

Type I error for a chi-square test when the response probability changes during a trial

FEIFANG HU^{*,†}, JIANDONG LU AND FENG TAI

Chi-square test is a commonly used method for testing null hypothesis of no difference between treatment groups in a binary response in a randomized clinical trial. The Type I error is considered well controlled with a homogeneous population. In this article, we investigate the type I error of the chi-square test when the underlying response probability changes in the middle of a trial. In particular, we derive the asymptotic properties of the chi-square test under such an assumption and show that the type I error of the chi-square test is well controlled. Therefore, the chi-square test is still valid under change in response probability. Additionally, we present the computation of the actual type I error and some numerical results in an example to illustrate impact on Type I error by a change in response probability.

KEYWORDS AND PHRASES: Response probability, Cochran-Mantel-Haenzsel test, Asymptotic property.

1. INTRODUCTION

Chi-square test is used for testing null hypothesis of no difference between treatment groups in a binary response in a randomized clinical trial. The Type I error is considered well controlled with a homogeneous population. However, in practice, the patient population may change over the period of enrollment which may span a period of months or even years.

The reasons for the change can be an expansion of the investigational sites, new external information about the investigational treatment, or even a change in study design. Because of small patient counts, it is common that studies enroll patients globally. Due to differing regulatory requirements and standard of care across countries, the characteristics of the patient population may differ over the period of enrollment. In addition some countries may start to enroll patients early, while some other countries may enroll patients later. For example, a global clinical trial may enroll patients from North America, South America, Europe, and Asia. For logistic reasons, such as time required for the

approval of the trial in different countries, patients enrolled first are from North America and Europe, then from South America, and finally from Asian countries.

Sometimes, new relevant information from external sources may become available during enrollment and it may trigger a change in the study design (e.g., dropping treatment arms). Because such a change to the protocol will be communicated to investigators, who determine the eligibility of patients enrolled and assess the outcome of patients, it is debatable whether such communication will influence the investigator's opinion of the study agent, and consequently will affect their decision making in terms of patients to be enrolled and then clinical assessment of patients.

One such example is the design of a phase III trial to evaluate the safety and efficacy of an investigational agent in patients with moderate to severe ulcerative colitis who are unresponsive to conventional therapy. The trial, a randomized double blinded placebo controlled study, has a dropping dose design. It starts with randomizing patients into 3 active dose groups and a placebo. The primary endpoint is the clinical response based on the improvement in Mayo score, a UC disease index. With an expectation that the results of a separate dose finding study becomes available to the sponsor during the enrollment, the protocol pre-specifies the dropping of 1 or 2 dose groups. The Chi-square test is to be used for the test of null hypothesis of no treatment effect between selected dose treatment group(s) and the placebo group by including all patients irrespective of whether they are enrolled before or after the change.

In both cases, the assumption of homogenous population may be violated. Consequently, it is unrealistic to assume a constant response rate in each treatment group over the time of enrollment. More examples and related discussions can be found in Altman and Royston (1988), Bai and Hu (1999), Coad (1991), Hu and Rosenberger (2000), Duan and Hu (2009), etc. Can the simple Chi-square test still be used to test the null hypothesis of no treatment effect?

This article intends to investigate the impact to Type I error of the chi-square test if there were such a change in underlying response probability. Section 2 defines the statistical problem, introduces notation and derives the asymptotic property of the chi-square test under such an assumption. Section 3 computes the actual type I error of the chi-square test for a finite sample. Numerical results are presented for the introductory example in comparison to the chi-square

*Professor Hu's research was partially supported by grant DMS-0907297 and DMS-0906661 from the National Science Foundation (USA).

†Corresponding author.

test in a homogenous population and CMH test with a stratum for before and after the decision of change. Finally the article concludes with a discussion of the practical implications to the analysis of such clinical trials.

2. STATISTICAL PROPERTY

2.1 Notations and framework

Without loss of generality, we consider the comparison of two treatment groups with an equal sample size. The variable of interest is a binary response variable. Let X_{11} and X_{12} represent the number of patients with response before the change is made to the trial for treatment groups 1 and 2, respectively. And let X_{21} and X_{22} be the number of patients with response after the change is made to the trial for treatment groups 1 and 2, respectively. Assume that X_{ij} is a random variable from a binomial distribution $B(n_i; p_{ij})$ where $i = 1, 2, j = 1, 2$, and n_1 and n_2 are the number of patients in each treatment group before and after the change is made. The observed data are presented in Table 1.

