

· 专题研究 ·

样本量估计及其在 nQuery 和 SAS 软件上的实现* ——相关分析

周立志¹ 姚阿玲¹ 陈平雁^{2△}

4. 相关分析

4.1. 单样本相关性分析

4.1.1. 差异性检验

4.1.1.1. kappa 系数检验(二分类变量)

方法:Donner 和 Eliasziw (1992)^[1]给出的单样本二分类变量 kappa 系数双侧检验的样本量估计方法,是建立在自由度为 1,非中心参数为 $\lambda(1, 1 - \beta, \alpha)$ 的非中心 χ^2 分布上的,其样本量的计算公式为:

$$n = \lambda(1, 1 - \beta, \alpha) \left\{ \frac{[\pi(1 - \pi)(\kappa_1 - \kappa_0)]^2}{\pi^2 + \pi(1 - \pi)\kappa_0} + \frac{2[\pi(1 - \pi)(\kappa_1 - \kappa_0)]^2}{\pi(1 - \pi)(1 - \kappa_0)} + \frac{[\pi(1 - \pi)(\kappa_1 - \kappa_0)]^2}{(1 - \pi)^2 + \pi(1 - \pi)\kappa_0} \right\} \quad (4-1)$$

式中, π 为研究对象被判为阳性的概率, κ_0 为原假设 kappa 系数, κ_1 为备择假设 kappa 系数。在自由度为 1 的情况下,非中心参数 $\lambda(1, 1 - \beta, \alpha)$ 近似等于 $(Z_{1-\alpha/2} + Z_{1-\beta})^2$, 在计算样本量时,将其代入(4-1)进行计算。

例 4-1 某项研究欲评价影像(CT)诊断和病理切片诊断对恶性肿瘤诊断的吻合程度。已知病理切片对恶性肿瘤诊断的阳性率为 65%。若 CT 诊断与病理切片诊断的吻合度达 0.6 以上为临床有效,即 $\kappa_0 = 0.6$, 根据预试验结果,两者的 kappa 系数为 0.8, 则假设 $\kappa_1 = 0.8$ 。试估计检验水准为 $\alpha = 0.05$, 检验效能为 $1 - \beta = 90\%$, 双侧检验欲验证 CT 诊断临床有效所需的样本量。

nQuery Advisor 7.0 实现: 设定检验水准 $\alpha = 0.05$; 双侧检验, 即 $s = 2$; 检验效能取 $1 - \beta = 90\%$ 。其他数据相应代入。

在 nQuery Advisor 7.0 主菜单选择:

Goal: Make Conclusion Using: Agreement

Number of Groups: One

Analysis Method: Test

方法框中选择: Kappa(binary outcome)。

在弹出的样本量计算窗口将各参数键入, 如图 4-

1 所示, 结果为 $n = 184$ 。即本试验所需的最少样本量为 184 例。

Agreement between two dichotomous ratings (intraclass kappa)			
	1	2	3
Test significance level, α	0.050		
1 or 2 sided test?	2		
Proportion successes, π	0.650		
Null hypothesis agreement, κ_0	0.600		
Alternative agreement, κ_1	0.800		
Power (%)	90		
n	184		

图 4-1 nQuery Advisor 7.0 关于例 4-1 样本量估计的参数设置与计算结果

SAS9.2 软件实现:

```
proc IML;
start AOT0(alpha, power, side, pi, kap0, kap1);
error = 0;
if(alpha > 1 | alpha < 0) then do;
error = 1; print "error" "Test significance level
must be in 0-1"; end;
if(power > 100 | power < 1) then do; error = 1;
print "error" "Power(%) must be in 1-100"; end;
if(side = 1 & side = 2) then do; error = 1; print
"error" "side = 1 or 2"; end;
if(pi > 1 | pi < 0) then do; error = 1; print "error"
"The probability of being a success must be in 0-1";
end;
if(kap0 < 0 | kap0 > 1 | kap1 < 0 | kap1 > 1) then
do; error = 1; print "error" "Kappa coefficient must be
in 0-1"; end;
if(error = 1) then stop;
if(error = 0) then do;
q = 1 - pi;
lambada = (probit(1 - alpha / side) + probit(power /
100)) * * 2;
f1 = 1 / (pi * q * (kap1 - kap0)) * * 2;
f2 = 1 / (1 / (pi * * 2 + pi * q * kap0) + 2 / (pi * q *
(1 - kap0)) + 1 / (q * * 2 + pi * q * kap0));
N = ceil(lambada * f1 * f2);
print alpha[ label = "Test Significance level"]
side[ label = "1 or 2 sided test"]
kap0[ label = "Null hypothesis Kappa"]
```

