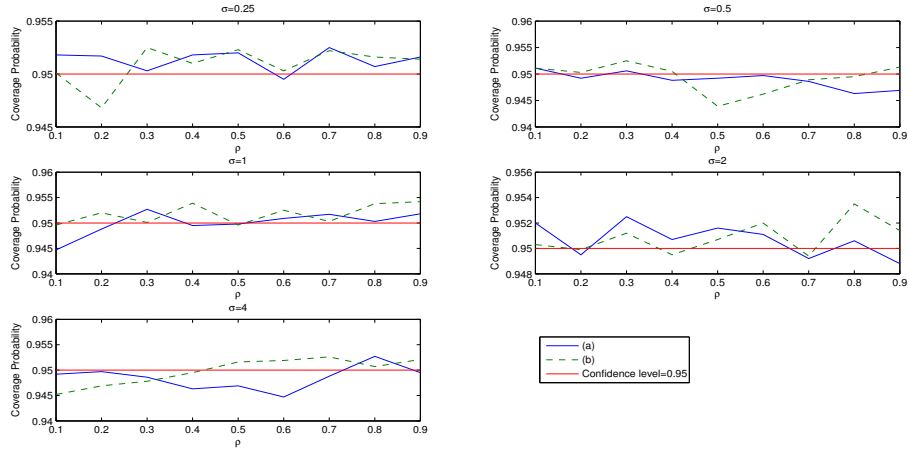
Figure 1: Comparison of Coverage Probabilities of Confidence Region of $\boldsymbol{\mu}$
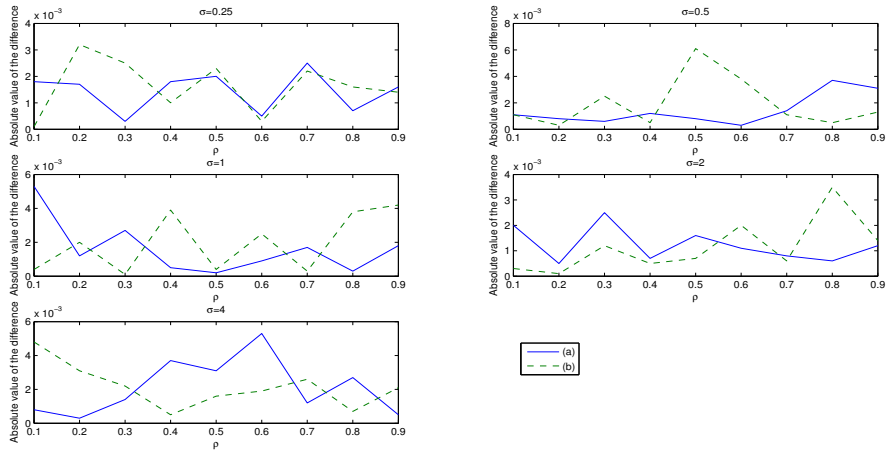
Figure 2: Absolute Value of the Difference Between the Coverage Probabilities and 0.95

It is evident from Figure 1 that, in both case (a) and (b), the coverage probabilities of the 95% confidence regions are close to 0.95. The absolute values of difference between the coverage probabilities and confidence level 0.95, as shown in Figure 2, are also small in both cases.

24

2. Compare the probability of type I error and power of testing

$H_0 : \boldsymbol{\mu} = \mathbf{0} \ vs. H_a : \boldsymbol{\mu} \neq \mathbf{0}$

2.1 Compare the probability of type I error

Table 2: The Probability of Type I error

	ρ	σ				
		0.25	0.5	1	2	4
(a)	0.1	0.0504	0.0449	0.0484	0.0490	0.0461
	0.2	0.0512	0.0508	0.0524	0.0514	0.0479
	0.3	0.0494	0.0497	0.0505	0.0514	0.0458
	0.4	0.0470	0.0511	0.0510	0.0511	0.0534
	0.5	0.0472	0.0484	0.0497	0.0515	0.0461
	0.6	0.0509	0.0529	0.0505	0.0487	0.0484
	0.7	0.0470	0.0464	0.0487	0.0481	0.0470
	0.8	0.0504	0.0487	0.0488	0.0509	0.0497
	0.9	0.0515	0.0471	0.0489	0.0513	0.0533
	ρ	σ				
		0.25	0.5	1	2	4
(b)	0.1	0.0510	0.0530	0.0523	0.0551	0.0513
	0.2	0.0489	0.0514	0.0537	0.0514	0.0502
	0.3	0.0513	0.0548	0.0501	0.0477	0.0482
	0.4	0.0513	0.0523	0.0479	0.0516	0.0540
	0.5	0.0485	0.0512	0.0524	0.0484	0.0472
	0.6	0.0483	0.0497	0.0463	0.0508	0.0449
	0.7	0.0439	0.0525	0.0515	0.0491	0.0481
	0.8	0.0490	0.0524	0.0498	0.0471	0.0527
	0.9	0.0501	0.0519	0.0478	0.0479	0.0525
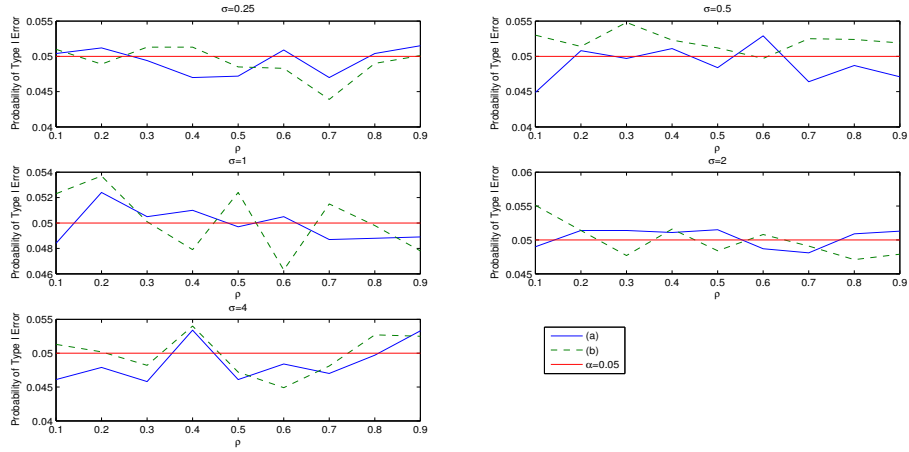
Figure 3: Comparison of Type I Error
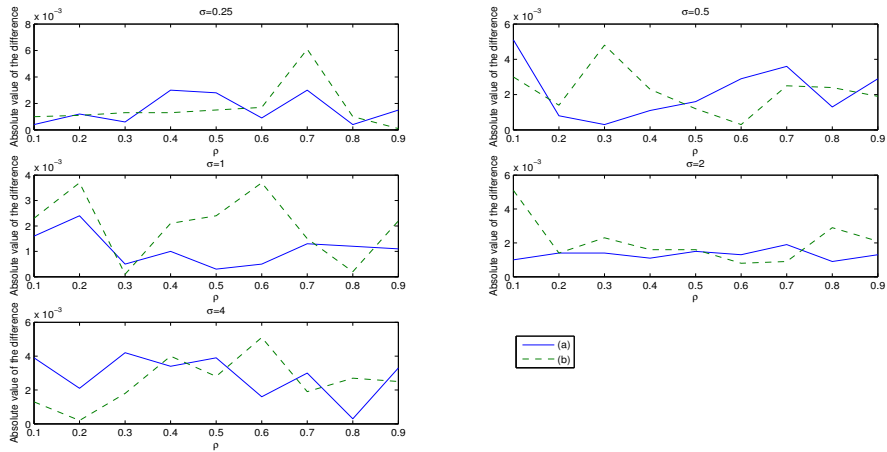
Figure 4: Absolute Value of the Difference Between the Type I Error and $\alpha = 0.05$

In the case of $\alpha = 0.05$, type I errors in both (a) and (b) are close to 0.05 at each level of σ. It can be noted more clearly from Figure 4 that the absolute values of difference between Type I Errors and $\alpha = 0.05$ are small in both cases.

