SUT Journal of Mathematics
Vol. 42, No. 1 (2006), 133–144

Simultaneous confidence intervals for all contrasts of the means in repeated measures with missing observations

Kazuyuki Koizumi and Takashi Seo

(Received January 14, 2006; Revised April 14, 2006)

Dedicated to Professor Minoru Siotani on his 80th birthday

Abstract. Simultaneous confidence intervals for all contrasts of the mean components in repeated measures with the intraclass correlation model are considered when the observations are missing at random. An exact test statistic for the equality of means and Scheffé, Bonferroni and Tukey types of simultaneous confidence intervals are given under the general case of missing observations. Finally, numerical examples by simulation are given to illustrate the method.

AMS 2000 *Mathematics Subject Classification.* 62J15, 62H15.

Key words and phrases. Intraclass correlation model, missing data, testing equality of means, Monte Carlo simulation.

§1. Introduction

The problem of data set $\{x_{ij}\}$, $i = 1, 2, \ldots, p$, $j = 1, 2, \ldots, n$ with missing observations at random arises frequently in practice. In this paper, we consider an exact test statistic for the equality of means and the simultaneous confidence intervals for all contrasts in the mean components when the missing observations are of the non-monotone type in repeated measures with the intraclass correlation model.

When the missing observations are of the monotone type, Seo, Kikuchi and Koizumi[10] has given an exact test statistic for the equality of means and Scheffé, Bonferroni and Tukey types of simultaneous confidence intervals for all contrasts in the means by an extension of the transformation in Bhargava and Srivastava[2]. In particular, Scheffé and Bonferroni types of simultaneous confidence intervals by Seo-Kikuchi-Koizumi procedure have exactly confidence level at $1 - \alpha$. In this paper these results are extended to the general case of missing observations.

In the intraclass correlation model, simultaneous confidence intervals of Scheffé and Tukey types for all contrasts in the means of a multivariate normal population has been given by Miller[6] and Scheffé[9] when the intraclass correlation coefficient is known. For the case that the intraclass correlation coefficient is unknown, Bhargava and Srivastava[2] has given an extension of the above results. When the missing observations are of the monotone type in the intraclass correlation model, Seo and Srivastava[11] gave an exact test statistic for the equality of means and simultaneous confidence intervals (Scheffé and Bonferroni types) for all contrasts in the means, which are different from those by Seo-Kikuchi-Koizumi procedure. Further, when the missing observations are of the non-monotone type, Seo and Srivastava[11] gave asymptotic tests and simultaneous confidence intervals based on the maximum likelihood estimates which are obtained numerically by iterative method. For the iterative method, see, Srivastava and Carter[14], Srivastava[12]. By using the same idea of Srivastava[12], an extension of Seo-Kikuchi-Koizumi procedure for the general case of missing observations can be obtained. The organization of the paper is as follows. In Section 2, we provide an exact test procedure for testing the equality of means under the general case of missing observations. In Section 3, Scheffé, Bonferroni and Tukey types of simultaneous confidence intervals for all contrasts in the means are given. Finally, numerical examples by simulation are presented to illustrate the method in Section 4.

§2. An exact test for the equality of mean components

Starting from complete data set $\{x_{ij}\}$, $i = 1, 2, \dots, p$, $j = 1, 2, \dots, n$ without missing observations:

$$\begin{pmatrix} x_{11} & x_{12} & \cdots & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & \cdots & x_{2n} \\ \vdots & \vdots & \cdots & \cdots & \vdots \\ x_{p1} & x_{p2} & \cdots & \cdots & x_{pn} \end{pmatrix}, \tag{2.1}$$

we have the intraclass correlation model in the following form. Let $\boldsymbol{x}_j = (x_{1j}, x_{2j}, \dots, x_{pj})'$, $j = 1, 2, \dots, n$ and let $\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_n$ be independently distributed as $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, where $\boldsymbol{\mu} = (\mu_1, \mu_2, \dots, \mu_p)'$, $\boldsymbol{\Sigma} = \sigma^2[(1-\rho)\boldsymbol{I}_p + \rho\boldsymbol{1}_p\boldsymbol{1}_p']$, $\boldsymbol{I}_p$ is $p \times p$ identity matrix and $\boldsymbol{1}_p$ is a p-vector, $\boldsymbol{1}_p = (1, 1, \dots, 1)'$. When the covariance matrix $\boldsymbol{\Sigma}$ is of the above structure, it is called an intraclass correlation model. Then we consider an exact test for testing the hypothesis $\boldsymbol{H}_0 : \mu_1 = \mu_2 = \dots = \mu_p$ against the alternative $\boldsymbol{H}_1 \neq \boldsymbol{H}_0$ when the missing observations are of the non-monotone type in repeated measures with the intraclass correlation model. Let n_i and p_j be the total numbers of the observed data for i-th row and j-th column, respectively. The data set is called a monotone type of missing observations if n_i and p_j satisfy $n = n_1 \geq n_2 \geq \cdots \geq n_p$ and $p = p_1 \geq p_2 \geq \cdots \geq p_n$, otherwise it is called a non-monotone type(general case) of missing observations. For the case of the non-monotone type as well as monotone type of missing observations, we can obtain a subvector without missing part by a transformation of a sample vector with missing components. As an example, suppose we have the observations $\boldsymbol{x}_j = (x_{1j}, *, *, x_{4j}, x_{5j})'$ for the j-th column, where "$*$" means a missing component. Then, we can define as $\boldsymbol{y}_j (= (y_{1j}, y_{2j}, y_{3j})') = \boldsymbol{B}_j \boldsymbol{x}_j = (x_{1j}, x_{4j}, x_{5j})'$, where

