Available online at www.sciencedirect.com
SCIENCE DIRECT®

Computational Statistics & Data Analysis 45 (2004) 697–707

COMPUTATIONAL STATISTICS & DATA ANALYSIS

www.elsevier.com/locate/csda

An exact method of testing equality of several binomial proportions to a specified standard

K. Krishnamoorthy[a,*], Jessica Thomson[b], Yong Cai[a]

[a]Department of Mathematics, University of Louisiana at Lafayette, Lafayette, LA 70504-1010, USA
[b]Department of Public Health and Preventive Medicine, Louisiana State University Health Sciences Center, 1600 Canal Street, Suite 800, New Orleans, LA 70112, USA

Received 1 December 2002; received in revised form 1 April 2003

Abstract

The problem of testing equality of several binomial proportions to a specified standard is considered. An exact method of testing based on the test statistic considered in Kulkarni and Shah (Statist. Probab. Lett. 25 (1995) 213) is proposed. Exact properties of the exact method and the approximate method due to Kulkarni and Shah are evaluated numerically for the two-sample case. Numerical studies show that the sizes of the approximate test due to Kulkarni and Shah often exceed the nominal level by considerable amount. For the one-sample case, numerical comparison shows that there is no clear cut winner between the new test and the usual exact test. A procedure for constructing simultaneous confidence intervals is also given. The methods are illustrated using two examples.

Keywords: Clopper–Pearson method; Multiple comparison; Powers; Simultaneous confidence intervals; Size

1. Introduction

The problem of testing equality of several binomial proportions is well-known, and has been considered in many text books in statistics. For example, see Rohatgi (1976) and Scheaffer and McClave (1994). The hypothesis testing of equality of several proportions to a prespecified quantity has been recently considered in Kulkarni and Shah (1995). These authors pointed out that such problem arises in practical situations and provided three examples. They considered a test statistic, which was mentioned in

* Corresponding author. Tel.: +1-337-482-5283; fax: +1-337-482-5346.
E-mail address: krishna@louisiana.edu (K. Krishnamoorthy).

0167-9473/$ - see front matter
doi:10.1016/S0167-9473(03)00095-1

Rohatgi (1976), whose null distribution is asymptotically chi-square. Kulkarni and Shah also derived a non-null distribution, and provided a method for testing hypotheses that involve one-sided alternatives.

To formulate the present problem, consider m independent binomial random variables $X_1,\ldots,X_m$ with $X_i \sim \text{binomial}(n_i, p_i)$, $0 < p_i < 1$, $i = 1,\ldots,m$. Let k_i be an observed value of X_i, $i = 1,\ldots,m$. The hypotheses of interest are

$$\mathrm{H}_0: p_1 = \cdots = p_m = p_0 \quad \text{vs.} \quad \mathrm{H}_\mathrm{a}: p_i \neq p_0 \text{ for some } i. \tag{1.1}$$

To test (1.1), Kulkarni and Shah (1995) consider the test statistic

$$T_X = \sum_{i=1}^{m} \frac{(X_i - n_i p_0)^2}{n_i p_0 q_0}, \tag{1.2}$$

where $q_0 = 1 - p_0$, whose null distribution is asymptotically chi-square with df $= m$ (see Rohatgi, 1976). Let k_i be an observed value of X_i, $i = 1,\ldots,m$, and let T_k be an observed value of T_X. The value of T_k can be computed using T_X in (1.2) with X_i replaced by k_i, $i = 1,\ldots,m$. The null hypothesis in (1.1) will be rejected whenever the p-value

$$P(\chi^2_m > T_k) \leqslant \alpha. \tag{1.3}$$

For testing

$$\mathrm{H}_0: p_1 = \cdots = p_m = p_0 \quad \text{vs.} \quad \mathrm{H}_\mathrm{a}: p_i > p_0 \text{ for some } i, \tag{1.4}$$

Kulkarni and Shah proposed the test statistic

$$T_X^+ = \sum_{i=1}^{m} \frac{(X_i - n_i p_0)^{2+}}{n_i p_0 q_0}, \tag{1.5}$$

where $x^{2+} = x^2$ if $x > 0$, and 0 otherwise. When $n_i = n$, $i = 1,\ldots,m$, an approximate formula to compute the p-value is given by

$$\sum_{i=1}^{m} P(Y = i) P(\chi^2_i > \text{observed } T_X^+), \tag{1.6}$$

where $Y \sim \text{binomial}(m, \theta)$, and $\theta = P(X > n p_0)$, with $X \sim \text{binomial}(n, p_0)$. These authors also provide an approximate formula for the p-value when the n_i's are not equal.

