第七章 假设检验

- 基本概念
- 正态总体均值的假设检验
- 正态总体方差的假设检验
- 置信区间与假设检验之间的关系
- 样本容量的选取
- 分布拟合检验
- 秩和检验

统计假设检验的基本任务

根据样本所提供的信息，对未知总体分布的某些方面（如总体均值、方差、分布本身等等）的假设作出合理的判断。

§1 假设检验的基本概念

一、假设检验问题

例1： 假定按国家规定，某种产品的次品率不得超过1%，现从一批产品中随机抽出200件，经检查发现有3件次品，试问：这批产品是否次品率符合国家标准。

问题：根据抽样的结果来判断是否$p \le 0.01$.

例2： 某厂生产的合金强度服从正态分布$N(\theta, 16)$，其中θ的设计值为$10(pa)$，某天从生产中随机取25块合金，测得强度值为$x_1, \cdots, x_{25}$，其均值$\bar{x} = 108$，

问当日生产是否正常 $\theta = 110.$

例3：某建筑材料，其抗断强度的分布以往一直服从正态分布，现改变配料方案，希望确定新产品的抗断强度 X 是否仍服从正态分布.

这类问题称为假设检验，有两个共同特点：

(1) 先根据实际问题的要求提出一个关于随机变量的一种论断，称为统计假设，记为 H_0

H_0： $p \le 0.01$

H_0： $\theta = 110$

H_0： $X \sim N(\mu, \sigma^2).$

（2） 抽取样本和集中样本的有关信息，要求对假设 H_0 的真伪进行判断，称为**检验假设**，最后对假设 H_0 作出拒绝或接受的决策。

假设检验问题的分类：

1. 参数假设检验：

若总体的分布函数 $F(x;\theta_1,\cdots,\theta_k)$ 或概率函数 $p(x;\theta_1,\cdots,\theta_k)$ 已知，参数未知，假设 H_0 针对未知参数提出并要求检验。

2. 非参数假设检验

若总体的分布函数或概率函数为未知，假设 H_0 针对总体的分布，分布的特征或总体的数字特征而提出并要求检验。

H_0 表示原来的假设，称为原假设或零假设。

所考察的问题的反面称为备择假设或对立假设，记为 H_1.

二、假设检验的基本原理

概率性质的反证法：

为了检验原假设H_0是否正确，先假定H_0为正确，看由此能推出什么结果，如果导致一个不合理现象出现，则表明“假设H_0 为正确”是错误的，拒绝原假设 H_0；否则，则接受原假设H_0。

概率性质的反证法的根据：

小概率事件在一次试验中几乎是不可能发生的。

注： 在假设检验中，我们作出接受 H_0 或拒绝 H_0 的决策，并不等于我们证明了原假设 H_0 正确或错误，而只是根据样本所提供的信息以一定的可靠程度认为 H_0 是正确或错误。

实例　某车间用一台包装机包装葡萄糖, 包得的袋装糖重是一个随机变量, 它服从正态分布. 当机器正常时, 其均值为0.5公斤, 标准差为0.015公斤. 某日开工后为检验包装机是否正常, 随机地抽取它所包装的糖9袋, 称得净重为(公斤):

0.497　0.506　0.518　0.524　0.498　0.511

0.520　0.515　0.512,　问机器是否正常?

分析: 用 μ 和 σ 分别表示这一天袋装糖重总体 X 的均值和标准差,

由长期实践可知, 标准差较稳定, 设 $\sigma = 0.015$,

则 $X \sim N(\mu,\ 0.015^2)$, 其中 μ 未知.

问题: 根据样本值判断 $\mu = 0.5$ 还是 $\mu \neq 0.5$.

提出两个对立假设 $H_0: \mu = \mu_0 = 0.5$ 和 $H_1: \mu \neq \mu_0$.

再利用已知样本作出判断是接受假设 H_0(拒绝假设 H_1), 还是拒绝假设 H_0(接受假设 H_1).

如果作出的判断是接受 H_0, 则 $\mu = \mu_0$,

即认为机器工作是正常的, 否则, 认为是不正常的.

由于要检验的假设涉及总体均值, 故可借助于样本均值来判断.

因为 $\overline{X}$ 是 μ 的无偏估计量,

所以若 H_0 为真, 则 $|\overline{x}-\mu_0|$ 不应太大,

当H_0为真时, $\dfrac{\overline{X}-\mu_0}{\sigma/\sqrt{n}} \sim N(0,1)$,

衡量 $|\overline{x}-\mu_0|$ 的大小可归结为衡量 $\dfrac{|\overline{x}-\mu_0|}{\sigma/\sqrt{n}}$ 的大小,

于是可以选定一个适当的正数k,

当观察值 $\bar{x}$ 满足 $\dfrac{|\bar{x}-\mu_0|}{\sigma/\sqrt{n}} \geq k$ 时，拒绝假设 H_0，

反之，当观察值 $\bar{x}$ 满足 $\dfrac{|\bar{x}-\mu_0|}{\sigma/\sqrt{n}} < k$ 时，接受假设 H_0.

当 H_0 为真时 $U = \dfrac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} \sim N(0,1)$,

检验统计量

根据小概率原理

因而当 H_0 为真，即 $\mu=\mu_0$ 时，$\left\{\left|\dfrac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}\right| \geq k\right\}$ 是一个小概率事件

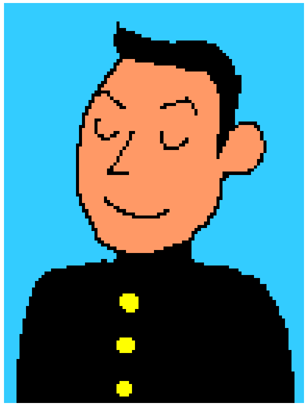
$$p\left\{\left|\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}\right| \geq k \,\middle|\, H_0\text{为真}\right\} = \alpha$$

取给定的常数α，通常α总是取得很小,一般取 $\alpha = 0.01, \alpha = 0.05$, 满足

$$p\left\{\left|\frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}\right| \ge k \middle| H_0\text{为真}\right\} = \alpha$$

由标准正态分布分位点的定义得 $k = z_{\alpha/2}$，

当$\dfrac{|\overline{x}-\mu_0|}{\sigma/\sqrt{n}} \ge z_{\alpha/2}$**时,拒绝**$H_0$, $\dfrac{|\overline{x}-\mu_0|}{\sigma/\sqrt{n}} < z_{\alpha/2}$**时, 接受**$H_0$.



在实例中若取定$\alpha = 0.05$,

则$k = z_{\alpha/2} = z_{0.025} = 1.96$,

又已知$n = 9$, $\sigma = 0.015$,

由样本算得 $\bar{x} = 0.511$, 即有 $\dfrac{|\bar{x} - \mu_0|}{\sigma / \sqrt{n}} = 2.2 > 1.96$,

于是拒绝假设H_0, 认为包装机工作不正常.

