显著性检验在择优方案中的运用

李军
辽宁工程技术大学理学院，辽宁阜新 (123000)
E-mail：lj0801@126.com

摘　要：运用显著性检验的方法来研究对两种或更多个体好坏评估的问题，先选用传统的假设检验来检验任意两个改革方案是否均有效，再通过构造新的统计变量并借助相应的分布形式进行。显著性检验，实践证明这种检验方法对改革方案的选取具有指导意义。

关键词：显著性检验；t 分布；距差
中图分类号：O0213

1.　引 言

发展成为当今社会的主流问题，人类社会的进步离不开发展。工业生产，教育教学，运动员训练，新药剂的研发，都要借助一定的改革来实现发展进步的目的。我们面对的很多的改革方案仅凭直觉来估测哪种更好还不够。

数理统计[1]中的假设检验问题是一种重要的统计推断方法，被经常用于对科学实验方案好坏检验。例如，验收一批产品重要的不是估测这批产品的不和格率，而是将其与规定的不合格率界限进行比较，不超过界限则接受，超过界限就拒绝。再如，通过两种不同的方案试销某种产品，需比较它们的经济效益，但这时我们讨论的问题的初始样本来自于同一总体。在现实问题中经常出现另一类问题，两种方案可能来自完全不同的总体，应用于不同总体时的假设检验就不能照搬同一总体的假设检验方法了。

为解决这一不足，本文提出了一套完整的比较方案。下面我们来引例说明。

2.　一个教学效果检验的实例

设某高校现有 A、B 两个试验班,上学期两班在相同的教学条件下教学,下学期对 A 班实行第一种教学改革，对 B 班实行第二种下学期教学改革，问这两种教学改革哪个更好。在这里，我们不能简单的利用 A、B 两个班下学期的成绩差做检验，因为在上学期 A、B 两个班的成绩可能就存在一定的差异。所以，最合适的检验方法应该是两个班级上学期差距与下学期差距间的比较。下面给出具体的检验方法和步骤。

2.1 样本的选取与相应特征值的计算

将上学期期末考试后两班的成绩作为其各自的总体，由于两班成员数目可能不同，从 A 班成绩总体 X_1 中抽取 n_1 个样本，从 B 班成绩总体 Y_1 中抽取 n_2 个样本，分别记两组样本为：

$$\text{A:}\quad x_{11}, x_{12}, \cdots, x_{1n_1}$$

$$\text{B:}\quad y_{11}, y_{12}, \cdots, y_{1n_2}$$

将下学期期末考试后两班所有成员的成绩作为改革后的总体，由于两班成员数目可能不同，从 A 班成绩总体 X_2 中抽取 n_1 个样本，从 B 班成绩总体 Y_2 中抽取 n_2 个样本，分别记为

$$\text{A:}\ x_{21}, x_{22}, \cdots, x_{2n_2}$$

$$\text{B:}\ y_{21}, y_{22}, \cdots, y_{2n_2}$$

设 $X_1 \sim N(\mu_{1x}, \sigma_{1x}^2)$，$Y_1 \sim N(\mu_{1y}, \sigma_{1y}^2)$，$X_2 \sim N(\mu_{2x}, \sigma_{2x}^2)$，$Y_2 \sim N(\mu_{2y}, \sigma_{2y}^2)$，上、

下学期在两班抽取样本的样本均值和方差分别为$\bar{x}_1,\bar{y}_1,\bar{x}_2,\bar{y}_2$和$S_{1x}^2,S_{1y}^2,S_{2x}^2,S_{2y}^2$，其中

$$\bar{x}_1=\sum_{i=1}^{n_1}x_{1i}\text{，}\bar{y}_1=\sum_{i=1}^{n_2}y_{1i}\text{，}\bar{x}_2=\sum_{i=1}^{n_1}x_{2i}\text{，}\bar{y}_2=\sum_{i=1}^{n_2}y_{2i} \tag{1}$$

$$S_{1x}^2=\frac{1}{n_1-1}\sum_{i=1}^{n_1}(x_{1i}-\bar{x}_1)^2\text{，}S_{1y}^2=\frac{1}{n_2-1}\sum_{i=1}^{n_2}(y_{1i}-\bar{y}_1)^2$$

$$S_{2x}^2=\frac{1}{n_1-1}\sum_{i=1}^{n_1}(x_{2i}-\bar{x}_2)^2\text{，}S_{2y}^2=\frac{1}{n_2-1}\sum_{i=1}^{n_2}(y_{2i}-\bar{y}_2)^2 \tag{2}$$

两班上学期的成绩水平就可能有差距，通过不同的教学改革也会产生差距成绩差距，这时不能用传统的检验方法，我们要比较通过改革产生的差距。

3. 统计量构造及有效性的显著性检验

3.1 有效性的检验

若知道两种改革方案是否有效，我们只需分别找到$\mu_{2x}-\mu_{1x}$与$\mu_{2y}-\mu_{1y}$有关的统计量及其分布即可。显然$\bar{x}_2-\bar{x}_1$，$\bar{y}_2-\bar{y}_1$ 是它们的无偏估计量[2]，再利用显著性检验的方法在一定的置信水平条件下，进行计算和比较分析，即可达到有效性检验目的，具体程本文不做过多解释。这里不妨假定满足方差齐性，即$\sigma_{1x}^2=\sigma_{2x}^2,\sigma_{1y}^2=\sigma_{2y}^2$.

