

# 如何正确运用 Z 检验——两总体率比较 优效性 Z 检验及 SAS 实现

胡纯严<sup>1</sup>, 胡良平<sup>1,2\*</sup>

(1. 军事科学院研究生院, 北京 100850;

2. 世界中医药学会联合会临床科研统计学专业委员会, 北京 100029

\*通信作者: 胡良平, E-mail: lphu927@163.com)

**【摘要】** 本文目的是介绍两总体率比较优效性 Z 检验的原理、方法和 SAS 实现。文中涉及两个方面的内容, 包括两总体率比较优效性 Z 检验的原理和方法以及基于 SAS 软件实现两总体率比较优效性 Z 检验。文中给出了 SAS 输出结果并对其进行解释, 并分别得出统计学结论和专业结论。

**【关键词】** 总体率; 正态分布; 优效性检验; 优效性界值

中图分类号: R195.1

文献标识码: A

doi: 10.11886/scjsws20201119003

## How to use Z test correctly——comparison of two population rates for the superiority Z test and the SAS implementation

Hu Chunyan<sup>1</sup>, Hu Liangping<sup>1,2\*</sup>

(1. Graduate School, Academy of Military Sciences PLA China, Beijing 100850, China;

2. Specialty Committee of Clinical Scientific Research Statistics of World Federation of Chinese Medicine Societies, Beijing 100029, China

\*Corresponding author: Hu Liangping, E-mail: lphu927@163.com)

**【Abstract】** The purpose of this paper was to introduce the principles, methods and SAS implementation of the comparison of two population rates by using the superiority Z test. Two special titles were involved as follows: one was to introduce the principles and methods of the superiority Z test for the comparison of two population rates, the other was to implement the superiority Z test for the comparison of two population rates based on the SAS software. The SAS output results were presented and explained, and also the statistical conclusions as well as the professional conclusions were given in this paper.

**【Keywords】** Population rate; Normal distribution; Superiority test; Superiority dividing value

虽然统计分析方法多种多样, 但临床科研工作者最常用的分析方法是一般差异性检验<sup>[1-2]</sup>。为了使假设检验的结果不仅具有统计学意义, 还具有临床实际意义, 在处理临床资料时常需要选用 3 种特殊的差异性检验, 即非劣效性检验、等效性检验和优效性检验<sup>[3]</sup>。本文将结合临床实际, 介绍与“两总体率比较优效性检验”有关的内容。

## 1 两总体率比较优效性 Z 检验的基本原理和方法

### 1.1 基本概念

优效性检验是指主要研究目的为显示试验药的治疗效果优效于对照药(安慰剂对照或阳性对照)的试验。在试验设计阶段需要设定一个界值  $\delta_0$ , 来界定试验药的优效性<sup>[3-4]</sup>。

### 1.2 问题与数据结构

**【例 1】**为观察复方薤白胶囊治疗慢性支气管炎急性发作期痰热郁肺证的效果。随机选取 240 例符合要求的患者, 均分为 2 组, 试验组口服复方薤白胶囊, 对照组口服芩暴红止咳胶囊, 疗程均为 10 d。两组的有效率分别为 96.67% (有效例数  $n=116$ ) 和 92.50% (有效例数  $n=111$ ), 如果试验组的有效率超过对照组 5%, 才能认为复方薤白胶囊优效于芩暴红止咳胶囊, 试评价复方薤白胶囊是否优效于芩暴红止咳胶囊<sup>[5]</sup>。临床资料见表 1。

表 1 两组患者的治疗效果

药物种类	例数( $n$ )	有效率(%)
复方薤白胶囊	120	96.67
芩暴红止咳胶囊	120	92.50

**【对数据结构的分析】**该资料的试验因素为“药物种类”, 其两个水平分别为“复方薤白胶囊”与“芩

暴红止咳胶囊”,观测指标为“有效率”,所以该资料为成组设计一元定性资料。

【统计分析方法的选择】研究者的研究目的是评价复方薤白胶囊治疗慢性支气管炎急性发作期痰热郁肺证的效果是否优于芬暴红止咳胶囊治疗效果,并且设定了优效性界值 $\delta_U=5%$ ,这时应选用优效性检验。

【例2】某一临床试验中观察某新药A对某种肿瘤临床症状的控制效果,以某标准药物B作对照行随访研究。所有患者同时入组,观察期为1年,主要评价指标为症状改善率。临床资料见表2,如果设定优效性界值为6%,试判断药物A是否优于药物B。临床资料见表2。

表2 两组患者的症状改善情况

药物种类	例数(n)	改善率(%)
A	130	62.31
B	130	45.38

注:本例作为练习,请读者完成

### 1.3 基本原理

#### 1.3.1 检验假设

给出检验假设,并规定检验水准 $\alpha$ 的值。

$$H_0: \pi_T - \pi_R \leq \delta_U;$$

$$H_1: \pi_T - \pi_R > \delta_U;$$

$$\alpha = 0.05(\text{单侧}).$$

#### 1.3.2 检验统计量

所需要的检验统计量如下式:

$$Z = \frac{P_T - P_R - \delta_U}{\sqrt{\frac{P_T(1-P_T)}{n_T} + \frac{P_R(1-P_R)}{n_R}}} \sim N(0,1) \quad (1)$$

