Research Article

Simultaneous Inference on All Linear Combinations of Means with Heteroscedastic Errors

Xin Yan[1] and Xiaogang Su[2]

[1] *Department of Statistics, University of Central Florida, Orlando, FL 32816, USA*
[2] *School of Nursing, University of Alabama at Birmingham, Birmingham, AL 35294, USA*

Correspondence should be addressed to Xin Yan, xin.yan@ucf.edu

We proposed a statistical method to construct simultaneous confidence intervals on all linear combinations of means without assuming equal variance where the classical Scheffé's simultaneous confidence intervals no longer preserve the familywise error rate (FWER). The proposed method is useful when the number of comparisons on linear combinations of means is extremely large. The FWERs for proposed simultaneous confidence intervals under various configurations of mean variances are assessed through simulations and are found to preserve the predefined nominal level very well. An example of pairwise comparisons on heteroscedastic means is given to illustrate the proposed method.

1. Introduction

Multiple comparisons on a large number of linear combinations of means is of general interest in many applications. If an inferential statistical procedure relies on the number of comparisons, it may be quite challenge as the number of comparisons is increasing. Additionally, oftentimes we may not be able to make the assumption that all variances of means are equal. Many authors proposed various methods for multiple comparison on means in the past. Scheffé [1] proposed a method to construct simultaneous confidence intervals for all linear combinations of means while keeping Type I error under control. Since Scheffé's method constructs simultaneous confidence intervals for all possible linear combinations of means, his method has its own advantage when dealing with a large number of comparisons on linear combinations of means. It is understood that there are three major assumptions for

Scheffé's simultaneous confidence intervals to be constructed correctly. (1) The samples are independent, (2) the populations are normally distributed, and (3) populations have an equal variance. The third assumption, often referred to as homoscedasticity, is most vulnerable. The violation of homoscedasticity often results in inflation of the familywise error rate (FWER). As pointed out by Scheffé [2], his method has certain robustness when the group sample sizes are the same even when the variances are not equal. However, the FWER is out of control in situation where both the variances and sample sizes are unequal. No explicit formula is available so far for simultaneous confidence intervals on all linear combinations of means in the case of unequal variances.

The problem of comparisons on two means in the case of unequal population variances is known as the Behrens-Fisher problem [3]. Dunnett [4, 5] and Nel and van der Merwe [6] published simulation-based results on assessing different pairwise mean comparison procedures in the unequal variance case. Kim [7] proposed a practical solution to the Behrens-Fisher problem using the geometry of confidence ellipsoids for two mean vectors. Wilcox [8] tackled the Behrens-Fisher problem via trimmed means. Christensen and Rencher [9] compared Type I error rates and power levels in the Behrens-Fisher problem. Fouladi and Yockey [10] conducted a Monte Carlo study to evaluate the performance of the tests on means under the conditions of normality and abnormality. Hoover [11] discussed behavioral interventions with heterogeneous subgroup effects in clinical trials. In this paper, a method for constructing simultaneous confidence intervals on all linear combinations of means with unequal variances is proposed. Since there is no limitation for the number of linear combinations of means the proposed method may be used in situation where the comparisons on a large number of linear combinations of means is deemed to be necessary. The proposed simultaneous confidence intervals, to which we refer as the generalized Scheffé's confidence intervals, have an explicit format that is similar to their classical counterparts. The equal mean variance assumption is no longer needed. In addition, these simultaneous confidence intervals become the classical Scheffé's confidence intervals when all population variances and sample sizes are equal. Most importantly, the proposed simultaneous confidence intervals preserve FWER in all configurations of variances and sample sizes.

2. Generalized Scheffé Confidence Intervals

Suppose that we have I populations and let (μ_i, σ_i^2) be the true mean and variance for population i. Let (n_i, D_i, S_i^2) be the sample size, sample mean, and sample variance of the ith population. In the case of equal variance among I populations, that is, $\sigma_i^2 \equiv \sigma^2$, Scheffé simultaneous confidence intervals on *all* linear combinations of means $\sum_{i=1}^{I} c_i \mu_i$ are given by:

$$\sum_{i=1}^{I} c_i D_i \pm \sqrt{I \cdot F_{\alpha, I, N-I}} \sqrt{\mathrm{MSE} \cdot \sum_{i=1}^{I} \frac{c_i^2}{n_i}}, \tag{2.1}$$

where the mean squared error $\mathrm{MSE} = \sum_{i=1}^{I} (n_i - 1) S_i^2 / (N - I)$ is the pooled estimate of the common variance from I populations; $F_{\alpha, I, N-I}$ is the upper αth quantile from the F distribution with degrees of freedom I, $N - I$; $N = \sum_{i=1}^{I} n_i$ is the total sample size. If I constants