$$\boldsymbol{B}_j = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

which is distributed as $N_3(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $\boldsymbol{\mu} = (\mu_1, \mu_4, \mu_5)'$ and $\boldsymbol{\Sigma} = \sigma^2[(1-\rho)\boldsymbol{I}_3 + \rho \mathbf{1}_3 \mathbf{1}_3']$. Therefore, in general, letting $\boldsymbol{y}_j = (y_{1j}, y_{2j}, \ldots, y_{p_j j})'$, then $\boldsymbol{y}_j$'s are independently distributed as $N_{p_j}(\boldsymbol{B}_j \boldsymbol{\mu}_j, \boldsymbol{\Sigma}_j)$, $j = 1, 2, \ldots, n$, where $\boldsymbol{B}_j$ is a $p_j \times p$ matrix, $\boldsymbol{\Sigma}_j = \sigma^2[(1-\rho)\boldsymbol{I}_{p_j} + \rho \mathbf{1}_{p_j} \mathbf{1}_{p_j}']$.

Next, as in Seo, Kikuchi and Koizumi[10], let $\boldsymbol{C}_j$ be a $p_j \times p_j$ matrix such that

$$\boldsymbol{C}_j = \boldsymbol{I}_{p_j} - \frac{\nu_j}{p_j} \mathbf{1}_j \mathbf{1}_j',$$

where $\nu_j = 1 \pm (1-\rho)^{\frac{1}{2}} \{1 + (p_j - 1)\rho\}^{-\frac{1}{2}}$. Then, by the transformation $\boldsymbol{w}_j = ((w_{1j}, w_{2j}, \ldots, w_{p_j j})') = \boldsymbol{C}_j \boldsymbol{y}_j$, we have

$$\boldsymbol{w}_j \sim N_{p_j}(\boldsymbol{C}_j \boldsymbol{B}_j \boldsymbol{\mu}_j, \gamma^2 \boldsymbol{I}_{p_j}),$$

where $\gamma^2 = \sigma^2(1-\rho)$. After this transformation, we can obtain $\boldsymbol{z}_j$ given by $\boldsymbol{z}_j = \boldsymbol{B}_j' \boldsymbol{w}_j$.

Without loss of generality, the observed original data set $\{x_{ij}\}$, $i = 1, 2, \ldots, p$, $j = 1, 2, \ldots, n$ and the transformed data set $\{z_{ij}\}$, $i = 1, 2, \ldots, p$, $j = 1, 2, \ldots, n$ can be grouped into s subsets of data with same missing pattern, respectively, where c-th group($c = 1, 2, \ldots, s \leq 2^p - 1$) consists of $n^{(c)}$ sample vectors such that $p^{(c)}$ observations are available in p components. We note that $p^{(c)}$ means the total number of components after excluding the missing part.

Let $y_{k\ell}^{(c)}$ and $w_{k\ell}^{(c)}$ be a (k, ℓ) component in the c-th group, respectively. Then we have the original sample means $\overline{y}_{k\cdot}^{(c)}$, $\overline{y}_{\cdot\ell}^{(c)}$ and $\overline{y}_{\cdot\cdot}^{(c)}$ are given by

$$\overline{y}_{k\cdot}^{(c)} = \frac{1}{n^{(c)}} \sum_{\ell=1}^{n^{(c)}} y_{k\ell}^{(c)}, \quad \overline{y}_{\cdot\ell}^{(c)} = \frac{1}{p^{(c)}} \sum_{k=1}^{p^{(c)}} y_{k\ell}^{(c)}, \quad \overline{y}_{\cdot\cdot}^{(c)} = \frac{1}{p^{(c)} n^{(c)}} \sum_{k=1}^{p^{(c)}} \sum_{\ell=1}^{n^{(c)}} y_{k\ell}^{(c)},$$

respectively. Similarly, the transformed sample means, $\overline{w}^{(c)}_{k\cdot}, \overline{w}^{(c)}_{\cdot\ell}$ and $\overline{w}^{(c)}_{\cdot\cdot}$ given by