拒绝域为 $Z > Z_{1-\alpha}$ 。

在式(1)中,检验统计量 $Z$ 服从标准正态分布; $(P_T, n_T)$ 与 $(P_R, n_R)$ 分别为试验组(T)与对照组(R)的“样本率、样本含量”; $\delta_U$ 为优效性界值。

#### 1.3.3 基于检验统计量 $Z$ 的取值计算单侧尾端概率

根据检验水准查表或借助相应的SAS函数进行计算,获得检验统计量 $Z$ 的临界值(针对手工计算而言)和(或)标准正态分布尾端的概率(针对统计软件而言)。

#### 1.3.4 得出统计学结论和专业结论

根据拒绝还是不能拒绝 $H_0$ 的结果,给出统计学结论,最后结合专业知识给出专业结论。

## 2 两总体率比较优效性检验的SAS实现

### 2.1 基于SAS语言按公式编程间接实现计算

基于SAS语言和标准正态分布函数等知识<sup>[6]</sup>并按式(1)进行编程,就可间接实现两总体率比较优效性 $Z$ 检验。

SAS程序如下:

```
%let n_t=120;
%let n_r=120;
%let p_t=0.966667;
%let p_r=0.925000;
%let delta_U=0.05;
data a;
nu=((&p_t-&p_r)-&delta_U);
A=&p_t*(1-&p_t)/&n_t;
B=&p_r*(1-&p_r)/&n_r;
nd=sqrt(A+B);
Z=nu/nd;
p=1-probnorm(Z);
run;
proc print noobs data=a;
var Z p;
run;
```

【程序说明】变量名后面的“\_t”与“\_r”分别代表“试验组”与“对照组”相应的量(包括样本含量与有效率);有效率用6位小数表示的目的是使计算结果尽可能精确一些,这样就与后面基于FREQ过程直接计算的结果具有一定的可比性。

【SAS输出结果】

Z	p
-0.28639	0.61271

【统计学结论】因 $Z=-0.28639$ , $P=0.61271$ (单侧概率) $>0.05$ ,所以,应接受 $H_0$ 。

【专业结论】在优效界值 $\delta_U=5%$ 的前提条件下,尚不能认为复方薤白胶囊治疗慢性支气管炎急性发作期痰热郁肺证的效果优于芬暴红止咳胶囊的治疗效果。

### 2.2 基于SAS中FREQ过程直接实现计算

由于SAS中FREQ过程包含了几乎绝大多数定性资料假设检验方法<sup>[7]</sup>,所以,此处将采用此过程来直接实现两总体率比较优效性 $Z$ 检验。

SAS程序如下:

```
data a;
```

```
do a=1 to 2;
do b=1 to 2;
input x @@;
output;
end;
end;
cards;
116 4
111 9
;
```

```
run;
proc freq data=a;
tables a*b/riskdiff(sup MARGIN=0.05);
weight x;
run;
```

【程序说明】各组中的数据应为“有效例数”与“无效例数”，而不适合直接采用“有效率”；程序中 sup 表示进行优效性检验；MARGIN=0.05 表示优效界值。

【SAS主要输出结果及解释】

比例(风险)差值的优效性分析

H0:P1-P2≤边际 Ha:P1-P2>边际  
 边际=0.05 Wald方法

比例差值	ASE(样本)	Z	Pr>Z	优效性限值	90% 置信区间
0.0417	0.0291	-0.2864	0.6127	0.0500	-0.0062 0.0895

因  $Z=-0.2864, P=0.6127$ , 说明尚不能认为复方薤白胶囊治疗慢性支气管炎急性发作期痰热郁肺证的效果优效于芩暴红止咳胶囊的治疗效果。若利用置信区间来判断, 由计算结果可知: 两有效率之差的 90% 置信区间为  $[-0.0062, 0.0895]$ , 因置信区间包含  $\delta_0=0.05$ , 说明优效性结论不成立。

切关系。现以本文例 2 数据为例, 若取显著性水平  $\alpha=0.05, \delta_0=6%$ , 得到的计算结果为  $Z=1.7925, P=0.0365$  (单侧概率)  $< \alpha=0.05$ , 即可以认为药物 A 优效于药物 B。若取显著性水平  $\alpha=0.05$ , 但取  $\delta_0=10%$ , 得到的计算结果为  $Z=1.1361, P=0.1280$  (单侧概率)  $> \alpha=0.05$ , 即不能认为药物 A 优效于药物 B。

3 讨论与小结

3.1 讨论

3.1.1 影响优效性检验结果的因素

3.1.1.1 概述

优劣效性检验结果不是绝对的, 随着条件的改变, 最终的检验结果也会发生相应的改变, 影响其结果的因素有“显著性水平  $\alpha$ ”“优效性界值  $\delta_0$ ”和“样本含量”。

