第五讲　假设检验

5.1 假设检验的基本概念

关于总体参数的推断统计有

- 参数估计
  - 点估计
  - 区间估计
- 假设检验

假设检验: 利用样本，对于总体中的某个参数或某种性质进行判断

例题：某十字路口每天早上 8点到 9点之间通过的车辆

从长期观察得知： $\mu = 2500;$

确定：新的信号灯系统是否明显增加了交通流量?

随机观察 25 天，得到： $\bar{x} = 2503$

$\bar{x} = 2800$

例题：某大学本科教学评估

本学期两位基础课教师的学生都为200人。在每一个班中随机抽取30人。经计算得到：

甲老师的平均分数是： **87**

乙老师的平均分数是： **86.9**

问题：

（1）是否可以确定甲老师的教学效果优于乙老师？

（2）为什么？

假设检验要解决的问题：判断我们所观察到的现象是随机因素造成的，还是系统因素造成的？

掷硬币：　硬币质地均匀

理想情况：

0	1
1/2	1/2

实验结果：（受随机因素的影响）

实验者	投掷次数n	正面m	m/n
蒲丰	4040	2048	0.5069
皮尔逊	12000	6019	0.5016
维尼	30000	14994	0.4998

例1：利用样本提供的信息，检验所提出的假设是否合理

某炼铁厂铁水含碳量：$X \sim N(4.55, 0.108^2)$

现改变工艺条件。检测5炉铁水，其含碳量为：

4.28，4.40，4.42，4.35，4.37

求得：

$$\bar{x} = \frac{1}{5}\sum_{i=1}^{5} x_i = 4.364 \neq 4.55$$

（假设方差无变化）

问题： $\bar{x} \neq \mu$

是因为测量造成的（偶然因素——随机误差），还是由于工艺条件改变造成的（系统性因素）？

假设检验的基本思路：

零假设 H_0 : 新工艺对含碳量无影响

[即： X 依然服从 $N(4.55, 0.108^2)$]

记： $\mu_0 = 4.55$

H_0: $\mu = \mu_0$ (所观察到的现象是随机误差造成的）

H_1: $\mu \neq \mu_0$ （所观察到的现象是真实的）

构造“检验统计量”：

$$Z = \frac{\overline{X} - \mu_0}{\sigma / \sqrt{n}}$$

问题： Z 服从什么分布？？？

问题： Z 在什么条件下服从标准正态分布？

$$X \sim N(\mu, 0.108^2) \leftarrow (X_1, \cdots, X_n)$$

当零假设为真时（ $\mu = \mu_0 = 4.55$ ）：

$$Z = \frac{\overline{X} - \mu_0}{\sqrt{\sigma^2 / n}} \overset{H_0}{\sim} N(0,1)$$

选择**检验水平**： $\alpha\ (= 0.05)$

由于

$$P\{|Z| \geq 1.96\} = 0.05$$

可知：在 H_0为真的情况下,事件 $\{|Z| \geq 1.96\}$ **是一个小概率事件。**

判断方法：

计算:

$$|Z| = \left|\frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}\right| = \left|\frac{4.364 - 4.55}{0.108 / \sqrt{5}}\right| = 3.85 > 1.96$$

一般认为，小概率事件是不经常发生的。如果在原假设成立的条件下，在一次实验中，小概率事件居然发生了！因此怀疑 H_0 假设的正确性，而认为工艺条件的改变使总体均值发生了显著变化。因此拒绝 H_0，认为 $\mu \neq \mu_0$；

注意正确的叙述方法

不准确的表述方法:“我们计算得到:

$$|Z| = \left|\frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}\right| = \left|\frac{4.364-4.55}{0.108/\sqrt{5}}\right| = 3.85 > 1.96$$

这是一个小概率事件，所以……”

而实际情况更可能是:

由于我们假定 $X \sim N(4.55,\ 0.108^2)$，因此大部分观测的 $\bar{x}$ 都应分布在 $4.55 \pm 1.96\dfrac{0.108}{\sqrt{5}}$ 的范围内。现在发现 $\bar{x} = 4.364$ 出了这个范围，说明样本可能来源于另一个分布。

概念:

— 零假设 **Null Hypothesis** H_0: $\theta = \theta_0$

— 备择假设 **Alternative Hypothesis**: H_1: $\theta \neq \theta_0$

— 检验统计量 **Test Statistic**:

用于确定是否拒绝 H_0 的统计量;

—检验水平 **Level of Significance** α

— 拒绝域 **Rejection Region** $(-\infty,-z_{\alpha/2}),\ (z_{\alpha/2},+\infty)$

— **接受域** **Acceptance Region** $(-z_{\alpha/2}, z_{\alpha/2})$

—**临界值** **Critical Value:** $-z_{\alpha/2}$ *and* $z_{\alpha/2}$

Separate the rejection region from acceptance region.

