JP Journal of Biostatistics **1**(1) (2007), 13-23

TESTS FOR EQUIVALENCE BASED ON PROPORTION RATIO IN CATEGORICAL PAIRED-SAMPLE DATA

KUNG-JONG LUI and WILLIAM G. CUMBERLAND

Department of Mathematics and Statistics
San Diego State University
San Diego, CA 92182-7720, U. S. A.
e-mail: kjl@rohan.sdsu.edu

Department of Biostatistics
University of California
Los Angeles, CA 90095, U. S. A.

Abstract

It is not uncommon that we encounter the situations in which our goal is to establish an equivalence rather than to detect a difference between the distributions of the two comparison groups. When the underlying proportions vary substantially between categories, we may wish to consider use of the proportion ratio (PR) in establishing equivalence. This paper develops test procedures for detecting equivalence based on the PR in paired-sample data for both nominal and ordinal scales. This paper further develops procedures to accommodate the case of testing for symmetric equivalence. Finally, this paper includes examples to illustrate the practical use of the procedures developed here.

1. Introduction

It is not uncommon that we encounter categorical paired-sample data in which it is much more interesting to test whether the distributions

2000 Mathematics Subject Classification: 62P10.

Keywords and phrases: equivalence test, proportion ratio, paired-sample, intersection-union test.

Received November 4, 2006

between two comparison groups are equivalent rather than equal. For example, consider the data (Table 1) consisting of 55981 subjects recording migration in the U.S. taken by Bureau of the Census (Agresti [1]). We classify each sampled subject according to the geographical area (Northeast, Midwest, South and West) of his/her residence in the years of 1980 and 1985. The residence locations in different years for the same subject are correlated and hence naturally form paired-sample data. When analyzing data with such a large sample size, we would expect to obtain a strongly significant finding of testing equality even for a tiny difference in the distributions of residence locations between these years. Therefore, testing the marginal homogeneity of the distributions regarding residence locations becomes a rather uninteresting problem. On the other hand, it can be of interest to find out whether the distributions of residence locations are equivalent between 1980 and 1985. As a second example, consider the data (Table 2) regarding the ordinal measurements (the highest grade, the second grade, the third grade, the lowest grade) of unaided distance vision on eyes over 7477 women (Agresti [1]; Stuart [15]). The measurements on eyes of the same subjects again naturally form paired-sample data. It is of importance to study whether the distributions of the unaided distance vision between the two eyes are equivalent. Note that Lui and Cumberland [12] considered use of the simple difference to establish equivalence for ordinal data with matched pairs. Although simple difference is probably the most commonly-used measure in clinical trials, we need to use different maximum acceptance level to define equivalence with respect to simple difference for different categories when the categorical proportions vary substantially. This may sometimes cause practical difficulty due to how to select the maximum acceptance levels for various categories. To alleviate this concern, we focus our discussion on use of the proportion ratio (PR) rather than simple difference for testing equivalence here.

In this paper, we have developed test procedures for testing equivalence with respect to the PR in paired-sample data. We have further developed procedures for testing equivalence between the cell proportions in a symmetric pattern for both nominal and ordinal data. Numerous papers on testing equivalence appear elsewhere (Dunnett and Gent [3]; Westlake [16]; Hauck and Anderson [4, 5]; Liu and Chow [7]; Lu and Bean [8]; Lui [9, 10, 13]; Nam [14]). None of them focuses discussion, however, on hypothesis testing in the situation described here.

2. Hypothesis Testing Procedures

Assume that our data consist of n matched pairs of responses, each response falling into exactly one of K categories. Assume further that the two responses within each pair correspond to the two comparison groups under investigation. Therefore, we can summarize our data in a $K \times K$ table with the rows corresponding to group one, and the columns corresponding to group two. Suppose that we want to test equivalence between the distributions of responses for the two comparison groups. Let n_{ij} (where $\sum_i \sum_j n_{ij} = n$) denote the observed frequency in cell (i, j) with the cell probability p_{ij}, where i and $j = 1, 2, \ldots, K$. Then, the random vector $\mathbf{n}' = (n_{11}, n_{12}, \ldots, n_{1K}, n_{21}, n_{22}, \ldots, n_{2K}, \ldots, n_{K1}, n_{K2}, \ldots, n_{KK})$ follows a multinomial distribution with parameters n and the probability vector $\mathbf{p}' = (p_{11}, p_{12}, \ldots, p_{1K}, p_{21}, p_{22}, \ldots, p_{2K}, \ldots, p_{K1}, p_{K2}, \ldots, p_{KK})$. We define the marginal probabilities: $p_{i.} = \sum_k p_{ik}$ and $p_{.i} = \sum_k p_{ki}$ for $i = 1, 2, \ldots, K$. Note that $p_{i.}$ and $p_{.i}$ simply represent the probabilities of a randomly selected subject falling into category i for the two comparison groups, respectively.

