

如何正确运用 Z 检验——总体方差已知时两算术均值比较一般差异性 Z 检验及 SAS 实现

胡良平^{1,2*}

(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】 本文的目的是介绍总体方差已知时两算术均值比较一般差异性 Z 检验及 SAS 实现。首先, 介绍了选用 Z 检验的前提条件; 其次, 介绍了单组设计和成组设计条件下进行均值比较的 Z 检验统计量, 并介绍了上单侧检验、下单侧检验和双侧检验时所对应的拒绝域; 最后, 结合两个实例, 基于 SAS 软件实现 3 种 Z 检验, 对输出结果做出了解释。并针对如何合理选择单侧检验还是双侧检验给予说明。

【关键词】 正态分布; Z 检验; 单侧检验; 双侧检验; 拒绝域

中图分类号: R195.1

文献标识码: A

doi: 10.11886/scjsws20200916004

How to use Z test correctly——comparison of two arithmetic mean values for the general difference Z test under the condition of known the population variance and the SAS implementation

Hu Liangping^{1,2*}

(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 comparison of two arithmetic mean values for the general difference Z test under the condition of known the population variance and the SAS implementation. Firstly, the preconditions of the Z test was introduced. Secondly, the Z statistics were given for the comparison of two arithmetic mean values for the general difference Z test under the designs of the single group and two groups. Some basic concepts, such as upper one-sided test, lower one-sided test, two-sided test and the rejection range, were introduced. Lastly, three kinds of Z tests were performed based on the SAS software combined with two real examples, and the output results were also explained. Furthermore, the necessary explanation was given about how to select one-sided test or two-sided test rationally.

【Keywords】 Normal distribution; Z test; One-sided test; Two-sided test; Rejection range

当研究者收集到单组设计或成组设计一元定量资料时, 若分析目的是对定量的效应指标(通常为均值)进行一般差异性检验, 此时, 人们依据定量资料所满足的前提条件, 拟从以下几种假设检验方法中选择一个最合适的, 即“Z 检验”“t 检验”或“秩和检验”。本文将重点阐述选择“Z 检验”的要领和用 SAS 实现计算的方法。

1 选用 Z 检验的前提条件

1.1 单组设计一元定量资料及选用 Z 检验的前提条件

【例 1】 文献[1]的目的是探讨新冠肺炎疫情期间封闭管理精神科医护人员焦虑抑郁症状及相关

因素, 为改善其心理健康状况提供参考。通过问卷调查的方式收集了包括“性别、年龄、职业、受教育程度”等 12 个变量(即项目)及其取值, 涉及到两个量表, 分别为“广泛性焦虑量表(简称 GAD-7)”和“患者健康问卷抑郁量表(简称 PHQ-9)”。共发放问卷 124 份, 回收有效问卷 121 份。现假定全部 121 例受试者的 PHQ-9 评分的样本均值 $\bar{x}=5.12$, 该定量指标总体方差 $\sigma^2=22$; 进一步假定 PHQ-9 评分的总体均值 $\mu_0=5.10$ 。

针对评价指标“PHQ-9”, 设文献[1]中样本(指 $n=121$ 例受试者的 PHQ-9 测定值)所抽自的总体均值为 μ (未知), 试回答下列 3 个问题。

问题 1: μ 是否小于或等于 μ_0 ?

问题2: μ 是否大于或等于 μ_0 ?

问题3: μ 是否等于 μ_0 ?

以上3个问题可以采用统计学上“零假设”与“备择假设”严格地呈现出来:

$$H_0: \mu \leq \mu_0 \leftrightarrow H_1: \mu > \mu_0 \quad (1)$$

$$H_0: \mu \geq \mu_0 \leftrightarrow H_1: \mu < \mu_0 \quad (2)$$

$$H_0: \mu = \mu_0 \leftrightarrow H_1: \mu \neq \mu_0 \quad (3)$$

式(1)叫做上(或右)单侧检验,式(2)叫做下(或左)单侧检验,而式(3)叫做双侧检验。这里的“名词概念(指上、下、双侧检验)”是基于“备择假设”而得出的。

解决上述3个问题都可采用Z检验,其检验统计量完全相同,但拒绝域(概念详见后文)是不同的;使用Z检验的前提条件:样本所抽自的总体为正态分布(但总体均值 μ 是未知的),总体方差为已知的常数(例如在题干中已假定 $\sigma^2=22$);同时,还需要给定待检验的总体均值(例如在题干中已假定 $\mu_0=5.10$)。