三、两类错误

检验的结果与真实情况可能吻合也可能不吻合，因此检验是可能犯错误的。

第一类错误：

H_0为真但由于随机性样本观测值落在拒绝域中，从而拒绝H_0。

拒真概率

$$\alpha = P(\text{拒绝}H_0 \mid H_0\text{为真}) = P_\theta(X \in W) \quad \theta \in \Theta_0$$

第二类错误：

H_0不真但由于随机性样本观测值落在接受域中，从而接受H_0。

受伪概率

$$\beta = P(\text{接受}H_0 \mid H_1\text{为真}) = P_\theta(X \in \overline{W}) \quad \theta \in \Theta_1$$

假设检验的两类错误

真实情况 (未知)	所 作 决 策	
	接受H_0	拒绝H_0
H_0为真	正确	犯第I类错误
H_0不真	犯第II类错误	正确

注：在样本量一定的条件下，不可能找到一个使α，β 都减小的检验。

一般来说,我们总是控制第I类错误的概率, 使它不大于 α . 这种只对犯第I类错误的概率加以控制, 而不考虑第II类错误的概率的检验, 称为**显著性检验.**

α 称为**显著性水平**.

说明:

控制α并制约β，通常选择$\alpha = 0.01/0.05/0.1$。

四、假设检验的基本步骤

（一） 建立假设：提出原假设 H_0 和备选假设 H_1.

（二） 构造检验统计量与确定拒绝域的形式

分布密度 $P(x,\theta)$ 的表达式为已知时，通常以 θ 的极大似然估计 $\hat{\theta}$ 为基础构造一个检验统计量 $T=T(X_1,\cdots,X_n)$，并且在 H_0 成立的条件下，确定 T 的精确分布或渐近分布

确定检验统计量后，根据原假设H_0与备选假设确定拒绝域W。

拒绝域：使原假设被拒绝的样本观测值所在的区域。

常用的拒绝域形式：

（1）单侧拒绝域

$$W=\{(x_1,\cdots,x_n):T(x_1,\cdots,x_n)\geq c\}$$

或

$$W=\{(x_1,\cdots,x_n):T(x_1,\cdots,x_n)\leq c\}$$

（2）双侧拒绝域

$$W=\{(x_1,\cdots,x_n):T(x_1,\cdots,x_n)\le c_1 \text{或} T(x_1,\cdots,x_n)\ge c_2\}$$

或

$$W=\{(x_1,\cdots,x_n):|T(x_1,\cdots,x_n)|\ge c\}$$

其中临界值c_1,c_2和c_3待定。

当拒绝域确定了，检验的判断准则也确定了

如果$(x_1,\cdots,x_n)\in W$，则认为H_0不成立

如果$(x_1,\cdots,x_n)\in \overline{W}$，则认为$H_0$成立。

（三）选定适当的显著水平 α ，并求出临界值。

（四）根据样本观测值确定是否拒绝H_0.

由样本值$(x_1,\cdots,x_n)$算得$T(x_1,\cdots,x_n)$，把它与临界值相比较，若$(x_1,\cdots,x_n)\in W$，则拒绝H_0，否则接受H_0。

§2 正态总体均值的假设检验

一、单个正态总体均值的检验

总体$X \sim N(\mu,\sigma^2)$，$X_1,\cdots,X_n$为X的样本，$x_1,\cdots,x_n$为观测值。

（一） 方差 σ^2 已知时均值 μ 的假设检验

当σ^2为已知时，在给定显著水平α下，关于μ的检验问题有以下三种

(1) $H_0:\mu=\mu_0$ ， $H_1:\mu\neq\mu_0$ （双侧检验）

(2) $H_0: \mu = \mu_0$, $H_1: \mu > \mu_0$ （右侧检验）

$H_0: \mu \le \mu_0$, $H_1: \mu > \mu_0$

(3) $H_0: \mu = \mu_0$, $H_1: \mu < \mu_0$ （左侧检验）

$H_0: \mu \ge \mu_0$, $H_1: \mu < \mu_0$

其中μ_0为常数。

(2)现讨论 $H_0: \mu = \mu_0$ ，$H_1: \mu > \mu_0$

对于正态总体，$\overline{X}$是μ的MLE。若$\overline{X}-\mu_0$比较大，则应该认定H_0为真时出现了小概率事件，应该拒绝H_0。

取

$$P\{\overline{X}-\mu_0 \geq k \mid H_0\text{为真}\} = \alpha,$$

因为在 H_0 下

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

根据标准正态分布的分位数定义可得

$$P\left\{\frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}} \geq z_\alpha\right\}=\alpha,$$

所以检验的拒绝域为

$$W=\left\{(x_1,\cdots,x_n):\frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} \geq z_\alpha\right\}.$$

再讨论 $H_0:\mu \le \mu_0$, $H_1:\mu > \mu_0$

在$\boldsymbol{H}_0$

$$= P_{\mu \le \mu_0}\left\{ \frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} \ge \frac{k-\mu_0}{\sigma/\sqrt{n}} \right\},$$

由于 $\mu \le \mu_0$

$$\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \ge \frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}$$

所以

$$\left\{\omega:\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}\ge\frac{k-\mu_0}{\sigma/\sqrt{n}}\right\}\supseteq\left\{\omega:\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}\ge\frac{k-\mu_0}{\sigma/\sqrt{n}}\right\}$$

即 $P\{当H_0为真拒绝H_0\}=P_{\mu\in H_0}\{\bar{X}\ge k\}$

$$=P_{\mu\le\mu_0}\left\{\omega:\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}\ge\frac{k-\mu_0}{\sigma/\sqrt{n}}\right\}\le P_{\mu\le\mu_0}\left\{\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}\ge\frac{k-\mu_0}{\sigma/\sqrt{n}}\right\}=\alpha,$$

由于总体$X\sim N(\mu,\sigma^2)$，因此 $\dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}}\sim N(0,1)$,

$$\frac{k-\mu_0}{\sigma/\sqrt{n}}=z_\alpha \quad \alpha\text{ 分位点}$$

故有

$$\alpha = P\left\{\omega : \frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \geq z_\alpha\right\} \geq P\left\{\omega : \frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}} \geq z_\alpha\right\}$$

也就是说$\left\{\omega : \frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}} \geq z_\alpha\right\}$是$H_0$为真时的小

概率事件，故拒绝域为

$$W = \left\{(x_1,\cdots,x_n) : \frac{\overline{x}-\mu_0}{\sigma/\sqrt{n}} \geq z_\alpha\right\}.$$

根据类似的讨论可以得到问题（3）

$$H_0:\mu \geq \mu_0, H_1:\mu < \mu_0$$

的拒绝域为

$$W=\left\{(x_1,\cdots,x_n):\frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}\leq -z_\alpha\right\}.$$

问题(1)的拒绝域为 (1) $H_0:\mu=\mu_0$ ， $H_1:\mu\neq\mu_0$

$$W=\left\{(x_1,\cdots,x_n):\left|\frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}\right|\geq z_{\alpha/2}\right\}.$$

Z检验法

例1： 有一批枪弹，其初速度 $\sim N(\mu,\sigma^2)$，其中 $\mu=950\,m/s$，$\sigma=10\,m/s$，经过较长时间存储后，现取出9发枪弹试射测其初速度，得样本值如下（单位：m/s）

$$914, 920, 910, 934, 953, 945, 912, 924, 940$$

给定显著性水平 $\alpha=0.05$，问这批枪弹的初速度是否起了变化（假定 σ 没有变化）？

解：　$H_0:\mu=950\quad H_1:\mu<950$

此检验的拒绝域为

$$W=\left\{(x_1,\cdots,x_n):\frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}\le -z_\alpha\right\}.$$

