Thailand Statistician
July 2011; 9(2) : 151-169
http://statassoc.or.th
Contributed paper

Jackknife Maximum Likelihood Estimates for a Bivariate Normal Distribution with Missing Data

Juthaphorn Sinsomboonthong
Department of Statistics, Faculty of Science, Kasetsart University, Bangkok 10900, Thailand.
E-mail: fscijps@ku.ac.th

Received: 10 February 2011
Accepted: 30 May 2011

Abstract

In this paper, the estimator of bivariate normal distribution for incomplete data called the J_Anderson estimator is proposed. The first r pairs of (X_1, X_2) were distributed as a bivariate normal distribution with mean vector (μ_1, μ_2) and a variance covariance matrix $\Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix}$, whereas the rest $n-r$ observations of X_1 were distributed as a normal distribution with mean μ_1 and variance σ_1^2. The factored maximum likelihood estimators, $\hat{\mu}_{1_{And}}$, $\hat{\mu}_{2_{And}}$, $\hat{\sigma}^2_{1_{And}}$ and $\hat{\sigma}^2_{2_{And}}$, were proposed by Anderson. In addition, the Anderson's method was found that $\hat{\sigma}^2_{1_{And}}$ and $\hat{\sigma}^2_{2_{And}}$ were biased estimators of σ_1^2 and σ_2^2 respectively. In this study, Jackknife's method of bias reduction was introduced to modify $\hat{\sigma}^2_{1_{And}}$ and $\hat{\sigma}^2_{2_{And}}$. In this case, the proposed estimators were denoted by $\hat{\sigma}^2_{1_{J_And}}$ and $\hat{\sigma}^2_{2_{J_And}}$. Moreover, the simulation study found that the absolute relative biases of $E\left(\hat{\sigma}^2_{1_{J_And}}\right)$ and $E\left(\hat{\sigma}^2_{2_{J_And}}\right)$ were smaller

than those of $E\left(\hat{\sigma}^2_{1_{And}}\right)$ and $E\left(\hat{\sigma}^2_{2_{And}}\right)$ respectively whatever sample size and percentage of missing data. Additionally, the mean square error of the proposed estimators seemed to decrease when the sample size was large whatever percentage of missing data.

Keywords: biased estimator, bivariate normal distribution, Jackknife's method, missing data, normal distribution.

1. Introduction

A sample from a bivariate normal population is such that not all the observations are observed with respect to all the variables and it is desired to estimate the population parameters from the sample. The sample data may be incomplete for various reasons: the cost of collecting the data may make it inadvisable to collect all of it; the collection of part of the data may be impossible. Moreover, incomplete data usually introduce bias and inefficiency in parameter estimation.

Principally, although incomplete data may possibly be analyzed using standard statistical methods through which missing data are ignored, an important limitation is that those methods are only appropriate for studies which contain fewer missing data. Moreover, standard statistical approaches can also cause deficiency of data, since incomplete cases are completely discarded. Data deficiency always causes imprecision and also escalation in biases [1]. In addition, Acock [2] mentioned that these may reduce or exaggerate statistical power. Likewise, Fitzmaurice [3], Gorelick [4], Roth et al. [5], Rotnitzky and Wypij [6] mentioned that these may possibly result in invalid conclusions, since degree of bias and loss of precision is dependent not only on the fraction of complete cases and patterns of missing data, but also differences between the complete and incomplete cases, and the parameters of interest.

Likelihood function-based methods have consequently been of interest and employed in the management of missing data, particularly the analysis of incomplete data sets. Scholars, such as Anderson [7], Edgett [8], and Lord [9] examined the application of estimating parameters of trivariate and multivariate normal population by using maximum likelihood estimation. In Particular, Lord [9], studied problems when estimating parameters of trivariate normal populations from file matching incomplete data by using maximum likelihood estimation. Subsequently, Edgett [8] proposed maximum likelihood estimators for the parameters of a trivariate normal distribution from univariate incomplete data and further applied the findings on the linear regression model. Later,

Anderson [7] introduced a factored likelihood function for normal data and estimated parameters of a multivariate normal population with incomplete data. Further, Little and Rubin [1], in reference to Anderson [7], presented an appropriate factorization of the likelihood function for a general monotone missing data pattern.

Generally, in the present study incomplete data with bivariate normal distribution were examined. On the first r pairs of (X_1, X_2) were completely observed with bivariate normal distribution but, the rest $n-r$ observations of X_2 were lost and there were only $n-r$ observations $(0 < r < n)$ of X_1 collected. The factored likelihood function of Anderson [7] was desired to estimate μ_1, μ_2, σ_1^2 and σ_2^2. In addition, the estimators of σ_1^2 and σ_2^2 from Anderson's method were biased estimators.

Efron and Tibshirani [10], Kim and Anderson [11], Smith and Pontius [12] have applied Jackknife's method [13] of bias reduction to the estimation of parameters. They shown the biases of the estimators were decreasing to zero.

In this study, the parameters μ_1, μ_2, σ_1^2 and σ_2^2 from a bivariate normal distribution with missing data were derived by Anderson's method [7] and the reduction of bias of Anderson's estimators using Jackknife's method [13]. In addition, the proposed estimators were called J_Anderson estimator.

2. The Proposed Estimators

2.1 The Factored Maximum Likelihood Estimators

Consider the incomplete bivariate sample from a bivariate normal distribution with mean vector (μ_1, μ_2), a variance covariance matrix $\Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix}$ and correlation coefficient ρ.

Observations:	**1**	**2**	**…**	**r**	**r+1**	**…**	**n**
Variable 1:	x_{11}	x_{12}	…	x_{1r}	$x_{1,r+1}$	…	x_{1n}
Variable 2:	x_{21}	x_{22}	…	x_{2r}			

Figure 1. Monotone missing data pattern for a bivariate normal distribution.

It may be noted that $\left(x_{1j}, x_{2j}\right)$, $j = 1,2,...,n$, are paired observations. The factored likelihood function of $\underline{\theta} = \left(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \sigma_{12}\right)$ can be written as

$$L\left(\underline{\theta} \mid X_{Obs}\right) = \prod_{j=1}^{n} f_{X_1}\left(x_{1j} \middle| \mu_1, \sigma_1^2\right) \prod_{j=1}^{r} f_{X_2 \mid X_1}\left(x_{2j} \middle| \beta_0 + \beta_1 x_{1j}, \sigma_{2|1}^2\right)$$

for $\beta_0 = \mu_2 - \beta_1\mu_1$, $\beta_1 = \rho\dfrac{\sigma_2}{\sigma_1}$ and $\sigma_{2|1}^2 = \left(1-\rho^2\right)\sigma_2^2$.