*: 教育部重点课题(DIA070113)资助

1. 南方医科大学 公共卫生与热带医学学院 生物统计学系 2008 级本科生

2. 南方医科大学 公共卫生与热带医学学院 生物统计学系

△通信作者

```

kap1 [ label = "Alternative Kappa" ]
power [ label = "Power ( % )" ]
N [ label = " n " ]; end; finish AOT0; run AOT0
( 0.05, 90, 2, 0.65, 0.6, 0.8 ); quit;
SAS 运行结果:
    
```

Test Significance level	Null hypothesis				n
	1 or 2 sided test	Kappa	Alternative Kappa	Power (%)	
0.05	2	0.6	0.8	90	184

图 4-2 SAS9.2 关于例 4-1 样本量估计的参数设置与计算结果

4.1.1.2. 相关系数检验(连续变量)

方法: Dixon 和 Massey^[2] 给出的两连续变量相关系数检验的样本量估计方法, 是基于大样本近似正态分布理论上的, 其样本量的计算公式为:

$$n = \frac{(Z_{1-\alpha/s} + Z_{1-\beta})^2}{[FZ(\rho_1) - FZ(\rho_0)]^2} + 3 \quad (4-2)$$

式中, FZ 为 Fisher's Z 转换, $FZ(\rho) = \frac{1}{2} \ln \left[\frac{1+\rho}{1-\rho} \right]$, s 为单双侧检验, ρ_0 为原假设相关系数, ρ_1 为备择假设相关系数。

例 4-2 某临床试验欲验证一款新研发的血气分析仪的临床有效性。根据美国国家临床实验室标准化委员会制定的 CLSI EP9-A2 标准, 试验仪器与阳性对照仪器的检测结果的相关系数不得低于 0.975。预期本试验的相关系数为 0.982 以上。试估计检验水准为 $\alpha = 0.05$, 检验效能为 $1-\beta = 80\%$, 单侧检验欲验证本款血气分析仪临床有效所需的样本量。

nQuery Advisor 7.0 实现: 根据题意, $\rho_0 = 0.975$, $\rho_1 = 0.982$, 检验水准 $\alpha = 0.05$, $s = 1$, 检验效能取 $1-\beta = 80\%$ 。

在 nQuery Advisor 7.0 主菜单选择:
 Goal: Make Conclusion Using: Agreement
 Number of Groups: One
 Analysis Method: Test

方法框中选择: Correlation coefficient (continue outcome)。

在弹出的样本量计算窗口将各参数键入, 如图 4-3 所示, 结果为 $n = 228$, 即本试验所需的最少样本量为 228 例。

	1	2
Test significance level, α	0.050	
1 or 2 sided test?	1	
Null hypothesis correlation, ρ_0	0.975	
Alternative correlation, ρ_1	0.982	
Power (%)	80	
n	228	

图 4-3 nQuery Advisor 7.0 关于例 4-2 样本量估计的参数设置与计算结果

SAS9.2 软件实现:
 AOT1; Test for value of correlation coefficient
 proc IML;

```

start AOT1 ( alpha, power, side, rho0, rho1 );
error = 0;
if ( alpha > 1 | alpha < 0 ) then do; error = 1; print
"error" "Test significance level must be in 0-1"; end;
if ( power > 100 | power < 1 ) then do; error = 1;
print "error" "Power ( % ) must be in 1-100"; end;
if ( side ^ = 1 & side ^ = 2 ) then do; error = 1; print
"error" "side = 1 or 2"; end;
if ( rho0 < 0 | rho0 > 1 | rho1 < 0 | rho1 > 1 ) then
do; error = 1; print "error" "Correlation coefficient must
be in 0-1"; end;
if ( error = 1 ) then stop;
if ( error = 0 ) then do;
lambda = ( probit ( 1-alpha/side ) + probit ( power/
100 ) ) * * 2;
beta = ( 1/2 * log ( ( 1 + rho1 ) / ( 1 - rho1 ) ) - 1/2 *
log ( ( 1 + rho0 ) / ( 1 - rho0 ) ) ) * * 2;
N = ceil ( lambda / beta + 3 );
print alpha [ label = "Test Significance level" ]
side [ label = "1 or 2 sided test" ]
rho0 [ label = "Null hypothesis Kappa" ]
rho1 [ label = "Alternative Kappa" ]
power [ label = "Power ( % )" ]
N [ label = " n " ];
end;
finish AOT1;
run AOT1 ( 0.05, 80, 1, 0.975, 0.982 );
quit;
SAS 运行结果:
    
```

