

## Estimation of sample size and testing power (Part 5)

Liang-ping Hu, Xiao-lei Bao, Xue Guan, Shi-guo Zhou

Consulting Center of Biomedical Statistics, Academy of Military Medical Sciences, Beijing 100850, China

**ABSTRACT:** Estimation of sample size and testing power is an important component of research design. This article introduced methods for sample size and testing power estimation of difference test for quantitative and qualitative data with the single-group design, the paired design or the crossover design. To be specific, this article introduced formulas for sample size and testing power estimation of difference test for quantitative and qualitative data with the above three designs, the realization based on the formulas and the POWER procedure of SAS software and elaborated it with examples, which will benefit researchers for implementing the repetition principle.

**KEYWORD:** statistics, medical; research design; sample size; testing power; parametric estimation

In this article, we will introduce the estimation of sample size and testing power of difference test for quantitative and qualitative data with the single-group design, the paired design or the crossover design.

### 1 Estimation of sample size and testing power of difference test for quantitative and qualitative data with the single-group design

The single-group design refers to the research on the purpose to observe the value of a quantitative or qualitative index of a random sample under specific conditions. The single-group design only involves a specific level of an experimental factor, and the research subjects are not divided into groups based on any experimental factor or block factor. If the observational index is quantitative,

the population mean or the standard value should be provided for hypothesis testing; if the observational index is qualitative, the population rate or the standard rate is needed<sup>[1]</sup>.

#### 1.1 Testing of the mean



**1.1.1 Formulas for sample size estimation** For one-sided test:

$$n = \left[ \frac{(t_{1-\alpha, df} + t_{1-\beta, df})S}{\theta} \right]^2 \quad (1)$$

For two-sided test:

$$n = \left[ \frac{(t_{1-\alpha/2, df} + t_{1-\beta, df})S}{\theta} \right]^2 \quad (2)$$

In formulas (1) and (2)<sup>[2]</sup>,  $\alpha$  and  $\beta$  are the allowed probabilities of making type I error and type II error, respectively;  $S$  represents the standard deviation;  $\theta$  is the difference of the population

 Open Access 开放获取	 Submission Guide 投稿指南
<p>DOI: 10.3736/jcim20120205  <a href="http://www.jcimjournal.com">http://www.jcimjournal.com</a></p> <p>Hu LP, Bao XL, Guan X, Zhou SG. Estimation of sample size and testing power (Part 5). <i>J Chin Integr Med</i>. 2012; 10(2): 154-159.                      胡良平, 鲍晓蕾, 关雪, 周诗国. 样本量估计与检验效能分析(五). <i>中西医结合学报</i>. 2012; 10(2): 154-159.</p> <p>Received December 2, 2011; accepted December 5, 2011; published online February 15, 2012.                      Full-text LinkOut at PubMed. Journal title in PubMed; <i>Zhong Xi Yi Jie He Xue Bao</i>.</p> <p><b>基金项目:</b> 国家科技重大专项“重大新药创制”资助项目 (No. 2009ZX09502-028)</p> <p><b>Correspondence:</b> Prof. Liang-ping Hu; Tel: 010-66931130; E-mail: lphu812@sina.com</p>	<p><i>Journal of Chinese Integrative Medicine (JCIM)</i> or <i>Zhong Xi Yi Jie He Xue Bao</i> is an international, peer-reviewed, open-access journal for the study of complementary and alternative medicine or integrative medicine from all regions of the world. <i>JCIM</i> is indexed in PubMed and Directory of Open Access Journals (DOAJ). <i>JCIM</i> is a member journal of CrossRef. Articles published in <i>JCIM</i> have maximum exposure to the international scholarly community.</p> <p><b>Submit your manuscript here:</b>  <a href="http://mc03.manuscriptcentral.com/jcim-en">http://mc03.manuscriptcentral.com/jcim-en</a> (for manuscripts written in English)  <a href="http://mc03.manuscriptcentral.com/jcim-cn">http://mc03.manuscriptcentral.com/jcim-cn</a> (for manuscripts written in Chinese)</p> <ul style="list-style-type: none"> <li>• No submission and page charges for manuscripts written in English</li> <li>• Quick decision and rapid publication</li> </ul> <p>Send your postal address by e-mail to <a href="mailto:jcim@163.com">jcim@163.com</a>, we will send you a complimentary print issue upon receipt.</p> <p>ISSN 1672-1977. Published by JCIM Press, Shanghai, China.</p>