双侧检验 $H_0:\theta=\theta_0$, $H_1:\theta\neq\theta_0$

（双边检验）

单侧检验

(1) $H_0:\theta=\theta_0$, $H_1:\theta>\theta_0$(右尾检验）

(2) $H_0:\theta=\theta_0$, $H_1:\theta<\theta_0$(左尾检验）

简单零假设问题

$$H_0 : \theta = \theta_0 \qquad H_1 : \begin{cases} \theta \neq \theta_0 \\ \theta > \theta_0 \\ \theta < \theta_0 \end{cases}$$

复合零假设问题

$$(1) \quad H_0 : \theta \leq \theta_0 , \qquad H_1 : \theta > \theta_0$$

$$(2) \quad H_0 : \theta \geq \theta_0 , \qquad H_1 : \theta < \theta_0$$

第一类错误、第二类错误

第一类错误（Type I Error）：弃真错误

当 H_0 为真时，拒绝 H_0 .

$$\alpha = P(\text{第一类错误}) = P(\text{拒绝 } H_0 | H_0 \text{ 为真})$$

$$P(\text{检验统计量落入拒绝域} | H_0 \text{ 为真})$$

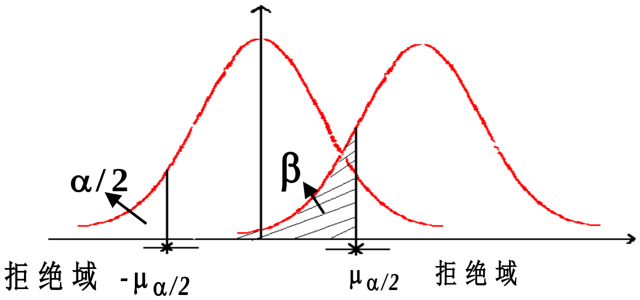
问题：为什么会犯第一类错误，犯第一类错误的概率是多大

第二类错误（Type II Error）：取伪错误

当 H_0 为错误时，接受 H_0 .

$$\beta = P(\text{第二类错误})$$

$$= P(\text{检验统计量落入接受域} | H_0 \text{ 为假})$$



问题：为什么会犯第二类错误，犯第二类错误的概率是多大

讨论问题：

（1）α 的取值是不是越小越好？

答案：不是！

因为真弃概率越小，有可能取伪概率会越大

英国统计学家Ronald Fisher把小概率的标准定为0.05 常被作为一个普遍适用的经验原则。

（2）怎样才能同时减少弃真概率和取伪概率？

答案：增加样本容量，实因检验是根据 $\bar{x}$ 的分布

5.2 关于总体均值 μ 的检验，σ^2已知

例题： Mataltech公司声称，其钻头寿命(feet)服从正态分布$X \sim N(32,16)$ 。

抽取一个容量为$n =$ 25的样本：$\bar{x} = 29.5$

(1) H_0: $\mu = 32$

H_1: $\mu < 32$ （左尾检验）

(2) 选择检验水平 $\alpha = 0.05$

(3) 选择统计量：

$$\boxed{Z = \frac{\overline{X} - \mu_0}{\sqrt{\sigma^2 / n}} \overset{H_0}{\sim} N(0,1)}$$

如果 H_0 为真, 则有

$$P\{Z < -z_\alpha\} = 0.05$$

$$\Rightarrow -z_\alpha = -1.645 \text{ (临界值)}$$

计算观察的 Z score:

$$z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} = \frac{29.5 - 32}{4 / \sqrt{25}} = -3.125 < -1.645$$

判断: 拒绝 H_0. 因此 Matalltech 过分宣传了其产品的寿命

计算临界值 $\bar{x}^*$：

由
$$z^* = \frac{\bar{x}^* - \mu_0}{\sigma / \sqrt{n}} = -1.645$$

临界值 $\bar{x}^*$ 等于：

$$\bar{x}^* = \mu_0 - z_\alpha \sigma / \sqrt{n}$$

$$= 32 - 1.645 \times 4 / \sqrt{25} = 30.684$$

我们拒绝 H_0，实因

$$\bar{x} = 29.5 < 30.684$$

5.3 关于总体均值 μ 的检验，大样本

对于大样本: $n \geq 30$

(1)根据**中心极限定理**

$$\bar{X} \sim N(\mu, \sigma^2 / n)$$

(2) 在大样本条件下:

$$s \approx \sigma$$

<u>检验统计量:</u>

$$\boxed{Z = \frac{\bar{X} - \mu_0}{S / \sqrt{n}} \overset{H_0}{\sim} N(0,1)}$$

例题：

一种罐装饮料采用自动生产线，每罐的容量是255ml。为检验每罐容量是否符合要求，质量检验人员在每天生产的饮料中随机抽取了40罐进行检验，测得每罐平均容量为255.8ml，**标准差是5ml** 。取显著水平为0.05，检验该天生产的饮料容量是否符合标准要求？

$$\mu_0 = 255,\quad n = 40,\quad \bar{x} = 255.8,\quad s = 5\ ,\alpha = 0.05$$

(1) $H_0: \mu = 255$

$H_1: \mu \neq 255$ （双尾检验）

(2)选择检验统计量: $Z = \dfrac{\bar{X} - \mu_0}{S/\sqrt{n}} \underset{\mu=\mu_0}{\overset{H_0}{\sim}} N(0,1)$

(3) 选择检验水平 $\alpha = 0.05$

(4) 拒绝域: $\mathrm{P}\left\{ |Z| > z_{\alpha/2} \right\} = 0.05$

查表： $z_{\alpha/2} = 1.96$

(5) 由 $\bar{x} = 255.8$, $s = 5$, 计算

$$z = \frac{255.8 - 255}{5/\sqrt{40}} = 1.01 < 1.96$$

(6) 判断: 不能拒绝 H_0 .

注意： 不能拒绝$H_0 \neq$ 接受H_0

复合零假设问题

如果检验问题为：

H_0: $\mu \geq \mu_0$

H_1: $\mu < \mu_0$

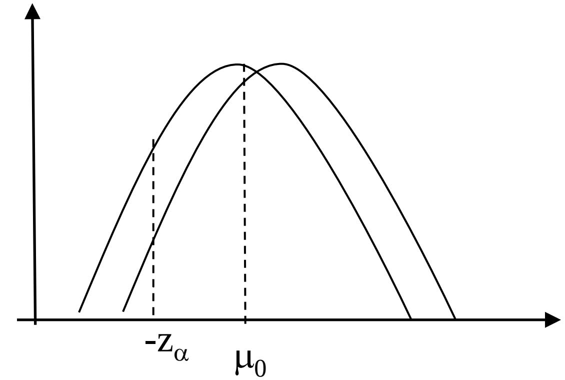
只需要采用和下面检验问题完全相同的研究步骤即可。

H_0 : $\mu = \mu_0$

H_1: $\mu < \mu_0$

问题① H_0: $\mu = \mu_0$

H_1: $\mu < \mu_0$



问题② H_0: $\mu \geq \mu_0$

H_1: $\mu < \mu_0$

如果在问题①中拒绝了H_0⇒在问题②中拒绝H_0

如果在问题①中接受了H_0⇒即：接受 $\mu = \mu_0$

在问题②中接受H_0

所以，检验问题①与检验问题②的结果等价。

例题：某小麦品种的平均产量一直低于5200kg/hm^2. 为检验改良后新品种的平均产量是否有显著提高，随机抽取了36个地块进行种植。得到的样本平均产量为5275kg/hm^2，标准差为120kg/hm^2。问改良品种的产量是否有显著提高？ α = 0.05

(1) $H_0: \mu \le 5200$

$H_1: \mu > 5200$ （右尾检验）

(2) 选择 Z 检验;

(3) 显著水平: α = 0.05 $z_\alpha = z_{0.05} = 1.645$

(4) 临界值 :

(5) 随机样本 n=100, $\bar{x} = 5275$

$$z = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{5275 - 5200}{120/\sqrt{36}} = 3.75 > 1.645$$

(6) 拒绝 H_0.（改良品种的产量有显著提高）

5.4 关于总体均值 μ 的检验，σ^2未知

例： 某汽车制造商声称该厂生产的汽车修理费用服从正态分布，并低于\$200。消费者协会随机抽取 $\boldsymbol{n = 8}$ 辆汽车，测得 $\bar{x} = 230, \quad s^2 = 2337.5$

可否判断制造商的宣传是过分的？ （先用直觉判断）

(1) $H_0: \mu \le \$200$

$H_1: \mu > \$200$ （右尾检验）

(2) 构造统计量

$$t = \frac{\bar{x} - \mu_0}{s / \sqrt{n}} \underset{\mu = \mu_0}{\overset{H_0}{\sim}} t(n-1)$$

(3) 选择检验水平 $\boldsymbol{\alpha = 0.05}$;

