

# 如何正确运用 $t$ 检验——两算术均值比较 一般差异性 $t$ 检验及 SAS 实现

张洪璐<sup>1</sup>, 刘媛媛<sup>1</sup>, 李长平<sup>1,2</sup>, 胡良平<sup>2,3\*</sup>

(1. 天津医科大学公共卫生学院, 天津 300070;

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;

3. 军事科学院研究生院, 北京 100850

\*通信作者: 胡良平, E-mail: lphu927@163.com)

**【摘要】** 本文目的是介绍算术均数以及单组设计、配对设计和成组设计一元定量资料  $t$  检验的基本原理及 SAS 实现。首先对算术均数的概念、适用条件和计算公式进行介绍; 其次, 介绍不同设计类型  $t$  检验的基本原理, 结合 SAS 实例进行分析, 并对相关结果进行解释; 最后, 讨论不同设计类型  $t$  检验应用的注意事项。

**【关键词】** 设计类型; 定量资料;  $t$  检验; 算术均数; 差异性分析

中图分类号: R195.1

文献标识码: A

doi: 10.11886/scjsws20200526004

## How to use $t$ test correctly——comparison of two arithmetic mean values for general difference $t$ test and SAS implementation

Zhang Honglu<sup>1</sup>, Liu Yuanyuan<sup>1</sup>, Li Changping<sup>1,2</sup>, Hu Liangping<sup>2,3\*</sup>

(1. Department of Epidemiology and Health Statistics, School of Public Health, Tianjin Medical University, Tianjin

300070, China;

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

100029, China;

3. Graduate School, Academy of Military Sciences PLA China, Beijing 100850, China

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

**【Abstract】** The purpose of this article was to introduce the arithmetic mean, as well as the basic principle and SAS implementation of single-group design, paired design and group design unary quantitative data  $t$  test. Firstly, the concept of arithmetic mean, applicable conditions and calculation formulas were introduced. Secondly, the basic principles of  $t$  tests of different design types were introduced, combined with SAS examples for analysis, and related results were explained. Finally, discuss application considerations for different design types of  $t$  test.

**【Keywords】** Design type; Quantitative data;  $t$  test; Arithmetic mean; Difference analysis

$t$  分布的发现使得小样本统计推断成为可能, 以  $t$  分布为基础的检验称为  $t$  检验<sup>[1]</sup>。在医学统计学中,  $t$  检验具有所需样本量小、计算简便及检验效能较高等特点, 成为应用较多的一类假设检验方法<sup>[2]</sup>。若统计分析的目的是比较算术均数, 根据试验设计类型,  $t$  检验可以分为三种: 单组设计一元定量资料  $t$  检验、配对设计一元定量资料  $t$  检验和成组设计一元定量资料  $t$  检验<sup>[3]</sup>。基于算术均数的  $t$  检验的应用条件一般包括: 随机样本(即独立性)、来自正态分布总体(即正态性)以及两总体方差相等(即方差齐性)<sup>[4]</sup>。本文将分别介绍上述三种类型的一般差异性  $t$  检验, 并结合实例, 采用 SAS 9.4 予以实现。

基金项目: 国家自然科学基金项目(项目名称: 贝叶斯生存分析方法在肝细胞癌肝移植患者预后预测中的应用研究, 项目编号: 81803333)

## 1 基本概念

### 1.1 算术均数及其计算公式

算术均数是大家最为熟悉的平均数指标<sup>[5]</sup>, 是指一个变量所有观测值的总和除以观测值的个数, 反映一个变量所有观测值的平均水平, 简称为均数或均值。样本均数用  $\bar{x}$  表示, 计算公式见式(1):

$$\bar{x} = \frac{x_1 + x_2 + \cdots + x_n}{n} \quad (1)$$

或写为  $\bar{x} = \frac{1}{n} \sum x_i$ , 其中  $\Sigma$  是求和符号,  $x_i$  表示第  $i$  个观测值,  $n$  表示观测值个数。通过计算可以发现, 均数并不一定等于这一组数据中的某个观测值。