means of the sample group and the standard group, which can be replaced by its estimated value if unknown;  $t_{1-\alpha,df}$ ,  $t_{1-\alpha/2,df}$  and  $t_{1-\beta,df}$  are the  $1-\alpha$  quantile,  $1-\alpha/2$  quantile and  $1-\beta$  quantile corresponding to the probability-density curve of the central  $t$  distribution with the degree of freedom  $df=n-1$ . The iterative algorithm is applied. First, let  $df=\infty$  and denote the  $n$  as  $n_1$ ; then let  $df=n_1-1$ , and denote the  $n$  as  $n_2\dots$ . The process continues until the nearest two “ $n$ ” are stable.

### 1.1.2 Formulas for testing power estimation

First, compute the value of  $t_{1-\beta,df}$  by formulas (3) and (4)<sup>[2]</sup>, then compute the value of  $1-\beta$  accordingly, namely, the testing power.

For one-sided test:

$$t_{1-\beta,df} = \frac{\theta\sqrt{n}}{S} - t_{1-\alpha,df}, \quad df = n - 1 \quad (3)$$

For two-sided test:

$$t_{1-\beta,df} = \frac{\theta\sqrt{n}}{S} - t_{1-\alpha/2,df}, \quad df = n - 1 \quad (4)$$

The symbols in formulas (3) and (4) have the same meanings as those in formulas (1) and (2).

### 1.1.3 Examples

**1.1.3.1 Example 1** The mean of the general closing time of the anterior fontanel of children in the northeast China is known as 14.1 months. A researcher planed to select a certain number of children with calcium deficiency in the northeast to survey whether the closing time of the anterior fontanel of the children in the northeast is longer. The pilot test showed that the closing time of the anterior fontanel of 36 children selected from the northeast was 14.3 months. and the standard deviation was 5.1 months. Suppose that the significance level  $\alpha$  was set to be 0.05 and the testing power was required to be 75%. How many children were needed so that the researcher could draw the conclusion that the closing time of the

anterior fontanel of the children in the northeast was longer?

Analysis: Example 1 deals with the sample size estimation of the one-sided test for quantitative data with the single-group design. The sample mean, the standard mean, the sample standard deviation, the significance level and the testing power are all given. Formula (1) can be used for sample size estimation.

The needed SAS program is as follows:


```
%let alpha=0.05;%let beta=0.25;%let miu_0=14.1;%let miu_1=14.3;%let s=5.1;%let side=1;
data a1;
theta=&miu_1-&miu_0;
if &side=1 then p1=1-&alpha;else p1=1-&alpha/2;p2=1-&beta;
df=10 * * 9;
t1=tinv(p1,df);t2=tinv(p2,df);
w=&s/theta;n1=((t1+t2)*w)* * 2;df=n1-1;
t1=tinv(p1,df);t2=tinv(p2,df);n2=((t1+t2)*w)* * 2;
do while(ABS(n1-n2)>1);
df=n2-1;n1=n2;t1=tinv(p1,df);t2=tinv(p2,df);
n2=((t1+t2)*w)* * 2;
end;
n=CEIL(MAX(n1,n2));file print;
if &side=1 then
PUT #3 @10 'mean comparison of the single-group design (the
one-sided test), the needed sample size' n=;
else
PUT #3 @10 'mean comparison of the single-group design (the
two-sided test), the needed sample size' n=;
run;quit;
```

Program explanation: In the beginning part, the six “%let” statements are used to specify the allowed probabilities of making type I error and type II error, the standard mean, the sample mean, the sample standard deviation and the one-sided test, which can be modified in specific situations.

