Research Article

Statistics in Medicine

Received 9 March 2012, Accepted 18 July 2012 Published online in Wiley Online Library

(wileyonlinelibrary.com) DOI: 10.1002/sim.5561

A non-iterative implementation of Tango's score confidence interval for a paired difference of proportions

Zhao Yang,[a*†] Xuezheng Sun[b] and James W. Hardin[c]

For matched-pair binary data, a variety of approaches have been proposed for the construction of a confidence interval (CI) for the difference of marginal probabilities between two procedures. The score-based approximate CI has been shown to outperform other asymptotic CIs. Tango's method provides a score CI by inverting a score test statistic using an iterative procedure. In this paper, we propose an efficient non-iterative method with closed-form expression to calculate Tango's CIs. Examples illustrate the practical application of the new approach.

Keywords: marginal probability; matched-pair binary data; confidence interval; diagnostic medicine; McNemar's test; Tango's method

1. Introduction

In biomedical research, there are many research studies for which a pair of outcomes (i.e., matched-pairs data) are collected from each sampling unit. For example, cross-over trials comparing two drugs on each patient or studies for which observations for each of the two treatment procedures are collected at pairs of body locations of each subject such as measurements on each eye or each ear. When the outcome measure is binary, McNemar's test [1] is a common non-parametric significance test for comparing whether the marginal frequencies are equal between the two different procedures. For statistical inference, the CI is easier to interpret and can provide additional information to statistical tests. A variety of approaches have been proposed for the construction of CIs for the difference of marginal probabilities in matched-pair binary data [2].

Of the various proposed CIs, a score-based CI developed by Tango [3] has been advocated by several researchers [4, 5] because it appears to be the best approximate CI and has been shown to consistently, although only slightly, outperform other existing asymptotic CIs. However, Tango's method involves inverting a score test statistic by using an iterative procedure to arrive at an interval estimate. Agresti and Min [6] provide an implementation in R by using an iterative procedure and comment that 'There is not a closed-form expression for the resulting interval, but it can be obtained using iterative methods.' Recently, Nam [7] proposed a non-iterative method to construct the CI for the ratio of marginal probabilities in matched-pair binary data. We apply this technique to derive an efficient method with closed-form expression for the calculation of Tango's CI for the difference of marginal probabilities. Thus, this presentation is not that which was termed 'not computationally friendly' by Tang *et al.* [5] in comments on Tango's CI.

[a] *Quintiles, Inc., 5927 South Miami Blvd., Morrisville, NC 27560, U.S.A.*
[b] *Department of Epidemiology, Gillings School of Global Public Health, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, U.S.A.*
[c] *Department of Epidemiology and Biostatistics, Arnold School of Public Health and Institute for Families in Society, University of South Carolina, Columbia, SC 29208, U.S.A.*
* *Correspondence to: Zhao Yang, Quintiles, Inc., 5927 South Miami Blvd., Morrisville, NC 27560, U.S.A.*
† *E-mail: tonyyangsxz@gmail.com*

2. Proposal

Suppose that there is a random sample of N pair of units, where each unit is measured under the ith procedure, $i = 1, 2$, where the index values 1 and 2 correspond to the new and standard procedures, respectively. Let 1 (success) and 0 (failure) indicate the kth unit's response to the ith procedure. Thus, there are four pairs of outcomes that can be tabulated in a 2×2 table: (1, 1), (1, 0), (0, 1), and (0, 0), where the first element is the response to the new procedure and the second is the response to the standard procedure. Let a, b, c, and d be the observed frequencies of (1, 1), (1, 0), (0, 1), and (0, 0), respectively. Clearly, a and d reflect counts of the concordant pairs, and b and c reflect counts of the discordant pairs. The 2×2 contingency in Table I presents the information for matched-pair data for which b and c are the frequencies of two types of discordant pairs; p_{11}, p_{10}, p_{01}, and p_{00} are the response probabilities for the pairs (1, 1), (1, 0), (0, 1), and (0, 0), respectively.

