概率论与数理统计

何志坚

华南理工大学数学学院

hezhijian@scut.edu.cn

第八章：假设检验

目录

女士品茶试验

奶茶是由牛奶与茶按一定比例混合而成，可以先倒茶后加奶，也可以先倒奶再倒茶。某女士声称她可以鉴别这两种混合方式，周围品茶的人对此产生了议论，都觉得不可思议。在场的费希尔也在思考这个问题，他提议做一项试验来检验如下命题是否可以接受：

假设H：该女士无此种鉴别能力

他准备了10杯调好的奶茶（两种顺序的都有）给该女士鉴别，结果那位女士竟然能够正确地分辨出10杯奶茶中的每一杯的调制顺序。

如何做出你的判断？ 假如假设H是正确的，即该女士无此鉴别能力，她只能靠猜，每次猜对的概率是1/2，连续10次猜对的概率为$2^{-10} < 0.001$，这是一个很小的概率，在一次试验中几乎不会发生，如今发生了，只能说明假设H不成立。假如该女士只猜对了9杯（或者8杯），又该如何判断？

参数假设检验基本概念

设有来自某一参数分布族$\{F(x,\theta)\}$, $\theta \in \Theta$, 其中Θ为参数空间。

原假设（零假设）：

$$H_0 : \theta \in \Theta_0$$

备选假设（对立假设，备择假设）：

$$H_1 : \theta \in \Theta_1$$

其中$\varnothing \neq \Theta_0, \Theta_1 \subset \Theta$，$\Theta_0 \cap \Theta_1 = \varnothing$. 最常见的情况$\Theta_1 = \Theta - \Theta_0$.

简单原假设：Θ_0只包含一个点，如$H_0 : \theta = \theta_0$, 备选假设H_1通常有三种可能：

- 双边假设：$H_1 : \theta \neq \theta_0$
- 单边假设：$H_1 : \theta > \theta_0$
- 单边假设：$H_1 : \theta < \theta_0$

其它情况称为**复杂（复合）原假设**。

假设检验

设X为通话时间，$X \sim N(\mu, \sigma^2)$.

$$H_0 : \mu = 2 \ vs. \ H_1 : \mu \neq 2$$

H_0为原假设（简单原假设）

H_1为备选假设

该假设为双边假设，如何检验?

假设检验基本原则

检验法则：检验本质上是把样本空间划分成两个互不相交的部分W和$\bar{W}$，当样本属于W时就拒绝H_0；否则接受H_0. 称W为该检验的拒绝域，而$\bar{W}$为接受域。

例：假设检验$H_0: \mu = 2$ vs. $H_1: \mu \neq 2$，拒绝域可以设为

$$W = \{(x_1, \ldots, x_n) : |\bar{x} - 2| \geq c\}.$$

c称为临界值。

小概率原理：小概率事件在一次试验中是几乎不发生的。换言之，如果H_0为真的前提下，样本落在拒绝域是一个小概率事件，不应该发生。如果发生，则推翻原命题，即拒绝原假设。

给定一个比较小的数$\alpha \in (0,1)$，通常$\alpha = \{0.01, 0.05, 0.001\}$，可求$c$使得

$$P((X_1, \ldots, X_n) \in W | H_0\text{为真}) = \alpha$$

称值α为显著性水平(或检验水平)，用来衡量原假设与实际情况差异是否明显的标准。

显著性水平的选择

人们自然会产生这样的问题：概率小到什么程度才当作"小概率事件"呢？这要据实际情况而定，例如即使下雨的概率为10%，仍有人会因为它太小而不带雨具。但某航空公司的事故率为1%，人们就会因为它太大而不敢乘坐该公司的飞机，通常把概率不超过0.05 (或0.01)的事件当作"小概率事件"。为此在假设检验时，必须先确定小概率即显著性的值α (即不超过α的概率认为是小概率)。

两类错误

第一类错误： H_0正确，但拒绝了它，这类错误也称“拒真错误”。

$$\alpha = P(\text{拒绝}H_0|H_0\text{为真}) = P((X_1,\ldots,X_n)\in W|H_0\text{为真})$$

第二类错误： H_0不正确(即H_1为真)，但接受了它，这类错误也称“受伪错误”。

$$\beta = P(\text{接受}H_0|H_1\text{为真}) = P((X_1,\ldots,X_n)\in \bar{W}|H_1\text{为真})$$

