概 率 统 计

第9章

假设检验

§9.1 假设检验的基本概念

我们将讨论不同于参数估计的另一类重要的统计推断问题. 这就是根据样本的信息检验关于总体的某个假设是否正确.

这类问题称作假设检验问题 .

△ 何为假设检验?

假设检验是指施加于一个或多个总体的概率分布或参数的假设. 所作假设可以是正确的, 也可以是错误的.

为判断所作的假设是否正确，从总体中抽取样本, 根据样本的取值, 按一定原则进行检验，然后作出接受或拒绝所作假设的决定.

若对参数一无所知 → 用参数估计的方法处理

若对参数有所了解 → 但有怀疑猜测需要证实之时 → 用假设检验的方法来处理

△ 假设检验的内容

假设检验

- 参数检验

  总体分布已知时检验关于未知参数的某个假设

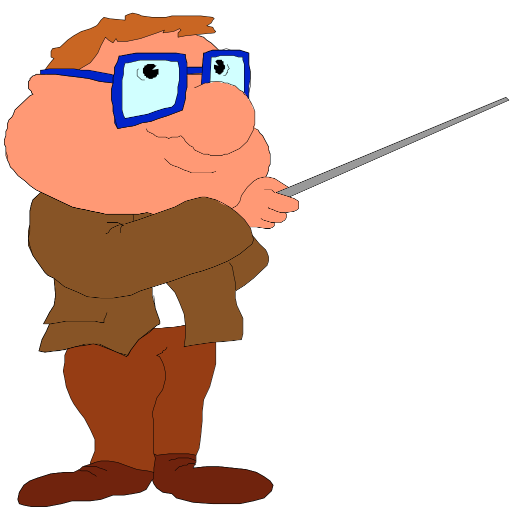
- 非参数检验

  总体分布未知时对分布类型的假设检验问题

△ 假设检验的理论依据

假设检验所以可行, 其理论背景为实际推断原理, 即"小概率原理"

人们在实践中普遍采用的一个原则:

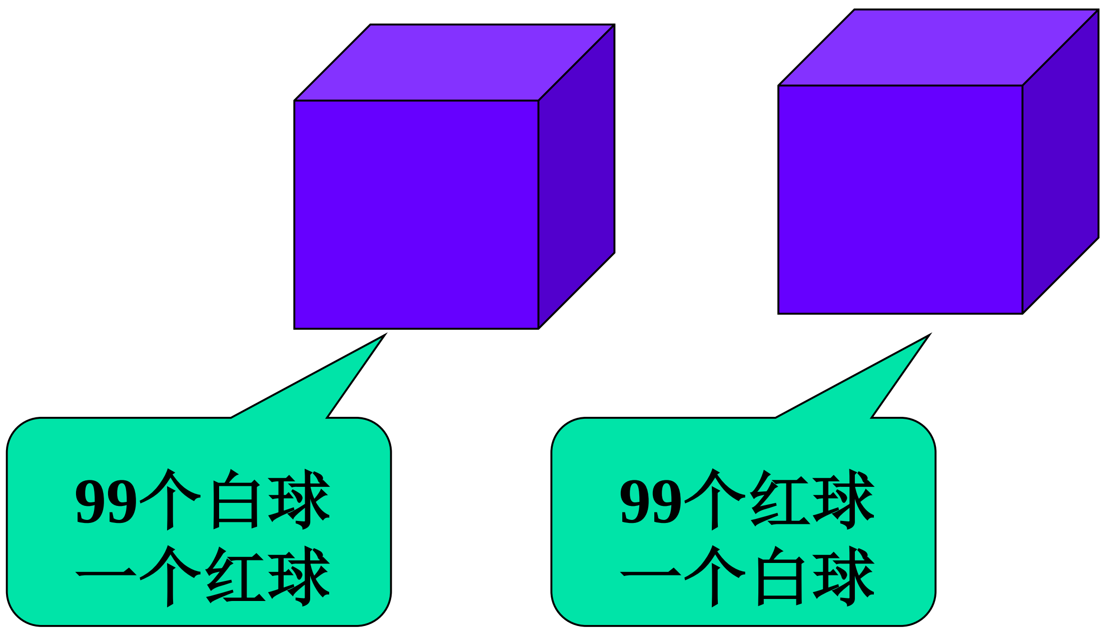


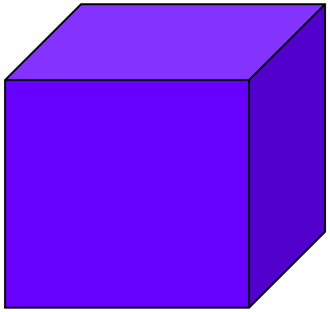
小概率事件在一次试验中基本上不会发生 .

小概率事件在一次试验中基本上不会发生.

下面我们用一例说明这个原则.

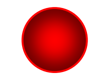
这里有两个盒子，各装有100个球.



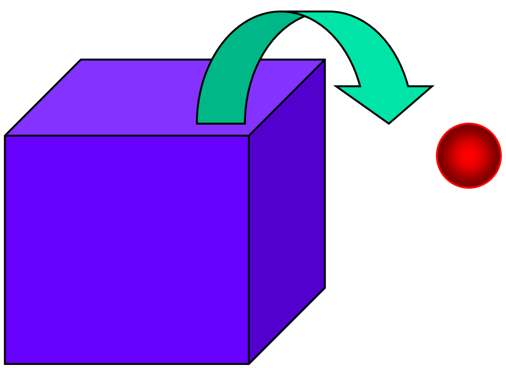


现从两盒中随机取出一个盒子，问这个盒子里是白球99个还是红球99个？

我们不妨先假设：这个盒子里有99个白球.

