

样本量估计及其在 nQuery 和 SAS 软件上的实现 ——均数比较(二)

邓居敏² 吕 朵¹ 陈平雁^{2△}

1.1.1.4 基于线性组合比较的单个重复测量因素方差分析

方法:Overall 和 Doyle (1994)^[4]给出的基于线性组合比较的单个重复测量因素方差分析的样本量估计方法,是建立在自由度为 1 和 $(M-1)(n-1)$,非中心参数为 $n(|c|/(\sigma \cdot D \cdot \sqrt{1-\rho}))^2$ 的非中心 F 分布上,其检验假设表达为:

$$H_0: c_1\mu_1 + c_2\mu_2 + \dots + c_M\mu_M = 0$$

其检验效能的计算公式为:

$$1 - \beta = 1 - \text{Prob } F(F_{1-\alpha, 1, (M-1)(n-1)}, 1, (M-1)(n-1), n(|c|/(\sigma \cdot D \cdot \sqrt{1-\rho}))^2) \quad (1-3)$$

式中,

$$c = c_1\mu_1 + c_2\mu_2 + \dots + c_M\mu_M,$$

$$D = \sqrt{\sum_{i=1}^M c_i^2}, \sum_{i=1}^M c_i = 0 \quad (1-4)$$

M 为水平数; ρ 为水平间的相关系数; σ 为每个水平的总体标准差; c_i 为线性系数, μ_i 为不同水平均值。

在计算样本量时,一般先设定样本量初始值,然后迭代样本量直到所得的检验效能满足条件为止。此时的样本量,即研究所需的样本量。

【例 1-4】引用例 1-3,水平间的多重比较采用线性组合比较方法,线性组合系数分别设为 -3, -1, 1, 3, 试估计检验效能为 90% 时所需样本量。

nQuery Advisor 7.0 实现:设定检验水准 $\alpha = 0.05$; 检验效能取 $1 - \beta = 90\%$ 。依据上述基础数据可知, $\rho = 0.7, \sigma = 10, M = 4$ 。在 nQuery Advisor 7.0 主菜单选择:

Goal: Make Conclusion Using: Means

Number of Groups: One

Analysis Method: Test

方法框中选择: One-way repeated measures contrast。

注意,这里首先应根据不同重复水平的观察结果对 c 和 D 进行估计,假设测量得分逐步均匀升高。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 1-8 所示,结果为 $c = 15, D = 4.472$ 。

Level	Mean	Coefficient
1	55.000	-3.000
2	56.500	-1.000
3	58.000	1.000
4	59.500	3.000
Contrast, C = $\sum c_i \mu_i$	15.000	$\sum C_i = 0.000$
Scale, D = $\text{SQRT}(\sum c_i^2)$	4.472	

图 1-8 nQuery Advisor 7.0 关于例 1-4 样本量估计的参数计算结果

点击 Transfer 按钮后 c 和 D 显示于主对话框,再键入其他参数,如图 1-9 所示,结果为 $n = 29$ 。需要指出,在选择不同的线性系数 c_i 时,所计算的样本量是不完全相同的,例如,从水平 1 到水平 4 线性系数分别为 -2、-1、2、1,求得的样本量为 40。

	1	2	3	4
Test significance level, α	0.050			
Number of levels, M	4			
Contrast, C = $\sum c_i \mu_i$	15.000			
Scale, D = $\text{SQRT}(\sum c_i^2)$	4.472			
Standard deviation at each level, σ	10.000			
Between level correlation, ρ	0.700			
Effect size, $\Delta = C / [D \sigma \text{SQRT}(1-\rho)]$	0.612			
Power (%)	90			
n	29			

图 1-9 nQuery Advisor 7.0 关于例 1-4 样本量估计的参数设置与计算结果

SAS 9.2 软件实现:

```

% let mean = {55 56.5 58 59.5} ;
% let coe = { -3 -1 1 3} ; * Coefficient;
proc IML;
start MOT3(a, M, sd, cor, power);
error = 0;
if(a > 1 | a < 0) then do; error = 1; print "error"
"Test significance level must be in 0-1"; end;
if(sd < 0) then do; error = 1; print "error" "Stand-
ard deviation must be >= 0"; end;
if(abs(cor) > 1) then do; error = 1; print "error"
"Between level correlation must be in [-1, 1]"; end;
if(M < 0 | ceil(M) ^ M) then do; error = 1; print
"error" "The Number of levels must be positive inte-
ger"; end;
if(power > 100 | power < 1) then do; error = 1;
print "error" "Power(%) must be in 1-100"; end;