The factored maximum likelihood estimators of μ_1 , σ_1^2 , β_0 , β_1 and $\sigma_{2|1}^2$ can be obtained by solving $\dfrac{\partial}{\partial\underline{\theta}} L\left(\underline{\theta} \middle| X_{Obs}\right) = 0$. Then, Anderson's estimators are given by

$$\hat{\mu}_{1_{And}} = \bar{x}_1 = \frac{1}{n}\sum_{j=1}^{n} x_{1j}, \quad \hat{\sigma}_{1_{And}}^2 = \frac{1}{n}\sum_{j=1}^{n}\left(x_{1j} - \bar{x}_1\right)^2 ,$$

$$\hat{\beta}_1 = \frac{\sum_{j=1}^{r}\left(x_{1j} - \bar{x}_1'\right)\left(x_{2j} - \bar{x}_2'\right)}{\sum_{j=1}^{r}\left(x_{1j} - \bar{x}_1'\right)^2}, \ \hat{\beta}_0 = \bar{x}_2' - \hat{\beta}_1\bar{x}_1' \quad \text{and} \quad \hat{\sigma}_{2|1}^2 = s_2'^2 - \frac{s_{12}'^2}{s_1'^2}$$

for $\bar{x}_1' = \dfrac{1}{r}\sum_{j=1}^{r} x_{1j}$, $\bar{x}_2' = \dfrac{1}{r}\sum_{j=1}^{r} x_{2j}$, $s_1'^2 = \dfrac{1}{r}\sum_{j=1}^{r}\left(x_{1j} - \bar{x}_1'\right)^2$, $s_2'^2 = \dfrac{1}{r}\sum_{j=1}^{r}\left(x_{2j} - \bar{x}_2'\right)^2$

and $s_{12}' = \dfrac{1}{r}\sum_{j=1}^{r}\left(x_{1j} - \bar{x}_1'\right)\left(x_{2j} - \bar{x}_2'\right)$.

Additionally, the Anderson's estimators of μ_2 and σ_2^2 are given by

$\hat{\mu}_{2_{And}} = \bar{x}_2' - \hat{\beta}_1\left(\bar{x}_1' - \bar{x}_1\right)$ and $\hat{\sigma}_{2_{And}}^2 = \hat{\sigma}_{2|1}^2 + \hat{\beta}_1^2\hat{\sigma}_1^2$ respectively.

Lord [9] mentioned that $\hat{\mu}_{1_{And}}$ and $\hat{\mu}_{2_{And}}$ were unbiased estimators of μ_1 and μ_2 respectively. However, $\hat{\sigma}_{1_{And}}^2$ and $\hat{\sigma}_{2_{And}}^2$ were biased estimators of σ_1^2 and σ_2^2 respectively. In section 2.2, the modifications of $\hat{\sigma}_{1_{And}}^2$ and $\hat{\sigma}_{2_{And}}^2$ using Jackknife's method [13] will be explained.

2.2 The Jackknife's Method for J_Anderson estimator

In this section, we will derive the J_Anderson estimator. In addition, two biased estimators, $\hat{\sigma}^2_{1_{And}}$ and $\hat{\sigma}^2_{2_{And}}$, will be modified by Jackknife's method. Moreover, these are denoted by $\hat{\sigma}^2_{1_{J_And}}$ and $\hat{\sigma}^2_{2_{J_And}}$ respectively.

2.2.1 The J_Anderson Estimator of σ^2_1

The estimator $\hat{\sigma}^2_{1_{And}}$ is modified by Jackknife's method as follows:

1) Suppose we have a sample $X = (x_{11}, x_{12}, \ldots, x_{1n}, x_{21}, x_{22}, \ldots, x_{2r})$ and an estimator $\delta(X)$ where $\delta(X) = \hat{\sigma}^2_{1_{And}} = \frac{1}{n}\sum_{j=1}^{n}(x_{1j} - \bar{x}_1)^2 = \frac{1}{n}\sum_{j=1}^{n} x^2_{1j} - \bar{x}^2_1$

2) The Jackknife focuses on the samples that leave out one observation at a time: for i = 1, 2, …, n,

$X_{(-i)} = (x_{11}, x_{12}, \ldots, x_{1(i-1)}, x_{1(i+1)}, \ldots, x_{1n}, x_{21}, x_{22}, \ldots, x_{2(i-1)}, x_{2(i+1)}, \ldots, x_{2r})$

are called Jackknife samples.

3) Let $\delta(X_{(-i)})$ be the i^{th} Jackknife replication of $\delta(X)$

where $\delta(X_{(-i)}) = \hat{\sigma}^2_{1_{And(-i)}} = \frac{1}{n-1}\sum_{j\neq i}^{n} x^2_{1j} - \left(\frac{1}{n-1}\sum_{j\neq i}^{n} x_{1j}\right)^2$

$$= \frac{1}{n-1}\left(\sum_{j=1}^{n} x^2_{1j} - x^2_{1i}\right) - \left(\frac{1}{n-1}\left(\sum_{j=1}^{n} x_{1j} - x_{1i}\right)\right)^2$$

$$= \frac{1}{n-1}\left(\sum_{j=1}^{n} x^2_{1j} - x^2_{1i}\right) - \frac{1}{(n-1)^2}\left(n^2\bar{x}^2_1 - 2n\bar{x}_1 x_{1i} + x^2_{1i}\right) \text{ for i=1, 2, …, n}$$

4) Calculate the pseudo values in the form of J_i

where $J_i = n\delta(X) - (n-1)\delta\left(X_{(-i)}\right)$

$$= n\hat{\sigma}^2_{1And} - (n-1)\hat{\sigma}^2_{1And(-i)}$$

$$= n\hat{\sigma}^2_{1And} - (n-1)\left[\frac{1}{n-1}\left(\sum_{j=1}^{n} x^2_{1j} - x^2_{1i}\right) - \frac{1}{(n-1)^2}\left(n^2\bar{x}^2_1 - 2n\bar{x}_1x_{1i} + x^2_{1i}\right)\right]$$

$$= n\hat{\sigma}^2_{1And} - \left(\sum_{j=1}^{n} x^2_{1j} - x^2_{1i}\right) + \frac{1}{(n-1)}\left(n^2\bar{x}^2_1 - 2n\bar{x}_1x_{1i} + x^2_{1i}\right)$$