现在我们从中随机摸出一个球，发现是 

此时你如何判断这个假设是否成立呢？



假设其中真有99个白球，
摸出红球的概率只有**1/100**，
这是小概率事件.

小概率事件在一次试验中竟然发生了，
不能不使人怀疑所作的假设.

例子中所使用的推理方法，可以称为

带概率性质的反证法

不妨称为概率反证法.

它不同于一般的反证法

一般的反证法要求在原假设成立的条件下导出的结论是绝对成立的，如果事实与之矛盾，则完全绝对地否定原假设.

概率反证法的逻辑是：

如果小概率事件在一次试验中居然发生，我们就以很大的把握否定原假设.

在假设检验中，我们称这个小概率为显著性水平，用 α 表示.

α 的选择要根据实际情况而定.

常取 $\alpha = 0.1, \alpha = 0.01, \alpha = 0.05.$

假设检验步骤

例 某工厂生产的一种螺钉，标准要求长度是32.5毫米. 实际生产的产品，其长度X假定服从正态分布$N(\mu,\sigma^2)$, σ^2未知，现从该厂生产的一批产品中抽取6件, 得尺寸数据如下:

32.56, 29.66, 31.64, 30.00, 31.87, 31.03

问这批产品是否合格?

分析：这批产品(螺钉长度)的全体组成问题的总体X. 现在要检验$E(X)$是否为32.5.

已知 $X \sim N(\mu, \sigma^2)$, σ^2 未知.

第一步：提出原假设和备择假设

$$H_0 : \mu = 32.5 \Leftrightarrow H_1 : \mu \neq 32.5$$

实际工作中，往往把不轻易否定的命题作为原假设.

第二步：取一检验统计量，在H_0成立下求出它的分布

$$t = \frac{\overline{X} - 32.5}{S/\sqrt{6}} \sim t(5)$$

能衡量差异大小且分布已知

第三步：

对给定的显著性水平α =**0.01**，查表确定临界值 $t_{\alpha/2}(5)=t_{0.005}(5)=4.0322$，使

$$P\{|t|>t_{\alpha/2}(5)\}=\alpha$$

即“$|t|>t_{\alpha/2}(5)$”是一个小概率事件.

小概率事件在一次试验中基本上不会发生.

得否定域 **W: $|t|>4.0322$**

拒绝域　W: $|t|>4.0322$

第四步:

将样本值代入算出统计量 t 的实测值,

$$|t|=2.997<4.0322$$

没有落入拒绝域

故接受H_0.

这并不意味着H_0一定对，只是差异还不够显著,不足以否定H_0.

所以假设检验又叫 “显著性检验”

假设检验步骤

(1) 建立假设

(2) 在H_0为真时，选择统计量

(3) 确定拒绝域

(4) 作出判断

§9.2-9.5 正态总体的参数检验

1. 一个正态总体

9.2 关于μ 的检验

给定显著性水平α与样本值$(x_1, x_2, \ldots, x_n)$

设 $X \sim N(\mu, \sigma^2)$，σ^2 已知，需检验：

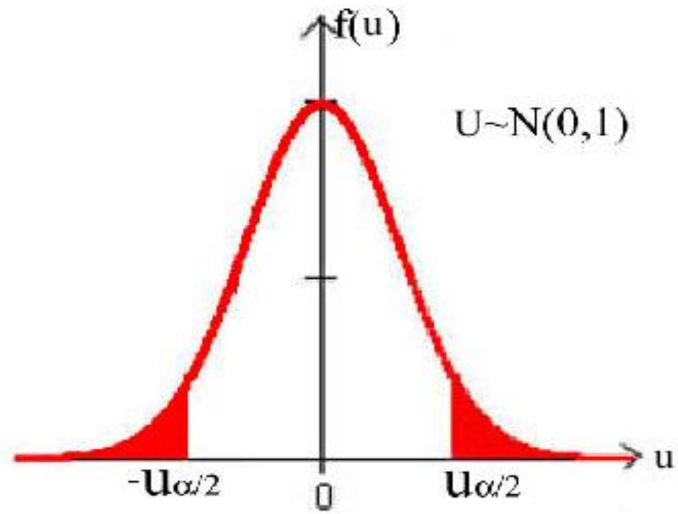
$$H_0: \mu = \mu_0; \quad H_1: \mu \neq \mu_0$$

构造统计量 $$U = \frac{\overline{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}} \sim N(0,1)$$

$$P\{|U|> z_{\alpha/2}\}=\alpha$$

也就是说,“ $|U|> z_{\alpha/2}$ ”是一个小概率事件.

故我们可以取拒绝域为:



$$W\text{：}\quad |U|> z_{\alpha/2}$$

如果由样本值算得该统计量的实测值落入区域W，则拒绝H_0；否则，不能拒绝H_0.

U 检验法

U 检验法 (σ^2 已知)

原假设 H_0	备择假设 H_1	检验统计量及其 H_0为真时的分布	拒绝域
$\mu = \mu_0$	$\mu \neq \mu_0$	$U = \dfrac{\bar{X} - \mu_0}{\sigma / \sqrt{n}} \sim N(0,1)$	$\lvert U \rvert \geq z_{\frac{\alpha}{2}}$
$\mu \geq \mu_0$	$\mu < \mu_0$		$U \leq -z_{\alpha}$
$\mu \leq \mu_0$	$\mu > \mu_0$		$U \geq z_{\alpha}$

T 检验法 (σ^2 未知)

原假设 H_0	备择假设 H_1	检验统计量及其 H_0为真时的分布	拒绝域
$\mu = \mu_0$	$\mu \neq \mu_0$	$T = \dfrac{\overline{X} - \mu_0}{\dfrac{S}{\sqrt{n}}} \sim t(n-1)$	$\lvert T \rvert \geq t_{\frac{\alpha}{2}}$
$\mu \geq \mu_0$	$\mu < \mu_0$		$T \leq -t_{\alpha}$
$\mu \leq \mu_0$	$\mu > \mu_0$		$T \geq t_{\alpha}$

例 化肥厂用自动包装机装化肥，规定每袋标准重量为100，设每袋重量X服从正态分布且标准差$\sigma=0.9$不变.某天抽取9袋，测得重量为

99.3, 98.7, 101.2, 100.5, 98.3, 99.7, 102.6, 100.5, 105.1

问机器工作是否正常$(\alpha=0.05)$?