In this article, using the fact that the common proportion under the H_0 is known, we point out an exact method for computing the p-value of the test based on the statistic T_X in (1.2). The test method, given in the following section, is exact in the sense that the size of the test never exceeds the nominal level. We also show that this method is applicable for one-sided alternatives as well. In Section 3, we compared the usual test and the new test for the one-sample case, and found no clear cut winner between them. In Section 4, we evaluated exact powers and sizes of the proposed tests and the asymptotic tests given above for the two-sample case. Our numerical study shows that the sizes of the approximate tests exceed the nominal level by a considerable amount. In particular, the sizes of the approximate test based on (1.6) (for one-sided alternatives) exceed the nominal level significantly even when the sample sizes are large. In Section 5, we outline a method of constructing simultaneous confidence intervals based on the Clopper–Pearson (1934) method for constructing one-sample confidence interval. The methods are illustrated using two examples in Section 6.

2. The proposed testing method

The probability mass function of a binomial(n, p) random variable is given by

$$f(x; n, p) = \binom{n}{x} p^x (1-p)^{n-x}, \quad x = 0, 1, \ldots, n. \tag{2.1}$$

Note that under the null hypothesis in (1.1), the common proportion p_0 is known, and hence $f(x_i; n_i, p_0)$, is known for $i = 1, \ldots, m$. Since $X_1, \ldots, X_m$ are independent, the joint probability mass function is given by $P(X_1 = x_1, \ldots, X_m = x_m) = f(x_1; n_1, p_0) \times \cdots \times f(x_m; n_m, p_0)$. Using this fact, and an indicator function $I(.)$, the p-value of the test statistic T_X can be computed. Indeed, for a given observed value T_k, the p-value can be expressed as

$$\begin{aligned} P(T_X \geqslant T_k|\mathrm{H}_0) &= E_{X_1,\ldots,X_m} I(T_X \geqslant T_k|\mathrm{H}_0) \\ &= \sum_{x_1=0}^{n_1} \cdots \sum_{x_m=0}^{n_m} P(X_1 = x_1, \ldots, X_m = x_m|\mathrm{H}_0) I(T_x \geqslant T_k) \\ &= \sum_{x_1=0}^{n_1} \cdots \sum_{x_m=0}^{n_m} \left(\prod_{i=1}^{m} f(x_i; n_i, p_0) I(T_x \geqslant T_k) \right), \end{aligned} \tag{2.2}$$

where $E_{X_1,\ldots,X_m}$ denotes the expectation under the joint distribution of $X_1, \ldots, X_m$. The null hypothesis in (1.1) will be rejected whenever the p-value in (2.2) is less than or equal to α.

Remark 2.1. At the outset one may think of the Fisher's exact test (the conditional test) for testing (1.1). We, however, note that the Fisher's method often tests the equality of proportions and not the equality of proportions to the specified standard. Hence, the Fisher's method is not applicable to the present problem.

The p-value for testing one-sided alternatives can be computed similarly. For testing (1.4), the p-value is given by

$$P(T_X^+ > T_k^+|\mathrm{H}_0) = \sum_{x_1=0}^{n_1} \cdots \sum_{x_m=0}^{n_m} \left(\prod_{i=1}^{m} f(x_i; n_i, p_0) I(T_x^+ > T_k^+) \right). \tag{2.3}$$

The value of T_k^+ can be computed using the expression for T_X^+ in (1.5) with X_i replaced by k_i, $i = 1, \ldots, m$.

For smaller m, say $m \leqslant 5$, the p-values in (2.2) and (2.3) can be computed in a straightforward manner using softwares such as SAS and Fortran with IMSL libraries. A faster computational approach is to compute each probability mass function at its mode (integer part of $n_i p_0$), and then to compute other terms using forward and backward recurrence relations for binomial probabilities. For moderate m one can use Monte Carlo method, as shown in the following algorithm, to compute the p-value.

Algorithm 1. For a given p_0, m, n_i's and k_i's:

Compute $T_k = \sum_{i=1}^{m} \frac{(k_i - n_i p_0)^2}{n_i p_0 q_0}$

Generate $X_{ij} \sim \text{binomial}(n_i, p_0)$, $i = 1, \ldots, k$, $j = 1, \ldots, N$

For each j, compute $T_{Xj} = \sum_{i=1}^{k} \frac{(X_{ij} - n_i p_0)^2}{n_i p_0 q_0}$, $j = 1, \ldots, N$.

