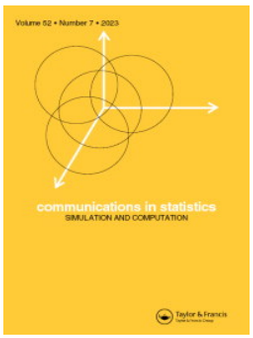

Taylor & Francis
Taylor & Francis Group

Communications in Statistics - Simulation and Computation

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/lssp20

Single-stage sampling procedure for heteroscedasticity in multiple comparisons with a control

Miin-Jye Wen, Chun-Che Wen & Wei-Ming Wang

To cite this article: Miin-Jye Wen, Chun-Che Wen & Wei-Ming Wang (2022): Single-stage sampling procedure for heteroscedasticity in multiple comparisons with a control, Communications in Statistics - Simulation and Computation, DOI: 10.1080/03610918.2022.2053863

To link to this article: https://doi.org/10.1080/03610918.2022.2053863

COMMUNICATIONS IN STATISTICS - SIMULATION AND COMPUTATION®
https://doi.org/10.1080/03610918.2022.2053863

Single-stage sampling procedure for heteroscedasticity in multiple comparisons with a control

Miin-Jye Wen[a,b], Chun-Che Wen[c], and Wei-Ming Wang[a]

[a]Department of Statistics and Institute of Data Science, College of Management, National Cheng Kung University, Tainan, Taiwan; [b]Institute of International Management, College of Management, National Cheng Kung University, Tainan, Taiwan; [c]Department of Public Health Sciences, College of Medicine, Medical University of South Carolina, Charleston, South Carolina, USA

ABSTRACT

Given $k(\geq 2)$ independent and normally distributed populations with unknown means and unknown (and possibly unequal) variances, we used the single-stage sampling procedure for multiple comparison with a control to the problem of unequal sample sizes. A simulation of the family-wise error type I was conducted to validate the quality of the procedure and a numerical example was provided for illustrating this procedure. We provided a user-friendly interface by *R* shiny to obtain the critical values for the practitioners.

ARTICLE HISTORY
Received 20 September 2021
Accepted 10 March 2022

KEYWORDS
Simultaneous confidence intervals; Student *t* distribution; Unknown and unequal variances

1. Introduction

Let $\pi_1, ..., \pi_k$ denoted $k(\geq 2)$ independent populations so that observations X_{ij} from π_i are normally distributed with unknown mean μ_i and unknown variance $\sigma_i^2, i = 1, ..., k$. Without loss of generality, assume that is the control population π_1 with mean μ_1 and variance σ_1^2. Our research objective is to find simultaneous confidence interval on the $k-1$ paired differences $\mu_i - \mu_1, i = 2, ..., k$. Paulson (1952) and Dunnett (1955) considered the case with equal but unknown variances σ_i^2. For an overview of these multiple-comparison procedures can refer to Miller (1981).

However, the assumption of homogeneous variances between groups is not always realistic and may lead to incorrect decisions. For example, dose finding research may have the problem of heteroscedasticity since the data's variance relies on the dose effects. (e.g., see the data in Westfall (1997) that are available in the *R* package *multcomp*). In case of unequal variances, Tamhane (1977) proposed approximated approaches for the multiple comparisons with all-pairwise and a control comparison. Under the heteroscedasticity for all-pairwise differences, Games and Howell (1976) provided the approximate simultaneous confidence intervals. Hasler and Hothorn (2008) proposed a procedure for the multiple contrast test that provided approximated simultaneous confidence intervals for all-contrast comparisons. Without the equal variance assumption, Li and Ning (2012) proposed three approximate simultaneous confidence intervals procedures for multiple comparison with a control, which were derived from Bonferroni approximation, Slepian's inequality, and multivariate *t* distribution.

CONTACT Wei-Ming Wang weiming3524@gmail.com, r28051012@mail.ncku.edu.tw Department of Statistics and Institute of Data Science, College of Management, National Cheng Kung University, Tainan, Taiwan.

Supplemental data for this article is available online at https://doi.org/10.1080/03610918.2022.2053863.

