P-ISSN 2586-9000
E-ISSN 2586-9027
Homepage : https://tci-thaijo.org/index.php/SciTechAsia
Science & Technology Asia
Vol.26 No.2 April - June 2021
Page: [16-24]

Original research article

Two Stage Approach Based on Welch Statistic for Multiple Comparisons of *k* Binomial Proportions for Small Sample

Chanokgan Sahatsathatsana, Nirun Nitisuk, Noppakun Tongmual*

Department of Science and Mathematics, Faculty of Sciences and Health Technology, Kalasin University, Kalasin 46000, Thailand

Received 22 February 2020; Received in revised form 29 December 2020
Accepted 12 January 2021; Available online 25 June 2021

ABSTRACT

Multiple comparisons of k independent binomial proportions ($k > 2$) are studied when the proportions p_i, $i = 1, 2, 3, ..., k$ are close to zero. Several tests perform poorly in terms of the pairwise error rate (PWER) and the familywise error rate (FWER) when the proportions p_i, $i = 1, 2, 3, ..., k$ are close to zero. This problem is an issue that seems to have been overlooked. Even though several tests have been proposed, they cannot perform well in terms of PWER and FWER. From the above problems, we proposed a procedure of multiple comparisons for examining the difference between k independent binomial proportions which is the proposed two-stage approach based on Welch statistic. For comparing the performance of test statistics for multiple comparisons, the proposed two stage approach is compared with the two-stage approach under PWER, FWER and the estimated pairwise powers. Our results were evaluated by using Monte Carlo simulation. The results indicated that the performance of the proposed two stage approach can protect PWER and FWER better than the two-stage approach. In cases of the estimated pairwise power, the proposed two stage approach and the two-stage approach have similar estimated pairwise power. Our study suggests the proposed approach for multiple comparisons because the proposed two-stage approach can protect PWER and FWER, and the estimated pairwise power is quite well.

Keywords: Pairwise error rate; Familywise error rate; Estimated pairwise power

1. Introduction

Multiple comparisons of k independent binomial proportions are studied in this research. Multiple comparisons are performed for comparing the differences between k independent binomial proportions. The problem of multiple comparisons in binomial distribution is

***Corresponding author:** Nop_stat@hotmail.com **doi:** 10.14456/scitechasia.2021.22

when the proportions p_i, $i = 1, 2, 3, ..., k$ are close to zero. This problem is found in the literature. For instance, Ana (2008) studied twenty methods for two-sided confidence intervals for the proportion parameter p and the results indicated poor coverage probability for p close to zero. Lawrence et al. [1] showed that the coverage properties of the Wald interval presented poorly for p close to zero. Brown et al. [2] reported that the actual coverage probability of the standard interval is poor for p near zero.

From the above problem, Kane [3] presented the two-stage approach for multiple comparisons for comparing the difference between k independent binomial proportions. For the procedure of the two-stage approach, assume that $X_i \sim BIN(n_i, p_i)$. In the first step, an analysis of variance is conducted for examining the equality of the proportion parameters p_i, $i = 1, 2, 3, ..., k$. The testing hypotheses are

$$H_0 : p_1 = ... = p_k \text{ vs. } H_1 : p_i \neq p_j, \quad (1)$$

for some $i \neq j$.

The test statistic in Eq. (1) is

$$G = \sum_{i=1}^{k} \frac{n_i \left(\breve{p}_i - \breve{p}\right)^2}{1/4}, \quad (2)$$

where $\breve{p}_i = \sin^{-1}\sqrt{\frac{X_i + 3/8}{n_i + 3/4}}$, $i = 1, 2, 3, ..., k$

and $\breve{p} = \sin^{-1}\sqrt{\frac{\sum_{i=1}^{k} X_i + 3/8}{\sum_{i=1}^{k} n_i + 3/4}}$. The G test in Eq. (2) has been distributed as an asymptotic chi-square with k-1 degrees of freedom.

Next, the pairwise tests will be conducted when G test rejects H_0 in Eq. (1) with $G = \sum_{i=1}^{k} \frac{n_i \left(\breve{p}_i - \breve{p}\right)^2}{1/4} > \chi^2_{k-1}$. The testing hypotheses for the pairwise testing are

$$H_{0ij} : p_i = p_j \text{ vs. } H_{1ij} : p_i \neq p_j, \quad (3)$$

for some $i \neq j$.