如果没有收集到个体观测值数据, 而只是获得了汇总的频数分布表数据, 则可用每个组段的组中

值作为该组段中各个观测值的估计值,并用式(2)计算均数的近似值:

$$\bar{x} = \frac{\sum f_i x_{Mi}}{\sum f_i} = \frac{\sum f_i x_{Mi}}{n} \tag{2}$$

式(2)中, $f_i$ 为第*i*组段的频数, $x_{Mi}$ 为对应组段的组中值, $x_{Mi}=($ 第*i*组段上限+第*i*组段下限) $/2$ 。

均数计算简便,易于理解,是应用最为广泛的平均数指标。但是均数对于特大或特小的观测值十分敏感。对于偏态分布数据,均数也会偏向拖尾一侧,不能很好地反映全部观测值的平均水平。因此,均数主要适用于描述不含极端值的对称分布变量的平均水平,这时均数位于分布的中心位置。

## 2 t 检验的基本原理

### 2.1 单组设计一元定量资料 t 检验的基本原理

单组设计一元定量资料 t 检验,实际上是推断该样本来自总体均数  $\mu$  与某一已知总体均数  $\mu_0$  (常为理论值或标准值)有无差别。零假设  $H_0:\mu = \mu_0$ ,备择假设  $H_1:\mu \neq \mu_0$ ;检验水准: $\alpha=0.05$ 。

单组设计一元定量资料 t 检验的检验统计量为:

$$t = \frac{\bar{x} - \mu_0}{S/\sqrt{n}} \tag{3}$$

式(3)中,分子是样本均数与零假设中  $\mu_0$  的差距,分母是样本均数的标准误,t 统计量实为用标准误来度量的样本均数与  $\mu_0$  的差距。这个差距小,将倾向于接受零假设;这个差距大,将倾向于拒绝零假设。数学家已证明:式(3)定义的检验统计量 t 是服从自由度为  $n-1$  的 t 分布的随机变量。故可以依据实际计算出来的 t 检验统计量的数值以及 t 分布,做出接受或拒绝零假设的决定<sup>[6]</sup>。

### 2.2 配对设计一元定量资料 t 检验

配对设计一元定量资料的分析着眼于每一对中两个观察值之差 d,这些差值构成一组资料,用 t 检验推断差值的总体均数是否为“0”。检验假设为:

$H_0:\mu_d = 0$ ,即差值的总体均数为0; $H_1:\mu_d \neq 0$ ,即差值的总体均数不为0;检验水准: $\alpha=0.05$ 。

当  $H_0$  成立时,检验统计量为:

$$t = \frac{\bar{d} - 0}{S_d/\sqrt{n}} \sim t_{(v)}, v = n - 1 \tag{4}$$

式(4)中, $\bar{d}$ 为差值的均数, $S_d$ 为差值的样本标

准差, $n$ 为对子数。同样,依据实际计算出来的 t 检验统计量的数值以及 t 分布(自由度为  $n-1$ ),给定一个小概率  $\alpha$  作为检验水准,如果与 t 值相应的 P 值小于给定的  $\alpha$ ,拒绝  $H_0$ ;否则,不拒绝  $H_0$ 。

### 2.3 成组设计一元定量资料 t 检验

将受试对象随机分为两组,每一组接受一种处理(注意:对照组也可被视为一种处理)。一般把这样获得的两组定量资料视为代表两个总体的两个独立样本,据此推断它们的总体均数是否相等。此类检验也基于 t 分布,必需假定定量资料应满足独立性、正态性和方差齐性。