The output: The needed sample size  $n = 3\ 500$ .

Sometimes, the output of the program will be slightly different from the result of the manual computation, because the  $t$  value computed by the computer is more accurate than that from the  $t$  table.

Below is the needed advanced SAS program:

 **Related Articles 推荐阅读**

卞兆祥, Moher D, 李幼平, 吴泰相, Dagenais S, 郑颂华, 李静, 李廷谦. 清晰报告中医药随机对照试验中的干预措施. 中西医结合学报. 2008; 6(7): 661-667.

Bian ZX, Moher D, Li YP, Wu TX, Dagenais S, Cheng CW, Li J, Li TQ. Precise reporting of traditional Chinese medicine interventions in randomized controlled trials. *J Chin Integr Med*. 2008; 6(7): 661-667.

**Free full text available at** <http://www.jcimjournal.com/FullText2.aspx?articleID=167219772008070661>

卞兆祥, 李幼平, Moher D, Dagenais S, 刘良, 吴泰相, 缪江霞, 关家伦, 宋丽. 提高中草药随机对照试验的质量 I: 临床试验设计和方法学. 中西医结合学报. 2006; 4(2): 120-129.

Bian ZX, Li YP, Moher D, Dagenais S, Liu L, Wu TX, Miao JX, Kwan AKL, Song L. Improving the quality of randomized controlled trials in Chinese herbal medicine, Part I: clinical trial design and methodology. *J Chin Integr Med*. 2006; 4(2): 120-129.

**Free full text available at** <http://www.jcimjournal.com/FullText2.aspx?articleID=502>

**More free related articles at** <http://www.jcimjournal.com/FullText2.aspx?articleID=jcim20120205>

```
proc power;
onesamplemeans test=t
sides=1 mean=0.2 stddev =5.1 ntotal = . power =0.75;
run;quit;
```

Program explanation: The option “test = t” requires to conduct the *t* test to the mean; “sides = 1” means to conduct the one-sided test; “mean = 0.2” specifies the difference between the population mean corresponding to the sample and the standard mean; “stddev = 5.1” specifies the sample standard value; “power = 0.75” specifies the testing power. The values in the program can be modified in specific situations.

The output is that 3 500 children were needed.

**1.1.3.2 Example 2** In example 1, suppose that the researcher used the data of the 36 children for the one-sided test instead of estimating the sample size, and drew the conclusion that there was no statistical significance between the population mean corresponding to the sample and the standard mean. Estimate the testing power.

The needed program based on formula (3) is as follows:

```
%let miu_0=14.1; %let miu_1=14.3; %let n=36; %let s=5.1;
%let alpha=0.05; %let side=1;
data a2;
theta=&miu_1-&miu_0;
df=&n-1;
if &side=1 then p1=1-&alpha; else p1=1-&alpha/2;
t_1_beta_df=theta * sqrt(&n)/&s-tinv(p1, df);
power=put(probt(t_1_beta_df, df), 7.3);
if power<0.75 then do; sig='<0.75,'; mean="the testing power is insufficient"; end;
if power>0.75 then do; sig='>0.75,'; mean="the testing power is sufficient"; end ;
file print;
PUT #3 @15 'power=' power sig mean;
run;
```

Program explanation: The six “%let” statements in the program specify the standard mean, the sample mean, the sample size, the sample standard deviation, the allowed probability of making type I error and the one-sided test. The values in the program can be changed in specific situations.

The output: Power = 0.077 < 0.75, the testing power is insufficient.

