

第8章 数理统计的基本概念

8.1 假设检验的基本概念

8.2 正态总体参数的假设检验

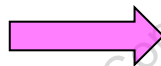
8.3 单侧假设检验

8.4 非正态总体的大样本检验及小样本检验

8.5 成对数据的假设检验

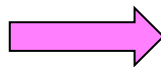
8.6 分布假设检验

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
的方法处理**

8.1 假设检验的基本概念

1. 假设检验的基本思想和方法

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**

➤ **参数假设检验：**总体分布已知，检验关于未知参数的某个假设

➤ **非参数假设检验：**总体分布未知时的建设检验问题

例1 某洗衣粉厂用自动包装机包装洗衣粉，每袋**标准重量**为 $\mu_0=500\text{g}$. 长期实践表明，袋装重量（单位：g）**服从正态分布**，且标准差 σ 为 2g 比较稳定. 某日，为了检验包装机是否正常，在装好的袋中随机抽取 **7袋**，称的净重为 501.8、502.4、499、500.3、504.5、498.2、505.6，问机器是否正常？

为此提出如下假设：

$$H_0 : \mu = 500$$

称为**原假设**或**零假设**

原假设的对立面：

$$H_1 : \mu \neq 500$$

称为**备择假设**

假设检验的任务

必须在原假设与备择假设之间作一选择

那么，如何判断原假设 H_0 是否成立呢？

由于 μ_0 是正态分布的期望值，它的估计量是样本均值 $\bar{X}$ ，因此，可以根据 $\bar{X}$ 与 μ_0 的差距

$|\bar{X} - \mu_0|$ 来判断 H_0 是否成立

当 $|\bar{X} - \mu_0|$ 较小时，可以认为 H_0 是成立的；

当 $|\bar{X} - \mu_0|$ 较大时，应认为 H_0 不成立，即生产已不正常

◆ 较大、较小是一个相对的概念，合理的界限在何处？应由什么原则来确定？

➤ 问题归结为对差异作定量的分析，以确定其性质

- 差异可能是由抽样的随机性引起的，称为“抽样误差”或 随机误差，这种误差反映偶然、非本质的因素所引起的随机波动
- 然而，这种随机性的波动是有一定限度的，如果差异超过了这个限度，则我们就不能用抽样的随机性来解释了
- 必须认为这个差异反映了事物的本质差别，即反映了生产已不正常，这种差异称作“系统误差”

➤ **问题是，根据所观察到的差异，如何判断它究竟是由于偶然性在起作用，还是生产确实不正常？**

- 即差异是“抽样误差”还是“系统误差”所引起的？
- 这里需要给出一个量的界限

➤ **问题是：如何给出这个量的界限？**

- 这里用到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

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

在提出原假设 H_0 后，如何作出接受和拒绝 H_0 的结论呢？

提出假设 $H_0 : \mu = 500 \quad \longleftrightarrow \quad H_1 : \mu \neq 500$

由于方差已知，选**检验统计量** $U = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}} \sim N(0, 1)$

它能衡量差异 $|\bar{X} - \mu_0|$ 大小且分布已知

对给定的显著性水平 α ，可以在 $N(0, 1)$ 表中查到分位点的值 $u_{\alpha/2}$ ，使

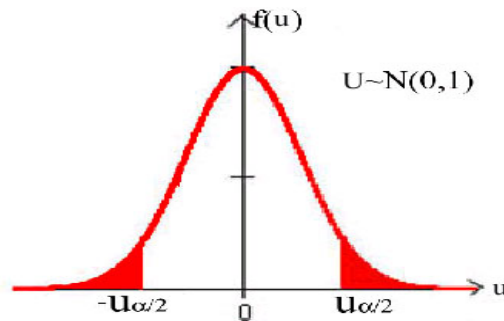
$$P\{|U| > u_{\alpha/2}\} = \alpha$$

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

也就是说，“ $|U| > \mu_{\alpha/2}$ ”是一个小概率事件

故我们可以取拒绝域为：

$$W: |U| > \mu_{\alpha/2}$$



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

这里所依据的逻辑是：

如果 H_0 是对的，那么衡量差异大小的某个统计量落入区域 W (拒绝域)是个小概率事件。如果该统计量的观测值落入 W ，也就是说， H_0 成立下的小概率事件发生了，那么就认为 H_0 不可信而否定它。否则我们就不能否定 H_0 （只好接受它）