H_0: $\mu=100$; H_1: $\mu\neq100$

U 检验法

构造统计量 $U=\dfrac{\overline{X}-100}{\dfrac{\sigma}{\sqrt{n}}}\sim N(0,1)$

$|U|>z_{\alpha/2}$

$U=\dfrac{100.66-100}{\dfrac{0.9}{\sqrt{9}}}=2.2 > z_{0.025}=1.96$

拒绝

例　金属锰的熔化点 X 服从正态分布，测量5次得熔化点为：1267， 1271， 1256， 1265， 1254

能否认为熔化点为 $1260(\alpha = 0.05)$?

$$X \sim N(\mu, \sigma^2), \ \boldsymbol{\sigma}^2 \text{ 未知.}$$

$$H_0: \mu = 1260; \quad H_1: \mu \neq 1260$$

接受

$$|T| > t_{\frac{\alpha}{2}}$$

$$t = \frac{\overline{X} - 1260}{S/\sqrt{5}} = 0.796 < t_{0.025}(4) = 2.776$$

9.3 关于 σ^2 的检验

χ^2检验法(μ已知)

原假设 H_0	备择假设 H_1	检验统计量及其在 H_0为真时的分布	拒绝域
$\sigma^2=\sigma_0^2$	$\sigma^2\neq\sigma_0^2$	$\chi^2=\dfrac{\sum_{i=1}^{n}(X_i-\mu)^2}{\sigma_0^2} \sim \chi^2(n)$	$\chi^2 \le \chi^2_{1-\frac{\alpha}{2}}(n)$ 或 $\chi^2 \ge \chi^2_{\frac{\alpha}{2}}(n)$
$\sigma^2\geq\sigma_0^2$	$\sigma^2<\sigma_0^2$		$\chi^2 \le \chi^2_{1-\alpha}(n)$
$\sigma^2\leq\sigma_0^2$	$\sigma^2>\sigma_0^2$		$\chi^2 \ge \chi^2_{\alpha}(n)$

χ^2 **检验法**(μ 未知)

原假设 H_0	备择假设 H_1	检验统计量及其在 H_0 为真时的分布	拒绝域
$\sigma^2=\sigma_0^2$	$\sigma^2\neq\sigma_0^2$	$\chi^2=\dfrac{(n-1)S^2}{\sigma_0^2}\sim\chi^2(n-1)$	$\chi^2\leq\chi^2_{1-\frac{\alpha}{2}}(n-1)$ 或 $\chi^2\geq\chi^2_{\frac{\alpha}{2}}(n-1)$
$\sigma^2\geq\sigma_0^2$	$\sigma^2<\sigma_0^2$		$\chi^2\leq\chi^2_{1-\alpha}(n-1)$
$\sigma^2\leq\sigma_0^2$	$\sigma^2>\sigma_0^2$		$\chi^2\geq\chi^2_{\alpha}(n-1)$

例　已知维尼纶纤度 X 在正常情况下服从正态分布 $N(\mu, 0.048^2)$. 现在测了5 根纤维，其纤度分别为:

1.44，　1.36，　1.40，　1.55，　1.32，

问：产品的精度是否有显著的变化 $(\alpha = 0.05)$?

$$H_0 : \sigma^2 = 0.048^2; \quad H_1 : \sigma^2 \neq 0.048^2.$$

拒绝

$\chi^2 \le \chi^2_{1-\frac{\alpha}{2}}(n-1)$ 或 $\chi^2 \ge \chi^2_{\frac{\alpha}{2}}(n-1)$

$$\chi^2 = \frac{(n-1)S^2}{\sigma_0^2} = \frac{0.03112}{0.048^2} = 13.51 > \chi^2_{0.025}(4) = 11.143$$

$$\chi^2_{0.975}(4) = 0.485$$

2. 单侧检验与双侧检验

前面各例的检验，拒绝域取在两侧，

——称为双侧检验.

单侧检验—拒绝域取在左侧或右侧.

下面看一个单侧检验的例子:

例 某织物强力指标X的均值 μ_0=21公斤. 改进工艺后生产一批织物，今从中取30件，测得 $\overline{X}$=21.55公斤. 假设强力指标服从正态分布 $N(\mu,\sigma^2)$, 且已知σ=1.2公斤， 问在显著性水平 α=0.01下，新生产织物比过去的织物强力是否有提高?

解:提出假设: $H_0:\mu\le 21;\quad H_1:\mu>21$

$$H_0:\mu=21;\quad H_1:\mu>21$$

取统计量$U=\dfrac{\overline{X}-21}{\sigma/\sqrt{n}}\sim N(0,1)$

$\{U>z_{0.01}\}$是一小概率事件

否定域为 W： $U>z_{0.01}$ =2.33

代入 σ =1.2, n=30，并由样本值计算得统计量U的实测值

U=2.51>2.33

落入否定域

故拒绝原假设H_0.

这时可能犯第一类错误，犯错误的概率不超过0.01.