(4)确定临界值: $t_{0.05}(8-1)=1.895$

(5) 由观测样本 : $\bar{x}$, s^2

$$\bar{x}=230, \quad s^2=2337.5$$

采用下式计算 t score

$$t=\frac{\bar{x}-\mu_0}{s/\sqrt{n}}=\frac{230-200}{48.35/\sqrt{8}}=1.755<1.895$$

(6) 决策: 不拒绝 H_0. （为什么？）

(7) 临界值:

$$\bar{x}^*=\mu_0+t_\alpha\frac{s}{\sqrt{n}}=200+1.895\frac{48.35}{\sqrt{8}}=232.394$$

5.5 关于总体比率的检验 (大样本)

大样本: $np_0 \geq 5$

某公司拟开发新产品，计算出盈亏平衡点为：市场占有率=10%。因此，只有当市场占有率大于10%才可获利。现抽取100个潜在用户，其中有14%的用户表示有意购买此新产品。问：**可否判断实际市场占有率大于10%（ α= 0.05）？**（请先用经验判断）

已知：$n = 100$, $p_0 = 10\%$, $\hat{p} = 14\%$

解：

(1)提出假设

$$H_0: p \le 10\%$$

$$H_1: p > 10\% \qquad \text{（右尾检验）}$$

(2)选择统计量

$$Z = \frac{\hat{p} - p_0}{\sqrt{p_0 q_0 / n}} \underset{p=p_0}{\overset{H_0}{\sim}} N(0,1)$$

(3)给α= 0.05,构造一“小概率事件” （拒绝域）

(4)确定拒绝域的边界值。

查表：

$$P\{Z > z_\alpha\} = 0.05$$

$$z_\alpha = 1.645$$

(5)计算 $$z=\frac{0.14-0.10}{\sqrt{0.1\times0.9/100}}=1.33<1.645$$

(6) 判断：**不能拒绝H_0**，不能判断市场占有率大于10%。

(7)问题：当 $\hat{p}$ 多大时，可以拒绝H_0?

可开发的前提：

即：$$\frac{\hat{p}-0.1}{\sqrt{0.1\times0.9/100}}\geq1.645$$

$$\hat{p}^*=p_0+z_\alpha\sqrt{p_0q_0/n}$$

$$=0.1+1.645\sqrt{0.1\times0.9/100}=15\%$$

所以，最低样本市场占有率应为 15%。

样本容量太低！如果： n =1000，则有z = 4.216>1.645

假设检验的工作程序总结

（1）叙述零假设 H_0 和备择假设H_1；

（2）**选择检验统计量（在H_0为真的情况下，该统计量服从某个标准分布，可以查表）。**

（3）选择检验水平 α，构造“小概率事件”**(拒绝域)**

（4）确定拒绝域的边界值；

（5）一次抽样，计算样本的检验统计量数值；

（6）判断：如果“小概率事件”发生，拒绝H_0；否则，不拒绝H_0；

不同的检验问题，选取不同的检验统计量。

5.6 进一步的讨论

—双侧检验与置信区间的联系

— p–Value

— 检验效率与样本容量的关系

讨论 I. 双侧检验与置信区间

对于 $X \sim N(\mu, \sigma^2)$, σ^2 已知

总体均值的置信区间为：

$$\left(\overline{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}, \overline{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \right)$$

接受域为：

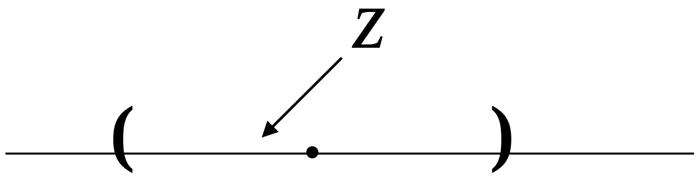
$$\left(\mu_0 - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}, \mu_0 + z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \right)$$

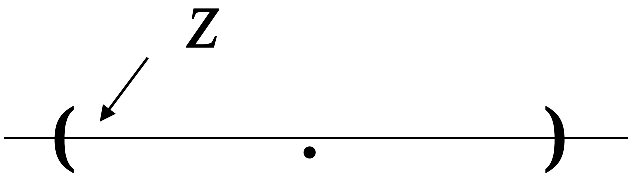
在H_0为真的条件下，构造一个概率为95%的区间, 使得 Z-value 超出此区间的可能性很小。