有$n=9$，易算得$\bar{x}=928$，则

$$u=\frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}=\frac{928-950}{10/\sqrt{9}}=-6.6$$

查标准正态表可得

$$-\mu_{0.05}=-1.96$$

因为

$$u=\frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}=\frac{928-950}{10/\sqrt{9}}=-6.6<-1.96$$

所以，应拒绝原假设，认为这批枪弹经过较长时间的存储后初速度已经发生了变化，变小了。

（二）方差 σ^2 未知时均值 μ 的假设检验

当σ^2为未知时，在给定显著性水平α下，关于μ的检验问题有以下三种:

(1) $H_0:\mu=\mu_0$ ， $H_1:\mu\neq\mu_0$ （双侧检验）

(2) $H_0:\mu=\mu_0$ ， $H_1:\mu>\mu_0$

$H_0:\mu\leq\mu_0$ ， $H_1:\mu>\mu_0$ （右侧检验）

(3) $H_0:\mu=\mu_0$ ， $H_1:\mu<\mu_0$

$H_0:\mu\geq\mu_0$ ， $H_1:\mu<\mu_0$ （左侧检验）

其中μ_0为常数。

现讨论 $H_0: \mu = \mu_0 \quad v.s. \quad H_1: \mu < \mu_0$

对于正态总体，$\bar{X}$是μ的MLE。若$\mu_0 - \bar{X}$比较大，则应该认定H_0为真时出现了小概率事件，应该拒绝H_0。

取

$$P\{\mu_0 - \bar{X} \geq k \mid H_0\text{为真}\} = \alpha,$$

因为在H_0下

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

根据t分布的分位数定义可得

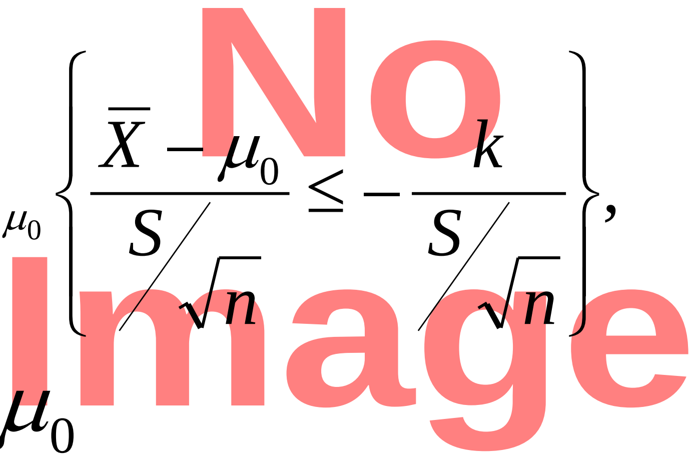
$$P\left\{\frac{\overline{X}-\mu_0}{S/\sqrt{n}} \le -t_\alpha(n-1)\right\} = \alpha,$$

所以检验的拒绝域为

$$W=\left\{(x_1,\cdots,x_n):\frac{\bar{x}-\mu_0}{S/\sqrt{n}} \le -t_\alpha(n-1)\right\}.$$

再讨论 $\quad H_0: \mu \geq \mu_0 \quad v.s. \quad H_1: \mu < \mu_0$

在H_0



$$= P_{\mu \geq \mu_0}\left\{ \frac{\bar{X} - \mu_0}{S/\sqrt{n}} \leq -\frac{k}{S/\sqrt{n}} \right\},$$

由于 $\quad \mu \geq \mu_0$

$$\frac{\bar{X} - \mu}{S/\sqrt{n}} \leq \frac{\bar{X} - \mu_0}{S/\sqrt{n}}$$

所以

$$\left\{\omega:\frac{\overline{X}-\mu}{S/\sqrt{n}}\le -\frac{k}{S/\sqrt{n}}\right\}\supseteq\left\{\omega:\frac{\overline{X}-\mu_0}{S/\sqrt{n}}\le -\frac{k}{S/\sqrt{n}}\right\}$$

即 $P\{当H_0为真拒绝H_0\}=P_{\mu\in H_0}\{\mu_0-\overline{X}\ge k\}$

$$=P_{\mu\ge\mu_0}\left\{\omega:\frac{\overline{X}-\mu_0}{S/\sqrt{n}}\le -\frac{k}{S/\sqrt{n}}\right\}\le P_{\mu\ge\mu_0}\left\{\frac{\overline{X}-\mu}{S/\sqrt{n}}\le -\frac{k}{S/\sqrt{n}}\right\}=\alpha,$$

由于总体$X\sim N(\mu,\sigma^2)$，因此 $\dfrac{\overline{X}-\mu}{S/\sqrt{n}}\sim t(n-1)$,

$$-\frac{k}{S/\sqrt{n}}=-t_\alpha(n-1)$$

故有

$$P\left\{\omega: \frac{\bar{X}-\mu_0}{S/\sqrt{n}} \le -t_\alpha(n-1)\right\} \le P\left\{\omega: \frac{\bar{X}-\mu}{S/\sqrt{n}} \le -t_\alpha(n-1)\right\} = \alpha$$

也就是说$\left\{\omega: \frac{\bar{X}-\mu_0}{S/\sqrt{n}} \le -t_\alpha(n-1)\right\}$是$H_0$为真时的小概率事件，故拒绝域为

$$W=\left\{(x_1,\cdots,x_n): \frac{\bar{x}-\mu_0}{s/\sqrt{n}} \le -t_\alpha(n-1)\right\}.$$

根据类似的讨论可以得到问题（2）的拒绝域为

$$W=\left\{(x_1,\cdots,x_n):\frac{\bar{x}-\mu_0}{s/\sqrt{n}}\geq t_\alpha(n-1)\right\}.$$

T检验法

问题(1)的拒绝域为

$$W=\left\{(x_1,\cdots,x_n):\left|\frac{\bar{x}-\mu_0}{s/\sqrt{n}}\right|\geq t_{\alpha/2}(n-1)\right\}.$$

例2: 某厂生产的某种铝材的长度服从正态分布，其均值设定为240cm,现从该厂抽取5件产品，测得其长度为（单位：cm）

239.7 239.6 239 240 239.2

试判断该厂此类铝材的长度是否满足设定要求?（取显著性水平为$\alpha = 0.05$）。

解：根据题意做检验 $H_0: \mu = 240 \quad H_1: \mu \neq 240$

拒绝域为

$$W = \left\{ (x_1, \cdots, x_n) : \left| \frac{\bar{x} - \mu_0}{s/\sqrt{n}} \right| \geq t_{\alpha/2}(n-1) \right\}.$$

现由样本计算可得 $\bar{x} = 239.5, \quad s = 0.4,$ 故

$$|t| = \frac{|\bar{x} - \mu_0|}{s/\sqrt{n}} = \sqrt{5} \cdot |239.5 - 240/0.4 = 2.795$$

查表可得 $t_{0.025}(4) = 2.776,$

由于**2.795>2.776**，故拒绝原假设，认为该厂生产的铝材的长度不满足设定要求。

例3 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? $(\alpha = 0.01)$

解: $H_0: \mu = 32.5, \quad H_1: \mu \neq 32.5$

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

例3（续）

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

拒绝域为

$$W = \left\{ (x_1, \cdots, x_n) \middle| \frac{|\overline{x} - \mu_0|}{s/\sqrt{n}} > t_{\alpha/2}(n-1) \right\}$$

其中

$$n = 6, \alpha = 0.01, \mu_0 = 32.5,$$

$$t_{\alpha/2}(n-1) = t_{0.005}(5) = 4.0322$$

将样本值代入算出统计量 T 的值, $|T| = 2.997 < 4.0322$

所以接受H_0, 即认为这批产品合格.