The pairwise tests are I test, L test and M test which are used for the pairwise testing in Eq. (3).

The I test for pairwise testing in Eq. (3) is

$$I = \frac{\left|\breve{p}_i - \breve{p}_j\right|}{\sqrt{\frac{1}{4}\left(\frac{1}{n_i} + \frac{1}{n_j}\right)}}, \quad (4)$$

where the I test rejects the pairwise null hypothesis with

$$I = \frac{\left|\breve{p}_i - \breve{p}_j\right|}{\sqrt{\frac{1}{4}\left(\frac{1}{n_i} + \frac{1}{n_j}\right)}} > \frac{q_{a,k-1}}{\sqrt{2}},$$

where $q_{a,k-1}$ is the studentized-range value.

The L test is based on the Wald test suggested by Agresti and Caffo [4]. The L test is

$$L = \frac{\left|\tilde{p}_i - \tilde{p}_j\right|}{\sqrt{\frac{\tilde{p}_i\left(1 - \tilde{p}_i\right)}{n_i} + \frac{\tilde{p}_j\left(1 - \tilde{p}_j\right)}{n_j}}}, \quad (5)$$

where $\tilde{p}_i = \left(X_i + 1\right)/\left(n_i + 2\right)$. The L test rejects the pairwise null hypothesis when

$$L = \frac{\left|\tilde{p}_i - \tilde{p}_j\right|}{\sqrt{\frac{\tilde{p}_i\left(1 - \tilde{p}_i\right)}{n_i} + \frac{\tilde{p}_j\left(1 - \tilde{p}_j\right)}{n_j}}} > Z_{a/2}. \quad (6)$$

The M test for pairwise testing in Eq. (3) is

$$M = \frac{\left|\tilde{p}_i - \tilde{p}_j\right|}{\sqrt{\frac{\tilde{p}_i\left(1-\tilde{p}_i\right)}{n_i} + \frac{\tilde{p}_j\left(1-\tilde{p}_j\right)}{n_j}}}. \qquad (7)$$

The M test rejects the pairwise null hypothesis when

$$M = \frac{\left|\tilde{p}_i - \tilde{p}_j\right|}{\sqrt{\frac{\tilde{p}_i\left(1-\tilde{p}_i\right)}{n_i} + \frac{\tilde{p}_j\left(1-\tilde{p}_j\right)}{n_j}}} > \frac{q_{\alpha,k-1,\infty}}{\sqrt{2}}. \qquad (8)$$

The details of the M test can be found in Hayter [5].

For testing, hypotheses in Eq. (1) are used for comparing the equality of the proportion parameters p_i, $i = 1,2,3,...,k$ with an analysis of variance. An alternative test for an analysis of variance is Welch test statistics. The results of Krishnamoorthy [6] and Noppakun et al. [7] indicated that the Welch test performs quite well in preventing Type I errors even for small sample sizes. Welch test is

$$W = \frac{\chi^2_{k-1}/(k-1)}{1+(2(k-2)/(k^2-1))\sum_{i=1}^{k}(1/(n_i-1))(1-w_i/\sum_{i=1}^{k}w_i)^2}, \qquad (9)$$

where Welch has distributed as F distribution with degree of freedom $f_1 = k-1$ and

$$f_2 = \left[\frac{3}{k^2-1}\sum_{i=1}^{k}\frac{1}{n_i-1}\left(1-w_i\Big/\sum_{i=1}^{k}w_i\right)^2\right]^{-1}.$$

In this research, we proposed the two-stage approach based on the Welch test, and compared it with the two-stage approach for multiple comparisons to see the difference between k independent binomial proportions with proportion parameter p_i, $i = 1,2,3,...,k$ close to zero based on PWER, FWER and the estimated pairwise power which reflect the performance of the test statistic. The paper is organized as follows. The proposed two-stage approach based on the Welch test is described in Section 2. Section 3 shows a comparison of the performance of the two-stage approach and the proposed two-stage approach based on PWER, FWER and the estimated pairwise power. Section 4 contains results and discussion. Finally, Section 5 contains the conclusion.