1.2 成组设计一元定量资料及选用Z检验的前提条件

【例2】假定文献[1]中全部121例受试者中“行为类型”为“B型和中间型”的PHQ-9评分的样本均值 $\bar{x}_1=4.41, n_1=98$;“A型”的PHQ-9评分的样本均值 $\bar{x}_2=8.18, n_2=23$ 。进一步假定:该定量指标在两个总体中的方差分别为 $\sigma_1^2=23.3, \sigma_2^2=6.4$ 。

针对评价指标“PHQ-9”,设文献[1]中两组样本分别为:

第1组样本(指 $n=98$ 例“B型和中间型”受试者的PHQ-9评分)所抽自的总体均值 μ_1 ;第2组样本(指 $n=23$ 例“A型”受试者的PHQ-9评分)所抽自的总体均值 μ_2 。

试回答下列3个问题:

问题4: μ_1 是否小于或等于 μ_2 ?

问题5: μ_1 是否大于或等于 μ_2 ?

问题6: μ_1 是否等于 μ_2 ?

以上3个问题可以采用统计学上“零假设”与“备择假设”严格地呈现出来:

$$H_0: \mu_1 \leq \mu_2 \leftrightarrow H_1: \mu_1 > \mu_2 \quad (4)$$

$$H_0: \mu_1 \geq \mu_2 \leftrightarrow H_1: \mu_1 < \mu_2 \quad (5)$$

$$H_0: \mu_1 = \mu_2 \leftrightarrow H_1: \mu_1 \neq \mu_2 \quad (6)$$

式(4)叫做上(或右)单侧检验,式(5)叫做下(或左)单侧检验,而式(6)叫做双侧检验。这里的

“名词概念(指上、下、双侧检验)”是基于“备择假设”而得出的。

解决上述3个问题都可以采用Z检验,其检验统计量是完全相同的,但拒绝域(概念详见后文)是不同的;使用Z检验的前提条件:各样本均取自正态分布总体(但各总体均值 μ_1, μ_2 都是未知的),各总体方差均为已知的常数(例如,在题干中已假定两个总体的方差分别为 $\sigma_1^2 = 23.3, \sigma_2^2 = 6.4$)。

2 单组设计一元定量资料Z检验及SAS实现

2.1 Z检验统计量

对服从正态分布的单组设计一元定量资料进行均值比较,若已知总体方差并给定定量评价指标的标准值(即理论均值 μ_0)时,其对应的检验统计量见下式:

$$Z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} \quad (7)$$

式(7)中的 $Z \sim N(0, 1)$ 。

2.2 拒绝域

依据随机变量的概率分布(本文指“标准正态分布”),在其横坐标轴上确定一个范围或区间,当检验统计量所代表的随机变量(本文指“Z”)的取值落入该区间内时,就有充足的理由(以一个较大的概率值来说话,单、双侧检验时均为“ $1-\alpha$ ”,其中,“ α ”为事先设定的检验水平)拒绝 H_0 ,从而接受 H_1 。在统计学上,就称该区间为与特定检验假设对应的“拒绝域”。

与式(1)对应的拒绝域W见下式^[2]:

$$W = (\bar{x}: Z > Z_{1-\alpha}) \quad (8)$$

与式(2)对应的拒绝域W见下式:

$$W = (\bar{x}: Z < Z_\alpha) \quad (9)$$

与式(3)对应的拒绝域W见下式:

$$W = (\bar{x}_1 - \bar{x}_2: |Z| > Z_{1-\alpha/2}) \quad (10)$$

【说明】在上面的三个式子中,不等号之后的内容均为标准正态分布横坐标轴上的“分位数”,其位置在标准正态分布曲线下左侧尾端概率分别为“ $1-\alpha$ ”、“ α ”和“ $1-\alpha/2$ ”的横坐标轴上。