Test Significance level	Null hypothesis				n
	1 or 2 sided test	Kappa	Alternative Kappa	Power (%)	
0.05	1	0.975	0.982	80	228

图 4-4 SAS9.2 关于例 4-2 样本量估计的参数设置与计算结果

4.1.1.3. Lin 和谐系数检验(连续变量)

方法: Lin 和 Hedayat 等 (2002)^[3] 给出的两连续变量 Lin 和谐系数检验的样本量估计方法, 是基于大样本近似正态分布理论上的, 其样本量的计算公式为:

$$n = \left[\frac{Z_{1-\beta} \sqrt{\eta_1} + Z_{1-\alpha} \sqrt{\eta_0}}{\zeta_0 - \zeta_1} \right]^2 + 2 \quad (4-3)$$

$$\zeta_m = \frac{1}{2} \ln \left[\frac{1 + \rho_{cm}}{1 - \rho_{cm}} \right], (m = 0, 1) \quad (4-4)$$

式中, ρ_{c0} 为原假设和谐系数, η_0 为原假设方差; ρ_{c1} 为备择假和谐系数, η_1 为备择假设方差。原假设及备择假设下的和谐系数及其方差按如下公式计算:

$$\rho_{cm} = \frac{2\rho_m}{v_m^2 + \tilde{\omega}_m + 1/\tilde{\omega}_m}, (m = 0, 1) \quad (4-5)$$

$$\eta_m = \frac{(1 - \rho_m^2) \rho_{cm}^2}{(1 - \rho_m^2) \rho_m^2} + \frac{2v_m^2(1 - \rho_{cm}) \rho_{cm}^3}{(1 - \rho_{cm}^2) \rho_m^2} - \frac{v_m^4 \rho_{cm}^4}{2(1 - \rho_{cm}^2)^2 \rho_m^2}, (m = 0, 1) \quad (4-6)$$

式中, ρ_m 为直线相关系数, $\tilde{\omega}_m$ 为尺度比 (Scale ratio), $\tilde{\omega}_m = \frac{\sigma_1}{\sigma_2}$; v_m^2 为均差 (mean difference), $v_m^2 = \frac{(\mu_1 - \mu_2)^2}{\sigma_1 \sigma_2}$; $m = 0$ 表示原假设条件下的参数, $m = 1$ 表示备择假设条件下的参数。

例 4-3 某研究欲研发一种自动测量血压的仪器, 以取代人工测量血压。研发时要求自动测压仪与人工测压的直线相关系数为 0.8, 均差 (mean difference) 为 0.15, 尺度比 (scale ratio) 为 1.15。预试验测得两者直线相关系数为 0.9, 均差为 0.20, 尺度比为 1.05。若想在检验水准为 $\alpha = 0.05$, 检验效能为 $1 - \beta = 90\%$ 的水平上, 经双侧检验得到 Lin 和谐系数有统计学差异的结论, 试估计样本量。

nQuery Advisor 7.0 实现: 设定检验水准 $\alpha = 0.05$; 双侧检验, 即 $s = 2$; 检验效能取 $1 - \beta = 90\%$ 。其他数据相应代入。

在 nQuery Advisor 7.0 主菜单选择:

Goal: Make Conclusion Using: Agreement

Number of Groups: One

Analysis Method: Test

方法框中选择: Lin's concordance coefficient (continuous outcome)。

这里, 和谐系数及其方差项可通过辅助计算获得。在菜单栏中选择:

Assistants: Compute Effect Size

在弹出的计算窗口将各参数键入, 如图 4-5 所示, 结果为 $c_0 = 0.74$, $v_0 = 0.84$; $c_1 = 0.82$, $v_1 = 0.77$ (结果保留两位小数)。

Transfer	Close	Reinput	Clear	Cut	Copy	Paste	Print
Input Item			Value (K_0)	Value (K_1)			
Correlation coefficient, ρ			0.800	0.900			
Mean difference, $(\mu_1 - \mu_2)^2 / \sigma_1 \sigma_2$			0.150	0.200			
Scale ratio, σ_1 / σ_2			1.150	1.050			
Null Hypothesis Concordance, K_0			0.737475				
Variance term, $V(K_0)$			0.835495				
Alternative Concordance, K_1				0.817297			
Variance term, $V(K_1)$				0.774097			