Therefore, for the new and standard procedures, we are interested in determining whether there are equal marginal probabilities; that is, the null hypothesis is $H_0 : P_1 = P_2$, where P_1 and P_2 are the true unknown marginal probabilities for new and standard procedures, respectively. Let $\delta = P_1 - P_2$ be the true difference between the two procedures. We assume that a larger success probability denotes greater efficacy. To calculate the score-based $100(1-\alpha)\%$ CI for δ developed by Tango [3], we need to solve Equation (1) by using an iterative algorithm to obtain the estimated confidence limits,

$$\frac{b - c - \delta N}{\sqrt{N\left[2\widehat{p}_{01} + \delta(1-\delta)\right]}} = \pm z_{1-\alpha/2}, \tag{1}$$

where $z_{1-\alpha/2}$ is the upper $100(1-\alpha/2)\%$ percentage point of the standard normal distribution, $\widehat{p}_{01}$ is the maximum likelihood estimator for p_{01} conditional on δ given by $\left(\sqrt{B^2 - 4AC} - B\right)/(2A)$, where $A = 2N$, $B = -b - c + (2N - b + c)\delta$, and $C = -c\delta(1-\delta)$.

The non-iterative efficient method by Nam [7] can be used to construct a $100(1-\alpha)\%$ CI for δ. Let $g = z^2_{1-\alpha/2}$, and then a series of intensive algebraic manipulations simplify Equation (1), resulting in a fourth-order polynomial equation of δ. From this, the $100(1-\alpha)\%$ CI is obtained by solving the equation $m_0\delta^4 + m_1\delta^3 + m_2\delta^2 + m_3\delta + m_4 = 0$, where the quantities m_i for $i = 0, \ldots, 4$ are defined as

$$\begin{aligned}
m_0 &= N^2\left(\frac{N}{g} + 1\right)^2, \\
m_1 &= -N(b-c)\left(\frac{N}{g} + 1\right)\left(\frac{4N}{g} + 1\right), \\
m_2 &= \frac{2N(b-c)^2}{g}\left(\frac{3N}{g} + 2\right) - N^2\left(\frac{b+c}{g} + 1\right), \\
m_3 &= N(b-c)\left(\frac{2(b+c)}{g} + 1\right) - \frac{(b-c)^3}{g}\left(\frac{4N}{g} + 1\right), \quad \text{and} \\
m_4 &= \frac{(b-c)^2}{g}\left(\frac{(b-c)^2}{g} - (b+c)\right),
\end{aligned}$$

where b and c are the frequencies of the two types of discordant pairs in Table I. Then, let $u_i = m_i/m_0$ for $i = 1, \ldots, 4$, and the fourth-order polynomial equation can be rewritten as $\delta^4 + u_1\delta^3 + u_2\delta^2 + u_3\delta + u_4 = 0$. Denote that $v_1 = -u_2$, $v_2 = u_1u_3 - 4u_4$, and $v_3 = -\left(u_3^2 + u_1^2u_4 - 4u_2u_4\right)$, and y_1 is an

Table I. Summary table for matched-pair binary data (parameters in parentheses indicate the expected probabilities).

	Procedure 2 (standard)		
Procedure 1 (new)	Success (1)	Failure (0)	Total
Success (1)	$a(p_{11})$	$b(p_{10})$	$a + b(P_1)$
Failure (0)	$c(p_{01})$	$d(p_{00})$	$c + d(1 - P_1)$
Total	$a + c(P_2)$	$b + d(1 - P_2)$	N

appropriate real root of the cubic equation $y^3 + v_1 y^2 + v_2 y + v_3 = 0$. For each real root y_1, calculate the roots for δ in the fourth-order polynomial equation as

$$\frac{1}{2}\left[-\left(\frac{u_1}{2}+h_1\right) \pm \sqrt{\left(\frac{u_1}{2}+h_1\right)^2 - 4\left(\frac{y_1}{2}+h_2\right)}\right] \tag{2}$$