The needed advanced SAS program is as follows:

```
proc power;
onesamplemeans test=t
side=1 mean =0.2 stddev =5.1 ntotal =36 power = . ;
run;quit;
```

Program explanation: The option “test = t” requires to conduct the *t* test to the mean; “sides = 1” specifies the one-sided test; “mean = 0.2” specifies the difference between the population mean corresponding to the sample and the standard mean; “stddev = 5.1” refers to the sample standard deviation; “total = 36” specifies the sample size.

The output: power = 0.079 < 0.75. The testing power is insufficient.

The result of the advanced SAS program is a little higher than that of the SAS program based on formulas (3) and (4) due to the fact that the advanced program adopted the approximate algorithm. The difference between the two algorithms is that the advanced program used the quantile of the standard normal distribution, while the other used the quantile of the *t* distribution. When the sample size is small, the result of the program based on formulas (3) and (4) is more accurate, while the result of the advanced program is slightly higher than the real value.

### 1.2 Testing of the rate

**1.2.1 Formulas for sample size estimation** For one-sided test:

$$n = \left[ \frac{t_{1-\alpha, df} \sqrt{p_0(1-p_0)} + t_{1-\beta, df} \sqrt{p(1-p)}}{\theta} \right]^2 \quad (5)$$

For two-sided test:

$$n = \left[ \frac{t_{1-\alpha/2, df} \sqrt{p_0(1-p_0)} + t_{1-\beta, df} \sqrt{p(1-p)}}{\theta} \right]^2 \quad (6)$$

In formulas (5) and (6)<sup>[2]</sup>,  $p_0$  and  $p$  stand for the standard rate and the sample rate, respectively;  $\theta$  stands for the difference of the population rates of the sample group and the standard group, which can be replaced by its estimated value if unknown;  $df = n - 1$ . The iterative algorithm is also adopted for sample size estimation based on formulas (5) and (6). First, let  $df = \infty$  and denote the  $n$  as  $n_1$ , then let  $df = n_1 - 1$  and denote the  $n$  as  $n_2$ . The process continues until the nearest two “ $n$ ” are stable.

### 1.2.2 Formulas for testing power estimation

First, compute the value of  $t_{1-\beta, df}$  by formulas (7) and (8)<sup>[2]</sup>, then compute the value of  $1 - \beta$  accordingly, namely, the testing power.

For one-sided test:

$$t_{1-\beta, df} = \frac{|\theta| \sqrt{n} - t_{1-\alpha, df} \sqrt{p_0(1-p_0)}}{\sqrt{p(1-p)}}, \quad df = n - 1 \quad (7)$$

For two-sided test:

$$t_{1-\beta, df} = \frac{|\theta| \sqrt{n} - t_{1-\alpha/2, df} \sqrt{p_0(1-p_0)}}{\sqrt{p(1-p)}}, \quad df = n - 1 \quad (8)$$

The symbols in formulas (7) and (8) have the same meanings as those in formulas (5) and (6).

### 1.2.3 Examples

**1.2.3.1 Example 3** According to the experience, the curative rate of traditional drug A in treatment of a certain disease is 60%. A researcher expected to test the curative rate of a new drug B. The pilot test showed that the curative rate of drug B was 75%. Suppose that the testing level  $\alpha$  was 0.05 and the testing power was set to be 0.75. How many patients were needed in order to draw the

conclusion that drug B was better than the traditional drug A?

Analysis: Example 3 deals with the sample size estimation of the one-sided test for qualitative data with the single-group design. The known information includes the sample rate, the standard rate, the significance level and the testing power. Formula (5) can be adopted for sample size estimation.