讨论 II.　*P*-Value

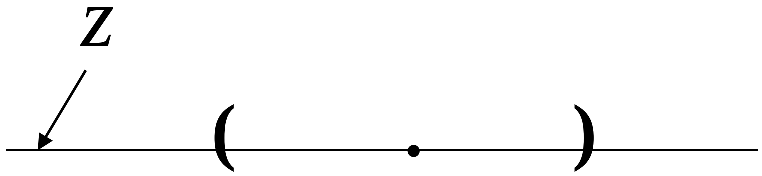
1. 拒绝域和接受域

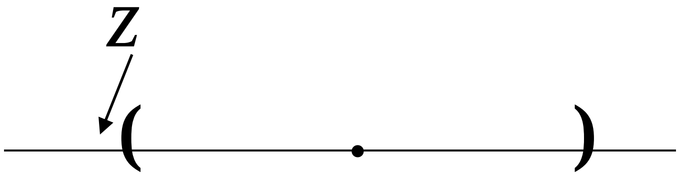
当 H_0 为真时：



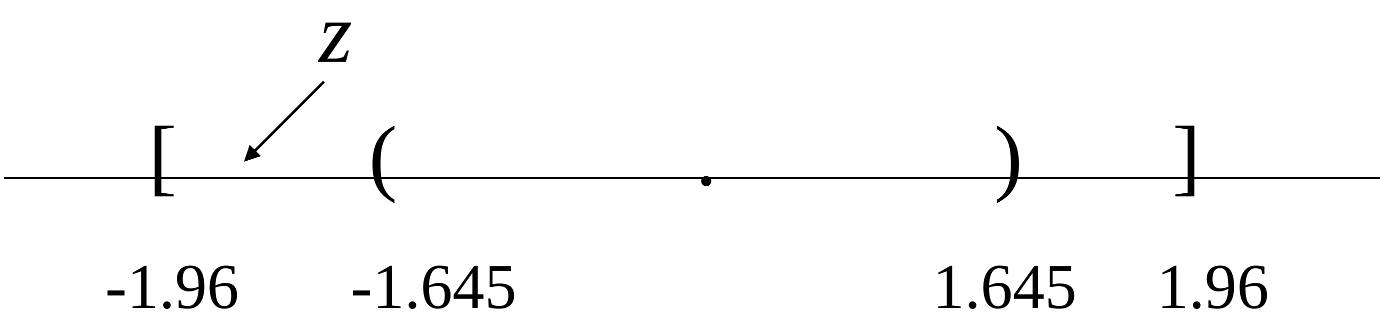


当 H_1 为真时:





2. 如何选择检验水平α?



α =0.10 : 拒绝 H_0

α =0.05 : 不拒绝 H_0

利用P值进行决策：

P- Value: 观察到的显著水平

犯第一类错误的概率

例题：钻头寿命

抽取一个随机样本 $n = 25$, $\bar{x} = 29.5$

H_0: $\mu = 32$

H_1: $\mu < 32$ （左尾检验）

观察到的 Z score 是

$$z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} = \frac{29.5 - 32}{4 / \sqrt{25}} = -3.125$$

p-Value **（NORMSDIS）**：

$$\mathrm{P}\{Z < -3.125\} = 0.000889093 < 0.05$$

这个概率值过分小了.因此我们拒绝 H_0.

例题：新产品开发的市场占有率

$$H_0: p \le 10\%$$

$$H_1: p > 10\%$$ （右尾检验）

$$p = 0.1, \quad \hat{p} = 0.14, \quad n = 100$$

$$Z = \frac{0.14 - 0.10}{\sqrt{0.1 \times 0.9 / 100}} = 1.33$$

$$\mathrm{P}\{Z \ge 1.33\} = 1 - 0.908 = 0.092$$

如果指定0.05为小概率事件，则这个事件的发生概率：0.092> 0.05。因此，我们不拒绝H_0.

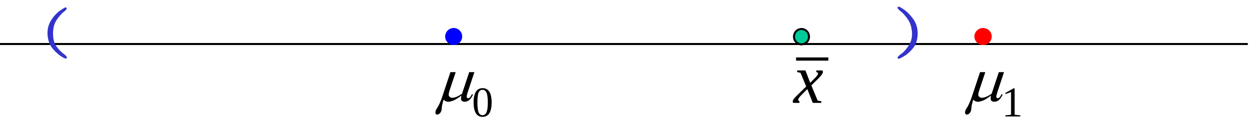
如果指定0.10为小概率事件，则这个事件的发生概率：0.092 < 0.10。因此，拒绝H_0

讨论III、检验效率与样本容量的关系

若$\bar{x}$来源于$N(\mu_1,\sigma^2)$

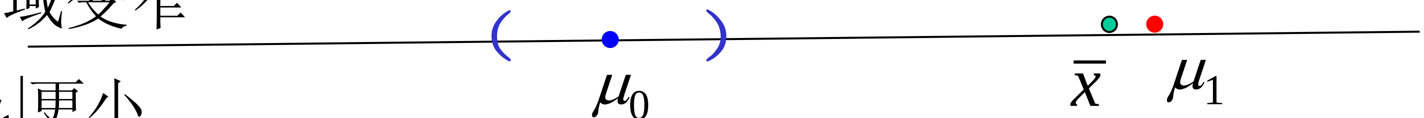
接受域的半长 $\boxed{z_{\alpha/2}\dfrac{\sigma}{\sqrt{n}}}$