When these variances $\sigma_i^2 (i = 1, ..., k)$ are unknown and possibly unequal, Dudewicz and Dalal (1983) considered a two-stage procedure for comparing all μ_i's with μ_1 to get exact distributions of the test statistics. Therefore, the distributions are free from unknown population variances. Stein (1945) first brought up the idea of two-stage procedure in his test for a fixed-width interval estimation of a normal mean and for a hypothesis. However, the two-stage procedure is design-oriented, which means a final sample size is to be determined with a fixed width or a controlled length. The drawback of the two-stage procedure is that one has to obtain equal sample sizes in the first stage and additional samples in the second stage. Nevertheless, in real scenarios, this may not be always practical. Sometimes, only one sampling procedure is available in an experiment due to limited time, budget, etc. The question of how to use these data in estimation or testing hypothesis under heteroscedasticity is worth our attention. Bishop (1976) used the concept of two-stage sampling procedure for a single sample by splitting them into two data sets in the analysis of variance problems. Bishop's (1976) single-stage sampling procedure was then adopted by Chen and Lam (1989) in the interval estimation of the largest normal mean. Further, Lam (1992) extended the single-stage to the subset selection procedure. In addition, Wen and Chen (1994) used the single-stage sampling procedure in the multiple comparison with the best and a control when the sample sizes are equal. For the applications of the single-stage sampling procedure, please see Chen and Chen (1998), Chen (2001), Chen, Chen, and Chang (2004), Wen, Chen, and Chen (2007), Chen, Wen, and Wang (2009), Chen, Wen, and Chuang (2011), and Wen, Huang, and Zhong (2017).

In this paper, the single-stage sampling procedure is introduced in Sec. 2. In Sec. 3, multiple comparisons with a control for one-sided and two-sided intervals are considered. In Sec. 4, the familywise error type I (FWE) was simulated to validate the quality of the procedure. In Sec. 5, a numerical example is demonstrated to show how to apply these procedures, and the last section is conclusion.

2. Single-stage sampling procedure

Let X_{ij} $(i = 1, ..., k; j = 1, ..., n_i)$ be an independent random sample from the normal population π_i with unknown mean μ_i and unknown and unequal variance $\sigma_i^2 (i = 1, ..., k)$. Among each population $\pi_i (i = 1, ..., k)$, we selected the first $n_i - 1$ observations to define the sample mean and the sample variance, respectively, by

$$\bar{X}_i = \frac{\sum_{j=1}^{n_i-1} X_{ij}}{n_i - 1} \text{ and } S_i^2 = \frac{\sum_{j=1}^{n_i-1} (X_{ij} - \bar{X}_i)^2}{n_i - 2}, i = 1, ..., k.$$

The sample size, n_i (≥ 3), of each population $\pi_i (i = 1, ..., k)$ is split into two parts: the first part is $n_i - 1$ observations, and the remaining observation to calculate the weighted sample mean $\tilde{X}_i$ as

$$\tilde{X}_i = \sum_{j=1}^{n_i} \omega_{ij} X_{ij} = U_i \sum_{j=1}^{n_i-1} X_{ij} + V_i X_{in_i}, \tag{1}$$

where

$$\omega_{ij} = \begin{cases} U_i, j = 1, 2, ..., n_i - 1; \\ V_i, j = n_i. \end{cases}$$

The weights of U_i and V_i of the n_i observations satisfy the following two conditions

$$(n_i - 1)U_i + V_i = 1 \tag{2}$$

$$(n_i - 1)U_i^2 + V_i^2 = \frac{z^*}{S_i^2} \tag{3}$$

where z^* is the maximum of $\left\{\frac{S_1^2}{n_1}, ..., \frac{S_k^2}{n_k}\right\}$. The advantage of Equation (2) implies $\sum_{j=1}^{n_i-1} U_i + V_i = \sum_{j=1}^{n_i} \omega_{ij} = 1$ that makes the weighted sample mean, $\tilde{X}_i$ in (1), an unbiased estimate of the population mean, μ_i, and Equation (3) implies $\sum_{j=1}^{n_i} \omega_{ij}^2 S_i^2 = z^*$ to eliminate the influence of the unknown and unequal variances. Based on the Equations (2) and (3), the solutions of the weights of U_i and V_i are derived as

$$U_i = \frac{1}{n_i} + \frac{1}{n_i}\sqrt{\frac{1}{n_i - 1}\left[n_i z^*/S_i^2 - 1\right]} \tag{4}$$

and

$$V_i = \frac{1}{n_i} - \frac{1}{n_i}\sqrt{(n_i - 1)\left[n_i z^*/S_i^2 - 1\right]}. \tag{5}$$

Therefore, the sampling distribution of weighted sample mean, $\tilde{X}_i$ $(i = 1, ..., k)$, is denoted by

$$\tilde{X}_i \sim N\left(\mu_i, \sum_{j=1}^{n_i} \omega_{ij}^2 \sigma_i^2\right), i = 1, ..., k.$$