2. The Proposed Two-Stage Approach

Consider the G test in Eq. (2) is $G = \sum_{i=1}^{k}\frac{n_i\left(\breve{p}_i - \breve{p}\right)^2}{1/4}$ which has chi-square distribution with k -1 degrees of freedom. Hence, we obtain the proposed test statistic based on the Welch test as

$$W_p = \frac{G/(k-1)}{1+(2(k-2)/(k^2-1))\sum_{i=1}^{k}(1/(n_i-1))(1-w_i^*/\sum_{i=1}^{k}w_i^*)^2} \qquad (10)$$

where W_p is distributed as F distribution with degree of freedom $f_1 = k-1$ and

$$f_2 = \left[\frac{3}{k^2-1}\sum_{i=1}^{k}\frac{1}{n_i-1}\left(1-w_i^*\Big/\sum_{i=1}^{k}w_i^*\right)^2\right]^{-1},$$

where $w_i^* = 4n_i$. The W_p test rejects H_0 in Eq. (1) when $W_P > F_{f_1,f_2}$. If W_P rejects H_0 in Eq. (1.1), we conduct the pairwise test as I test, L test and M test for hypotheses testing in Eq. (3). This process is the proposed two-stage approach based on Welch test

3. Comparison of the Performance

In this study, we perform a Monte Carlo simulation by using the R statistical package [8] consisting of 100,000 iterations to compute PWER, FWER and the estimated pairwise power of the two-stage approach and the proposed two stage approach. We consider number of groups $k = 3$ and 5 populations at the significance level of 0.05 with equal sample sizes of $\mathbf{n} = (25, 25, ..., 25)$. For the proportion p_i, $i = 1, 2, 3, ..., k$ are in a wide range $0.02 \le p_i \le 0.5$. In the simulation, we generate k random variates at a time from $BIN(n_i, p_i)$. The two-stage approach and the proposed two-stage approach are conducted in the first step where $G > \chi^2_{\alpha,k-1}$ and $W_P > F_{f_1,f_2}$ rejects H_0 in Eq. (1), respectively. Next, I test, L test and M test for pairwise tests are conducted when G test and W_P test reject H_0 in Eq. (1). We repeat 100,000 times and calculate the proportion of times of rejecting H_{0ij} in Eq. (3).

In Table 1 and Table 2, I test, M test and W test are the two-stage approach based on G test, and IW test, LW test and MW test are the proposed two-stage approach based on W_P test. The columns " $p_i = p_j$ " and " $p_i \ne p_j$ " are PWER and the estimated pairwise power, respectively, and FWER is presented in the column " FWER" . The columns " p_i, p_j " show the pairwise power. The column "Global" shows the estimated probability of G test and W_P test for rejecting H_0 in Eq. (1).

4. Results and Discussion

4.1 Results

Table 1 shows the estimated value of PWER, FWER and the estimated pairwise power of six tests as I test, M test, L test, IW test, LW test and MW test by Monte Carlo simulation for $k = 3$.

For configuration 1, the estimated values of PWER and FWER of six tests are exceedingly small in terms of the estimated value of PWER with 0.0000-0. 0201 under the proportion vectors $\mathbf{p} = (0.02, 0.02, 0.02)$, $\mathbf{p}$ = (0.05, 0.05, 0.05) and $\mathbf{p}$ = (0.1, 0.1, 0.1). For the estimated value of FWER, six tests show the estimated value of FWER ranging from 0.0001-0.0514. Again, it was observed that FWER of the I test exceeds the nominal level of 0.05.

For configuration 2, the estimated values of PWER and FWER, the results of simulation show that the estimated value of PWER and the estimated value of FWER of six tests are lower than the nominal level of 0.05 with the estimated value of PWER and the estimated value of FWER ranging from 0.0001-0.0468 under the proportion vectors $\mathbf{p}$ = $(0.02, 0.02, 0.27)$, $\mathbf{p}$ = $(0.05, 0.05, 0.30)$ and $\mathbf{p}$ = (0.1, 0.1, 0.35). For the proportion vector $\mathbf{p}$ = $(0.05, 0.05, 0.35)$ it was found that PWER and FWER of the I test and the IW test are nearly the nominal level of 0.05. Considering the estimated pairwise power of six tests, the simulation studies show that I test and IW test higher than M, L, LW and MW tests.