3.1.1.2  $\alpha$  的取值对优效性检验结果的影响

优效性检验结果不是绝对的, 与显著性水平  $\alpha$  的取值有密切关系。现以本文例 2 数据为例, 若取显著性水平  $\alpha=0.05, \delta_0=6%$ , 得到的计算结果为  $Z=1.7925, P=0.0365$  (单侧概率)  $< \alpha=0.05$ , 即可以认为药物 A 优效于药物 B。若取显著性水平  $\alpha=0.025, \delta_0=6%$ , 得到的计算结果为  $Z=1.7925, P=0.0365$  (单侧概率)  $> \alpha=0.025$ , 即不能认为药物 A 优效于药物 B。

3.1.1.3  $\delta_0$  的取值对优效性检验结果的影响

优效性检验结果与优效性界值  $\delta_0$  的取值有密切

3.1.1.4 样本含量对优效性检验结果的影响

现以本文例 2 数据为例, 若取显著性水平  $\alpha=0.05, \delta_0=6%$ , 若直接按表 2 中的数据进行优效性检验, 其结论为“药物 A 优效于药物 B”。若将各组的样本含量各减少一半, 即 A 药物组有效例数为 41 例 (有效率为 62.12%)、无效例数为 25 例; B 药物组有效例数为 30 例 (有效率为 45.45%)、无效例数为 36 例, 基本保持两组的有效率分别与表 2 中的接近。此时, 优效性检验的结果为  $Z=1.2466, P=0.1063$  (单侧概率)  $> \alpha=0.05$ , 即不能认为药物 A 优效于药物 B, 与前面的结论正好相反。需注意的是, 在拟开展优效性试验研究之初期, 就应当按照已知的条件 [包括两组估计的有效率、显著性水平  $\alpha$  的值、检验效能  $(1-\beta)$  的值和优效性界值  $\delta_0$  的值], 找到相应的计算公式, 估计出合适的样本含量<sup>[3]</sup>。

3.1.2 主要疗效评价指标的临床意义

在进行优效性检验时, 一个常被忽视的问题是评价指标的临床意义。具体地说, 就是应明确当前的评价指标究竟属于“高优指标 (指标取值越大越好, 如治愈率)”还是“低优指标 (指标取值越小越

好,如死亡率)”。当面对高优指标时,通常情况下,试验组指标的取值高于对照组指标的取值时才适合考虑选用“优效性检验”;如果在较大样本含量的预试验中,已获得的试验结果是“试验组指标的取值明显低于对照组指标的取值”,若此时仍坚持选用“优效性检验”就不太符合常理了。按正常思维,此时应当考虑选用“非劣效性检验”或“等效性检验”。

### 3.1.3 $\delta_{\nu}$ 是否总是取正值

在进行优效性检验时,统计学教科书上通常都默认把 $\delta_{\nu}$ 取为“正值”,代表试验药疗效比对照药疗效“好”。事实上,这样做有一个隐含的前提,即疗效的评价指标为“高优指标”。而当疗效的评价指标为低优指标时,若拟采用优效性检验,此时,其优效界值 $\delta_{\nu}$ 应取负值。

## 3.2 小结

本文介绍了两总体率比较优效性 $Z$ 检验的原理、方法和SAS实现。在SAS实现方面,介绍了基于SAS语言编程间接实现两总体率比较优效性 $Z$ 检验以及基于SAS中FREQ过程直接实现与前面提及的相同计算。在后者中,还给出了两总体率之差的置

信区间。基于置信区间也可以判断优效性结论是否成立。

## 参考文献

- [1] 杜韵华,傅礼洪,吴伟博,等. 喹硫平与奥氮平治疗癫痫性精神障碍的对照研究[J]. 四川精神卫生, 2018, 31(3): 226-229.
- [2] 肖爱祥,叶君荣,于林,等. 穴位埋线治疗非典型抗精神病药物所致肥胖的随机对照研究[J]. 四川精神卫生, 2018, 31(3): 230-235.
- [3] 胡良平,陶丽新. 临床试验设计与统计分析[M]. 北京: 军事医学科学出版社, 2013: 71-74, 88-100, 111-118.
- [4] 姚婷婷,李长平,胡良平,等. 如何正确运用 $t$ 检验——两算术均值比较优效性 $t$ 检验及SAS实现[J]. 四川精神卫生, 2020, 33(4): 297-301.
- [5] 胡良平. SAS常用统计分析教程[M]. 2版. 北京: 电子工业出版社, 2015: 388-390.
- [6] 胡良平,胡纯严. SAS语言基础与高级编程技术[M]. 北京: 电子工业出版社, 2014: 37-97.
- [7] SAS Institute Inc. SAS/STAT®15.1 user's guide[M]. Cary, NC: SAS Institute Inc, 2018: 2997-3216.
- [8] 胡良平. 现代医学统计学[M]. 北京: 科学出版社, 2020: 244-257.

(收稿日期:2020-11-19)

(本文编辑:戴浩然)