Combine the observations from before and after together, let $n = n_1 + n_2$,

$$S_1 = X_{11} + X_{21} \quad \text{and} \quad S_2 = X_{12} + X_{22}$$

be the total number of response of treatment 1 and 2 respectively. Also let

$$S = S_1 + S_2 = X_{11} + X_{21} + X_{12} + X_{22}$$

be the total number of response from both treatments. The null hypothesis is $H_0: p_{11} = p_{12} = p_1$ and $p_{21} = p_{22} = p_2$.

The chi-square test statistic for combined samples can be written as,

$$T_{chi} = (S_1 - S/2)^2/V_1$$

where

$$V_1 = \frac{n^2 S(2n - S)}{(2n)^2(2n)} = \frac{S(2n - S)}{8n},$$

for Pearson chi-square test, while

$$V_1 = \frac{n^2 S(2n - S)}{(2n)^2(2n - 1)} = \frac{S(2n - S)}{4(2n - 1)},$$

for Mantel-Haenszel chi-square test. For brevity of discussion below, we only consider Pearson chi-square since they have the same limiting properties.

On the other hand, the Cochran-Mantel-Haenszel (CMH) test statistics with stratum of before or after the change is made (Stokes, Davis and Koch, 2000), is

$$T_{cmh} = (S_1 - S/2)^2/V_2,$$

where

$$\begin{aligned} V_2 &= \frac{n_1^2 d_1(2n_1 - d_1)}{(2n_1)^2(2n_1 - 1)} + \frac{n_2^2 d_2(2n_2 - d_2)}{(2n_2)^2(2n_2 - 1)} \\ &= \frac{d_1(2n_1 - d_1)}{4(2n_1 - 1)} + \frac{d_2(2n_2 - d_2)}{4(2n_2 - 1)}. \end{aligned}$$

For both cases ($p_1 = p_2$ and $p_1 \neq p_2$), it is known that T_{cmh} has a limit (for large n) χ^2 distribution with 1 degree of freedom under $H_0: p_{11} = p_{12} = p_1$ and $p_{21} = p_{22} = p_2$ is true. When $p_1 = p_2$ (the population does not change), it is well known that T_{chi} has a limit χ^2 distribution with 1 degree of freedom H_0 . However, it is unclear about the limiting distribution of T_{chi} under H_0 for the case $p_1 \neq p_2$ (the population does change). We will discuss this in the next subsection.

2.2 Asymptotic property of chi-square test statistic when $p_1 \neq p_2$

To discuss the properties of T_{chi} and T_{cmh} , we first write

$$S_1 - S/2 = (X_{11} - X_{12})/2 + (X_{21} - X_{22})/2,$$

which is the sum of independent random variables. Its corresponding mean and variance are

$$\begin{aligned} E(S_1 - S/2) &= E[(X_{11} - X_{12}) + (X_{21} - X_{22})]/2 \\ &= [n_1(p_1 - p_1) + n_2(p_2 - p_2)]/2 \\ &= 0 \end{aligned}$$

and

$$\begin{aligned} \text{Var}(S_1 - S/2) &= \text{Var}([(X_{11} - X_{12}) + (X_{21} - X_{22})]/2) \\ &= [n_1 p_1(1 - p_1) + n_2 p_2(1 - p_2)]/2. \end{aligned}$$

By the central limit theorem (for sum of independent random variables, not for independent and identical distributed (iid) random variables, see appendix of Hu and Rosenberger, 2006), we have

$$\frac{S_1 - S/2}{(n[\rho p_1(1 - p_1) + (1 - \rho)p_2(1 - p_2)]/2)^{1/2}} \rightarrow N(0, 1)$$

Table 1. Data structure and observations

	Before			After			Combined		
	Trt 1	Trt 2	Total	Trt 1	Trt 2	Total	Trt 1	Trt 2	Total
response	X_{11}	X_{12}	d_1	X_{21}	X_{22}	d_2	S_1	S_2	S
no response	$n_1 - X_{11}$	$n_1 - X_{12}$	$2n_1 - d_1$	$n_2 - X_{21}$	$n_2 - X_{22}$	$2n_2 - d_2$	$n - S_1$	$n - S_2$	$2n - S$
	n_1	n_1		n_2	n_2		n	n	

in distribution, when $n \rightarrow \infty$ and $n_1/n \rightarrow \rho \in (0, 1)$. Therefore,

$$\frac{(S_1 - S/2)^2}{(n[\rho p_1(1 - p_1) + (1 - \rho)p_2(1 - p_2)]/2)} \rightarrow \chi_{(1)}^2$$

in distribution, where $\chi_{(1)}^2$ represents Chi-squared distribution with 1 degree of freedom. Because $n_1/n \rightarrow \rho$, $d_1/n_1 \rightarrow 2p_1$ and $d_2/n_2 \rightarrow 2p_2$ in probability, we have

$$\frac{V_2}{(n[\rho p_1(1 - p_1) + (1 - \rho)p_2(1 - p_2)]/2)} \rightarrow 1$$

in probability by Slutsky's Theorem (Bickel and Doksum, 1977, page 461). This reiterates the asymptotic property of CMH test, whose type I error is generally controlled in large sample theory.