For configuration 3, the estimated values of PWER and FWER, six tests have the estimated value of PWER ranging from 0.0415-0.0519 and the estimated value of FWER ranging from 0.0415-0.0519 under the proportion vectors $\mathbf{p}$ = (0.02, 0.27, 0.27), $\mathbf{p}$ = (0.05, 0.3, 0.3) and $\mathbf{p}$ = (0.1, 0.35, 0.35). Again, it was observed that PWER and FWER of I test and IW test exceed the nominal level of 0. 05 but IW test is nearly the nominal level of 0.05. For PWER and FWER of L, M, LW and MW tests can produce the nominal level of 0.05. Considering the estimated pairwise power of I, L, M, IW, LW and MW tests, the results indicate that I test and IW test have higher than L, M, LW and MW tests.

For Configuration 4, the estimated values of PWER and FWER of six tests cannot protect the estimated value of PWER and the estimated value of FWER under the proportion vectors $\mathbf{p}$ = (0.02, 0.06, 0.1), $\mathbf{p}$ = (0.05, 0.15, 0.25) and $\mathbf{p}$ $= (0.10, 0.30, 0.50)$. In the estimated pairwise power of six tests the results indicated that I test and IW test are higher than L test, M test, LW test and MW test under the proportion vectors $\mathbf{p}=(0.02,0.06,0.10)$, $\mathbf{p}=(0.05,0.15,0.25)$ and $\mathbf{p}=(0.10,0.30,0.50)$.

Table 1. PWER, FWER, estimated pairwise power and Estimated probabilities for $k=3$.

Configuration 1						Configuration 2					
$\mathbf{p}=(0.02,0.02,0.02)$						$\mathbf{p}=(0.02,0.02,0.27)$					
Test	$p_i=p_j$	$p_i\neq p_j$	FWER	Global G	Global W_P	Test	$p_i=p_j$	$p_i\neq p_j$	FWER	Global G	Global W_P
I	0.0013	-	0.0020	0.0020	-	I	0.0020	0.7646	0.0020	0.8268	-
L	0.0000	-	0.0001	0.0020	-	L	0.0001	0.6991	0.0001	0.8268	-
M	0.0000	-	0.0001	0.0020	-	M	0.0001	0.6991	0.0001	0.8268	-
IW	0.0013	-	0.0020	-	0.0020	IW	0.0018	0.7599	0.0018	-	0.8167
LW	0.0000	-	0.0001	-	0.0020	LW	0.0001	0.6945	0.0001	-	0.8167
MW	0.0000	-	0.0001	-	0.0020	MW	0.0001	0.6945	0.0001	-	0.8167
$\mathbf{p}=(0.05,0.05,0.05)$						$\mathbf{p}=(0.05,0.05,0.30)$					
I	0.0091	-	0.0179	0.0179	-	I	0.0198	0.6421	0.0198	0.7299	-
L	0.0046	-	0.0116	0.0179	-	L	0.0033	0.6109	0.0033	0.7299	-
M	0.0046	-	0.0116	0.0179	-	M	0.0033	0.6109	0.0033	0.7299	-
IW	0.0088	-	0.0173	-	0.0173	IW	0.0181	0.6180	0.0181	-	0.6892
LW	0.0045	-	0.0115	-	0.0173	LW	0.0033	0.5871	0.0033	-	0.6892
MW	0.0045	-	0.0115	-	0.0173	MW	0.0033	0.5871	0.0033	-	0.6892
$\mathbf{p}=(0.10,0.10,0.10)$						$\mathbf{p}=(0.10,0.10,0.35)$					
I	0.0238	-	0.0514	0.0514	-	I	0.0468	0.5172	0.0468	0.6039	-
L	0.0175	-	0.0456	0.0514	-	L	0.0213	0.4948	0.0213	0.6039	-
M	0.0175	-	0.0456	0.0514	-	M	0.0213	0.4948	0.0213	0.6039	-
IW	0.0201	-	0.0448	-	0.0448	IW	0.0463	0.5002	0.0463	-	0.5654
LW	0.0164	-	0.0430	-	0.0448	LW	0.0213	0.4696	0.0213	-	0.5654
MW	0.0164	-	0.0430	-	0.0448	MW	0.0213	0.4696	0.0213	-	0.5654
Configuration 3						Configuration 4					
Test	$p_i=p_j$	$p_i\neq p_j$	FWER	Global						Global	
$\mathbf{p}=(0.02,0.27,0.27)$				G	W_P	$\mathbf{p}=(0.02,0.06,0.10)$				G	W_P
I	0.0514	0.7520	0.0514	0.8261	-	I	0.0162	0.0478	0.0824	0.0970	-
L	0.0423	0.6694	0.0423	0.8261	-	L	0.0088	0.0239	0.0584	0.0970	-
M	0.0423	0.6694	0.0423	0.8261	-	M	0.0088	0.0239	0.0584	0.0970	-
IW	0.0506	0.7431	0.0506	-	0.8135	IW	0.0127	0.0472	0.0779	-	0.0920
LW	0.0415	0.6658	0.0415	-	0.8135	LW	0.0088	0.0233	0.0574	-	0.0920
MW	0.0415	0.6658	0.0415	-	0.8135	MW	0.0088	0.0233	0.0574	-	0.0920
$\mathbf{p}=(0.05,0.30,0.30)$						$\mathbf{p}=(0.05,0.15,0.25)$					
I	0.0519	0.6247	0.0519	0.7087	-	I	0.1764	0.1330	0.4142	0.4352	-
L	0.0455	0.5737	0.0455	0.7087	-	L	0.1090	0.1138	0.3876	0.4352	-
M	0.0455	0.5737	0.0455	0.7087	-	M	0.1090	0.1138	0.3876	0.4352	-
IW	0.0507	0.6110	0.0507	-	0.6848	IW	0.1702	0.1305	0.4092	-	0.4051
LW	0.0443	0.5623	0.0443	-	0.6848	LW	0.1079	0.1043	0.3687	-	0.4051
MW	0.0443	0.5623	0.0443	-	0.6848	MW	0.1079	0.1043	0.3687	-	0.4051
$\mathbf{p}=(0.10,0.35,0.35)$						$\mathbf{p}=(0.10,0.30,0.50)$					
I	0.0474	0.5063	0.0474	0.5977	-	I	0.4356	0.3005	0.8260	0.8305	-
L	0.0494	0.4743	0.0494	0.5977	-	L	0.3788	0.3256	0.8231	0.8305	-
M	0.0494	0.4743	0.0494	0.5977	-	M	0.3788	0.3256	0.8231	0.8305	-
IW	0.0461	0.4965	0.0461	-	0.5502	IW	0.4303	0.3001	0.8246	-	0.7997
LW	0.0481	0.4499	0.0481	-	0.5502	LW	0.3788	0.3116	0.8201	-	0.7997
MW	0.0481	0.4499	0.0481	-	0.5502	MW	0.3788	0.3116	0.8201	-	0.7997