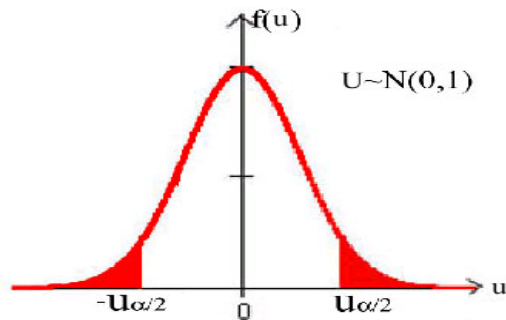
不否定 H_0 并不是肯定 H_0 一定对，而只是说差异还不够显著，还没有达到足以否定 H_0 的程度。

所以假设检验又叫“显著性检验”

如果显著性水平 α 取得很小，则拒绝域 W 也会比较小。

其产生的后果是： H_0 难于被拒绝

如果在 α 很小的情况下 H_0 仍被拒绝了，则说明实际情况很可能与之有显著差异。



基于这个理由，人们常把 $\alpha=0.05$ 时拒绝 H_0 称为是显著的，而把在 $\alpha=0.01$ 时拒绝 H_0 称为是高度显著的。

2. 假设检验的一般步骤

在上面的例子的叙述中，我们已经初步介绍了假设检验的基本思想和方法. 我们再结合例1，进一步说明

现取容量为7的样本，其均值计算为 $\bar{x} = 501.69$

若原假设正确，则 $\bar{X} \sim N(500, 2^2 / 7)$

因而 $E(\bar{X}) = 500$ ，即 $\bar{X}$ 偏离500不应该太远

故 $\left| \frac{\bar{X} - 500}{2 / \sqrt{7}} \right|$ 取较大值是小概率事件. 因此

可以确定一个常数 k 使得 $P\left(\left| \frac{\bar{X} - 500}{2 / \sqrt{7}} \right| > k\right) = \alpha$

取 $\alpha = 0.05$ ，则 $k = u_{\alpha/2} = u_{0.025} = 1.96$

由 $\left| \frac{\bar{X} - 500}{2 / \sqrt{7}} \right| > 1.96 \longrightarrow \bar{X} > 501.48 \text{ 或 } \bar{X} < 498.52$

即区间 $(-\infty, 498.52)$ 与 $(501.48, +\infty)$ 为检验的拒绝域

称 $\bar{X}$ 的取值区间 $(498.52, 501.48)$ 为检验的接受域 (实际上没有理由拒绝)

现 $\bar{x} = 501.69$ 落入拒绝域, 则拒绝原假设 $H_0: \mu = 500$

假设检验的基本步骤:

1. 根据实际问题的要求提出原假设 H_0 及备择假设 H_1 ;
2. 确定检验统计量;
3. 对给定显著水平 α (一般取 $\alpha=0.05$, 或 0.01) , 根据统计量的分布求出拒绝域;
4. 决策: 将样本值代入统计量, 若统计量的值落入拒绝域, 则拒绝 H_0 , 否则接受 H_0 .

3. 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
是下述小概率原理

不是一定不发生

小概率事件在一次试验中基本上不会发生

- ◆ 如果 H_0 成立，但统计量的实测值落入拒绝域，从而作出拒绝 H_0 的结论，那就犯了“以真为假”的错误
- ◆ 如果 H_0 不成立，但统计量的实测值未落入拒绝域，从而没有作出拒绝 H_0 的决策，即接受了错误的 H_0 ，那就犯了“以假为真”的错误