- $\beta \neq 1-\alpha$, 可以证明，在样本容量一定时，不可能同时缩小两类错误。
- 当样本容量一定时，犯第一类错误的概率越小，则犯第二类错误的概率越大。
- 当现实中样本容量不可能无限制的大，从而同时控制两类错误就不可能。

显著性检验

实际中常用的是只控制第一类错误而不控制第二类错误的检验方法，即显著性检验。当想用显著性检验对某一猜测结论作强有力的支持时，应该将猜测结论的反面作为原假设。这是反证法的思想。

例

例：设总体$X \sim N(\mu, \sigma^2)$, σ已知，$(X_1, \ldots, X_n)$为其样本，对假设检验问题

$$H_0 : \mu = \mu_0, \; H_1 : \mu = \mu_1$$

做检验，其中$\mu_0 < \mu_1$, 解释该检验的第一类错误的概率α与第二类错误的概率β之间的关系。

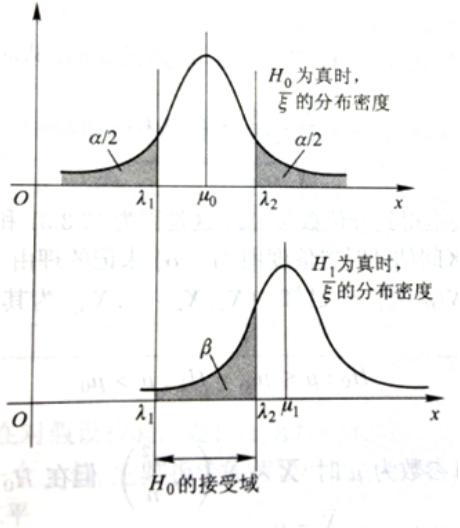
解：选择拒绝域为$W = \{\left|\bar{X} - \mu_0\right| \geq c\}$. 则有

$$\alpha = P(\left|\bar{X} - \mu_0\right| \geq c | \mu = \mu_0) = P\left(\frac{\left|\bar{X} - \mu_0\right|}{\sigma/\sqrt{n}} \geq u_{1-\alpha/2}\right)$$

接受域$\bar{W} = \{\mu_0 - u_{1-\alpha/2}\sigma/\sqrt{n} < \bar{X} < \mu_0 + u_{1-\alpha/2}\sigma/\sqrt{n}\} =: \{\lambda_1 < \bar{X} < \lambda_2\}$.

$$\beta = P(\bar{X} \in \bar{W} | \mu = \mu_1) = \Phi((\lambda_2 - \mu_1)\sqrt{n}/\sigma) - \Phi((\lambda_1 - \mu_1)\sqrt{n}/\sigma).$$

两类错误关系的图示



H_0 为真时，
ξ 的分布密度
$\alpha/2$
$\alpha/2$
O
λ_1
μ_0
λ_2
x
H_1 为真时，
ξ 的分布密度
β
O
λ_1
λ_2
μ_1
x
H_0 的接受域

假设检验的基本步骤

1. 提出假设
2. 选择检验统计量，给出拒绝域形式
3. 求临界值（确定接受域/拒绝域）
4. 算出观察值
5. 作出判断

。

目录

单个正态总体期望的假设检验（双边）

(1)已知方差σ^2

$$H_0: \mu = \mu_0,\ H_1: \mu \neq \mu_0$$

选择检验统计量:

$$U = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \sim N(0,1)$$

求临界值:

$$P(|U| \geq u_{1-\alpha/2}) = \alpha.$$

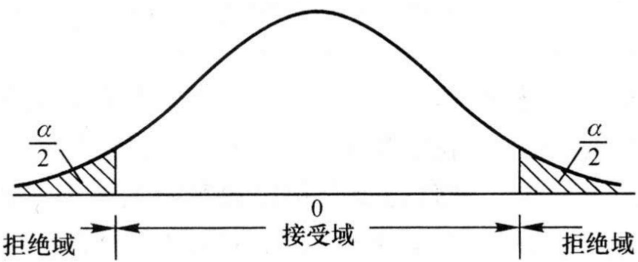
接受域:

$$\bar{W} = \{|U| < u_{1-\alpha/2}\} = \{\mu_0 - u_{1-\alpha/2}\sigma/\sqrt{n} < \bar{X} < \mu_0 + u_{1-\alpha/2}\sigma/\sqrt{n}\}$$

算出观察值：根据样本的观测值算出统计量U的观测值u

作出判断：若$u < u_{1-\alpha/2}$，则接受原假设，若$u \geq u_{1-\alpha/2}$，则拒绝原假设，接受备选假设。这种检验方法称为u检验法。

图例



注：α越大越容易拒绝原假设。

例

例： 已知OPPO手机充电五分钟通话时间$X \sim N(\mu, 0.07^2)$. 随机抽取6部手机测试通话时间（单位：小时）为

$$1.6,\ 2.1,\ 1.9,\ 1.8,\ 2.2,\ 2.1,$$

检验OPPO手机的平均通话时间为2小时是否成立，显著性水平取为5%.