2.2 Compare the power of testing $H_0 : \boldsymbol{\mu} = \mathbf{0} \ vs. H_a : \boldsymbol{\mu} \neq \mathbf{0}$

Table 3: The Power of testing $H_0 : \boldsymbol{\mu} = (0,0,0) \ vs. H_a : \boldsymbol{\mu} = (1,2,3)'$

	ρ	σ				
		6	8	10	12	14
(a)	0.1	0.3614	0.2168	0.1473	0.1094	0.1009
	0.2	0.3334	0.2003	0.1404	0.1065	0.0901
	0.3	0.3034	0.1837	0.1249	0.1076	0.0905
	0.4	0.2992	0.1826	0.1302	0.1048	0.0891
	0.5	0.3108	0.1825	0.1289	0.1029	0.0874
	0.6	0.3120	0.1863	0.1363	0.1039	0.0936
	0.7	0.3578	0.2096	0.1482	0.1142	0.0930
	0.8	0.4394	0.2562	0.1714	0.1335	0.1045
	0.9	0.6456	0.3976	0.2608	0.1958	0.1577
	ρ	σ				
		6	8	10	12	14
(b)	0.1	0.9888	0.8653	0.6705	0.4926	0.3724
	0.2	0.9835	0.8483	0.6398	0.4685	0.3606
	0.3	0.9818	0.8377	0.6286	0.4641	0.3523
	0.4	0.9791	0.8318	0.6152	0.4492	0.3434
	0.5	0.9758	0.8265	0.6151	0.4524	0.3375
	0.6	0.9791	0.8317	0.6261	0.4519	0.3521
	0.7	0.9860	0.8510	0.6535	0.4712	0.3559
	0.8	0.9924	0.8982	0.6977	0.5347	0.4108
	0.9	0.9994	0.9696	0.8453	0.6901	0.5399
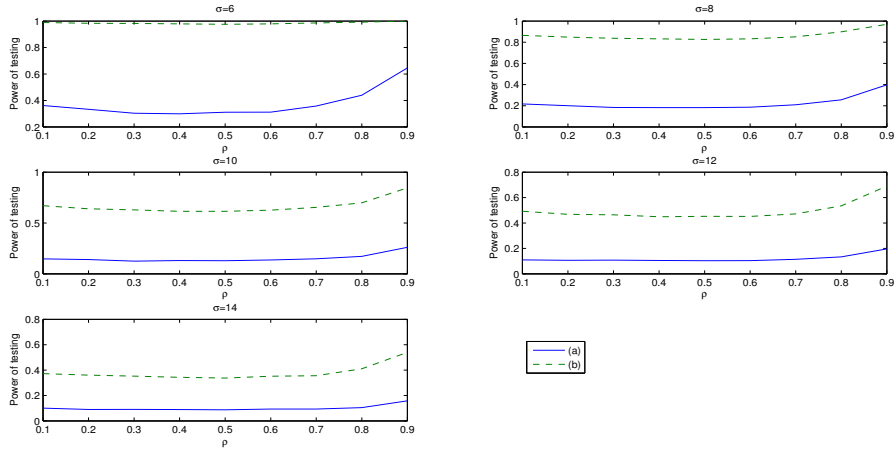
Figure 5: Compare the Power of Testing

In order to compare the power of testing, our levels of standard deviation have been chosen as $\sigma = 6, 8, 10, 12, 14$ to make the results more observable. It can be observed from Figure 5 that, at each level of σ, the power of testing is much higher in case (b) than that in (a).

3. Compare the confidence intervals of μ_1

3.1 Compare the coverage probabilities of confidence intervals of μ_1

It can be noted from Figure 6 and Figure 7 that, in both case (a) and (b), the coverage probabilities of the 95% confidence intervals of μ_1 are close to 0.95. The absolute values of differences between the coverage probabilities and confidence level 0.95 are small in both cases.

Table 4: The Coverage probabilities of confidence intervals of $\mu_1 = 1$

|	ρ	σ				
		0.25	0.5	1	2	4
(a)	0.1	0.9538	0.9489	0.9512	0.9484	0.9502
	0.2	0.9513	0.9492	0.9429	0.9521	0.9520
	0.3	0.9472	0.9470	0.9506	0.9514	0.9505
	0.4	0.9538	0.9531	0.9497	0.9534	0.9491
	0.5	0.9484	0.9502	0.9535	0.9492	0.9488
	0.6	0.9521	0.9520	0.9533	0.9489	0.9512
	0.7	0.9514	0.9505	0.9513	0.9492	0.9429
	0.8	0.9534	0.9491	0.9472	0.9470	0.9506
	0.9	0.9492	0.9488	0.9538	0.9531	0.9497
(b)	ρ	σ				
		0.25	0.5	1	2	4
	0.1	0.9484	0.9491	0.9524	0.9490	0.9510
	0.2	0.9509	0.9486	0.9484	0.9483	0.9475
	0.3	0.9509	0.9513	0.9524	0.9520	0.9466
	0.4	0.9491	0.9509	0.9486	0.9481	0.9491
	0.5	0.9479	0.9492	0.9482	0.9535	0.9484
	0.6	0.9505	0.9473	0.9485	0.9552	0.9507
	0.7	0.9508	0.9475	0.9515	0.9491	0.9495
	0.8	0.9494	0.9468	0.9540	0.9495	0.9530
	0.9	0.9501	0.9498	0.9513	0.9478	0.9499
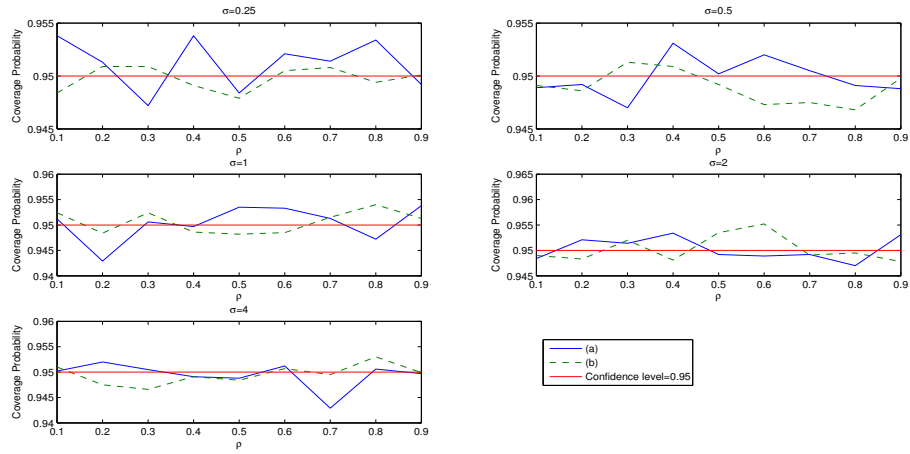
Figure 6: Compare the Coverage Probabilities of the confidence intervals of μ_1

Figure 7: Absolute Value of the Difference Between the Coverage Probabilities and 0.95

3.2 Compare the average width of confidence intervals of μ_1

Table 5: The average width of confidence intervals of $\mu_1 = 1$

	ρ	σ				
		0.25	0.5	1	2	4
(a)	0.1	0.1549	0.3099	0.6198	1.2396	2.4792
	0.2	0.1549	0.3099	0.6198	1.2396	2.4792
	0.3	0.1549	0.3099	0.6198	1.2396	2.4792
	0.4	0.1549	0.3099	0.6198	1.2396	2.4792
	0.5	0.1549	0.3099	0.6198	1.2396	2.4792
	0.6	0.1549	0.3099	0.6198	1.2396	2.4792
	0.7	0.1549	0.3099	0.6198	1.2396	2.4792
	0.8	0.1549	0.3099	0.6198	1.2396	2.4792
	0.9	0.1549	0.3099	0.6198	1.2396	2.4792
	ρ	σ				
		0.25	0.5	1	2	4
(b)	0.1	0.0631	0.1262	0.2524	0.5049	1.0097
	0.2	0.0627	0.1254	0.2507	0.5014	1.0028
	0.3	0.0620	0.1239	0.2479	0.4957	0.9914
	0.4	0.0610	0.1219	0.2438	0.4877	0.9754
	0.5	0.0596	0.1193	0.2386	0.4771	0.9542
	0.6	0.0580	0.1159	0.2318	0.4637	0.9274
	0.7	0.0559	0.1117	0.2234	0.4468	0.8937
	0.8	0.0532	0.1065	0.2129	0.4258	0.8516
	0.9	0.0499	0.0998	0.1997	0.3994	0.7987
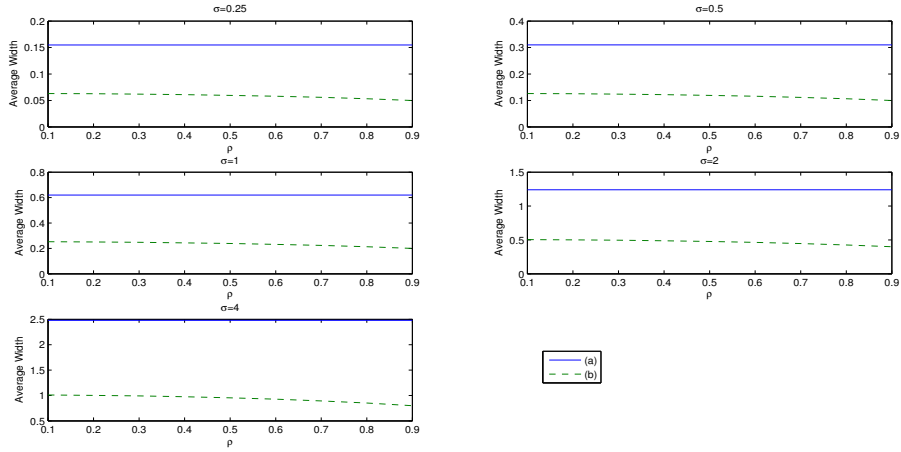
Figure 8: Compare the Average Widths of the Confidence Intervals of μ_1

As presented in Figure 8, at each level of σ, the average width of the confidence intervals is smaller in case (b) than that in (a).

3.3 Compare the average variance of $\hat{\mu}_1$

As presented in Figure 9, at each level of σ, the average variance of $\hat{\mu}_1$ is smaller in case (b) than that in (a). This result explains our finding in section 3.2: the average width of the confidence intervals is smaller in case (b) than that in (a).