2.1. Equivalence testing for nominal data

First, note that if the distributions of the marginal probabilities were equivalent, we would expect $-\Delta < (p_{i.} - p_{.i})/p_{i.} < \Delta$, where $\Delta(> 0)$ is the maximum acceptable level and is predetermined by the investigator. Note that $-\Delta < (p_{i.} - p_{.i})/p_{i.} < \Delta$ if and only if $\delta_l < p_{.i}/p_{i.} < \delta_u$, where $\delta_l = 1 - \Delta$ and $\delta_u = 1 + \Delta$. Therefore, we consider testing the null hypothesis $H_0 : p_{.i}/p_{i.} \geq \delta_u$ or $p_{.i}/p_{i.} \leq \delta_l$ for some i $(i = 1, 2, \ldots, K)$ versus the alternative hypothesis $H_a : \delta_l < p_{.i}/p_{i.} < \delta_u$ for all i $(i = 1, 2, \ldots, K)$. When we reject the null hypothesis at a nominal α-level, we can state that the two marginal distributions for the two comparison groups are equivalent. Let $\hat{p}_{ij} = n_{ij}/n$ denote the cell sample proportion, which is, in fact, the maximum likelihood estimator (MLE) of p_{ij}. We define $\hat{p}_{i.} = \sum_k \hat{p}_{ik}$ and $\hat{p}_{.i} = \sum_k \hat{p}_{ki}$. Note that the MLE of $p_{.i}/p_{i.}$ is $\hat{p}_{.i}/\hat{p}_{i.}$, for which the sampling distribution is generally skewed, especially when the sample size n is not large. To improve the normal approximation to the statistic $\hat{p}_{.i}/\hat{p}_{i.}$, we consider use of the logarithmic transformation (Katz et al. [6]; Lui [11]). Using the delta method (Casella and Berger [2]), we obtain the estimated asymptotic variance $\log(\hat{p}_{.i}/\hat{p}_{i.})$ to be

$$\hat{Var}(\log(\hat{p}_{.i}/\hat{p}_{i.})) = (\hat{p}_{i.} + \hat{p}_{.i} - 2\hat{p}_{ii})/(n\hat{p}_{.i}\hat{p}_{i.}). \tag{1}$$

Define $N_i(\delta_s) = [(\log(\hat{p}_{.i}/\hat{p}_{i.})) - \log(\delta_s)]/(\hat{Var}(\log(\hat{p}_{.i}/\hat{p}_{i.})))^{1/2}$, for $s = l$ and u. On the basis of the intersection-union principle (Casella and Berger [2]), for a given large sample size n, we would reject H_0 at the nominal α-level if

$$N_i(\delta_u) < -Z_\alpha \text{ and } N_i(\delta_l) > Z_\alpha \tag{2}$$

held for all i $(i = 1, 2, ..., K)$, where Z_α is the upper 100(α)th percentile of the standard normal distribution.

Note that it may also be interesting to find out whether there is an equivalence in a symmetric pattern of the response proportions between the two comparison groups. In other words, we want to test the null hypothesis $H_0 : p_{ji}/p_{ij} \geq \delta_u$ or $p_{ji}/p_{ij} \leq \delta_l$ for some (i, j) versus the alternative hypothesis $H_a : \delta_l < p_{ji}/p_{ij} < \delta_u$ for all (i, j), where $(j > i, i$ and $j = 1, 2, ..., K)$. Again, employing the delta method, we obtain the estimated asymptotic variance of $\log(\hat{p}_{ji}/\hat{p}_{ij})$ to be

$$\hat{Var}(\log(\hat{p}_{ji}/\hat{p}_{ij})) = (\hat{p}_{ij} + \hat{p}_{ji})/(n\hat{p}_{ji}\hat{p}_{ij}). \tag{3}$$

Define $R_{ij}(\delta_s) = [(\log(\hat{p}_{ji}/\hat{p}_{ij})) - \log(\delta_s)]/(\hat{Var}(\log(\hat{p}_{ji}/\hat{p}_{ij})))^{1/2}$, for $s = l$ and u. We would reject H_0 at the nominal α-level if

$$R_{ij}(\delta_u) < -Z_\alpha \text{ and } R_{ij}(\delta_l) > Z_\alpha \tag{4}$$

held for all (i, j), where $(i = 1, 2, ..., K-1$, and $j = i+1, i+2, ..., K)$.