假设检验的两类错误

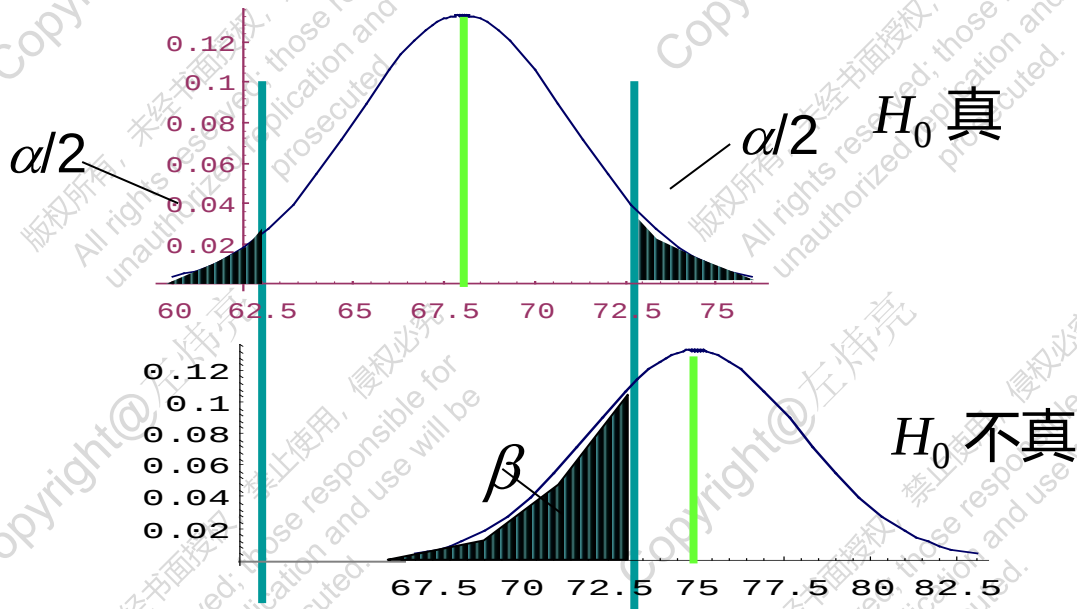
决策	实际情况	
	H_0 为真	H_0 不真
拒绝 H_0	第一类错误	正确
接受 H_0	正确	第二类错误

犯两类错误的概率:

$$P\{\text{拒绝}H_0 \mid H_0\text{为真}\} = \alpha$$

$$P\{\text{接受}H_0 \mid H_0\text{不真}\} = \beta$$

显著性水平 α 为犯第一类错误的概率



两类错误的概率关系

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

说明：关于检验标准(显著水平) α

- 一般说来，原假设 H_0 是经过周密调查和考虑的，所以否定 H_0 要谨慎，这就要求 α 取的要小些。反过来，也表明小概率 α 取得越小，一旦否定 H_0 ，就越有说服力
- 检验标准 α ，取值不同，就决定着不同的拒绝域。因此对于同一组样本值 $x_1, x_2, \dots, x_n$ 来说，可以得到相反的决策
- 只对犯第I类错误的概率加以控制，而不考虑犯第II类错误的检验问题，称为显著性检验问题

8.2正态总体参数的假设检验

1. 单个总体 $N(\mu, \sigma^2)$ 均值 μ 的检验

设总体 $X \sim N(\mu, \sigma^2)$, μ, σ^2 未知, $X_1, X_2, \dots, X_n$ 是来自 X 的样本, 给定显著性水平 α (显著水平)

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

1). σ^2 已知 (为 σ_0^2)

取统计量:

$$U = \frac{\bar{X} - \mu_0}{\sigma_0 / \sqrt{n}}$$

—— u 检验法

拒绝域:

$$|u| \geq u_{\alpha/2}$$

例2 在某年级学生中抽测9名跳远成绩，得样本均值 $\bar{x}=4.38$ 。设跳远成绩 X 服从正态分布且 $\sigma=0.3$ ，问是否可以认为该年级学生跳远平均成绩为 $\mu=4.40$ 米 ($\alpha=0.10$)。

解: 根据题意提出假设 $H_0: \mu = 4.40$ $H_1: \mu \neq 4.40$

因为 σ 已知，取统计量: $U = \frac{\bar{X} - 4.40}{\sigma / \sqrt{n}}$ 拒绝域为 $|u| \geq u_{\alpha/2}$

查标准正态分布表得 $u_{\alpha/2} = u_{0.05} = 1.645$

经计算得 $\left| \frac{\bar{x} - 4.40}{0.3/3} \right| = 0.2 < u_{0.05} = 1.645$

所以接受 H_0 ，认为 $\mu=4.40$ 米，即可以认为该年级学生跳远平均成绩为4.40米

例3 某工厂生产的固体燃料推进器的燃烧率服从正态分布

$N(\mu, \sigma^2)$, $\mu=40\text{cm/s}$, $\sigma=2\text{cm/s}$

现在用新方法生产了一批推进器. 从中随机取 $n=25$ 只, 测得燃烧率的样本均值为 $\bar{x} = 41.25\text{cm/s}$

设在新方法下总体均方差仍为 2cm/s , 问这批推进器的燃烧率是否较以往生产的推进器的燃烧率有显著的提高?