Given the sample variance, $S_i^2, i = 1, ..., k,$ the sampling distribution of the test statistic

$$T_i = \frac{\tilde{X}_i - \mu_i}{\sqrt{\sum_{j=1}^{n_i} \omega_{ij}^2 S_i^2}} = \frac{\tilde{X}_i - \mu_i}{\sqrt{z^*}}, \ i = 1, ..., k, \tag{6}$$

is a conditional normal distribution with mean zero and variance σ_i^2/S_i^2, $i = 1, ..., k,$ and is also an unconditional Student's t distribution with degrees of freedom $n_i - 2$ (Wen and Chen (1994); Chen and Chen (1998)). Therefore, $T_i (i = 1, ..., k)$ is not affected by the unknown population variance $\sigma_i^2 (i = 1, ..., k)$ and $T_1, ..., T_k$ are mutually independent by Wen and Chen (1994).

In the following section one-sided and two-sided intervals I for all mean μ_i with the control mean μ_1 are proposed

$$\inf_{\Omega} P(\mu_i - \mu_1 \in I, i = 2, ..., k) \geq P^* \tag{7}$$

where $P^* \in \left(\frac{1}{k-1}, 1\right)$ and $\Omega = \left\{\underset{\sim}{\mu} = (\mu_1, ..., \mu_k)', \underset{\sim}{\sigma}^2 = (\sigma_1^2, ..., \sigma_k^2)'\right\}$.

3. Simultaneous confidence intervals

3.1. One-sided confidence interval

In this subsection, simultaneous lower and upper confidence intervals for $\mu_i - \mu_1, i = 2, ..., k$ are considered.

Let

$$I_L = \left(\tilde{X}_i - \tilde{X}_1 - a_i\sqrt{z^*}, \infty\right) \tag{8}$$

be a lower confidence intervals for $\mu_i - \mu_1$, where a_i's are constants chosen to satisfy the requirement of probability (7).

Theorem 1. *The constants a_i's are the solution of*

$$\int_{-\infty}^{\infty} \prod_{i=2}^{k} \left[F_{v_i}(t_1 + a_i)\right] f_{v_1}(t_1)\, dt_1 = P^*, \tag{9}$$

where $F_{v_i}(\cdot)$ and $f_{v_i}(\cdot)$ are the distribution and density functions respectively of a student's t random variables with $v_i (= n_i - 2), i = 1, ..., k,$ degrees of freedom.

Proof.

$$
\begin{aligned}
&P(\mu_i - \mu_1 \in I_L, i = 2, ..., k)\\
&= P\big(\mu_i - \mu_1 \geq \tilde{X}_i - \tilde{X}_1 - a_i\sqrt{z^*}, i = 2, ..., k\big)\\
&= P\big(\tilde{X}_i - \mu_i \leq \tilde{X}_1 - \mu_1 + a_i\sqrt{z^*}, i = 2, ..., k\big)\\
&= P\left(\frac{\tilde{X}_i - \mu_i}{\sqrt{z^*}} \leq \frac{\tilde{X}_1 - \mu_1}{\sqrt{z^*}} + a_i, i = 2, ..., k\right)\\
&= P(T_i \leq T_1 + a_i, i = 2, ..., k) \text{ by Equation (6)}\\
&= P(T_i \leq t_1 + a_i, i = 2, ..., k | T_1 = t_1)P(t_1)\\
&= P(T_i \leq t_1 + a_i, i = 2, ..., k)P(t_1) \ (\because T_1 \perp T_i)\\
&= \prod_{i=2}^{k} P(T_i \leq t_1 + a_i)P(t_1) \ (\because T_2, ..., T_k \text{ are independent})\\
&= \int_{-\infty}^{\infty} \prod_{i=2}^{k} [F_{v_i}(t_1 + a_i)] f_{v_1}(t_1) dt_1
\end{aligned}
$$

In order to find a simultaneous confidence interval, a_i's need to satisfy

$$\int_{-\infty}^{\infty} \prod_{i=2}^{k} \big[F_{v_i}(t_1 + a_i)\big] f_{v_1}(t_1)\, dt_1 = P^*.$$

Similarly, the upper confidence interval for $\mu_i - \mu_1, i = 2, ..., k,$ let

$$I_U = \big(-\infty, \tilde{X}_i - \tilde{X}_1 + b_i\sqrt{z^*}\big).$$

Theorem 2. *The constants b_i's are the solution of*

$$\int_{-\infty}^{\infty} \prod_{i=2}^{k} \big[1 - F_{v_i}(t_1 - b_i)\big] f_{v_1}(t_1)\, dt_1 = P^*. \tag{10}$$