接受域过宽，
无法拒绝H_0



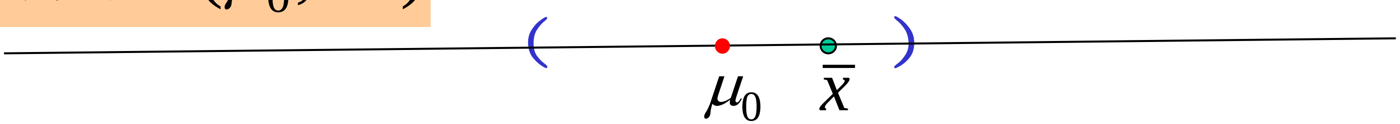
样本容量增大:

(1)接受域变窄



(2) $|\bar{x}-\mu_1|$更小

若$\bar{x}$来源于$N(\mu_0,\sigma^2)$



95%的置信区间：$|\bar{x}-\mu_0|\le z_{\alpha/2}\dfrac{\sigma}{\sqrt{n}}$,

当$n\to+\infty, |\bar{x}-\mu_0|\le z_{\alpha/2}\dfrac{\sigma}{\sqrt{n}}$会变得很小

如何确定假设检验的样本容量(讨论):

接受域的半长：

$$\boxed{z_{\alpha/2}\frac{\sigma}{\sqrt{n}}=D}$$

$$\frac{D}{\mu_0}\le 5\% \quad\Rightarrow\quad z_{\alpha/2}\frac{\sigma}{\sqrt{n}}=0.05\mu_0$$

$$n=\left(\frac{z_{\alpha/2}\cdot\sigma}{0.05\mu_0}\right)^2$$

这时接受域已经充分狭窄,如果这时 $\bar{x}$ 还在接受域内,说明 $\bar{x}$ 与μ_0之间的误差也在这个范围内。而我们认为,当它们之间的**相对误差在0.05以内**, $\bar{x}$ 就是μ_0的很好的估计了.

5.7 两个总体的假设检验

关于以下零假设的检验:

$(\mu_1 - \mu_2) = D_0$ σ_1^2 and σ_2^2 已知

$(\mu_1 - \mu_2) = D_0$ 大样本

$(\mu_1 - \mu_2) = D_0$ σ_1^2 and σ_2^2 未知

$(p_1 - p_2) = 0$

$\sigma_1^2 / \sigma_2^2 = 1$

一、检验统计量的选择依据

1、两个总体均值之差的样本分布：

(1) 正态总体, σ_1 和 σ_2 已知, 独立样本；

$$\frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} \sim N(0,1)$$

(2) σ_1 和 σ_2 未知, 大样本；

$$\sigma_1 \approx s_1 \qquad \sigma_2 \approx s_2$$

(3) 正态总体, 小样本,

σ_1 , σ_2 未知, 但是 $\sigma_1 = \sigma_2 = \sigma$;

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

$$t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t(n_1 + n_2 - 2)$$

2、两个总体比率之差的样本分布（大样本）

$$(\hat{p}_1 - \hat{p}_2) \sim N\left((p_1 - p_2), (\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}) \right)$$

则有：

$$Z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{p_1 q_2 / n_1 + p_2 q_2 / n_2}} \sim N(0,1)$$

3、两个总体方差之比的样本分布

情况: X 与 Y 都是正态总体

X 与 Y 相互独立

$$F=\frac{S_X^2/\sigma_X^2}{S_Y^2/\sigma_Y^2}\sim F(n_1-1,n_2-1)$$

$(n_1$-1): 分子自由度

$(n_2$-1): 分母自由度

问题: $H_0:\sigma_X^2=\sigma_Y^2$, 如何选择研究统计量?

$$F=\frac{S_X^2}{S_Y^2}\sim F(n_1-1,n_2-1)$$

4、独立样本和配对样本

— 独立样本：

在两个样本中不包含共同的样本点;

— 配对样本:

来自两个总体的所以样本点都是成对的或者是有关联的

二、几种常用的检验方法

1、 总体均值之差的检验

正态总体, σ_1^2 and σ_2^2 已知

<u>独立样本</u>:

两个样本中不包括同样的观测个体.

$$H_0: \mu_1 = \mu_2$$

$$H_1: \mu_1 \neq \mu_2$$

统计量: $$Z = \frac{(\overline{X}_1 - \overline{X}_2) - 0}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} \overset{H_0}{\sim} N(0,1)$$