$P216.9$ 已知某溶液中的水分 X 服从正态分布

10个测定值给出 $\overline{X}=0.452\%$，$S=0.037\%$,

检验假设 $(\alpha=0.05)$:

书上答案 ？？？

$$H_0:\sigma\geq 0.04\%;\quad H_1:\sigma<0.04\%.$$

拒绝域 $\chi^2<\chi^2_{1-\alpha}(n-1)$

$$\chi^2=\frac{(n-1)S^2}{\sigma_0^2}=7.7 > \chi^2_{1-\alpha}(9)=\chi^2_{0.95}(9)=3.325$$

故接受H_0 不符合

$P216.9$　已知某溶液中的水分 X 服从正态分布

10个测定值给出 $\overline{X}=0.452\%$，$S=0.037\%$,

检验假设 $(\alpha=0.05)$：

$$H_0:\sigma\le 0.04\%;\quad H_1:\sigma>0.04\%.$$

拒绝域　$\chi^2>\chi^2_\alpha(n-1)$

$$\chi^2=\frac{(n-1)S^2}{\sigma_0^2}=7.7<\chi^2_\alpha(9)=\chi^2_{0.05}(9)=16.919$$

故接受H_0　　**符合**　　？？？

例 某厂生产小型马达,说明书上写着:在正常负载下平均消耗电流不超过0.8 安培.

随机测试16台马达, 平均消耗电流为0.92安培， 标准差为0.32安培.

设马达所消耗的电流 服从正态分布,取显著性水平为$\alpha = 0.05$, 问根据此样本,能否否定厂方的断言?

$H_0: \mu \le 0.8$；　　$H_1: \mu > 0.8$

假设

$H_0: \mu \ge 0.8$；　　$H_1: \mu < 0.8$

解一 $H_0: \mu \le 0.8$； $H_1: \mu > 0.8$

σ 未知, 选检验统计量: $T = \dfrac{\bar{X} - \mu}{S/\sqrt{16}} \sim T(15)$

拒绝域为 $T = \dfrac{\bar{x} - 0.8}{s/\sqrt{n}} > 1.753 = t_{0.05}(15)$

将 $\bar{x} = 0.92$, $s = 0.32$，代入得

$T = 1.5 < 1.735$, 落在拒绝域外

故接受原假设 H_0, 即不能否定厂方断言.

解二 $H_0: \mu \geq 0.8$； $H_1: \mu < 0.8$

选用统计量 $T = \dfrac{\bar{X} - \mu}{S/\sqrt{16}} \sim T(15)$

拒绝域 $T = \dfrac{\bar{x} - 0.8}{s/\sqrt{n}} < -1.753 = -t_{0.05}(15)$

现 $T = 1.5 > -1.735,$ 落在拒绝域外

故接受原假设，即否定厂方断言.

上述两种解法得到不同的结论

第一种假设是不轻易否定厂方的结论；

第二种假设是不轻易相信厂方的结论.

由例可见：对问题的提法不同(把哪个假设作为原假设)，统计检验的结果也会不同.

为何用假设检验处理同一问题会得到截然相反的结果?

这里固然有把哪个假设作为原假设从而引起检验结果不同这一原因；除此外还有一个根本的原因，即样本容量不够大．

若样本容量足够大，则不论把哪个假设作为原假设所得检验结果基本上应该是一样的．否则假设检验便无意义！

由于假设检验是控制犯第一类错误的概率, 使得拒绝原假设 H_0 的决策变得比较慎重, 也就是 H_0 得到特别的保护. 因而, 通常把有把握的, 经验的结论作为原假设, 或者尽量使后果严重的错误成为第一类错误.

9.4–9.5 两个正态总体

设 $X \sim N(\mu_1, \sigma_1^{\,2}),\ Y \sim N(\mu_2, \sigma_2^{\,2})$

两样本 X, Y 相互独立,

样本 $(X_1, X_2, \ldots, X_{n1}), (Y_1, Y_2, \ldots, Y_{n2})$

显著性水平α

（1）关于均值差 $\mu_1 - \mu_2$ 的检验

原假设 H_0	备择假设 H_1	检验统计量及其在 H_0为真时的分布	拒绝域
$\mu_1 - \mu_2 = \delta$	$\mu_1 - \mu_2 \neq \delta$	$U = \dfrac{\bar{X} - \bar{Y} - \delta}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} \sim N(0,1)$ （ σ_1^2， σ_2^2 已知）	$\lvert U \rvert \geq z_{\frac{\alpha}{2}}$
$\mu_1 - \mu_2 \geq \delta$	$\mu_1 - \mu_2 < \delta$		$U \leq -z_\alpha$
$\mu_1 - \mu_2 \leq \delta$	$\mu_1 - \mu_2 > \delta$		$U \geq z_\alpha$

原假设 H_0	备择假设 H_1	检验统计量及其在 H_0为真时的分布	拒绝域
$\mu_1-\mu_2=\delta$	$\mu_1-\mu_2\neq\delta$	$\left(\begin{array}{c}\sigma_1^2,\sigma_2^2\text{未知}\\ \text{但 }\sigma_1^2=\sigma_2^2\end{array}\right)$ $t=\dfrac{\bar{X}-\bar{Y}-\delta}{S_w\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}$ $\sim t(n_1+n_2-2)$	$\lvert t\rvert>t_{\frac{\alpha}{2}}$
$\mu_1-\mu_2\geq\delta$	$\mu_1-\mu_2<\delta$		$t<-t_\alpha$
$\mu_1-\mu_2\leq\delta$	$\mu_1-\mu_2>\delta$		$t>t_\alpha$