取显著性水平 $\alpha=0.05$

解: 提出假设 $H_0: \mu \leq 40, H_1: \mu > 40$

取统计量
$$U = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}}$$

拒绝域为 $W: u_{\alpha/2} = u_{0.05} = 1.645$

并由样本值计算得统计

$$U = \frac{41.25 - 40}{2 / \sqrt{25}} = 3.125 > 1.645$$

落入拒绝域

故拒绝原假设 H_0 , 即燃烧率较以往生产的推进器的燃烧率有显著的提高

例4 某林场内造了杨树丰产林，五年后调查50株得平均树高为 $\bar{x}=9.2\text{m}$ ，假设树高服从正态分布，其标准差为 1.6m 时，在 $\alpha=0.05$ 的显著性水平下，推断丰产林的树高是否低于10米？

解：提出假设 $H_0: \mu \geq 10$, $H_1: \mu < 10$

取统计量
$$U = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}}$$

拒绝域为 $W: u < -u_{0.05} = -1.645$

落入拒绝域

并由样本值计算得统计
$$U = \frac{9.2 - 10}{1.6 / \sqrt{49}} = -3.5 < -1.645$$

故拒绝**原假设** H_0 ，即丰产林的树高低于10米

2). σ^2 未知

取统计量
$$t = \frac{\bar{X} - \mu_0}{S / \sqrt{n}}$$

—— t 检验法

拒绝域 $|t| \geq t_{\alpha/2}(n-1)$

例5 对一批新的某种液体存贮罐进行耐裂试验,抽测5个,得爆破压力数据为 (单位: 斤/寸²) : 545, 530, 545, 550, 545. 根据经验, 爆压可认为是服从正态分布的, 且过去该种液体存贮罐的平均爆压为549斤/寸², 问这批新罐的平均爆压与过去有无显著差别 ($\alpha = 0.05$)

解: 提出假设: $H_0: \mu = \mu_0 = 549$, $H_1: \mu \neq 549$

σ^2 未知, 故检验统计量为 $t = \frac{\bar{X} - \mu_0}{S / \sqrt{n}}$

拒绝域为 $|t| \geq t_{\alpha/2}(n-1) = t_{0.025}(4) = 2.776$

由样本算得 $\bar{x} = 543, s^2 = (7.58)^2$

$$|t| = \left| \frac{\bar{x} - 549}{s / \sqrt{n}} \right| = \left| \frac{543 - 549}{7.58 / \sqrt{5}} \right| \approx 1.77 < 2.776$$

所以接受 H_0 , 认为新罐的平均爆压与旧罐的无显著差异

例6 已知精料养鸡时， 经若干天鸡的平均重量为4斤,今对一批鸡改用粗料饲养， 同时改善饲养方法， 经同样长的饲养期， 随机抽测10只， 得重量(斤)数据如下： 3.7, 3.8, 4.1, 3.9, 4.6, 4.7, 5.0, 4.5, 4.3, 3.8. 经验表明， 同一批鸡的重量 X 服从正态分布， 试推断这一批鸡的平均重量是否显著提高 ($\alpha = 0.10$)

解: 提出假设 $H_0: \mu \leq 4, H_1: \mu > 4$

σ^2 未知, 故取检验统计量为 $t = \frac{\bar{X} - \mu_0}{S / \sqrt{n}}$

得拒绝域 $t = \frac{\bar{x} - \mu_0}{s / \sqrt{n}} \geq t_{\alpha}(n-1)$ 查表得 $t_{0.1}(9) = 1.383$

经计算得 $\bar{x} = 4.24, s^2 = 0.448^2$

$$t = \frac{\bar{x} - 4}{s / \sqrt{n}} \approx 1.694 > 1.383$$

所以拒绝 H_0 , 接受 H_1 , 即在显著水平 $\alpha = 0.10$ 下, 认为这批鸡的平均重量显著提高

3). σ^2 的检验问题

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

取检验统计量: $\chi^2 = \frac{(n-1)S^2}{\sigma_0^2}$

—— χ^2 检验法

拒绝域: $\chi^2 \leq \chi_{1-\alpha/2}^2(n-1) \quad \chi^2 \geq \chi_{\alpha/2}^2(n-1)$