For the Chi-square test statistics T_{chi} , we have the following results.

Theorem 2.1. *When $n \rightarrow \infty$ and $n_1/n \rightarrow \rho \in (0, 1)$, then*

$$C_1 T_{chi} \rightarrow \chi_{(1)}^2$$

in distribution under the H_0 : $p_{11} = p_{12} = p_1$ and $p_{21} = p_{22} = p_2$, where

$$C_1 = \frac{(\rho p_1 + (1 - \rho)p_2)[\rho(1 - p_1) + (1 - \rho)(1 - p_2)]}{[\rho p_1(1 - p_1) + (1 - \rho)p_2(1 - p_2)]}.$$

Proof can be found in the Appendix. The theorem establishes the relationship between the chi-square test and CMH test via a constant C_1 . Theoretically we have the following result about the constant C_1 .

Lemma 2.1. *For all $p_1, p_2 \in (0, 1)$ and $\rho \in (0, 1)$,*

$$C_1 = \frac{(\rho p_1 + (1 - \rho)p_2)[\rho(1 - p_1) + (1 - \rho)(1 - p_2)]}{[\rho p_1(1 - p_1) + (1 - \rho)p_2(1 - p_2)]} \geq 1.$$

Proof can be found in the Appendix. Lemma 2.1 shows that the constant $C_1 \geq 1$ for any p_1, p_2 , and ρ , indicating the chi-square test statistic T_{chi} is always less than T_{cmh} under the null hypothesis for large sample size. Therefore, chi-square test is more conservative than the CMH test under the null hypothesis H_0 . Figure 1 illustrates the value of C_1 for various values of ρ, p_1 , and p_2 . Since $C_1(p_1, p_2) = C_1(1 - p_1, 1 - p_2)$, we only need to consider the cases where $p_2 - p_1 \geq 0$. C_1 increases as ρ or $p_2 - p_1$ increases, indicating that the chi-square test becomes more conservative if the change in the response rate is more substantial and if the proportion of subjects before the change increases.

3. ACTUAL TYPE I ERROR FOR CHI-SQUARE AND CMH TEST

In section 2, we showed that the chi-square test has a smaller type I error asymptotically. It is important to know

its type I error for finite sample. To understand the impact to the type I error due to the difference of p_1 and p_2 , we extended the formula provided by Garside and Mack (1976) to the situation under study. Let S_{chi} (S_{cmh}) be the set of all distinct vector of $(X_{11}, X_{12}, X_{21}, X_{22})$ (see the observed data above), for which the chi-square test (CMH test) rejects the null hypothesis H_0 at the significance level of α (say, 0.05). Then the exact Type I error for the test, for any pair of (p_1, p_2) , is given by

$$\begin{aligned} E_{chi}(p_1, p_2) &= \sum_{(X_{11}, X_{12}, X_{21}, X_{22}) \in S_{chi}} \binom{n_1}{X_{11}} \binom{n_1}{X_{12}} \binom{n_2}{X_{21}} \binom{n_2}{X_{22}} \\ &\times p_1^{d_1} (1 - p_1)^{2n_1 - d_1} p_2^{d_2} (1 - p_2)^{2n_2 - d_2} \end{aligned}$$

and,

$$\begin{aligned} E_{cmh}(p_1, p_2) &= \sum_{(X_{11}, X_{12}, X_{21}, X_{22}) \in S_{cmh}} \binom{n_1}{X_{11}} \binom{n_1}{X_{12}} \binom{n_2}{X_{21}} \binom{n_2}{X_{22}} \\ &\times p_1^{d_1} (1 - p_1)^{2n_1 - d_1} p_2^{d_2} (1 - p_2)^{2n_2 - d_2}. \end{aligned}$$

In addition, because for every vector of $(X_{11}, X_{12}, X_{21}, X_{22})$ in a rejection region, there is the vector of $(n_1 - X_{11}, n_1 - X_{12}, n_2 - X_{21}, n_2 - X_{22})$ in the rejection region, we have

$$E_{chi}(p_1, p_2) = E_{chi}(1 - p_1, 1 - p_2)$$

and,

$$E_{cmh}(p_1, p_2) = E_{cmh}(1 - p_1, 1 - p_2).$$

Therefore, we only consider the cases when $p_2 - p_1 \geq 0$.