The SAS program based on formula (5) is as follows:

```
%let p=0.75;%let p0=0.60;%let theta=0.15;%let alpha=0.05;%let beta=0.25;%let side=1;
data a3;
df=10 * * 9;
if &side=1 then do
n1=(((tinv(1-&alpha,df)) * sqrt(&p0 * (1-&p0)))+(tinv(1-&beta,df)) * sqrt(&p * (1-&p)))/&theta * * 2;
df=n1-1;
n2=(((tinv(1-&alpha,df)) * sqrt(&p0 * (1-&p0)))+(tinv(1-&beta,df)) * sqrt(&p * (1-&p)))/&theta * * 2;
do while(ABS(n1-n2)>1);
df=n2-1;n1=n2;
n2=(((tinv(1-&alpha,df)) * sqrt(&p0 * (1-&p0)))+(tinv(1-&beta,df)) * sqrt(&p * (1-&p)))/&theta * * 2;
end;end;
else do
n1=(((tinv(1-&alpha/2,df)) * sqrt(&p0 * (1-&p0)))+(tinv(1-&beta,df)) * sqrt(&p * (1-&p)))/&theta * * 2;
df=n1-1;
n2=(((tinv(1-&alpha/2,df)) * sqrt(&p0 * (1-&p0)))+(tinv(1-&beta,df)) * sqrt(&p * (1-&p)))/&theta * * 2;
do while(ABS(n1-n2)>1);
df=n2-1;n1=n2;
n2=(((tinv(1-&alpha/2,df)) * sqrt(&p0 * (1-&p0)))+(tinv(1-&beta,df)) * sqrt(&p * (1-&p)))/&theta * * 2;
end;end;
n=CEIL(MAX(n1,n2));
file print;
PUT #3 @15 n 'patients were needed';
run;quit;
```

Program explanation: The six “%let” statements in the program are used to set the sample rate, the standard rate, the rate difference, the probability of making type I error, the probability of making type II error and the one-sided test. In similar situations, readers only need to change the values of the parameters in the six “%let” statements to get the result.

The output: Totally 56 patients were needed.

There is no available advanced SAS program yet.

**1.2.3.2 Example 4** In example 3, suppose that the researcher did not estimate the sample size. Instead, he performed the one-sided test using the data of the 50 patients and drew the conclusion that drug B was not better than the traditional drug A. Estimate the testing power.

Below is the corresponding SAS program based on formula (7).

```
%let p=0.75;%let p0=0.60;%let theta=0.15;%let alpha=0.05;%let n=50;%let side=1;
data a4;
df=&n-1;
if &side=1 then p1=1-&alpha;else p1=1-&alpha/2;
t_1_beta_df=(abs(&theta) * sqrt(&n)-tinv(p1,df) * sqrt(&p0 * (1-&p0)))/sqrt(&p * (1-&p));
power=put(probt(t_1_beta_df,df),7.3);
if power<0.75 then do; sig='<0.75, ' ; mean="the testing power is insufficient. "; end;
if power>0.75 then do; sig='>0.75, ' ; mean="the testing power is sufficient. "; end ;
file print;
PUT #3 @15 'power= ' power sig mean;
run;
```

Program explanation: The six “%let” statements are used to specify the sample rate, the standard rate, the rate difference, the probability of making type I error, the probability of making type II error and the one-sided test. In similar situations, readers only need to change the values of the parameters in the six “%let” statements to get the result.

The output: power = 0.709 < 0.75, the test power is insufficient.

The needed advanced SAS program is as follows:

```
proc power;
onesamplefreq test=exact
nullproportion = 0.60
proportion = 0.75
sides = U
ntotal = 50
power = .;
run;
proc power;
onesamplefreq test=adjz
nullproportion = 0.60
proportion = 0.75
sides = U
ntotal = 50
power = .;
run;
proc power;
onesamplefreq test=z
nullproportion = 0.60
proportion = 0.75
sides = U
ntotal = 50
power = .;
run;
```

Program explanation: The option “test = exact | adjz | z” requires to perform the Fisher exact test or the adjusted Z test or the Z test; “sides = 1” requires to conduct the one-sided test; “nullproportion =” and “proportion =” are used to set the standard rate and the sample rate; “sides = U | L | 1 | 2” requires to conduct the upper one-sided test, the lower one-sided test or the two-sided test, respectively; “ntotal = 50” specifies the sample size; “power = .” means to estimate the testing power. The values of the parameters in the program can be altered in similar situations.