and

$$\frac{1}{2}\left[\left(h_1-\frac{u_1}{2}\right) \pm \sqrt{\left(h_1-\frac{u_1}{2}\right)^2 - 4\left(\frac{y_1}{2}-h_2\right)}\right], \tag{3}$$

where $h_1 = \left(u_1^2/4 - u_2 + y_1\right)^{1/2}$ and $h_2 = (u_1 y_1/2 - u_3)/(2h_1)$. The smallest and largest real numbers from the aforementioned calculations are the lower and upper limits for δ, denoted by δ_l and δ_u, respectively. For detailed information about the derivation of these quantities, refer to the Appendix A. If $b = c$, then the fourth-order polynomial equation of δ reduces to $m_0\delta^4 + m_2\delta^2 = 0$, and the lower and upper limits for δ can be calculated as $\delta_l = -\sqrt{-u_2}$ and $\delta_u = \sqrt{-u_2}$, respectively, where $u_2 = -[(b+c)/g + 1]/(N/g+1)^2$. In addition, note that if we use $g = z_{1-\alpha}^2$, the $100(1-2\alpha)\%$ CI for the difference of marginal probabilities can also be calculated for the statistic to test bio-equivalence/non-inferiority developed by Tango [3] and Nam [8].

Remark 1

Arbitrarily, the upper limit δ_u is set to 1 for $b = N$ and $c = 0$, and the lower limit δ_l is set to -1 for $b = 0$ and $c = N$. These manipulations can avoid the awkward situation that the point estimate $\widehat{\delta} = \pm 1$ falls outside the CI for δ.

3. Examples

To illustrate the application of the non-iterative calculation of Tango's CI, we used the new approach for CI construction of the data examples from [4, 5] and for some examples with very small sample size. The first part results from Table II are verified to be an exact match with those reported in [4]. Appendix B provides sample R codes for interested readers, and to emphasize the ease of construction of the CI by using the new method, to be useful to practitioners and researchers, the final version of the R code is available at http://works.bepress.com/zyang/1/. On the basis of the results from Table II, we summarize two remarks as follows.

Table II. Tango's score-based 95% CI for several data examples.

Data example					Data example				
N	$a+d$	b	c	95% CI	N	$a+d$	b	c	95% CI
Part 1: examples from [4, 5]									
44	43	0	1	(−0.11808, 0.05940)	100	2	98	0	(0.90675, 0.99450)
14	10	3	1	(−0.16697, 0.43266)	30	0	30	0	(0.77297, 1.00000)
32	20	9	3	(−0.02709, 0.38970)	54	54	0	0	(−0.06641, 0.06641)
50	36	12	2	(0.06111, 0.34471)	350	94	254	2	(0.66875, 0.76537)
50	36	14	0	(0.17474, 0.41665)	350	50	297	3	(0.79391, 0.87620)
100	2	97	1	(0.86984, 0.98659)	605	242	290	73	(0.30266, 0.41207)
30	0	29	1	(0.66659, 0.98818)	350	29	101	220	(−0.42991, −0.24309)
Part 2: examples to investigate performance under very small sample size									
10	1	9	0	(0.37269, 0.98212)	10	9	1	0	(−0.20529, 0.40415)
5	1	4	0	(0.01793, 0.96378)	5	4	1	0	(−0.32138, 0.62447)
3	1	2	0	(−0.26916, 0.93851)	3	2	1	0	(−0.41533, 0.79234)
3	1	1	1	(−0.69240, 0.69240)	2	1	1	0	(−0.48643, 0.90547)
10	10	0	0	(−0.27753, 0.27753)	9	9	0	0	(−0.29915, 0.29915)
7	7	0	0	(−0.35433, 0.35433)	5	5	0	0	(−0.43448, 0.43448)
3	3	0	0	(−0.56150, 0.56150)	2	2	0	0	(−0.65762, 0.65762)

Remark 2
The function `scoreci.mp` in the R package `PropCIs` can calculate Tango's CI by using an iterative procedure [9]. The results from the proposed R code in Appendix B match the results from `scoreci.mp` to the seventh decimal places. However, because the proposed method is a non-iterative procedure based on closed-form expression, it is expected to provide more reliable and accurate results. Also, for $b = c$, the function `scoreci.mp` gives the asymmetrical lower and upper limits of CI about 0, and the proposed R code is void of this concern.