Proof. This proof is similar to that of Theorem 1. Note that the property of symmetry for t distribution, $1 - F_{v_i}(t_1 - b_i) = F_{v_i}(t_1 + b_i)$, Therefore, Equation (10) becomes

$$\int_{-\infty}^{\infty} \prod_{i=2}^{k} \big[F_{v_i}(t_1 + b_i)\big] f_{v_1}(t_1)\, dt_1 = P^*. \tag{11}$$

3.2. Two-sided confidence interval

Let

$$I_s = \big(\tilde{X}_i - \tilde{X}_1 - a_i\sqrt{z^*}, \tilde{X}_i - \tilde{X}_1 + b_i\sqrt{z^*}\big)$$

be a two-sided confidence intervals for $\mu_i - \mu_1, i = 2, ..., k,$ where the constants a_i's and b_i's are selected to satisfy the requirement of probability (7).

Theorem 3. *The constants a_i's and b_i's are the solution of*

$$\int_{-\infty}^{\infty} \prod_{i=2}^{k} \big[F_{v_i}(t_1 + a_i) - F_{v_i}(t_1 - b_i)\big] f_{v_1}(t_1)\, dt_1 = P^*. \tag{12}$$

Proof. This proof is similar to that of Theorem 1. From Equation (12), if we choose all a_i's to go infinite, it reduces the equation down to Equation (10). Similarly, if we choose all b_i's to go infinite, it reduces the equation down to Equation (9).

3.3. Optimal confidence intervals

In practice, there are multiple choices of critical values a_i and $b_i (i = 1, ..., k)$ for the different purposes. In this section, we present an approach to find the optimal confidence intervals for the further analysis. Guenther (1969) and Casella and Berger (2002) have both shown that the shortest length of confidence interval occurs when the critical values $a = -b$ in both known and unknown variance in normal model. That is, the shortest confidence interval is the same as equal-tails interval if the distribution is symmetric. Wen and Chen (1994) have shown that the optimal solution for the Equation (12) with the maximized coverage probability is when all critical values of t are the same under the equal sample sizes. Therefore, the optimal confidence interval based on the symmetrical normal and Student t distributions is the shortest.

When sample sizes are unequal, Wen and Chen (1994) have shown that $T_1, ..., T_k$ are mutually independent. Each confidence interval has the shortest expected length of confidence of t distribution by Casella and Berger (2002). Therefore, we rewrite the Equations (9) (10) and (12) as follows.

$$\int_{-\infty}^{\infty} \prod_{i=2}^{k} \left[F_{v_i}(t_1 + a)\right] f_{v_1}(t_1)\, dt_1 = P^*, \tag{13}$$

$$\int_{-\infty}^{\infty} \prod_{i=2}^{k} \left[1 - F_{v_i}(t_1 - b)\right] f_{v_1}(t_1)\, dt_1 = P^*, \tag{14}$$

and

$$\int_{-\infty}^{\infty} \prod_{i=2}^{k} \left[F_{v_i}(t_1 + h) - F_{v_i}(t_1 - h)\right] f_{v_1}(t_1)\, dt_1 = P^*. \tag{15}$$

It is noted that the critical values a and b for Equations (13) and (14) are the same. When sample sizes are equal $v_1 = ... = v_k = v$, Equations (13) and (15) become

$$\int_{-\infty}^{\infty} \left[F_v(t + a)\right]^{k-1} f_v(t)\, dt = P^*, \tag{16}$$

and

$$\int_{-\infty}^{\infty} \left[F_v(t + h) - F_v(t - h)\right]^{k-1} f_v(t)\, dt = P^*. \tag{17}$$

The critical values tables of a or h for Equations (16) and (17) were given by Dudewicz, Ramberg, and Chen (1975), Dudewicz and Dalal (1983), and FORTRAN program of Wen and Chen (1994). When sample sizes are not equal, it becomes impossible to list critical value tables for all situations. In order to get the critical value of Equations (13) and (15) user-friendly, we built an interface by using *R* shiny. The interface provides the one-sided and two-sided critical values after inputting parameters and the URL of the interface can be accessed at https://weiming3524.shinyapps.io/mc_control/.

Table 1. Simulated FWEs for the single-stage sampling procedure (SS) and the plug-in procedure (PI), respectively, with unequal sample sizes and three variance scenarios.