```

1. 南方医科大学 公共卫生与热带医学学院 生物统计学系 2007 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
△通讯作者

```

if(sum(&coe)^=0) then do;error = 1;print "er-
ror" "The sum of coe must be 1"; end;
if(error = 1) then stop;
if(error = 0) then do;
C = sum(&coe#&mean);
SDD = sqrt(sum(&coe##2));
es = abs(C)/(SDD#sd#sqrt(1-cor));
n = 2;
do until(pw >= power/100);
df1 = 1; df2 = (n - 1)#(M - 1);
ncp = n#es##2;
f = finv(1 - a, df1, df2);
pw = 1 - probf(f, df1, df2, ncp);
n = n + 0.01; end;

```

```

n = ceil(n - 0.01);
print a[ label = "Test significance level" ]
M[ label = "Number of levels" ]
C[ label = "Contrast, C" ]
SDD[ label = "Scale, D = sqrt( MSE )" ]
sdlabel = [ "Standard deviation at each level" ]
cor[ label = "Between level correlation" ]
es[ label = "Effect size" ]
power[ label = "Power( % )" ]
n[ label = "n" ]; end;
finish MOT3;
run MOT3(0.05, 4, 10, 0.7, 90);
quit;
SAS 运行结果:

```

Test significance level	Number of levels	Contrast, C	Scale, D = sqrt(MSE)	Standard deviation at each level	Between level correlation	Effect size	Power(%)	n
0.05	4	15	4.472136	10	0.7	0.6123724	90	29

图 1-10 SAS 9.2 关于例 1-4 样本量估计的参数设置与计算结果

1.1.1.5 基于 Greenhouse-Geisser 校正的单个重复测量因素方差分析

方法: Muller 和 Barton (1989)^[5] 等给出基于 Greenhouse-Geisser 校正的单个重复测量因素方差分析的样本量估计方法,它是建立在经 ϵ 系数调整后的自由度和非中心参数的非中心 F 分布上,其检验效能的计算公式为:

$$1 - \beta = 1 - \text{Prob}F(F_{1-\alpha, v_1, v_2}, v_1, v_2, nM \left(\frac{V}{\xi_w}\right)^2 \epsilon) \quad (1-5)$$

式中,

$$v_1 = (M - 1) \left(\epsilon + \frac{g_1}{n - 1} \right),$$

$$v_2 = (M - 1)(n - 1) \left(\epsilon + \frac{g_1}{n - 1} \right) \quad (1-6)$$

式中, M 为水平数, V 为各个水平均数的方差, ξ_w 为组内误差, ϵ 为 Greenhouse-Geisser“球对称(sphericity)”系数; g_1 为偏性系数。

在计算样本量时,一般先设定样本量初始值,然后迭代样本量直到所得的检验效能满足条件为止。此时的样本量,即研究所需的样本量。

【例 1-5】仍以例 1-3 为例,假设分析中采用 Greenhouse-Geisser 校正,试估计在检验效能为 90% 的情况下所需样本量。

nQuery Advisor7.0 实现:设定检验水准 $\alpha = 0.05$; 检验效能取 $1 - \beta = 90\%$ 。其他数据相应代入。

- 在 nQuery Advisor7.0 主菜单选择:
 Goal: Make Conclusion Using: Means
 Number of Groups: One
 Analysis Method: Test

方法框中选择:
 Univariate one-way repeated measures analysis of variance (Greenhouse-Geisser)。

弹出主对话框,如图 1-11 所示。均数的方差通过辅助计算获得,在 Assistants 菜单选择 Compute Effect Size,弹出对话框如图 1-12,键入相应的数据,求得均数的方差为 2.813。组内误差和 Greenhouse-Geisser 校正系数 ϵ 的计算亦需要通过辅助计算获得,点击按钮 Σ ,弹出对话框如图 1-13,键入共同标准差 10 和相关系数 0.7,获得相应的数据传输到主对话框(图 1-11),结果为 $n = 54$ 。

	1	2	3	4	5	6
Test significance level, α	0.050					
Number of levels, M	4					
Variance of means, $V = \sum(u_i - \bar{u})^2 / M$	2.813					
Within-group error term	6.11					
Measure of "sphericity", ϵ	0.88					
Bias term multiplier, g	-1.98					
Effect size, Δ^2	0.0754					
Power (%)	90					
n	54					

图 1-11 nQuery Advisor7.0 关于例 1-5 样本量估计的参数设置与计算结果

Level	Mean
1	55.000
2	56.500
3	58.000
4	59.500
Variance of means, $V = \sum(u_i - \bar{u})^2 / M$	2.813

图 1-12 nQuery Advisor7.0 关于例 1-5 样本量估计的参数计算结果

SAS 9.2 软件实现:
 PROC IML;
 start MOT4(a, M, V, inerr, spher, g1, power);
 error = 0;