$c_1, c_2, \ldots, c_I$ satisfy $\sum_{i=1}^{I} c_i = 0$, Scheffé's simultaneous confidence intervals on *all* contracts $\sum_{i=1}^{I} c_i \mu_i$ are given by:

$$\sum_{i=1}^{I} c_i D_i \pm \sqrt{(I-1) \cdot F_{\alpha, I-1, N-I}} \sqrt{\text{MSE} \cdot \sum_{i=1}^{I} \frac{c_i^2}{n_i}}. \tag{2.2}$$

If pairwise comparisons are of interest, we can set one pair of (c_i, c_j) to be $(1, -1)$ and rest c_is to be zero. This is a special case of contrast. Note that Scheffé's intervals are useful when dealing with a large number of linear combinations of means. When the total number of observations and the number of populations are determined, the quantity $F_{\alpha, I, N-I}$ stays the same regardless the number of simultaneous confidence intervals. For the Bonferroni approach, the width of the confidence intervals tend to be wider if the number of linear combinations of means is increasing. Suppose that we have 10 populations each with a sample size 10. If we have 100 simultaneous confidence intervals for the linear combinations of means, the $\sqrt{F}$ in Scheffé's method is $\sqrt{10 \times F_{0.05, 10, 100-10}} = 1.9635$. If we apply Bonferroni's approach the $|t(0.05/200, 100-10)| = 3.6118$. This means that the width of Scheffé's intervals may be shorter than the width of the Bonferroni's intervals. There is a breakdown point such that Scheffé's intervals may be shorter than the Bonferroni's intervals when the number of linear combinations of means gets larger. This alerts the common perception that Scheffé's intervals are more conservative than Bonfferoni's intervals.

We now consider the problem of constructing simultaneous intervals without assuming equal variance. Let $a_i = \sigma_i^2 / \sum_{i=1}^{I} \sigma_i^2$ and define

$$R_1 = \sum_{i=1}^{I} a_i \left(\frac{D_i - \mu_i}{\sigma_i / \sqrt{n_i}} \right)^2 = \sum_{i=1}^{I} a_i Y_i, \qquad R_2 = \sum_{i=1}^{I} \frac{a_i}{n_i - 1} \frac{(n_i - 1) S_i^2}{\sigma_i^2} = \sum_{i=1}^{I} \frac{a_i}{n_i - 1} Z_i. \tag{2.3}$$

Note that $Y_i \sim \chi_1^2$ and $Z_i \sim \chi_{(n_i-1)}^2$. Therefore, R_1 and R_2 are linear combinations of χ^2 variables with $E(R_1) = E(R_2) = 1$.

Finding the exact distribution of linear combination of χ^2 variables, known as Satterthwaite's problem, is rather difficult. Satterthwaite tried to approximate this type of variable as a χ_ν^2 random variable divided by its degrees of freedom ν (see [12]). This degree of freedom ν is then solved via the method of moment estimation. As noted in Casella and Berger [12], for a variable $Y \sim \chi_\nu^2 / \nu$, we have $E(Y) = 1$. Hence

$$\nu = \frac{2(EY)^2}{\text{Var}(Y)} = \frac{2}{\text{Var}(Y)}. \tag{2.4}$$

We then set $R_1 \sim \chi_{\nu_1}^2 / \nu_1$, and $R_2 \sim \chi_{\nu_2}^2 / \nu_2$, where ν_1 and ν_2 are the respective degrees of freedom for R_1 and R_2. By applying the results above we can estimate ν_1 and ν_2. First we consider ν_1, which can be found as

$$\nu_1 = \frac{2}{\sum_{i=1}^{I} a_i^2 \text{Var}(Y_i)} = \frac{1}{\sum_{i=1}^{I} a_i^2} = \frac{\left(\sum_{i=1}^{I} \sigma_i^2 \right)^2}{\sum_{i=1}^{I} \sigma_i^4}. \tag{2.5}$$