2.2. Equivalence testing for ordinal data

When the underlying responses are on an ordinal scale, it is more natural and appealing to consider using the tail distributions (or equivalently, the cumulative distributions), accounting for the ordered responses in the data, to establish equivalence between the two comparison groups. Thus, we define $p_{(i).} = \sum_{k=i}^{K} p_{k.}$ and $p_{.(i)} = \sum_{k=i}^{K} p_{.k}$ for $i = 1, 2, \ldots, K$. We define equivalence between the two tail marginal distributions when $\delta_l < p_{.(i)}/p_{(i).} < \delta_u$ holds for all i. Note that when $i = 1$, both $\hat{p}_{(i).}$ and $\hat{p}_{.(i)}$ are, by definition, equal to 1 and hence $p_{.(1)}/p_{(1).}$ always falls between δ_l and δ_u. We consider testing $H_0 : p_{.(i)}/p_{(i).} \geq \delta_u$ or $p_{.(i)}/p_{(i).} \leq \delta_l$ for some $i\,(i = 2, \ldots, K)$ versus $H_a : \delta_l < p_{.(i)}/p_{(i).} < \delta_u$ for all $i\,(i = 2, \ldots, K)$. When we reject the null hypothesis, we state that the two tail marginal distributions of the two comparison groups are equivalent.

Following the functional invariance property, we can see that the MLE of $p_{.(i)}/p_{(i).}$ is $\hat{p}_{.(i)}/\hat{p}_{(i).}$, where $\hat{p}_{(i).} = \sum_{k=i}^{K} \hat{p}_{k.}$ and $\hat{p}_{.(i)} = \sum_{k=i}^{K} \hat{p}_{.k}$. Again, to improve the normal approximation to the statistic $\hat{p}_{.(i)}/\hat{p}_{(i).}$, we consider use of the logarithmic transformation. We obtain the estimated asymptotic variance $\log(\hat{p}_{.(i)}/\hat{p}_{(i).})$ to be

$$\hat{Var}(\log(\hat{p}_{.(i)}/\hat{p}_{(i).})) = (\hat{p}_{(i).} + \hat{p}_{.(i)} - 2\hat{p}_{(i)(i)})/(n\hat{p}_{(i).}\hat{p}_{.(i).}), \tag{5}$$

where $\sum_{k=i}^{K}\sum_{k'=i}^{K} \hat{p}_{kk'}$, $(i = 2, \ldots, K)$.

Define the test statistic

$$Q_i(\delta_s) = [\log(\hat{p}_{.(i)}/\hat{p}_{(i).}) - \log(\delta_s)]/(\hat{Var}(\log(\hat{p}_{.(i)}/\hat{p}_{(i).})))^{1/2}$$

for $s = l$ and u. Therefore, for a given large sample size n, we would reject H_0 at the nominal α-level if

$$Q_i(\delta_u) < -Z_\alpha \text{ and } Q_i(\delta_l) > Z_\alpha \tag{6}$$

held for all $i\,(i = 2, \ldots, K)$.

To test whether there is a symmetric pattern of equivalence in the ordinal data, we define $p_{i(j)} = \sum_{k=j}^{K} p_{ik}$ and $p_{(j)i} = \sum_{k=j}^{K} p_{ki}$ for i and $j = 1, 2, ..., K$. When the inequalities: $\delta_l < p_{(j)i}/p_{i(j)} < \delta_u$ hold for all i and $j = 1, 2, ..., K$, we define there is an equivalence between the two comparison groups in a symmetric pattern. We consider testing $H_0 : p_{(j)i}/p_{i(j)} \geq \delta_u$ or $p_{(j)i}/p_{i(j)} \leq \delta_l$ for some (i, j) versus the alternative hypothesis $H_a : \delta_l < p_{(j)i}/p_{i(j)} < \delta_u$ for all (i, j). Note that by definition $p_{i(1)} = p_{i.}$ and $p_{(1)i} = p_{.i}$. Thus, a symmetric equivalence pattern defined here for the ordinal data would imply an equivalence for the marginal proportions: $\delta_l < p_{.i}/p_{i.} < \delta_u$ for $i = 1, 2, ..., K$. By employing the delta method, we obtain the estimated variance to be

