Published in final edited form as:

Biometrics. 2011 December ; 67(4): 1666–1668. doi:10.1111/j.1541-0420.2011.01593.x.

A Note on Monotonicity Assumptions for Exact Unconditional Tests in Binary Matched-pairs Designs

Xiaochun Li^{*}, Mengling Liu^{**}, and Judith D. Goldberg^{***}

Division of Biostatistics, Department of Environmental Medicine, New York University School of Medicine, New York, U.S.A

Summary

Exact unconditional tests have been widely applied to test the difference between two probabilities for 2×2 matched-pairs binary data with small sample size. In this context, Lloyd (2008, *Biometrics* **64**, 716–723) proposed an $E + M$ p -value, that showed better performance than the existing M p -value and C p -value. However, the analytical calculation of the $E + M$ p -value requires that the Barnard convexity condition be satisfied; this can be challenging to prove theoretically. In this paper, by a simple reformulation, we show that a weaker condition, *conditional monotonicity*, is sufficient to calculate all three p -values (M , C and $E + M$) and their corresponding exact sizes. Moreover, this conditional monotonicity condition is applicable to non-inferiority tests.

Keywords

Barnard convexity; conditional monotonicity; exact unconditional test; matched-pairs; non-inferiority test

1. Introduction

Lloyd (2008) considered exact unconditional tests for binary matched-pairs data when the sample size is small. A 2×2 table (as Table 1 in Lloyd (2008)) can summarize this type of data in four random cell counts $\mathbf{X} = (X_{00}, X_{01}, X_{10}, X_{11})$ where X_{ij} , $i, j = 0, 1$, is the number of matched-pairs with responses i for Treatment 1 and j for Treatment 2. Vector $\mathbf{X}$ is assumed to have a multinomial distribution with probabilities $\boldsymbol{\pi} = (\pi_{00}, \pi_{01}, \pi_{10}, \pi_{11})$. The problem of interest is to test the difference between success probabilities in the two groups, i.e., $H_0: \pi_{1+} \leq \pi_{+1}$ versus $H_1: \pi_{1+} > \pi_{+1}$, where $\pi_{1+} = \pi_{11} + \pi_{10}$ for Treatment 1 and $\pi_{+1} = \pi_{11} + \pi_{01}$ for Treatment 2. This is equivalent to testing

$$H_0: \pi_{01} \geq \pi_{10} \text{ versus } H_1: \pi_{01} < \pi_{10}. \quad (1)$$

To calculate p -values (M , C and $E + M$) and their exact sizes for testing (1), Lloyd (2008) used the Barnard Convexity Condition (BCC) to reduce the two-dimensional nuisance parameter space to one dimension. However, a rigorous proof that the BCC is satisfied is not always available. Therefore, numerical methods are often used to verify the BCC; this is the case for the E p -value when calculating the $E + M$ p -value (Lloyd, 2008) and for the C p -value when calculating its exact size (Berger and Sidik, 2003).

* xiaochun.li@nyu.edu

** mengling.liu@nyu.edu

*** jd.goldberg@nyumc.org

In this paper, we show that the *conditional monotonicity condition (CMC)*, a weaker condition than *BCC*, is sufficient for the dimension reduction in the calculation of *p*-values (*M, C* and *E + M*) and their exact sizes for testing (1) and that this dimension reduction technique can be extended to the non-inferiority case.

2. The *BCC* and the *CMC*

Denote observed values of $\mathbf{X}$ by $(x_{00}, x_{01}, x_{10}, x_{11})$ and let $k = x_{10} + x_{01}$ be the number of discordant pairs and $x = x_{01}$. Following Lloyd (2008), we consider three statistics: McNemar's test statistic, the likelihood ratio test statistic and the conditional sign test statistic. For any of the three test statistics above, denoted by $T(x, k)$, the *BCC* (Lloyd, 2008) is satisfied if

$$T(x, k) \leq T(x+1, k) \quad \text{and} \quad T(x, k) \geq T(x, k+1) \quad (2)$$

for all (x, k) . That is, $T(x, k)$ is nondecreasing in x for any fixed k , and is nonincreasing in k for any fixed x . Note that Berger and Sidik (2003) used a slightly different definition of the *BCC*: $T(x_{01}, x_{10}) \triangleq T(x_{01}, x_{01} + x_{10})$ is nondecreasing in x_{01} and nonincreasing in x_{10} . It is easy to prove that this definition is stronger than Lloyd's definition in (2).