In the introductory example, the protocol specifies to enroll a total of 157 patients in each treatment group chosen to be continued. The sponsor projects that the dose selection decision will occur when $n_1 = 29$ enrolled in each of the 4 treatment groups, which leaves $n_2 = 128$. To investigate the impact of timing of such a decision (i.e., $n_1:n_2$), we also compute the type I error for the cases where $n_1 = 15$ and $n_1 = 58$. Figure 2 showed the actual type I error for the situation, where $p_2 - p_1 = 0, 0.05, 0.1, 0.2$, respectively.

When $p_2 - p_1 = 0$ (in this case, $C_1 = 1$), the chi-square test and the CMH test have a similar type I error, because they have the same asymptotic distribution. When $p_2 - p_1$ increases, the chi-square test tends to have a smaller type I error than the CMH test, especially in the case of $p_2 - p_1 = 0.2$. This agrees with our theoretical results in Section 2. Comparing the CMH test, the actual type I error of the chi-square test are similar when $|p_2 - p_1| \leq 0.1$. Consistent with the asymptotic results, the chi-square test becomes more conservative when the proportion of sample size before the change is $\geq 20\%$, or the change in the response rate is greater than 10%.

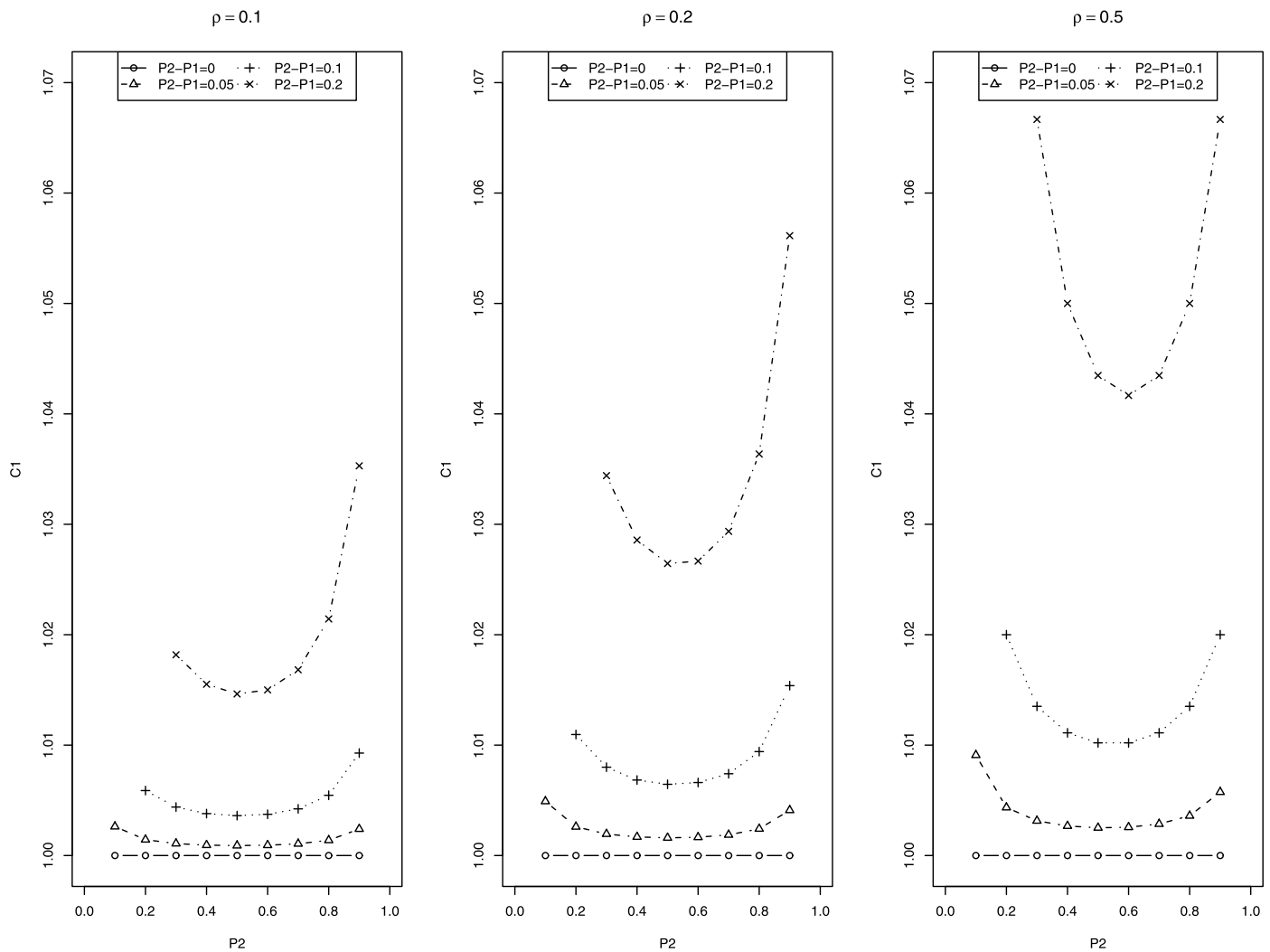


Figure 1. Constant C_1 for different p_1 and p_2 .

In addition to the sample size of $n = 157$, we have also computed the type I error for varying n s, the results are generally consistent with those in Figure 2. In cases with the small ($n = 50$) to moderate ($n = 200$) sample size, the type I error of chi-square test and a CMH test is similar until $p_2 - p_1 \geq 0.2$. Due to the discreteness of multinomial distribution, however, both chi-square and CMH test may result in a type I error as large as 0.056 for some n s, when $p_1 \approx p_2 \approx 0.5$.

4. CONCLUSION AND DISCUSSION

This article is intended to address the type I error issue when homogeneity assumption may be violated, particularly with a binary response variable. Due to the complexity of the contemporary trial designs, it is unrealistic to assume a homogenous population for patients enrolled over the entire enrollment.

Both asymptotic and numerical results have demonstrated that even when response rate changes during the course of a trial, the chi-squared test based on combining data before and after the change is conservative in terms of control of type I error, when n is large.

As pointed out by Garside and Mack (1976), the chi-square test is “approximate since it replaces multinomial probabilities by a continuous multi-normal function” and “the uncorrected chi-squared test gives actual error probabilities which usually exceed α for some value of p ”. The same remark applies to the CMH test. Our actual type I error computation showed that the type I error of both tests can be as large as 0.056 for some n and $p_1 \approx p_2 \approx 0.5$. However, when $p_1 \neq p_2$, the error probability is asymptotically less than the error probability of the homogeneity case with $p_1 = p_2$. All these observations and the asymptotic results lead us to conclude that the type I error of the chi-square test is generally controlled (if not conservative) even

