两组率同为 100% 或 0% 时率差置信区间估计的 SAS 实现*

南方医科大学生物统计学系(510515) 黄耀华 唐欣然 段重阳 陈平雁△

【提 要】 **目的** 通过 SAS 编程实现两组事件发生率均为 0% 或 100% 时率差置信区间的估计。**方法** 针对事件发生率均为 100% 或 0% 时率差置信区间的估计问题,采用 SAS9.4 编程,使置信区间估计的 Miettinen Nurminen 法、Newcombe-Wilson 法及校正 Newcombe-Wilson 法等三种方法得以实现,并通过实例进行说明。**结果** 所编程序实现了三种方法的置信区间估计,便于专业和非专业人员使用。实例中两组样本量分别为 59,56,结果两组事件发生率均为 100%,三种方法的 95% 置信区间: Miettinen Nurminen 法为 [-6.16%,6.47%]; Newcombe 法为 [-6.11%,6.42%]; 校正 Newcombe 法为 [-7.62%,8.00%]。**结论** 本文所提供的 SAS 宏程序可以简便地实现两组事件发生率均为 0% 或 100% 时三种常用的率差置信区间的估计方法。

【关键词】 率差置信区间 SAS 宏程序 Newcombe 法 Miettinen Nurminen 法

SAS Implements of Calculating Rate Differences Confidence Intervals in Clinical Trials with Rates of 0% or 100% in Both Groups

Huang Yaohua, Tang Xinran, Duan Chongyang, et al (*Biometrics Department, School of Public Health and Tropical Medicine, South Medical University* (510515), *Guangzhou*)

【**Abstract**】 **Objective** To estimate confidence intervals of clinical trials with success rates of 0% or 100% in both treatment and controlled groups using SAS programming. **Methods** To resolve the issue of calculating confidence intervals of rate differences in clinical trials with both rates of 0% or 100%, programs were drafted using SAS 9.4. Miettinen and Nurminen, Newcombe-Wilson Score and Continuity-corrected Newcombe-Wilson methods could all be implemented with these programs. In addition, one example was displayed to illustrate the convenience of the programs. **Results** Confidence intervals in trials with both success rates of 0% or 100% could be resolved using the 3 methods, and it can be used feasibly by professionals and non-professionals. In the given example, with sample size of 59, 56, both of two groups had the success rate of 100%. 95% CI of rate difference was [-6.16%, 6.47%] calculated by Miettinen Nurminen, [-6.11%, 6.42%] by Newcombe-Wilson Score and [-7.62%, 8.00%] by Continuity-corrected Newcombe-Wilson Score. **Conclusion** Miettinen and Nurminen, Newcombe-Wilson Score and Continuity-corrected Newcombe-Wilson methods could all be implemented easily to calculate confidence intervals of rate differences in clinical trials with both rates of 0% or 100% by invoking the developed programs.

【**Key words**】 Proportion difference confidence interval; SAS macro procedure; Newcome-Wilson score; Miettinen Nurminen

医学研究领域,有时会遇到一种极端的结果,即两个比较组的事件发生率均为 100% 或 0%,如 CT 成像的优良率、关节置换的成功率、使用脑膜贴片的脑脊液渗漏率等,此时两组的率差为 0。目前,常用的两组事件发生率均为 100% 或 0% 时率差的置信区间估计方法有三种,分别是 Miettinen Nurminen 法[1]、Newcombe 法和校正 Newcombe 法[2-4]。然而,最新版本的 SAS 软件尚未提供上述三种方法的计算模块,既不便于专业人员的操作,又阻碍了非专业人员的应用。因此,本研究将编制 SAS 9.4 宏程序,为此种类型的数据处理提供方便可靠的工具。

方法介绍

1. Miettinen Nurminen 法

若用 x_1、x_2 分别表示两组的事件数,p_1、p_2 为两组事件发生率,n_1,n_2 分别为两组样本量,$N = n_1 + n_2$ 为总样本。对于率差 θ,Miettinen Nurminen 法[1] 先构建如下统计量:

$$T_{MN} = \frac{p_1 - p_2 - \theta}{\sqrt{\frac{N}{N-1}\left(\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}\right)}}$$