A natural estimate of v_1 is given by $\widehat{v}_1 = (\sum_{i=1}^{I} S_i^2)^2 / \sum_{i=1}^{I} S_i^4$. For v_2, we have

$$v_2 = \frac{2}{\sum_{i=1}^{I} \left(a_i^2 / (n_i - 1)^2 \right) \mathrm{Var}(Z_i)} = \frac{1}{\sum_{i=1}^{I} a_i^2 / (n_i - 1)} = \frac{\left(\sum_{i=1}^{I} \sigma_i^2 \right)^2}{\sum_{i=1}^{I} \sigma_i^4 / (n_i - 1)}. \tag{2.6}$$

It can be estimated by $\widehat{v}_2$ and $\widehat{v}_2 = (\sum_{i=1}^{I} S_i^2)^2 / \sum_{i=1}^{I} S_i^4 / n_i - 1$. Furthermore, note that R_1 is independent of R_2, therefore, $R = R_1 / R_2$ has approximately the F distribution with degrees of freedom v_1 and v_2. It turns out that $R = R_1 / R_2$ has a very simple form

$$R = \frac{R_1}{R_2} = \frac{\sum_{i=1}^{I} n_i (D_i - \mu_i)^2}{\sum_{i=1}^{I} S_i^2} \sim F_{v_1, v_2}. \tag{2.7}$$

Note that if the I populations have equal variance, $\sigma_i^2 \equiv \sigma^2$, we have $v_1 = I$; additionally, if all populations have the same sample size, that is, $n_i \equiv n$, then $v_2 = N - I$.

To derive the generalized Scheffé's interval we would need the following projection lemma (see [13] pages 231-232). For I real numbers $z_1, z_2, \ldots, z_I$ and all $\mathbf{a} = (a_1, a_2, \ldots, a_I) \in \mathfrak{R}^I$ to satisfy the following inequality:

$$\sum_{i=1}^{I} a_i y_i - r \left(\sum_{i=1}^{I} a_i^2 \right)^{1/2} \leq \sum_{i=1}^{I} a_i z_i \leq \sum_{i=1}^{I} a_i y_i + r \left(\sum_{i=1}^{I} a_i^2 \right)^{1/2}, \tag{2.8}$$

the necessary and sufficient condition is $\sum_{i=1}^{I} (z_i - y_i)^2 \leq r^2$. We then choose $z_i = \sqrt{n_i} \mu_i$ and let $\mathbf{z} = (z_1, z_2, \ldots, z_I)$ satisfy $\sum_{i=1}^{I} (z_i - \sqrt{n_i} D_i)^2 \leq F_{\alpha, \widehat{v}_1, \widehat{v}_2} \sum_{i=1}^{I} S_i^2$ which constitutes the interior of a I-dimensional sphere centered at the point $(\sqrt{n_1} D_1, \sqrt{n_2} D_2, \ldots, \sqrt{n_I} D_I)$ with radius $\sqrt{F_{\alpha, \widehat{v}_1, \widehat{v}_2} \sum_{i=1}^{I} S_i^2}$. By applying the projection lemma to vector $\mathbf{a}$, where $\mathbf{a} = (c_1 / \sqrt{n_1}, c_2 / \sqrt{n_2}, \ldots, c_I / \sqrt{n_I})$, we have

$$\left\{ \sum_{i=1}^{I} (\sqrt{n_i} D_i - \sqrt{n_i} \mu_i)^2 \leq F_{\alpha, \widehat{v}_1, \widehat{v}_2} \sum_{i=1}^{I} S_i^2 \right\}$$

$$= \left\{ \sum_{i=1}^{I} \frac{c_i}{\sqrt{n_i}} \sqrt{n_i} \mu_i \in \sum_{i=1}^{I} \frac{c_i}{\sqrt{n_i}} \sqrt{n_i} D_i \pm \sqrt{F_{\alpha, \widehat{v}_1, \widehat{v}_2}} \sqrt{\sum_{i=1}^{I} S_i^2} \sqrt{\sum_{i=1}^{I} \frac{c_i^2}{n_i}} \right\} \tag{2.9}$$

$$= \left\{ \sum_{i=1}^{I} c_i \mu_i \in \sum_{i=1}^{I} c_i D_i \pm \sqrt{F_{\alpha, \widehat{v}_1, \widehat{v}_2}} \sqrt{\sum_{i=1}^{I} S_i^2} \sqrt{\sum_{i=1}^{I} \frac{c_i^2}{n_i}} \right\}.$$