Remark 3
As one of the referees suggests the use of a very small sample size to test if there is a flaw in the R code, we used the second part results in Table II to investigate the performance of the R code. We note that the proposal can give appropriate CIs and is robust to different data scenarios.

Appendix A. Derivation of non-iterative closed-form CI

Following closely the derivation of Nam [7], the simplified fourth-order polynomial equation (quartic equation) $\delta^4 + u_1\delta^3 + u_2\delta^2 + u_3\delta + u_4 = 0$ can be rewritten using Ferrari's formulation as

$$\left(\delta^2 + \frac{1}{2}u_1\delta\right)^2 = \left(\frac{1}{4}u_1^2 - u_2\right)\delta^2 - u_3\delta - u_4, \tag{A.1}$$

where $u_i = m_i/m_0$ for $i = 1, \ldots, 4$. Adding $\left(\delta^2 + \frac{1}{2}u_1\delta\right)y + \frac{1}{4}y^2$ to both sides of Equation (A.1) leads to

$$\left(\delta^2 + \frac{1}{2}u_1\delta + \frac{1}{2}y\right)^2 = \left(\frac{1}{4}u_1^2 - u_2 + y\right)\delta^2 + \left(\frac{1}{2}u_1 y - u_3\right)\delta + \left(\frac{1}{4}y^2 - u_4\right). \tag{A.2}$$

The discriminant of the right-hand side of Equation (A.2) being zero results in a perfect square; that is, $(u_1y/2 - u_3)^2 - 4\left(u_1^2/4 - u_2 + y\right)\left(y^2/4 - u_4\right) = 0$. This equation of the discriminant can be expressed as a cubic equation of y,

$$y^3 + v_1y^2 + v_2y + v_3 = 0, \tag{A.3}$$

where $v_1 = -u_2$, $v_2 = u_1u_3 - 4u_4$, and $v_3 = -\left(u_3^2 + u_1^2u_4 - 4u_2u_4\right)$. The cubic equation (A.3) is the resolvent for the quartic equation. Following standard methods, we find an appropriate root of Equation (A.3) such that the right-hand side of Equation (A.2) can be expressed as a square of a linear function of δ. Let $y = x - v_1/3$, and Equation (A.3) can be then reduced to

$$f(x) = x^3 + d_1x + d_2 = 0, \tag{A.4}$$

where $d_1 = v_2 - v_1^2/3$ and $d_2 = 2v_1^3/27 - v_1v_2/3 + v_3$. Defining $\Delta = (d_1/3)^3 + (d_2/2)^2$ and $i = \sqrt{-1}$, Cardan's solution of the cubic equation (A.4) is expressed as

$$x_1 = A^{1/3} + B^{1/3}, \quad x_2 = \omega A^{1/3} + \omega^2 B^{1/3}, \quad x_3 = \omega^2 A^{1/3} + \omega B^{1/3}, \tag{A.5}$$

where $A = -d_2/2 + \sqrt{\Delta}$, $B = -d_2/2 - \sqrt{\Delta}$, $\omega = \left(-1 + \mathrm{i}\sqrt{3}\right)/2$, and $\omega^2 = \left(-1 - \mathrm{i}\sqrt{3}\right)/2$. Any real root of Equation (A.4) can be used to solve the quartic equation, and Δ can be used to determine the number of real roots from Equation (A.4). That is, there is only one real root for $\Delta > 0$, two real roots for $\Delta = 0$, and three real roots for $\Delta < 0$. If x_1 is a real root of Equation (A.4), then the corresponding $y_1 = x_1 - v_1/3$ is the resolvent of the quartic equation. Substituting y_1 into Equation (A.2) gives