5) The J_Anderson estimator of σ^2_1 is given by $\delta_J(X)$

where $\delta_J(X) = \hat{\sigma}^2_{1J_And} = \frac{1}{n}\sum_{i=1}^{n} J_i$

$$= \frac{1}{n}\sum_{i=1}^{n}\left[n\hat{\sigma}^2_{1And} - \left(\sum_{j=1}^{n} x^2_{1j} - x^2_{1i}\right) + \frac{1}{(n-1)}\left(n^2\bar{x}^2_1 - 2\bar{x}_1x_{1i} + x^2_{1i}\right)\right]$$

$$= n\hat{\sigma}^2_{1And} - \left(\sum_{j=1}^{n} x^2_{1j} - \frac{1}{n}\sum_{i=1}^{n} x^2_{1i}\right) + \left(\left(\frac{n^2-2n}{n-1}\right)\bar{x}^2_1 + \frac{1}{n(n-1)}\sum_{i=1}^{n} x^2_{1i}\right)$$

$$= n\left(\frac{1}{n}\sum_{j=1}^{n} x^2_{1j} - \bar{x}^2_1\right) - \left(\sum_{j=1}^{n} x^2_{1j} - \frac{1}{n}\sum_{i=1}^{n} x^2_{1i}\right) + \left(\left(\frac{n^2-2n}{n-1}\right)\bar{x}^2_1 + \frac{1}{n(n-1)}\sum_{i=1}^{n} x^2_{1i}\right)$$

$$= \left(\frac{n^2-2n}{n-1} - n\right)\bar{x}^2_1 + \left(\frac{1}{n} + \frac{1}{n(n-1)}\right)\sum_{i=1}^{n} x^2_{1i}$$

$$= \frac{1}{n-1}\sum_{i=1}^{n}(x_{1i} - \bar{x}_1)^2$$

Further, we know that $E\left[\hat{\sigma}^2_{1J_And}\right] = E\left[\frac{1}{n-1}\sum_{i=1}^{n}(x_{1i} - \bar{x}_1)^2\right] = \sigma^2_1$. Hence, $\hat{\sigma}^2_{1J_And} = \frac{1}{n-1}\sum_{i=1}^{n}(x_{1i} - \bar{x}_1)^2$ is unbiased estimator of σ^2_1. In this case, the modification in section 2.2.1 indicates that the bias of $\hat{\sigma}^2_{1And}$ reduces to zero by using Jackknife's method.

2.2.2 The J_Anderson Estimator of σ_2^2

The estimator $\hat{\sigma}^2_{2_{And}}$ is modified by Jackknife's method as follows:

1) Suppose we have a sample $X = \left(x_{11}, x_{12}, \ldots, x_{1n}, x_{21}, x_{22}, \ldots, x_{2r}\right)$ and an estimator

$$\delta(X) = \hat{\sigma}^2_{2_{And}} = \hat{\sigma}^2_{2|1} + \hat{\beta}_1^2\hat{\sigma}^2_{1_{And}} \tag{1}$$

where $\hat{\beta}_1 = \dfrac{\sum_{j=1}^{r}\left(x_{1j}-\bar{x}'_1\right)\left(x_{2j}-\bar{x}'_2\right)}{\sum_{j=1}^{r}\left(x_{1j}-\bar{x}'_1\right)^2}$, $\hat{\sigma}^2_{1_{And}} = \dfrac{1}{n}\sum_{j=1}^{n}\left(x_{1j}-\bar{x}_1\right)^2$,

$\hat{\sigma}^2_{2|1} = s'^2_2 - \dfrac{s'^2_{12}}{s'^2_1}$, $s'_{12} = \dfrac{1}{r}\sum_{j=1}^{r}\left(x_{1j}-\bar{x}'_1\right)\left(x_{2j}-\bar{x}'_2\right)$,

$s'^2_1 = \dfrac{1}{r}\sum_{j=1}^{r}\left(x_{1j}-\bar{x}'_1\right)^2$, $s'^2_2 = \dfrac{1}{r}\sum_{j=1}^{r}\left(x_{2j}-\bar{x}'_2\right)^2$,

$\bar{x}_1 = \dfrac{1}{n}\sum_{j=1}^{n}x_{1j}$, $\bar{x}'_1 = \dfrac{1}{r}\sum_{j=1}^{r}x_{1j}$ and $\bar{x}'_2 = \dfrac{1}{r}\sum_{j=1}^{r}x_{2j}$.

2) The i^{th} Jackknife sample consists of the data set with the i^{th} observation removed.

$X_{(-i)} = \left(x_{11}, x_{12}, \ldots, x_{1(i-1)}, x_{1(i+1)}, \ldots, x_{1n}, x_{21}, x_{22}, \ldots, x_{2(i-1)}, x_{2(i+1)}, \ldots, x_{2r}\right)$ for i = 1, 2, …, n.

3) Let $\delta\left(X_{(-i)}\right)$ be the i^{th} Jackknife replication of $\delta(X)$ for i = 1, 2, …, r, r+1, r+2,…, n

$$\delta\left(X_{(-i)}\right) = \hat{\sigma}^2_{2_{And(-i)}} = \hat{\sigma}^2_{2|1(-i)} + \hat{\beta}^2_{1(-i)}\hat{\sigma}^2_{1_{And(-i)}} \tag{2}$$

where $\hat{\beta}_{1(-i)} = \begin{cases} \dfrac{\sum_{j\neq i}^{r}\left(x_{1j}-\bar{x}'_{1(-i)}\right)\left(x_{2j}-\bar{x}'_{2(-i)}\right)}{\sum_{j\neq i}^{r}\left(x_{1j}-\bar{x}'_{1(-i)}\right)^2} & \text{for } i = 1, 2, \ldots, r \\ \hat{\beta}_1 & \text{for } i = r+1, r+2, \ldots, n \end{cases}$

$$\hat{\sigma}^2_{1_{And(-i)}} = \frac{1}{n-1}\sum_{j\neq i}^{n}\left(x_{1j}-\bar{x}_{1(-i)}\right)^2$$

$$\hat{\sigma}^2_{2|1(-i)} = \begin{cases} s'^2_{2(-i)} - \dfrac{s'^2_{12(-i)}}{s'^2_{1(-i)}} & \text{for } i = 1, 2, \ldots, r \\ \hat{\sigma}^2_{2|1} & \text{for } i = r+1, r+2, \ldots, n \end{cases}$$