图 4-5 nQuery Advisor 7.0 关于例 4-3 样本量估计的参数计算结果

点击 Transfer 按钮后将参数结果传送至主菜单, 再键入其他参数, 如图 4-6 所示, 结果为 $n = 208$ 。即本试验的最少样本量为 208 例。

SAS9.2 软件实现:

```
proc IML; /* rho: Pearson Correlation coefficient
* /
/* nu: Mean difference */
/* omega: Scale ratio */
```

Large sample test that Lin's concordance coefficient = K_0		
	1	2
Test significance level, α	0.050	
1 or 2 sided test?	2	
Null hypothesis concordance, K_0	0.737	
Variance term, $V(K_0)$	0.8355	
Alternative concordance, K_1	0.817	
Variance term, $V(K_1)$	0.7741	
Power (%)	90	
n	208	

图 4-6 nQuery Advisor 7.0 关于例 4-3 样本量估计的参数设置与计算结果

```
start AOT2(alpha, power, side, rho0, nu0, omega0,
rho1, nu1, omegal);
error = 0;
if(alpha > 1 | alpha < 0) then do; error = 1; print
"error" "Test significance level must be in 0-1"; end;
if(power > 100 | power < 1) then do; error = 1;
print "error" "Power(%) must be in 1-100"; end;
if(side ^= 1 & side ^= 2) then do; error = 1; print
"error" "side = 1 or 2"; end;
if(nu0 < 0 | omega0 < 0 | nu1 < 0 | omegal < 0)
then do; error = 1; print "error" "nu and omega must be
larger than 0"; end;
if(rho0 < 0 | rho0 > 1 | rho1 < 0 | rho1 > 1) then
do; error = 1; print "error" "Correlation coefficient must
be in 0-1"; end;
if(error = 1) then stop;
if(error = 0) then do; CC0 = 2 * rho0 / (nu0 + omega0 +
1/omega0);
Eta0 = (1 - rho0 ** 2) * CC0 ** 2 / (1 - CC0 ** 2) / rho0 ** 2 + 2 * nu0 * (1 - CC0) * CC0 ** 3 / (1 - CC0 ** 2) ** 2 / rho0 - nu0 ** 2 * CC0 ** 4 / (1 - CC0 ** 2) ** 2 / rho0 ** 2;
Zeta0 = 0.5 * log((1 + CC0) / (1 - CC0));
CC1 = 2 * rho1 / (nu1 + omegal + 1/omegal);
Eta1 = (1 - rho1 ** 2) * CC1 ** 2 / (1 - CC1 ** 2) / rho1 ** 2 + 2 * nu1 * (1 - CC1) * CC1 ** 3 / (1 - CC1 ** 2) ** 2 / rho1 - nu1 ** 2 * CC1 ** 4 / (1 - CC1 ** 2) ** 2 / rho1 ** 2;
Zeta1 = 0.5 * log((1 + CC1) / (1 - CC1));
N = ((probit(power/100) * sqrt(Eta1) + probit(1 - alpha / side) * sqrt(Eta0)) / (Zeta0 - Zeta1)) ** 2 + 2;
CC0 = round(CC0, .001);
Eta0 = round(Eta0, .0001);
CC1 = round(CC1, .001); Eta1 = round(Eta1, .0001);
N = ceil(N); print alpha[label = "Test Significance level"]
```

(下转第 622 页)

学生,让学生根据所提供的基础数值进行相应的分析。这样,相应的可节约一部分试验课时,笔者对同一资料,借助电脑,通过 SPSS 软件演示其分析过程,这样不仅可以完成本次课的实习任务,同时也可给学生介绍统计软件的具体操作过程及对结果的合理解释,教学效果良好。