$$\left(\delta^2 + \frac{1}{2}u_1\delta + \frac{1}{2}y\right)^2 = (h_1\delta + h_2)^2,$$

where $h_1 = \left(u_1^2/4 - u_2 + y_1\right)^{1/2}$ and $h_2 = (u_1 y_1/2 - u_3)/(2h_1)$. The four roots of the quartic equation are

$$\frac{1}{2}\left[-\left(\frac{u_1}{2}+h_1\right) \pm \sqrt{\left(\frac{u_1}{2}+h_1\right)^2 - 4\left(\frac{y_1}{2}+h_2\right)}\right]$$

and

$$\frac{1}{2}\left[\left(h_1-\frac{u_1}{2}\right) \pm \sqrt{\left(h_1-\frac{u_1}{2}\right)^2 - 4\left(\frac{y_1}{2}-h_2\right)}\right].$$

The smallest and largest roots correspond to the lower and upper limits of the $100(1-\alpha)\%$ CI for δ.

Appendix B. R code to calculate Tango's CI

```
### R code to calculate Tango's score-based CI using a non-iterative method;
TangoCI <- function(N, b, c, alpha = 0.05) {
  options(digits = 12)
  g <- qnorm(1 - alpha/2)^2

  m0 <- N^2*(N/g + 1)^2
  m1 <- -N*(b - c)*(N/g + 1)*(4*N/g + 1)
  m2 <- 2*N*(b - c)^2*(3*N/g + 2)/g - N^2*((b + c)/g + 1)
  m3 <- N*(b - c)*(2*(b + c)/g + 1) - (b - c)^3*(4*N/g + 1)/g
  m4 <- (b - c)^2*((b - c)^2/g - (b + c) )/g

  u1 <- m1/m0;  u2 <- m2/m0;  u3 <- m3/m0; u4 <- m4/m0;

  if (b != c){
    nu1 <- -u2
    nu2 <- u1*u3 - 4*u4
    nu3 <- -(u3^2 + u1^2*u4 - 4*u2*u4)

    d1 <- nu2 - nu1^2/3
    d2 <- 2*nu1^3/27 - nu1*nu2/3 + nu3
    CritQ <- d2^2/4 + d1^3/27

    y1A <- h1 <- h2 <- h3 <- h4 <- root1 <- root2 <- root3 <- root4 <- c()

    if (CritQ > 0) { ## keep one real root;
        BigA <- -d2/2 + sqrt(CritQ)
        BigB <- -d2/2 - sqrt(CritQ)
        x1 <- sign(BigA)*abs(BigA)^(1/3) + sign(BigB)*abs(BigB)^(1/3)
        y1A <- x1 - nu1/3
    }

    if (CritQ == 0) { ## keep two real roots;
        BigA <- -d2/2 + sqrt(CritQ)
        BigB <- -d2/2 - sqrt(CritQ)
        Omega <- complex(real = -1/2, imaginary = sqrt(3)/2)
        Omega2 <- complex(real = -1/2, imaginary = -sqrt(3)/2)
        x1 <- sign(BigA)*abs(BigA)^(1/3) + sign(BigB)*abs(BigB)^(1/3)
        x2 <- Omega*sign(BigA)*abs(BigA)^(1/3) + Omega2*sign(BigB)*abs(BigB)^(1/3)
        y1A[1] <- x1 - nu1/3
        y1A[2] <- x2 - nu1/3
    }
```

```
    if (CritQ <0) { ## keep three real roots;
        BigA <- -d2/2 + sqrt(as.complex(CritQ))
        BigB <- -d2/2 - sqrt(as.complex(CritQ))
        Omega <- complex(real = -1/2, imaginary = sqrt(3)/2)
        Omega2 <- complex(real = -1/2, imaginary = -sqrt(3)/2)
        x1 <- BigA^(1/3) + BigB^(1/3)
        x2 <- Omega*BigA^(1/3) + Omega2*BigB^(1/3)
        x3 <- Omega2*BigA^(1/3) + Omega*BigB^(1/3)
        y1A[1] <- x1 - nu1/3
        y1A[2] <- x2 - nu1/3
        y1A[3] <- x3 - nu1/3
    }

    y1 <- Re(y1A) #keep the real part;
    ny <- length(y1)

    for (i in 1:ny){
      h1[i] <- sqrt(u1^2/4 - u2 + y1[i])
      h2[i] <- (u1 * y1[i]/2 - u3)/(2*h1[i])
      h3[i] <- (u1/2 + h1[i])^2 - 4*(y1[i]/2 + h2[i])
      h4[i] <- (h1[i] - u1/2)^2 - 4*(y1[i]/2 - h2[i])
      if (h3[i] >= 0){
        root1[i] <- ( - (u1/2 + h1[i]) + sqrt( h3[i] ) )/2
        root2[i] <- ( - (u1/2 + h1[i]) - sqrt( h3[i] ) )/2
      }
      if (h4[i] >= 0){
        root3[i] <- ( (h1[i] - u1/2) + sqrt( h4[i] ) )/2
        root4[i] <- ( (h1[i] - u1/2) - sqrt( h4[i] ) )/2
      }
    }
    lower <- max(-1,min(root1, root2, root3, root4, na.rm = TRUE))
    upper <- min(max(root1, root2, root3, root4, na.rm = TRUE), 1)

    if (b == N & c == 0) root <- c(lower, 1)
      else if (b == 0 & c == N) root <- c(-1, upper)
        else root <- c(lower, upper)
  }

  if (b == c){
    root1 <- -sqrt(-u2)
    root2 <- sqrt(-u2)
    root <- c(root1, root2)
  }
  return(root)
}
> TangoCI(350,254,2)
[1] 0.668750178889 0.765374630554
```