The output of the first procedure: power =

0.637<0.75. The testing power is insufficient.

The output of the second procedure: power =

0.637<0.75. The testing power is insufficient.

The output of the third procedure: power =

0.748<0.75. The testing power is insufficient.

The above three results are not exactly the same, and are slightly different from the result of the program based on formulas (7) and (8). The first two procedures have the same result, and are relatively conservative. The result of the third procedure may be slightly higher than the real value. The difference is caused by different hypothesis premise and different algorithms. When the sample size is small, the difference of the results of the above algorithms will be relatively larger. Therefore, the conservative result should be used. When the sample size is large, the results of the different algorithms will have slight difference. In that case, either algorithm is fine.

### 2 Estimation of sample size and testing power of difference test for quantitative and qualitative data with the paired design

The paired design refers to the design of which only two levels of one experimental factor is involved and the data obtained are paired<sup>[1]</sup>.

#### 2.1 Testing of the mean

**2.1.1 Formulas for sample size estimation** In order to analyze the quantitative data of the paired design, first, calculate the difference of each pair, and then compare the mean of the difference and 0 using statistical method of quantitative data with the single-group design. Thus, the sample size estimation of the paired design is the same as that of the single-group design, except that  $\theta$  in formulas (1) and (2) is changed into the mean of the sample difference,  $S$  is changed into the standard deviation of the sample difference and  $n$  is the number of the pairs.

**2.1.2 Formulas for testing power estimation** The testing power estimation of quantitative data with the paired design is the same as that with the single-group design, except that  $S$  in formulas (3) and (4) is changed into the standard deviation of the sample difference, and  $n$  is changed into the number of the pairs.

**2.1.3 Example** Since the formulas for sample size estimation and testing power estimation of quantitative data with the paired design is the same as that with the single-group design, we do not give examples in this part due to the space limit.

#### 2.2 Testing of the rate

**2.2.1 Formulas** Below is the standard 2×2 table of the paired design<sup>[2]</sup>:

Method B	Method A	
	+	-
+	a	b
-	c	d

Let  $\pi_{+-} = \frac{b}{a+b}$ ,  $\pi_{-+} = \frac{c}{a+c}$  and  $\pi_c = \frac{\pi_{+-} + \pi_{-+}}{2}$ , the needed number of the pairs can be computed by formula (9):

$$n = \left[ \frac{u_{1-\alpha/2} \sqrt{2\pi_c} + u_{1-\beta} \sqrt{\frac{2\pi_{-+}\pi_{+-}}{\pi_c}}}{\pi_{-+} - \pi_{+-}} \right]^2 \quad (9)$$

**2.2.2 Example 5** A certain kind of bacteria was inoculated in two mediums (A and B). The probability of A being positive and B being negative was  $\pi_{+-} = 0.04$ , and the probability of A being negative and B being positive was  $\pi_{-+} = 0.24$ . Suppose  $\alpha = 0.05$  and  $\beta = 0.10$  and the two-sided test was required. How many pairs were needed?

Analysis: Example 5 deals with the sample size estimation for qualitative data with the paired design having a binary responding variable. The inconsistent rates of the two methods and the probabilities of making type I error and type II error are given. The minimum number of the pairs is expected to estimate.

The SAS program based on formula (9) is as follows:

```

%let alpha=0.05;%let beta=0.10;%let pi10=0.04;%let pi01=0.24;
data a5;
pi=(&pi10+&pi01)/2;
u1=PROBIT(1-&alpha/2);
u2=PROBIT(1-&beta);
n=CEIL(((u1 * SQRT(2 * pi) + u2 * SQRT(2 * &pi10 * &pi01/pi))/(&pi10-&pi01)) * * 2);
file print;
put #2 @15 'the needed number of pairs is ' n';
run;quit;

```

Program explanation: The four “%let” statements are used to specify the probabilities of making type I error and type II error and the probabilities of the two inconsistent rates. In similar situations, readers only need to alter the values of the parameters in the four “%let” statements to get the result.