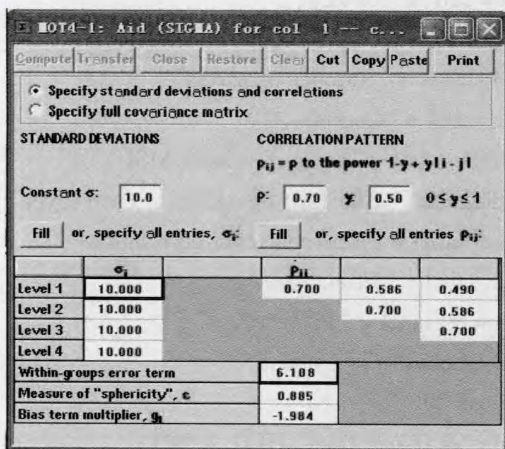


图 1-13 nQuery Advisor7.0 关于例 1-5 样本量估计的参数计算结果

```

if(a > 1 | a < 0) then do; error = 1; print "error"
"Test significance level must be in 0 - 1"; end;
if(spher > 1 | spher < 0) then do; error = 1; print
"error" "Measure of sphericity must be in 0 - 1"; end;
if(M < 0 | ceil(M)^= M) then do; error = 1; print
"error" "The Number of levels must be positive integer"
; end;
if(power > 100 | power < 1) then do; error = 1;
print "error" "Power(%) must be in 1 - 100"; end;
if(error = 1) then stop;
if(error = 0) then do;
es = V/inerr##2; n = M + 1;
do until(pw >= power/100);
df1 = (M - 1)#(spher + g1/(n - 1));
df2 = (n - 1) * (M - 1)#(spher + g1/(n - 1));

```

Test Significance level	Variance of means	Within-group error term	Measure of sphericity	Bias term multiplier	Effect size	Power (%)	n
0.05	2.813	6.11	0.88	-1.98	0.0753507	90	54

图 1-14 SAS 9.2 关于例 1-5 样本量估计的参数设置与计算结果

【例 1-6】某项研究欲评价两种治疗骨关节炎疼痛的方法是否等效,采用配对设计,用 VAS 疼痛量表进行测量。假定临床允许的等效差异为 10mm,并预期两种方法差值的均数为 0mm,标准差为 30mm。若 α 取 0.025(双单侧),检验效能取 90% 时,试估计样本量。

nQuery Advisor7.0 实现:设定检验水准 $\alpha = 0.025$ (双单侧);检验效能取 $1 - \beta = 90%$ 。依据上述基础数据可知, $\Delta_0 = 10, \Delta_1 = 0, \sigma = 30$ 。

在 nQuery Advisor7.0 主菜单选择:
 Goal: Make Conclusion Using: Means
 Number of Groups: One
 Analysis Method: Equivalence

```

ncp = n#M#es#(spher + g1/(n - 1));
qf = finv(1 - a, df1, df2);
pw = 1 - probf(qf, df1, df2, ncp);
n = n + 0.01; end;
n = floor(n - 0.01);
print a[ label = "Test Significance level" ]
V[ label = "Variance of means" ]
inerr[ label = "Within - group error term" ]
spher[ label = "Measure of sphericity" ]
g1[ label = "Bias term multiplier" ]
es[ label = "Effect size" ]
power[ label = "Power(%)" ]
n[ label = "n" ]; end;
finish MOT4;
run MOT4 (0.05, 4, 2.813, 6.11, 0.88, -1.98,
90); quit;

```

SAS 运行结果:

1. 1. 2 等效性检验

1. 1. 2. 1 等效性配对 t 检验

方法: Machin 和 Campbell (1987)^[6] 等给出等效性配对 t 检验的样本量估计是建立在自由度为 $n - 1$, 非中心参数为 $\sqrt{n}|\Delta_0 - \Delta_1|/\sigma$ 的非中心 t 分布上。其检验效能的计算公式为:

$$1 - \beta = 1 - \text{Prob } t(t_{1-\alpha, n-1}, n-1, \sqrt{n}|\Delta_0 - \Delta_1|/\sigma) \quad (1-7)$$

式中, Δ_0 代表等效的界值, Δ_1 代表期望的差值; σ 为总体标准差。

在计算样本量时,一般先设定样本量初始值,然后迭代样本量直到所得的检验效能满足条件为止。此时的样本量,即研究所需的样本量。

方法框中选择: Paired t test of equivalence of means。

在弹出的样本量计算窗口将各参数键入,如图 1-15 所示,结果为 $n = 1053$ 。

Paired t-test of equivalence of means				
	1	2	3	4
Test significance level, α (one-sided)	0.025			
Equivalence limit difference, Δ_0	10.000			
Expected difference, Δ_1	0.000			
$\Delta_0 - \Delta_1$	10.000			
Standard deviation of differences, σ_d	30.000			
Effect size, $\delta = \Delta_0 - \Delta_1 /\sigma_d$	0.333			
Power (%)	90			
n	97			