3.2 统计量的构造

如果这两种方案均有效，我们来比较哪种改革方案更好，这两班改革前后成绩直观图有四种：

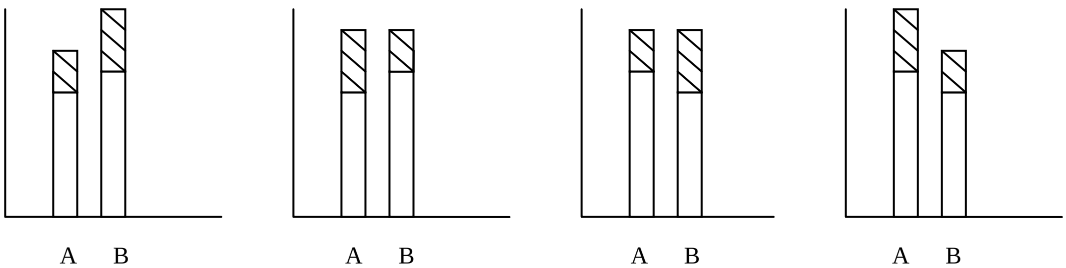

图 1 成绩水平直观图

Fig.1 result sdandard audio-visual figue

其中白条框代表原始成绩水平，斜条框代表改革以后成绩提高的程度，其实就是$\mu_{2x}-\mu_{1x}$，$\mu_{2y}-\mu_{1y}$的直观反映。$\Delta\mu_x=\mu_{2x}-\mu_{1x}$，$\Delta\mu_y=\mu_2y-\mu_1y$，$\Delta\mu=\mu_x-\mu_y$，我们称$\Delta\mu$为距差，表示由于改革方法不同造成成绩提高的差距。要比较哪中改革方案更好，只需找到与$\Delta\mu$有关的统计量及其分布即可。

令 $M_x=\bar{x}_2-\bar{x}_1,M=\bar{y}_2-\bar{y}_1$，则$\Delta M=M_x-M_y$，本文称$\Delta M$为样本距差。显然$\Delta M$是$\Delta\mu$的无偏估计量。

$$\Delta M = N[(\mu_{2x} - \mu_{1x}) - (\mu_{2y} - \mu_{1y}), \frac{\sigma_{1x}^2 + \sigma_{2x}^2}{n_1} + \frac{\sigma_{1y}^2 + \sigma_{2y}^2}{n_2}]; \quad (3)$$

令

$$\sigma_{1x}^2 + \sigma_{2x}^2 = \sigma_{1y}^2 + \sigma_{2y}^2 = \sigma^2$$

则可以推出

$$\frac{\Delta M - [(\mu_{2x} - \mu_{1x}) - (\mu_{2y} - \mu_{1y})]}{\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim N(0,1) \quad (4)$$

分别将M_x, M_y各项按顺序作差再求和，即得到M_x, M_y的样本均值$\overline{M}_x, \overline{M}_y$.

$$M_x = \frac{1}{n_1}(\sum_{i=1}^{n_1} x_{2i} - \sum_{i=1}^{n_1} x_{1i}) = \frac{1}{n_1}\sum_{i=1}^{n_1}(x_{2i} - x_{1i}) = \overline{M}_x;$$

$$M_y = \frac{1}{n_2}(\sum_{i=1}^{n_2} y_{2i} - \sum_{i=1}^{n_2} y_{1i}) = \frac{1}{n_2}\sum_{i=1}^{n_2}(x_{2i} - x_{1i}) = \overline{M}_y; \quad (5)$$

由上述两式可得到如下性质 1：两个样本均值的差等于样本差的样本均值。

设S_1^2, S_2^2分别为M_x, M_y的样本方差。

$$S_1^2 = \frac{1}{n_1 - 1}\sum_{i=1}^{n_1}[(x_{2i} - x_{1i}) - \overline{M}_x]^2;$$

$$s_2^2 = \frac{1}{n_2 - 1}\sum_{i=1}^{n_2}[(y_{2i} - y_{1i}) - \overline{M}_y]^2; \quad (6)$$

$$\frac{(n_1 - 1)S_1^2}{\sigma^2} \sim \chi^2(n_1 - 1), \quad \frac{(n_2 - 1)S_2^2}{\sigma^2} \sim \chi^2(n_2 - 1);$$