Acknowledgements

The authors are grateful to the editor, the associate editor, Prof. Robert G. Newcombe, and two other anonymous referees for their careful reading of the manuscript. Their valuable, constructive criticisms and suggestions greatly improved the quality and presentation of this material.

References

1. McNemar Q. Note on the sampling error of the difference between correlated proportions or percentages. *Psychometrika* 1947; **12**(2):153–157. DOI: 10.1007/BF02295996.

2. Newcombe RG. Improved confidence intervals for the difference between binomial proportions based on paired data. *Statistics in Medicine* 1998; **17**(22):2635–2650. DOI: 10.1002/(SICI)1097-0258(19981130)17:22<2635::AID-SIM954>3.0.CO;2-C.
3. Tango T. Equivalence test and confidence interval for the difference in proportions for the paired sample design. *Statistics in Medicine* 1998; **17**(8):891–908. DOI: 10.1002/(SICI)1097-0258(19980430)17:8<891::AID-SIM780>3.0.CO;2-B.
4. Newcombe RG. Confidence intervals for the mean of a variable taking the values 0, 1 and 2. *Statistics in Medicine* 2003; **22**(17):2737–2750. DOI: 10.1002/sim.1479.
5. Tang M, Ling M, Ling L, Tian G. Confidence intervals for a difference between proportions based on paired data. *Statistics in Medicine* 2010; **29**(1):86–96. DOI: 10.1002/sim.3738.
6. Agresti A, Min Y. Simple improved confidence intervals for comparing matched proportions. *Statistics in Medicine* 2005; **24**(5):729–740. DOI: 10.1002/sim.1781.
7. Nam J. Efficient interval estimation of a ratio of marginal probabilities in matched-pair data: non-iterative method. *Statistics in Medicine* 2009; **28**(23):2929–2935. DOI: 10.1002/sim.3685.
8. Nam J. Establishing equivalence of two treatments and sample size requirements in matched-pairs design. *Biometrics* 1997; **53**(4):1422–1430. DOI: 10.2307/2533508.
9. Scherer R. PropCIs: various confidence interval methods for proportions. R package version 0.1-8, 2012. http://cran.r-project.org/web/packages/PropCIs/.