图 1-15 nQuery Advisor 7.0 关于例 6 样本量估计的参数设置与计算结果

求,将数据集按每个病例或每个对照一个记录展开,共得到 1834 (= 262 × 7, 其中 262 是在 9 层中病例记录的死亡数总和) 记录,从而完成了整个按设计要求的资料整理过程,最终的数据集为 eee。例如,第一层的资料整理结果列于表 3 中,第一层共有 28 × 7 = 196 个记录。

```
data eee;set ddd;retain dno;
set ddd;by stratum index;do i = 1 to dno;output;end;
run;
proc print data = eee;format index indfmt. ;run;
```

讨 论

病例交叉研究是继横断面研究、病例对照研究和队列研究后发展的一种新的流行病学方法,在环境污染与健康关系的研究中应用较广。但由于通常记录中的资料是按时间循序排列的,而且每一个记录包含有多个反应个体,其资料整理比较复杂,在实际工作中应用比较费时费力。本文以双向 1:3 病例配比的设计方案为泛例,用 SAS 数据步对原始数据整理,对整理后的资料可直接采用通用统计软件配合配对 logistic 回

归模型。本文为病例交叉研究资料提供了比较简单实用的 SAS 程序。如果实际设计方案不是如本文介绍的双向 1:3 病例配比的设计方案,只要在第二步中加以调整即可。例如,实际设计方案为回顾性 1:3 配比设计时,只要将队列配比的三个数据子集 coh_1, coh_2, coh_3 去掉即可。如果时间间隔不是 7, 14 和 21 (天),即可将 dif1, dif2 和 dif3 重新定义即可。本文提供了一个简便灵活的 SAS 程序,为实际工作者选用。

参 考 文 献

1. Maclure M. The case- crossover design: a method for studying transient effects on the risk of acute events. *Am J Epidemiol*, 1991, 133:144-153.
2. 张政,詹思延. 病例交叉设计, *中华流行病学杂志*, 2001, 22(4):304-306.
3. 胡以松. 病例交叉研究. *疾病控制杂志*, 2001, 5(4):341-343.
4. 阚海东,陈秉衡,贾健. 上海市大气污染与居民每日死亡关系的病例交叉研究. *中华流行病学杂志*, 2003, 24(10):863-867.
5. 贾健,阚海东,陈秉衡,等. 上海市闸北区大气污染与死亡率的病例交叉研究. *中华流行病学杂志*, 2003, 24(10):863-867.
6. SAS Institute Inc. SAS/STAT User's Guide. Version 9. 2, Cary, NC: SAS Institute Inc., 2008.

(上接第 134 页)

SAS 9.2 软件实现:

PROC IML;

start MOE0(a, eld, ed, sd, power); error = 0;

if(a > 1 | a < 0) then do; error = 1; print "error"

"Test significance level must be in 0 - 1"; end;

if(sd < 0) then do; error = 1; print "error" "Stand-

ard deviation must be > = 0"; end;

if(eld < 0) then do; error = 1; print "error" "E-

quivalence limit difference must be > = 0"; end;

if(ed < 0) then do; error = 1; print "error" "Expec-

ted difference must be > = 0"; end;

if(eld < = ed) then do; error = 1; print "error" "E-

quivalence limit difference must be > = Expected difference"; end;

if(power > 100 | power < 1) then do; error = 1;

print "error" "Power(%) must be in 1 - 100"; end;

if(error = 1) then stop;

if(error = 0) then do;

es = abs(ed - eld)/sd;

n = 2;

do until(pw > = power/100);

ncp = sqrt(n) # es;

df = n - 1; t = tinv(1 - a, df);

pw = 1 - probt(t, df, ncp); n = n + 0.01;

end;

n = ceil(n - 0.01);

print a[label = "Test Significance level"]

eld[label = "Equivalence limit difference"]

ed[label = "Expected difference"]

sd[label = "Standard deviation of differences"]

es[label = "Effect size"]

power[label = "Power(%)"]

n[label = "n"]; end;

finish MOE0;

run MOE0(0.025, 10, 0, 100, 90);

quit;

SAS 运行结果:

Test Significance level	Equivalence limit difference	Expected difference	Standard deviation of differences	Effect size	Power (%)	n
0.025	10	0	30	0.3333333	90	97

图 1-16 SAS 9.2 关于例 1-6 样本量估计的参数设置与计算结果