#### 2.3.1 两样本所属总体方差相等,即具有方差齐性

将两个正态分布总体分别记为  $N(\mu_1, \sigma_1^2)$  和  $N(\mu_2, \sigma_2^2)$ ,假定  $\sigma_1^2 = \sigma_2^2$ ,检验假设  $H_0:\mu_1 = \mu_2$ ,即两样本所属的两个总体均数相等; $H_1:\mu_1 \neq \mu_2$ ,即两样本所属的两个总体均数不相等;检验水准  $\alpha=0.05$ 。

检验统计量为:

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{S_c^2 \left( \frac{1}{n_1} + \frac{1}{n_2} \right)}} \tag{5}$$

其中, $S_c^2$ 是利用两样本联合估计的方差:

$$S_c^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2} = \frac{\sum (X_1 - \bar{X}_1)^2 + \sum (X_2 - \bar{X}_2)^2}{n_1 + n_2 - 2}$$

式(5)中的分子是两个样本均数之差,分母是样本均数之差的标准误,检验统计量 t 实为用标准误差度量的均数之差。

可以证明,当  $H_0$  成立时,这个检验统计量服从自由度为  $v = n_1 + n_2 - 2$  的 t 分布。

根据式中算得检验统计量的数值后,利用这个 t 分布可得相应的 P 值。同样,依据实际计算出来的 t 检验统计量的数值以及 t 分布(自由度为  $n_1 + n_2 - 2$ ),给定一个小概率  $\alpha$  作为检验水准,如果与 t 值相应的 P 值小于给定的  $\alpha$ ,拒绝  $H_0$ ;否则,不拒绝  $H_0$ 。

#### 2.3.2 两样本所属总体方差不等

两正态总体分别记为  $N(\mu_1, \sigma_1^2)$  和  $N(\mu_2, \sigma_2^2)$ ,假定  $\sigma_1^2 \neq \sigma_2^2$ ,检验假设同前,即  $H_0:\mu_1 = \mu_2$ ,即两样本所属的两个总体均数相等; $H_1:\mu_1 \neq \mu_2$ ,即两样本所属的两个总体均数不相等。

此时,可采用  $t'$  作为检验统计量:

$$t' = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \quad (6)$$

式(6)中,分子仍是两样本均数之差,分母是均数之差的标准误, $t'$ 仍是用标准误差度量的均数之差。

$H_0$ 成立时, $t'$ 的分布比较复杂,可用如下自由度的  $t$  分布来近似  $t'$  的分布。

$$v = \frac{\left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right)^2}{\frac{\left(\frac{S_1^2}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{S_2^2}{n_2}\right)^2}{n_2 - 1}} \quad (7)$$

利用该式算得统计量  $t'$  的数值后,据此近似的  $t$  分布可以得到相应的  $P$  值。同样,给定一个小概率  $\alpha$  作为检验水准,如果与  $t$  值相应的  $P$  值小于给定的  $\alpha$ ,拒绝  $H_0$ ;否则,不拒绝  $H_0$ 。

### 3 基于 SAS 的实例分析

#### 3.1 单组设计一元定量资料 $t$ 检验

【例1】大量研究显示,正常人的血清钙标准值是 2.50 mmol/L,现有 20 名骨质疏松患者的血清钙测量值如下(单位:mmol/L):2.39、2.41、2.47、2.28、2.19、2.50、2.37、2.49、2.18、2.09、2.29、2.34、2.55、2.27、2.31、2.39、2.29、2.20、2.41、2.42。试比较骨质疏松患者血清钙与正常人是否不同。

SAS 程序如下:

```
data calcium;
input x @@;
cards;
2.39 2.41 2.47 2.28 2.19 2.50 2.37 2.49 2.18
2.09 2.29 2.34 2.55 2.27 2.31 2.39 2.29 2.20
2.41 2.42
;
run;
proc univariate normal data=calcium;
var x;
run;
proc ttest data=calcium h0=2.50;
var x;
run;
```