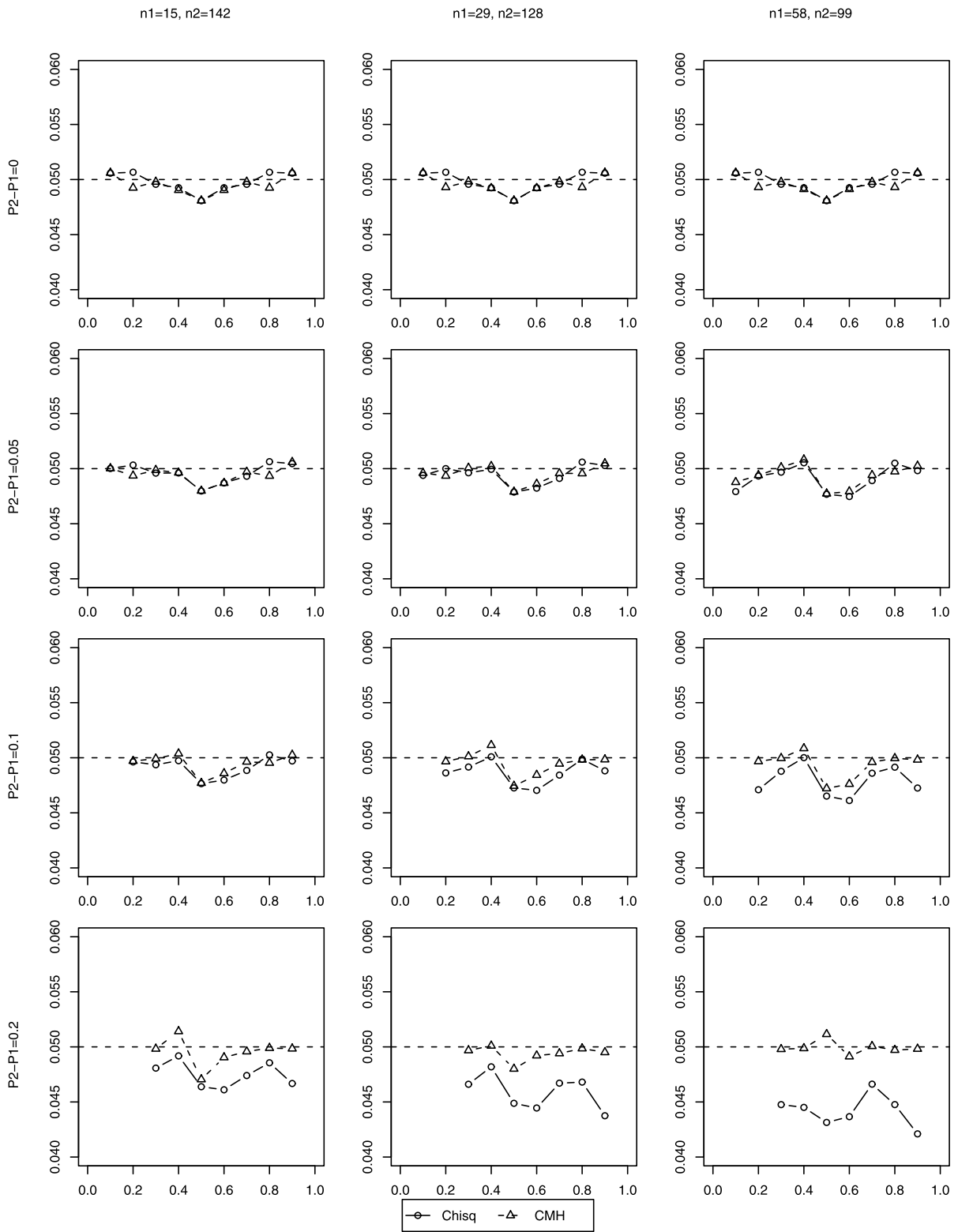


Figure 2. Actual Type I error of chi-square and CMH test for the introductory example when $|p_1 - p_2| = 0, 0.1, 0.2$, where x-axis represents the value of p_2 and y-axis represents the value of actual type I error.

when the response rate may change in the middle of the trial.

However, the conservative nature of the Chi-square test in such cases would result in a reduced power under the assumption of homogeneous treatment effect (i.e., same odds ratio irrespective of the change). Theoretically, the CMH test with the stratum before and after change is preferred when the response rate is considered different across the stratum. However, in many cases, it is not clear when the change occurred during the trial, thus, difficult to classify the changes as strata. For example, time trend is quite common in clinical studies (Altman and Royston, 1988, Duan and Hu, 2009, etc.), it is often impossible to classify the trend as strata. One cannot use the CMH test with the strata in those cases. In these cases, the simple Chi-square test could be applied.

Chow and Chang (2007, chapter 2) discussed power properties of statistical tests for clinical trials with continuous endpoints. It is important to study the properties of commonly used statistical tests when patient population changes during the conduct of clinical trials. There are several papers which discussed this issue for clinical trials with continuous responses (Chow and Shao, 2005; Feng, Shao and Chow, 2007; Losch and Neuhauser, 2008). In this paper, we discussed some properties of the simple Chi-square test with binary response. It remains a further research problem to study how to adjust the Chi-square test to increase its power when patient population changes.