2.3 例1的SAS实现

SAS程序如下:

```

%let n=121; /*样本含量*/
%let sm=5.12; /*样本均值*/
%let mu=5.10; /*已知的总体均值*/
%let pv=22; /*已知的总体方差*/
%let a=0.05; /*显著性水平*/
data abc;
z=(&sm-&mu)/(sqrt(&pv/&n));
absz=abs(z);
za=probit(&a);
z1_a=probit(1-&a);
z1_ha=probit(1-&a/2);
PU=1-probnorm(Z);
PL=probnorm(Z);
PT=2*(1-probnorm(absz));
title1 '上单侧检验结果';
proc print data=abc;
var z z1_a PU;
footnote1 '上单侧检验结果的判定:';
footnote2 '若 z>z1_a,则接受(H1:μ>μ0)';
footnote3 '否则,就接受(H0:μ≤μ0)';
footnote4 'PU为上单侧概率。';
run;
title1 '下单侧检验结果';
proc print data=abc;
var z za PL;
footnote1 '下单侧检验结果的判定:';
footnote2 '若 z<za,则接受(H1:μ<μ0)';
footnote3 '否则,就接受(H0:μ≥μ0)';
footnote4 'PL为下单侧概率。';
run;
title1 '双侧检验结果';
proc print data=abc;
var absz z1_ha PT;
footnote1 '双侧检验结果的判定:';
footnote2 '若 |z|>z1_ha,则接受(H1:μ≠μ0)';
footnote3 '否则,就接受(H0:μ=μ0)';
footnote4 'PT为双侧概率。';
run;

```

【程序说明】“%let”定义“宏变量”，用户可根据资料的实际情况修改宏变量的取值；“&”后接宏变量名，表示引用宏变量；“absz”代表z的绝对值；“za、z1_a、z1_ha”都是标准正态分布横坐标轴上的分位数（对应于标准正态分布曲线下左侧的概率分别为 α 、 $1-\alpha$ 、 $1-\alpha/2$ ）；“ α ”为事先给定的显著性水平；PU、

PL和PT分别代表上单侧、下单侧和双侧检验的概率。

【SAS主要输出结果及解释】

上单侧检验结果

Obs	z	z1_a	PU
1	0.046904	1.64485	0.48129

输出结果的第1部分[回答上述式(1)所提出的问题]是“上单侧检验的结果”。上单侧检验结果的判定：若 $z > z_{1-a}$ ，则接受 $(H_1: \mu > \mu_0)$ ；否则，就接受 $(H_0: \mu \leq \mu_0)$ ；PU为上单侧概率。因检验统计量 $z = 0.046904 < 1.64485$ ，故只能接受“ $H_0: \mu \leq \mu_0$ ”，即认为样本所抽自的总体均值 μ 小于等于给定的均值 $\mu_0 = 5.10 (P = 0.48129 > 0.05)$ 。

下单侧检验结果

Obs	z	za	PL
1	0.046904	-1.64485	0.51871

输出结果的第2部分[回答上述式(2)所提出的问题]是“下单侧检验的结果”。下单侧检验结果的判定：若 $z < z_a$ ，则接受 $(H_1: \mu < \mu_0)$ ；否则，就接受 $(H_0: \mu \geq \mu_0)$ ；PL为下单侧概率。因检验统计量 $z = 0.046904 > -1.64485$ ，故只能接受“ $H_0: \mu \geq \mu_0$ ”，即认为样本所抽自的总体均值 μ 大于等于给定的均值 $\mu_0 = 5.10 (P = 0.51871 > 0.05)$ 。

双侧检验结果

Obs	absz	z1_ha	PT
1	0.046904	1.95996	0.96259

输出结果的第3部分[回答上述式(3)所提出的问题]是“双侧检验的结果”。双侧检验结果的判定：若 $|z| > z_{1-\alpha/2}$ ，则接受 $(H_1: \mu \neq \mu_0)$ ；否则，就接受 $(H_0: \mu = \mu_0)$ ；PT为双侧概率。因检验统计量 $|z| = 0.046904 < 1.95996$ ，故只能接受“ $H_0: \mu = \mu_0$ ”，即认为样本所抽自的总体均值 μ 等于给定的均值 $\mu_0 = 5.10 (P = 0.96259 > 0.05)$ 。