Table 6: The average variance of $\hat{\mu}_1$

	ρ	σ				
		0.25	0.5	1	2	4
(a)	0.1	0.0063	0.0250	0.1000	0.4000	1.6000
	0.2	0.0063	0.0250	0.1000	0.4000	1.6000
	0.3	0.0063	0.0250	0.1000	0.4000	1.6000
	0.4	0.0063	0.0250	0.1000	0.4000	1.6000
	0.5	0.0063	0.0250	0.1000	0.4000	1.6000
	0.6	0.0063	0.0250	0.1000	0.4000	1.6000
	0.7	0.0063	0.0250	0.1000	0.4000	1.6000
	0.8	0.0063	0.0250	0.1000	0.4000	1.6000
	0.9	0.0063	0.0250	0.1000	0.4000	1.6000
	ρ	σ				
		0.25	0.5	1	2	4
(b)	0.1	0.0010	0.0041	0.0166	0.0664	0.2654
	0.2	0.0010	0.0041	0.0164	0.0654	0.2618
	0.3	9.99e-04	0.0040	0.0160	0.0640	0.2559
	0.4	9.67e-04	0.0039	0.0155	0.0619	0.2476
	0.5	9.26e-04	0.0037	0.0148	0.0593	0.2370
	0.6	8.75e-04	0.0035	0.0140	0.0560	0.2239
	0.7	8.12e-04	0.0032	0.0130	0.0520	0.2079
	0.8	7.37e-04	0.0029	0.0118	0.0472	0.1888
	0.9	6.49e-04	0.0026	0.0104	0.0415	0.1661
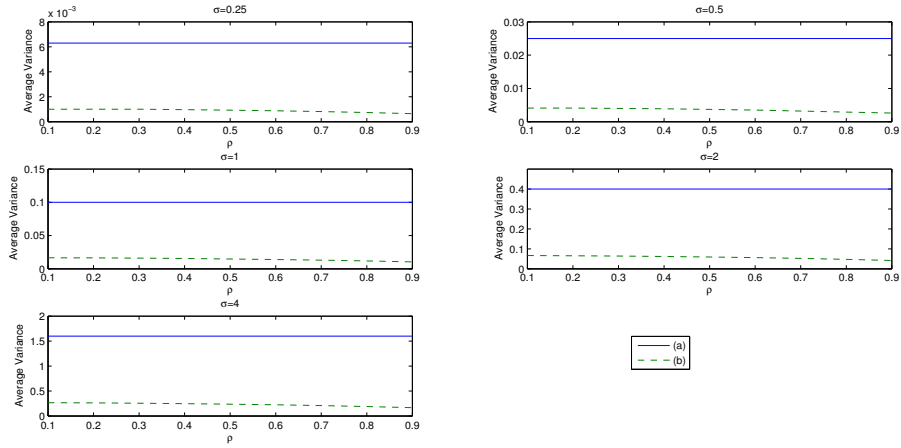
Figure 9: Compare the average variance of $\hat{\mu}_1$

4. Compare the confidence intervals of $\mu_2 - \mu_1$

4.1 The coverage probabilities of confidence intervals of $\mu_2 - \mu_1$ ($\mu_2 = 2, \mu_1 = 1$)

It can be observed from Figure 10 and Figure 11 that, in both case (a) and (b), the coverage probabilities of the 95% confidence intervals of $\mu_2 - \mu_1$ are close to 0.95. The absolute values of differences between the coverage probabilities and confidence level 0.95 are small in both cases.

Table 7: The coverage probability of confidence intervals of $\mu_2 - \mu_1$ ($\mu_2 = 2, \mu_1 = 1$)

	ρ	σ				
		0.25	0.5	1	2	4
(a)	0.1	0.9514	0.9498	0.9472	0.9522	0.9529
	0.2	0.9494	0.9518	0.9469	0.9486	0.9459
	0.3	0.9491	0.9512	0.9524	0.9480	0.9486
	0.4	0.9523	0.9458	0.9535	0.9497	0.9499
	0.5	0.9519	0.9502	0.9542	0.9530	0.9444
	0.6	0.9505	0.9479	0.9494	0.9507	0.9469
	0.7	0.9502	0.9491	0.9496	0.9497	0.9495
	0.8	0.9516	0.9472	0.9534	0.9494	0.9531
	0.9	0.9496	0.9455	0.9541	0.9443	0.9521
(b)	ρ	σ				
		0.25	0.5	1	2	4
	0.1	0.9497	0.9489	0.9523	0.9526	0.9495
	0.2	0.9500	0.9509	0.9504	0.9496	0.9483
	0.3	0.9519	0.9528	0.9528	0.9451	0.9511
	0.4	0.9503	0.9465	0.9509	0.9516	0.9479
	0.5	0.9509	0.9492	0.9476	0.9499	0.9472
	0.6	0.9512	0.9450	0.9482	0.9494	0.9528
	0.7	0.9460	0.9493	0.9491	0.9506	0.9532
	0.8	0.9527	0.9490	0.9500	0.9497	0.9525
	0.9	0.9497	0.9505	0.9541	0.9462	0.9542
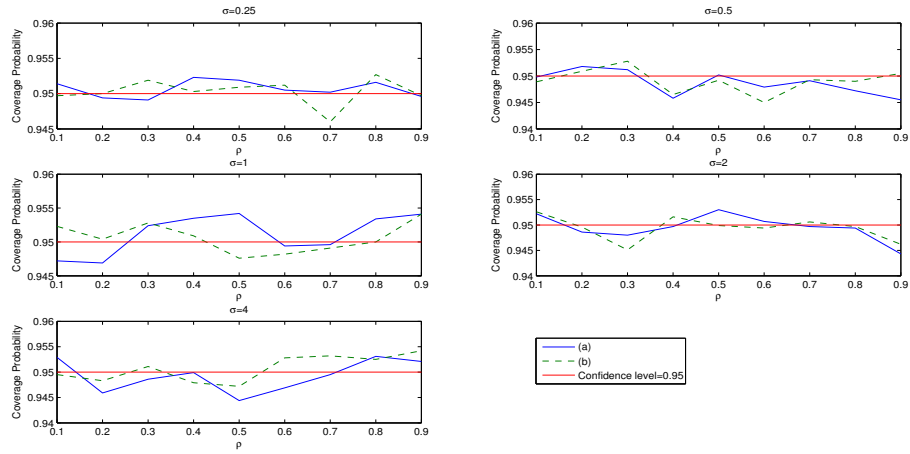
Figure 10: Compare the Coverage Probabilities of the confidence intervals of $\mu_2 - \mu_1$
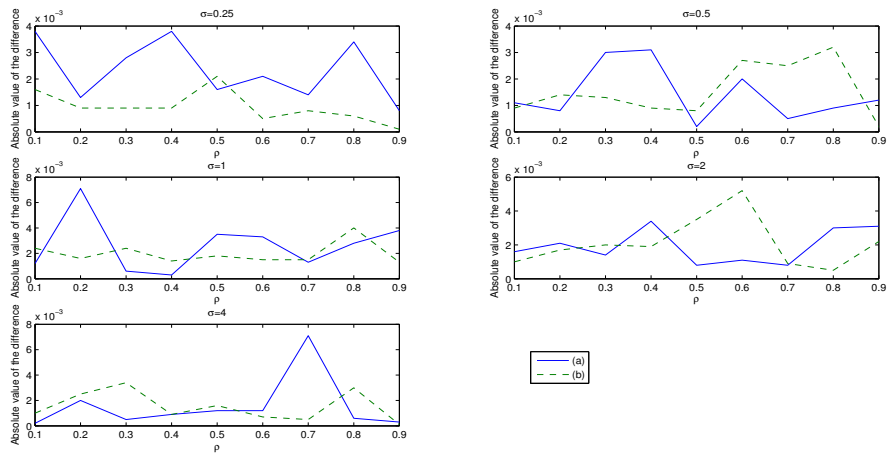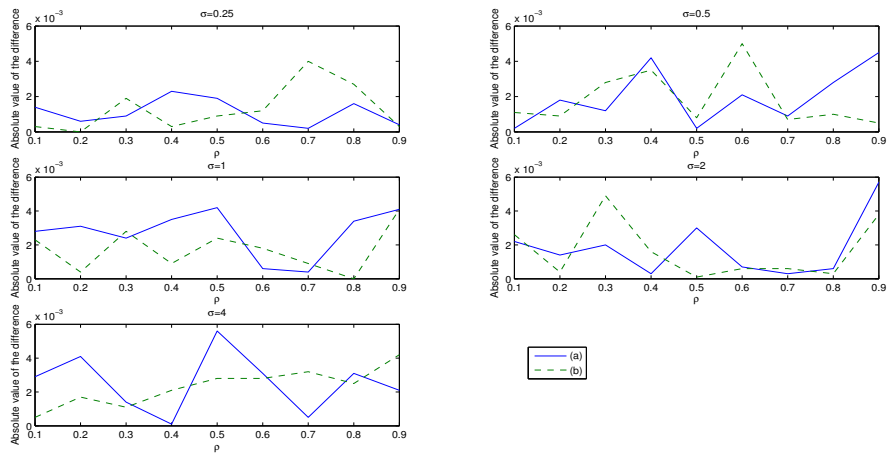

Figure 11: Absolute value of the difference between Coverage Probabilities and 0.95

4.2 Average widths of confidence intervals of $\mu_2 - \mu_1 (\mu_2 = 2, \mu_1 = 1)$

Table 8: Average widths of confidence intervals of $\mu_2 - \mu_1 (\mu_2 = 2, \mu_1 = 1)$