式中 $\hat{p}_1$、$\hat{p}_2$ 为给定 θ 情况下的限制性极大似然估计值 (restricted maximum likelihood estimates):

$$\hat{p}_1 = \hat{p}_2 + \theta$$

$$\hat{p}_2 = 2B\cos(A) - L_2/(3L_3)$$

其中

$$A = \frac{\pi + \cos^{-1}(C/B^3)}{3},$$

$$B = sign(C)\sqrt{\left(\frac{L_2}{3L_3}\right)^2 - \frac{L_1}{3L_3}},$$

$$C = \left(\frac{L_2}{3L_3}\right)^3 - \frac{L_1 L_2}{6L_3^2} + \frac{L_0}{2L_3}$$

$$L_0 = x_2\theta(1-\theta),\ L_1 = (n_2\theta - N - 2x_2)\theta + x_1 + x_2$$

* 国家自然科学基金项目资助(81673270)

△通信作者: 陈平雁,E-mail: chenpy99@126.com

$L_2 = (n_1 + 2n_2)\theta - N - x_1 - x_2$, $L_3 = N$

率差置信区间(L, U)分别为如下两个等式的解:

L: $T_{MN} = -z_{\alpha/2}$

U: $T_{MN} = z_{\alpha/2}$

Newcombe[2]对上述算法重新表示为如下表达式:

$$|\hat{\theta} - \theta| = z_{\alpha/2} \times \sqrt{\frac{N}{N-1}\left[\frac{(\psi_\theta + \theta/2)(1 - \psi_\theta - \theta/2)}{n_1} + \frac{(\psi_\theta - \theta/2)(1 - \psi_\theta + \theta/2)}{n_2}\right]}$$

率差置信区间(L, U)为上述表达式的两个解,$\hat{\theta} = p_1 - p_2$,ψ_θ为给定θ情况下两组率加和的限制性极大似然估计值。

当两组事件发生率均为0%时,$\hat{\theta} = 0$,Newcombe[2]给出$\psi_\theta = \frac{|\theta|}{2}$,上述置信区间构建过程可以简化为:

$$L = -\frac{z^2_{\alpha/2}}{n_2} \times \frac{N}{N-1} \Big/ \left(1 + \frac{z^2_{\alpha/2}}{n_2} \times \frac{N}{N-1}\right)$$

$$U = \frac{z^2_{\alpha/2}}{n_1} \times \frac{N}{N-1} \Big/ \left(1 + \frac{z^2_{\alpha/2}}{n_1} \times \frac{N}{N-1}\right)$$

当两组事件发生率均为100%时,$\hat{\theta} = 0$,我们推出$\psi_\theta = 1 - \frac{|\theta|}{2}$,上述置信区间构建过程可以简化为:

$$L = -\frac{z^2_{\alpha/2}}{n_1} \times \frac{N}{N-1} \Big/ \left(1 + \frac{z^2_{\alpha/2}}{n_1} \times \frac{N}{N-1}\right)$$

$$U = \frac{z^2_{\alpha/2}}{n_2} \times \frac{N}{N-1} \Big/ \left(1 + \frac{z^2_{\alpha/2}}{n_2} \times \frac{N}{N-1}\right)$$