		$k=3$			$k=4$			$k=5$		
		homosc.	heterosc. I	heterosc. II	homosc.	heterosc. I	heterosc. II	homosc.	heterosc. I	heterosc. II
n_1=4	SS	0.078	0.077	0.079	0.080	0.081	0.081	0.085	0.084	0.084
	PI	0.047	0.050	0.051	0.050	0.052	0.047	0.049	0.051	0.046
n_1=8	SS	0.059	0.056	0.059	0.056	0.058	0.056	0.056	0.059	0.057
	PI	0.051	0.050	0.051	0.049	0.052	0.047	0.046	0.051	0.046
n_1=12	SS	0.051	0.054	0.053	0.052	0.055	0.051	0.054	0.056	0.055
	PI	0.049	0.048	0.046	0.053	0.050	0.047	0.052	0.051	0.053
n_1=16	SS	0.055	0.051	0.054	0.050	0.050	0.052	0.055	0.056	0.054
	PI	0.050	0.051	0.052	0.051	0.047	0.052	0.050	0.047	0.046
n_1=20	SS	0.051	0.046	0.052	0.051	0.047	0.051	0.052	0.053	0.050
	PI	0.048	0.050	0.049	0.050	0.047	0.048	0.049	0.050	0.050

4. α-Simulations

The single-stage sampling procedure provides an exact distribution for its statistics employed in the multiple comparison with a control under heteroscedasticity, but simulation research is also necessary for validating their quality. The quality of a procedure varies depending on the number of treatments, variance scenarios and sample size allocations. According to Hasler (2016), we chose $k = 3, 4, 5$ treatments respectively, in order to be compared by the one-sided multiple comparison with a control where the first treatment was regarded as control. The familywise error type I (FWE) was simulated and the nominal level was $\alpha = 0.05$. Following Hasler's (2016) scenarios, the expected value of all treatments was identical, $\mu_i = 100(i = 1, ..., k)$. Five sample size allocations were chosen, i.e., $n_1 = 4, 8, 12, 16, 20,$ and also, unbalanced sample sizes per allocations, i.e., $n_i = n_{i-1} + 2(i = 2, ..., k)$. Note that, as per theory, an initial sample of size greater than or equal to 3 ($n_1 \geq 3$) will work, but for a better result, 10 or more would be more suitable according to Bishop and Dudewicz (1978). Three different variance scenarios were considered as follows:

- the homoscedastic case (homosc.): $\sigma_i = 10, i = 1, ..., k$,
- the treatment with the smallest standard deviation has the smallest sample size (heterosc. I): $\sigma_i = 10 + 40(i-1)/(k-1), i = 1, ..., k$,
- the treatment with the highest standard deviation has the smallest sample size (heterosc. II): $\sigma_i = 50 - 40(i-1)/(k-1), i = 1, ..., k$.

All simulation results were obtained by running the simulation 10,000 times using the *R* simulation program. Here, SS indicates the single-stage sampling procedure. The plug-in (PI) procedure was provided in the Hasler and Hothorn (2008). They applied the idea from Games and Howell (1976) to deal with multiple comparison procedure by using the multiple contrast test (MCT) techniques. The degree of freedom was specific group comparison based on Welch-Satterthwaite equation and group variance estimators were plugged into the correlation matrix among contrast tests. The primary advantage of the PI is that it can maintain the familywise error rate at level α according to Hasler's simulation study and the PI procedure is available in the package *SimComp* (Hasler (2014)) of the statistical software *R*.

Table 1 shows the simulation results with unequal sample sizes correspond to the different variance scenarios. Overall, the SS procedure is always more liberal than PI procedure for the smaller sample size allocations, and independent of the variance scenarios. However, SS produces equal comparison wise error type I, and it keeps the FWE under the sample size greater than 8. In contrast, the PI procedure is slightly more conservative sometimes than the SS procedure

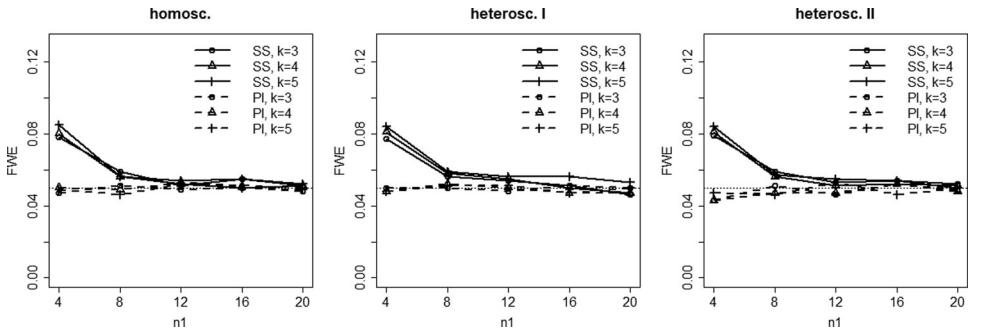

Figure 1. Simulated FWEs for $k = 3$, 4, 5 groups, different sizes of sample and variance scenarios respectively.

when the samples size is large. As the number of treatments increases, it may have a slight influence on SS. The more the number of treatments, the better the chance to form extreme variation, although it does not happen easily.