Choosing $F_{\alpha, \widehat{v}_1, \widehat{v}_2}$, the $1 - \alpha$ quantile of an F distribution with $\widehat{v}_1$ and $\widehat{v}_2$ degrees of freedom, based on the results in (2.7), we have

$$P \left\{ \sum_{i=1}^{I} (\sqrt{n_i} D_i - \sqrt{n_i} \mu_i)^2 \leq F_{\alpha, \widehat{v}_1, \widehat{v}_2} \sum_{i=1}^{I} S_i^2 \right\} = 1 - \alpha. \tag{2.10}$$

Applying the projection lemma this probability can be pivoted to give the following generalized $1 - \alpha$ simultaneous confidence intervals for $\sum_{i=1}^{I} c_i \mu_i$,

$$\sum_{i=1}^{I} c_i D_i \pm \sqrt{F_{\alpha, \widehat{\nu}_1, \widehat{\nu}_2}} \sqrt{\sum_{i=1}^{I} S_i^2} \sqrt{\sum_{i=1}^{I} \frac{c_i^2}{n_i}}. \tag{2.11}$$

For population mean μ_i's and their pairwise differences $\mu_i - \mu_j$, the generalized Scheffé's confidence intervals are

$$D_i \pm \sqrt{F_{\alpha, \widehat{\nu}_1, \widehat{\nu}_2}} \sqrt{\sum_{i=1}^{I} S_i^2} \sqrt{\frac{1}{n_i}}, \tag{2.12}$$

$$D_i - D_j \pm \sqrt{F_{\alpha, \widehat{\nu}_1, \widehat{\nu}_2}} \sqrt{\sum_{i=1}^{I} S_i^2} \sqrt{\frac{1}{n_i} + \frac{1}{n_j}}, \tag{2.13}$$

where $1 \leq i \neq j \leq I$. By comparing (2.1) with (2.11), it can be seen that the generalized Scheffé's confidence intervals are very similar to their classical counterparts.

3. Assessment of Familywise Error Rate

The Type I error in multiple comparisons is referred to as the probability of incorrectly rejecting at least one of the null hypotheses that make up the family. The validity of the proposed generalized Scheffé's confidence intervals largely lies in successfully controlling the FWER at a given nominal level α.

There are two major factors, population sample sizes and variances, which affect the performance of the Scheffé's confidence intervals. We will show through simulation that the FWER will be inflated in the situation where population variances are unequal.

A variety of configurations of variances and sample sizes will be selected to assess the performance of the generalized Scheffé method. To this end, the number of groups is chosen to be $I = 4$. Without loss of generality, we use 0 for all population means, that is, $(\mu_1, \mu_2, \mu_3, \mu_4) = (0, 0, 0, 0)$. The specification of sample sizes and variances is given in Table 1.

Although Scheffé's intervals apply to inference on all linear combinations, for simplicity, we have focused on two sets of inferences only: population means and their pairwise differences. For each configuration we conducted 5,000 simulation runs and for each run 95% Scheffé's intervals and generalized Scheffé's intervals on both population means and pairwise mean differences were computed. We then obtained the coverage rates that the proposed intervals contain the true means, which all equal 0.

Table 1 reports the coverage rates based on both methods. Note that the empirical FWER would be one minus the coverage rate. Clearly, in the case of equal variances, both methods give very similar rates of coverage for balanced design or unbalanced design. In the unequal variance case, the coverage rate of Scheffé's method drops. However, its FWER still stays well within the nominal level, that is, around $\alpha = 0.05$, for balanced designs. This confirms the notion of Scheffé that his method is robust to heteroscedasticity when sample sizes from populations are equal. We notice that the FWER is

Table 1: Coverage rates of 95% Scheffé's intervals (S) and generalized Scheffé (GS) intervals: two sets of inferences are considered, the population means and pairwise mean differences.