The proportion of T_{Xj}'s that are greater than or equal to T_k is a Monte Carlo estimate of the p-value in (2.2).

The p-value in (2.3) can be computed similarly. If Algorithm 1 is used with $N = 100,000$, then the maximum simulation error will be $2 \times \sqrt{0.5(1-0.5)/100,000} = 0.0032$. Thus, the Monte Carlo results based on 100,000 runs maybe quite satisfactory, for practical applications.

3. Size and power properties for the one-sample case

For the one-sample case, the proposed test is different from the usual one, and hence it is of interest to compare them with respect to size and power. For an observed value k_1 of n_1, the usual exact test rejects the H_0 in (1.1), whenever the p-value

$$2 \min\{P(X_1 \geqslant k_1 | n_1, p_0), P(X_1 \leqslant k_1 | n_1, p_0)\} \leqslant \alpha.$$

The sizes and powers of both tests can be computed using the expression

$$\sum_{k_1=0}^{n_1} f(k_1; n_1, p_1) I((p\text{-value} \mid n_1, p_0) \leqslant \alpha).$$

Note that when $p_1 = p_0$, the above expression gives exact size. We computed the sizes of both exact tests and presented them in Table 1. We observe from the table values that the sizes of the new test are, in most cases, closer to the nominal level than are the sizes of the usual test. Comparison of powers in Table 2 indicates that there is no clear cut winner between these two exact tests. We observe that the test that has larger size has more power than the other test. For example, we see in Table 1 that the sizes of the new test are greater than or equal to those of the usual test, and in this case the new test seems to be more powerful than the usual one (see Table 2, $p_0 = 0.1$). On the other hand, the usual test has better size property than the new test for $p_1 = p_0 = 0.25, n_1 = 14$ and 20. In these situations, the usual test appears to be more powerful than the new test (see Table 2, $p_0 = 0.25, n_1 = 14$ and 20).

4. Size and power properties for the two-sample case

For a given n_1, n_2, p_1, p_2, p_0, recall that the power of a test is the probability of rejecting H_0 when H_a is true. The exact power of the approximate test in (1.3) can be

Table 1
Exact sizes of the usual test (1) and the new test (2) when $\alpha = 0.05$

$p = p_0$	$n = 10$		14		20		27		50	
	1	2	1	2	1	2	1	2	1	2
0.05	0.012	0.012	0.004	0.030	0.016	0.016	0.010	0.044	0.012	0.038
0.10	0.013	0.013	0.009	0.044	0.011	0.043	0.015	0.047	0.030	0.030
0.15	0.010	0.050	0.012	0.047	0.022	0.022	0.026	0.026	0.027	0.044
0.20	0.006	0.033	0.012	0.044	0.022	0.044	0.030	0.049	0.033	0.049
0.25	0.020	0.020	0.028	0.010	0.038	0.017	0.042	0.042	0.033	0.048
0.30	0.011	0.011	0.015	0.038	0.025	0.025	0.034	0.034	0.031	0.043
0.35	0.018	0.039	0.045	0.045	0.032	0.032	0.041	0.041	0.037	0.037
0.40	0.018	0.018	0.026	0.026	0.037	0.037	0.029	0.049	0.029	0.029
0.45	0.028	0.007	0.028	0.028	0.040	0.040	0.032	0.032	0.045	0.045
0.50	0.021	0.021	0.013	0.013	0.041	0.041	0.019	0.019	0.033	0.033

Table 2
Exact powers of the usual test (1) and the new test (2) when $H_0 : p = p_0$ vs. $H_a : p \neq p_0$ and $\alpha = 0.05$