	ρ	σ				
		0.25	0.5	1	2	4
(a)	0.1	0.2079	0.4158	0.8315	1.6631	3.3262
	0.2	0.1960	0.3920	0.7840	1.5680	3.1359
	0.3	0.1833	0.3667	0.7334	1.4667	2.9334
	0.4	0.1697	0.3395	0.6790	1.3579	2.7158
	0.5	0.1549	0.3099	0.6198	1.2396	2.4792
	0.6	0.1386	0.2772	0.5544	1.1087	2.2174
	0.7	0.1200	0.2400	0.4801	0.9602	1.9204
	0.8	0.0980	0.1960	0.3920	0.7840	1.5680
	0.9	0.0693	0.1386	0.2772	0.5544	1.1087
	ρ	σ				
		0.25	0.5	1	2	4
(b)	0.1	0.0874	0.1747	0.3494	0.6988	1.3976
	0.2	0.0847	0.1695	0.3389	0.6779	1.3558
	0.3	0.0816	0.1631	0.3262	0.6524	1.3049
	0.4	0.0777	0.1554	0.3108	0.6216	1.2433
	0.5	0.0730	0.1461	0.2922	0.5843	1.1687
	0.6	0.0673	0.1347	0.2693	0.5386	1.0772
	0.7	0.0601	0.1203	0.2406	0.4812	0.9623
	0.8	0.0507	0.1014	0.2029	0.4058	0.8115
	0.9	0.0371	0.0742	0.1484	0.2968	0.5935
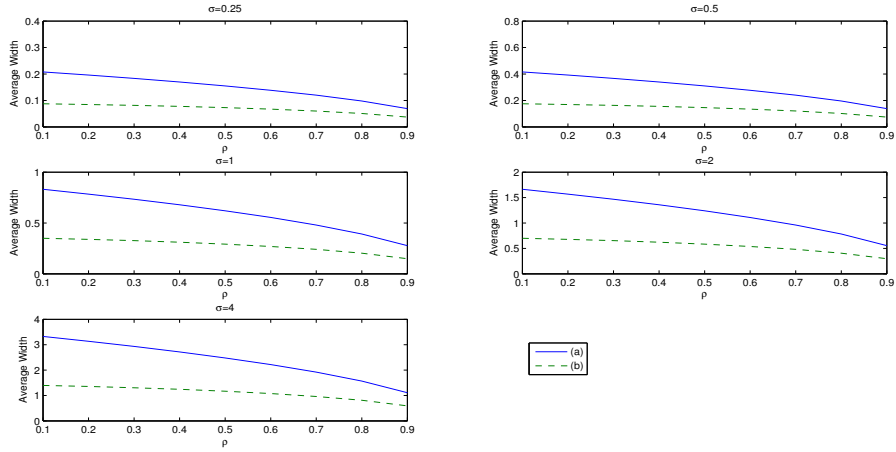
Figure 12: Compare Average Widths of Confidence Intervals of $\mu_2 - \mu_1$

Figure 12 shows that the average width of confidence intervals of $\mu_2 - \mu_1$ is decreasing as the ρ is increasing. The average widths in case (b) are always smaller than those in (a) at each level of σ.

4.3 Compare the average variance of confidence intervals of $\mu_2 - \mu_1$

As presented in Figure 13, at each level of σ, the average variance of $\hat{\mu}_2 - \hat{\mu}_1$ is smaller in case (b) than that in (a). This result explains our finding in section 4.2: the average width of the confidence intervals is smaller in case (b) than that in (a).

38

Table 9: The average variance of $\hat{\mu}_2 - \hat{\mu}_1$

	ρ	σ				
		0.25	0.5	1	2	4
(a)	0.1	0.0113	0.045	0.180	0.720	2.880
	0.2	0.0100	0.040	0.160	0.640	2.560
	0.3	0.0088	0.035	0.140	0.560	2.240
	0.4	0.0075	0.030	0.120	0.480	1.920
	0.5	0.0063	0.025	0.100	0.400	1.600
	0.6	0.0050	0.020	0.080	0.320	1.280
	0.7	0.0037	0.015	0.060	0.240	0.960
	0.8	0.0025	0.010	0.040	0.160	0.640
	0.9	0.0013	0.005	0.020	0.080	0.320
	ρ	σ				
		0.25	0.5	1	2	4
(b)	0.1	0.0020	0.0079	0.0318	0.1271	0.5085
	0.2	0.0019	0.0075	0.0299	0.1196	0.4785
	0.3	0.0017	0.0069	0.0277	0.1108	0.4432
	0.4	0.0016	0.0063	0.0251	0.1006	0.4024
	0.5	0.0014	0.0056	0.0222	0.0889	0.3556
	0.6	0.0012	0.0047	0.0189	0.0755	0.3021
	0.7	9.42e-04	0.0038	0.0151	0.0603	0.2411
	0.8	6.70e-04	0.0027	0.0107	0.0429	0.1714
	0.9	3.58e-04	0.0014	0.0057	0.0229	0.0917
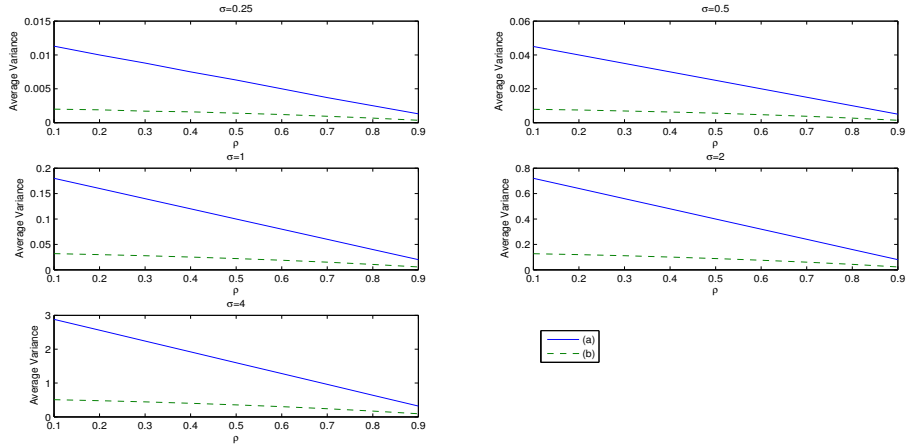
Figure 13: Compare the average variance of $\hat{\mu}_2 - \hat{\mu}_1$

VII. Simulation Study with Unknown ρ

It is assumed that ρ is unknown in this section, so the estimate of ρ is used. The formula to estimate ρ is shown by Orjuela (2013) using MLE method with three-dimensional complete data:

$$\hat{\rho} = \frac{\sum_{i=1}^{n_0}[(x_{i1} - \bar{x}_{.1})(x_{i2} - \bar{x}_{.2}) + (x_{i1} - \bar{x}_{.1})(x_{i3} - \bar{x}_{.3}) + (x_{i2} - \bar{x}_{.2})(x_{i3} - \bar{x}_{.3})]}{\sum_{i=1}^{n_0}[(x_{i1} - \bar{x}_{.1})^2 + (x_{i2} - \bar{x}_{.2})^2 + (x_{i3} - \bar{x}_{.3})^2]}.$$

Everything else is the same as that in Section VI, specifically, simulation study in this part is conducted to compare the performance of estimates in case (a) only three-dimensional observations and (b) both three-dimensional observations and all possible incomplete observations.