The conditional monotonicity condition (*CMC*) is defined as

$$T(x, k) \leq T(x+1, k), \quad (3)$$

which is only one of the constraints in (2), and thus, is a weaker condition than either Lloyd's *BCC* or Berger and Sidik's *BCC*. For example, while the test statistic $T(x, k) = (2x - k)/(k - \sqrt{k})$ satisfies the *CMC*, $T(x, k)$ does not satisfy the *BCC* by either definition. Note that McNemar's, the likelihood ratio and the conditional sign test statistics all satisfy the *CMC*.

3. The *p*-value calculation

Following Lloyd (2008), we reparameterize π_{01} and π_{10} by φ and η as $\varphi := \pi_{10} + \pi_{01}$ and $\eta := \pi_{01}/\varphi$. Then the null space $H_0 = \{(\pi_{01}, \pi_{10}): \pi_{01} \geq \pi_{10}\} = \{(\varphi, \eta): 0 \leq \varphi \leq 1, \eta \geq 1/2\}$. When the test statistic $T(x, k)$ satisfies the *CMC*, let $h(t, k) = \sup\{x: T(x, k) \leq t\}$, which is also nondecreasing in t . We define $h(t, k)$ to be -1 if the supremum does not exist. Using the definition of $h(t, k)$, for a given observed statistic t_{obs} , the rejection set can be rewritten as $\{(x', k'): T(x', k') \leq t_{obs}\} = \{(x', k'): 0 \leq k' \leq n, 0 \leq x' \leq h(t_{obs}, k')\}$.

3.1 *M p*-value and *C p*-value

We adopt the notion that smaller values of test statistics favor H_1 and note that the probability mass function of (X, K) can be decomposed into $P_{(\varphi, \eta)}(X = x, K = k) = b(k; n, \varphi)b(x; k, \eta)$ where $b(x; n)$ is the binomial probability mass function (Lloyd, 2008). We also note that the *M p*-value is a special case of the *C p*-value; therefore, we focus our discussion on the *C p*-value.

Berger and Boos (1994) proposed a *C p*-value that confines the maximization within a well-behaved confidence interval in a general setting. Berger and Sidik (2003) adopted this *C p*-value definition for 2×2 matched-pairs data. Let I_γ be a $100(1 - \gamma)\%$ confidence interval for φ (Berger and Sidik, 2003). Now we can reformulate the *C p*-value as

$$\begin{aligned}
C(x, k) &= \sup_{\varphi \in I_\gamma, \eta \geq 1/2} P_{(\varphi, \eta)}\{T(X, K) \leq t_{obs}\} + \gamma \\
&= \sup_{\varphi \in I_\gamma, \eta \geq 1/2} \sum_{(x', k'): T(x', k') \leq t_{obs}} P_{(\varphi, \eta)}(X=x', K=k') + \gamma \\
&= \sup_{\varphi \in I_\gamma, k'=0} \sum_{n=0}^n \left\{ b(k'; n, \varphi) \sup_{n \geq 1/2} \sum_{x' \leq h(t_{obs}, k')} b(x'; k', \eta) \right\} + \gamma \\
&= \sup_{\varphi \in I_\gamma, k'=0} \sum_{n=0}^n b(k'; n, \varphi) \sup_{n \geq 1/2} B\{h(t_{obs}, k'); k', \eta\} + \gamma \\
&= \sup_{\varphi \in I_\gamma, k'=0} \sum_{n=0}^n b(k'; n, \varphi) B\{h(t_{obs}, k'); k', 1/2\} + \gamma
\end{aligned} \tag{4}$$