其中 $S_w=\sqrt{\dfrac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}}$

(2) **两正态总体方差比 σ_1^2/σ_2^2 的检验** **F 检验法**

原假设 H_0	备择假设 H_1	检验统计量及其在 H_0 为真时的分布	拒绝域
$\sigma_1^2=\sigma_2^2$	$\sigma_1^2\neq\sigma_2^2$	$F=\dfrac{\sum_{i=1}^{n_1}(X_i-\mu_1)^2/n_1}{\sum_{i=1}^{n_2}(Y_i-\mu_2)^2/n_2}\sim F(n_1,n_2)$ μ_1,μ_2均已知	$F<F_{1-\frac{\alpha}{2}}(n_1,n_2)$ 或 $F>F_{\frac{\alpha}{2}}(n_1,n_2)$
$\sigma_1^2\geq\sigma_2^2$	$\sigma_1^2<\sigma_2^2$		$F<F_{1-\alpha}(n_1,n_2)$
$\sigma_1^2\leq\sigma_2^2$	$\sigma_1^2>\sigma_2^2$		$F>F_{\alpha}(n_1,n_2)$

两正态总体方差比σ_1^2/σ_2^2的检验　F 检验法

原假设 H_0	备择假设 H_1	检验统计量及其在H_0为真时的分布	拒绝域
$\sigma_1^2 = \sigma_2^2$	$\sigma_1^2 \neq \sigma_2^2$	$F = \dfrac{S_1^2}{S_2^2} \sim$	$F < F_{1-\frac{\alpha}{2}}(n_1-1, n_2-1)$ 或 $F > F_{\frac{\alpha}{2}}(n_1-1, n_2-1)$
$\sigma_1^2 \geq \sigma_2^2$	$\sigma_1^2 < \sigma_2^2$	$F(n_1-1, n_2-1)$	$F < F_{1-\alpha}(n_1-1, n_2-1)$
$\sigma_1^2 \leq \sigma_2^2$	$\sigma_1^2 > \sigma_2^2$	μ_1, μ_2 均未知	$F > F_{\alpha}(n_1-1, n_2-1)$

$P143.25.$ 货车有A,B两条行车路线，行车时间分别服从$N(\mu_i,\sigma_i^2)$，$i=1,2.$ 每条路各跑50次，在A线上$\overline{X}=95$，$S_x=20$；在B线上$\overline{Y}=76$，$S_y=15.$ 问方差是否一样，均值是否一样?$(\alpha=0.05)$

第一阶段 $H_0:\sigma_1^2=\sigma_2^2;\quad H_1:\sigma_1^2\neq\sigma_2^2$

$$F=\frac{S_x^2}{S_y^2}\sim F(49,49)\quad F=\frac{S_x^2}{S_y^2}=1.78$$

$$F_{0.975}(49,49)<F<F_{0.025}(49,49)$$

故接受H_0

第二阶段

$$H_0: \mu_1 = \mu_2 \;; \qquad H_1: \mu_1 \neq \mu_2$$

$$取统计量 T = \frac{\overline{X} - \overline{Y}}{\sqrt{\frac{1}{n_1} + \frac{1}{n_2}} S_w} \sim t(n_1 + n_2 - 2)$$

$$|T| \geq t_{0.025}(98)$$

拒绝H_0

例 为比较两台自动机床的精度，分别取容量为**11**和**9**的两个样本，测量某个指标的尺寸(假定服从正态分布)，得到下列结果：

车床甲：**6.2, 5.7, 6.0, 6.3, 6.5, 6.0, 5.7, 5.8, 6.0, 5.8, 6.0;**

车床乙：**5.6, 5.7, 5.9, 5.5, 5.6, 6.0, 5.8, 5.5, 5.7.**

在 α =**0.05**时， 问这两台机床是否有同样的精度?

解:设两台自动机床的方差分别为$\boldsymbol{\sigma}_1^2,\boldsymbol{\sigma}_2^2$,在$\alpha$=0. 05下检验假设:

$$H_0:\sigma_1^2=\sigma_2^2;\quad H_1:\sigma_1^2\neq\sigma_2^2$$

取统计量 $F=\dfrac{S_1^2}{S_2^2}\sim F(10,8)$

其中$\boldsymbol{S}_1^2,\boldsymbol{S}_2^2$为两样本的样本方差

否定域为 $\boldsymbol{W}$: $F\leq F_{1-\alpha/2}(10,8)$ 或

$$F\geq F_{\alpha/2}(10,8)$$

由样本值可计算得$\boldsymbol{F}$的实测值为:

$$\boldsymbol{F=2.13}$$

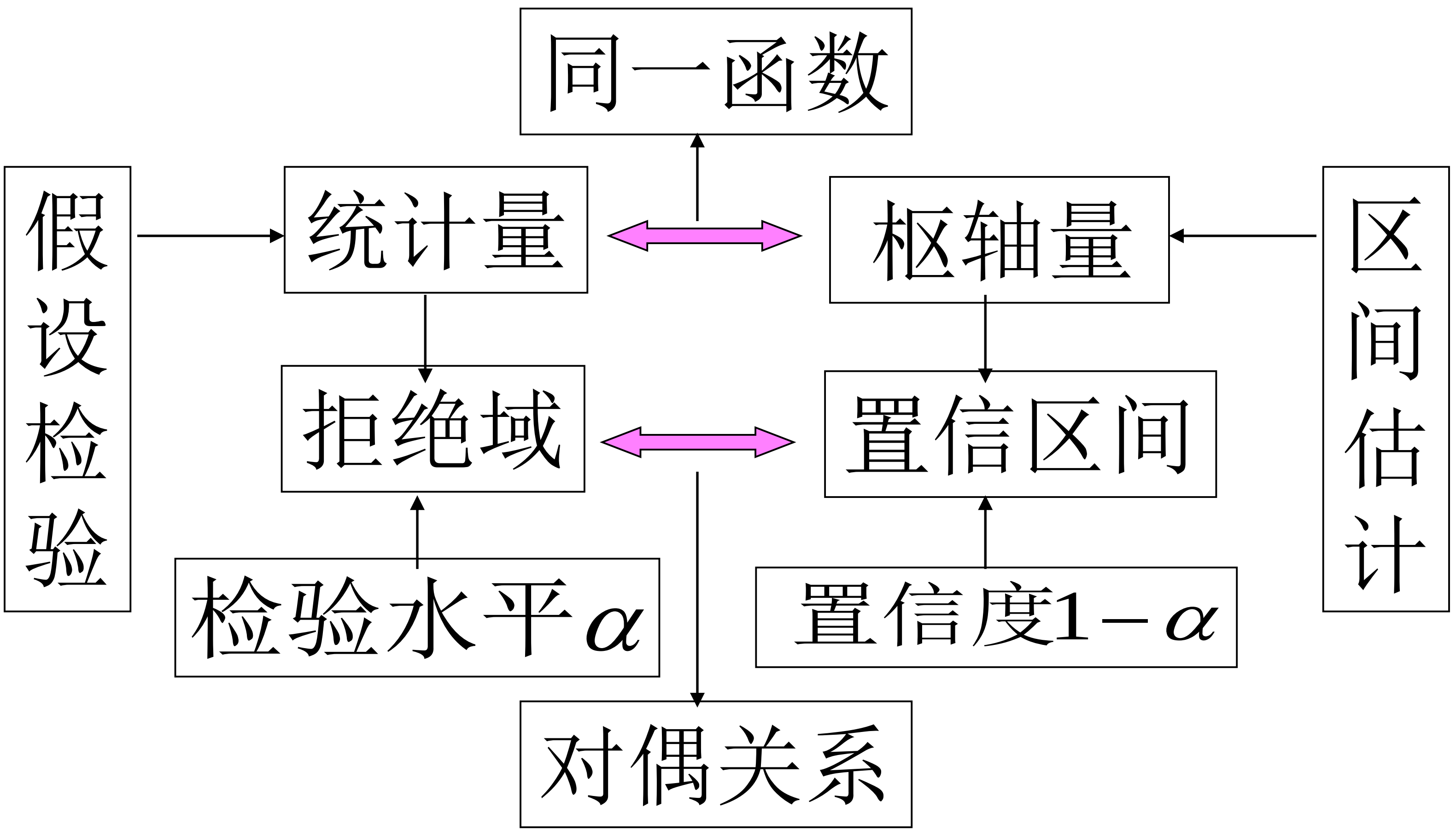
查表得 $F_{\alpha/2}(10,8)=F_{0.025}(10,8)=4.3$

$$F_{1-\alpha/2}(10,8)=F_{0.975}(10,8)=1/F_{0.025}(8,10)$$

$$=1/3.85=0.26$$

由于 $\mathbf{0.26<2.13<4.3}$， 故接受$\boldsymbol{H_0}$.

假设检验与区间估计的联系



假设检验与置信区间对照

原假设 H_0	备择假设 H_1	检验统计量及其在 H_0为真时的分布	接受域
$\mu=\mu_0$	$\mu \neq \mu_0$	$U=\dfrac{\overline{X}-\mu_0}{\sigma/\sqrt{n}} \sim N(0,1)$ （σ^2 已知)	$\left\lvert \dfrac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} \right\rvert < z_{\frac{\alpha}{2}}$

待估参数		枢轴量及其分布	置信区间
μ		$U=\dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$ （σ^2 已知)	$(\bar{x}-z_{\frac{\alpha}{2}}\dfrac{\sigma}{\sqrt{n}},\bar{x}+z_{\frac{\alpha}{2}}\dfrac{\sigma}{\sqrt{n}})$

原假设 H_0	备择假设 H_1	检验统计量及其在 H_0为真时的分布	接受域
$\mu=\mu_0$	$\mu\neq\mu_0$	$T=\dfrac{\overline{X}-\mu_0}{S/\sqrt{n}}\sim t(n-1)$ （σ^2未知）	$\left\lvert\dfrac{\overline{x}-\mu_0}{s/\sqrt{n}}\right\rvert<t_{\frac{\alpha}{2}}(n-1)$
待估参数		枢轴量及其分布	置信区间
μ		$T=\dfrac{\overline{X}-\mu}{S/\sqrt{n}}\sim t(n-1)$ （σ^2未知）	$(\overline{x}-t_{\frac{\alpha}{2}}\dfrac{s}{\sqrt{n}},$ $\overline{x}+t_{\frac{\alpha}{2}}\dfrac{s}{\sqrt{n}})$

原假设 H_0	备择假设 H_1	检验统计量及其在 H_0为真时的分布	接受域
$\sigma^2=\sigma_0^2$	$\sigma^2\neq\sigma_0^2$	$\chi^2=\dfrac{(n-1)S^2}{\sigma_0^2}\sim\chi^2(n-1)$ (μ未知)	$\chi^2_{1-\frac{\alpha}{2}}<\dfrac{(n-1)S^2}{\sigma_0^2}<\chi^2_{\frac{\alpha}{2}}$

待估参数	枢轴量及其分布	置信区间
σ^2	$\chi^2=\dfrac{(n-1)S^2}{\sigma^2}\sim\chi^2(n-1)$ (μ未知)	$\left(\dfrac{(n-1)s^2}{\chi^2_{\frac{\alpha}{2}}(n-1)},\dfrac{(n-1)s^2}{\chi^2_{1-\frac{\alpha}{2}}(n-1)}\right)$

9.6 假设检验的两类错误

假设检验会不会犯错误呢？

由于作出结论的依据是　小概率原理

小概率事件在一次试验中基本上不会发生．

不是一定不发生

在给定α的前提下，接受还是拒绝原假设完全取决于样本值，因此所作检验可能导致以下两类错误的产生：

第一类错误 —— 弃真错误

第二类错误 —— 取伪错误

假设检验的两类错误

决定	实际情况	
	H_0为真	H_0不真
拒绝H_0	第一类错误	正确
接受H_0	正确	第二类错误

犯两类错误的概率:

显著性水平

$P\{$第一类错误$\}= P\{$拒绝$H_0|H_0$为真$\}=\alpha$,

$P\{$第二类错误$\}= P\{$接受$H_0|H_0$不真$\}=\beta$.