$$\overline{w}^{(c)}_{k\cdot} = \frac{1}{n^{(c)}}\sum_{\ell=1}^{n^{(c)}} w^{(c)}_{k\ell}, \quad \overline{w}^{(c)}_{\cdot\ell} = \frac{1}{p^{(c)}}\sum_{k=1}^{p^{(c)}} w^{(c)}_{k\ell}, \quad \overline{w}^{(c)}_{\cdot\cdot} = \frac{1}{p^{(c)}n^{(c)}}\sum_{k=1}^{p^{(c)}}\sum_{\ell=1}^{n^{(c)}} w^{(c)}_{k\ell},$$

respectively. Hence, we have an unbiased estimator of γ^2 as

$$\begin{aligned} \widehat{\gamma}^{(c)^2} &= \frac{1}{f^{(c)}}\sum_{k=1}^{p^{(c)}}\sum_{\ell=1}^{n^{(c)}}\left(w^{(c)}_{k\ell} - \overline{w}^{(c)}_{k\cdot} - \overline{w}^{(c)}_{\cdot\ell} + \overline{w}^{(c)}_{\cdot\cdot}\right)^2 \\ &= \frac{1}{f^{(c)}}\sum_{k=1}^{p^{(c)}}\sum_{\ell=1}^{n^{(c)}}\left(y^{(c)}_{k\ell} - \overline{y}^{(c)}_{k\cdot} - \overline{y}^{(c)}_{\cdot\ell} + \overline{y}^{(c)}_{\cdot\cdot}\right)^2, \end{aligned}$$

where $f^{(c)} = (p^{(c)}-1)(n^{(c)}-1)$. Then under the hypothesis $\boldsymbol{H}_0$, $(f^{(c)}\widehat{\gamma}^{(c)^2})/\gamma^2$ has a χ^2 distribution with $f^{(c)}$ degrees of freedom, and hence,

$$\sum_{c=1}^{s}\frac{f^{(c)}\widehat{\gamma}^{(c)^2}}{\gamma^2} \tag{2.2}$$

has a χ^2 distribution with $f = \sum_{c=1}^{s} f^{(c)}$ degrees of freedom.

On the other hand, letting the values of missing observations be zero, then we have

$$\overline{x}_{i\cdot} = \frac{1}{n_i}\sum_{j=1}^{n} x_{ij}, \ \tilde{x}_{\cdot\cdot} = \frac{1}{p}\sum_{i=1}^{p}\overline{x}_{i\cdot}, \ \overline{z}_{i\cdot} = \frac{1}{n_i}\sum_{j=1}^{n} z_{ij}, \ \tilde{z}_{\cdot\cdot} = \frac{1}{p}\sum_{i=1}^{p}\overline{z}_{i\cdot}.$$

Since we note that $n_i(\overline{z}_{i\cdot} - \tilde{z}_{\cdot\cdot}) = n_i(\overline{x}_{i\cdot} - \tilde{x}_{\cdot\cdot})$,

$$\sum_{i=1}^{p}\left(\frac{\sqrt{n_i}(\overline{z}_{i\cdot} - \tilde{z}_{\cdot\cdot})}{\gamma}\right)^2 = \sum_{i=1}^{p}\left(\frac{\sqrt{n_i}(\overline{x}_{i\cdot} - \tilde{x}_{\cdot\cdot})}{\gamma}\right)^2,$$

which has a χ^2 distribution with $p-1$ degrees of freedom and is independent of (2.2). Thus, we can obtain the following theorem.

Theorem 1. *Suppose that a data set has the general missing observations at random in the intraclass correlation model. Then a test statistic for the null hypothesis* $\boldsymbol{H}_0$ *is given by*

$$F_0 = \frac{\displaystyle\sum_{i=1}^{p} n_i(\overline{x}_{i\cdot} - \tilde{x}_{\cdot\cdot})^2/(p-1)}{\displaystyle\sum_{c=1}^{s} f^{(c)}\widehat{\gamma}^{(c)^2}/f}, \tag{2.3}$$

where the distribution of F_0 *under the null hypothesis* $\boldsymbol{H}_0$ *is an* F *distribution with* $p-1$ *and* $f = \sum_{c=1}^{s}(p^{(c)}-1)(n^{(c)}-1)$ *degrees of freedom.*

Large value of this statistic (2.3) leads to the rejection of hypothesis $\boldsymbol{H}_0$. It may be noted that the value of F_0 is directly calculated from the original data set. Also, when the missing observations are of the monotone type, the statistic F_0 in (2.3) reduces to the test statistic by Seo, Kikuchi and Koizumi[10].