Figure 1 visually shows the simulation results of which sub-figures correspond to the different variance scenarios. We provide an interface by *R* shiny to obtain the critical values for the SS procedure for the problems caused by multiple comparisons with a control problem.

5. A numerical example

The mutagenicity data set shown in Table 2 is the same as in Hasler (2016). They are available in the *R* package *mratios* (Djira et al. (2020)), and they were taken from Adler and Kliesch (1990). Mutagenicity assay for 4 doses of a compound (hydroquinone) against a vehicle control. Hydroquinone was applied in doses of 30, 50, 70, 100 mg/kg. After 24 h, counts of micronuclei are used in polychromatic erythrocytes as a measure for the potency to cause chromosome damage. The objective is to test if chromosome damage can be caused by the underlying substance. The statistic summary for the number of micronuclei per male mice is detailed in Table 2.

Although the purpose of the trial was different, the data are employed as an example here for a multiple comparison to a control problem. We used the Shapiro-Wilk test to verify the normality of the data producing a p-value of at least 0.05 except 0.045 of the 100 mg/kg group. Although the data slightly departures from normality, a statistical models with a Gaussian error structure is generally robust to violations of the normality assumption. As long as the distributions are not extremely different form a normal distribution, the level of significance of a test is usually not great affected (Knief and Forstmeier (2021)). Dudewicz and Van Der Meulen (1983) also shown robustness results that make it reasonable to use these procedures for general non-normal distribution. To test the homogeneity of variances, we conducted a Levene's test which produced a p-value of 0.028. Hence, the data are with unequal variances.

Here, we applied the single-stage sampling procedure to this example. We calculated the initial sample mean $\bar{X}_i, i = 1, ..., 5,$ and the sample standard deviation S_i with the first $n_i - 1$ observations. Then, we used Equation (1) to calculate the weighted sample mean $\tilde{X}_i, i = 1, ..., 5$ by the weights of U_i and V_i from Equations (4) and (5) for a control and treatment i. The summary statistics are shown in Table 3.

For Table 3, the simultaneous lower confidence intervals of the SS procedure are as follows.

$$I_L = \left(\tilde{X}_i - \tilde{X}_1 - a\sqrt{z^*}, \infty\right), i = 30, 50, 75, 100\,\text{mg/kg}$$

where $z^* = \max\left\{\frac{0.753^2}{7}, \frac{1^2}{5}, \frac{1.258^2}{5}, \frac{3.742^2}{5}, \frac{4.272^2}{5}\right\} = 3.650$ and the critical value $a = 4.834$ from *R* shiny interface is shown in the Appendix. The lower confidence limits are as follows:

Table 2. Summary statistics of the mutagenicity data set.

Treatment group	Mean	Standard deviation	Sample size
Vehicle control	2.571	1.272	7
30 mg/kg	3.800	1.095	5
50 mg/kg	6.200	1.483	5
75 mg/kg	14.000	3.937	5
100 mg/kg	20.000	4.062	5

Table 3. Summary statistics for the mutagenicity data applying the SS procedure and lower confidence intervals of the SS procedure and the PI procedure.

						Lower conf. intervals	
Treatment group	$\bar{X}_i$	S_i	U_i	V_i	$\tilde{X}_i$	SS	PI
Vehicle control	2.167	0.753	0.530	−2.181	−4.012		
30 mg/kg	3.500	1.000	0.615	−1.461	1.308	$(-3.915, \infty)$	$(-0.557, \infty)$
50 mg/kg	5.750	1.258	0.524	−1.098	3.280	$(-1.943, \infty)$	$(1.391, \infty)^*$
75 mg/kg	13.000	3.742	0.255	−0.020	12.898	$(7.675, \infty)^*$	$(5.495, \infty)^*$
100 mg/kg	19.250	4.272	0.200	0.200	20.000	$(14.776, \infty)^*$	$(11.319, \infty)^*$

$$IL_1 = 1.308 - (-4.012) - 4.834 \times \sqrt{3.65} = -3.915$$
$$IL_2 = 3.280 - (-4.012) - 4.834 \times \sqrt{3.65} = -1.943$$
$$IL_3 = 12.898 - (-4.012) - 4.834 \times \sqrt{3.65} = 7.675$$
$$IL_4 = 20.000 - (-4.012) - 4.834 \times \sqrt{3.65} = 14.776.$$