两类错误是互相关联的， 当样本容量固定时，一类错误概率的减少导致另一类错误概率的增加.

要同时降低两类错误的概率α, β，或者要在 α 不变的条件下降低 β，需要增加样本容量.

例1 某厂生产的螺钉, 按标准强度为 $68/\text{mm}^2$, 而实际生产的强度X 服$N(\mu, 3.6^2)$. 若$E(X)=\mu=68$, 则认为这批螺钉符合要求, 否则认为不符合要求. 为此提出如下假设:

$$H_0: \mu = 68 \qquad H_1: \mu \neq 68$$

现从整批螺钉中取容量为**36**的样本, 其均值为 $\bar{x} = 68.5$, 问原假设是否正确?

$$\alpha = 0.05$$

由 $\left|\dfrac{\overline{X}-68}{3.6/6}\right|>1.96 \Longrightarrow$ $\overline{X}>69.18$ 或 $\overline{X}<66.824$

即区间($-\infty$,66.824) 与 (69.18 , $+\infty$) 为检验的拒绝域

称 $\overline{X}$ 的取值区间（ 66.824 , 69.18 ）

为检验的接受域 (实际上没理由拒绝),

现 $\overline{x}=68.5$ 落入接受域, 则接受原假设

$$H_0: \mu = 68$$

例1 中，犯第一类错误的概率

$$P(\text{拒绝}\boldsymbol{H_0} \mid \boldsymbol{H_0}\text{为真})$$

$$= P(\overline{X} < 66.824 \cup \overline{X} > 69.18)$$

$$= \alpha = 0.05$$

若$\boldsymbol{H_0}$为真，则 $\overline{X} \sim N(68\ ,\ 3.6^2/36)$

所以, 拒绝 $\boldsymbol{H_0}$ 的概率为$\boldsymbol{\alpha}$,

下面计算犯第二类错误的概率 β

$$\beta = P(\text{接受}H_0 | H_0\text{不真})$$

H_0不真, 即$\mu \neq 68$, μ可能小于68, 也可能大于68, β 的大小取决于 μ 的真值的大小.

设 $\mu = 66,\ n = 36,\quad \bar{X} \sim N(66, 3.6^2/36)$

$$\beta_{\mu=66} = P(66.82 \le \bar{X} \le 69.18 \mid \mu = 66)$$

$$= \Phi\left(\frac{69.18-66}{0.6}\right) - \Phi\left(\frac{66.82-66}{0.6}\right)$$

$$= \Phi(5.3) - \Phi(1.37) = 1 - 0.9147 = 0.0853$$

若 $\mu = 69$, $n = 36$, $\bar{X} \sim N(69, 3.6^2/36)$

$$\beta_{\mu=69} = P(66.82 \le \bar{X} \le 69.18 \mid \mu = 69)$$

$$= \Phi\left(\frac{69.18-69}{0.6}\right) - \Phi\left(\frac{66.82-69}{0.6}\right)$$

$$= \Phi(0.3) - \Phi(-3.63)$$

$$= 0.6179 - 0.0002 = 0.6177$$

取伪的概率较大.

现增大样本容量, 取$n = 64,\ \mu = 66$, 则

$$\bar{X} \sim N(66, 3.6^2/64)$$

仍取α=0.05, 则 $c = z_{\frac{\alpha}{2}} = z_{0.025} = 1.96$

由$\left|\dfrac{\bar{X}-68}{3.6/8}\right| > 1.96$ 可以确定拒绝域为

$(-\infty,\ 67.118)$ 与 $(68.882,\ +\infty)$

因此，接受域为$(67.118,\ 68.882)$

$$\beta_{\mu=66} = P\left(67.118 \le \bar{X} \le 68.882 \,\middle|\, \mu = 66 \right)$$

$$\approx \Phi\left(\frac{68.88-66}{0.45}\right) - \Phi\left(\frac{67.12-66}{0.45}\right)$$

$$= \Phi(6.4) - \Phi(2.49)$$

$$\approx 1 - 0.9936 = 0.0064 < 0.0853$$

$$\beta_{\mu=69} = P\left(67.12 \le \bar{X} \le 68.88 \,\middle|\, \mu = 69 \right)$$

$$= 0.3936 < 0.6177$$

$$(\mu \to \mu_0, \beta \to 1-\alpha)$$

9.7 非正态总体参数的假设检验

前面我们介绍了假设检验的基本思想，并讨论了当总体分布为正态时，关于其中未知参数的假设检验问题．

但有时总体服从的分布不是正态分布，要求我们直接对总体中的未知参数提出一个假设并作出检验．

非正态总体 (大样本)

P209.1

原假设 H_0	备择假设 H_1	检验统计量及其 H_0为真时的分布	拒绝域
$\mu = \mu_0$	$\mu \neq \mu_0$	$T = \dfrac{\overline{X} - \mu_0}{\dfrac{S}{\sqrt{n}}}$ 近似 $\sim N(0,1)$	$\lvert T \rvert \geq z_{\frac{\alpha}{2}}$
$\mu \geq \mu_0$	$\mu < \mu_0$		$T \leq -z_\alpha$
$\mu \leq \mu_0$	$\mu > \mu_0$		$T \geq z_\alpha$