Sample size	Equal variances				Unequal variances			
	(0.1, 0.1, 0.1, 0.1)		(1, 1, 1, 1)		(0.3, 0.3, 0.1, 0.1)		(3, 3, 1, 1)	
	S	GS	S	GS	S	GS	S	GS
Balanced								
(5, 5, 5, 5)	98.00	98.60	98.45	99.00	93.60	96.85	94.05	97.35
(10, 10, 10, 10)	97.90	98.45	98.20	98.65	94.75	97.10	95.10	97.30
(20, 20, 20, 20)	97.70	97.90	98.20	98.35	93.90	96.25	94.80	96.45
(50, 50, 50, 50)	97.90	97.95	98.35	98.35	94.35	96.75	94.55	96.60
Unbalanced								
(5, 5, 10, 10)	98.20	98.75	98.20	98.70	87.70	97.50	87.20	97.40
(5, 5, 20, 20)	98.40	99.10	97.90	98.45	73.00	96.20	76.50	96.65
(10, 10, 20, 20)	97.95	98.05	98.35	98.35	88.40	97.30	87.05	96.65
(10, 10, 50, 50)	98.60	98.80	98.70	98.65	73.95	96.70	72.55	97.10

inflated when sample sizes are different among the populations. It can be found from Table 1 that when $(\sigma_1, \sigma_2, \sigma_3, \sigma_4) = (0.3, 0.3, 0.1, 0.1)$, for sample sizes $(n_1, n_2, n_3, n_4) = (5, 5, 10, 10), (5, 5, 20, 20), (10, 10, 20, 20), (10, 10, 50, 50)$ the FWERs are 12.3%, 27%, 11.6%, and 26.5%, respectively. When $(\sigma_1, \sigma_2, \sigma_3, \sigma_4) = (3, 3, 1, 1)$, for sample sizes $(n_1, n_2, n_3, n_4) = (5, 5, 10, 10), (5, 5, 20, 20), (10, 10, 20, 20), (10, 10, 50, 50)$, the FWERs are 12.8%, 23.5%, 12.95%, and 27.45%, respectively. Note that these FWERs are all significantly greater than the nominal level $\alpha = 0.05\%$. It can be seen that the greater the difference in sample sizes is the larger the corresponding FWER will be. On the other hand, the performance of the generalized Scheffé method is much more robust. For the same configuration settings, the FWERs based on the generalized Scheffé's intervals are between 0.025% and 0.038%. Although it is conservative, but it stays well within the nominal level of $\alpha = 0.05$.

It would also be interesting to see how different in width the two types of intervals are. Comparing (2.1) with (2.12), one can see that the difference between them are due to the following two terms:

$$Q_1 = \sqrt{F_{\alpha, I, N-I} \cdot I \cdot \text{MSE}},$$

$$Q_2 = \sqrt{F_{\alpha, \widehat{\nu}_1, \widehat{\nu}_2} \cdot \sum S_i^2}.$$

(3.1)

The averaged Q_1 and Q_2 from 5,000 simulation runs are presented in Table 2.

It can be seen that they are very close to each other in the case of equal variances. However, in the case of unequal variances, Q_1 becomes over optimistically smaller than Q_2, which leads to the inflation of FWER. Finally, Scheffé's intervals are derived from the fact that the F statistic

$$F = \frac{\sum_{i=1}^{I} (D_i - \mu_i)^2 / I}{\text{MSE}},$$

(3.2)

follows the $F_{I,N-I}$ distribution under a number of assumptions. When these assumptions are violated, the performance of Scheffé's intervals would depend on how the above F statistic deviates from the distribution $F_{I,N-I}$. For the generalized Scheffé's intervals, the FWER

Table 2: Comparison of interval widths between Scheffé's and generalized Scheffé's methods. Their interval widths differ in quantities: $Q_1 = \sqrt{I \cdot F_{\alpha,I,N-I} \cdot \mathrm{MSE}}$ in Scheffé's method and $Q_2 = \sqrt{F_{\alpha,\hat{\nu}_1,\hat{\nu}_2} \cdot \sum S_i^2}$ in the generalized Scheffé's method ($\alpha = 0.05$).

Sample size	Equal variances				Unequal variances			
	(0.1, 0.1, 0.1, 0.1)		(1, 1, 1, 1)		(0.3, 0.3, 0.1, 0.1)		(3, 3, 1, 1)	
Balanced	Q_1	Q_2	Q_1	Q_2	Q_1	Q_2	Q_1	Q_2
(5, 5, 5, 5)	0.343	0.370	3.422	3.680	0.754	0.909	7.598	9.182
(10, 10, 10, 10)	0.322	0.331	3.229	3.323	0.718	0.813	7.166	8.105
(20, 20, 20, 20)	0.315	0.319	3.153	3.194	0.703	0.778	7.032	7.780
(50, 50, 50, 50)	0.310	0.312	3.105	3.120	0.694	0.759	6.945	7.595
Unbalanced								
(5, 5, 10, 10)	0.329	0.350	3.284	3.490	0.602	0.905	6.052	9.125
(5, 5, 20, 20)	0.318	0.340	3.196	3.423	0.489	0.905	4.894	9.093
(10, 10, 20, 20)	0.318	0.326	3.173	3.250	0.597	0.812	5.951	8.092
(10, 10, 50, 50)	0.312	0.321	3.128	3.218	0.466	0.810	4.669	8.138