$$s'_{12(-i)} = \begin{cases} \dfrac{1}{r-1}\sum_{j\neq i}^{r}\left(x_{1j} - \bar{x}'_{1(-i)}\right)\left(x_{2j} - \bar{x}'_{2(-i)}\right) & \text{for } i = 1, 2, \ldots, r \\ s'_{12} & \text{for } i = r+1, r+2, \ldots, n \end{cases}$$

$$s'^2_{1(-i)} = \begin{cases} \dfrac{1}{r-1}\sum_{j\neq i}^{r}\left(x_{1j} - \bar{x}'_{1(-i)}\right)^2 & \text{for } i = 1, 2, \ldots, r \\ s'^2_1 & \text{for } i = r+1, r+2, \ldots, n \end{cases}$$

$$s'^2_{2(-i)} = \begin{cases} \dfrac{1}{r-1}\sum_{j\neq i}^{r}\left(x_{2j} - \bar{x}'_{2(-i)}\right)^2 & \text{for } i = 1, 2, \ldots, r \\ s'^2_2 & \text{for } i = r+1, r+2, \ldots, n \end{cases}$$

$$\bar{x}_{1(-i)} = \frac{1}{n-1}\sum_{j\neq i}^{n} x_{1j}$$

$$\bar{x}'_{1(-i)} = \begin{cases} \dfrac{1}{r-1}\sum_{j\neq i}^{r} x_{1j} & \text{for } i = 1, 2, \ldots, r \\ \bar{x}'_1 & \text{for } i = r+1, r+2, \ldots, n \end{cases}$$

$$\bar{x}'_{2(-i)} = \begin{cases} \dfrac{1}{r-1}\sum_{j\neq i}^{r} x_{2j} & \text{for } i = 1, 2, \ldots, r \\ \bar{x}'_2 & \text{for } i = r+1, r+2, \ldots, n \end{cases}$$

4) Calculate of the pseudo values in the form of J_i

where $J_i = n\delta(X) - (n-1)\,\delta\left(X_{(-i)}\right) = n\hat{\sigma}^2_{2And} - (n-1)\,\hat{\sigma}^2_{2And(-i)}$

5) The J_Anderson estimator of σ_2^2 is given by $\delta_J(X)$

where $\delta_J(X) = \hat{\sigma}^2_{2J_And} = \frac{1}{n}\sum_{i=1}^{n} J_i$

$$= \frac{1}{n}\sum_{i=1}^{n}\left(n\hat{\sigma}^2_{2And} - (n-1)\hat{\sigma}^2_{2And(-i)}\right)$$

$$= \frac{1}{n}\sum_{i=1}^{n} n\hat{\sigma}^2_{2And} - \frac{(n-1)}{n}\sum_{i=1}^{n}\hat{\sigma}^2_{2And(-i)}$$

$$= n\hat{\sigma}^2_{2And} - \frac{(n-1)}{n}\sum_{i=1}^{n}\hat{\sigma}^2_{2And(-i)}$$

In conclusion, the estimators of μ_1, μ_2, σ_1^2 and σ_2^2 from bivariate normal distribution with incomplete data are $\hat{\mu}_{1And} = \frac{1}{n}\sum_{j=1}^{n} x_{1j}$, $\hat{\mu}_{2And} = \bar{x}'_2 - \hat{\beta}_1(\bar{x}'_1 - \bar{x}_1)$, $\hat{\sigma}^2_{1J_And} = \frac{1}{n-1}\sum_{i=1}^{n}(x_{1i} - \bar{x}_1)^2$ and $\hat{\sigma}^2_{2J_And} = n\hat{\sigma}^2_{2And} - \frac{(n-1)}{n}\sum_{i=1}^{n}\hat{\sigma}^2_{2And(-i)}$ where $\hat{\sigma}^2_{2And}$ and $\hat{\sigma}^2_{2And(-i)}$ are given by the form of equation (1) and (2) respectively.

3. Simulation Study

In order to empirically evaluate the validity and reliability of the proposed estimator, a simulation study was conducted by SAS 9.1 programming. In the study, populations of (X_1, X_2) at a size of N = 100,000 were generated in the form of a bivariate normal distribution with $\mu_1 = 2$, $\mu_2 = 3$, $\sigma_1^2 = 4$ and $\sigma_2^2 = 9$. Correlations of (X_1, X_2) at 0, 0.1, 0.2, …, 1 with sample sizes of n = 20, 30 and 60 were made using a simple random sampling with replacement method with 2,000 times repetitions and missing data were set at 0, 10, 20 and 30 percent of the total cases, thus creating 132 situations for the simulation study. After that, absolute relative bias (ARB) and mean square error (MSE) comparisons of the J_Anderson estimator and the Anderson's estimator were empirically performed for $\hat{\sigma}_{1And}^2$ versus $\hat{\sigma}_{1J_And}^2$ and $\hat{\sigma}_{2And}^2$ versus $\hat{\sigma}_{2J_And}^2$. The results of the study are shown in the following.

The measure of biasedness for the estimator was computed in terms of the absolute relative biases of $E\left(\hat{\sigma}_{1And}^2\right)$, $E\left(\hat{\sigma}_{1J_And}^2\right)$, $E\left(\hat{\sigma}_{2And}^2\right)$ and $E\left(\hat{\sigma}_{2J_And}^2\right)$. These were calculated using the method of Horton et al. [14].

$$ARB\left[E\left(\hat{\sigma}_{1And}^2\right)\right] = \frac{\left|E\left(\hat{\sigma}_{1And}^2\right) - \sigma_1^2\right|}{\sigma_1^2} \times 100$$

$$ARB\left[E\left(\hat{\sigma}_{1J_And}^2\right)\right] = \frac{\left|E\left(\hat{\sigma}_{1J_And}^2\right) - \sigma_1^2\right|}{\sigma_1^2} \times 100$$

$$ARB\left[E\left(\hat{\sigma}_{2And}^2\right)\right] = \frac{\left|E\left(\hat{\sigma}_{2And}^2\right) - \sigma_2^2\right|}{\sigma_2^2} \times 100 \quad \text{and}$$

$$ARB\left[E\left(\hat{\sigma}_{2J_And}^2\right)\right] = \frac{\left|E\left(\hat{\sigma}_{2J_And}^2\right) - \sigma_2^2\right|}{\sigma_2^2} \times 100.$$