例4 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$\bar{x}$=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： $H_0: \mu = 21, \quad H_1: \mu > 21$

取统计量 $Z = \dfrac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \sim N(0,1)$

拒绝域为 $W=\left\{(x_1,\cdots,x_n)\middle|\ \dfrac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}>z_\alpha\right\}$

这里$\mu_0=21$, $\sigma=1.2, n=30, \alpha=0.01$, 由样本值算出

Z=2.51, 而 $z_{0.01}=2.33$, Z=2.51>2.33,

所以拒绝原假设H_0,

即认为新生产织物比过去的织物强力有提高.

二、 两个正态总体均值差的检验

考虑 $X \sim N(\mu_1, \sigma_1^2), X_1, \cdots, X_{n_1}$ 为X的一个样本

$Y \sim N(\mu_2, \sigma_2^2), Y_1, \cdots, Y_{n_2}$ 为Y的一个样本

（一）当方差已知时 $\mu_1 - \mu_2$ 的假设检验

当σ_1^2, σ_2^2已知时，在水平α下，$\mu_1 - \mu_2$的假设检验问题

$$(1)\quad H_0:\mu_1-\mu_2=\delta \quad v.s. \quad H_1:\mu_1-\mu_2\neq\delta$$

（双侧检验）

$$(2)\quad H_0:\mu_1-\mu_2=\delta \quad v.s. \quad H_1:\mu_1-\mu_2>\delta$$

$$H_0:\mu_1-\mu_2\leq\delta \quad v.s. \quad H_1:\mu_1-\mu_2>\delta$$

（右侧检验）

$$(3)\quad H_0:\mu_1-\mu_2=\delta \quad v.s. \quad H_1:\mu_1-\mu_2<\delta$$

$$H_0:\mu_1-\mu_2\geq\delta \quad v.s. \quad H_1:\mu_1-\mu_2<\delta$$

（左侧检验）

其中δ为任意已知的常数

现讨论 $H_0: \mu_1 - \mu_2 = \delta \quad v.s. \quad H_1: \mu_1 - \mu_2 > \delta$

当H_0为真时，$(\overline{X} - \overline{Y}) - \delta$比较大就应该拒绝$H_0$，于是

$$P\{(\overline{X} - \overline{Y}) - \delta \geq k \mid H_0\text{为真}\} = \alpha,$$

其中k为适当大的正数

当H_0为真时，

$$U = \frac{(\overline{X} - \overline{Y}) - \delta}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0,1),$$

由$N(0,1)$的分位数知，当H_0为真时，

$$P\left\{\frac{\overline{X}-\overline{Y}-\delta}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}\geq z_\alpha\right\}=\alpha,$$

所以，此检验的拒绝域为

$$W=\left\{(x_1,\cdots,x_n):\frac{\overline{X}-\overline{Y}-\delta}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}\geq z_\alpha\right\}.$$

再讨论 $H_0: \mu_1 - \mu_2 \le \delta \quad v.s. \quad H_1: \mu_1 - \mu_2 > \delta$

当H_0为真时，有

$$\frac{\bar{X} - \bar{Y} - \delta}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \le \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

由此可得

$$\left\{ \frac{\bar{X} - \bar{Y} - \delta}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \ge z_\alpha \right\} \subseteq \left\{ \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \ge z_\alpha \right\}$$

所以，在H_0为真时，

$$P\left\{\frac{\overline{X}-\overline{Y}-\delta}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}\geq z_\alpha\right\}\leq P\left\{\frac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}\geq z_\alpha\right\}=\alpha,$$

因此，右侧检验的拒绝域都是

$$W=\left\{(x_1,\cdots,x_n):\frac{\overline{X}-\overline{Y}-\delta}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}\geq z_\alpha\right\}.$$

同样，检验（1）的拒绝域为

$$W=\left\{(x_1,\cdots,x_n):\left|\frac{\overline{X}-\overline{Y}-\delta}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}\right|\geq z_{\frac{\alpha}{2}}\right\}.$$

检验（3）的拒绝域为

$$W=\left\{(x_1,\cdots,x_n):\frac{\overline{X}-\overline{Y}-\delta}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}\leq -z_{\alpha}\right\}.$$

（二）方差未知但相等时 $\mu_1-\mu_2$ 的假设检验

当 σ_1^2,σ_2^2 为未知，但 $\sigma_1^2=\sigma_2^2$ 时，均值差 $\mu_1-\mu_2$ 的常见的假设检验问题,与 σ_1^2,σ_2^2 为已知时的情况一样，仍然是那三个问题。

统计量应选为

$$T=\frac{\overline{X}-\overline{Y}-(\mu_1-\mu_2)}{S_W\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}\sim t(n_1+n_2-2).$$

其中

$$S_W=\sqrt{\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}}$$

把T作为检验统计量，可得下面的水平为α的拒绝域

（1） 双侧检验

$$W=\left\{(x_1,\cdots,x_n):\frac{|(\bar{X}-\bar{Y})-\delta|}{S_W\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}\ge t_{\frac{\alpha}{2}}(n_1+n_2-2)\right\}$$

（2） 右侧检验

$$W=\left\{(x_1,\cdots,x_n):\frac{(\bar{X}-\bar{Y})-\delta}{S_W\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}\ge t_{\alpha}(n_1+n_2-2)\right\}$$

（3） 左侧检验

$$W=\left\{(x_1,\cdots,x_n):\frac{(\overline{X}-\overline{Y})-\delta}{S_W\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}\le -t_\alpha(n_1+n_2-2)\right\}$$

例5： 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，试验是在一个平炉上进行的。每炼一炉钢时，除操作方法外，其它条件都尽可能做的相同。先用标准方法炼一炉，然后用建议的新方法炼一炉，各炼了**10**炉其得率分别为

(1)标准方法：**78.1 72.4 76.2 74.3 77.4 78.4 76.0 75.5 76.7 77.3**

(2)新方法：**79.1 81.0 77.3 79.1 80.0 79.1 79.1 77.3 80.2 82.1**

设这两个样本独立，且分别来自正态总体$N(\mu_1,\sigma^2)$和$N(\mu_2,\sigma^2)$，μ_1,μ_2,σ^2均未知，问建议的新操作方法能否提高得率？(取$\alpha=0.05$。)

解：需要检验假设

$$H_0:\mu_1-\mu_2\geq 0 \quad H_1:\mu_1-\mu_2<0$$

分别求出标准方法和新方法下的样本均值和样本方差如下

$$n_1=10,\quad \bar{x}=76.23,\quad s_1^2=3.325$$

$$n_2=10,\quad \bar{y}=79.43,\quad s_2^2=2.225$$

又$s_w^2=\dfrac{(10-1)s_1^2+(10-1)s_2^2}{10+10-2}=2.775,\quad t_{0.05}(18)=1.7341$

