Estimation of sample size and testing power (Part 4)

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: Sample size estimation is necessary for any experimental or survey research. An appropriate estimation of sample size based on known information and statistical knowledge is of great significance. This article introduces methods of sample size estimation of difference test for data with the design of one factor with two levels, including sample size estimation formulas and realization based on the formulas and the POWER procedure of SAS software for quantitative data and qualitative data with the design of one factor with two levels. In addition, this article presents examples for analysis, which will play a leading role for researchers to implement the repetition principle during the research design phase.

KEYWORDS: statistics, medical, research design, sample size, testing power, parametric estimation

In the last two issues of J Chin Integr Med, we have introduced sample size estimation of three special tests, namely, the equivalence test, the superiority test and the non-inferiority test. This article will introduce sample size estimation of difference test for quantitative and qualitative data with the design of one factor with two levels. The difference test refers to the research design of which the objective is to analyze whether two population means or two population rates represented by two samples are equal. As to quantitative data, if the conditions for parametric test, namely, independence, normality and equal variance, are satisfied, the \( t \) test of quantitative data with the design of one factor with two levels is adopted; otherwise, the rank-sum test is adopted. In terms of qualitative data, the \( \chi^2 \) test or the Fisher exact test should be used.

1 How to estimate the sample size of difference test for quantitative data with the design of one factor with two levels

1.1 Estimation formula For one-sided test, let \( n_T = k n_R \), then

\[
n_R = \left[ \frac{1}{2} \left( \frac{1}{\alpha} \right) \frac{s_2}{\sigma_2} + \frac{1}{2} \left( \frac{1}{\beta} \right) \frac{s_2}{\sigma_2} \right]^2 \left( \frac{\sigma^2 / k + \sigma^2}{\delta_1 - \delta_0} \right)^2
\]

For two-sided test, let \( n_T = 2 n_R \), then
\[ n_R = \left[ t_{1-\alpha/2, (k+1)\sigma_R^2} + t_{1-\beta, (k+1)\sigma_R^2} \right]^2 \left( \sigma_1^2 / (k + \sigma_R^2) \right) / (\delta_1 - \delta_0)^2 \] (2)

Analysis: Example 1 involves the issue of sample size estimation for quantitative data with the design of one factor with two levels. The known information includes the estimated difference of the two efficacies, the sample standard deviation, the allowed probabilities of making type I error and type II error in the hypothesis testing and the sample size ratio between the two groups. The minimum sample size is required to be estimated. The SAS program based on formulas (1) and (2) is as follows:

```sas
%let alpha = 0.05; %let beta = 0.10; %let k = 1; %let xbar_t = 139; %let xbar_r = 104; %let s = 20; %let s_r = 25; %let delta_0 = 0; %let side = 2; data a1;
  delta_1=xbar_t-xbar_r;
  theta=delta_1-delta_0;
  if theta=0 then do; file print;
  put #3 'note: because theta=0, it is unable to estimate the sample size. A new reasonable value for delta_1 is needed for further estimation';
goto warning;end;
else do;
  if side=1 then p1=1-alpha; else p1=1-alpha/2; p2=1-beta; df=10; t=tinv(p1,df); t2=tinv(p2,df);
  n=ceil(((t+t2) * 2 + (delta_t * 2 + delta_r * 2) / (theta * 2));
  df=(delta_t+1)+n-2; t=tinv(p1,df); t2=tinv(p2,df); n2=ceil(((t+t2) * 2 + (delta_t * 2 + delta_r * 2) / (theta * 2));
  do while (abs(n1-n2)>1);
  n=(n1+n2)/2; df=(delta_t+1)+n1-2; t=tinv(p1,df); t2=tinv(p2,df); n2=ceil(((t+t2) * 2 + (delta_t * 2 + delta_r * 2) / (theta * 2));
  end;
  n_r=max(n1,n2); n_t=delta_r*n_r; file print;
  if side=1 then put #10 'mean comparison for quantitative data with the design of one factor with two levels the one-sided test: the needed sample size: n_r=' n_r=' n_t='; else put #10 'mean comparison for quantitative data with the design of one factor with two levels the two-sided test: the needed sample size: n_r=' n_t='; end;
warning:;
runtime quitting;
```

1.2 Example 1 A study expected to investigate the effects of Wuxin capsule on heart rate variability of patients suffering from coronary heart disease. The observational index was standard deviation of NN intervals (SDNN). The researchers planed to select a certain number of patients with coronary heart disease and divide them into experimental group and control group. The result of pilot study showed that the mean of SDNN of the experimental group was 139 and the standard deviation was 20, while the mean of SDNN of the control group was 104 and the standard deviation was 25. Suppose that the result would be analyzed by two-sided difference test and \( a \) was set as 0.05. The sample size ratio between the experimental group and the control group was 1:1. The testing power was required to be 90%. How many patients were needed for this clinical study?