§3. Scheffé, Bonferroni and Tukey types of simultaneous confidence intervals for all contrasts

In this section, we consider Scheffé, Bonferroni and Tukey types of simultaneous confidence intervals for all contrasts in the means, that is, simultaneous confidence intervals of $\boldsymbol{a}'\boldsymbol{\mu}$ for non-null vector $\boldsymbol{a}$ such that $\boldsymbol{a}'\boldsymbol{1} = (a_1, a_2, \ldots, a_p)\boldsymbol{1} = 0$.

Since the test statistic F_0 in (2.3) has an F distribution with $p-1$ and f degrees of freedom, we have Scheffé type of simultaneous confidence intervals given by

$$\boldsymbol{a}'\boldsymbol{\mu} \in \left[\sum_{i=1}^{p} a_i \overline{x}_{i\cdot} \pm \sqrt{(p-1)F_{p-1,f,\alpha} \sum_{c=1}^{s} \frac{f^{(c)}\widehat{\gamma}^{(c)2}}{f} \sum_{i=1}^{p} \frac{a_i^2}{n_i}} \right], \tag{3.1}$$

where $F_{p-1,f,\alpha}$ is the upper $100\alpha\%$ of an F distribution with $p-1$ and f degrees of freedom.

We can also obtain simultaneous confidence intervals for h linear contrasts $\boldsymbol{a}_1'\boldsymbol{\mu}, \boldsymbol{a}_2'\boldsymbol{\mu}, \ldots, \boldsymbol{a}_h'\boldsymbol{\mu}$ by Bonferroni's inequality. Let $\boldsymbol{a}_j = (a_1^{(j)}, a_2^{(j)}, \ldots, a_p^{(j)})'$, $j = 1, 2, \ldots, h$ such that $\boldsymbol{a}_j'\boldsymbol{1} = 0$. Then we have Bonferroni type of simultaneous confidence intervals given by

$$\boldsymbol{a}_j'\boldsymbol{\mu} \in \left[\sum_{i=1}^{p} a_i^{(j)} \overline{x}_{i\cdot} \pm t_{f,\frac{\alpha}{2h}} \sqrt{\sum_{c=1}^{s} \frac{f^{(c)}\hat{\gamma}^{(c)2}}{f} \sum_{i=1}^{p} \frac{{a_i^{(j)}}^2}{n_i}} \right], j = 1, 2, \ldots, h, \tag{3.2}$$

where $t_{f,\alpha/(2h)}$ is the upper $100\alpha/(2h)\%$ of a t distribution with f degrees of freedom.

In the same as Seo, Kikuchi and Koizumi[10], we note that Bonferroni type of simultaneous confidence intervals should be used only if

$$(p-1)F_{p-1,f,\alpha} \geq t^2_{f,\frac{\alpha}{2h}},$$

otherwise Scheffé type of simultaneous confidence intervals should be used. It holds that $(p-1)F_{p-1,f,\alpha} < t^2_{f,\alpha/(2h)}$ if h is considerably bigger than $p-1$ (see, Miller[6]).

Further, Tukey type of simultaneous confidence intervals for all contrasts are given by

$$(3.3)\quad \boldsymbol{a}'\boldsymbol{\mu} \in \left[\sum_{i=1}^{p} a_i \overline{x}_{i\cdot} \pm q_{p,f,\alpha} \sum_{i=1}^{p} \frac{|a_i|}{2} \sqrt{\sum_{c=1}^{s} \frac{f^{(c)}\widehat{\gamma}^{(c)2}}{f} \sum_{i=1}^{p} \frac{1}{n_i|a_i|} \left(\sum_{i=1}^{p} \frac{1}{|a_i|} \right)^{-1}} \right],$$

where $q_{p,f,\alpha}$ is the upper $100\alpha\%$ of a Studentized range on p and f degrees of freedom. It may be noted that the these confidence intervals (3.1), (3.2) and (3.3) are directly calculated from the original data set. Also, when observed data is monotone type of missing data, these confidence intervals reduce to those results by Seo, Kikuchi and Koizumi[10].

§4. Numerical example

In this section, in order to investigate the accuracy of the procedure developed in this paper, we give numerical examples for some selected parameters by simulation study. We generate an artificial complete data set as (2.1) from the multivariate normal population in the intraclass correlation model. A data set with missing observations at random is composed by deleting some data from the above artificial complete data set.

The values of p, n, $\boldsymbol{\mu}$, σ^2, ρ and α were selected as follows:

$p = 4, 6$;
$n = 20, 40$;
$\boldsymbol{\mu} = (1, 1, 5, 5)', (1, 5, 10, 15)', (1, 1, 5, 5, 10, 10)'$;
$\sigma^2 = 1, 9$;
$\rho = 0.1, 0.5, 0.9$ and $\alpha = 0.05$.