40

1. Compare the confidence regions of μ

Table 10: Coverage Probabilities of confidence regions of $\mu = (1, 2, 3)'$

	ρ	σ				
		0.25	0.5	1	2	4
(a)	0.1	0.9286	0.9273	0.9290	0.9302	0.9262
	0.2	0.9296	0.9273	0.9271	0.9301	0.9309
	0.3	0.9301	0.9266	0.9336	0.9318	0.9255
	0.4	0.9313	0.9257	0.9315	0.9322	0.9267
	0.5	0.9296	0.9269	0.9294	0.9288	0.9239
	0.6	0.9250	0.9272	0.9283	0.9231	0.9257
	0.7	0.9305	0.9263	0.9258	0.9180	0.9256
	0.8	0.9277	0.9212	0.9287	0.9192	0.9289
	0.9	0.9240	0.9199	0.9259	0.9195	0.9255
	ρ	σ				
		0.25	0.5	1	2	4
(b)	0.1	0.9370	0.9372	0.9377	0.9378	0.9328
	0.2	0.9341	0.9364	0.9391	0.9339	0.9330
	0.3	0.9372	0.9398	0.9361	0.9362	0.9320
	0.4	0.9374	0.9318	0.9387	0.9324	0.9289
	0.5	0.9322	0.9287	0.9278	0.9326	0.9319
	0.6	0.9288	0.9251	0.9323	0.9305	0.9327
	0.7	0.9296	0.9271	0.9254	0.9267	0.9314
	0.8	0.9264	0.9194	0.9291	0.9303	0.9275
	0.9	0.9246	0.9217	0.9289	0.9259	0.9263
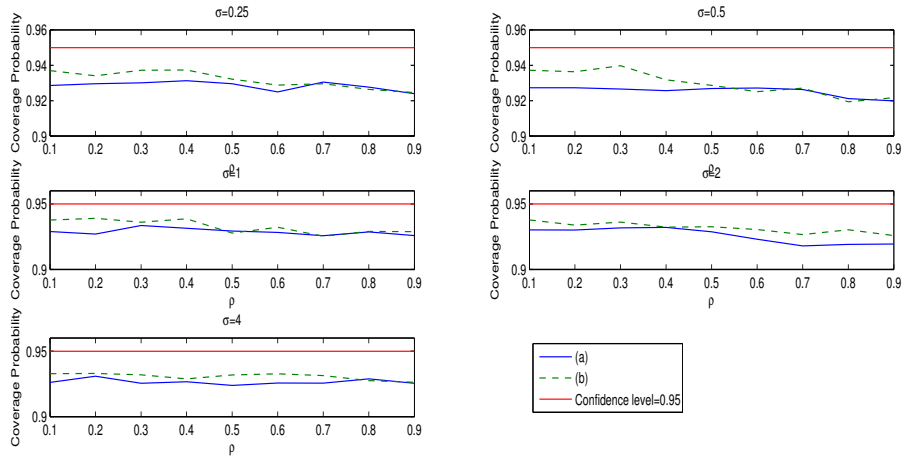
Figure 14: Compare the Coverage Probabilities of the Confidence Regions of $\boldsymbol{\mu}$
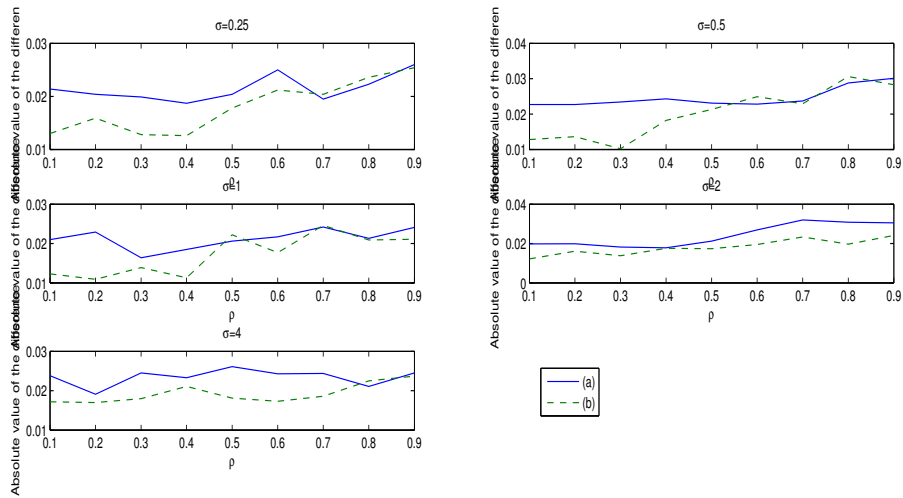
Figure 15: Absolute value of the difference between Coverage Probabilities and 0.95

Using the estimates of ρ would make coverage probabilities a little lower than the confidence level 0.95 for both case (a) and (b). It can be noted from Figure 15 that the absolute values of the difference between Coverage Probabilities and 0.95 are around 0.02 for each level of σ.

2. Compare the probability of type I error and power of testing

$H_0 : \boldsymbol{\mu} = \mathbf{0} \ vs. H_a : \boldsymbol{\mu} \neq \mathbf{0}$

2.1 Compare the probability of type I error

Table 11: The Probability of Type I error

| | ρ | σ | | | | |
		0.25	0.5	1	2	4
	0.1	0.0680	0.0627	0.0648	0.0692	0.0705
	0.2	0.0665	0.0667	0.0681	0.0743	0.0695
	0.3	0.0647	0.0662	0.0693	0.0742	0.0692
(a)	0.4	0.0666	0.0671	0.0680	0.0700	0.0751
	0.5	0.0647	0.0679	0.0658	0.0747	0.0678
	0.6	0.0701	0.0721	0.0687	0.0726	0.0711
	0.7	0.0683	0.0654	0.0732	0.0746	0.0700
	0.8	0.0733	0.0697	0.0781	0.0743	0.0768
	0.9	0.0716	0.0698	0.0775	0.0784	0.0821
	ρ	σ				
		0.25	0.5	1	2	4
	0.1	0.0607	0.0644	0.0619	0.0657	0.0646
	0.2	0.0581	0.0627	0.0656	0.0676	0.0639
	0.3	0.0628	0.0682	0.0621	0.0653	0.0636
(b)	0.4	0.0641	0.0637	0.0601	0.0702	0.0688
	0.5	0.0648	0.0637	0.0678	0.0659	0.0646
	0.6	0.0651	0.0667	0.0657	0.0730	0.0626
	0.7	0.0618	0.0729	0.0688	0.0723	0.0702
	0.8	0.0679	0.0701	0.0774	0.0720	0.0759
	0.9	0.0706	0.0731	0.0753	0.0736	0.0730

Figure 16: Compare the Type I Error
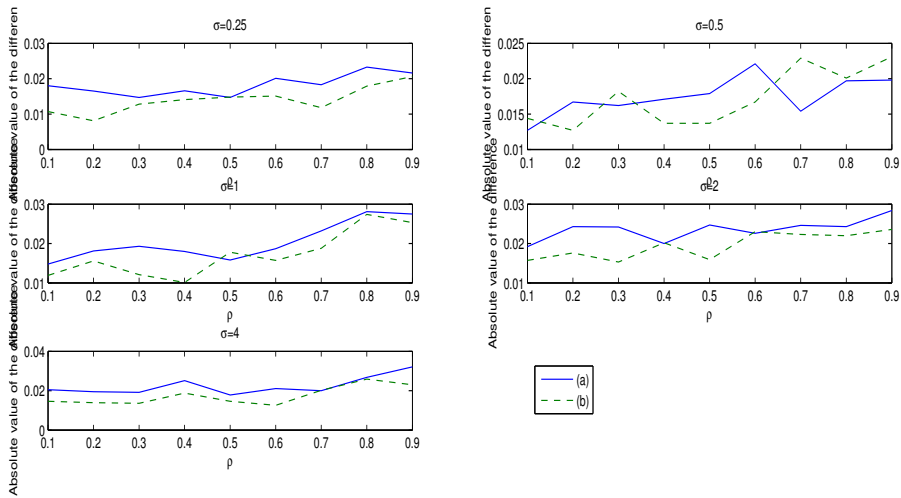
Figure 17: Absolute value of difference between Type I Error and $\alpha = 0.05$

When using the estimate of ρ, type I errors in both (a) and (b) are a little higher than $\alpha = 0.05$ at each level of σ. It can be observed more clearly from Figure 17 that the absolute values of difference between Type I Errors and $\alpha = 0.05$ are around 0.02 in both cases.

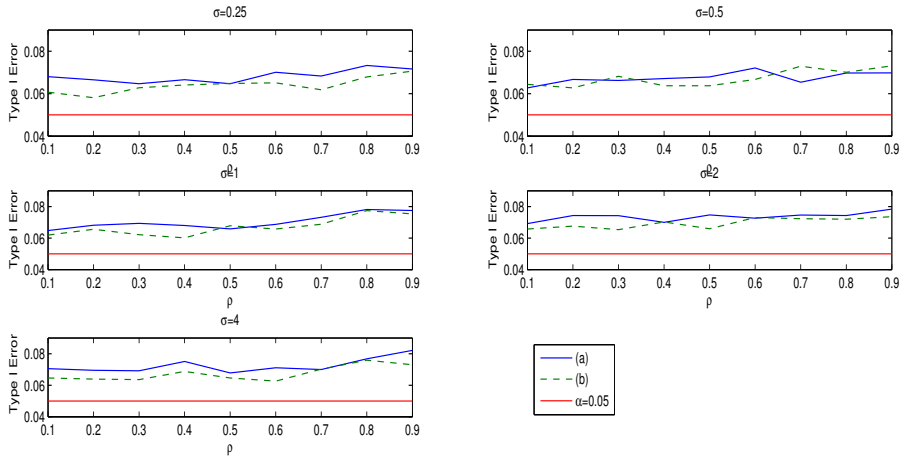**2.2 Compare the power of testing** $H_0 : \boldsymbol{\mu} = \boldsymbol{0} \; vs. H_a : \boldsymbol{\mu} \neq \boldsymbol{0}$

Table 12: The Power of testing $H_0 : \boldsymbol{\mu} = \boldsymbol{0} \; vs. H_a : \boldsymbol{\mu} = (1, 2, 3)'$