largely depends on how accurately R_1/R_2 approximates F_{ν_1,ν_2}. Figure 1 plots the empirical distribution function of R_1/R_2 and the F statistic in (2.13), along with their designated F distribution. We selected the following four different configurations of variances and sample sizes, which correspond to homoscedastic/heteroscedastic and balanced/unbalanced cases:

(1) $(\sigma_1, \sigma_2, \sigma_3, \sigma_4) = (1, 1, 1, 1)$, $(n_1, n_2, n_3, n_4) = (10, 10, 10, 10)$, $(10, 10, 50, 50)$,

(2) $(\sigma_1, \sigma_2, \sigma_3, \sigma_4) = (3, 3, 1, 1)$, $(n_1, n_2, n_3, n_4) = (10, 10, 10, 10)$, $(10, 10, 50, 50)$.

The configuration (1) indicates the equal variance for the 4 means with equal or different sample sizes. The configuration (2) indicates the unequal variances for the 4 means with equal or different sample sizes. We calculate the empirical distribution function of R_1/R_2, and it can be seen that they are nearly overlaps with F_{ν_1,ν_2} in all four cases of configurations of variance and sample sizes (1(a)–4(a) in Figure 1). The overlapping between edf of R_1/R_2 and F_{ν_1,ν_2} suggests an excellent approximation of the F distribution to the ratio of R_1 and R_2. In addition, the edf of the F statistic also matches well with the distribution $F_{I,N-I}$ (1(b)–3(b) in Figure 1), except in the unbalanced heteroscedastic case where Scheffé's method fails (4(b) in Figure 1). This explains why the FWER is inflated in the case of unequal variances.

One last comment, the above simulation results suggest that the widths of the generalized Scheffé intervals tend to be wider than that of the Scheffé intervals. This is our overall impression, but may not always be true in general. In the simulations, from time to time, we observed narrower generalized Scheffé intervals. We will see this feature from the data analysis example in the next section.