According to the results of the SS procedure, it is concluded that the two lower doses 30 and 50 mg/kg are non-significant to the control. The two higher doses 75 and 100 mg/kg reveal significance higher than the control. In contrast to the PI procedure, dose 50 mg/kg reveals significance higher than the control. Although the inference of the SS procedure that takes the largest variance to eliminate the affection of unknown population is more conservative, the SS procedure provides an exact distribution for handling the multiple comparison with a control under heteroscedasticity problems.

6. Conclusion

In this paper, we proposed a single-stage sampling procedure to multiple comparison with a control without the assumption of equal variance. The single-stage sampling procedure also can provide an exact distribution for its statistics employed in the multiple comparison with a control under heteroscedasticity. The single-stage sampling procedure is advantageous since it is data-analysis oriented, and works for any given data set whether the variances are equal or unequal.

According to the FWE simulation results, if sample sizes are small, SS is more liberal than PI considering all kinds of situations. However, SS maintains the FWE for the sample size greater than 8. It should be noted that Bishop and Dudewicz (1978) suggested that the sample size of the first stage in the two-stage sampling procedure at least 10 giving better results in practical studies. Thus, we suggested that the initial sample of size in the single-stage sampling procedure can be chosen at least 8.

We provided a real numerical example to illustrate how to find the lower confidence intervals by applying the procedure. In order for people to use the procedures easily, a web application using *R* Shiny was developed to simulate the one-sided and two-sided critical values for single-stage multiple comparisons with a control. The URL of the interface can be accessed at https://weiming3524.shinyapps.io/mc_control/.

Conflict of interest statement

The authors declare no conflicts of interest.

Funding

Wen's and Wang's research was sponsored by the Ministry of Science and Technology, Taiwan (MOST 109-2118-M-006-004-MY2).

Data availability statement

All authors make sure that all data support their published claims and comply with field standards.

References

Adler, I, and U. Kliesch. 1990. Comparison of single and multiple treatment regimens in the mouse bone marrow micronucleus assay for hydroquinone (HQ) and cyclophosphamide (CP). *Mutation Research* 234 (3–4):115–23. doi:10.1016/0165-1161(90)90002-6.

Bishop, T. A. 1976. Heteroscedastic ANOVA, MANOVA, and multiple-comparisons. Unpublished Ph.D. thesis, The Ohio State University.

Bishop, T. A, and E. J. Dudewicz. 1978. Exact analysis of variance with unequal variances: Test procedures and tables. *Technometrics* 20 (4):419–30. doi:10.1080/00401706.1978.10489696.

Casella, G, and R. L. Berger. 2002. *Statistical inference. Section 9.3: Methods of evaluating interval estimation*. 2nd ed., 443. CA: Duxbury Press.

Chen, H. J, and K. Lam. 1989. Single-stage interval estimation of the largest normal mean under heteroscedasticity. *Communications in Statistics - Theory and Methods* 18 (10):3703–18. doi:10.1080/03610928908830118.

Chen, H. J., M. J. Wen, and A. M. Wang. 2009. On testing the bioequivalence of several treatments using the measure of distance. *Statistics* 43 (5):513–30. doi:10.1080/02331880802497447.

Chen, H. J., M. J. Wen, and C. J. Chuang. 2011. On testing the equivalence of treatments using the measure of range. *Computational Statistics & Data Analysis* 55 (1):603–14. doi:10.1016/j.csda.2010.06.002.

Chen, S. Y, and H. J. Chen. 1998. Single-stage analysis of variance under heteroscedasticity. *Communications in Statistics - Simulation and Computation* 27 (3):641–66. doi:10.1080/03610919808813501.

Chen, S. Y. 2001. One-stage and two-stage statistical inference under heteroscedasticity. *Communications in Statistics: Simulation and Computation* 30 (4):991–1009. doi:10.1081/SAC-100107792.

Chen, S. Y., H. J. Chen, and H. F. Chang. 2004. A one-stage procedure for testing homogeneity of means against an ordered alternative under unequal variances. *Communications in Statistics - Simulation and Computation* 33 (1):49–67. doi:10.1081/SAC-120028433.