Table 2 shows the estimated value of PWER, FWER and the estimated pairwise power of six tests as I test, M test, L test, IW test, LW test and MW test by Monte Carlo simulation for $k = 5$.

For configuration 1, the estimated values of PWER and FWER, the results indicate that PWER and FWER of six tests are less than the nominal level of 0.05, ranging from 0.0000-0.0067 under **p** = (0.02, 0.02, 0.02, 0.02, 0.02), **p** = (0.05, 0.05, 0.05, 0.05, 0.05) and **p** = (0.1, 0.1, 0.1, 0.1, 0.1). The estimated values of PWER of six tests are less than the nominal level of 0.05 but FWER of I, L, IW and LW tests are close to the nominal level of 0. 05 for $\mathbf{p} = (0.01, 0.01, 0.01, 0.01, 0.01)$.

For Configuration 2, the estimated values of PWER of six tests are very small at the nominal level of 0.05 under **p** = (0.02, 0.02, 0.02, 0.02, 0.27), **p** = (0.05, 0.05, 0.05, 0.05, 0.3) and **p** = (0.1, 0.1, 0.1, 0.1, 0.35). For the results of the estimated value of PWER and the estimated value of FWER under $\mathbf{p} = (0.10, 0.10, 0.10, 0.10, 0.35)$ it was found that FWER of L test and LW test are more than the nominal level of 0. 05 compared with the other tests, but PWER of I, M, IW and MW tests are less than the nominal level of 0. 05. Considering the estimated pairwise power of I, L, M, IW, LW and MW tests, it appears that the best test statistics are I, L, IW and LW where the estimated pairwise powers of I, L, IW and LW tests present higher than the M test and MW test.