p	$n = 10$		14		20		27		50	
	1	2	1	2	1	2	1	2	1	2
$p_0 = 0.1$										
0.05	0.001	0.001	0.000	0.004	0.000	0.003	0.000	0.002	0.077	0.077
0.10	0.013	0.013	0.009	0.044	0.011	0.043	0.015	0.047	0.030	0.030
0.15	0.050	0.050	0.047	0.147	0.067	0.170	0.099	0.210	0.209	0.209
0.20	0.121	0.121	0.130	0.302	0.196	0.370	0.287	0.461	0.556	0.556
0.25	0.224	0.224	0.258	0.479	0.383	0.585	0.529	0.701	0.836	0.836
0.30	0.350	0.350	0.416	0.645	0.584	0.762	0.744	0.864	0.960	0.960
0.35	0.486	0.486	0.577	0.780	0.755	0.882	0.885	0.949	0.993	0.993
0.40	0.618	0.618	0.721	0.876	0.874	0.949	0.958	0.985	0.999	0.999
0.45	0.734	0.734	0.833	0.937	0.945	0.981	0.987	0.996	1.00	1.00
0.50	0.828	0.828	0.910	0.971	0.979	0.994	0.997	0.999	1.00	1.00
$p_0 = 0.25$										
0.05	0.000	0.000	0.488	0.000	0.736	0.358	0.850	0.850	0.988	0.988
0.10	0.000	0.000	0.229	0.000	0.392	0.122	0.485	0.485	0.770	0.770
0.15	0.001	0.001	0.103	0.000	0.176	0.039	0.208	0.208	0.361	0.361
0.20	0.006	0.006	0.046	0.002	0.072	0.014	0.075	0.075	0.104	0.106
0.25	0.020	0.020	0.028	0.010	0.038	0.017	0.042	0.042	0.033	0.048
0.30	0.047	0.047	0.038	0.031	0.056	0.049	0.085	0.085	0.087	0.143
0.35	0.095	0.095	0.078	0.075	0.124	0.122	0.203	0.203	0.274	0.379
0.40	0.166	0.166	0.151	0.150	0.245	0.245	0.387	0.387	0.554	0.664
0.45	0.262	0.262	0.259	0.259	0.409	0.409	0.597	0.597	0.803	0.873
0.50	0.377	0.377	0.395	0.395	0.588	0.588	0.779	0.779	0.941	0.968

expressed as

$$P(\chi_2^2 > T_k|\mathrm{H_a}) = \sum_{k_1=0}^{n_1}\sum_{k_2=0}^{n_2} f(k_1; n_1, p_1) f(k_2; n_2, p_2) I(P(\chi_2^2 > T_k|\mathrm{H_0}) \leqslant \alpha). \quad (4.1)$$

The power of test (1.6) for one-sided alternative can be expressed similarly.

The power of the exact test in (2.2) can be written as

$$\sum_{k_1=0}^{n_1}\sum_{k_2=0}^{n_2} f(k_1; n_1, p_1) f(k_2; n_2, p_2) I$$
$$\times \left(\sum_{x_1=0}^{n_1}\sum_{x_2=0}^{n_2} f(x_1; n_1, p_0) f(x_2; n_2; p_0) I(T_x \geqslant T_k) \leqslant \alpha\right). \quad (4.2)$$

The power of the exact test in (2.3) for testing one-sided alternatives can be computed using (4.2) with (T_x, T_k) replaced by (T_x^+, T_k^+).

Notice that expressions (4.1) and (4.2) give sizes of the tests when $p_1 = p_2 = p_0$. It is clear from the above expressions that the power functions can be extended to the case of $m \geqslant 3$ in a straightforward manner. Even though the power functions can be written for any m, they are computationally intensive when $m \geqslant 3$. In these cases, we again recommend Monte Carlo method similar to the one presented in Algorithm 1.

The sizes of the approximate test and the exact test are computed using (4.1) and (4.2) for the case of $m=2$, and various values of sample sizes and parameter configurations. The sizes are plotted in Figs. 1(a)–(f) for testing two-sided hypotheses in (1.1), and in Figs. 2(a)–(f) for one-sided alternatives in (1.4). It is clear from Figs. 1(a)–(e) that the sizes of the approximate test based on (1.3) are larger than the nominal level 0.05 when the proportions are small. When the alternatives are one-sided, the sizes of the approximate test exceed the nominal level considerably even for moderate samples (see Figs. 2(a)–(e)). The sizes of the exact tests are either below or very close to the nominal level 0.05 for all the cases considered.

To understand the power properties of the tests, we plotted the powers of the approximate test and the exact test in Fig. 3. Both tests satisfy some natural requirements of a test. In particular, we see from Figs. 3(a)–(d) that the power of the tests increases as the sample size increases, and/or when the difference between the true parameter and the specified parameter increases. The power of the approximate test is inflated (see Fig. 3(a)) because of its larger sizes.