解：

(1)提出假设：$H_0 : \mu = 2$ vs. $H_1 : \mu \neq 2$

(2)选择检验统计量：$U = \frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} \sim N(0,1)$

(3)求临界值：$P(|U| \geq u_{1-\alpha/2}) = \alpha = 0.05$, $u_{1-\alpha/2} = u_{0.975} = 1.96$.

(4)算出观察值：$\bar{x} = (1.6 + \cdots + 2.1)/6 = 1.95$, $u = \frac{|1.95-2|}{0.07/\sqrt{6}} = 1.75 < 1.96$

(5)作出判断：接受原假设

思考：如果$\alpha = 10\%$呢？$u_{0.95} = 1.65$ 拒绝原假设

检验的p值

从上个例子中发现，对于同一个的样本观测值，在一个较小($\alpha = 5\%$)的显著水平下得到接受原假设的结论，而在一个较大($\alpha = 10\%$)的显著水平下得到拒绝原假设的结论。

问题：是否存在一个数$p \in (0,1)$，使得

- 当$\alpha \geq p$时，拒绝H_0
- 当$\alpha < p$时，接受H_0

定义：在一个假设检验问题中，利用样本观测值能够作出拒绝原假设的最小显著水平称为检验的p值。

对于上例，$p = \min\{\alpha : u_{1-\alpha/2} \leq |u|\}$，即$u_{1-p/2} = |u| = 1.75$，于是

$$p = P(|U| \geq 1.75) = 2(1 - \Phi(1.75)) = 8\%$$

注：统计软件一般只给出p值，而不是针对给定的显著性水平α进行判断。

R软件中p值的解释

```
Call:
lm(formula = Balance ~ ., data = ccs)

Residuals:
    Min      1Q  Median      3Q     Max 
-204.86  -79.28  -12.15   70.47  296.61 

Coefficients:
                    Estimate Std. Error t value Pr(>|t|)    
(Intercept)       -549.31402   35.08452 -15.657   <2e-16 ***
Income              -7.77460    0.24389 -31.878   <2e-16 ***
Rating               3.97896    0.05501  72.332   <2e-16 ***
Cards                3.96537    3.79288   1.045   0.2965    
Age                 -0.64159    0.30614  -2.096   0.0367 *  
Education           -0.37986    1.65922  -0.229   0.8190    
GenderFemale       -10.71056   10.32498  -1.037   0.3002    
StudentYes         416.43756   17.33606  24.021   <2e-16 ***
MarriedYes         -15.10961   10.72822  -1.408   0.1598    
EthnicityAsian      21.76158   14.67762   1.483   0.1390    
EthnicityCaucasian  10.64919   12.71571   0.837   0.4028    
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 102.9 on 389 degrees of freedom
Multiple R-squared:  0.9512,	Adjusted R-squared:  0.9499 
F-statistic: 757.8 on 10 and 389 DF,  p-value: < 2.2e-16
```

单个正态总体期望的假设检验（双边）

(2)未知方差σ^2

$$H_0: \mu = \mu_0,\ H_1: \mu \neq \mu_0$$

选择检验统计量：

$$T = \frac{\bar{X} - \mu_0}{S_n/\sqrt{n-1}} \sim t(n-1)$$

求临界值：

$$P(|T| \geq t_{1-\alpha/2}) = \alpha.$$

接受域：

$$\bar{W} = \{|T| < t_{1-\alpha/2}(n-1)\}$$

算出观察值：根据样本的观测值算出统计量T的观测值t

作出判断：若$t < t_{1-\alpha/2}(n-1)$, 则接受原假设，若$t \geq t_{1-\alpha/2}(n-1)$, 则拒绝原假设，接受备选假设。这种检验方法称为t检验法。

例

例： 已知OPPO手机充电五分钟通话时间$X \sim N(\mu, \sigma^2)$. 随机抽取6部手机测试通话时间（单位：小时）为

$$1.6，2.1，1.9，1.8，2.2，2.1，$$

检验OPPO手机的平均通话时间为2小时是否成立，显著性水平取为5%.