For Configuration 3, I test, M test, IW test and MW test are lower than the nominal level of 0.05 compared with L test and LW test in terms of the estimated values of PWER and the estimated value of FWER under **p** = (0.02, 0.27, 0.27, 0.27, 0.27), **p** = (0.05, 0.3, 0.3, 0.3, 0.3) and **p** = (0.1, 0.35, 0.35, 0.35, 0.35).
Considering the estimated pairwise power of I, L, M, IW, LW and MW tests, the results indicate that the L test and LW test have higher estimated pairwise power

For Configuration 4, we observe that I, L, IW and LW tests appear to have the highest estimated pairwise power of the tests in terms of the estimated values of PWER and the estimated values of FWER under **p** = (0.02, 0.04, 0.06, 0.08, 0.1), **p** = (0.05, 0.1, 0.15, 0.2, 0.25) and **p** = (0.1, 0.2, 0.3, 0.4, 0.5). Again, it was observed that the estimated pairwise power of M and MW is less than the other tests.

Table 2. PWER, FWER, estimated pairwise power and estimated probabilities for $k = 5$.

Configuration 1						Configuration 2					
$\mathbf{p} = (0.02, 0.02, 0.02, 0.02, 0.02)$						$\mathbf{p} = (0.02, 0.02, 0.02, 0.02, 0.27)$					
Test	$p_i = p_j$	$p_i \neq p_j$	FWER	Global G	Global W_P	Test	$p_i = p_j$	$p_i \neq p_j$	FWER	Global G	Global W_P
I	0.0000	-	0.0000	0.0003	-	I	0.0000	0.5301	0.0000	0.7349	-
L	0.0001	-	0.0002	0.0003	-	L	0.0001	0.6537	0.0003	0.7349	-
M	0.0000	-	0.0000	0.0003	-	M	0.0000	0.3865	0.0000	0.7349	-
IW	0.0000	-	0.0000	-	0.0002	IW	0.0000	0.5294	0.0000	-	0.6916
LW	0.0001	-	0.0002	-	0.0002	LW	0.0001	0.6343	0.0003	-	0.6916
MW	0.0000	-	0.0000	-	0.0002	MW	0.0000	0.3864	0.0000	-	0.6916
$\mathbf{p} = (0.05, 0.05, 0.05, 0.05, 0.05)$						$\mathbf{p} = (0.05, 0.05, 0.05, 0.05, 0.30)$					
I	0.0005	-	0.0020	0.0101	-	I	0.0007	0.8159	0.0024	0.6696	-
L	0.0020	-	0.0067	0.0101	-	L	0.0040	0.9196	0.0130	0.6696	-
M	0.0001	-	0.0004	0.0101	-	M	0.0001	0.7971	0.0005	0.6696	-
IW	0.0004	-	0.0017	-	0.0049	IW	0.0007	0.8151	0.0023	-	0.6228
LW	0.0013	-	0.0037	-	0.0049	LW	0.0038	0.9195	0.0122	-	0.6228
MW	0.0001	-	0.0004	-	0.0049	MW	0.0001	0.7970	0.0005	-	0.6228

(*continued*)

Table 2. Continued.

