

R12 Hypothesis Testing

Introduction to the Steps of Hypothesis Testing

设假设

Null Hypothesis: 拒绝、怀疑的

Alternative Hypothesis: 希望得到的结论

=放在原假设中，因为连续型随机变量分布的原因，例如， $P(x=3)=0$ ，其发生概率为零（不可能发生），因此我们假设其为假的事件，如果放在备择假设中，则备择假设中就包含了一个假的事件，从而产生悖论

1个 $\mu$  — z/t

独立 — t

2个 $\mu$

不独立 — 成对数检验

Test type	Assumptions	H <sub>0</sub>	Test-statistic	Critical value
Mean hypothesis testing	Normally distributed population, <u>known population variance</u>	$\mu=0$	$Z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$	N(0,1)
	Normally distributed population, <u>unknown population variance</u>	$\mu=0$	$t = \frac{\bar{x} - \mu_0}{s / \sqrt{n}}$	t(n-1)
	Independent populations, <u>unknown population variances assumed equal</u>	$\mu_1 - \mu_2 = 0$	$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{s_p^2 / n_1 + s_p^2 / n_2}}$ , where $s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$	t(n <sub>1</sub> + n <sub>2</sub> - 2)
	Independent populations, <u>unknown population variances not assumed equal</u>	$\mu_1 - \mu_2 = 0$	$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{s_1^2 / n_1 + s_2^2 / n_2}}$	t*
	<u>Samples not independent</u> paired comparisons test	$\mu_d = 0$	$t = \frac{\bar{d}}{s_d / \sqrt{n}}$	t(n-1)

1个方差 — Chi-squared分布

2个方差 — F分布

Test type	Assumptions	H <sub>0</sub>	Test-statistic	Critical value
Variance hypothesis is testing	Normally distributed population	$\sigma^2 = \sigma_0^2$	$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2}$	$\chi^2(n-1)$
	Two independent normally distributed populations	$\sigma_1^2 = \sigma_2^2$	$F = \frac{s_1^2}{s_2^2}$	$F(n_1 - 1, n_2 - 1)$

双尾：不等号

拒绝域在右边：大于号

单尾

拒绝域在左边：小于号

画出拒绝域（本质：根据备择假设判断双尾还是单尾）

找出拒绝域的面积（本质：确定显著性水平）

找出k值（本质：根据分布和显著性水平找出对应的关键值）

Test Statistic =  $\frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}}$

Test Statistic =  $\frac{\bar{X} - \mu_0}{S / \sqrt{n}}$

计算检验统计量

需要掌握对1个总体均值检验的检验统计量的计算方式

比较k值和检验统计量

Reject H<sub>0</sub> if |test statistic| > critical value

Fail to reject H<sub>0</sub> if |test statistic| < critical value

总结陈词

reject ..... is significantly different from .....

fail to reject cannot say "accept the null hypothesis", only can say "cannot reject"

..... is not significantly different from .....

定义

P-value值就是能够使得原假设被拒绝的最小的显著性水平

P-Value

判断：越小越拒绝

P-value <  $\alpha$  — reject H<sub>0</sub>

P-value >  $\alpha$  — do not reject H<sub>0</sub>

定义

Type I error: 拒真（“错杀好人”），即H<sub>0</sub>是真的，反而被拒绝

Type II error: 取伪（“放走坏人”），即H<sub>0</sub>是假的，反而没有被拒绝

P(Type I error) =  $\alpha$

Power of test = 1 - P(Type II error)

假设检验的势，正确的拒绝一个错误的原假设的概率

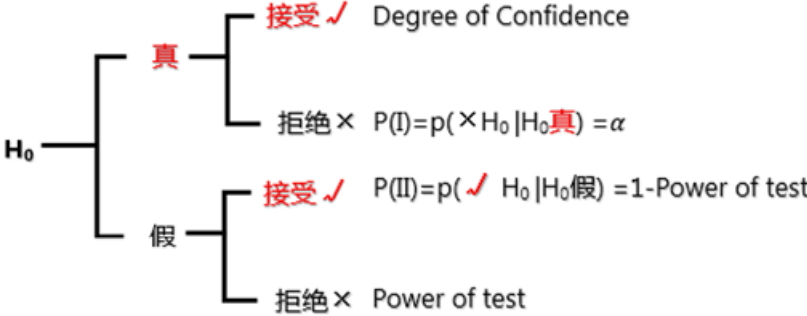
其他条件不变时，P(I)和P(II)此消彼长

n增大，P(I)和P(II)都减小

一类错误的概率和二类错误的概率相加并不一定等于1

二类错误更为危险，应该尽量避免二类错误的发生

Type I Error and Type II Error



Parametric Test: 对总体参数进行检验

Parametric and Nonparametric Tests

Nonparametric Test

不满足分布的假设

非正态分布小样本

序数排列

第1名，第2名

检验的不是参数

(++--+-++++)