解:

(1)提出假设：$H_0 : \mu = 2$ $vs.$ $H_1 : \mu \neq 2$

(2)找检验统计量：$\frac{\bar{X}-2}{S_n/\sqrt{5}} \sim t(5)$

(3)求临界值：$P(|T| \geq t_{1-\alpha/2})(5) = \alpha = 0.05$，$t_{0.975}(5) = 2.5706$.

(4)算出观察值：$\bar{x} = 1.95$，$s_n = 0.206$，$T = \left|\frac{1.95-2}{0.206/\sqrt{5}}\right| = 0.54 < 2.5706$

(5)作出判断：接受原假设

单个正态总体期望的假设检验对比

序号	H_0	H_1	σ^2已知	σ^2未知
I	$\mu=\mu_0$	$\mu\neq\mu_0$	$\frac{\lvert\bar{x}-\mu_0\rvert}{\sigma/\sqrt{n}}\geq u_{1-\alpha/2}$	$\frac{\lvert\bar{x}-\mu_0\rvert}{S_n/\sqrt{n-1}}\geq t_{1-\alpha/2}(n-1)$
II	$\mu=\mu_0$	$\mu>\mu_0$	$\frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}\geq u_{1-\alpha}$	$\frac{\bar{X}-\mu_0}{S_n/\sqrt{n-1}}\geq t_{1-\alpha}(n-1)$
III	$\mu\leq\mu_0$	$\mu>\mu_0$		
IV	$\mu=\mu_0$	$\mu<\mu_0$	$\frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}\leq u_{\alpha}$	$\frac{\bar{X}-\mu_0}{S_n/\sqrt{n-1}}\leq t_{\alpha}(n-1)$
V	$\mu\geq\mu_0$	$\mu<\mu_0$		

Table: 单个正态总体均值μ的假设检验的拒绝域（显著性水平为α）

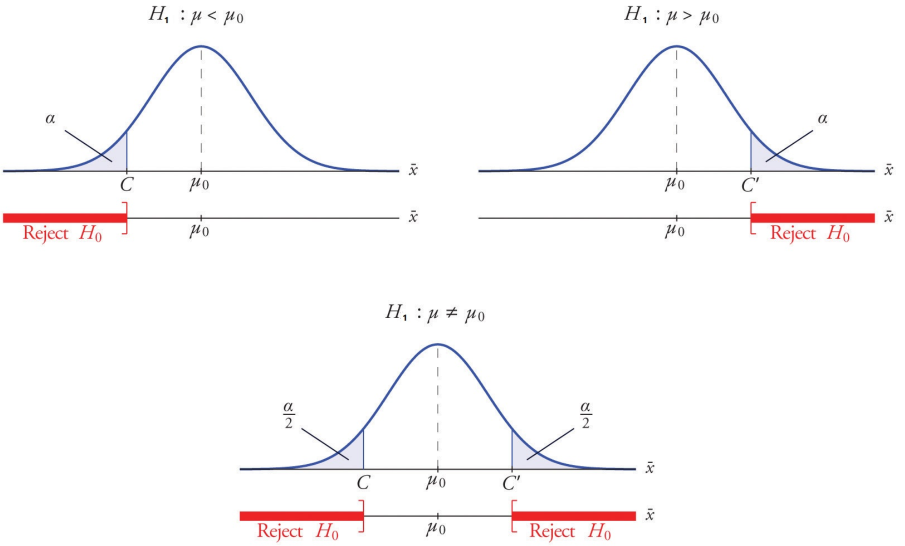
注： 对于情形III, 犯第一类错误的概率为

$$P\left(\frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}\geq u_{1-\alpha}\Big|\mu\leq\mu_0\right)\leq P\left(\frac{\bar{x}-\mu}{\sigma/\sqrt{n}}\geq u_{1-\alpha}\Big|\mu\leq\mu_0\right)=\alpha$$

同理，对于情形V, 犯第一类错误的概率同样不超过α.

思考： 各种情况下的p值如何计算?