$\mathbf{p}=(0.10,0.10,0.10,0.10,0.10)$						$\mathbf{p}=(0.10,0.10,0.10,0.10,0.35)$					
I	0.0036	-	0.0155	0.0382	-	I	0.0054	0.3134	0.0198	0.5711	-
L	0.0072	-	0.0232	0.0382	-	L	0.0170	0.4422	0.0577	0.5711	-
M	0.0013	-	0.0063	0.0382	-	M	0.0012	0.2897	0.0044	0.5711	-
IW	0.0026	-	0.0146	-	0.0222	IW	0.0052	0.3104	0.0190	-	0.5011
LW	0.0051	-	0.0164	-	0.0222	LW	0.0161	0.4028	0.0545	-	0.5011
MW	0.0012	-	0.0055	-	0.0222	MW	0.0012	0.2756	0.0044	-	0.5011
		Configuration 3						Configuration 4			
Test	$p_i=p_j$	$p_i\neq p_j$	FWER	Global		Test	p_1,p_2	p_2,p_3	p_3,p_4	p_4,p_5	p_1,p_3
$\mathbf{p}=(0.02,0.27,0.27,0.27,0.27)$				G	W_P	$\mathbf{p}=(0.02,0.04,0.06,0.08,0.10)$					
I	0.0108	0.5325	0.0108	0.8322	-	I	0.0004	0.0007	0.0027	0.0058	0.0013
L	0.0431	0.6548	0.0431	0.8322	-	L	0.0005	0.0032	0.0098	0.0125	0.0041
M	0.0097	0.3944	0.0097	0.8322	-	M	0.0001	0.0002	0.0006	0.0017	0.0004
IW	0.0108	0.5182	0.0108	-	0.7857	IW	0.0004	0.0006	0.0025	0.0058	0.0011
LW	0.0431	0.6285	0.0431	-	0.7857	LW	0.0004	0.0027	0.0073	0.0093	0.0032
MW	0.0097	0.3884	0.0097	-	0.7857	MW	0.0001	0.0002	0.0006	0.0017	0.0004
$\mathbf{p}=(0.05,0.30,0.30,0.30,0.30)$						$\mathbf{p}=(0.05,0.10,0.15,0.20,0.25)$					
I	0.0115	0.4306	0.0115	0.6961	-	I	0.0089	0.0191	0.0166	0.0181	0.0526
L	0.0478	0.5385	0.0478	0.6961	-	L	0.0269	0.0433	0.0437	0.0484	0.1034
M	0.0116	0.3360	0.0116	0.6961	-	M	0.0025	0.0085	0.0088	0.0137	0.0231
IW	0.0114	0.4118	0.0114	-	0.6334	IW	0.0083	0.0180	0.0157	0.0162	0.0483
LW	0.0463	0.5015	0.0463	-	0.6334	LW	0.0231	0.0395	0.0388	0.0446	0.0925
MW	0.0115	0.3246	0.0115	-	0.6334	MW	0.0026	0.0085	0.0089	0.0139	0.0236
$\mathbf{p}=(0.10,0.35,0.35,0.35,0.35)$						$\mathbf{p}=(0.10,0.20,0.30,0.40,0.50)$					
I	0.0085	0.3205	0.0085	0.5657	-	I	0.0528	0.0352	0.0287	0.0240	0.2162
L	0.0457	0.4298	0.0457	0.5657	-	L	0.1198	0.1062	0.1098	0.1244	0.3705
M	0.0103	0.2840	0.0103	0.5657	-	M	0.0285	0.0315	0.0340	0.0411	0.1740
IW	0.0079	0.3024	0.0079	-	0.4959	IW	0.0525	0.351	0.0285	0.0239	0.2147
LW	0.0422	0.3983	0.0422	-	0.4959	LW	0.1176	0.1045	0.1063	0.1218	0.3614
MW	0.0097	0.2671	0.0097	-	0.4959	MW	0.0287	0.0316	0.0340	0.0412	0.1755

Configuration 4

Test	p_2,p_4	p_3,p_5	p_1,p_4	p_2,p_5	p_1,p_5	G	W_P
$\mathbf{p}=(0.02,0.04,0.06,0.08,0.10)$							
I	0.0039	0.0082	0.0061	0.0124	0.0156	0.0650	-
L	0.0119	0.0171	0.0158	0.0267	0.0267	0.0650	-
M	0.0010	0.0022	0.0018	0.0038	0.0038	0.0650	-
IW	0.0039	0.0081	0.0057	0.0118	0.0142	-	0.0398
LW	0.0092	0.0135	0.0118	0.0196	0.0196	-	0.0398
MW	0.0010	0.0022	0.0018	0.0038	0.0038	-	0.0398
$\mathbf{p}=(0.05,0.10,0.15,0.20,0.25)$							
I	0.0504	0.0462	0.1384	0.1156	0.2579	0.4201	-
L	0.1005	0.0930	0.2079	0.1901	0.1901	0.4201	-
M	0.0307	0.0351	0.0793	0.0859	0.0859	0.4201	-
IW	0.0482	0.0433	0.1275	0.1079	0.2338	-	0.3548
LW	0.0918	0.0846	0.1846	0.1737	0.1737	-	0.3548
MW	0.0312	0.0355	0.0805	0.0859	0.0870	-	0.3548
$\mathbf{p}=(0.10,0.20,0.30,0.40,0.50)$							
I	0.1430	0.1151	0.4716	0.3620	0.7250	0.8327	-
L	0.3116	0.3161	0.6479	0.5994	0.5994	0.8327	-
M	0.1453	0.1490	0.4468	0.3820	0.3820	0.8327	-
IW	0.1420	0.1146	0.4645	0.3577	0.7105	-	0.7854
LW	0.3046	0.3078	0.6244	0.5761	0.5782	-	0.7854
MW	0.1458	0.1494	0.4515	0.4515	0.3802	-	0.7854