5. Multiple comparison

If the null hypothesis in (1.1) is rejected, then it is important to identify the population proportions which caused the rejection. This can be done by examining the simultaneous confidence intervals for $p_1, \ldots, p_m$. Toward this end, we describe the Clopper and Pearson's (1934) confidence interval (see Johnson and Kotz (1969, p. 58)) for the one-sample case. Let $\mathrm{Beta}(c; a, b)$ denote the cth quantile of a beta distribution with

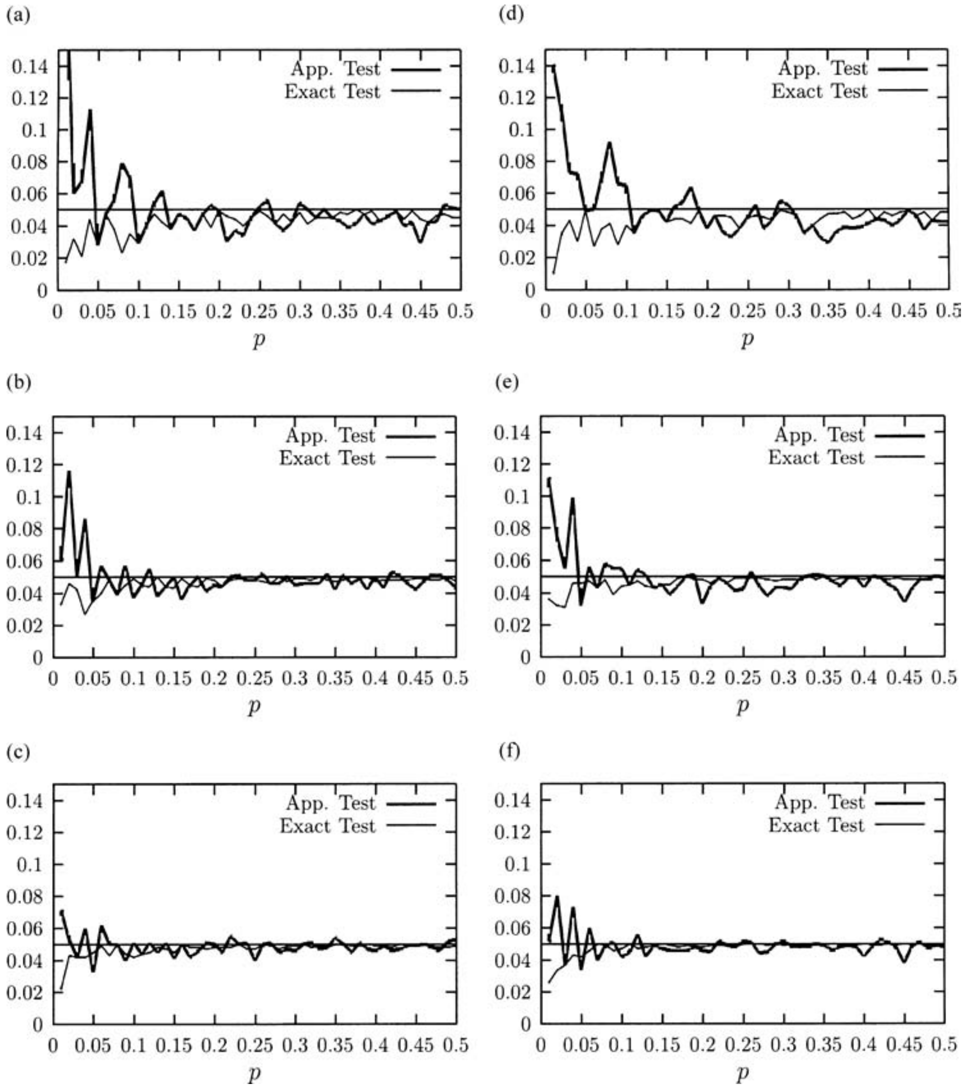

Fig. 1. Exact sizes of the two-sided tests as a function of the proportions at the nominal level $\alpha = 0.05$; 'App. Test' is the test based on (1.3), 'Exact Test' is the test based on (2.2): (a) $n_1 = n_2 = 10$, (b) $n_1 = n_2 = 20$, (c) $n_1 = n_2 = 30$, (d) $n_1 = 10,\ n_2 = 5$, (e) $n_1 = 20,\ n_2 = 10$, and (f) $n_1 = 30,\ n_2 = 20$.