【程序说明】创建名为 calcium 的数据集,首先调用 UNIVARIATE 过程进行正态性检验,VAR 语句后面的变量为待检验的血清钙测量值,然后调用

TTEST 过程,设置评价指标的理论均值  $h_0$  为 2.50。

【主要输出结果及解释】正态性检验的结果  $W=0.9785, P=0.9136$ ,可以认为评价指标的测定值符合正态性的要求。 $t$  检验结果如下: $t=-5.87, P<0.0001$ ,骨质疏松患者血清钙均值与正常人不同,骨质疏松患者血清钙浓度均值 2.34 mmol/L 低于正常人血清钙浓度均值 2.50 mmol/L。

#### 3.2 配对设计一元定量资料 $t$ 检验

【例2】采集 10 只正常成年大耳白兔下肢骨标本,共 10 对胫骨,行三点弯力学测试。从实验兔胫骨三点弯载荷挠度曲线得最大挠度,见表 1。试比较双侧胫骨最大挠度是否不同。表 1 资料对应的实验设计类型为自身配对设计。

表 1 10 只家兔双侧胫骨最大挠度

序 号	最大挠度(mm)	
	左侧	右侧
1	2.91	2.73
2	2.56	2.41
3	2.75	2.65
4	2.66	2.90
5	2.65	2.90
6	2.58	2.49
7	2.60	2.88
8	3.04	2.79
9	2.46	2.39
10	3.19	3.24

SAS 程序如下:

```
data bending;
input x1 x2 @@;
d=x2-x1;
cards;
2.91 2.73 2.56 2.41 2.75 2.65 2.66 2.90 2.65
2.90 2.58 2.49 2.60 2.88 3.04 2.79 2.46 2.39
3.19 3.24
;
run;
proc univariate normal data=bending;
var d;
run;
proc ttest data=bending;
paired x1*x2;
run;
```

【程序说明】创建名为 bending 的数据集。首先生成一个新的变量 d,代表每对中同一评价指标取

值的差量;调用 UNIVARIATE 过程对其进行正态性检验;然后调用 TTEST 过程。与单组设计一元定量资料  $t$  检验的区别是, PAIRED 语句后面为变量的对子,对子间用\*连接。

【主要输出结果及解释】差值  $d$  正态性检验的结果为  $W=0.8808, P=0.1334$ , 可以认为符合正态性的要求,  $t$  检验结果如下:  $t=0.03, df=9, P=0.9748$ , 双侧胫骨最大挠度均值之间差异无统计学意义。

### 3.3 成组设计一元定量资料 $t$ 检验

【例 3】在一项氯氮平血浓度对白细胞总数的影响的研究中<sup>[7]</sup>, 研究者选取受试组 28 例, 给予氯氮平治疗, 对照组 28 例, 比较受试组治疗一周末与三周末的血药浓度是否存在差异。假设测定的患者血药浓度(单位:  $\mu\text{g/L}$ )如下。治疗一周末血药浓度: 49、417、229、342、445、700、273、50、286、212、516、444、334、176、170、434、791、219、384、213、96、327、217、16、256、438、102、495; 治疗三周末血药浓度: 1189、517、534、293、708、588、374、691、694、1025、278、611、854、597、721、479、530、755、610、583、672、71、194、457、132、570、721、421。试分析两组受试者血药浓度差异是否有统计学意义。

SAS 程序如下:

```
data blood;
input g$ ;
input x @@;
output;
end;
cards;
FIRST
49 417 229 342 445 700 273 50 286 212 516 444
334 176 170 434 791 219 384 213 96 327 217 16 256
438 102 495
THIRD
1189 517 534 293 708 588 374 691 694 1025 278
611 854 597 721 479 530 755 610 583 672 71 194
457 132 570 721 421
;
run;
proc univariate normal data=blood;
var x;
class g;
run;
proc ttest data=blood;
```

```
var x;
class g;
run;
```