2. Newcombe-Wilson 法

Newcombe-Wilson方法已被FDA指南推荐,作为差置信区间计算方法的首选[2-5]。其计算方法是通过Wilson法分别得到两单样本率的可信区间上下限[3]。Wilson法单样本率置信区间上下限为等式$(z^2_{\alpha/2} + n)\pi^2 - (z^2_{\alpha/2} + 2np)\pi + np^2 = 0$中$\pi$的两个解,$n$表示单组的样本量,$p$表示单组的事件发生率。求解可得单组率置信区间$(l, u)$的计算公式为:$(2np + z^2 \pm z\sqrt{(z^2 + 4npq)})/2(n + z^2)$,式中$q = 1 - p$。Newcombe-Wilson法通过杂交方式构建出率差置信区间上下限(L, U)如下:

$$L = p_1 - p_2 - \sqrt{(p_1 - l_1)^2 + (u_2 - p_2)^2}$$

$$U = p_1 - p_2 + \sqrt{(p_2 - l_2)^2 + (u_1 - p_1)^2}$$

其中l_1,u_1,l_2,u_2分别为两组率Wilson得分方法计算得到的置信区间上下限[2-4]。

3. 校正Newcombe-Wilson法

连续校正Newcombe-Wilson得分方法相对较为保守[2-3],其率差计算公式杂交方法同Newcombe-Wilson得分方法,区别在于Wilson单组置信区间的计算公式有所调整,采用了连续性校正后的结果,具体计算公式为

$$l = [2np + z^2 - 1 - z\sqrt{z^2 - 2 - 1/n + 4p(qn + 1)}]/[2(n + z^2)]$$

$$u = [2np + z^2 + 1 + z\sqrt{z^2 + 2 - 1/n + 4p(qn - 1)}]/[2(n + z^2)]$$

连续校正方法因单组率计算的调整而增加可信区间的宽度,从而更加保守地估计组间差异[2-4]。

率差置信区间估计的SAS实现[6-8]

```
% macroratediff( n1 = ,n1_event = ,n2 = ,n2_event = ,alpha = ,side = );
/* 近似正态方法1,此方法无法计算两组率均100%率差,因此将两组率保守设为99% */
data CMHChisq;
n1 = &n1.; p1 = 0.995;
n2 = &n2.; p2 = 0.995;
d = p1-p2;
l_diff = d-probit(1-&alpha./&side.) * sqrt(p1*(1-p1)/n1 + p2*(1-p2)/n2);
u_diff = d + probit(1-&alpha./&side.) * sqrt(p1*(1-p1)/n1 + p2*(1-p2)/n2);
run;
/* Miettinen Nurminen 方法* /
data mienur;
n1 = &n1.; a1 = &n1_event.; a2 = n1-a1; p1 = a1/n1;
n2 = &n2.; a3 = &n2_event.; a4 = n2-a3; p2 = a3/n2;
z = probit(1-&alpha./&side.);
d = p1-p2;
* 率差置信区间下限;
l_diff = -(z * * 2*(a1 + a3)/((a1 + a3-1) * a1))/((z * * 2*(a1 + a3)/((a1 + a3-1) * a1)) + 1);
* 率差置信区间上限;
u_diff = (z * * 2*(a1 + a3)/((a1 + a3-1) * a3))/((z * * 2*(a1 + a3)/((a1 + a3 - 1) * a3)) + 1);
run;
/* Newcombe - Wilson 得分方法,提交FDA报告中常见的方法 * /
data Newcombe;
n1 = &n1.; a1 = &n1_event.; a2 = n1 - a1; p1 = a1/n1;
n2 = &n2.; a3 = &n2_event.; a4 = n2 - a3; p2 = a3/n2;
```

```
z = probit( 1 - &alpha. /&side. );
* 单样本率置信区间下限;
l1 = ( 2* a1 + z * *2 - z* sqrt( z * *2 +4* a1* a2/n1) ) /( 2* ( n1 + z * *2) );
l2 = ( 2* a3 + z * *2 - z* sqrt( z * *2 +4* a3* a4/n2) ) /( 2* ( n2 + z * *2) );
* 单样本率置信区间上限;
u1 = ( 2* a1 + z * *2 + z* sqrt( z * *2 +4* a1* a2/n1) ) /( 2* ( n1 + z * *2) );
u2 = ( 2* a3 + z * *2 + z* sqrt( z * *2 +4* a3* a4/n2) ) /( 2* ( n2 + z * *2) );
d = p1 - p2;
单样本率置信区间下限
l_diff = d - sqrt( ( p1 - l1) * *2 + ( u2 - p2) * * 2);
单样本率置信区间上限
u_diff = d + sqrt( ( p2 - l2) * *2 + ( u1 - p1) * * 2);
run;
/* Newcombe - Wilson 得分连续校正方法,所有计算方法中最保守 * /
data NewcombeCC;
n1 = &n1. ; a1 = &n1_event. ; a2 = n1 - a1; p1 = a1/n1;
n2 = &n2. ; a3 = &n2_event. ; a4 = n2 - a3; p2 = a3/n2;
z = probit( 1 - &alpha. /&side. );
* 单样本率置信区间下限;
l1 = ( 2* a1 + z * *2 - 1 - z* sqrt( z * *2 - 2 - 1/n1 +4* p1* ( n1* a2/n1 + 1) ) ) /( 2* ( n1 + z * * 2) );
l2 = ( 2* a3 + z * *2 - 1 - z* sqrt( z * *2 - 2 - 1/n2 +4* p2* ( n2* a4/n2 + 1) ) ) /( 2* ( n2 + z * * 2) );
* 单样本率置信区间上限;
u1 = ( 2* a1 + z * *2 + 1 + z* sqrt( z * *2 + 2 - 1/n1 +4* p1* ( n1* a2/n1 - 1) ) ) /( 2* ( n1 + z * * 2) );
u2 = ( 2* a3 + z * *2 + 1 + z* sqrt( z * *2 + 2 - 1/n2 +4* p2* ( n2* a4/n2 - 1) ) ) /( 2* ( n2 + z * * 2) );
d = p1 - p2;
* 率差置信区间下限;
l_diff = d - sqrt( ( p1 - l1) * *2 + ( u2 - p2) * * 2);
* 率差置信区间上限;
u_diff = d + sqrt( ( p2 - l2) * *2 + ( u1 - p1) * * 2);
run;
data ratediff;
length method $ 200;
setmienur( in = mienur) CMHChisq( in = CMH-Chisq)
Newcombe( in = Newcombe) NewcombeCC( in = NewcombeCC);
if mienur then method = 'Miettinen Nurminen (仅限两组均为 100%)';
if CMHChisq then method = '近似正态(两组率均为 99.5%)';
if Newcombe then method = 'Newcombe';
if NewcombeCC then method = 'Newcombe 连续校正';
l_diff = l_diff* 100;
u_diff = u_diff* 100;
run;
proc print data = ratediff;
var method n1 a1 p1 n2 a3 p2 d l_diff u_diff;
format l_diff u_diff8. 2;
run;
% mend;
```