shape parameters a and b. Then, for a given $0 < c < 1$, the Clopper–Pearson $1 - c$ confidence interval for p_i based only on (k_i, n_i) is given by

$$(L_i = \mathrm{Beta}(c/2; k_i, n_i - k_i + 1), U_i = \mathrm{Beta}(1 - c/2; k_i + 1, n_i - k_i)) \quad \text{and}$$

$$(L_i, U_i) = (0, (1-c)^{1/n_i}), \text{ if } k_i = 0; \quad (L_i, U_i) = (c^{1/n_i}, 1), \text{ if } k_i = n_i. \qquad (5.1)$$

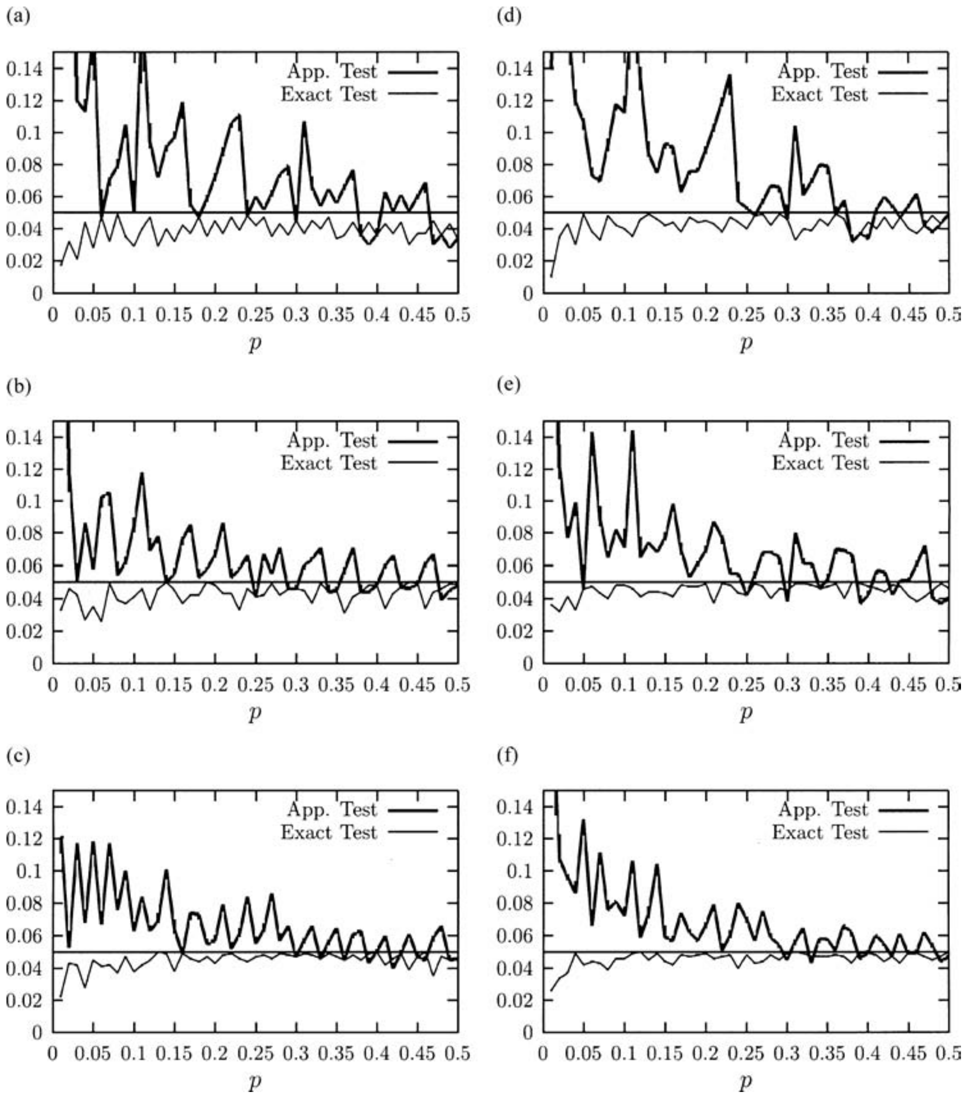

Fig. 2. Exact sizes of the one-sided tests as a function of the proportions at the nominal level $\alpha = 0.05$; 'App. Test' is the test based on (1.6), 'Exact Test' is the test based on (2.3): (a) $n_1 = n_2 = 10$, (b) $n_1 = n_2 = 20$, (c) $n_1 = n_2 = 30$, (d) $n_1 = 10,\ n_2 = 5$, (e) $n_1 = 20,\ n_2 = 10$, and (f) $n_1 = 30,\ n_2 = 20$.