$$\hat{V}ar(\log(\hat{p}_{(j)i}/\hat{p}_{i(j)})) = (\hat{p}_{(j)i} + \hat{p}_{i(j)})/(n\hat{p}_{(j)i}\hat{p}_{i(j)}) \text{ for } j > i \text{ and}$$

$$= (\hat{p}_{(j)i} + \hat{p}_{i(j)} - 2\hat{p}_{ii})/(n\hat{p}_{(j)i}\hat{p}_{i(j)}) \text{ for } j \leq i, \qquad (7)$$

where $\hat{p}_{i(j)} = \sum_{k=j}^{K} \hat{p}_{ik}$ and $\hat{p}_{(j)i} = \sum_{k=j}^{K} \hat{p}_{ki}$.

Define $T_{ij}(\delta_s) = [\log(\hat{p}_{(j)i}/\hat{p}_{i(j)}) - \log(\delta_s)]/(\hat{V}ar(\log(\hat{p}_{(j)i}/\hat{p}_{i(j)})))^{1/2}$ for $s = l$ and u. We would reject the null hypothesis H_0 at the nominal α-level if

$$T_{ij}(\delta_u) < -Z_\alpha \text{ and } T_{ij}(\delta_l) > Z_\alpha \qquad (8)$$

held for all (i, j), where i and $j = 1, 2, ..., K$.

3. Examples

To illustrate the use of test procedures (2) and (4), we consider the data (Table 1) consisting of $(n =)$ 55981 subjects about their residence locations in 1980 and 1985 over the geographic areas: Northeast, Midwest, South and West (Agresti [1]). Table 1 summarizes the observed frequency and the sampled proportions $\hat{p}_{ij}$ of subjects, as well as the corresponding marginal proportions $\hat{p}_{i.}$ and $\hat{p}_{.i}$. For illustration purposes, suppose we choose the maximum acceptable level of Δ to be 0.20 and wish to determine whether the two marginal distributions of residence locations are equivalent over different geographical areas between 1980 and 1985 (i.e., $0.80 < p_{.i}/p_{i.} < 1.20$ for all $i = 1, 2, 3, 4$). When applying test procedure (2), we find the values of test statistics $N_i(1.2) < -1.645$ and $N_i(0.80) > 1.645$ for all $i = 1, 2, 3, 4$. Thus, we conclude that there is a significant evidence at 5% to support that the distributions of the residence locations in 1980 and 1985 are essentially equivalent. This conclusion is certainly consistent with the fact that the two marginal proportions between years of 1980 and 1985 under consideration are similar to each other: $(\hat{p}_{1.}, \hat{p}_{.1}) = (0.218, 0.213)$, $(\hat{p}_{2.}, \hat{p}_{.2}) = (0.260, 0.253)$, $(\hat{p}_{3.}, \hat{p}_{.3}) = (0.330, 0.339)$, $(\hat{p}_{4.}, \hat{p}_{.4}) = (0.191, 0.194)$; these give the sample proportion ratio $\hat{p}_{.i}/\hat{p}_{i.}$ to be $0.977, 0.973, 1.027$, and 1.016, which all lie around of the ratio 1.0. To further investigate whether there is an equivalence between p_{ij} and p_{ji} in a symmetric pattern, we employ test procedure (4). We have found that there is no significant evidence to reject $H_0 : p_{ji}/p_{ij} \geq 0.80$ or $p_{ji}/p_{ij} \leq 1.20$ for some (i, j) at 5% level and hence the direction of movement is probably not equivalent with respect to symmetric pattern. In fact, this finding is consistent with the observation that the number of subjects moving from Northeast to the West is almost twice of those moving from the West to the Northeast (Table 1).

To illustrate the use of procedures (6) and (8), we consider the data (Table 2) regarding the ordinal measurements (highest grade, second grade, third grade and lowest grade) of unaided distance vision on eyes over 7477 women with ages ranging from 30 and 39 years old (Agresti [1]; Stuart [15]). Suppose that we wish to study if the ratio between the tail marginal proportion of the left eyes relative to that of the right eyes at the ordinal measurements i, $p_{.(i)}/p_{(i).}$, falls into the acceptable range [0.80, 1.20] for all i. When applying test procedure (6), we find out there is a significant evidence to claim that the two tail marginal proportions are equivalent between the left and the right eyes. Again, this finding is consistent with the observations on the empirical proportion ratios $(p_{.(i)}/p_{(i).}$ for $i = 1, 2, 3, 4)$: 1.0, 1.01, 1.03, and 1.06, that are all in the neighborhood of 1.0. When applying test procedure (8) to detect the symmetric equivalence, we have found that there is no evidence to support that the measurements of unaided distance vision on the two eyes are symmetrically equivalent. This is because, for example, the ratio $\hat{p}_{14}/\hat{p}_{41} = 1.83$ is much larger than the upper acceptable limit $\delta_u (= 1.20)$ for equivalence.