	ρ	σ				
		6	8	10	12	14
(a)	0.1	0.3991	0.2495	0.1766	0.1372	0.1274
	0.2	0.3682	0.2307	0.1690	0.1316	0.1125
	0.3	0.3381	0.2141	0.1507	0.1317	0.1155
	0.4	0.3240	0.2126	0.1584	0.1269	0.1126
	0.5	0.3401	0.2106	0.1522	0.1247	0.1069
	0.6	0.3380	0.2114	0.1607	0.1251	0.1157
	0.7	0.3665	0.2321	0.1676	0.1357	0.1172
	0.8	0.4361	0.2778	0.1964	0.1579	0.1256
	0.9	0.6122	0.3983	0.2768	0.2196	0.1824
(b)	ρ	σ				
		6	8	10	12	14
	0.1	0.9884	0.8725	0.6868	0.5139	0.3915
	0.2	0.9858	0.8588	0.6586	0.4867	0.3807
	0.3	0.9826	0.8419	0.6481	0.4854	0.3742
	0.4	0.9804	0.8408	0.6323	0.4691	0.3638
	0.5	0.9773	0.8363	0.6310	0.4748	0.3558
	0.6	0.9796	0.8387	0.6373	0.4732	0.3708
	0.7	0.9841	0.8536	0.6620	0.4884	0.3776
	0.8	0.9900	0.8907	0.6945	0.5423	0.4248
	0.9	0.9985	0.9568	0.8241	0.6726	0.5358
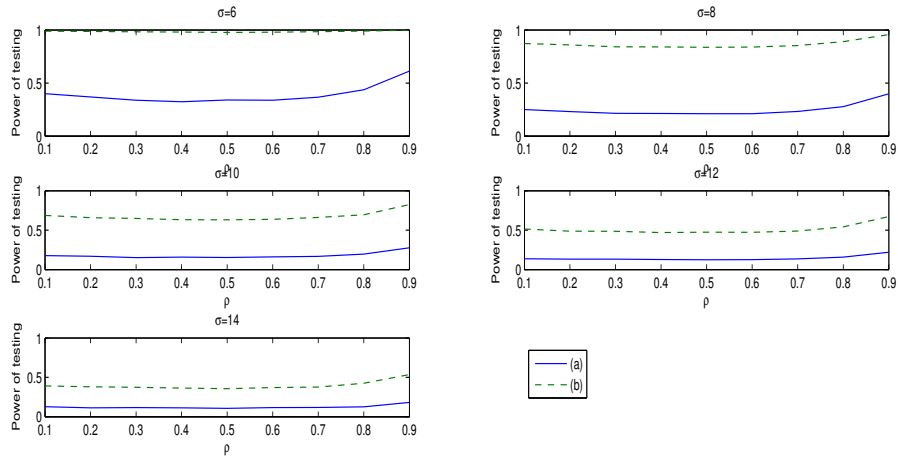
Figure 18: Compare the Power of Testing $H_0 : \boldsymbol{\mu} = \mathbf{0} \ vs. H_a : \boldsymbol{\mu} = (1, 2, 3)'$

With regard to the power of testing, it can be observed that, when the level of σ is fixed, case (b) has higher power than that of case (a).

3. Compare the confidence intervals of μ_1

3.1 Compare the coverage probabilities of confidence intervals of μ_1

The coverage probabilities of the 95% confidence intervals are close to 0.95 for all cases, and the differences between coverage probabilities and confidence level 0.95 are small.

Table 13: The Coverage probabilities of confidence intervals of $\mu_1 = 1$

|	ρ	σ				
		0.25	0.5	1	2	4
(a)	0.1	0.9538	0.9489	0.9512	0.9484	0.9502
	0.2	0.9513	0.9492	0.9429	0.9521	0.9520
	0.3	0.9472	0.9470	0.9506	0.9514	0.9505
	0.4	0.9538	0.9531	0.9497	0.9534	0.9491
	0.5	0.9484	0.9502	0.9535	0.9492	0.9488
	0.6	0.9521	0.9520	0.9533	0.9489	0.9512
	0.7	0.9514	0.9505	0.9513	0.9492	0.9429
	0.8	0.9534	0.9491	0.9472	0.9470	0.9506
	0.9	0.9492	0.9488	0.9538	0.9531	0.9497
	ρ	σ				
		0.25	0.5	1	2	4
(b)	0.1	0.9440	0.9443	0.9485	0.9442	0.9446
	0.2	0.9453	0.9430	0.9428	0.9439	0.9440
	0.3	0.9476	0.9460	0.9510	0.9485	0.9415
	0.4	0.9443	0.9456	0.9436	0.9446	0.9432
	0.5	0.9442	0.9451	0.9446	0.9472	0.9445
	0.6	0.9463	0.9460	0.9459	0.9490	0.9481
	0.7	0.9488	0.9472	0.9507	0.9452	0.9478
	0.8	0.9482	0.9468	0.9525	0.9479	0.9539
	0.9	0.9516	0.9512	0.9529	0.9497	0.9513
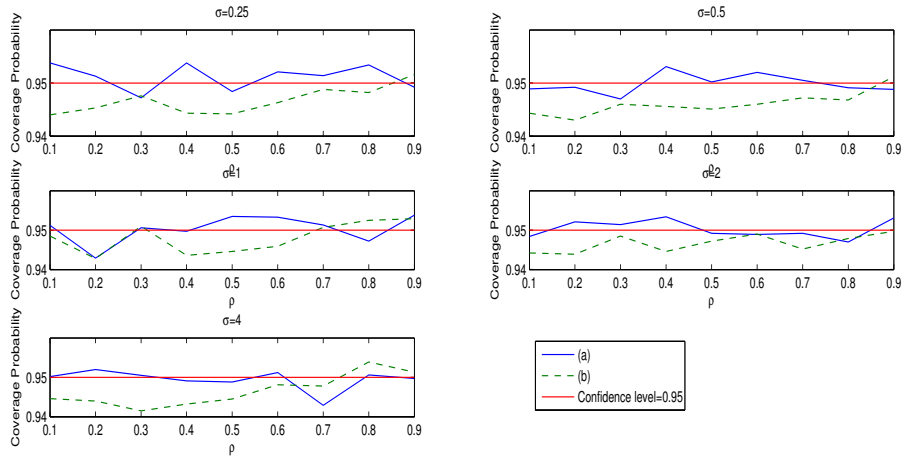
Figure 19: Compare the Coverage Probabilities of the Confidence Intervals of $\mu_1 = 1$
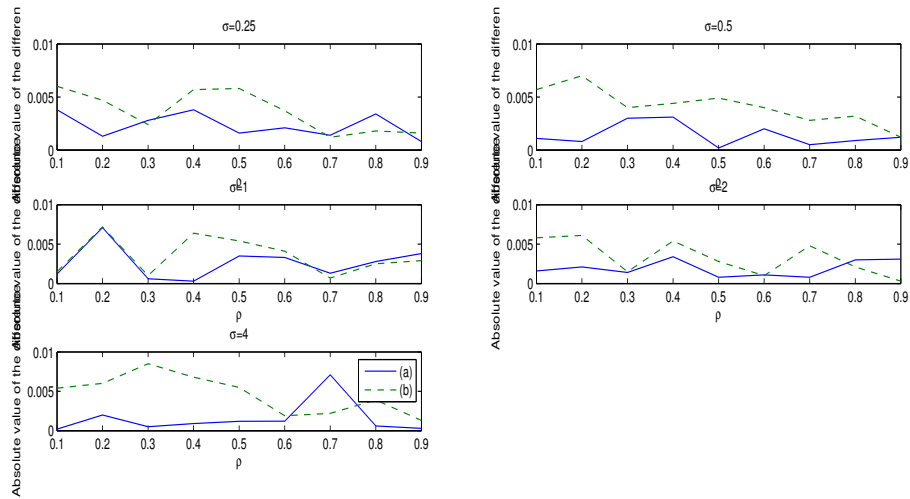
Figure 20: Absolute value of the difference between Coverage Probabilities and 0.95

3.2 Compare the average width of confidence intervals of μ_1

Table 14: The average width of confidence intervals of $\mu_1 = 1$

	ρ	σ				
		0.25	0.5	1	2	4
(a)	0.1	0.1549	0.3099	0.6198	1.2396	2.4792
	0.2	0.1549	0.3099	0.6198	1.2396	2.4792
	0.3	0.1549	0.3099	0.6198	1.2396	2.4792
	0.4	0.1549	0.3099	0.6198	1.2396	2.4792
	0.5	0.1549	0.3099	0.6198	1.2396	2.4792
	0.6	0.1549	0.3099	0.6198	1.2396	2.4792
	0.7	0.1549	0.3099	0.6198	1.2396	2.4792
	0.8	0.1549	0.3099	0.6198	1.2396	2.4792
	0.9	0.1549	0.3099	0.6198	1.2396	2.4792
	ρ	σ				
		0.25	0.5	1	2	4
(b)	0.1	0.0625	0.1251	0.2501	0.5002	1.0003
	0.2	0.0621	0.1243	0.2485	0.4970	0.9942
	0.3	0.0615	0.1230	0.2462	0.4919	0.9835
	0.4	0.0606	0.1213	0.2426	0.4850	0.9700
	0.5	0.0595	0.1189	0.2378	0.4757	0.9516
	0.6	0.0580	0.1161	0.2320	0.4644	0.9279
	0.7	0.0561	0.1123	0.2245	0.4493	0.8992
	0.8	0.0537	0.1074	0.2149	0.4297	0.8612
	0.9	0.0505	0.1009	0.2021	0.4045	0.8084
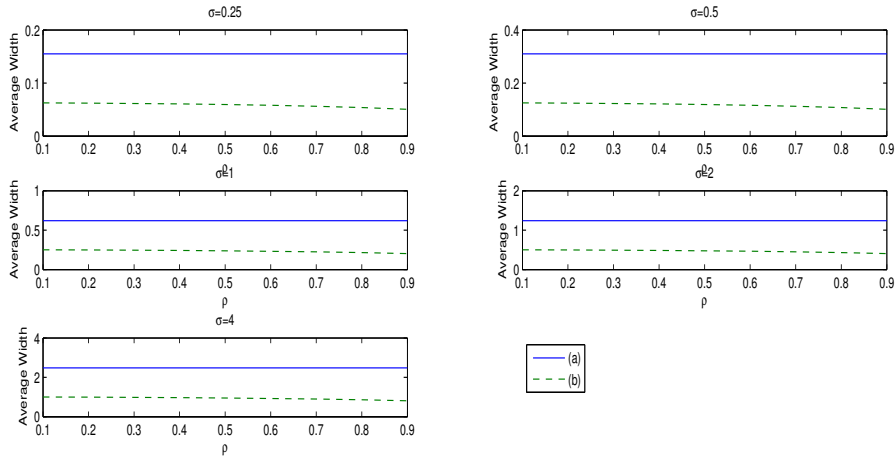
Figure 21: Compare the Average Widths of the Confidence Intervals of $\mu_1 = 1$

With regard to the average width of the confidence intervals of $\mu_1 = 1$, the confidence intervals in (b) are always narrower than those in (a).