下面介绍比率p的假设检验

在实际应用中，通常要对比率 p（如产品的不合格率等）作出判断，这种类型的总体服从二项分布$B(1, p)$，其分布列为：

$$P(X=1)=p\,,\quad P(X=0)=1-p$$

这里$(X=1)$表示事物具有某种属性，

$(X=0)$表示事物不具有某种属性。

设样本 $X_1, \ldots, X_n$ 来自两点分布 $X \sim B(1,p)$，关于参数 p 的检验有二种类型：

1）双侧检验：$H_0: p = p_0 \quad \leftrightarrow \quad H_1: p \neq p_0$

2）单侧检验：$\begin{cases} H_0: p \leq p_0 \quad \leftrightarrow \quad H_1: p > p_0 \\ H_0: p \geq p_0 \quad \leftrightarrow \quad H_1: p < p_0 \end{cases}$

大样本情况下取检验统计量:

$$U=\frac{\overline{X}-p_0}{\sqrt{p_0(1-p_0)/n}}\overset{\text{近似}}{\sim}N(0,1)$$

原假设	拒绝域
$H_0: p=p_0 \quad \leftrightarrow \quad H_1: p\neq p_0$	$W=\left\{\lvert U\rvert>z_{\alpha/2}\right\}$
$H_0: p\leq p_0 \quad \leftrightarrow \quad H_1: p>p_0$	$W=\{U>z_{\alpha}\}$
$H_0: p\geq p_0 \quad \leftrightarrow \quad H_1: p<p_0$	$W=\{U<-z_{\alpha}\}$

例 某厂生产的电冰箱，以往优等品率不超过**50%**，现技术革新后抽查了**50**台，其中**35**台为优等品，问技术革新后优等品率有无显著提高（α=0.05）？

解 $H_0: p \le 0.5 \quad \leftrightarrow \quad H_1: p > 0.5$

$$t = \frac{\bar{x} - p_0}{\sqrt{p_0(1-p_0)/n}} = \frac{0.7-0.5}{\sqrt{0.5\times(1-0.5)/50}}$$

$$= 2.828 > 1.645 = z_{0.05}$$ 故拒绝 H_0

即认为技术革新后优等品率有显著提高.

9.8　非参数假设检验

-------卡方拟合优度检验

在此之前面介绍的假设检验均是在总体分布已知情况下关于未知参数的假设检验.

然而在实际情况中经常遇到总体服从何种分布并不知道的问题，要求直接对总体的分布提出一个假设并作出检验 .

解决这类问题的工具是英国统计学家 **K.**皮尔逊在**1900**年发表的一篇文章中引进的所谓χ^2检验法.

这是一项很重要的工作，不少人把它视为近代统计学的开端.

K.皮尔逊

χ^2检验法是在总体X的分布未知时，根据来自总体的样本，检验关于总体分布的假设的一种检验方法.

使用 χ^2检验法对总体分布进行检验时,

我们先提出原假设:

H_0: 总体X的分布函数为$F(x)$

然后根据样本的经验分布和所假设的理论分布之间的吻合程度来决定是否接受原假设.

这种检验通常称作拟合优度检验，它是一种非参数检验.

在用 χ^2 检验法检验假设 H_0 时，若在 H_0 下分布类型已知，但其参数未知，这时需要先用极大似然估计法估计参数，然后作检验.

分布拟合的 χ^2 检验法的基本原理和步骤如下:

1. 将总体X的取值范围分成k个互不重迭的小区间,记作$A_1, A_2, \ldots, A_k$.

2.把落入第i个小区间A_i的样本值的个数记作m_i， 称为实测频数. 所有实测频数之和 $m_1+ m_2+ \ldots+ m_k$等于样本容量n.

3.根据所假设的理论分布,可以算出总体X的值落入每个A_i的概率p_i,于是np_i就是落入A_i的样本值的理论频数.

实测频数

理论频数

$$m_i - np_i$$

标志着经验分布与理论分布之间的差异的大小.

皮尔逊引进如下统计量表示经验分布与理论分布之间的差异:

在理论分布已知的条件下，np_i是常量

$$\chi^2 = \sum_{i=1}^{k} \frac{(m_i - np_i)^2}{np_i}$$

统计量 χ^2 的分布是什么?

定理(皮尔逊) 若原假设中的理论分布$F(x)$已经完全给定，那么当$n\to\infty$时，统计量

$$\chi^2=\sum_{i=1}^{k}\frac{(m_i-np_i)^2}{np_i}$$

的分布渐近$(k\text{-}1)$个自由度的χ^2分布.

如果理论分布$F(x)$中有r个未知参数需用相应的估计量来代替，那么当$n\to\infty$时，统计量χ^2的分布渐近$(k\text{-}r\text{-}1)$个自由度的χ^2分布.

根据这个定理，对给定的显著性水平α，查χ^2分布表可得临界值χ_α^2，使得

$$P(\chi^2 > \chi_\alpha^2) = \alpha$$

得拒绝域：$\chi^2 > \chi_\alpha^2(k-1)$ (不需估计参数)

$$\chi^2 > \chi_\alpha^2(k-r-1)$$ (估计r个参数)

如果根据所给的样本值$X_1, X_2, \ldots, X_n$算得统计量χ^2的实测值落入拒绝域，则拒绝原假设，否则就认为差异不显著而接受原假设.

皮尔逊定理是在n无限增大时推导出来的，因而在使用时要注意n要足够大，以及np_i不太小这两个条件.

根据计算实践，要求n不小于50，以及np_i都不小于 5. 否则应适当合并区间，使np_i满足这个要求 .

P210.1（自习）

这一节我们介绍了拟合优度的 χ^2 检验法. 在对总体的分布进行检验时经常使用.

教材上的其他一些非参数假设检验方法留给同学们自己看. 由于这些检验的计算量相对较大，一般要用统计软件包来实现.