由于 $t=\dfrac{\bar{x}-\bar{y}}{s_w^2\sqrt{\dfrac{1}{10}+\dfrac{1}{10}}}=-4.295<-1.7341$

所以拒绝H_0，即认为建议的新操作方法较原来的方法为优。

三、基于成对数据的检验

逐对比较法：在相同的条件下作对比试验，得到一批成对的观察值，然后分析观察数据作出判断。

举例说明：

1. 比较甲、乙两种橡胶轮胎的耐磨性的试验：从甲
2. 乙两种轮胎中各抽取n个，各取一个组成一对。
3. 再随机抽取n架飞机，将n对轮胎分配给n架飞机
4. 飞行了一定时间的起落后，测得轮胎磨损量的n
5. 对数据。

2. 比较两种测量方法所得的结果是否有显著性的差异。对同一试验品分别用两种方法测量得到成对数据。

设有n对相互独立的观察结构$(X_1,Y_1),(X_2,Y_2),\cdots,$ (X_n,Y_n),令$D_1=X_1-Y_1,\cdots,D_n=X_n-Y_n$,则$D_1,D_2,\cdots,$ D_n相互独立，又由于$D_1,D_2,\cdots,D_n$是由同一因素所引起的,可认为它们服从同一分布.假设$D_i\sim N(\mu_D,\sigma_D^2)$, $i=1,2,\cdots,n$。这就是说$D_1,D_2,\cdots,D_n$构成正态总体$N(\mu_D,\sigma_D^2)$的一个样本，其中μ_D,σ_D^2未知。

基于这一样本检验假设:

(1) $H_0:\mu_D=0,\quad H_1:\mu_D\neq 0;$

(2) $H_0:\mu_D\leq 0,\quad H_1:\mu_D>0;$

(3) $H_0:\mu_D\geq 0,\quad H_1:\mu_D<0.$

分别记$D_1, D_2, \cdots, D_n$的样本均值和样本方差的观察值为$\bar{d}, s_D^2$，则此问题转化为关于单个正态总体均值的t检验。检验问题(1),(2),(3)的拒绝域分别为(检验水平为α)：

$$|t| = \left| \frac{\bar{d}}{s_D / \sqrt{n}} \right| \geq t_{\alpha/2}(n-1),$$

$$t = \frac{\bar{d}}{s_D / \sqrt{n}} \geq t_{\alpha}(n-1),$$

$$t = \frac{\bar{d}}{s_D / \sqrt{n}} \leq -t_{\alpha}(n-1)$$

例6 有两台光谱仪，用来测量材料中某种金属的含量，为鉴定他们的测量结果有无显著的差异，制备了9件试块（它们的成份、金属含量、均匀性等均各不相同），现在分别用这两台仪器对每一试件测量一次，得到9对观察值如下：

x%	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90	1.00
y%	0.10	0.21	0.52	0.32	0.78	0.59	0.68	0.77	0.89
d=x% -y%	0.10	0.09	-0.12	0.18	-0.18	0.11	0.12	0.13	0.11

问能否认为这两台仪器的测量结果有显著的差异？）

（取$\alpha=0.01$）

解：$H_0: \mu_0 = 0, \quad H_1: \mu_0 \neq 0$

取统计量 $T = \dfrac{\overline{D} - 0}{S_D / \sqrt{n}}$

拒绝域为 W：$|T| > t_{\alpha/2}(n-1)$

由样本值计算得：$\overline{D} = 0.06, \quad S_D = 0.1227, T = 1.467$

查表得 $t_{\alpha/2}(n-1) = t_{0.005}(8) = 3.3554$

由于 $|T| = 1.467 < 3.3554$， 故接受H_0，即认为这两台仪器的测量结果无显著差异.

例7 假设有A, B两种药, 试验者欲比较它们在服用2小时后血液中的含量是否一样. 对药品A, 随机抽取了8个病人, 测得他们服药2小时后血液中药的浓度为

1.23, 1.42, 1.41, 1.62, 1.55, 1.51, 1.60, 1.76.

对药品B, 随机抽取了6个病人, 测得他们服药2小时后血液中药的浓度为

1.76, 1.41, 1.87, 1.49, 1.67, 1.81.

假定这两组观测值是具有相同方差的正态分布, 试在显著水平$\alpha=0.10$下, 检验病人血液中这两种药的浓度是否有显著不同?

例7解: $H_0: \mu_1 - \mu_2 = 0, \quad H_1: \mu_1 - \mu_2 \neq 0$

取统计量 $T = \dfrac{(\overline{X} - \overline{Y}) - 0}{S_w\sqrt{\dfrac{1}{m} + \dfrac{1}{n}}} \sim t(m+n-2)$

其中 $S_w^2 = \dfrac{(m-1)S_1^2 + (n-1)S_2^2}{m+n-2}$

拒绝域为 $W = \{t : |t| \geq t_{\alpha/2}(m+n-2)\}$

其中 $m = 8, n = 6, \alpha = 0.1,$

由样本值可计算得

$\overline{x} = 1.51, \ \overline{y} = 1.66, \ s_1^2 = 0.03, \ s_2^2 = 0.034, \quad s_w = 0.18$

例7解: 拒绝域为 W: $|t| \geq t_{\alpha/2}(m+n-2)$

$$t = \frac{(\bar{x}-\bar{y})-0}{S_w\sqrt{\frac{1}{m}+\frac{1}{n}}} = -1.54$$

查表得 $t_{\alpha/2}(m+n-2) = t_{0.05}(12) = 1.78$

由于 $|t| = 1.54 < 1.78$， 故接受 H_0 ，

即认为血液中这两种药的浓度没有显著不同.

§3 正态总体方差的假设检验

（一）单个总体的情况

设总体$X\sim N(\mu,\sigma^2)$，$X_1,X_2,\cdots,X_n$是来自X的样本。当μ为未知时，在给定显著水平α下，关于σ^2的检验问题有以下三种：

(1) $H_0:\sigma^2=\sigma_0^2 \quad v.s. \quad H_1:\sigma^2\neq\sigma_0^2$ （双侧检验）

(2) $H_0:\sigma^2=\sigma_0^2 \quad v.s. \quad H_1:\sigma^2>\sigma_0^2$

$H_0:\sigma^2\le\sigma_0^2 \quad v.s. \quad H_1:\sigma^2>\sigma_0^2$ （右侧检验）

(3) $H_0:\sigma^2=\sigma_0^2 \quad v.s. \quad H_1:\sigma^2<\sigma_0^2$

$H_0:\sigma^2\ge\sigma_0^2 \quad v.s. \quad H_1:\sigma^2<\sigma_0^2$ （左侧检验）

其中σ_0^2为常数。

先讨论 $H_0:\sigma^2=\sigma_0^2 \quad v.s. \quad H_1:\sigma^2\neq\sigma_0^2$

对于正态总体，$\dfrac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2$是$\sigma^2$的无偏估计

因此当$H_0:\sigma^2=\sigma_0^2$为真时，$\dfrac{\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2}{\sigma_0^2}$应接近1，

当比值接近0或比1大很多时，都应当拒绝H_0。

$$P\left\{\frac{\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2}{\sigma_0^2}\le k_1 \,\middle|\, H_0\text{为真}\right\}+P\left\{\frac{\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2}{\sigma_0^2}\ge k_2 \,\middle|\, H_0\text{为真}\right\}=\alpha$$