Djira, G. M. Hasler, D. Gerhard, and F. Schaarschmidt. 2020. mratios: Ratios of coefficients in the general linear model. https://carn.r-project.org/package=mratios. *R package version,* 1.4.2.

Dudewicz, E. J, and E. C. Van Der Meulen. 1983. Entropy-based statistical inference, II: Selection-of-the-best/complete ranking for continuous distributions on (0, 1), with applications to random number generators. *Statistics and Risk Modeling* 1 (2):131–46.

Dudewicz, E. J, and S. R. Dalal. 1983. Multiple-comparisons with a control when variances are unknown and unequal. *American Journal of Mathematical and Management Sciences* 3 (4):275–95. doi:10.1080/01966324.1983.10737128.

Dudewicz, E. J., J. S. Ramberg, and H. J. Chen. 1975. New tables for multiple comparisons with a control (unknown variances). *Biometrische Zeitschrift* 17 (1):13–26. doi:10.1002/bimj.19750170103.

Dunnett, C. W. 1955. A multiple comparison procedure for comparing several treatments with a control. *Journal of the American Statistical Association* 50 (272):1096–121. doi:10.1080/01621459.1955.10501294.

Games, P. A, and J. F. Howell. 1976. Pairwise multiple comparison procedures with unequal n's and/or variances: A Monte Carlo study. *Journal of Educational Statistics* 1 (2):113–25.

Guenther, W. C. 1969. Shortest confidence intervals. *The American Statistician* 23 (1):22–5.

Hasler, M, and L. A. Hothorn. 2008. Multiple contrast tests in the presence of heteroscedasticity. *Biometrical Journal. Biometrische Zeitschrift* 50 (5):793–800. doi:10.1002/bimj.200710466.

Hasler, M. 2014. SimComp: Simultaneous comparisons for multiple endpoints. http://cran.r-project.org/package=SimComp. *R package version 3.3*.

Hasler, M. 2016. Heteroscedasticity: Multiple degrees of freedom vs. sandwich estimation. *Statistical Papers* 57 (1): 55–68. doi:10.1007/s00362-014-0640-4.

Knief, U, and W. Forstmeier. 2021. Violating the normality assumption may be the lesser of two evils. *Behavior Research Methods* 53 doi: (6):2576–90. doi:10.3758/s13428-021-01587-5.

Lam, K. 1992. Subset selection of normal populations under heteroscedasticity. In *Proceedings and Discussions of the Second International Conference on Inference Procedures Associated with Statistal Ranking and Selection on The Frontiers of Modern Statistical Inference Procedures, II*, Sydney, Australia.

Li, H, and W. Ning. 2012. Multiple comparisons with a control under heteroscedasticity. *Journal of Applied Statistics* 39 (10):2275–83. doi:10.1080/02664763.2012.706269.

Miller, R. G. Jr.1981. *Simultaneous statistical inference*. 2nd ed. NY: Springer.

Paulson, E. 1952. On the comparison of several experimental categories with a control. *The Annals of Mathematical Statistics* 23 (2):239–46. doi:10.1214/aoms/1177729440.

Stein, C. 1945. A two-sample test for a linear hypothesis whose power is independent of the variance. *The Annals of Mathematical Statistics* 16 (3):243–58. doi:10.1214/aoms/1177731088.

Tamhane, A. C. 1977. Multiple comparisons in model I one-way ANOVA with unequal variances. *Communications in Statistics - Theory and Methods* 6 (1):15–32. doi:10.1080/03610927708827466.

Wen, M. J, and H. J. Chen. 1994. Single-stage multiple comparison procedures under heteroscedasticity. *American Journal of Mathematical and Management Sciences* 14 (1–2):1–48. doi:10.1080/01966324.1994.10737368.

Wen, M. J., L. C. Huang, and J. Zhong. 2017. Single-stage sampling procedure of the t best populations under heteroscedasticity. *Communications in Statistics - Theory and Methods* 46 (18):9265–73. doi:10.1080/03610926.2016.1206935.

Wen, M. J., S. Y. Chen, and H. J. Chen. 2007. On testing a subset of regression parameters under heteroskedasticity. *Computational Statistics & Data Analysis* 51 (12):5958–76. doi:10.1016/j.csda.2006.11.018.

Westfall, P. H. 1997. Multiple testing of general contrasts using logical constraints and correlations. *Journal of the American Statistical Association* 92 (437):299–306. doi:10.1080/01621459.1997.10473627.