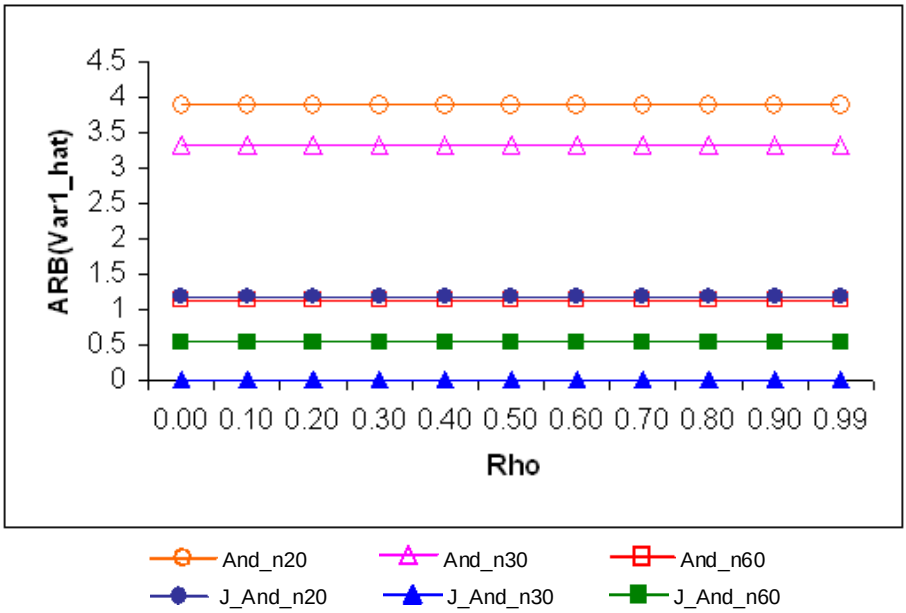


Figure 2. The comparison of ARB of $E\left(\hat{\sigma}^2_{1And}\right)$ and $E\left(\hat{\sigma}^2_{1J_And}\right)$ when n=20, 30 and 60 for non-missing data.

For this simulation study, the ARB of $E\left(\hat{\sigma}^2_{1J_And}\right)$ is smaller than that of $E\left(\hat{\sigma}^2_{1And}\right)$ whatever the sample size (see Table 1 and Figure 2). In addition, the ARB of $E\left(\hat{\sigma}^2_{1And}\right)$ seems to reduce to zero when large sample size. However, the ARB of $E\left(\hat{\sigma}^2_{1J_And}\right)$ seems to reduce to zero whatever the sample size. It is found that no difference for the ARB of $E\left(\hat{\sigma}^2_{1And}\right)$ when the increasing of ρ for all sample sizes. Furthermore, the conclusion is that the ARB of $E\left(\hat{\sigma}^2_{1J_And}\right)$ is similar to that of $E\left(\hat{\sigma}^2_{1And}\right)$.

Table 1. The absolute relative biases (ARB) of $E\left(\hat{\sigma}^2_{1And}\right)$ and $E\left(\hat{\sigma}^2_{1J_And}\right)$ for non-missing data

n	Method	ρ										
		0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
20	Anderson	3.88	3.88	3.88	3.88	3.88	3.88	3.88	3.88	3.88	3.88	3.88
	J_Anderson	1.18	1.18	1.18	1.18	1.18	1.18	1.18	1.18	1.18	1.18	1.18
30	Anderson	3.32	3.32	3.32	3.32	3.32	3.32	3.32	3.32	3.32	3.32	3.32
	J_Anderson	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
60	Anderson	1.13	1.13	1.13	1.13	1.13	1.13	1.13	1.13	1.13	1.13	1.13
	J_Anderson	0.54	0.54	0.54	0.54	0.54	0.54	0.54	0.54	0.54	0.54	0.54

The results of the simulation study indicate that the mean square errors of $\hat{\sigma}^2_{1And}$ and $\hat{\sigma}^2_{1J_And}$ seem to decrease whenever the sample size increases whatever the correlation coefficient between X_1 and X_2 (see Figure 3). In addition, the mean square error of $\hat{\sigma}^2_{1And}$ seems to have no difference from that of $\hat{\sigma}^2_{1J_And}$ in each situation.

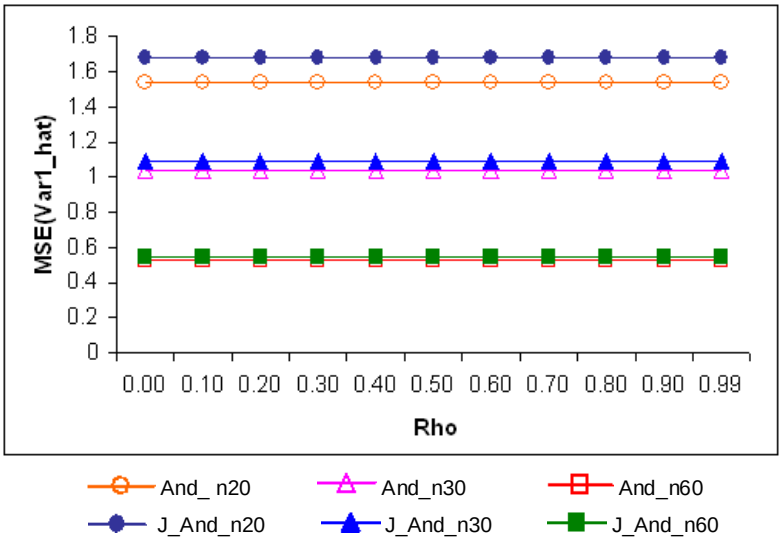


Figure 3. The MSE of $\hat{\sigma}^2_{1And}$ and $\hat{\sigma}^2_{1J_And}$ when n=20, 30 and 60 for non-missing data.