The last equality in (4) results from the fact that $B(x; n, p)$ is nonincreasing in p for fixed x and n . Note that when $\gamma = 0$, the confidence-interval region is the null space H_0 and thus the C p -value becomes the M p -value which is defined as $\sup_{(\varphi, \eta) \in H_0} P_{(\varphi, \eta)}\{T(X, K) \leq t_{obs}\}$

(Suisa and Shuster, 1991) and equals $\sup_{0 \leq \varphi \leq 1} \sum_{k'=0}^n b(k'; n, \varphi) B\{h(t_{obs}, k'); k', 1/2\}$. Thus, the maximization over two nuisance parameters (φ, η) in the M p -value and the C p -value can be reformulated into the maximization over the single parameter φ by using only the CMC property of the test statistic.

3.2 E + M p-value

The $E + M$ p -value proposed by Lloyd (2008) consists of two steps. First the E p -value is defined as

$$E(x, k) = \sup_{\eta \geq 1/2} P_{(\hat{\varphi}, \eta)}\{T(X, K) \leq t_{obs}\} = \sum_{k'=0}^n b(k'; n, \hat{\varphi}) B\{h(t_{obs}, k'); k', 1/2\}, \tag{5}$$

where $\hat{\varphi} = k/n$ is both the restricted and unrestricted MLE of φ under H_0 . Since the E p -value is not necessarily valid (for the definition of a valid p -value, see Berger and Boos, 1994), Lloyd (2008) further proposed a maximization step on the E p -value to obtain a valid p -value

$$(E+M)(x, k) = \sup_{(\varphi, \eta) \in H_0} P_{(\varphi, \eta)}\{E(X, K) \leq e_{obs}\}, \tag{6}$$

where $e_{obs} = E(x, k)$.

It is difficult to prove that the E p -value is non-increasing for k when x is fixed and thus the E p -value may not satisfy the BCC . Lloyd (2008) numerically verified the BCC for the E p -value. Our Theorem 1 guarantees that the E p -value satisfies the CMC , so the dimension reduction technique in (4) can be applied on the E p -value to obtain the $E + M$ p -value.

3.3 Exact size of the p-values

Let $\tilde{P}(x, k)$ denote any of the M , C and $E + M$ p -values. The exact size of $\tilde{P}$ is defined as

$$\text{size}(\alpha) = \sup_{(\varphi, \eta) \in H_0} P\{\tilde{P}(X, K) \leq \alpha\} + \sup_{0 \leq \varphi \leq 1} \sum_{k'=0}^n b(k'; n, \varphi) B\{\tilde{h}(\alpha, k'); k', 1/2\},$$

where $\tilde{h}(\alpha; k) = \sup\{x': \tilde{P}(x', k) \leq \alpha\}$. Berger and Sidik (2003) numerically checked the *BCC* for the *C p*-value when using the dimension reduction to calculate the exact size for the *C p*-value. By the following Theorem 1, we can show that all of the *p*-values (*M*, *C* and *E + M*) satisfy the *CMC* and thus the same dimension reduction technique in (4) can be applied to calculate their exact sizes.

Theorem 1—If a test statistic $T(x, k)$ satisfies the conditional monotonicity condition, i.e., it is nondecreasing in x for any fixed k , then the associated *p*-values, $M(x, k)$, $C(x, k)$, $E(x, k)$ and $(E + M)(x, k)$, also satisfy the conditional monotonicity condition.

Proof: For a fixed k and $x_1 \leq x_2$, let $t_{obs,1} = T(x_1, k)$, $t_{obs,2} = T(x_2, k)$, and then $t_{obs,1} \leq t_{obs,2}$ as $T(x, k)$ is nondecreasing in x . Note that $h(t, k)$ is also nondecreasing in t , then $h(t_{obs,1}, k') \leq h(t_{obs,2}, k')$ for any $k' \in \{0, \dots, n\}$. Thus we have $B\{h(t_{obs,1}, k'); k', 1/2\} \leq B\{h(t_{obs,2}, k'); k', 1/2\}$. Therefore, $M(x, k)$ and $C(x, k)$ are nondecreasing in x for fixed k . Note that when k is fixed, $\hat{\varphi}$ is also fixed; therefore, $E(x, k)$ in (5) and $(E + M)(x, k)$ in (6) are nondecreasing in x .