3.3 Compare the average variance of $\hat{\mu}_1$

As presented in Figure 22, at each level of σ, the average variance of $\hat{\mu}_1$ is smaller in case (b) than that in (a). This result explains our finding in section 3.2: the average width of the confidence intervals is smaller in case (b) than that in (a).

Table 15: The average variance of $\hat{\mu}_1$

	ρ	σ				
		0.25	0.5	1	2	4
(a)	0.1	0.0063	0.0250	0.1000	0.4000	1.6000
	0.2	0.0063	0.0250	0.1000	0.4000	1.6000
	0.3	0.0063	0.0250	0.1000	0.4000	1.6000
	0.4	0.0063	0.0250	0.1000	0.4000	1.6000
	0.5	0.0063	0.0250	0.1000	0.4000	1.6000
	0.6	0.0063	0.0250	0.1000	0.4000	1.6000
	0.7	0.0063	0.0250	0.1000	0.4000	1.6000
	0.8	0.0063	0.0250	0.1000	0.4000	1.6000
	0.9	0.0063	0.0250	0.1000	0.4000	1.6000
(b)	ρ	σ				
		0.25	0.5	1	2	4
	0.1	0.0010	0.0041	0.0163	0.0652	0.2606
	0.2	0.0010	0.0040	0.0161	0.0644	0.2573
	0.3	9.86e-04	0.0039	0.0158	0.0631	0.2522
	0.4	9.57e-04	0.0038	0.0153	0.0613	0.2452
	0.5	9.23e-04	0.0037	0.0148	0.0590	0.2362
	0.6	8.78e-04	0.0035	0.0140	0.0562	0.2248
	0.7	8.23e-04	0.0033	0.0132	0.0528	0.2107
	0.8	7.53e-04	0.0030	0.0121	0.0483	0.1928
	0.9	6.64e-04	0.0027	0.0106	0.0425	0.1704
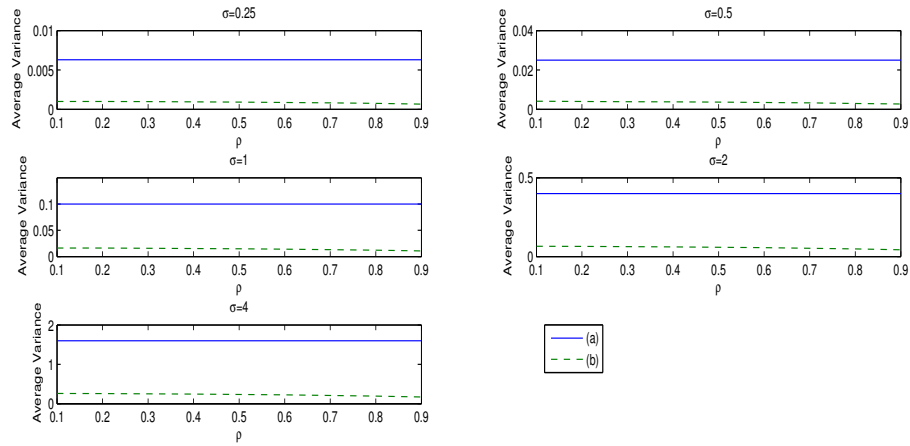
Figure 22: Compare the average variance of $\hat{\mu}_1$

4. Compare the confidence intervals of $\mu_2 - \mu_1$

4.1 The coverage probabilities of confidence intervals of $\mu_2 - \mu_1$ ($\mu_2 = 2, \mu_1 = 1$)

Using the estimates of ρ would make coverage probabilities a little lower than the confidence level 0.95. It can be noted from Figure 24 that the absolute values of the difference are around 0.02 for each level of σ.

Table 16: The coverage probabilities of confidence intervals of $\mu_2 - \mu_1$ ($\mu_2 = 2, \mu_1 = 1$)

|		σ				
	ρ	0.25	0.5	1	2	4
(a)	0.1	0.9464	0.9433	0.9424	0.9444	0.9442
	0.2	0.9426	0.9429	0.9378	0.9426	0.9385
	0.3	0.9414	0.9398	0.9431	0.9388	0.9389
	0.4	0.9414	0.9364	0.9408	0.9406	0.9384
	0.5	0.9404	0.9394	0.9433	0.9425	0.9358
	0.6	0.9394	0.9357	0.9374	0.9358	0.9345
	0.7	0.9342	0.9332	0.9363	0.9333	0.9344
	0.8	0.9346	0.9346	0.9381	0.9344	0.9362
	0.9	0.9359	0.9320	0.9368	0.9247	0.9308
		σ				
	ρ	0.25	0.5	1	2	4
(b)	0.1	0.9423	0.9407	0.9465	0.9457	0.9408
	0.2	0.9406	0.9422	0.9446	0.9422	0.9401
	0.3	0.9423	0.9428	0.9442	0.9363	0.9390
	0.4	0.9401	0.9374	0.9402	0.9400	0.9385
	0.5	0.9411	0.9343	0.9354	0.9394	0.9339
	0.6	0.9343	0.9326	0.9347	0.9385	0.9380
	0.7	0.9294	0.9346	0.9324	0.9354	0.9397
	0.8	0.9337	0.9296	0.9351	0.9353	0.9362
	0.9	0.9330	0.9316	0.9382	0.9335	0.9376
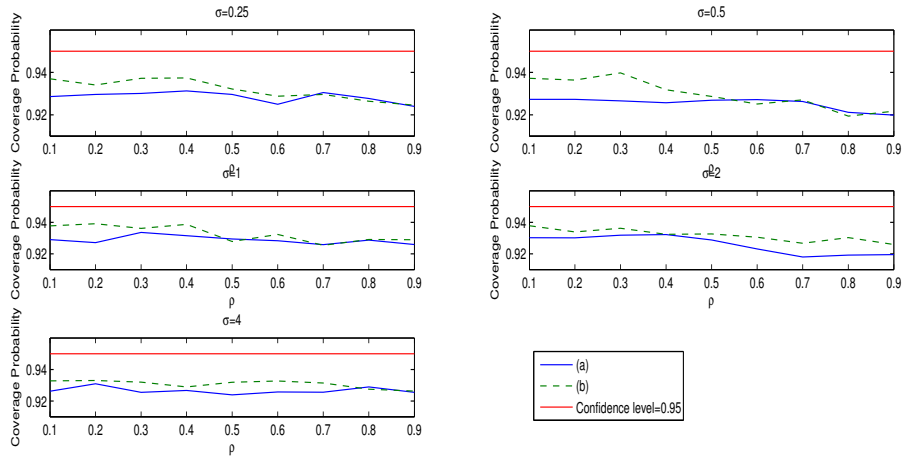
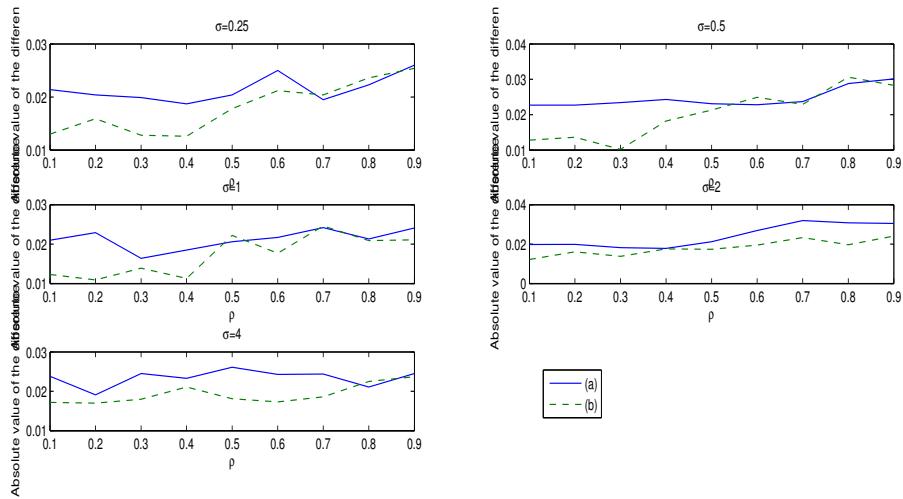Figure 23: Compare the Coverage Probabilities of the Confidence Intervals of $\mu_2 - \mu_1$

Figure 24: Absolute value of the difference between Coverage Probabilities and 0.95