单边假设与双边假设


$H_1: \mu < \mu_0$
α
C
μ_0
$\bar{x}$
Reject H_0
$H_1: \mu > \mu_0$
α
μ_0
C'
$\bar{x}$
Reject H_0
$H_1: \mu \neq \mu_0$
$\frac{\alpha}{2}$
$\frac{\alpha}{2}$
C
μ_0
C'
$\bar{x}$
Reject H_0
Reject H_0

假设检验与置信区间的关系

考虑双侧检验(σ^2已知)，显著性水平为α的接受域为

$$\bar{W} = \{\frac{|\bar{X} - \mu_0|}{\sigma/\sqrt{n}} \le u_{1-\alpha/2}\} = \{\bar{X} - u_{1-\alpha/2}\sigma/\sqrt{n} \le \mu_0 \le \bar{X} + u_{1-\alpha/2}\sigma/\sqrt{n}\}$$

为μ的$1-\alpha$置信区间。这表明，假设检验与置信区间是一一对应的。

单个正态总体方差的假设检验（双边）

(1)已知期望μ

$$H_0 : \sigma^2 = \sigma_0^2,\ H_1 : \sigma^2 \neq \sigma_0^2$$

选择检验统计量:

$$\chi^2 = \frac{1}{\sigma_0^2}\sum_{i=1}^{n}(X_i - \mu)^2 \sim \chi^2(n)$$

求临界值:

$$P(\chi^2 \geq \chi^2_{1-\alpha/2}(n)) = P(\chi^2 \leq \chi^2_{\alpha/2}(n)) = \alpha/2.$$

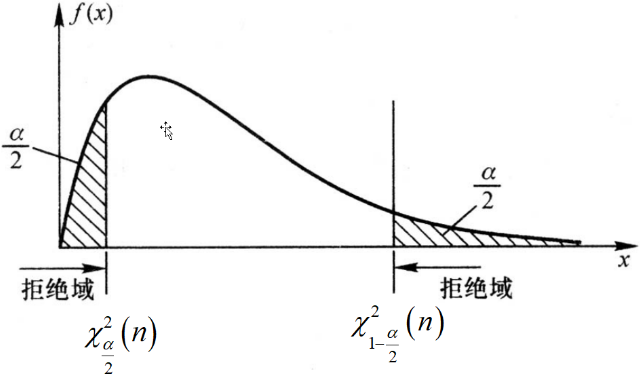
接受域:

$$\bar{W} = \{\chi^2_{\alpha/2}(n) < \chi^2 < \chi^2_{1-\alpha/2}(n)\}$$

算出观察值：根据样本的观测值算出统计量χ^2的观测值χ_1^2

作出判断：若$\chi^2_{\alpha/2}(n) < \chi_1^2 < \chi^2_{1-\alpha/2}(n)$, 则接受原假设，否则拒绝原假设，接受备选假设。这种检验方法称为χ^2检验法。

图例


$f(x)$
$\frac{\alpha}{2}$
$\frac{\alpha}{2}$
x
拒绝域
拒绝域
$\chi^2_{\frac{\alpha}{2}}(n)$
$\chi^2_{1-\frac{\alpha}{2}}(n)$

单个正态总体方差的假设检验（双边）

(2)未知期望μ

$$H_0:\sigma^2=\sigma_0^2,\ H_1:\sigma^2\neq\sigma_0^2$$

选择检验统计量:

$$\chi^2=\frac{1}{\sigma_0^2}\sum_{i=1}^{n}(X_i-\bar{X})^2\sim\chi^2(n-1)$$

求临界值:

$$P(\chi^2\geq\chi^2_{1-\alpha/2}(n-1))=P(\chi^2\leq\chi^2_{\alpha/2}(n-1))=\alpha/2.$$

接受域:

$$\bar{W}=\{\chi^2_{\alpha/2}(n-1)<\chi^2<\chi^2_{1-\alpha/2}(n-1)\}$$

算出观察值: 根据样本的观测值算出统计量χ^2的观测值χ_1^2

作出判断: 若$\chi^2_{\alpha/2}(n-1)<\chi_1^2<\chi^2_{1-\alpha/2}(n-1)$, 则接受原假设, 否则拒绝原假设, 接受备选假设。

例

例： 已知OPPO手机充电五分钟通话时间$X \sim N(\mu, \sigma^2)$. 随机抽取6部手机测试通话时间（单位：小时）为

$$1.6, \ 2.1, \ 1.9, \ 1.8, \ 2.2, \ 2.1,$$

检验OPPO手机的通话时间方差为0.06是否成立，显著性水平取为5%:

- 已知$\mu = 2$.
- 未知μ.