4. Extension to Non-inferiority Trials

In practice, the non-inferiority test of $H_0^*: \pi_{01} - \pi_{10} \geq \delta$ versus $H_1^*: \pi_{01} - \pi_{10} < \delta$ is of interest. Several test statistics have been proposed to test H_0^* ; these include the score statistic (Nam, 1997; Tango, 1998) and the likelihood ratio statistic (Lloyd and Moldovan, 2008). To the best of our knowledge, rigorous proofs for the *BCC* of the score and the likelihood ratio statistics are not available in the literature, neither are the proofs for the *BCC* of the *E p*-value and the *C p*-value. Here, we can prove that both of these test statistics satisfy the weaker condition *CMC* (the proof for the likelihood ratio test statistic is given in the Appendix), and so do the *E p*-value and the *C p*-value. We applied the same transformations $\varphi = \pi_{01} + \pi_{10}$ and $\eta = \pi_{01}/\varphi$. Note that φ has a lower bound δ under H_0^* . Following the derivation in Section 3, the *M p*-value for the non-inferiority test is $M(x, k) = \sup_{\varphi \in (\delta, 1)} \Phi(x, k, \varphi)$, where

$$\Phi(x, k, \varphi) = \sum_{k'=0}^n b(k'; n, \varphi) B\{h(t_{obs}; k'); k', (\delta + \varphi)/(2\varphi)\}.$$

The *C p*-value is $C(x, k) = \sup_{\varphi \in I_\gamma} \Phi(x, k, \varphi) + \gamma$, where I_γ is the $100(1-\gamma)\%$ confidence interval for φ in Sidik (2003). In Lloyd and Moldovan (2008), the *E p*-value was based on the restricted *MLE* $\hat{\varphi}$ of φ . We define the *E p*-value to be $E(x, k) = F(x, k, \hat{\varphi})$, where $\hat{\varphi} = (1 - \delta)k/n + \delta$; $\hat{\varphi}$ is always within $[\delta, 1]$ and is a function of k . Thus, we obtain the *CMC* of the *E p*-value that ensures the *M* step required for the *E + M p*-value. The *E + M p*-values obtained from these two different *E p*-values are practically identical from our numerical evaluations. Similar to the superiority case, the *CMC* is sufficient to calculate *p*-values and their exact sizes for non-inferiority tests with one dimensional maximization.

Acknowledgments

We would like to thank the associate editor and two referees for providing valuable comments that greatly improved the paper.

References

- Berger RL, Sidik K. Exact unconditional tests for a 2×2 matched-pairs design. *Statistical Methods in Medical Research*. 2003; 12:91–108. [PubMed: 12665205]
- Berger RL, Boos DD. P values maximized over a confidence set for the nuisance parameter. *Journal of the American Statistical Association*. 1994; 89:1012–1016.
- Lloyd CJ. A new exact and more powerful unconditional test of no treatment effect from binary matched pairs. *Biometrics*. 2008; 64:716–723. [PubMed: 18047530]
- Lloyd CJ, Moldovan MV. A more powerful exact test of noninferiority from binary matched-pairs data. *Statistics in Medicine*. 2008; 27:3540–3549. [PubMed: 18314932]
- Nam J. Establishing equivalence of two treatments and sample size requirements in matched-pairs design. *Biometrics*. 1997; 53:1422–1430. [PubMed: 9423257]
- Sidik K. Exact unconditional tests for testing non-inferiority in matched-pairs design. *Statistics in Medicine*. 2003; 22:265–278. [PubMed: 12520561]
- Suissa S, Shuster JJ. The 2×2 matched pairs trial: Exact unconditional design and analysis. *Biometrics*. 1991; 47:361–372. [PubMed: 1912252]
- Tango T. Equivalence test and confidence interval for the difference in proportions for the paired-sample design. *Statistics in Medicine*. 1998; 17:891–908. [PubMed: 9595618]