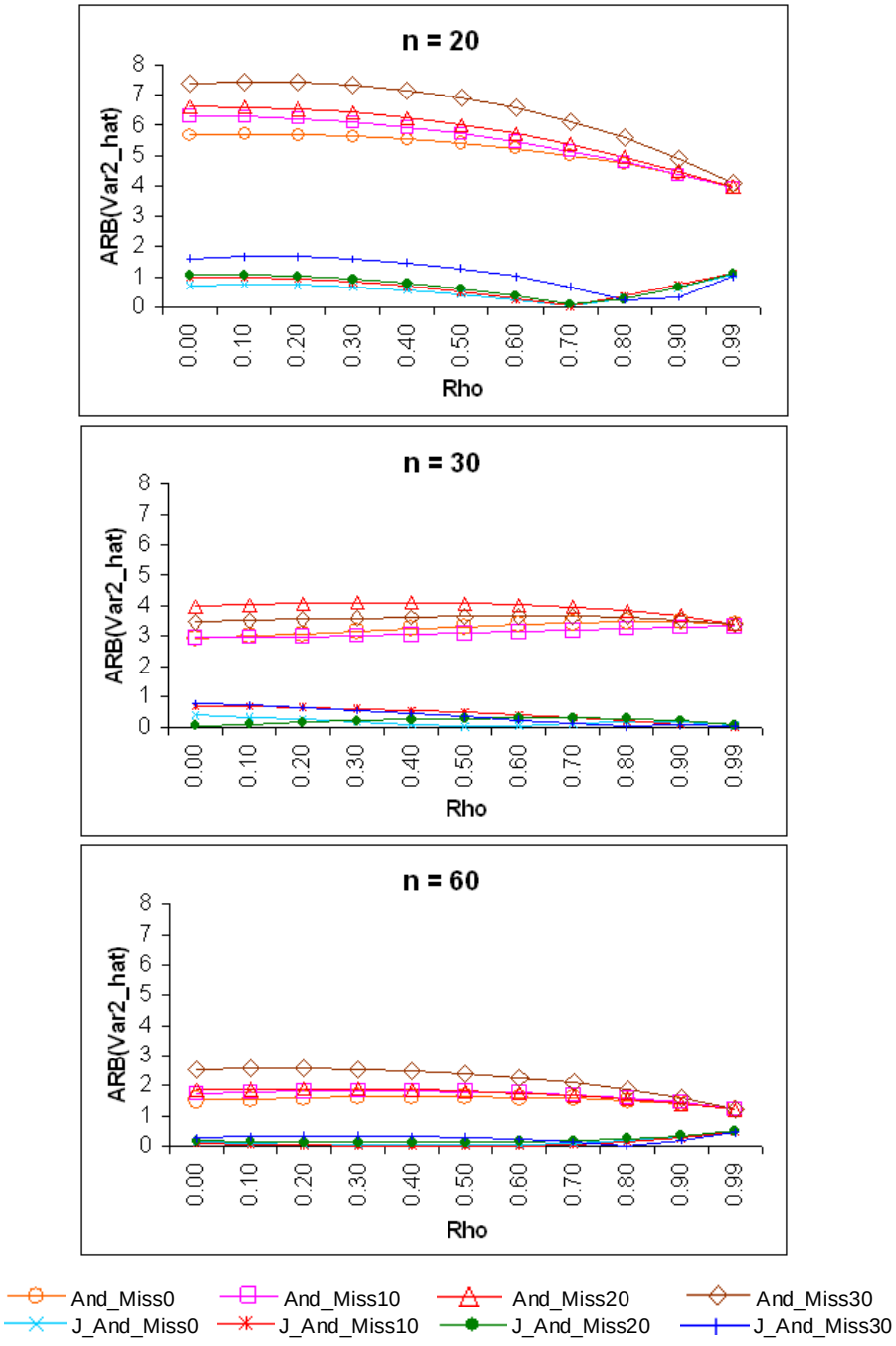


Figure 4. The comparison of ARB of $E\left(\hat{\sigma}^2_{2\text{And}}\right)$ and $E\left(\hat{\sigma}^2_{2\text{J_And}}\right)$ when n=20, 30 and 60.

Likewise, the ARB of $E\left(\hat{\sigma}^2_{2_{J_And}}\right)$ is smaller than that of $E\left(\hat{\sigma}^2_{2_{And}}\right)$ whatever the sample size and percentage of missing data (see Table 2 and Figure 4). In addition, it was found that the ARB of $E\left(\hat{\sigma}^2_{2_{And}}\right)$ seems to decrease to zero for all percentages of missing data, and values of ρ when large sample size. However, the ARB of $E\left(\hat{\sigma}^2_{2_{J_And}}\right)$ seems to reduce to zero whatever the sample size (see Table 2 and Figure 5).

Table 2. The absolute relative bias (ARB) of $E\left(\hat{\sigma}^2_{2_{And}}\right)$ and $E\left(\hat{\sigma}^2_{2_{J_And}}\right)$

n	ρ	Non-missing data		Missing 10%		Missing 20%		Missing 30%	
		And[1]	J_And[2]	And	J_And	And	J_And	And	J_And
20	0	5.67	0.71	6.31	1.00	6.61	1.05	7.38	1.61
	0.1	5.70	0.74	6.29	0.99	6.61	1.06	7.44	1.66
	0.2	5.69	0.72	6.22	0.93	6.54	1.01	7.42	1.66
	0.3	5.64	0.67	6.10	0.83	6.42	0.92	7.33	1.60
	0.4	5.55	0.57	5.94	0.69	6.25	0.79	7.16	1.47
	0.5	5.41	0.43	5.72	0.50	6.02	0.60	6.91	1.27
	0.6	5.24	0.25	5.46	0.26	5.72	0.37	6.56	1.01
	0.7	5.01	0.01	5.14	0.02	5.37	0.08	6.13	0.66
	0.8	4.73	0.28	4.78	0.35	4.96	0.26	5.58	0.23
	0.9	4.39	0.64	4.36	0.73	4.48	0.66	4.91	0.32
	1	3.97	1.08	3.93	1.13	3.96	1.10	4.08	1.01
30	0	2.95	0.40	2.96	0.69	3.99	0.05	3.47	0.78
	0.1	3.02	0.33	2.97	0.68	4.04	0.10	3.52	0.72
	0.2	3.09	0.25	2.99	0.65	4.08	0.16	3.56	0.64
	0.3	3.16	0.18	3.02	0.60	4.10	0.21	3.60	0.55
	0.4	3.24	0.10	3.05	0.55	4.10	0.25	3.63	0.45
	0.5	3.31	0.02	3.09	0.48	4.08	0.28	3.65	0.35
	0.6	3.38	0.05	3.14	0.40	4.03	0.30	3.66	0.23
	0.7	3.44	0.11	3.19	0.31	3.96	0.30	3.65	0.12
	0.8	3.49	0.16	3.24	0.21	3.85	0.28	3.62	0.02
	0.9	3.49	0.17	3.29	0.10	3.68	0.22	3.55	0.06
	1	3.40	0.07	3.33	0.01	3.42	0.07	3.39	0.05
60	0	1.52	0.15	1.73	0.11	1.85	0.18	2.52	0.26
	0.1	1.57	0.10	1.77	0.06	1.88	0.15	2.56	0.31
	0.2	1.60	0.07	1.80	0.02	1.89	0.13	2.57	0.34
	0.3	1.63	0.04	1.82	0.00	1.88	0.12	2.54	0.34
	0.4	1.64	0.03	1.82	0.01	1.86	0.12	2.48	0.33
	0.5	1.63	0.03	1.80	0.01	1.81	0.13	2.39	0.29
	0.6	1.61	0.05	1.76	0.02	1.75	0.15	2.26	0.23
	0.7	1.57	0.10	1.69	0.06	1.66	0.19	2.09	0.13
	0.8	1.50	0.17	1.59	0.14	1.55	0.25	1.88	0.00
	0.9	1.40	0.28	1.44	0.26	1.40	0.34	1.59	0.18
	1	1.21	0.47	1.21	0.46	1.19	0.49	1.23	0.45