在大样本的情况下

统计量：$$Z=\frac{(\overline{X}_1-\overline{X}_2)-0}{\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}}\overset{H_0}{\sim}N(0,1)$$

例题：

某家庭日用食品商店在六个月内作了两次调查，以了解家庭每月平均消费量有无变化。结果如下

第一次：$n_1 = 200,\ \ \bar{x}_1 = 4.91,\ \ s_1^2 = 0.4$

第二次：$n_2 = 250,\ \ \bar{x}_2 = 4.97,\ \ s_2^2 = 0.5$

$(1) H_0: \mu_1 = \mu_2$

$H_1: \mu_1 \neq \mu_2$

(2) 统计量: $Z = \dfrac{\bar{X}_1 - \bar{X}_2}{\sqrt{\dfrac{S_1^2}{n_1} + \dfrac{S_2^2}{n_2}}} \overset{H_0}{\sim} N(0,1)$

$(3) \alpha = 0.05 \Rightarrow \pm z_{\alpha/2} = \pm 1.96$

$(4) z = \dfrac{4.91 - 4.97}{\sqrt{\dfrac{0.4}{200} + \dfrac{0.5}{250}}} = -0.9487 > -1.96$

(5) 不能拒绝 H_0

2、 总体均值之差的检验

小样本，总体方差未知，$\sigma_1^2 = \sigma_2^2$

$$H_0: \mu_1 - \mu_2 = 0$$

$$H_1: \mu_1 - \mu_2 \neq 0$$

统计量： $$t = \frac{(\overline{X}_1 - \overline{X}_2) - 0}{\sqrt{s_p^2 / n_1 + s_p^2 / n_2}} \overset{H_0}{\sim} t(n_1 + n_2 - 2)$$

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

例题：

现有两种不同的热处理方法对金属材料做抗拉强度实验，得到实验数据（单位：公斤/厘米^2）为：

甲方法：$n_1 = 12$, $\bar{x}_1 = 31.75$, $(n_1 - 1)s_1^2 = 112.25$

乙方法：$n_2 = 12$, $\bar{x}_2 = 28.67$, $(n_2 - 1)s_2^2 = 66.64$

设两总体均为正态总体，且方差相等。在给定显著性水平0.05时，问两种方法得到的抗拉强度有无显著差异？

$$H_0: \mu_1 = \mu_2$$

$$H_0: \mu_1 \neq \mu_2$$

$$\alpha = 0.05, \quad (n_1 + n_2 - 2) = 22 \Rightarrow \pm t_{\alpha/2}(22) = 2.074$$

$$t = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\dfrac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}}$$

$$= \frac{31.75 - 28.67}{\sqrt{\dfrac{112.25 + 66.64}{12 + 12 - 2}\left(\dfrac{1}{12} + \dfrac{1}{12}\right)}} = 2.646 > 2.074$$

拒绝 H_0.

3、 总体均值之差的检验
—— 配对样本

配对样本是不独立的！

假设随机选取10辆汽车。首先将10辆汽车加满标号1的汽油，测量它们的行驶公里数。然后，再将这10辆车加满标号2的汽油。数据见下表：

汽车	1	2	3	4	5	6	7	8	9	10
标号1	14	21	19	11	15	16	8	32	37	10
标号2	16	24	20	15	17	19	10	33	39	11
差值D_0	-2	-3	-1	-4	-2	-3	-2	-1	-2	-1

错误地使用 t 检验

汽车	标号 1	标号 2	差值D_0
Mean	**18.3**	**20.4**	**-2.1**
Std.	**89.79**	**86.27**	**0.9889**

随机选取10辆汽车

标号 1: $\bar{x}_1 = 18.3, \quad s_1^2 = 89.79$

标号 2: $\bar{x}_1 = 20.4, \quad s_1^2 = 86.27$

(1) $H_0: \mu_1 - \mu_2 = 0$

$H_1: \mu_1 - \mu_2 \neq 0$

$$(2) \quad t=\frac{(\bar{x}_1-\bar{x}_2)-0}{\sqrt{s_p^2/n_1+s_p^2/n_2}} \overset{H_0}{\sim} t(n_1+n_2-2)$$

$$s_p^2=\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}$$

$$(3)\alpha=0.05 \Rightarrow \text{临界值}: \pm t_{\alpha/2}(18)=\pm 2.101$$

$$P\left\{t<-2.101\right\}+P\left\{t>2.101\right\}=0.05$$

$$s_p^2=\frac{9\times 89.79+9\times 86.27}{10+10-2}=88.03$$

$$(4) \quad t=\frac{18.3-20.4}{\sqrt{88.03/10+88.03/10}}=-0.5$$

(5)不能拒绝 H_0 .