实例分析

某项用于骨折患者的骨钉临床试验,由于产品技术成熟,所有随访到的受试者在最终的临床评价中都为有效,即试验组和对照组事件发生率皆为 100%。其中试验组有效例数为 59,对照组有效例数为 56,计算两组事件发生率差值的点估计和置信区间估计。

该研究符合两组率都为 100% 条件,可以调用之前所编写程序,获得三种方法下计算得到的率差的点估计和置信区间估计。

```
% ratediff( n1 = 59, n1_event = 59, n2 = 56, n2_event = 56, alpha = 0. 05, side = 2);
```

表 1 为调用该宏程序后得到的结果,其中 n1 为试验组样本量,a1 为试验组有效的例数,p1 为试验组事件发生率,n2 为对照组样本量,a2 为对照组有效的例数,p2 为对照组事件发生率,d 为试验组和对照组两组率差点估计,l_diff 和 u_diff 分别为率差置信区间的下限和上限。

三种方法算得的点估计都为 0,Miettinen Nurminen 计算的率差置信区间估计为 [-6.16%,6.47%];近似正态方法因无法估算两组率差的置信区间,将两组事件发生率保守估计为 99.5%,获得率差置信区间估计为 [-2.58%,2.58%];Newcombe 方法计算的率差置信区间估计为 [-6.11%,6.42%];而 Newcombe