The coverage probability of this confidence interval is at least $1-c$ for any sample size. It should be noted that it is not possible to construct an exact confidence interval (that is, the coverage probability is equal to the confidence level) because of the discreteness of the distribution. By choosing $c = 1 - (1 - \alpha)^{1/m}$, we see that (L_i, U_i) contains p_i with probability at least $(1 - \alpha)^{1/m}$, and hence

$$P(p_i \in (L_i, U_i),\ i = 1, \ldots, m) \geqslant 1 - \alpha. \tag{5.2}$$

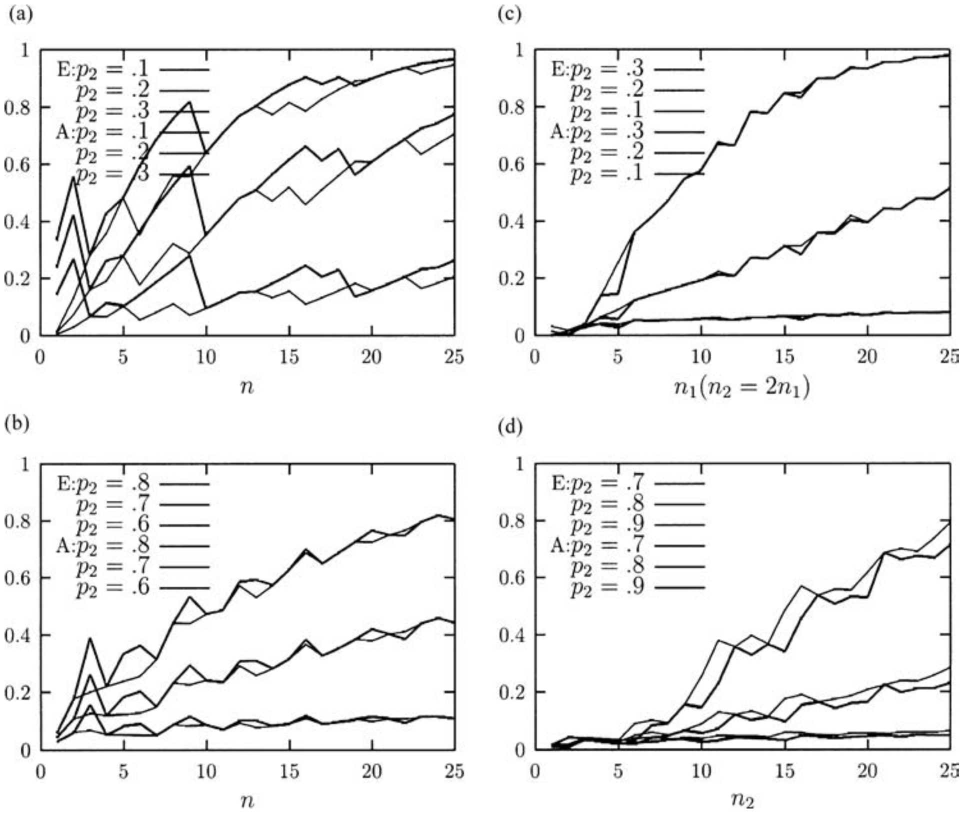

Fig. 3. Actual power of the tests as a function of the sample sizes at the nominal level $\alpha = 0.05$; A=the approximate test based on (1.3); E = the exact test based on (2.2): (a) $p_0 = p_1 = 0.05$, (b) $p_0 = p_1 = 0.85$, (c) $p_0 = p_1 = 0.35$, and (d) $p_0 = p_1 = 0.65$, $n_1 = 5$.

Thus, if $p_0 \notin (L_j, U_j)$, $j=1,\ldots,m_1 \leqslant m$, then we conclude that the p_j's are significantly different from p_0 at the level α.

One-sided intervals for $p_1,\ldots,p_m$ can also be constructed similarly. Suppose we want to find the population proportions that are greater than p_0. Then, taking $c=1-(1-\alpha)^{1/m}$, we see that

$$P(p_i \in (L_i, 1),\ i = 1,\ldots,m) \geqslant 1 - \alpha, \tag{5.3}$$

where $L_i = \mathrm{Beta}(c; k_i, n_i - k_i + 1)$, $L_i = 0$ if $k_i = 0$ and $L_i = c^{1/n_i}$ if $k_i = n_i$, $i = 1,\ldots,m$. If $L_j > p_0$, $j = 1,\ldots,m_1 \leqslant m$, then we conclude that the corresponding p_j's exceed p_0 at the level α.