解： 假设检验:

$$H_0 : \sigma^2 = 0.06, \ H_1 : \sigma^2 \neq 0.06$$

(1)检验统计量的观测值为$\frac{1}{0.06}\sum_{i=1}^{n}(x_i - 2)^2 = 4.5 \in (1.24, 14.45)$. 接受$H_0$

(2)检验统计量的观测值为$\frac{1}{0.06}\sum_{i=1}^{n}(x_i - \bar{x})^2 = 4.2 \in (0.83, 12.83)$. 接受$H_0$

单个正态总体方差的假设检验对比

序号	H_0	H_1	μ已知	μ未知
I	$\sigma^2=\sigma_0^2$	$\sigma^2\neq\sigma_0^2$	$\frac{\sum_{i=1}^n(x_i-\mu)^2}{\sigma_0^2}\leq\chi^2_{\frac{\alpha}{2}}(n)$ 或 $\frac{\sum_{i=1}^n(x_i-\mu)^2}{\sigma_0^2}\geq\chi^2_{1-\frac{\alpha}{2}}(n)$	$\frac{\sum_{i=1}^n(x_i-\mu)^2}{\sigma_0^2}\leq\chi^2_{\frac{\alpha}{2}}(n-1)$或 $\frac{\sum_{i=1}^n(x_i-\mu)^2}{\sigma_0^2}\geq\chi^2_{1-\frac{\alpha}{2}}(n-1)$
II	$\sigma^2=\sigma_0^2$	$\sigma^2>\sigma_0^2$	$\frac{\sum_{i=1}^n(x_i-\mu)^2}{\sigma_0^2}\geq\chi^2_{1-\alpha}(n)$	$\frac{\sum_{i=1}^n(x_i-\mu)^2}{\sigma_0^2}\geq\chi^2_{1-\alpha}(n-1)$
III	$\sigma^2\leq\sigma_0^2$	$\sigma^2>\sigma_0^2$		
IV	$\sigma^2=\sigma_0^2$	$\sigma^2<\sigma_0^2$	$\frac{\sum_{i=1}^n(x_i-\mu)^2}{\sigma_0^2}\leq\chi^2_{\alpha}(n)$	$\frac{\sum_{i=1}^n(x_i-\mu)^2}{\sigma_0^2}\leq\chi^2_{\alpha}(n-1)$
V	$\sigma^2\geq\sigma_0^2$	$\sigma^2<\sigma_0^2$		

Table: 单个正态总体均值σ^2的假设检验的拒绝域（显著性水平为α）

两个独立正态总体均值差的假设检验（双边）

设总体$X \sim N(\mu_1, \sigma_1^2)$, 另有与$X$独立的总体$Y \sim N(\mu_2, \sigma_2^2)$.

$$H_0: \mu_1 - \mu_2 = \delta,\ H_1: \mu_1 - \mu_2 \neq \delta$$

(1) σ_1^2, σ_2^2已知

选择检验统计量:

$$U = \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{\sigma_1^2/m + \sigma_2^2/n}} = \frac{\bar{X} - \bar{Y} - \delta}{\sqrt{\sigma_1^2/m + \sigma_2^2/n}} \sim N(0, 1)$$

若$|U| \geq u_{1-\alpha/2}$拒绝H_0，否则接受H_0.

(2) σ_1^2, σ_2^2未知，已知$\sigma_1^2 = \sigma_2^2$

选择检验统计量:

$$T = \frac{\bar{X} - \bar{Y} - \delta}{S_w\sqrt{1/m + 1/n}} \sim t(m + n - 2)$$

若$|T| \geq t_{1-\alpha/2}(n + m - 2)$拒绝$H_0$，否则接受$H_0$.

例

例: 已知OPPO手机充电五分钟通话时间$X \sim N(\mu_1, \sigma_1^2)$, VIVO手机充电五分钟通话时间$Y \sim N(\mu_2, \sigma_2^2)$. 随机抽取6部手机测试通话时间（单位：小时）为

OPPO	1.6	2.1	1.9	1.8	2.2	2.1
VIVO	1.8	2.2	1.5	1.4	2.0	1.7

问这两种手机通话时间平均水平无显著差异（显著性水平$\alpha = 0.05$）：

- 已知$\sigma_1^2 = 0.06$, $\sigma_2^2 = 0.08$.
- 已知$\sigma_1^2 = \sigma_2^2$.