3 成组设计一元定量资料Z检验及SAS实现

3.1 Z检验统计量

对服从正态分布的成组设计一元定量资料进行均值比较，若已知两总体方差时，其对应的检验统计量见下式：

$$Z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \quad (11)$$

式(11)中的 $Z \sim N(0, 1)$ 。

3.2 拒绝域

与式(4)对应的拒绝域 W 见下式^[2]:

$$W = (\bar{x}_1 - \bar{x}_2; Z > Z_{1-\alpha}) \tag{12}$$

与式(5)对应的拒绝域 W 见下式:

$$W = (\bar{x}_1 - \bar{x}_2; Z < Z_\alpha) \tag{13}$$

与式(6)对应的拒绝域 W 见下式:

$$W = (\bar{x}_1 - \bar{x}_2; |Z| > Z_{1-\alpha/2}) \tag{14}$$

3.3 例 2 的 SAS 实现

SAS 程序如下:

```
%let n1=98; /*第 1 组样本含量*/
%let n2=23; /*第 2 组样本含量*/
%let sm1=4.41; /*第 1 组样本均值*/
%let sm2=8.18; /*第 2 组样本均值*/
%let pv1=23.3; /*第 1 组已知的总体方差*/
%let pv2=6.4; /*第 2 组已知的总体方差*/
%let a=0.01; /*显著性水平*/
data abc;
z=(&sm1-&sm2)/sqrt(&pv1/&n1+&pv2/&n2);
absz=abs(z);
za=probit(&a);
z1_a=probit(1-&a);
z1_ha=probit(1-&a/2);
PU=1-probnorm(Z);
PL=probnorm(Z);
PT=2*(1-probnorm(absz));
title1 '上单侧检验结果';
proc print data=abc;
var z z1_a PU;
footnote1 '上单侧检验结果的判定:';
footnote2 '若 z>z1_a,则接受(H1:μ1>μ2)';
footnote3 '否则,就接受(H0:μ1≤μ2)';
footnote4 'PU为上单侧概率。';
run;
title1 '下单侧检验结果';
proc print data=abc;
var z za PL;
footnote1 '下单侧检验结果的判定:';
footnote2 '若 z<za,则接受(H1:μ1<μ2)';
footnote3 '否则,就接受(H0:μ1≥μ2)';
footnote4 'PL为下单侧概率。';
run;
```

title1 '双侧检验结果';

```
proc print data=abc;
var absz z1_ha PT;
footnote1 '双侧检验结果的判定:';
footnote2 '若|z|>z1_ha,则接受(H1:μ1≠μ2)';
footnote3 '否则,就接受(H0:μ1=μ2)';
footnote4 'PT为双侧概率。';
run;
```

【程序说明】显著性水平 $\alpha=0.01$,其他同“例 1 中的 SAS 程序说明”,从略。

【SAS 主要输出结果及解释】

上单侧检验结果

Obs	z	z1_a	PU
1	-5.24819	2.32635	1.00000

输出结果的第 1 部分[回答上述式(4)所提出的问题]是“上单侧检验的结果”。上单侧检验结果的判定:若 $z > z_{1-a}$,则接受 $(H_1: \mu_1 > \mu_2)$; 否则,就接受 $(H_0: \mu_1 \leq \mu_2)$; PU 为上单侧概率。因检验统计量 $z = -5.24819 < 2.32635$,故只能接受“ $H_0: \mu_1 \leq \mu_2$ ”,即认为受试者中“行为类型”为“B 型和中间型”的 PHQ-9 评分的总体均值小于或等于“A 型”的 PHQ-9 评分的总体均值($P=1 > 0.05$)。