4.2 Discussion

The performance of the proposed two-stage approach (IW test, LW test and MW test) and the two-stage approach (I test, L test and M test) for multiple comparisons are compared with PWER, FWER and the estimated pairwise power for the sample size of $\mathbf{n} = (25, 25, ..., 25)$ when the proportions p_i, $i = 1, 2, 3, ..., k$ are close to zero. From the results in Table 1 for $k = 3$, the IW test, LW test and MW test approach can protect PWER and FWER better than the I test, L test and M test. From the results in Table 2 for $k = 5$, the IW test, LW test and MW test can protect PWER and FWER better than the I test, L test and M test. Also, the IW test is a statistic of the proposed test which can protect PWER and FWER quite well compared with I, L, M, LW and MW. In terms of the estimated pairwise power, the IW test has the estimated pairwise power similar to the other tests. The IW test, LW test and MW test are the test statistics for multiple comparisons based on the Welch statistic, and can protect PWER and FWER quite well as shown by the results of Krishnamoorthy [6] and Noppakun et al. [7]. Therefore, the proposed two-stage approach is an alternative statistic for multiple comparison.

5. Conclusion

The IW test, LW test and MW test are compared with I test, L test and M test in terms of PWER, FWER and the estimated pairwise power. We observed that for k = 3, IW test, LW test and MW test tend to protect PWER and FWER quite well compared with I test, L test and M test. Moreover, I test and IW test are near the nominal level of 0.05. Considering the estimated pairwise power, when the estimated pairwise powers are observed for k = 3, the I test, L test, IW test and LW have high estimated pairwise power. Meanwhile, for k =5, the IW test, LW test and MW test appear to provide superb protection against PWER and FWER. However, results of the estimated pairwise power show that the I test, L test, IW test and LW test have higher estimated pairwise power compared with M and MW. Our results suggest that the IW test has the highest estimated pairwise power, and it shows nearly the nominal level of 0.05 in terms of PWER and FEWR compared with the I, L, M, LW and MW tests. Therefore, the IW test should be used as an alternative approach.

Acknowledgements

This study is supported by Kalasin University, Thailand.

References

[1] Lawrence DB, Tony T, Anirban DG. Confidence Intervals for a Binomial Proportion and Asymptotic Expansions. The Annals of Statistics 2002; 30(1): 160-201.

[2] Brown LD, Cai, TT, DasGupta A. Confidence interval for a binomial proportion and asymptotic expansions. *Annals of Statistics* 2002; 30 (1): 160-201.

[3] Kane N, Kristin MH, Christopher MT. Multiple comparisons of k binomial proportions. Computational Statistics and Data Analysis 2013; 68: 202-12.

[4] Agresti A, Bini M, Bertaccini B, Ryu E. Simultaneous confidence intervals for comparing binomial parameters. Biometrics 2008; 64: 1270-5.

[5] Hayter AJ. The maximum familywise error rate of Fisher’ s least significant difference test. Journal of the American Statistical Association 1986; 81: 1000-4.

[6] Krishnamoorthy K, Lu F, Mathew T. A parametric bootstrap approach for ANOVA with unequal variances: Fixed and random models. Computational Statistics and Data Analysis 2007; 51: 5731-42.

[7] Noppakun T, Wuttichai S, Angkana BA. Bayesian Approach to the One Way ANOVA under Unequal Variance. Sains Malaysiana 2016; 45(10): 1565-72.

[8] R Core Team. R: a language and environment for statistical computing. R Foundation for Statistical Computing. Vienna: Austria; 2018 [cited 2019 Jan 20]. Available from: https://www.R-project.org.

[9] Ana MP, Conceicao A, 2008. Interval estimations for a binomial proportion: comparison of twenty methods, Statistical Journal 2008; 6(2): 165-97.