[1] Anderson's method
[2] J_Anderson's method

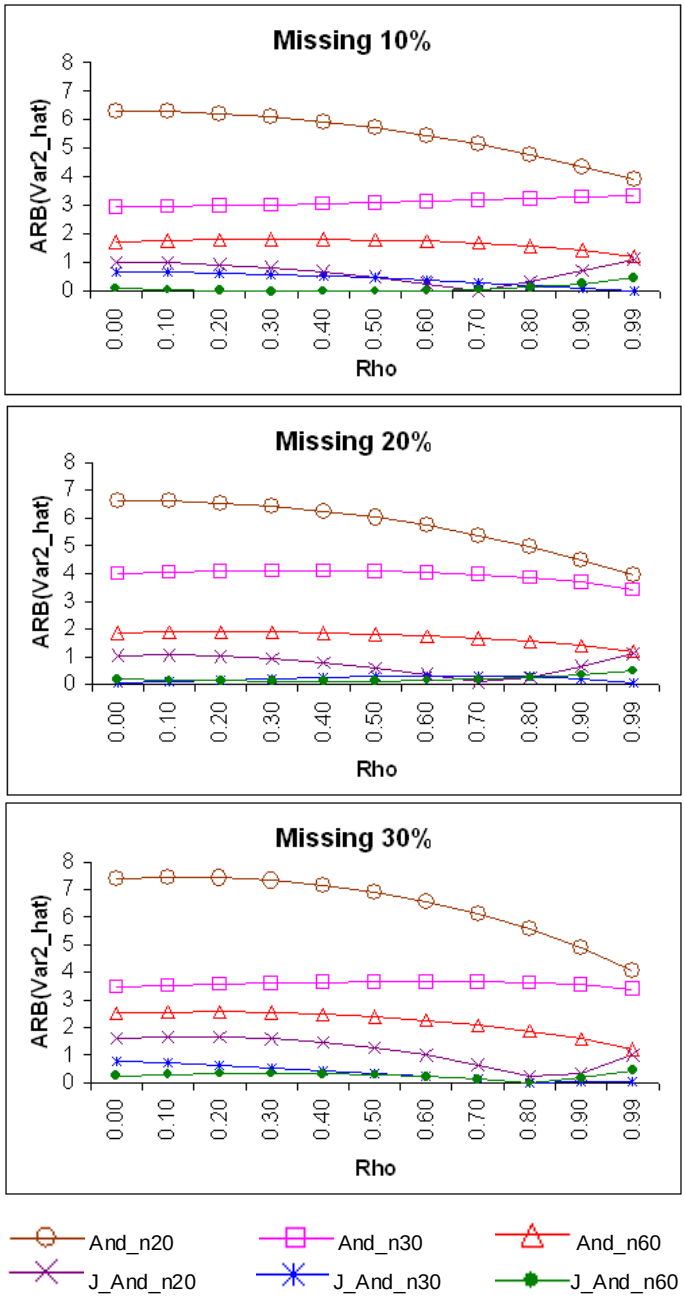


Figure 5. The comparison of ARB of $E\left(\hat{\sigma}^2_{2And}\right)$ and $E\left(\hat{\sigma}^2_{2J_And}\right)$ when percentages of missing data are 10, 20 and 30.

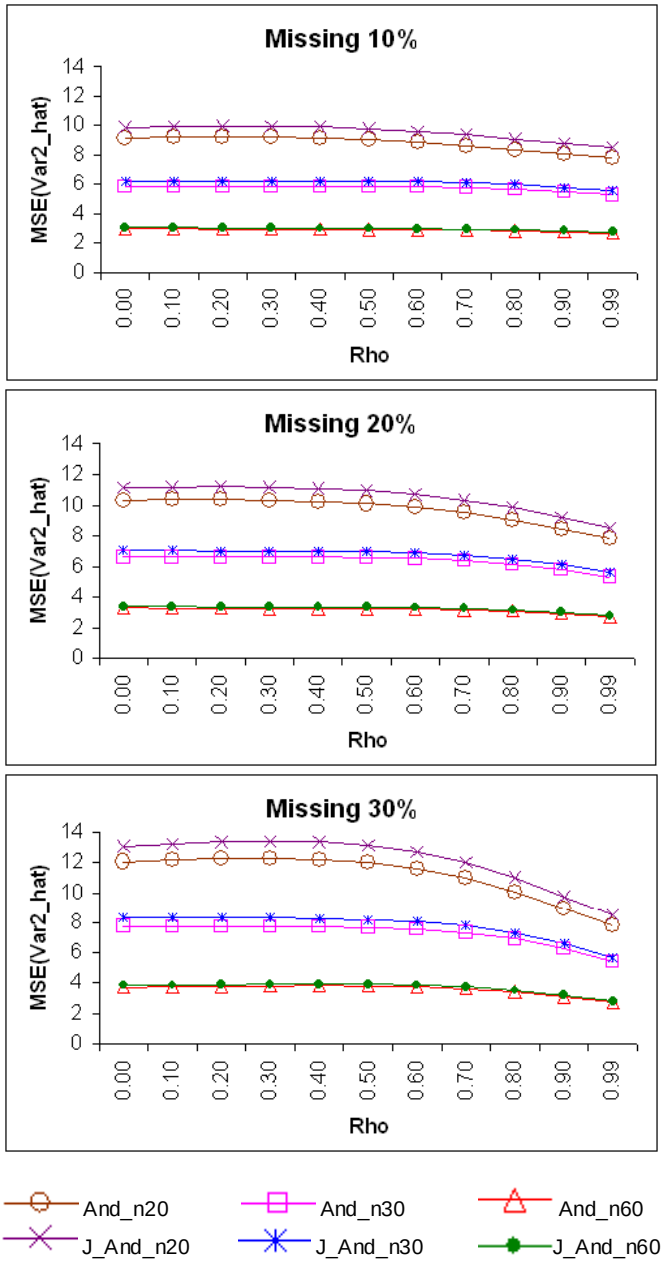


Figure 6. The MSE of $\hat{\sigma}^2_{2\text{And}}$ and $\hat{\sigma}^2_{2\text{J_And}}$ when percentages of missing data are 10, 20 and 30.