4.2 The average width of confidence intervals of $\mu_2 - \mu_1 (\mu_2 = 2, \mu_1 = 1)$

Table 17: The average width of confidence intervals of $\mu_2 - \mu_1 (\mu_2 = 2, \mu_1 = 1)$

|	ρ	σ				
		0.25	0.5	1	2	4
(a)	0.1	0.2075	0.4156	0.8292	1.6584	3.3135
	0.2	0.1962	0.3924	0.7851	1.5691	3.1371
	0.3	0.1841	0.3679	0.7403	1.4729	2.9458
	0.4	0.1717	0.3441	0.6872	1.3722	2.7433
	0.5	0.1576	0.3149	0.6290	1.2585	2.5186
	0.6	0.1420	0.2840	0.5678	1.1398	2.2663
	0.7	0.1241	0.2486	0.4955	0.9938	1.9924
	0.8	0.1024	0.2046	0.4092	0.8175	1.6526
	0.9	0.0732	0.1464	0.2947	0.5930	1.1801
	ρ	σ				
		0.25	0.5	1	2	4
(b)	0.1	0.0865	0.1731	0.3458	0.6915	1.3822
	0.2	0.0839	0.1678	0.3357	0.6711	1.3425
	0.3	0.0808	0.1616	0.3244	0.6467	1.2928
	0.4	0.0773	0.1548	0.3094	0.6181	1.2362
	0.5	0.0729	0.1458	0.2913	0.5828	1.1664
	0.6	0.0676	0.1353	0.2704	0.5422	1.0805
	0.7	0.0609	0.1220	0.2434	0.4879	0.9779
	0.8	0.0520	0.1040	0.2080	0.4155	0.8383
	0.9	0.0387	0.0774	0.1556	0.3128	0.6231
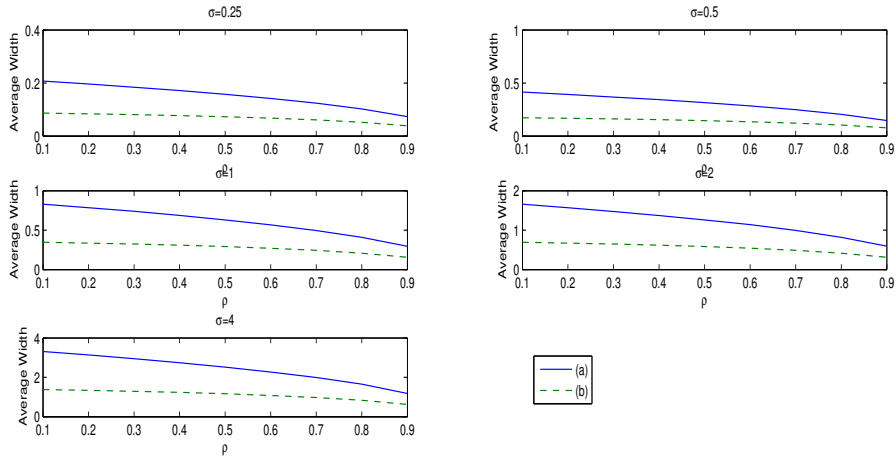
Figure 25: Compare the Average Widths of the Confidence Intervals of $\mu_2 - \mu_1$

With regard to the average width of the confidence intervals of $\mu_2 - \mu_1$, the confidence intervals in (b) are always narrower than those in (a).

4.3 Compare the average variance of $\hat{\mu}_2 - \hat{\mu}_1$

As presented in Figure 26, at each level of σ, the average variance of $\hat{\mu}_1$ is smaller in case (b) than that in (a). This result explains our finding in section 4.2: the average width of the confidence intervals is smaller in case (b) than that in (a).

Table 18: The average variance of $\hat{\mu}_1$

|		σ				
	ρ	0.25	0.5	1	2	4
(a)	0.1	0.0113	0.0453	0.1815	0.7256	2.9070
	0.2	0.0102	0.0408	0.1630	0.6511	2.6049
	0.3	0.0090	0.0361	0.1438	0.5782	2.3182
	0.4	0.0079	0.0314	0.1262	0.5029	2.0145
	0.5	0.0067	0.0270	0.1071	0.4275	1.7121
	0.6	0.0054	0.0220	0.0874	0.3498	1.4001
	0.7	0.0043	0.0168	0.0678	0.2714	1.0812
	0.8	0.0029	0.0116	0.0466	0.1867	0.7402
	0.9	0.0015	0.0061	0.0243	0.0969	0.3915
		σ				
	ρ	0.25	0.5	1	2	4
(b)	0.1	0.0020	0.0078	0.0312	0.1250	0.5001
	0.2	0.0018	0.0074	0.0295	0.1179	0.4714
	0.3	0.0017	0.0069	0.0274	0.1099	0.4399
	0.4	0.0016	0.0063	0.0252	0.1005	0.4025
	0.5	0.0014	0.0057	0.0226	0.0900	0.3604
	0.6	0.0012	0.0049	0.0195	0.0780	0.3117
	0.7	0.0010	0.0040	0.0160	0.0641	0.2555
	0.8	7.39e-04	0.0029	0.0118	0.0473	0.1878
	0.9	4.18e-04	0.0017	0.0067	0.0267	0.1075
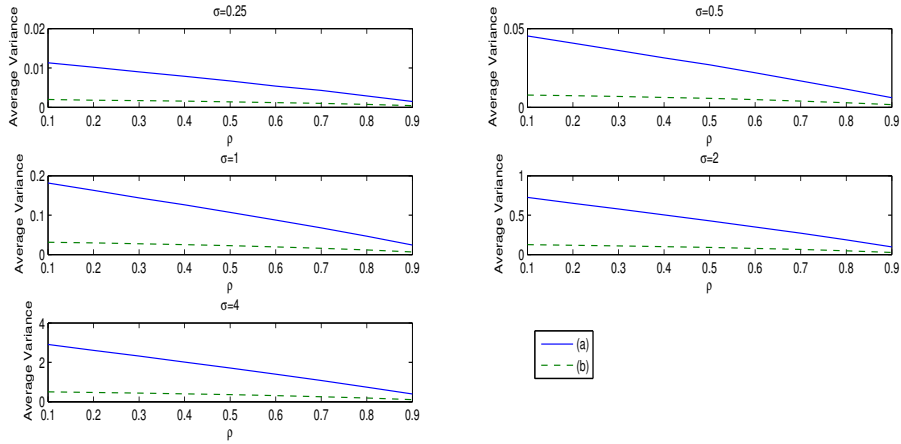
Figure 26: Compare the average variance of $\hat{\mu}_1$

VIII. Conclusion

It can be concluded from the numerical study that, compared to estimators in the case of omitting observations with missing data, using the considered estimators in this thesis would generally lead to better performance in the following aspects: higher power of testing and shorter confidence intervals, while keeping the coverage probabilities and type I errors close to the desired α level. Furthermore, when ρ is unknown, using the estimate of ρ would lead to the same conclusion.

REFERENCES

Anderson, T. W. (1957). *Maximum Likelihood Estimates for a Multivariate Normal Distribution when some Observations are Missing.* Journal of the American Statistical Association, 52: 200-203.

Chang, Wan-Ying and Richards, Donald St.P. (2009). *Finite-sample inference with monotone incomplete multivariate normal data, I.* Journal of Multivariate Analysis, 100: 1883-1899.

Edgett, G. L. (1956). *Multiple Regression with Missing Observations Among the Independent Variables.* Journal of the American Statistical Association 51. 122-131.

Hocking, R. R. and Smith, William B. (1968). *Estimation of Parameters in the Multivariate Normal Distribution with Missing Observations.* Journal of the American Statistical Association, 63: 159-173.

J. N. Srivastava and M. K. Zaatar(1973). *A Monte Carlo Comparison of Four Estimators of the Dispersion Matrix of a Bivariate Normal Population, Using Incomplete Data.* Journal of the American Statistical Association , Vol. 68, No. 341 (Mar., 1973), pp. 180-183.

Kim, Kevin H. and Bentler, Peter M. (2002). *Tests of homogeneity of means and covariance matrices for multivariate incomplete data.* Psychometrika, 67: 609-624.

Little, Roderick J. A. (1994). *A Class of Pattern-Mixture Models for Normal Incomplete Data.* Biometrika, 81: 471-483.

Lord. F. M. (1955). *Estimation of Parameters from Incomplete Data.* Journal of the American Statistical Association, 50, 870-876.

Mara del Pilar Orjuela garavito (2013). *A study on correlation of Bivariate and Trivariate Normal Models.*

Marios A. G. Viana, Hak-Myung Lee (2006). *Correlation Analysis of Ordered Symmetrically Dependent Observations and Their concomitants of Order Statistics.* Canadian Journal of Statistics 34 (2):327 - 340.

Matthai, A. (1951). *Estimation of Parameters from Incomplete Data with Application to Design of Sample Surveys.* Sankhya 2, 11, 145-152.

Nicholas J. Horton and Ken P. Kleinman (2007). *Much A do About Nothing: A Comparison of Missing Data Methods and Software to Fit Incomplete Data Regression Models.* Am Stat. Feb, 61 (1):79-90.

Nicholson, George E., Jr (1957). *Estimation of Parameters from Incomplete Multivariate Samples.* Journal of the American Statistical Association, 52: 523-526.

Philip K. Hopke, Chuanhai Liu, and Donald B. Rubin.(2001). *Multiple imputation for multivariate data with missing and below-threshold measurements: Time-series concentrations of pollutants in the Arctic.* Biometrics 57: 22-33.

Richard A. Johnson, Dean W. Wichern (2007). *Applied Multivariate Statistical Analysis (6th Edition).*

S. F. Buck (1960). *A Method of Estimation of Missing Values in Multivariate Data Suitable for use with an Electronic Computer.* Journal of the Royal Statistical Society. Series B (Methodological), 22, No. 2: 302-306.

Steffen Lauritzen (2009). *Wishart and Inverse Wishart Distributions.* University of Oxford, BS2 Statistical Inference, Lecture 9, Hilary Term 2009.

Wilks, Samuel S. (1932). *Moments and distributions of estimates of population parameters from fragmentary samples.* The Annals of Mathematical Statistics, Vol.3, No.3,163-195.

Y.L Tong (1990). *The Multivariate Normal Distribution.* Springer-Verlag New York Inc.