Also letting m be a total number of available observations, we consider the following five cases for $p = 4$, $n = 20$:

(I) A complete data set($m = 80$), that is $s = 1$, given by

$$\{x_{ij}\} = \begin{pmatrix} x_{11} & \cdots & x_{1,10} & \cdots & x_{1,20} \\ x_{21} & \cdots & x_{2,10} & \cdots & x_{2,20} \\ x_{31} & \cdots & x_{3,10} & \cdots & x_{3,20} \\ x_{41} & \cdots & x_{4,10} & \cdots & x_{4,20} \end{pmatrix}.$$

(II) A complete data set($m = 40$), that is $s = 1$, given by

$$\{x_{ij}\} = \begin{pmatrix} x_{11} & \cdots & x_{1,10} & & \\ x_{21} & \cdots & x_{2,10} & & \\ x_{31} & \cdots & x_{3,10} & & * \\ x_{41} & \cdots & x_{4,10} & & \end{pmatrix}.$$

(III) A missing data set($m = 60$) with $s = 2$. Practically, the numbers of components for first and third rows are twenty and those for second and fourth rows are ten. We delete each of the ten observations in the second and fourth rows, that is, we have two kinds of missing patterns: $(x_{1j}, x_{2j}, x_{3j}, x_{4j})'$ and $(x_{1j}, *, x_{3j}, *)'$.

(IV) A missing data set($m = 60$) with $s = 3$. Practically, the numbers of components for first and second rows are twenty and those for third and fourth are five and fifteen, respectively. In this case, we delete the fifteen observations in the third row and the five observations in the fourth row, that is, we have three kinds of missing patterns: $(x_{1j}, x_{2j}, x_{3j}, x_{4j})'$, $(x_{1j}, x_{2j}, *, x_{4j})'$ and $(x_{1j}, x_{2j}, *, *)'$.

(V) A missing data set($m = 50$) with $s = 4$. Practically, the numbers of components for the first, second, third and fourth rows are twenty, fifteen, five and ten, respectively. In this case, we have four kinds of the missing patterns: $(x_{1j}, x_{2j}, x_{3j}, x_{4j})'$, $(x_{1j}, x_{2j}, *, x_{4j})'$, $(x_{1j}, *, *, x_{4j})'$ and $(x_{1j}, *, *, *)'$.

Further, we consider the following five cases for $p = 6$, $n = 20$:

(I) A complete data set($m = 120$), that is $s = 1$, given by

$$\{x_{ij}\} = \begin{pmatrix} x_{11} & \cdots & x_{1,10} & \cdots & x_{1,20} \\ x_{21} & \cdots & x_{2,10} & \cdots & x_{2,20} \\ x_{31} & \cdots & x_{3,10} & \cdots & x_{3,20} \\ x_{41} & \cdots & x_{4,10} & \cdots & x_{4,20} \\ x_{51} & \cdots & x_{5,10} & \cdots & x_{5,20} \\ x_{61} & \cdots & x_{6,10} & \cdots & x_{6,20} \end{pmatrix}.$$

(II) A complete data set($m = 60$), that is $s = 1$, given by

$$\{x_{ij}\} = \begin{pmatrix} x_{11} & \cdots & x_{1,10} & \\ x_{21} & \cdots & x_{2,10} & \\ x_{31} & \cdots & x_{3,10} & \\ x_{41} & \cdots & x_{4,10} & * \\ x_{51} & \cdots & x_{5,10} & \\ x_{61} & \cdots & x_{6,10} & \end{pmatrix}.$$

(III) A missing data set($m = 90$) with $s = 2$. Practically, the numbers of components for first, second, third and fifth rows are twenty and those for fourth and sixth rows are ten. We delete each of the ten observations in the fourth and sixth rows, that is, we have two kinds of missing patterns: $(x_{1j}, x_{2j}, x_{3j}, x_{4j}, x_{5j}, x_{6j})'$ and $(x_{1j}, x_{2j}, x_{3j}, *, x_{5j}, *)'$.

(IV) A missing data set($m = 90$) with $s = 3$. Practically, the numbers of components for first, second, third and fourth rows are twenty and those for fifth and sixth are five and fifteen, respectively. In this case, we delete the fifteen observations in the fifth row and the five observations in the sixth row, that is, we have three kinds of missing patterns: $(x_{1j}, x_{2j}, x_{3j}, x_{4j}, x_{5j}, x_{6j})'$, $(x_{1j}, x_{2j}, x_{3j}, x_{4j}, *, x_{6j})'$ and $(x_{1j}, x_{2j}, x_{3j}, x_{4j}, *, *)'$.

(V) A missing data set($m = 90$) with $s = 4$. Practically, the numbers of components for first, second and third rows are twenty and those for fourth, fifth and sixth rows are fifteen, ten and five, respectively. In this case, we have four kinds of the missing patterns: $(x_{1j}, x_{2j}, x_{3j}, x_{4j}, x_{5j}, x_{6j})'$, $(x_{1j}, x_{2j}, x_{3j}, x_{4j}, *, x_{6j})'$, $(x_{1j}, x_{2j}, x_{3j}, *, *, x_{6j})'$ and $(x_{1j}, x_{2j}, x_{3j}, *, *, *)'$.