且它们相互独立，由χ^2分布的可加性可知：

$$V = \frac{(n_1 - 1)S_1^2}{\sigma^2} + \frac{(n_2 - 1)S_2^2}{\sigma^2} \sim \chi^2(n_1 + n_2 - 2);$$

作假设

$$H_0:\ \mu_{2x} - \mu_{1x} = \mu_{2y} - \mu_{1y},\ （H_1:\mu_{2x} - \mu_{1x} \neq \mu_{2y} - \mu_{1y}）$$

当H_0成立时，有 $U = \dfrac{\Delta M}{\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim N(0,1)$，则有

$$T = \frac{U}{\sqrt{V/(n_1 + n_2 - 2)}} = \frac{(M_x - M_y)}{S_\omega\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t_{(n_1+n_2-2)} \quad (7)$$

其中

$$S_\omega^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2},\quad S_\omega = \sqrt{S_2^2} \quad (8)$$

3.3 方案的选取

对于给定的显著性水平[3]，我们可以求出检验的临界值$t_{\frac{\alpha}{2}}(n_1+n_2-2)$，于是

(1). 若$T=\dfrac{\Delta M}{S_\omega\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}>t_{\frac{\alpha}{2}}(n_1+n_2-2)$,即$\Delta M>\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}\cdot S_\omega t_{\frac{\alpha}{2}}(n_1+n_2-2)$ 时，应拒绝H_0，说明第一种改革比第二种更有效；

(2). 若$T=\dfrac{\Delta M}{S_\omega\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}<-t_{\frac{\alpha}{2}}(n_1+n_2-2)$,即$\Delta M>-\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}\cdot S_\omega t_{\frac{\alpha}{2}}(n_1+n_2-2)$ 时 ，应拒绝H_0，说明第二种改革比第一种更有效；

(3) 若时$\left|\dfrac{\Delta M}{S_\omega\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}\right|\le t_{\frac{\alpha}{2}}(n_1+n_2-2)$，即$|\Delta M|\le S_\omega\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}\cdot t_{\frac{\alpha}{2}}(n_1+n_2-2)$，应接受原假设$H_0$，即说明两种改革没有显著的差别。

4. 实例分析

现有 A, B 两块水稻的实验田，我们要利用两块试验田来检验两种不同的杀虫剂的效果。但是，由于两块试验田原有的土质条件本来就不同，亩产量也不相同，所以，不能仅仅以喷洒药剂后的产量来区分杀虫剂的好坏，而只能采用本文提出的两阶段对比假设检验方法。

第一年不对两试验田使用任何杀虫剂，分别从两块实验田取出 5 块等面积区域采集水稻进行称量，分别记为A_1，B_1.

$$A_1: x_{11}, x_{12}, x_{13}, x_{14}, x_{15}$$

$$B_1: y_{11}, y_{12}, y_{13}, y_{14}, y_{15}$$

第二年对 A 田使用甲杀虫剂，对 B 田使用乙杀虫剂，再分别从两块实验田的 5 块等面积区域采集水稻进行称量，记为A_2，B_2.

$$A_2: x_{21}, x_{22}, x_{23}, x_{24}, x_{25}$$

$$B_2: y_{21}, y_{22}, y_{23}, y_{24}, y_{25}$$

我们在取显著性水平$\alpha=0.05$的条件下分析哪种杀虫剂更有效。表 1 是具体观测数据。

由式(1)得: $\bar{x}_1=4.16, \bar{y}_1=3.50, \bar{x}_2=5.12, \bar{y}_2=4.26$

$$S_{1x}^2=\frac{1}{n_1-1}\sum_{i=1}^{n_1}(x_{1i}-\bar{x}_1)^2=\frac{1}{4}[(4-4.16)^2+(4.2-4.16)^2+(4.3-4.16)^2$$
$$+(4.4-4.16)^2+(3.9-4.16)^2]=0.172$$

$$S_{2x}^2=\frac{1}{n_1-1}\sum_{i=1}^{n_1}(x_{2i}-\bar{x}_2)^2=\frac{1}{4}[(5-5.12)^2+(5.4-5.12)^2+(4.9-5.12)^2$$
$$+(5.1-5.12)^2+(5.0-5.12)^2]=0.148$$

$$S_{\omega_x}^2=\frac{(n_1-1)S_{1x}^2+(n_1-1)S_{2x}^2}{2n_1-2}=\frac{S_{1x}^2+S_{2x}^2}{2}=\frac{0.172+0.148}{2}=0.16$$