连续校正方法计算的率差置信区间估计为　[-7.62%, 8.00%]。

表1　调用%ratediff宏获得三种方法下两组率差点估计和置信区间估计

Obs	method	n1	a1	p1	n2	a3	p2	d	I_diff	u_diff
1	Miettinen Nurminen(仅限两组均为100%)	59	59	1	56	56	1	0	-6.16	6.47
2	Newcombe	59	59	1	56	56	1	0	-6.11	6.42
3	Newcomber连续校正	59	59	1	56	56	1	0	-7.62	8.00

研究者可以根据试验预先设定的评价方法选择恰当的一种，结合临床和统计评价标准判断试验研究假设是否成立。

讨　论

率差置信区间估计最常用的方法是CMH(Cochran-Mantal Haenszel)法[5,9-11]，但该法对于两组率同为0%或100%的情况无法进行置信区间估计，应用中虽然有将0%或100%用接近的数据替代(如0.5%或99.5%)，但毕竟导致数据失真，不宜提倡。

从实例看，Miettinen Nurminen法和Newcombe法的结果相近，而校正Newcombe法的结果最为保守，且精度较差。关于这三种方法的统计性能究竟如何，尚有待我们进一步的研究予以明确[9-11]。

参 考 文 献

[1] Miettinen O, Nurminen M. Comparative analysis of two rates Stat Med, 1985, 4(2): 213-226.

[2] Newcombe RG. Interval estimation for the difference between independent proportions: comparison of eleven methods. Stat Med, 1998, 17(8): 873-890.

[3] Newcombe RG. Improved confidence intervals for the difference between binomial prportions based on paired data. Statist. Stat Med, 1998, 17(6): 2635-2650.

[4] Newcombe RG. Two-sided confidence intervals for the single proportion: comparison of seven methods. Stat Med, 1998, 17(8): 857-872.

[5] FDA. Guidance for Industry and FDA Staff - Statistical Guidance on Reporting Results from Studies Evaluating Diagnostic Tests. (2007-03-13) http://www.fda.gov/medicaldevices/deviceregulationandguidance/guidancedocuments/ucm071148.htm

[6] SAS Institute Inc. SAS/IML® 9.2User's Guide. Second Edition. Cary, North Carolina, USA: SAS Institute Inc, 2009.

[7] Barker N. A Practical Introduction to the Bootstrap Using the SAS System. Heidelberg, 2005.

[8] Mehrotra D, Railkar R. Minimum risk weights for comparing treatments in stratified binomial trials. Statistics in Medicine, 2000, 19: 811-825.

[9] 张高魁. 阳性药对照临床试验有效性的可信区间评价方法. 中华临床药学, 2005, (5): 389-391.

[10] 刘沛. 总体率可信区间计算的一次近似法及其特征. 中国卫生统计, 2004, 21(5): 297-299.

[11] 刘沛. 四种方法计算总体率可信区间的比较研究. 中国卫生统计, 2005, 22(6): 354-358.

(责任编辑：郭海强)

(上接第10页)

[14] Demicheli R, Abbattista A, Miceli R, et al. Time distribution of the recurrence risk for breast cancer patients undergoing mastectomy: further support about the concept of tumor dormancy. Breast cancer research and treatment, 1996, 41(2): 177-185.

[15] Demicheli R, Retsky MW, Hrushesky WJ, et al. Tumor dormancy and surgery-driven interruption of dormancy in breast cancer: learning from failures. Nature Clinical Practice Oncology, 2007, 4(12): 699-710.

[16] 周力恒，殷文瑾，陆劲松，等. 乳腺癌患者术后不同部位复发转移的风险分布. 中国癌症杂志, 2008, 18(2): 124-127.

[17] Demicheli R, Biganzoli E, Ardoino I, et al. Recurrence and mortality dynamics for breast cancer patients undergoing mastectomy according to estrogen receptor status: different mortality but similar recurrence. Cancer science, 2010, 101(3): 826-830.

(责任编辑：郭海强)