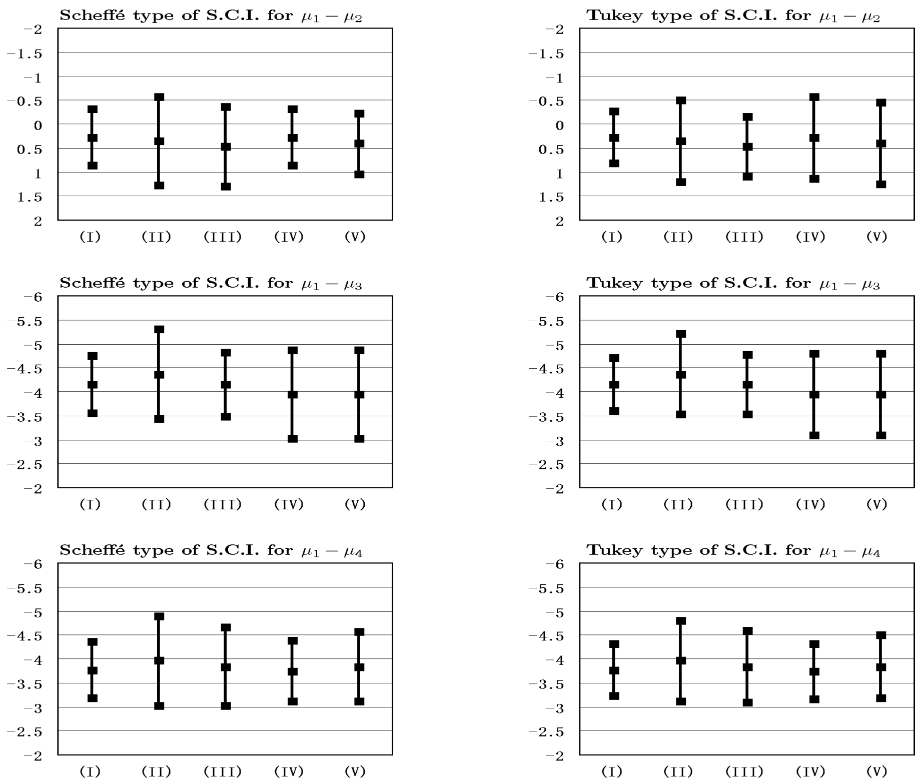

Figure 1. Scheffé and Tukey types of S.C.I.(simultaneous confidence intervals) for $p = 4$, $\sigma^2 = 1$, $\rho = 0.5$ and $\boldsymbol{\mu} = (1, 1, 5, 5)'$

In the same way, we selected the missing patterns for $p = 4, n = 40$ and $p = 6, n = 40$ respectively. Figures 1, 2 and 3 give the simultaneous confidence intervals for each parameter sets of $p = 4, n = 20, \boldsymbol{\mu} = (1, 1, 5, 5)', \sigma^2 = 1, \rho = 0, 5$ and $\alpha = 0.05$. For the cases of the complete data sets (I) and (II), we note that Scheffé type of simultaneous confidence intervals for $\mu_i - \mu_j, 1 \le i < j \le 4$ have the same width. Similarly, we note that Tukey type of simultaneous confidence intervals have the same width for the complete data sets (I) and (II).

For all cases (I) $\sim$ (V), it may be noted that the simultaneous confidence intervals for $\mu_1 - \mu_2$ and $\mu_3 - \mu_4$ include zero, and those for $\mu_1 - \mu_3$, $\mu_1 - \mu_4$, $\mu_2 - \mu_3$ and $\mu_2 - \mu_4$ do not include zero. In particular, comparing case (III) with (IV), it can be seen that the width for the case (IV) is shorter than that for case (III) in $\mu_1 - \mu_3$, $\mu_2 - \mu_3$ and $\mu_3 - \mu_4$. Concerning the width of intervals, it may be seen from the figures that the width tends to depend on total number of the observed data. Figure 3 gives Scheffé and Tukey types of simultaneous confidence intervals for some contrasts in the means, that is, $2\mu_1 - (\mu_2 + \mu_3), 3\mu_1 - (\mu_2 + \mu_3 + \mu_4), \mu_1 - \mu_2 - (\mu_3 - \mu_4)$ and $\mu_1 - \mu_3 - (\mu_2 - \mu_4)$.