$$S_{\omega_x}=\sqrt{S_{\omega_x}^2}=\sqrt{0.16}=0.4$$

$$T_x=\frac{(\bar{x}_2-\bar{x}_1)}{S_{\omega_x}\sqrt{2/n_1}}=\frac{5.12-4.16}{0.4\times0.63}=3.809$$

$$S_{1y}^2=\frac{1}{n_2-1}\sum_{i=1}^{n_2}(y_{1i}-\overline{X_1})^2=\frac{1}{4}[(3.1-3.5)^2+(3.4-3.5)^2+(3.9-3.5)^2$$

$$+(3.8-3.5)^2+(3.3-3.5)^2]=0.115,$$

$$S_{2x}^2=\frac{1}{n_2-1}\sum_{i=1}^{n_2}(y_{2i}-\overline{y_2})^2=\frac{1}{4}[(4.6-4.26)^2+(4.2-4.26)^2$$

$$+(4.1-4.26)^2+(4.4-4.26)^2+(4.0-4.26)^2]=0.058$$

$$S_{\omega_y}^2=\frac{(n_2-1)S_{1y}^2+(n_2-1)S_{2y}^2}{2n_2-2}=\frac{S_{1y}^2+S_{2y}^2}{2}=\frac{0.115+0.058}{2}=0.87$$

$$S_{\omega_y}=\sqrt{S_{\omega_y}^2}=\sqrt{0.87}=0.29$$

$$T_y=\frac{(\bar{y}_2-\bar{y}_1)}{S_{\omega_y}\sqrt{2/n_2}}=\frac{4.26-3.50}{0.29\times0.63}=4.16$$

查表可知：$t_{0.025}8=2.306,-t_{0.025}8=-2.306$ $\because$ 3.809>2.306,4.16>2.306．所以两种杀虫剂都能起到增产作用。

接下来我们计算与距差有关的特征值。

由（5）得:$\bar{M}_x=\bar{x}_2-\bar{x}_1=5.12-4.16=0.96$

$$\bar{M}_y=\bar{y}_2-\bar{y}_1=4.26-3.50=0.76$$

$$\Delta M=M_x-M_y=0.96-0.76=0.2$$

$$S_1^2=\frac{1}{4}[2\times(0.96-1.2)^2+(0.96-0.6)^2$$

$$+(0.96-0.7)^2+(0.96-1.1)^2]=0.083$$

$$S_2^2=\frac{1}{4}[(0.76-1.5)^2+(0.76-1.8)^2+(0.76-0.2)^2$$

$$+(0.76-0.6)^2+(0.76-0.7)^2]=0.223$$

$$S_\omega^2=\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}=\frac{4S_1^2+4S_2^2}{8}=\frac{0.083+0.223}{2}$$

$$=0.153, S_\omega=\sqrt{S_\omega^2}=\sqrt{0.153}=0.391$$

$$T=\frac{(\bar{M}_x-\bar{M}_y)}{S_\omega\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}=\frac{0.96-0.76}{0.391\times0.63}=0.812,$$

查表可知：$t_{0.025}8=2.306,-t_{0.025}8=-2.306,\because 2.306>0.812$

$\therefore$ 应接受H_0，说明这两种沙虫剂效果相同。

表 1　等面积区域水稻产量(单位:千克)

Tab.1 the output of rice equal area

序号	实验田	1	2	3	4	5	样本均值
使用杀虫剂之前	A_1	4．0	4．2	4．3	4．4	3．9	4．16
	B_1	3．1	3．4	3．9	3．8	3．3	3．50
使用杀虫剂之后	A_2	5．2	5．4	4．9	5．1	5．0	5．12
	B_2	4．6	4．2	4．1	4．4	4．0	4．26

5.　结论

本文通过选取两个不同总体的样本进行检验分析，比传统检验方法更具有普遍性，为实际生产生活提供一种好坏评判标准。这种方法也存在不足，它忽略了基底对增长速度的影响，有待进一步研究，但对生产的指导还有实际意义。

参考文献

[1]W.J.Conover.实用非参数统计[M].北京:人民邮电出版社,2006.
[2]邰淑彩,何娟娟.应用数理统计[M].武汉:武汉大学出版社,2005.
[3]高雷阜,柴岩.概率论与数理统计[M].沈阳:东北大学出版社,2006.

The Application of Significance Test in Selecting The Superior Scheme

Li Jun
Colleage of Science, Liaoning Technical University, Fuxin, Liaoning (123000)

Abstract
Use of significant test method to study the quality of two or more individual assessment of the problem, first choice of the traditional hypothesis testing to test whether any two reform proposals are valid, and by constructing new statistical variables and with the appropriate distribution Form. Significant test, practice tests to prove that this reform programme with the selection of guiding significance.
Keywords: significance test; t- distribution; range-variaty