In clinical trials involving important covariates (prognostic factors), stratified randomization procedures are usually used to balance treatments. Rosenberger and Sverdlov (2008) provided a good review of both the design and statistical inference about covariate-adaptive randomized clinical trials (Taves, 1974; Pocock and Simon, 1975; etc.) and covariate-adjusted response-adaptive randomized clinical trials (Zhang, Hu, Cheung and Chan, 2007). As stated in the paper, very little is known about the theoretical properties of the covariate-adaptive randomization procedures. The response probabilities may vary in different strata. When the number of strata is relative large and the sample size in each stratum is small, CMH test may not be a good test. This is because the value of V_2 could be unstable. In these cases, a simple Chi-square test could be a better choice. The theoretical results in this paper can be extended to these situations. This remains a further research problem.

APPENDIX

Proof of Theorem 2.1. From Section 2, we have $E(S_1 - S/2) = 0$ and $Var(S_1 - S/2) = [n_1p_1(1-p_1) + n_2p_2(1-p_2)]/2$. Now we calculate the mean of V_1 (asymptotic mean). It is easy to see that

$$E(S) = E(X_{11} + X_{12} + X_{21} + X_{22}) = 2n_1p_1 + 2n_2p_2.$$

By the law of large number (Casella and Berger, 2002, page 232 and 235 for independent and identical distributed (iid) random variables, Shao, 2003, page 65 (Theorem 1.14) for independent random variables with finite expectations), we have

$$\begin{aligned} \frac{S}{n} &= \frac{X_{11} + X_{12} + X_{21} + X_{22}}{n} \\ &\sim \frac{2n_1p_1 + 2n_2p_2}{n} \rightarrow 2\rho p_1 + 2(1-\rho)p_2 \end{aligned}$$

in probability, when $n \rightarrow \infty$ and $n_1/n \rightarrow \rho \in (0, 1)$.