The output: The needed number of the pairs is 58.

The needed advanced SAS program is as follows:

```

PROC POWER;
pairedfreq dist=NORMAL discproppdiff = 0.2 totalproppdisc = 0.28 method=CONNOR npairs = . power = 0.90;
run;quit;
PROC POWER;
pairedfreq dist=NORMAL discproppdiff = 0.2 totalproppdisc = 0.28 method=MIETTINEN npairs = . power = 0.90;
run;quit;

```

Program explanation: The option “dist=NORMAL” specifies the normal distribution as the distribution of the testing statistics; “discproppdiff = 0.20”

specifies the difference of the two inconsistent rates  $\pi_{-+}$  and  $\pi_{+-}$ ; “totalprodisc = 0.28” stands for the sum of the two inconsistent rates  $\pi_{-+}$  and  $\pi_{+-}$ ; the options “method=CONNOR” and “method=MIETTINEN” specify the two different normal approximate algorithms (the results are slightly different).

The output: The needed number of the pairs is 70 (based on the Connor method) or 61 (based on the Miettinen method).

### 3 Estimation of sample size and testing power of difference test for quantitative data with the two-stage crossover design

The two-stage crossover design is adopted when there is a two-level experimental factor, and the researcher expects the two levels of the factor to be conducted on the same research subject in order. If the two levels of the experimental factor are performed on subjects that are paired according to similar conditions, it is called the paired two-stage crossover design. If the two levels of the experimental factor are performed on two groups of subjects, it is defined as the two-stage crossover design with two groups<sup>[1]</sup>.

The formulas for sample size estimation and testing power estimation for quantitative data

with the single-group design, namely, formulas (1), (2), (3) and (4) can also be applied to the two-stage crossover design. Here,  $\theta = \bar{x}_1 - \bar{x}_2$ ,  $\bar{x}_1$  and  $\bar{x}_2$  stand for the sample means of the two treatments, and  $S$  is the standard deviation of the difference of the two treatments<sup>[3]</sup>.

### REFERENCES

- 1 Hu LP. The application of statistical triple-type theory in experimental designs. Beijing: People's Military Medical Press. 2006: 46, 48, 77. Chinese.  
胡良平. 统计学三型理论在实验设计中的应用. 北京: 人民军医出版社. 2006: 46, 48, 77.
- 2 Hu LP. SAS research design and statistical analysis. Beijing: People's Medical Publishing House. 2010: 193, 219. Chinese.  
胡良平. SAS 实验设计与统计分析. 北京: 人民卫生出版社. 2010: 193, 219.
- 3 Jin PH, Chen F. Medical statistics methods. 3rd ed. Shanghai: Fudan University Press. 2009: 565. Chinese.  
金丕焕, 陈峰. 医学统计方法. 第 3 版. 上海: 复旦大学出版社. 2009: 565.

## 样本量估计与检验效能分析(五)

胡良平, 鲍晓蕾, 关雪, 周诗国

军事医学科学院生物医学统计学咨询中心, 北京 100850

**摘要:** 样本含量与检验效能估计是研究设计的重要内容之一。本文介绍了拟作单组设计、配对设计和交叉设计定量与定性资料差异性检验时的样本含量和检验效能估计方法, 具体地说, 就是针对前述的 3 种设计且为差异性检验的情形, 介绍了效应指标为定量和定性资料时估计样本含量和检验效能的计算公式, 以及基于公式和借用 SAS 软件中的 POWER 过程分别实现样本含量和检验效能估计的方法, 并通过实例进行了演示和讲解, 对实际工作者在试验设计时正确落实重复原则具有很好的指导作用。

**关键词:** 统计学; 医学; 研究设计; 样本大小; 检验效能; 参数估计