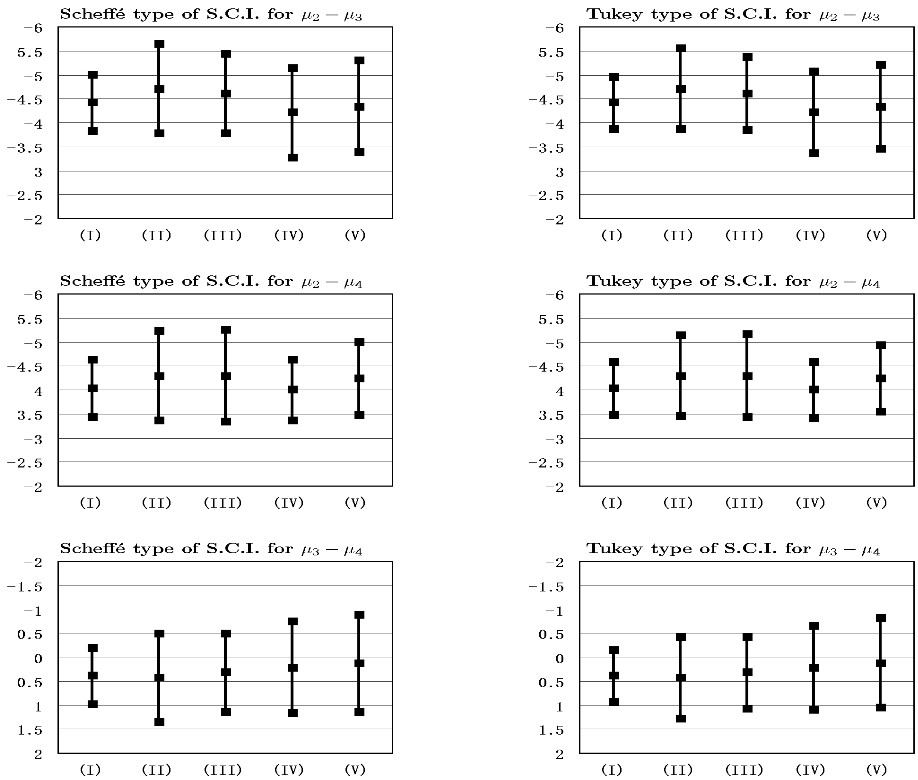

Figure 2. (continued)

It may be also noted from simulation results that Tukey type of simultaneous confidence intervals in (3.3) are shorter than Scheffé type of simultaneous confidence intervals in (3.1) if the cases are pairwise differences, otherwise Scheffé's tends to be shorter though it must be checked by the simulation study for the other cases. The mention of the above results is the same as the monotone type of missing observations (see, Seo, Kikuchi and Koizumi[10]). In conclusion, the proposed procedure in this paper is useful for the simultaneous confidence intervals under the general case of missing observations.