4. Example of Data Analysis

Solomon et al. [14] studied smoking behavior in pregnant women. They examined the women's determination to quit smoking while pregnant. They interviewed 349 women at their first prenatal visit, all of whom were smokers when they became pregnant, and were classified into four groups: precontemplation (PC), contemplation (C), preparation (P), and action (A). Their intention was to look at the subsequent smoking behavior of these subjects during the course of pregnancy, but one important consideration was how much
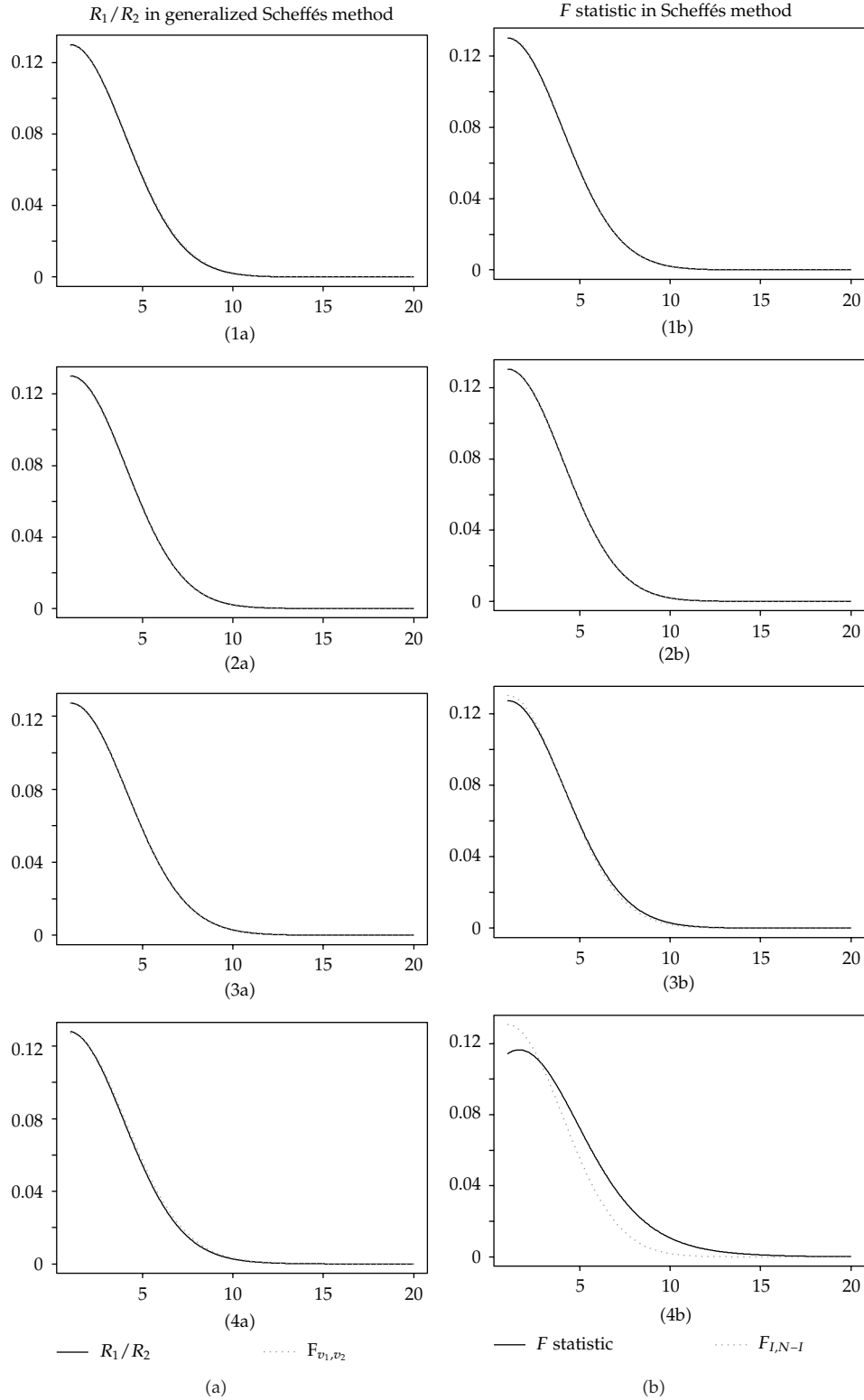
Figure 1: Empirical density plots: each density curve is generated from 5000 simulation runs. The solid line is for the F statistic or R_1/R_2 and the dotted line is for their designated F distribution.

Table 3: The simultaneous Scheffé intervals and Generalized Scheffé's intervals on means and pairwise mean differences in the cigarette example.

Parameters	Scheffé	Generalized Scheffé
Mean		
μ_{PC}	(20.66, 28.94)	(20.76, 28.84)
μ_{C}	(10.95, 22.25)	(11.08, 22.12)
μ_{P}	(26.02, 31.58)	(26.09, 31.51)
μ_{A}	(10.08, 17.32)	(10.16, 17.24)
Pairwise comparisons		
$\mu_{PC} - \mu_{C}$	(1.19, 15.21)	(1.36, 15.04)
$\mu_{PC} - \mu_{P}$	(−8.98, 0.98)	(−8.87, 0.87)
$\mu_{PC} - \mu_{A}$	(5.59, 16.60)	(5.73, 16.47)
$\mu_{C} - \mu_{P}$	(−18.49, −5.90)	(−18.35, −6.05)
$\mu_{C} - \mu_{A}$	(−3.81, 9.61)	(−3.65, 9.45)
$\mu_{P} - \mu_{A}$	(10.53, 19.67)	(10.64, 19.56)

Table 4: Sample sizes, means, and sample standard deviations of 349 women who stopped smoking during pregnancy period [14].

Label	Condition	Description	n_i	$\overline{y}_i$	s_i
PC	Precontemplation	Smokes and has no plan to quit smoking	69	24.8	13.3
C	Contemplation	Smokes but is thinking of quitting	37	16.6	5.2
P	Preparation	Smokes but has made some effort at quitting	153	28.8	12.2
A	Action	Has already quit	90	13.7	8.8

these women smoked when they became pregnant. The sample sizes, means, and standard deviations of these four groups, in terms of cigarettes smoked per day when they became pregnant, are given in Table 4. Noting that the smallest sample size is 37, we do not need to worry about the normality assumption even if the response of interest is count or integer.

Table 3 presents the 95% Scheffé's intervals and the generalized Scheffé intervals for the four group means and their differences. Since both sample sizes and variances are quite different from each other, the generalized Scheffé intervals are more reliable.