Simultaneous upper limits can be constructed similarly. Taking c as above in the construction of lower limits, we see that

$$P(p_i \in (0, U_i),\ i = 1,\ldots,m) \geqslant 1 - \alpha,$$

where $U_i = \mathrm{Beta}(1 - c; k_i + 1, n_i - k_i)$, $U_i = 1$ if $k_i = n_i$ and $U_i = (1 - c)^{1/n_i}$ if $k_i = 0$.

6. Examples

We shall now illustrate the methods using the two examples given in Kulkarni and Shah (1995).

Example 1. In this example, we are interested in estimating the percentage of gopher tortoise active burrows. The data were obtained by experience judgement and camera judgement during the summers of 1987–1989. The data along with 95% simultaneous confidence intervals are given in the following table.

Judgement type and period	n_i	k_i	90% CI	95% CI
Experience judgement				
Summer 1987 and 1988	81	59	(0.6097, 0.8274)	(0.5945, 0.8380)
Experience judgement				
Summer 1989	151	107	(0.6230, 0.7846)	(0.6120, 0.7933)
Camera judgement				
Summer 1989	114	48	(0.3226, 0.5243)	(0.3109, 0.5371)

The established percentage of active burrows is 62%, and hence we want to test $H_0 : p_1 = p_2 = p_3 = 0.62$ vs. $H_a : p_i \neq 0.62$ for some i. Both the approximate test in (1.3) and the exact test in (2.2) produced p-values very close to 0. To find out the proportions that are not equal to $p_0 = 0.62$, we constructed 90% and 95% simultaneous confidence intervals for p_1, p_2, p_3, and presented in the above table. These simultaneous confidence intervals are constructed using (5.1) with $c = 1 - (0.95)^{1/3} = 0.01695$. Examining the 95% confidence intervals, we see that the third interval does not contain $p_0 = 0.62$, and hence we conclude that p_3 is different from 0.62 at the significance level of 0.05.

Example 2. A manufacturing firm produces bolts using six different processes. The firm inspects the bolts for defects periodically. The production process will be stopped for correction if there is an indication that at least one of the machines has started production higher than 5% of defective bolts. The data set along with 90% simultaneous lower limits for the proportions are given in the following table.

Machine	k_i	n_i	90% Lower limits	95% Lower limits
1	11	100	0.05317	0.04791
2	6	100	0.02054	0.01752
3	2	100	0.00200	0.00137
4	8	100	0.03283	0.02884
5	8	100	0.03283	0.02884
6	4	100	0.00984	0.00794

The hypotheses of interest are

$$H_0 : p_1 = \cdots = p_6 = 0.05 \quad \text{vs.} \quad H_a : p_i > 0.05 \text{ for some } i.$$

Kulkarni and Shah's (1995) approximate method produced a p-value of 0.008 whereas the Monte Carlo method in Algorithm 1 produced a p-value of 0.029. Kulkarni and Shah also reported the p-value for two-sided alternative, which is 0.036; our Monte Carlo method also yielded 0.036. It is clear that Monte Carlo method p-values for one-sided alternatives and two-sided alternatives are in agreement, whereas those of approximate methods are not. We also provided 90% simultaneous confidence intervals and 95% simultaneous confidence intervals for $p_1, \ldots, p_6$ using (5.2). Based on these intervals, we can conclude, at the level 0.10, that machine 1 produces defective bolts at a rate higher than 5%.

Acknowledgements

The authors are thankful to a referee and an associate editor for their comments.

References

Clopper, C.J., Pearson's, E.S., 1934. The use of confidence or fiducial limits illustrated in the case of the binomial. Biometrika 26, 404–413.

Johnson, N.L., Kotz, S., 1969. Discrete Distributions. Wiley, New York, NY.

Kulkarni, P.M., Shah, A.K., 1995. Testing the equality of several binomial proportions to a prespecified standard. Statist. Probab. Lett. 25, 213–219.

Rohatgi, V.K., 1976. An Introduction to Probability Theory and Mathematical Statistics. Wiley, New York.

Scheaffer, R.L., McClave, J.T., 1994. Probability and Statistics for Engineers. Duxbury Press, New York.