其中k_1为适当小的正数,k_2为适当大的正数为简便取

$$P\left\{\frac{\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2}{\sigma_0^2}\le k_1\right\}=P\left\{\frac{\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2}{\sigma_0^2}\ge k_2\right\}=\frac{\alpha}{2}$$

当H_0为真时，知

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

所以

$$P\left\{\frac{\sum_{i=1}^{n}(X_i-\overline{X})^2}{\sigma_0^2}\le \chi^2_{1-\frac{\alpha}{2}}(n-1)\right\}=P\left\{\frac{\sum_{i=1}^{n}(X_i-\overline{X})^2}{\sigma_0^2}\ge \chi^2_{\frac{\alpha}{2}}(n-1)\right\}=\frac{\alpha}{2}$$

因此拒绝域为

$$W=\left\{(x_1,\cdots,x_n):\frac{\sum_{i=1}^{n}(x_i-\overline{x})^2}{\sigma_0^2}\le \chi^2_{1-\frac{\alpha}{2}}(n-1)\text{或}\frac{\sum_{i=1}^{n}(x_i-\overline{x})^2}{\sigma_0^2}\ge \chi^2_{\frac{\alpha}{2}}(n-1)\right\}\text{。}$$

现讨论 $H_0:\sigma^2=\sigma_0^2 \quad v.s. \quad H_1:\sigma^2>\sigma_0^2$

对于正态总体，当μ未知时，$\dfrac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2$是$\sigma^2$的无偏估计

在H_0为真时，

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

若$\dfrac{\sum_{i=1}^{n}(X_i-\overline{X})^2}{\sigma_0^2}$比较大，就应该认为在假设$H_0$为真时

出现了小概率事件，从而应该拒绝H_0。于是取

$$P\left\{\frac{\sum_{i=1}^{n}(X_i-\overline{X})^2}{\sigma_0^2}\geq\chi_\alpha^2(n)\right\}=\alpha,$$

即得所讨论的假设检验问题的拒绝域

$$W=\left\{(x_1,\cdots,x_n):\frac{\sum_{i=1}^{n}(x_i-\overline{x})^2}{\sigma_0^2}\geq\chi_\alpha^2(n-1)\right\}。$$

现讨论 $\boldsymbol{H_0:\sigma^2\le\sigma_0^2 \quad v.s. \quad H_1:\sigma^2>\sigma_0^2}$

当 H_1 为真时, S^2 的观察值 s^2 往往很大,因此拒绝域的形式为

$$s^2\ge k$$

$$P\{当H_0为真拒绝H_0\}=P_{\sigma^2\le\sigma_0^2}\{S^2\ge k\} =$$

$$P_{\sigma^2\le\sigma_0^2}\left\{\frac{\sum_{i=1}^{n}(X_i-\overline{X})^2}{\sigma_0^2}\ge\frac{(n-1)k}{\sigma_0^2}\right\}=P_{\sigma^2\le\sigma_0^2}\left\{\frac{\sum_{i=1}^{n}(X_i-\overline{X})^2}{\sigma_0^2}\ge\chi_\alpha^2(n-1)\right\}$$

当H_0为真时，有$\sigma^2 \le \sigma_0^2$，从而得

$$\frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{\sigma_0^2} \le \frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{\sigma^2}$$

由此可得

$$\left\{ \frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{\sigma_0^2} \ge \chi_\alpha^2(n-1) \right\} \subseteq \left\{ \frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{\sigma^2} \ge \chi_\alpha^2(n-1) \right\}$$

由于总体$X \sim N(\mu,\sigma^2)$，得

$$\chi^2 = \frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{\sigma^2} \sim \chi^2(n-1),$$

故当H_0为真时，

$$P\left\{\frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{\sigma_0^2} \geq \chi_\alpha^2(n-1)\right\} \leq P\left\{\frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{\sigma^2} \geq \chi_\alpha^2(n-1)\right\} = \alpha,$$

所以此检验问题的拒绝域为

$\boxed{\chi^2\text{检验法}}$

$$W=\left\{(x_1,\cdots,x_n):\frac{\sum_{i=1}^{n}(x_i-\overline{x})^2}{\sigma_0^2}\geq\chi_{\alpha}^2(n-1)\right\}。$$

同样，检验问题（3）的拒绝域为

$$W=\left\{(x_1,\cdots,x_n):\frac{\sum_{i=1}^{n}(x_i-\overline{x})^2}{\sigma_0^2}\leq\chi_{1-\alpha}^2(n-1)\right\}。$$

例1：某厂生产的某种型号的电池，其寿命（以小时计）长期以来服从方差 $\sigma^2 = 5000$ 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，测出其寿命的样本方差 $s^2 = 9200$。问根据这一数据能否推断这批电池的寿命的波动性较以往的有显著的变化 $(取\ \alpha = 0.02)$?

解：由题意知要求在 $\alpha = 0.02$ 下检验假设

$$H_0 : \sigma^2 = 5000, \quad H_1 : \sigma^2 \neq 5000.$$

现在 $n = 26, \chi^2_{\alpha/2}(n-1) = \chi^2_{0.01}(25) = 44.314,$

$\chi^2_{1-\alpha/2}(n-1) = \chi^2_{0.99}(25) = 11.524, \sigma_0^2 = 5000,$

检验的拒绝域为

$$\frac{(n-1)s^2}{\sigma_0^2} \geq 44.314 \quad 或 \quad \frac{(n-1)s^2}{\sigma_0^2} \leq 11.524$$

由观察值 $s^2 = 9200$ 得 $\dfrac{(n-1)s^2}{\sigma_0^2} = 46 > 44.314$，所以

拒绝 H_0，认为这批电池寿命的波动性较以往的有显著性的变化。

（二）、两个总体的情况

考虑 $X \sim N(\mu_1, \sigma_1^2), X_1, \cdots, X_{n_1}$ 为X的一个样本

$Y \sim N(\mu_2, \sigma_2^2), Y_1, \cdots, Y_{n_2}$ 为Y的一个样本

当μ_1, μ_2未知时，在水平α下，σ_1^2与σ_2^2的检验问题为

(1) $H_0: \sigma_1^2 = \sigma_2^2 \quad v.s. \quad H_1: \sigma_1^2 \neq \sigma_2^2$

（双侧检验）

$$(2) \quad H_0: \sigma_1^2 = \sigma_2^2 \quad v.s. \quad H_1: \sigma_1^2 > \sigma_2^2$$

$$H_0: \sigma_1^2 \le \sigma_2^2 \quad v.s. \quad H_1: \sigma_1^2 > \sigma_2^2$$

（右侧检验）

$$(3) \quad H_0: \sigma_1^2 = \sigma_2^2 \quad v.s. \quad H_1: \sigma_1^2 < \sigma_2^2$$

$$H_0: \sigma_1^2 \ge \sigma_2^2 \quad v.s. \quad H_1: \sigma_1^2 < \sigma_2^2$$

（左侧检验）

下面我们来讨论（1）的检验法则

由于S_1^2和S_2^2分别是σ_1^2和σ_2^2的无偏估计，因此当H_0为真时，$\dfrac{S_1^2}{S_2^2}$应接近1。若$\dfrac{S_1^2}{S_2^2}$接近于0或比1大得多时，就应该认为H_0为真时出现了小概率事件。于是取

$$P\left\{\frac{S_1^2}{S_2^2}\le k_1 \mid H_0\text{为真}\right\}+P\left\{\frac{S_1^2}{S_2^2}\ge k_2 \mid H_0\text{为真}\right\}=\alpha$$