One may make a number of inferences with a joint confidence level of 95%. For example, women in the preparation (P) group have an average number of cigarettes every day ranging from 26.09 to 31.51, which seems to be the most frequent smoker group. There is no significant difference found between group P and group PC, because their difference has a confidence interval (−8.87, 0.87) that includes 0. It is also quite interesting to notice that the generalized Scheffé's intervals are even narrower than the Scheffé intervals.

5. Discussion

Among others, the Scheffé method is one of the commonly-used method to make simultaneous inference on all linear combinations of means. Scheffé intervals are for all possible linear combinations of means and this brings benefit if a large number of linear combinations of means need to be compared. Assumption of equal variance for all means is needed to control type I error. When this assumption is violated the proposed method can be conveniently used for constructing simultaneous confidence intervals where type I error

is controlled at a prespecified nominal level. Results from simulations show that the FWER of the proposed simultaneous confidence intervals are well preserved at a nominal level and the equal variance assumption can be simply ignored.

References

[1] H. Scheffé, "A method for judging all contrasts in the analysis of variance," *Biometrika*, vol. 40, pp. 87–104, 1953.

[2] H. Scheffé, *The Analysis of Variance*, John Wiley & Sons, New York, NY, USA, 1959.

[3] H. Scheffé, "Practical solutions of the Behrens-Fisher problem," *Journal of the American Statistical Association*, vol. 65, pp. 1501–1504, 1970.

[4] C. W. Dunnett, "Pairwise multiple comparison in the unequal variance case," *Journal of American Statistical Association*, vol. 75, pp. 796–800, 1980.

[5] C. W. Dunnett, "Pairwise multiple comparisons in the homogeneous variance, unequal sample size case," *Journal of American Statistical Association*, vol. 75, pp. 789–795, 1980.

[6] D. G. Nel and C. A. van der Merwe, "A solution to the multivariate Behrens-Fisher problem," *Communications in Statistics. Simulation and Computation*, vol. 15, no. 12, pp. 3719–3735, 1986.

[7] S. J. Kim, "A practical solution to the multivariate Behrens-Fisher problem," *Biometrika*, vol. 79, no. 1, pp. 171–176, 1992.

[8] R. R. Wilcox, "Simulation results on solutions to the multivariate Behrens-Fisher problem via trimmed means," *The Statistician*, vol. 44, pp. 213–225, 1995.

[9] W. F. Christensen and A. C. Rencher, "A comparison of type I error rates and power levels for seven solutions to the multivariate Behrens-Fisher problem," *Communications in Statistics. Simulation and Computation*, vol. 26, no. 4, pp. 1251–1273, 1997.

[10] R. T. Fouladi and R. D. Yockey, "Type I error control of two-group multivariate tests on means under conditions of heterogeneous correlation structure and varied multivariate distributions," *Communications in Statistics. Simulation and Computation*, vol. 31, no. 3, pp. 375–400, 2002.

[11] D. R. Hoover, "Clinical trials of behavioural interventions with heterogeneous teaching subgroup effects," *Statistics in Medicine*, vol. 30, pp. 1351–1364, 2002.

[12] G. Casella and R. L. Berger, *Statistical Inference*, Duxbury, 2001.

[13] J. C. Hsu, *Multiple Comparisons, Theory and Methods*, Chapman and Hall/CRC, London, UK, 1999.

[14] L. J. Solomon, R. H. Secker-Walker, J. M. Skelly, and B. S. Flynn, "Stages of change in smoking during pregnancy in low-income women," *Journal of Behavioral Medicine*, vol. 19, no. 4, pp. 333–344, 1996.

Advances in
Operations Research

Advances in
Decision Sciences

Mathematical Problems
in Engineering

Algebra

Journal of
Probability and Statistics

The Scientific
World Journal

International Journal of
Differential Equations

International Journal of
Combinatorics

Advances in
Mathematical Physics

Journal of
Complex Analysis

Journal of
Mathematics

International Journal of
Stochastic Analysis

Abstract and
Applied Analysis

Discrete Dynamics in
Nature and Society

International
Journal of
Mathematics and
Mathematical
Sciences

Journal of
Discrete Mathematics

Journal of
Function Spaces

Journal of
Applied Mathematics

Journal of
Optimization

Submit your manuscripts at
http://www.hindawi.com

Hindawi