Related Articles 推荐阅读


Output: the needed sample size \( n_r = 10 \) and \( n_t = 10 \).

Program explanation: The 9 "\%let" statements in the beginning part of the program respectively specify the following information: the allowed probability of making type I error, the allowed probability of making type II error, the sample size ratio between the experimental group and the control group, the mean of the experimental group, the mean of the control group, the standard deviation of the experimental group, the standard deviation of the control group, the population efficacy difference between the two groups in the null hypothesis and the two-sided test. In similar situations, readers only need to alter the parameter values in the 9 "\%let" statements to get the result.

In addition, we can also use the POWER procedure in SAS to estimate the sample size. Before invoking the POWER procedure, we should assume that the standard deviations of the two groups are equal. Here, we may assume that the standard deviations of the two groups are 20 and 25, respectively. The needed program is as follows:

```sas
proc power;
twosamplemeans test=diff
sides=2
meandiff = 35
stddev = 20 25
groupweights = (1 1)
ntotal = .
power = 0.90;
run; quit;
```

\[
n_R = \left[ t_{1-\alpha/(1+k)n_R-\alpha} + t_{1-\beta/(1+k)n_R-\beta} \right]^2 \left[ P_T(1-P_T)/k + P_R(1-P_R)/(\delta_1 - \delta_0)^2 \right] (3)
\]

For two-sided test\(^{(1)}\), let \( n_T = k n_R \), then

\[
n_R = \left[ t_{a/2,(1+k)n_R-\alpha} + t_{1-\beta/(1+k)n_R-\beta} \right]^2 \left[ P_T(1-P_T)/k + P_R(1-P_R)/(\delta_1 - \delta_0)^2 \right] (4)
\]

In formulas (3) and (4), \( P_T \) and \( P_R \) represent the sample size rates of the experimental group and the control group, respectively. \( \delta_0 \) is the difference of curative effects between the experimental group and the control group in the null hypothesis, which depends on specialty. If \( \delta_0 \) is not specified, it equals to 0. \( \delta_1 \) is the difference of curative effects between the experimental group and the control group in the alternative hypothesis, which also equals to 0 if not specified.

For each test, \( \delta_1 - \delta_0 \) is a parameter whose expected value is a new reasonable value before the sample size estimation begins. In addition, when \( \delta_0 = 0 \), \( \delta_1 - \delta_0 = \delta_1 - 0 \) = \( \delta \).

2.2 Example 2 A researcher expected to study the efficacy of urinary kallidinogenase in treatment of cerebral infarction. He planned to select a certain number of patients suffering from cerebral infarction and divide them into experimental group and control group. The result of the pilot test showed that the efficient rate of the experimental drug was 96.2\%, while that of the control drug was 70.35\%. Suppose that the two-sided difference test would be adopted and \( \alpha \) was set to be 0.05. The sample size ratio between the experimental group and the control group was 1:1. The testing power was required to be 80\%. How many patients were needed for the clinical research?

Analysis: Example 2 involves the issue of sample size estimation for qualitative data with the design of one factor with two levels. The known information includes the two efficient rates of the two groups, the allowed probabilities of making type I error and type II error in the hypothesis testing and the sample size ratio of the two groups. The minimum sample size is required to be estimated. The SAS program based on formulas (3) and (4) is as follows:
control group, the population efficacy difference between the two groups in the null hypothesis and the two-sided test. In similar research, readers only need to alter the parameter values in the "%let" statements to get the result.

Furthermore, we can also use the POWER procedure in SAS to estimate the sample size.

```
PROC POWER;
  TwoSampleFreq test=PChi
  GroupProportions=(0.962 0.703)
  sides=2
  GroupWeights=(1 1)
  total=.
  alpha=0.05
  power=0.80;
run;quit;
```

Output:

```
  Actual Power       N total
  0.807       64
```

The result shows that the actual testing power reaches 0.807 and 64 patients were needed, with 32 for each group.

Program explanation: The option "test=PChi" specifies the Pearson χ² test for data analysis; the option "GroupProportions = (0.962 0.703)" specifies the two efficient rates of the two groups; the other options are the same as in example 1. The values in the program can be altered in similar situations.

REFERENCES


样本量估计与检验效能分析(四)

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 1000850

摘要: 样本含量估计是任何一项实验研究或调查研究都不可回避的问题。根据已知的基本情况和统计学知识，估计出合适的样本含量在科研实践上是十分有意义的。本文介绍了拟作成组设计定量资料与定性资料差异性检验时的样本含量估计方法。具体地说，就是根据成组设计且为差异性检验的情形，介绍了寿命指数为定量资料和定性资料时估计样本含量的计算公式，以及基于公式和借用 SAS 软件中 POWER 过程分别实现样本含量估计的方法，并通过实例进行了讲解，对科研工作者在实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