Substitute above results to V_1 , we have

$$\begin{aligned} V_1 &= S(2n - S)/[4(2n - 1)] \\ &\sim \frac{(n_1p_1 + n_2p_2)(n_1(1-p_1) + n_2(1-p_2))}{(2n - 1)}. \end{aligned}$$

When $n \rightarrow \infty$ and $n_1/n \rightarrow \rho$, it is not difficult to see that

$$\begin{aligned} &\frac{Var([(X_{11} - X_{12}) + (X_{21} - X_{22})]/2)}{n} \\ &= \frac{[n_1p_1(1-p_1) + n_2p_2(1-p_2)]}{2n} \\ &\rightarrow [\rho p_1(1-p_1) + (1-\rho)p_2(1-p_2)]/2. \end{aligned}$$

Because

$$\frac{S}{n} \rightarrow 2\rho p_1 + 2(1-\rho)p_2,$$

and Slutsky's Theorem (Bickel and Doksum, 1977, page 461), we also have

$$(4.1) \quad V_1/n \rightarrow (\rho p_1 + (1-\rho)p_2)[\rho(1-p_1) + (1-\rho)(1-p_2)]/2,$$

in probability.

By the Lindeberg's central limit theorem (for sum of independent random variables, not for independent and identical distributed (iid) random variables, Shao (2003), Theorem 1.15, page 67; also see appendix of Hu and Rosenberger, 2006, page 166 for more general results), we have

$$\frac{(S_1 - S/2)}{(n[\rho p_1(1-p_1) + (1-\rho)p_2(1-p_2)]/2)^{1/2}} \rightarrow N(0, 1)$$

in distribution. Therefore

$$(4.2) \quad \frac{(S_1 - S/2)^2}{(n[\rho p_1(1-p_1) + (1-\rho)p_2(1-p_2)]/2)} \rightarrow \chi_{(1)}^2$$

in distribution, where $\chi_{(1)}^2$ represents Chi-squared distribution with 1 degree of freedom.

Let

$$C_1 = \frac{(\rho p_1 + (1-\rho)p_2)[\rho(1-p_1) + (1-\rho)(1-p_2)]}{[\rho p_1(1-p_1) + (1-\rho)p_2(1-p_2)]}.$$

Therefore,

$$\begin{aligned} C_1 T_{chi} &= \frac{(\rho p_1 + (1 - \rho)p_2)[\rho(1 - p_1) + (1 - \rho)(1 - p_2)]}{[\rho p_1(1 - p_1) + (1 - \rho)p_2(1 - p_2)]} \\ &\quad \times \frac{(S_1 - S/2)^2}{V_1} \\ &= \frac{(S_1 - S/2)^2}{(n[\rho p_1(1 - p_1) + (1 - \rho)p_2(1 - p_2)]/2)} \\ &\quad \times \frac{n(\rho p_1 + (1 - \rho)p_2)[\rho(1 - p_1) + (1 - \rho)(1 - p_2)]/2}{V_1}. \end{aligned}$$

From the result (4.1), we have

$$\frac{n(\rho p_1 + (1 - \rho)p_2)[\rho(1 - p_1) + (1 - \rho)(1 - p_2)]/2}{V_1} \rightarrow 1$$

in probability. Based on (4.2) and Slutsky's Theorem, we then have

$$C_1 T_{chi} \rightarrow \chi_{(1)}^2$$

in distribution, when $n \rightarrow \infty$ and $n_1/n \rightarrow \rho$. \square

Proof of Lemma 2.1. Because

$$(\rho p_1 + (1 - \rho)p_2)[\rho(1 - p_1) + (1 - \rho)(1 - p_2)] \geq 0$$

and

$$[\rho p_1(1 - p_1) + (1 - \rho)p_2(1 - p_2)] \geq 0,$$

we just need to prove

$$\begin{aligned} f(\rho) &= (\rho p_1 + (1 - \rho)p_2)[\rho(1 - p_1) + (1 - \rho)(1 - p_2)] \\ &\quad - [\rho p_1(1 - p_1) + (1 - \rho)p_2(1 - p_2)] \\ &\geq 0 \end{aligned}$$

for all $p_1, p_2 \in (0, 1)$ and $\rho \in (0, 1)$. To do this, we calculate the derivative of f first. After some simple calculation, we have

$$f'(\rho) = \frac{\partial f}{\partial \rho} = (p_1 - p_2)^2(1 - 2\rho).$$

It is easy to see that $f'(\rho) \geq 0$ for $\rho \in (0, 1/2)$ and $f'(\rho) \leq 0$ for $\rho \in (1/2, 1)$. Also $f'(\rho = 1/2) = 0$. Therefore $f(\rho)$ is increasing for $\rho \in (0, 1/2)$ and then decreasing for $\rho \in (1/2, 1)$. Now because $f(\rho = 0) = 0$ and $f(\rho = 1) = 0$, we have $f(\rho) \geq 0$. Therefore, $C_1 \geq 1$. \square