Appendix

The CMC of likelihood ratio test statistic in non-inferiority trials

By using transformations $\varphi = \pi_{01} + \pi_{10}$ and $\eta = \pi_{01}/\varphi$, the null space

$H_0^* = \{(\pi_{01}, \pi_{10}) : \pi_{01} - \pi_{10} \geq \delta\}$ becomes $\{(\varphi, \eta) : \delta \leq \varphi \leq 1, \eta \geq (\delta + \varphi)/2\varphi\}$, where $0 \leq \delta \leq 1$. The log likelihood function can be written as $l = l(\varphi, \eta) = k \log \varphi + (n - k) \log(1 - \varphi) + x \log \eta + (k - x) \log(1 - \eta) + C$, where C is independent of the parameters (φ, η) . The restricted MLE of φ satisfies $\partial l / \partial \varphi = 0$ under $\pi_{01} - \pi_{10} = \delta$, i.e.,

$$x/(\tilde{\varphi} + \delta) + (k - x)/(\tilde{\varphi} - \delta) - (n - k)/(1 - \tilde{\varphi}) = 0. \quad (\text{A.1})$$

A simple algebra manipulation verifies that $\tilde{\varphi}$ is the greater of the two solutions for (A.1) and is no less than δ . Furthermore, by taking derivative on $\tilde{\varphi}$ for (A.1), we obtain $\partial \tilde{\varphi} / \partial x \cdot (-U) = 1/(\tilde{\varphi} - \delta) - 1/(\tilde{\varphi} + \delta)$, where $U = x/(\tilde{\varphi} + \delta)^2 + (k - x)/(\tilde{\varphi} - \delta)^2 + (n - k)/(1 - \tilde{\varphi})^2$ is nonnegative. Therefore, $\partial \tilde{\varphi} / \partial x \leq 0$.

Let $(\hat{\varphi}, \hat{\eta}) = \{k/n, x/k\}$ be the unrestricted MLE of (φ, η) . Define $D = l(\hat{\varphi}, \hat{\eta}) - l(\tilde{\varphi}, (\delta + \tilde{\varphi})/2\tilde{\varphi})$, then the likelihood ratio statistic $Z_{LR} = \text{sign}(2x - k - n\delta) \sqrt{2D}$. With (A.1), we obtain $d = \partial D / \partial x = \log x(\tilde{\varphi} - \delta)/(k - x)(\tilde{\varphi} + \delta)$. And $\partial d / \partial x = \{1/(\tilde{\varphi} - \delta) - 1/(\tilde{\varphi} + \delta)\} \cdot \partial \tilde{\varphi} / \partial x + 1/x + 1/(k - x)$. Here, $\partial d / \partial x \geq 0$ can be established if $U \geq x(k - x)/k \{1/(\tilde{\varphi} + \delta)^2 + 1/(\tilde{\varphi} - \delta)^2 - 2/(\tilde{\varphi} + \delta)(\tilde{\varphi} - \delta)\}$, which is true by observing that $x/(\tilde{\varphi} + \delta)^2 \geq x(k - x)/k(\tilde{\varphi} + \delta)^2$, and $(k - x)/(\tilde{\varphi} - \delta)^2 \geq x(k - x)/k(\tilde{\varphi} - \delta)^2$.

Denote $x_0 = (k + n\delta)/2$. When $x = x_0$, we have $\hat{\varphi} = \tilde{\varphi}$, so $d(x_0)$ is zero. Note that d is nondecreasing in x . If $x \geq x_0$, then $d(x) \geq 0$ and thus D is nondecreasing, and so is Z_{LR} . If $x < x_0$, then $d(x) \leq 0$ and thus D is nonincreasing, and then Z_{LR} is nondecreasing due to the negative sign of $2x - k - n$. Therefore, Z_{LR} is nondecreasing in x for any fixed k .