其中k_1为适当小的正数，k_2为适当大的正数

为方便，我们取

$$P\left\{\frac{S_1^2}{S_2^2}\le k_1 \mid H_0\text{为真}\right\}=P\left\{\frac{S_1^2}{S_2^2}\ge k_2 \mid H_0\text{为真}\right\}=\frac{\alpha}{2}$$

当H_0为真时，

统计量

$$F=\frac{S_1^2}{S_2^2}\sim F(n_1-1,n_2-1),$$

由F分布的分位数可得

$$k_1=F_{1-\frac{\alpha}{2}}(n_1-1,n_2-1)\quad k_2=F_{\frac{\alpha}{2}}(n_1-1,n_2-1).$$

由此检验的拒绝域为

$$W=\left\{(x_1,\cdots,x_{n_1},y_1,\cdots,y_{n_2}):\frac{S_1^2}{S_2^2}\le F_{1-\frac{\alpha}{2}}(n_1-1,n_2-1)\right.$$

$$\left.\frac{S_1^2}{S_2^2}\ge F_{\frac{\alpha}{2}}(n_1-1,n_2-1)\right\}.$$

同样，我们可以得到右侧检验的拒绝域为

$$W=\left\{(x_1,\cdots,x_{n_1},y_1,\cdots,y_{n_2}):\frac{S_1^2}{S_2^2}\ge F_{\alpha}(n_1-1,n_2-1)\right\}.$$

左侧检验的拒绝域为

$$W=\left\{(x_1,\cdots,x_{n_1},y_1,\cdots,y_{n_2}):\frac{S_1^2}{S_2^2}\le F_{1-\alpha}(n_1-1,n_2-1)\right\}.$$

例5为比较两台自动机床的精度，分别取容量为10和8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：1.08, 1.10, 1.12, 1.14, 1.15, 1.25, 1.36, 1.38,1.40,1.42

车床乙：1.11, 1.12, 1.18, 1.22, 1.33, 1.35, 1.36, 1.38

在 α =0.1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 σ_1^2, σ_2^2，在 α=0.1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 $F = \dfrac{S_1^2}{S_2^2} \sim F(9,7)$

拒绝域为 W： $F \leq F_{1-\alpha/2}(9,7)$ 或 $F \geq F_{\alpha/2}(9,7)$

由样本值可计算得F的值为： $F=1.51$

查表得 $F_{\alpha/2}(9,7) = F_{0.05}(9,7) = 3.68$

$$F_{1-\alpha/2}(9,7) = F_{0.95}(9,7) = 1/3.29 = 0.304$$

由于 0.304<1.51<3.68， 故接受H_0，即认为这两台机床有同样的精度.

正态总体均值的检验

原假设H_0	检验统计量	备择假设H_1	拒绝域
$\mu \le \mu_0$ $\mu \ge \mu_0$ $\mu = \mu_0$ （σ^2已知）	$Z = \dfrac{\overline{X} - \mu_0}{\sigma / \sqrt{n}}$	$\mu > \mu_0$ $\mu < \mu_0$ $\mu \ne \mu_0$	$Z > z_\alpha$ $Z < -z_\alpha$ $\lvert Z \rvert > z_{\alpha/2}$
$\mu \le \mu_0$ $\mu \ge \mu_0$ $\mu = \mu_0$ （σ^2未知）	$T = \dfrac{\overline{X} - \mu_0}{S / \sqrt{n}}$	$\mu > \mu_0$ $\mu < \mu_0$ $\mu \ne \mu_0$	$T > t_\alpha(n-1)$ $T < -t_\alpha(n-1)$ $\lvert T \rvert > t_{\alpha/2}(n-1)$

正态总体均值的检验

原假设 H_0	检验统计量	备择假设 H_1	拒绝域
$\mu_1-\mu_2\le\delta$ $\mu_1-\mu_2\ge\delta$ $\mu_1-\mu_2=\delta$ (σ^2,σ_2^2已知)	$Z=\dfrac{\left(\overline{X}-\overline{Y}\right)-\delta}{\sqrt{\dfrac{\sigma_1^2}{m}+\dfrac{\sigma_2^2}{n}}}$	$\mu_1-\mu_2>\delta$ $\mu_1-\mu_2<\delta$ $\mu_1-\mu_2\ne\delta$	$Z>z_\alpha$ $Z<-z_\alpha$ $\lvert Z\rvert>z_{\alpha/2}$
$\mu_1-\mu_2\le\delta$ $\mu_1-\mu_2\ge\delta$ $\mu_1-\mu_2=\delta$ ($\sigma_1^2=\sigma_2^2$,未知)	$T=\dfrac{\left(\overline{X}-\overline{Y}\right)-\delta}{S_w\sqrt{\dfrac{1}{m}+\dfrac{1}{n}}}$	$\mu_1-\mu_2>\delta$ $\mu_1-\mu_2<\delta$ $\mu_1-\mu_2\ne\delta$	$T>t_\alpha(m+n-2)$ $T<-t_\alpha(m+n-2)$ $\lvert T\rvert>t_{\alpha/2}(m+n-2)$

其中 $S_w^2=\dfrac{(m-1)S_1^2+(n-1)S_2^2}{m+n-2}$

正态总体方差的检验

原假设H_0	检验统计量	备择假设 H_1	拒绝域
$\sigma^2 \le \sigma_0^2$ $\sigma^2 \ge \sigma_0^2$ $\sigma^2 = \sigma_0^2$ $(\mu$未知$)$	$\chi^2 = \dfrac{(n-1)S^2}{\sigma_0^2}$	$\sigma^2 > \sigma_0^2$ $\sigma^2 < \sigma_0^2$ $\sigma^2 \ne \sigma_0^2$	$\chi^2 > \chi_\alpha^2(n-1)$ $\chi^2 < \chi_{1-\alpha}^2(n-1)$ $\chi^2 > \chi_{\alpha/2}^2(n-1)$或 $\chi^2 < \chi_{1-\alpha/2}^2(n-1)$
$\sigma_1^2 \le \sigma_2^2$ $\sigma_1^2 \ge \sigma_2^2$ $\sigma_1^2 = \sigma_2^2$ $(\mu_1,\mu_2$未知$)$	$F = \dfrac{S_1^2}{S_2^2}$	$\sigma_1^2 > \sigma_2^2$ $\sigma_1^2 < \sigma_2^2$ $\sigma_1^2 \ne \sigma_2^2$	$F > F_\alpha(n_1-1, n_2-1)$ $F < F_{1-\alpha}(n_1-1, n_2-1)$ $F > F_{\alpha/2}(n_1-1, n_2-1)$或 $F < F_{1-\alpha/2}(n_1-1, n_2-1)$

正态总体方差的检验

原假设H_0	检验统计量	备择假设 H_1	拒绝域
$\mu_0 \le 0$ $\mu_0 \ge 0$ $\mu_0 = 0$ (成对数据)	$T = \dfrac{\overline{D} - 0}{S_D / \sqrt{n}}$	$\mu_0 > 0$ $\mu_0 < 0$ $\mu_0 \ne 0$	$T > t_\alpha(n-1)$ $T < t_{1-\alpha}(n-1)$ $\lvert T \rvert > t_{\alpha/2}(n-1)$