ACKNOWLEDGEMENTS

Special thanks go to the anonymous referee and the editor for their constructive comments, which led to a much improved version of the paper.

Received 12 June 2010

REFERENCES

- ALTMAN, D. G. and ROYSTON, J. P. (1988). The hidden effect of time. *Statistics in Medicine* **7** 629–637.
- BAI, Z. D. and HU, F. (1999). Asymptotic theorems for urn models with nonhomogeneous generating matrices. *Stochastic Process. Appl.* **80** 87–101. [MR1670107](#)
- BICKEL, P. J. and DOKSUM, K. A. (1977). *Mathematical Statistics: Basic Ideas and Selected Topics*. Prentice Hall, New Jersey. [MR0443141](#)
- CHOW, S. C. and CHANG, M. (2007). *Adaptive Design Methods in Clinical Trials*. Chapman & Hall/CRC.
- CHOW, S. C. and SHAO, J. (2005). Inference for clinical trials with some protocol amendments. *Journal of Biopharmaceutical Statistics* **15** 659–666. [MR2190576](#)
- COAD, D. S. (1991). Sequential tests for an unstable response variable. *Biometrika* **78** 113–121. [MR1118236](#)
- DUAN, L. and HU, F. (2009). Doubly adaptive biased coin designs with heterogeneous responses. *Journal of Statistical Planning and Inference* **139** 3220–3230. [MR2535195](#)
- FENG, H., SHAO, J., and CHOW, S. C. (2007). Adaptive group sequential test for clinical trials with changing patient population. *Journal of Biopharmaceutical Statistics* **17** 1227–1238. [MR2414572](#)
- GARSDIE, G. R. and MACK, C. (1976). Actual type 1 error probabilities for various tests in the homogeneity case of the 2×2 contingency table. *The American Statistician* **30** 18–21.
- HU, F. and ROSENBERGER, W. F. (2000). Analysis of time trends in adaptive designs with application to a neurophysiology experiment. *Statistics in Medicine* **19** 2067–2075.
- HU, F. and ROSENBERGER, W. F. (2006). *The Theory of Response-Adaptive Randomization in Clinical Trials*. John Wiley and Sons, Inc., New York. [MR2245329](#)
- LOSCH, C. and NEUHAUSER, M. (2008). The statistical analysis of a clinical trial when a protocol amendment changed the inclusion criteria. *BMC Medical Research Methodology* **8** 16.
- MANTEL, N. and HAENSZEL, W. (1959). Statistical aspects of the analysis of data from retrospective studies of disease. *Journal of the National Cancer Institute* **22** 719–748.
- POCOCK, S. J. and SIMON, R. (1975). Sequential treatment assignment with balancing for prognostic factors in the controlled clinical trial. *Biometrics* **31** 103–115.
- ROSENBERGER, W. F. and SVERDLOV, O. (2008). Handling covariates in the designs of clinical trials. *Statistical Science* **23** 404–419. [MR2483911](#)
- SHAO, J. (2003). *Mathematical Statistics*. Springer. [MR2002723](#)
- STOKES, M. E., DAVIS, C. S. and KOCH, G. G. (2000). *Categorical Data Analysis Using the SAS System*. SAS Publishing.
- TAVES, D. R. (1974). Minimization: A new method of assigning patients to treatment and control groups. *Clin. Pharmacol. Therap.* **15** 443–453.
- ZHANG, L.-X., HU, F., CHEUNG, S. H. and CHAN, W. S. (2007). Asymptotic properties of covariate-adjusted response-adaptive designs. *The Annals of Statistics* **35** 1166–1182. [MR2341702](#)

Feifang Hu
 Department of Statistics
 University of Virginia
 Charlottesville, VA 22904
 USA
 E-mail address: fh6e@virginia.edu

Jiandong Lu
Clinical Biostatistics
Janssen Research and Development
Springhouse, PA
USA
E-mail address: jlu5@its.jnj.com

Feng Tai
Department of Biostatistics
University of Minnesota
Minneapolis, MN
USA
E-mail address: fengtai@biostat.umn.edu