但在原始数据表：标号2的数值均大于标号1

正确地检验方法

随机选取 n 对样本点:

$$(x_{11}, x_{12}), (x_{21}, x_{22}), \cdots, (x_{n1}, x_{n2})$$

对于每一对 $i = 1, 2, \cdots, n$, 计算： $d_i = x_{i1} - x_{i2}$

$$\bar{d} = \frac{1}{n}\sum_{i=1}^{n} d_i, \quad S_d^2 = \frac{1}{n-1}\sum_{i=1}^{n}(d_i - \bar{d})^2$$

记： $\mu_d = \mu_1 - \mu_2$

(1) $H_0 : \mu_d = 0$

$H_1 : \mu_d \neq 0$

(2) 使用检验统计量:

$$t=\frac{\bar{d}-0}{s_d/\sqrt{n}} \underset{\mu_d=0}{\overset{H_0}{\sim}} t(n-1)$$

(3) 选择 α:

$$\mathrm{P}\left\{|t|<t_{\alpha/2}\right\}=0.05$$

(4) 计算 t 的观测值.

(5) 拒绝 H_0 假设，如果有:

$$t>t_{\alpha/2} \quad \text{or} \quad t<-t_{\alpha/2}.$$

例题：两种标号汽油行驶公里是否一致

从数据表看到：差值 d_i

$$\bar{d} = -2.1, \quad s_d^2 = 0.9889$$

选择： $\alpha = 0.05, \ \upsilon = n\text{-}1: \quad \pm t_{0.025}(9) = \pm 2.262$

$$\mathrm{t} = \frac{(-2.1) - 0}{\sqrt{0.9889/10}} = -6.68 < -2.262$$

拒绝 H_0 。

阅读与练习

关于总体方差的假设检验

检验两个总体的的方差是否相等

一. 关于总体方差的假设检验

设总体$X \sim N(\mu,\sigma^2)$, $X_1, X_2, \cdots, X_n$是来自X的样本.

检验 $H_0:\sigma^2=\sigma_0^2$, $H_1:\sigma^2\neq\sigma_0^2$, σ_0^2是已知常数

例. 某厂生产的某种型号的电池，其寿命长期以来服从方差$\sigma_0^2=5000$(小时^2)的正态分布，现有一批这种电池，从它的生产情况来看，寿命的**波动性有所改变**. 现随机取26只电池，测得其寿命样本方差为$s^2=9200$(小时^2). 问根据这一数据能否推断这批电池寿命的波动性较以往的有显著的变化？（取$\alpha=0.05$）

$$\boxed{U_0=\frac{(n-1)S^2}{\sigma_0^2}\sim\chi^2(n-1)}$$

二. 检验两个总体的的方差是否相等

设： $X_1, X_2, \cdots, X_{n_1}$ 是来自正态总体 $N(\mu_1, \sigma_1^2)$ 的样本,

$Y_1, Y_2, \cdots, Y_{n_2}$ 是来自正态总体 $N(\mu_2, \sigma_2^2)$ 的样本,

且设两样本独立,

分别记它们的样本均值为 $\overline{X}, \overline{Y}$； 样本方差 S_1^2, S_2^2,

设： $\mu_1, \mu_2, \sigma_1^2, \sigma_2^2$ 均未知.

检验问题：$H_0: \sigma_1^2 = \sigma_2^2, H_1: \sigma_1^2 > \sigma_2^2$.

$$\boxed{F = \frac{s_1^2}{s_2^2} \sim F(n_1 - 1, n_2 - 1)}$$

例2. 在平炉上进行一项试验以确定改变操作方法的建议是否会增加钢的得率, 试验是在同一只平炉上进行的. 每炼一炉钢时除操作方法外, 其它条件都尽可能做到相同. 先用标准方法炼一炉, 然后手建议的方法炼一炉, 以后交替进行, 各炼了10炉, 其得率分别为:标准方法:

78.1 72.4 76.2 74.3 77.4 78.4 76.0 75.5 76.7 77.3

新方法:

79.1 81.0 77.3 79.1 80.0 79.1 79.1 77.3 80.2 82.1

设这两个样本相互独立, 且分别来自正态总体$N(\mu_1,\sigma_1^2)$和$N(\mu_2,\sigma_2^2)$, μ_1, μ_2,σ_1^2, σ_2^2均未知.

试对数据检验假设(α=0.01),H_0: $\sigma_1^2=\sigma_2^2$, H_1: $\sigma_1^2\neq\sigma_2^2$.