§7.4 置信区间与假设检验之间的关系

置信区间与假设检验之间存在着密切的关系，我们仅考察置信区间与双边检验之间的对应关系。

设$X_1,X_2,\cdots,X_n$是一个来自总体的样本，$x_1,x_2,\cdots,x_n$是相应的样本值，Θ是参数θ的可能取值范围。

设$\left(\underline{\theta}(X_1,\cdots,X_n),\overline{\theta}(X_1,\cdots,X_n)\right)$是参数$\theta$的一个置信水平为$1-\alpha$的置信区间，则对于$\theta\in\Theta$有

$$P_\theta\left(\underline{\theta}(X_1,\cdots,X_n)<\theta<\overline{\theta}(X_1,\cdots,X_n)\right)\geq 1-\alpha \quad (1)$$

考虑显著水平为α的双边检验

$$H_0:\theta=\theta_0 \quad v.s. \quad H_1:\theta\neq\theta_0 \qquad (2)$$

由（1）

$$P_{\theta_0}\left(\underline{\theta}(X_1,\cdots,X_n)<\theta_0<\overline{\theta}(X_1,\cdots,X_n)\right)\geq 1-\alpha$$

即有

$$P_{\theta_0}\left\{\left(\theta_0\leq\underline{\theta}(X_1,\cdots,X_n)\right)\cup\left(\theta_0\geq\overline{\theta}(X_1,\cdots,X_n)\right)\right\}\leq\alpha$$

按显著水平为α的假设检验的拒绝域的定义，检验（2）的拒绝域为

$$\theta_0\leq\underline{\theta}(X_1,\cdots,X_n) \quad \text{或} \quad \theta_0\geq\overline{\theta}(X_1,\cdots,X_n)$$

这就是说，当我们要检验（2）时，先求出θ的置信水平为$1-\alpha$的置信区间$(\underline{\theta},\overline{\theta})$。若$\theta_0 \in (\underline{\theta},\overline{\theta})$，则接受$H_0$；若$\theta_0 \notin (\underline{\theta},\overline{\theta})$，则拒绝$H_0$.

反之，对于$\forall \theta_0 \in \Theta$，考虑显著水平为$\alpha$的假设检验问题

$$H_0:\theta=\theta_0 \quad v.s. \quad H_1:\theta \neq \theta_0$$

假设它的接受域为

$$\underline{\theta}(X_1,\cdots,X_n) < \theta_0 < \overline{\theta}(X_1,\cdots,X_n)$$

即有

$$P_{\theta_0}\left(\underline{\theta}(X_1,\cdots,X_n)<\theta_0<\bar{\theta}(X_1,\cdots,X_n)\right)\geq 1-\alpha$$

由θ_0的任意性，由上式知，对$\forall\theta\in\Theta$，有

$$P_{\theta}\left(\underline{\theta}(X_1,\cdots,X_n)<\theta<\bar{\theta}(X_1,\cdots,X_n)\right)\geq 1-\alpha$$

因此，$\left(\underline{\theta}(X_1,\cdots,X_n),\bar{\theta}(X_1,\cdots,X_n)\right)$是参数$\theta$的置信水平为$1-\alpha$的置信区间。

这就是说，为要求出参数θ的置信水平为$1-\alpha$的置信区间，我们先求出显著性水平为α的假设检验问题：$H_0:\theta=\theta_0 \quad H_1:\theta\neq\theta_0$的接受域

$$\underline{\theta}(X_1,\cdots,X_n)<\theta_0<\overline{\theta}(X_1,\cdots,X_n),$$

那么$\left(\underline{\theta}(X_1,\cdots,X_n),\ \overline{\theta}(X_1,\cdots,X_n)\right)$就是$\theta$的置信水平为$1-\alpha$的置信区间。

同样地，我们也可以推出单侧置信区间与单侧假设检验也有类似的关系。

(1) 置信水平为$1-\alpha$的单侧置信区间$(-\infty,\overline{\theta}(X_1,\cdots,X_n))$与显著性水平为$\alpha$的左侧检验问题

$$H_0:\theta\ge\theta_0,H_1:\theta<\theta_0$$

有类似的对应关系。即若已求得单侧置信区间$(-\infty,\overline{\theta}(X_1,\cdots,X_n))$，则当$\theta_0\in(-\infty,\overline{\theta}(X_1,\cdots,X_n))$时接受$H_0$，当$\theta_0\notin(-\infty,\overline{\theta}(X_1,\cdots,X_n))$时拒绝$H_0$。反之，若已求得检验问题$H_0:\theta\ge\theta_0,H_1:\theta<\theta_0$的接受域为$-\infty<\theta_0\le\overline{\theta}(X_1,\cdots,X_n)$，则可得$\theta$的一个单侧置信区间$(-\infty,\overline{\theta}(X_1,\cdots,X_n))$。

(2) 置信水平为$1-\alpha$的单侧置信区间$(\underline{\theta}(X_1,\cdots,X_n),\infty)$与显著性水平为$\alpha$的右侧检验问题

$$H_0:\theta\le\theta_0, H_1:\theta>\theta_0$$

有类似的对应关系。即若已求得单侧置信区间$(\underline{\theta}(X_1,\cdots,X_n),\infty)$，则当$\theta_0\in(\underline{\theta}(X_1,\cdots,X_n),\infty)$时接受$H_0$，当$\theta_0\notin(\underline{\theta}(X_1,\cdots,X_n),\infty)$时拒绝$H_0$。反之，若已求得检验问题$H_0:\theta\le\theta_0, H_1:\theta>\theta_0$的接受域为$\underline{\theta}(X_1,\cdots,X_n)\le\theta_0<\infty$，则可得$\theta$的一个单侧置信区间$(\underline{\theta}(X_1,\cdots,X_n),\infty)$。

例1：设$X \sim N(\mu,1)$, μ未知，$\alpha=0.05, n=16$，且由一样本$\bar{x}=5.20$，于是得到参数μ的一个置信水平为0.95的置信区间

$$(\bar{x}-\frac{1}{\sqrt{16}}z_{0.025}, \bar{x}+\frac{1}{\sqrt{16}}z_{0.025})=(4.71, 5.69).$$

现在考虑检验问题$H_0: \mu=5.5$, $H_1: \mu \neq 5.5$。

由于$5.5 \in (4.71, 5.69)$，故接受H_0。

例2： 设$X \sim N(\mu,1)$, μ未知，$\alpha = 0.05, n = 16$，且由一样本$\bar{x} = 5.20$。试求右侧检验问题$H_0: \mu \le \mu_0$, $H_1: \mu > \mu_0$的接受域，并求μ的单侧置信下限$(\alpha = 0.05)$.

解：

检验问题的拒绝域为$z = \dfrac{\bar{x} - \mu_0}{1/\sqrt{16}} \ge z_{0.05}$，或即$\mu_0 \le 4.79$。

于是检验问题的接受域$\mu_0 > 4.79$。这样就得到μ的单侧置信区间$(4.79, \infty)$，单侧置信下限$\underline{\mu} = 4.79$。