下单侧检验结果

Obs	z	za	PL
1	-5.24819	-2.32635	7.6799E-8

输出结果的第 2 部分[回答上述式(5)所提出的问题]是“下单侧检验的结果”。下单侧检验结果的判定:若 $z < z_a$,则接受 $(H_1: \mu_1 < \mu_2)$; 否则,就接受 $(H_0: \mu_1 \geq \mu_2)$; PL 为下单侧概率。因检验统计量 $z = -5.24819 < -2.32635$,故拒绝“ $H_0: \mu_1 \geq \mu_2$ ”,接受“ $H_1: \mu_1 < \mu_2$ ”,即认为受试者中“行为类型”为“B 型和中间型”的 PHQ-9 评分的总体均值小于“A 型”的 PHQ-9 评分的总体均值($P=7.6799E-8 < 0.01$)。

双侧检验结果

Obs	absz	z1_ha	PT
1	5.24819	2.57583	0.000000154

输出结果的第 3 部分[回答上述式(6)所提出的问题]是“双侧检验的结果”。双侧检验结果的判定:若 $|z| > z_{1-a}$,则接受 $(H_1: \mu_1 \neq \mu_2)$; 否则,就接受 $(H_0: \mu_1 = \mu_2)$; PT 为双侧概率。因检验统计量 $|z| = 5.24819 > 2.57583$,故拒绝“ $H_0: \mu_1 = \mu_2$ ”,接受“ $H_1: \mu_1 \neq \mu_2$ ”,即认为受试者中“行为类型”为“B 型和中间型”的 PHQ-9 评分的总体均值不等于“A 型”的

PHQ-9评分的总体均值。再结合实际数据,可得出结论:前者均值小于后者均值($P=0.000000154<0.01$)。

4 讨论与小结

4.1 讨论

在解决实际问题时,一般对同一个问题只需要实施上述三个检验(指“上单侧”“下单侧”和“双侧”检验)中的某一个,而不是三个都同时要实施。一般来说,仅在已知总体服从正态分布且总体方差已知的条件下,均值比较才适合选用“Z检验”,否则,应根据具体情况,选择“t检验”或“秩和检验”^[3]。

通常情况下,进行一般差异性检验时都采取“双侧检验”,因为双侧检验比单侧检验保守一些。而仅在在有“专业知识为依据的场合下”,才可以采取“单侧检验”。例如,若在临床上可以认为“某两种药合用的效果”肯定比“单用其中一种药的效果”好,进行一般差异性检验时,可以考虑选用“单侧检验”(若拟进行非劣效性检验或优效性检验,其本身就要求进行单侧检验)。在选择采取“上单侧检验”还是“下单侧检验”时,需结合具体问题以及将何者视为第1组,在此基础上,再结合“备择假设”来确定“具体方向(指‘上’还是‘下’单侧检验)”。实际问题提示的方向应与“备择假设的方向”一致。例如,在本文例2中,若有理由进行单侧检验,则宜选用“下单侧检验”。

4.2 小结

本文结合两个实例,介绍了单组设计和成组设计条件下一元定量资料一般差异性Z检验及SAS实现。对常规的“双侧检验”及两种“单侧检验”都进行了介绍,以期为实际工作者合理选择假设检验方法开阔了思路、提供更多的备选方案。虽然本文所介绍的方法计算量并不大,但不便于直接采用国际通用的统计软件(SAS、SPSS、R等)^[4-7]来实现。

参考文献

- [1] 孙振晓,于相芬.新冠肺炎疫情期间封闭管理精神科医护人员焦虑抑郁症状及相关因素调查[J].四川精神卫生,2020,33(2):102-106.
- [2] 茆诗松.统计手册[M].北京:科学出版社,2006:78-85.
- [3] 高飞,刘媛媛,李长平,等.如何正确运用t检验——t检验的基本概念与前提条件[J].四川精神卫生,2020,33(3):211-216.
- [4] 胡良平.SAS常用统计分析教程[M].2版.北京:电子工业出版社,2015:264-418.
- [5] 杜智敏,樊文强.SPSS在社会调查中的应用[M].北京:电子工业出版社,2015:171-246.
- [6] 吴喜之.复杂数据统计方法——基于R的应用[M].3版.北京:中国人民大学出版社,2015:18-146.
- [7] 胡良平.现代医学统计学[M].北京:科学出版社,2020:235-362.

(收稿日期:2020-09-16)

(本文编辑:戴浩然)