解: 假设检验:

$$H_0: \mu_1 = \mu_2,\ H_1: \mu_1 \neq \mu_2$$

(1)检验统计量的观测值为$\frac{\bar{x}-\bar{y}}{\sqrt{0.06/6+0.08/6}} = 1.18 < u_{0.975} = 1.96$. 接受$H_0$

(2)检验统计量的观测值为$\frac{\bar{x}-\bar{y}}{s_w\sqrt{1/6+1/6}} = 1.24 < t_{0.975}(10) = 2.23$. 接受$H_0$

两个独立正态总体均值差的假设检验对比

序号	H_0	H_1	σ_1^2, σ_2^2已知	σ_1^2, σ_2^2未知
I	$\mu_1 - \mu_2 = \delta$	$\mu_1 - \mu_2 \neq \delta$	$\frac{\lvert\bar{x}-\bar{y}-\delta\rvert}{\sqrt{\frac{\sigma_1^2}{m}+\frac{\sigma_2^2}{n}}} \geq u_{1-\frac{\alpha}{2}}$	$\frac{\lvert\bar{x}-\bar{y}-\delta\rvert}{s_w\sqrt{\frac{1}{m}+\frac{1}{n}}} \geq t_{1-\frac{\alpha}{2}}$
II	$\mu_1 - \mu_2 = \delta$	$\mu_1 - \mu_2 > \delta$	$\frac{\bar{x}-\bar{y}-\delta}{\sqrt{\frac{\sigma_1^2}{m}+\frac{\sigma_2^2}{n}}} \geq u_{1-\frac{\alpha}{2}}$	$\frac{\bar{x}-\bar{y}-\delta}{s_w\sqrt{\frac{1}{m}+\frac{1}{n}}} \geq t_{1-\alpha}$
III	$\mu_1 - \mu_2 \leq \delta$	$\mu_1 - \mu_2 > \delta$		
IV	$\mu_1 - \mu_2 = \delta$	$\mu_1 - \mu_2 < \delta$	$\frac{\bar{x}-\bar{y}-\delta}{\sqrt{\frac{\sigma_1^2}{m}+\frac{\sigma_2^2}{n}}} \leq u_{\alpha}$	$\frac{\bar{x}-\bar{y}-\delta}{s_w\sqrt{\frac{1}{m}+\frac{1}{n}}} \leq t_{\alpha}$
V	$\mu_1 - \mu_2 \geq \delta$	$\mu_1 - \mu_2 < \delta$		

Table: 单个正态总体均值μ的假设检验的拒绝域（显著性水平为α），这里t分布的分位数对应的自由度为$m+n-2$. $s_w = \sqrt{(ms_{1m}^2 + ns_{2n}^2)/(n+m-2)}$

两个独立正态总体方差的假设检验（双边）

设总体$X \sim N(\mu_1, \sigma_1^2)$, 另有与$X$独立的总体$Y \sim N(\mu_2, \sigma_2^2)$.

$$H_0 : \sigma_1^2 = \sigma_2^2, \ H_1 : \sigma_1^2 \neq \sigma_2^2$$

(1) μ_1, μ_2已知

选择检验统计量:

$$F = \frac{\frac{1}{m}\sum_{i=1}^{m}(X_i - \mu_1)^2}{\frac{1}{n}\sum_{i=1}^{n}(Y_i - \mu_2)^2}\frac{\sigma_2^2}{\sigma_1^2} = \frac{\frac{1}{m}\sum_{i=1}^{m}(X_i - \mu_1)^2}{\frac{1}{n}\sum_{i=1}^{n}(Y_i - \mu_2)^2} \sim F(m, n)$$

若$F \geq F_{1-\alpha/2}(m, n)$或者$F < F_{\alpha/2}(m, n)$拒绝$H_0$，否则接受$H_0$.

(2) μ_1, μ_2未知

选择检验统计量:

$$F = \frac{\frac{1}{m-1}\sum_{i=1}^{m}(X_i - \bar{X})^2}{\frac{1}{n-1}\sum_{i=1}^{n}(Y_i - \bar{Y})^2} = \frac{S_{1m}^{*2}}{S_{2n}^{*2}} \sim F(m-1, n-1)$$

若$F \geq F_{1-\alpha/2}(m-1, n-1)$或者$F < F_{\alpha/2}(m-1, n-1)$拒绝$H_0$，否则接受$H_0$.