【程序说明】创建名为 blood 的数据集, 设置分组变量  $g$ , 利用循环语句 DO-END 分别输入两组变量值。调用 UNIVARIATE 过程分别对两组定量数据进行正态性检验, CLASS 语句后加分类变量。然后, 调用 TTEST 过程, 与前两种  $t$  检验的程序相比, 增加了 CLASS 语句。

【主要输出结果及解释】两组定量资料均符合正态性的要求(一周末:  $W=0.9558, P=0.2757$ ; 三周末:  $W=0.9687, P=0.5463$ )。方差齐性检验的结果:  $F=1.75, P=0.1542$ , 符合方差齐性的要求。 $t$  检验的结果见表 2。 $t$  检验结果:  $t=-4.42, P<0.0001$ , 两组血药浓度均值不同, 治疗一周末血药浓度均值  $308.3 \mu\text{g/L}$  低于三周末血药浓度均值  $566.8 \mu\text{g/L}$ 。

表 2 成组设计一元定量资料  $t$  检验分析结果

方法	方差	自由度	$t$ value	$Pr> t $
汇总	等于	54	-4.42	<0.0001
Satterthwaite	不等于	50.292	-4.42	<0.0001

## 4 讨论与小结

### 4.1 讨论

数据的集中位置表示大多数观测值所在的中心位置。算术均数是一个良好的描述集中位置的指标, 具有计算简单、简明易懂等优点。但每个数据或大或小的变化都会对均数造成影响。因此, 均数更适用于描述呈正态分布数据的集中位置。在实际应用中, 均数常和标准差相结合用于描述数据的集中趋势和离散程度。

$t$  检验是在统计学中较常应用于定量资料差异性分析的一类假设检验方法。值得注意的是, 对于定量资料而言, 在数据分析时, 应先考察该数据是否符合  $t$  检验的应用条件: ①结合调查研究方式及样本来源等判断是否符合随机样本的要求; ②调用 UNIVARIATE 过程进行正态性检验, 若数据为非正态分布, 则考虑采用变量变换的方式使其满足正态或采用非参数假设检验的方法; ③在对成组设计一元定量资料进行分析时, 对于正态分布的数据, 要求其两总体方差相等, 即满足方差齐性, 若经检验方差不齐, 则可使用  $t'$  检验。

此外, 应根据专业知识确定选择单侧检验还是双侧检验。对于单因素  $k$  水平设计的定量资料或多

因素设计的定量资料,应严格检查是否满足参数检验的前提条件,进而选择合适的检验方法<sup>[2]</sup>。

#### 4.2 小结

本文基于 3 个实例,采用 SAS 软件实现了对单组设计、配对设计和成组设计下一元定量资料的  $t$  检验。在此过程中,借助直观判断法,考察了“独立性”的前提条件;借助 UNIVARIATE 过程中的选项“NORMAL”实现了正态性检验;借助 TTEST 过程中自身带有的方差齐性检验功能,实现了方差齐性检验。此外,还交代了基于均值比较的  $t$  检验的基本原理和应用时的注意事项。

#### 参考文献

[1] 董秀玥. 配对  $t$  检验与成组  $t$  检验优选方法研究[J]. 数理医药

学杂志, 2010, 23(1): 11-14.

[2] 胡良平, 高辉. 如何正确运用  $t$  检验[J]. 中西医结合学报, 2008, 6(2): 209-212.

[3] 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: 264-272.

[4] 高辉, 胡良平, 金松华, 等. 常见试验设计定量资料统计分析错误辨析[J]. 中西医结合学报, 2008, 6(9): 979-982.

[5] 李晓松. 卫生统计学[M]. 8 版. 北京: 人民卫生出版社, 2017: 12-15.

[6] 方积乾. 卫生统计学[M]. 7 版. 北京: 人民卫生出版社, 2012: 105-112.

[7] 喻东山, 刘阳, 高小宁. 氯氮平血浓度对白细胞总数的影响[J]. 四川精神卫生, 2000, 13(1): 18-20.

(收稿日期: 2020-05-26)

(本文编辑: 吴俊林)