4. Discussion

First, note that $\mathbf{p}$ is in $\{\mathbf{p} \mid p_{i.} = p_{.i}$ for all $i = 2, ..., K\}$ if and only if $\mathbf{p}$ is in $\{\mathbf{p} \mid p_{(i).} = p_{.(i)}$ for all $i = 2, ..., K\}$. Thus, testing hypothesis of the marginal homogeneity using the marginal proportions is equivalent to that using the tail distribution of the marginal proportions. By contrast, we can easily show that the set of inequalities: $\delta_l < p_{.i}/p_{i.} < \delta_u$ for $i = 1, 2, 3, ..., K$, implies that the set of inequalities: $\delta_l < p_{.(i)}/p_{(i).} < \delta_u$ for $i = 2, 3, ..., K$, but the converse is no longer true. Thus, an equivalence with respect to the marginal proportions would implicitly suggest that an equivalence with respect to the tail distribution of the marginal proportions, but not vice versa.

Note also that the inequality: $1 - \Delta < p_{.i}/p_{i.} < 1 + \Delta$ is not exactly the same as the inequality: $1 - \Delta < p_{i.}/p_{.i} < 1 + \Delta$. Thus, when considering use of the ratio in establishing equivalence, we need to decide which parameter $p_{.i}/p_{i.}$ and $p_{i.}/p_{.i}$ for use. However, because the value Δ is usually chosen to be small, $1/(1 - \Delta) \approx 1 + \Delta$ and $1/(1 + \Delta) \approx 1 - \Delta$. Therefore, the concern of choosing which ratio $p_{.i}/p_{i.}$ or $p_{i.}/p_{.i}$ for use should not be an issue of important concern in practice.

As noted elsewhere (Hauck and Anderson [5]; Lui [9]), we note that the test procedures proposed here are actually equivalent to the procedure, in which we state that there is an equivalence whenever the corresponding 100(1-2α) percent confidence interval of PR are all contained in the acceptable range $[\delta_l, \delta_u]$. For example, consider use of test procedure (2), in which we reject H_0 at the nominal α-level if $N_i(\delta_u) < -Z_\alpha$ and $N_i(\delta_l) > Z_\alpha$ held for all i ($i = 1, 2, ..., K$). Note that we can easily show that the conditions: $N_i(\delta_u) < -Z_\alpha$ and $N_i(\delta_l) > Z_\alpha$ hold for all i if and only if the corresponding 100(1-2α) percent confidence $[\hat{p}_{.i}\exp(-Z_\alpha(\hat{V}ar(\log(\hat{p}_{.i}/\hat{p}_{i.})))^{1/2})/\hat{p}_{i.},\ \hat{p}_{.i}\exp(Z_\alpha(\hat{V}ar(\log(\hat{p}_{.i}/\hat{p}_{i.})))^{1/2})/\hat{p}_{i.}]$ lie entirely in the acceptable range $[\delta_l,\ \delta_u]$ for the ratio $p_{.i}/p_{i.}$. Therefore, when all the resulting 90% confidence interval of $p_{.i}/p_{i.}$ are contained in $[\delta_l, \delta_u]$, we may then reject $H_0 : p_{.i}/p_{i.} \geq \delta_u$ or $p_{.i}/p_{i.} \leq \delta_l$ at 0.05-level.

In summary, this paper has developed test procedures for testing equivalence with respect to the proportion ratios between categories. This paper has further showed that we can easily develop test procedures to accommodate the case for detecting a symmetric equivalence as well. This paper has also included examples to illustrate the practical usefulness of these procedures developed here. The results and the findings presented here should have use for biostatisticians and epidemiologists when they wish to do hypothesis testing for equivalence in categorical paired-sample data.

Table 1. The frequency and the frequency proportions (in parenthesis) of residence in years of 1980 and 1985 for a sample selected by the U. S. Bureau of the Census, Current Population Reports, Series P-20, No. 420, Geographical Mobility: 1985, U. S. Government Printing office, Washington, D. C.