由此可见,电脑试验在卫生统计学的试验教学中,不仅可加强学生的实践能力,同时也为以后的科研工作奠定了一定的基础。

4. 卫生统计学非实时课堂学习平台建设

卫生统计学的大多数专业术语较为抽象,概念性、实践性很强,课后组织小组讨论,利用讨论反复强化;师生之间、生生之间互相交流,互相学习,相互启发对卫生统计学教学效果具有积极的促进作用。然而,课时教学有限,仅通过实时课堂教学,学生对卫生统计学内容的学习效果较差,可考虑以实时课堂为基础,搭建学生的非实时课堂学习平台。学生通过实时课堂学习,可了解课程的重点、难点、资源情况和学习方法等。通过非实时课堂学习平台,加强教师和学生之间的有效交流,可提高学生的学习能力,增强学生的自信心。非实时课堂是指不受时间限制的授课形式。非实时课堂主要用于导学,包括网上课程辅导、BBS 答疑、网上讨论、网上作业、课件练习、查找资料、期末复习、网上毕业论文指导等教学活动。学生可以自主地选择时间、地点和媒体进行自主学习活动。因此,加强卫生统计学资源平台建设,可增强师生的互动性,提高学生学习卫生统计学能力,对卫生统计学教学效果将起到一定的支撑作用。

小 结

随着医学科研工作的蓬勃开展,卫生统计学方法日益受到重视,并得到广泛应用。恰当、正确地应用卫生统计学的理论和方法可提高医学生的科研水平。然而,在实际应用过程中,未用、误用卫生统计学方法者比比皆是,导致研究设计不良、浪费人力物力、事倍功半。据此,针对统计学抽象难懂,内容枯燥,公式复杂等特点,采用案例教学、电脑试验、非实时课堂教学等多种教学方法,可缩短教学与实践的距离,达到教与学相长的目的,对提高医学生的科研水平也具有积极作用。

参 考 文 献

1. 徐勇勇,赵清波,颜艳. 医学院校统计教学值得商榷的几个问题. 中国卫生统计,2000,17(3):181-182.
2. 陈雅,邓青,姚佳红. 浅析医学统计学教学改革的必要性及措施. 中国卫生统计,2009,26(3):317-322.
3. 王玖,孙红卫,韩春蕾,等. 案例+任务驱动法在医学生统计软件教学中的应用. 中国卫生统计,2012,29(1):153-154.
4. Barrows HS. A taxonomy of problem-based learning methods. Med Edu, 1986, (20):481-486.
5. 姜萍,杨振宁,商庆新,等. PBL 教学模式在高等医学教学改革中的应用分析. 中国中医药信息杂志,2005,12(3):104-105.
6. 田冶,吴玲,王小丹. PBL 模式在卫生统计实践教学中的应用. 中国卫生统计,2010,27(6):661-665.
7. Dochy F, Segers M, Bossche PV, et al. Effects of problem-based learning: a meta-analysis. Learning and Instruction,2003,13(5):533-568.
8. Tarmizia RA, Tarmizib MAA, Loginina NI, et al. Problem-based learning: engaging students in acquisition of mathematical competency. Procedia Social and Behavioral Sciences,2010(2):4683-4688.
9. 方积乾. 医学统计学与电脑试验. 上海:上海科学技术出版社,1997.

(责任编辑:郭海强)

(上接第 619 页)

```

side[ label = "1 or 2 sided test" ]
CC0[ label = "Null hypothesis concordance" ]
Eta0[ label = "Null hypothesis variance term" ]
CC1[ label = "Alternative concordance" ]
Eta1[ label = "Alternative variance term" ]
power[ label = "Power(%)" ]
N[ label = "n" ]; end; finish AOT2; run AOT2
(0.05,90,2,0.80,0.15,1.15,0.90,0.20,1.05);quit;
SAS 运行结果:

```

Test Significance level	1 or 2 sided test	Null hypothesis concordance	Null hypothesis variance term	Alternative concordance
0.05	2	0.737	0.8355	0.817

Alternative variance term	Power (%)	n
0.7741	90	208

图 4-7 SAS9.2 关于例 4-2 样本量估计的参数设置与计算结果

(致谢:我们在此感谢 nQuery Advisor 软件全球经销商“爱尔兰 Statistical Solutions 有限公司”为本研究提供的支持和帮助。)

参 考 文 献

1. Donner A, Eliasziw M. A goodness-of-fit approach to inference procedures for the kappa statistic; confidence interval construction, significance-testing and sample size estimation. Statistics in Medicine, 1992, 11:1511-1519.
2. Dixon WJ, Massey FJ. Introduction to Statistical Analysis. 4th Edition. New York: McGraw-Hill, 1983.
3. Lin L, Hedayat AS, Sinha B, et al. Statistical methods in assessing agreement; models, issues, and tools. Journal of the American Statistical Association, 2002, 97:257-270.

(责任编辑:郭海强)