这种检验称为F检验法。

例

例： 已知OPPO手机充电五分钟通话时间$X \sim N(\mu_1, \sigma_1^2)$, VIVO手机充电五分钟通话时间$Y \sim N(\mu_2, \sigma_2^2)$. 随机抽取6部手机测试通话时间（单位：小时）为

OPPO	1.6	2.1	1.9	1.8	2.2	2.1
VIVO	1.8	2.2	1.5	1.4	2.0	1.7

求检验两种手机通话时间方差是否相等（显著性水平$\alpha = 0.1$）：

- 已知$\mu_1 = 2$, $\mu_2 = 1.8$.
- μ_1, μ_2未知.

解： 假设检验:

$$H_0 : \sigma_1^2 = \sigma_2^2,\ H_1 : \sigma_1^2 \neq \sigma_2^2$$

(1)检验统计量的观测值为$\frac{\frac{1}{m}\sum_{i=1}^{m}(X_i-2)^2}{\frac{1}{n}\sum_{i=1}^{n}(Y_i-1.8)^2} = 0.59 \in (0.23, 4.28)$. 接受$H_0$

(2)检验统计量的观测值为$\frac{\frac{1}{m}\sum_{i=1}^{m}(X_i-\bar{X})^2}{\frac{1}{n}\sum_{i=1}^{n}(Y_i-\bar{Y})^2} = 0.43 \in (0.2, 5.05)$. 接受$H_0$

两个独立正态总体方差的假设检验对比

序号	H_0	H_1	μ已知	μ未知
I	$\sigma_1^2=\sigma_2^2$	$\sigma_1^2\neq\sigma_2^2$	$\frac{\sum_{i=1}^m(x_i-\mu_1)^2/m}{\sum_{i=1}^n(y_i-\mu_2)^2/n}\leq F_{\frac{\alpha}{2}}$ 或 $\frac{\sum_{i=1}^m(x_i-\mu_1)^2/m}{\sum_{i=1}^n(y_i-\mu_2)^2/n}\geq F_{1-\frac{\alpha}{2}}$	$\frac{S_{1m}^{*2}}{S_{2n}^{*2}}\leq F_{\frac{\alpha}{2}}$ 或 $\frac{S_{1m}^{*2}}{S_{2n}^{*2}}\geq F_{1-\frac{\alpha}{2}}$
II	$\sigma_1^2=\sigma_2^2$	$\sigma_1^2>\sigma_2^2$	$\frac{\sum_{i=1}^m(x_i-\mu_1)^2/m}{\sum_{i=1}^n(y_i-\mu_2)^2/n}\geq F_{1-\alpha}$	$\frac{S_{1m}^{*2}}{S_{2n}^{*2}}\geq F_{1-\alpha}$
III	$\sigma_1^2\leq\sigma_2^2$	$\sigma_1^2>\sigma_2^2$		
IV	$\sigma_1^2=\sigma_2^2$	$\sigma_1^2<\sigma_2^2$	$\frac{\sum_{i=1}^m(x_i-\mu_1)^2/m}{\sum_{i=1}^n(y_i-\mu_2)^2/n}\leq F_{\alpha}$	$\frac{S_{1m}^{*2}}{S_{2n}^{*2}}\leq F_{\alpha}$
V	$\sigma_1^2\geq\sigma_2^2$	$\sigma_1^2<\sigma_2^2$		

Table: 两个独立正态总体方差假设检验的拒绝域（显著性水平为α），倒数第二列F分布为$F(m,n)$, 最后一列F分布为$F(m-1,n-1)$

目录

非正态总体均值的假设检验

与非正态总体的区间估计思想类型，当样本量充分大时，以下检验统计量可近似看出标准正态分布：

$$\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}=\frac{\sqrt{n}(\bar{X}-\mu)}{\sigma}\sim N(0,1).$$

如果σ未知，可以用样本标准差S_n替代σ.

非参假设检验

前面讨论的关于参数的假设检验，都是事先假定总体的分布类型为已知的。但有些时候，事先并不知道总体服从什么分布，需要对总体的分布类型进行推断。这类检验称为非参数假设检验。

$$H_0 : X \sim F_0(\cdot;\boldsymbol{\theta}),\ H_1 : X \nsim F_0(\cdot;\boldsymbol{\theta})$$

例：比如检验某总体是否为正态总体

常用的方法：χ^2拟合优度检验——构造皮尔逊(Pearson)统计量

自学

最后一次作业

根据期中考试成绩数据回答调查问卷中的问题，6月22日（周五）前提交！

数据链接：https://www.jianguoyun.com/p/DenoTWkQpvLJBhj-y1o

问卷连接：https://ks.wjx.top/jq/24951942.aspx