1980 | 1985	Northeast	Midwest	South	West	Marginal Distribution
Northeast	11607 (0.207)	100 (0.002)	366 (0.007)	124 (0.002)	12197 (0.218)
Midwest	87 (0.002)	13677 (0.244)	515 (0.009)	302 (0.005)	14581 (0.260)
South	172 (0.003)	225 (0.004)	17819 (0.318)	270 (0.005)	18486 (0.330)
West	63 (0.001)	176 (0.003)	286 (0.005)	10192 (0.182)	10717 (0.191)
Marginal Distribution	11929 (0.213)	14178 (0.253)	18986 (0.339)	10888 (0.194)	55981 (1.00)

Table 2. The frequency and the frequency proportions (in parenthesis) of measurements on eyes over a sample of 7477 women with ages ranging from 30 to 39 years old

Right | Left Eyes | Eyes	Highest Grade	Second Grade	Third Grade	Lowest Grade	Marginal Distribution	Tail Distribution
Highest Grade	1520 (0.203)	266 (0.036)	124 (0.017)	66 (0.009)	1976 (0.264)	7477 (1.00)
Second Grade	234 (0.031)	1512 (0.202)	432 (0.058)	78 (0.010)	2236 (0.302)	5501 (0.736)
Third Grade	117 (0.016)	362 (0.048)	1772 (0.237)	205 (0.027)	2456 (0.328)	3245 (0.434)
Lowest Grade	36 (0.005)	82 (0.011)	179 (0.024)	492 (0.066)	789 (0.106)	789 (0.106)
Marginal Distribution	1907 (0.255)	2222 (0.297)	2507 (0.335)	841 (0.112)	7477 (1.00)	
Tail Distribution	1907 (1.00)	5570 (0.744)	3348 (0.447)	841 (0.112)		

References

[1] A. Agresti, Categorical Data Analysis, Wiley & Sons, New York, 1990.

[2] G. Casella and R. L. Berger, Statistical Inference, Wadsworth, Belmont, CA, 1990.

[3] C. W. Dunnett and M. Gent, Significance testing to establish equivalence between treatments with special reference to data in the form of 2×2 tables, Biometrics 33 (1977), 593-602.

[4] W. W. Hauck and S. Anderson, A new statistical procedure for testing equivalence in two-group comparative bioavailability trials, J. Pharmacokinetics and Biopharmaceutics 12 (1984), 83-91.

[5] W. W. Hauck and S. Anderson, A proposal for interpreting and reporting negative studies, Statistics in Medicine 5 (1986), 203-209.

[6] D. Katz, J. Baptista, S. P. Azen and M. C. Pike, Obtaining confidence intervals for the risk ratio in cohort studies, Biometrics 34 (1978), 469-474.

[7] J. P. Liu and S.-C. Chow, Sample size determination for the two one-sided tests procedure in bioequivalence, J. Pharmacokinetics and Biopharmaceutics 20 (1992), 101-104.

[8] Y. Lu and J. Bean, On the sample size for one-sided equivalence of sensitivities based upon McNemar's test, Statistics in Medicine 14 (1995), 1831-1839.

[9] K.-J. Lui, Exact equivalence test for risk ratio and its sample size determination under inverse sampling, Statistics in Medicine 16 (1997a), 1777-1786.

[10] K.-J. Lui, Sample size determination for repeated measurements in bioequivalence test, J. Pharmacokinetics and Biopharmaceutics 25 (1997b), 507-513.

[11] K.-J. Lui, A note on interval estimation of the relative difference in data with matched pairs, Statistics in Medicine 17 (1998), 1509-1515.

[12] K.-J. Lui and W. G. Cumberland, A test procedure of equivalence in ordinal data with matched-pairs, Biometrical J. 43 (2001), 977-983.

[13] K.-J. Lui, Sample size calculation for testing non-inferiority and equivalence under Poisson distribution, Statistical Methodology 2 (2005), 37-48.

[14] J.-M. Nam, Establishing equivalence of two treatments and sample size requirements in matched-pairs design, Biometrics 53 (1997), 1422-1430.

[15] A. Stuart, A test for homogeneity of the marginal distributions in a two-way classification, Biometrika 42 (1955), 412-416.

[16] W. J. Westlake, Statistical aspects of comparative bioavailability trials, Biometrics 35 (1979), 273-280.

■