Figure 6 illustrates that the mean square error of $\hat{\sigma}^2_{2_{And}}$ seems to have no difference from that of $\hat{\sigma}^2_{2_{J_And}}$ in each situation. In particular, the mean square error of $\hat{\sigma}^2_{2_{And}}$ and that of $\hat{\sigma}^2_{2_{J_And}}$ seem to decrease when the sample size was large whatever percentage of missing data.

4. Discussion

The simulation results presented that $\hat{\sigma}^2_{1_{And}}$ and $\hat{\sigma}^2_{2_{And}}$ seemed to be biased estimators. Hence, biases of $\hat{\sigma}^2_{1_{And}}$ and $\hat{\sigma}^2_{2_{And}}$ can be reduced by Jackknife's method as reported in research studies of (Efron and Tibshirani [10]; Smith and Pontius [12]). These findings can be applied in researches in education, psychology, medicine and other fields. Jackknife's method can be applied in the elimination of biases of the parameter estimation (Lehmann and Casella [15]) for incomplete samples from bivariate normal populations.

5. Summary and Conclusions

This paper proposes the estimators for μ_1, μ_2, σ_1^2 and σ_2^2 of the monotone incomplete bivariate sample from a bivariate normal distribution. The proposed estimators were derived from Anderson's method [7] whose approach was to construct a factored likelihood function. Then, Anderson's estimators were given by $\hat{\mu}_{1_{And}}$, $\hat{\mu}_{2_{And}}$, $\hat{\sigma}^2_{1_{And}}$ and $\hat{\sigma}^2_{2_{And}}$. However, $\hat{\sigma}^2_{1_{And}}$ and $\hat{\sigma}^2_{2_{And}}$ were biased estimators of σ_1^2 and σ_2^2 respectively . In this paper, a Jackknife's method of bias reduction was introduced to modify the estimators $\hat{\sigma}^2_{1_{And}}$ and $\hat{\sigma}^2_{2_{And}}$. Moreover, these were denoted by $\hat{\sigma}^2_{1_{J_And}}$ and $\hat{\sigma}^2_{2_{J_And}}$ respectively. The simulation study found that the absolute relative bias of $E\left(\hat{\sigma}^2_{1_{J_And}}\right)$ was smaller than that of $E\left(\hat{\sigma}^2_{1_{And}}\right)$. Likewise, the absolute relative bias of $E\left(\hat{\sigma}^2_{2_{J_And}}\right)$ was smaller than that of $E\left(\hat{\sigma}^2_{2_{And}}\right)$ whatever sample size and percentage of missing data. Furthermore, the mean square errors of $\hat{\sigma}^2_{1_{And}}$ and $\hat{\sigma}^2_{1_{J_And}}$ seemed to decrease whenever the sample size increases whatever the percentage of missing

data and correlation coefficient between X_1 and X_2. Moreover, the mean square error of $\hat{\sigma}^2_{2_{And}}$ seemed to have no difference from that of $\hat{\sigma}^2_{2_{J_And}}$ in each situation of this simulation study. In particular, the mean square error of $\hat{\sigma}^2_{2_{And}}$ and that of $\hat{\sigma}^2_{2_{J_And}}$ seemed to decrease when the sample size was large whatever percentage of missing data.

6. Acknowledgements

The author would like to thank the Department of Statistics, Faculty of Science, Kasetsart University for financial and software program support during the research.

References

[1] Little, R.J.A. and Rubin, D.B., *Statistical Analysis with Missing Data,* New Jersey: Wiley, 2002.

[2] Acock, A.C., Working with Missing Values, *Journal of Marriage and Family*, 2005; 67: 1012-1028.

[3] Fitzmaurice, G., Missing Data: Implications for Analysis, *Nutrition*, 2008; 24: 200-202.

[4] Gorelick, M.H., Bias Arising from Missing Data in Predictive Models, *Journal of Clinical Epidemiology,* 2006; 59: 1115-1123.

[5] Roth, P.L., Campion, J.E. and Jones, S.D. The Impact of Four Missing Data Techniques on Validity Estimates in Human Resource Management, *Journal of Business and Psychology,* 1996; 11: 101-112.

[6] Rotnitzky, A. and Wypij, D., A Note on the Biased of Estimators with Missing Data, *Biometrics,* 1994; 50: 1163-1170.

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

[8] Edgett, G.L., Multiple Regression with Missing Observations among the Independent Variables, *Journal of the American Statistical Association*, 1956; 51: 122-131.

[9] Lord, F.M., Estimation of Parameter from Incomplete Data, *Journal of the American Statistical Association*, 1955; 50: 870-876.

[10] Efron, B. and Tibshirani, R.J., *An Introduction to the Bootstrap*, USA: Chapman & Hall/CRC, 1993.

[11] Kim, J.M. and Anderson, J.E., Jackknife Variance Estimation for Two Samples after Imputation under Two-Phase Sampling, *ASA Section on Survey Research Methods*.

[12] Smith, C.D. and Pontius, J.S., Jackknife Estimator of Species Richness with S-PLUS, *Journal of Statistical Software*, 2006; 15: 1-12.

[13] Quenouille, M.H., Notes on Bias in Estimation, *Biometrika*, 1956; 43: 353-360.

[14] Horton, N.J., Lipsitz, S.R. and Parzen, M., A Potential for Bias when Rounding in Multiple Imputation, *The American Statistician*, 2003; 57: 229-232.

[15] Lehmann, E.L. and Casella,G., *Theory of Point Estimation*, New York: Springer-Verlag, 1998.