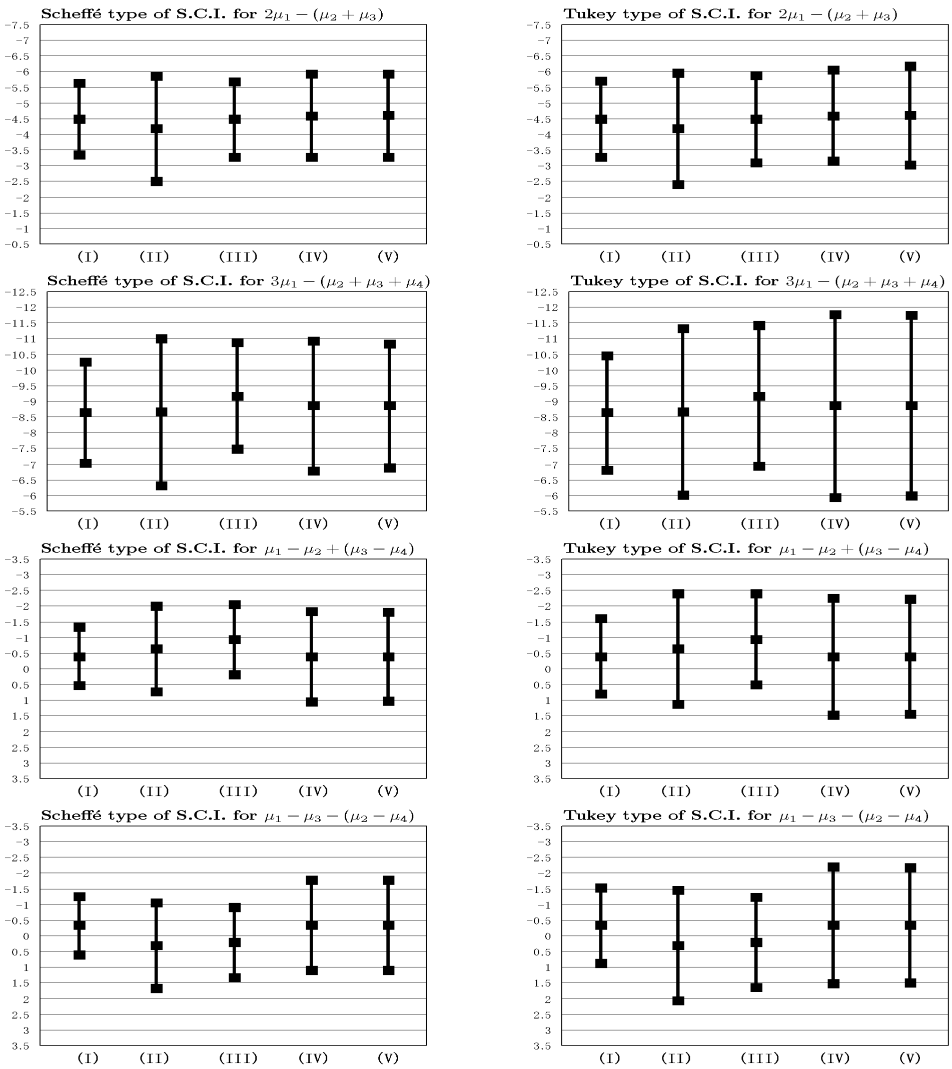

Figure 3. Scheffé and Tukey types of S.C.I.(simultaneous confidence intervals) for $p = 4$, $\sigma^2 = 1$, $\rho = 0.5$ and $\boldsymbol{\mu} = (1, 1, 5, 5)'$

Acknowledgements

The authors would like to thank the referee for his useful comments. The authors are also grateful to Hiromichi Isowa, Takao Kimura and Yuichi Koike for their help with drawing the graphs. The research of the second author was supported in part by Grant-in-Aid for Encouragement of Young Scientists under Contract Number 15700238.

References

[1] T. W. Anderson, Maximum Likelihood Estimates for a Multivariate Normal Distribution when some Observations are Missing, *J. Amer. Statist. Assoc.* 52 (1957) 200-203.

[2] R. P. Bhargava and M. S. Srivastava, On Tukey's Confidence Intervals for the Contrasts in the Means of the Intraclass Correlation Model, *J. Royal Statist. Soc.* B35 (1973) 147-152.

[3] T. Kanda and Y. Fujikoshi, Some Basic Properties of the MLE's for Multivariate Normal Distribution with Monotone Missing Data, *Amer. J. Math. Management Sci.* 18 (1998) 161-190.

[4] S. Konishi and K. Shimizu, Maximum Likelihood Estimation of an Intraclass Correlation in a Bivariate Normal Distribution with Missing Observations, *Comm. Statist. Theory Methods* 23 (1994) 1593-1604.

[5] R. J. A. Little and D. B. Rubin, Statistical Analysis with Missing Data, Second edition, Wiley Series in Probability and Statistics, New York, 2002.

[6] R. G. Miller, Simultaneous Statistical Inference, McGraw-Hill, New York, 1966.

[7] M. Minami and K. Shimizu, Estimation for a Common Intraclass Correlation in Bivariate Normal Distributions with Missing Observations, *Amer. J. Math. Management Sci.* 17 (1997) 3-14.

[8] C. R. Rao, Analysis of Distribution with Missing Observations, *J. Royal Statist. Soc.* B18 (1956) 259-264.

[9] H. Scheffé, The Analysis of Variance, Wiley, New York, 1959.

[10] T. Seo, J. Kikuchi and K. Koizumi, On simultaneous confidence intervals for all contrasts in the means of the intraclass correlation model with missing data, to appear in *Journal of Multivariate Analysis* (2006).

[11] T. Seo and M. S. Srivastava, Testing Equality of Means and Simultaneous Confidence Intervals in Repeated Measures with Missing Data, *Biometrical J.* 42 (2000) 981-993.

[12] M. S. Srivastava, Multivariate Data with Missing Observations, *Comm. Statist.* A 14 (1985) 775-792.

[13] M. S. Srivastava, Methods of Multivariate Statistics, Wiley, New York, 2002.

[14] M. S. Srivastava and E. M. Carter, The Maximum Likelihood Method for Non-Response in Sample Survey, *Survey Methodology* 12 (1986) 61-72.

[15] L. J. Wei and J. M. Lachin, Two Sample Asymptotically Distribution Free Tests for Incomplete Multivariate Observations, *J. Amer. Statist. Assoc.* 79 (1984) 653-661.

Kazuyuki Koizumi
Department of Mathematical Information Science, Tokyo University of Science
1-3, Kagurazaka, Shinjuku-ku, Tokyo 162-8601, Japan
E-mail: j1106702@ed.kagu.tus.ac.jp

Takashi Seo
Department of Mathematical Information Science, Tokyo University of Science
1-3, Kagurazaka, Shinjuku-ku, Tokyo 162-8601, Japan
E-mail: seo@rs.kagu.tus.ac.jp