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

由于 S^2 是 σ^2 的无偏估计，自然想到将 S^2 与 σ_0^2 作比较。比值一般来说应在1附近摆动，而不应过分大于1或过分小于1

由定理6.2知，当 H_0 为真时 $\chi^2 = \frac{(n-1)S^2}{\sigma_0^2} \sim \chi^2(n-1)$

$$P\{\text{拒}H_0 | H_0 \text{真}\} = P\left\{\left(\frac{(n-1)S^2}{\sigma_0^2} \leq k_1\right) \cup \left(\frac{(n-1)S^2}{\sigma_0^2} \geq k_2\right)\right\} = \alpha$$

故得 $k_1 = \chi_{1-\alpha/2}^2(n-1), \quad k_2 = \chi_{\alpha/2}^2(n-1)$

得拒绝域为 $\frac{(n-1)s^2}{\sigma_0^2} \leq \chi_{1-\alpha/2}^2(n-1)$ 或 $\frac{(n-1)s^2}{\sigma_0^2} \geq \chi_{\alpha/2}^2(n-1)$

例6 某厂生产的某种型号的电池，其寿命长期以来服从方差为 $\sigma^2=5000$ (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，测出寿命的样本方差 $S^2=9200$ (小时²)。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: 根据题意提出假设

$$H_0: \sigma^2 = \sigma_0^2 = 5000, H_1: \sigma^2 \neq \sigma_0^2$$

检验统计量: $\chi^2 = \frac{(n-1)S^2}{\sigma_0^2} \quad n=26, \sigma_0^2=5000$

拒绝域为 $\chi^2 \leq \chi_{n-1}^2(1-\alpha/2)$ 或 $\chi^2 \geq \chi_{n-1}^2(\alpha/2)$

查表得: $\chi_{n-1}^2(1-\alpha/2) = \chi_{25}^2(0.99) = 11.523$ $\chi_{n-1}^2(\alpha/2) = \chi_{25}^2(0.01) = 44.313$

由样本观察值算得 $\chi^2 = \frac{(n-1)s^2}{\sigma_0^2} = 46 > 44.313$

所以拒绝 H_0 , 在显著水平 $\alpha=0.02$ 下, 可以认为这批电池的寿命的波动性较以往有显著的变化

2. 两个总体 $N(\mu_1, \sigma_1^2)$ 与 $N(\mu_2, \sigma_2^2)$ 的情形

1) 两总体均值差的检验($\sigma_1^2 = \sigma_2^2$ 未知)

取检验统计量
$$t = \frac{(\bar{X} - \bar{Y}) - \delta}{S_{\omega} \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \quad \text{—— } t\text{检验法}$$

检验假设 $H_0: \mu_1 - \mu_2 = \delta, \quad H_1: \mu_1 - \mu_2 \neq \delta$

拒绝域: $|t| \geq t_{\alpha/2}(n_1 + n_2 - 2)$

例 7 某地某年高考后随机抽得15名男生、12名女生的物理考试成绩如下：

男生：49 48 47 53 51 43 39 57 56 46 42 44 55 44 40

女生：46 40 47 51 43 36 43 38 48 54 48 34

这27名学生的成绩能说明这个地区男、女生的物理考试成绩不相上下吗？（显著水平 $\alpha=0.05$ ）

解：设男女生的物理考试成绩分别近似服从正态分布

$$X \sim N(\mu_1, \sigma^2), Y \sim N(\mu_2, \sigma^2)$$

检验 $H_0: \mu_1 = \mu_2, H_1: \mu_1 \neq \mu_2 \quad t_{25}(0.025) = 2.060$

算得 $\bar{x} = 47.6, \bar{y} = 44, (n_1 - 1)s_1^2 = 469.6, (n_2 - 1)s_2^2 = 412,$

$$|t| = 1.566 < t_{25}(0.025)$$

接受原假设

例8 一药厂生产一种新的止痛片，厂方希望验证服用新药片后至开始起作用的时间间隔较原有止痛片至少缩短一半，因此厂方提出需检验假设

$$H_0: \mu_1 \leq 2\mu_2, \quad H_1: \mu_1 > 2\mu_2$$

这里 μ_1 、 μ_2 分别是服用原有止痛片和服用新止痛片后至起作用的时间间隔的总体的均值. 设两总体均为正态且方差分别为已知值 σ_1 、 σ_2 . 现分别在两总体中取一样本 $X_1, X_2, \dots, X_n$ 和 $Y_1, Y_2, \dots, Y_m$, 设两个样本独立. 试给出上述假设检验问题的检验统计量及拒绝域. (取显著性水平为 α)

检验统计量

$$Z = \frac{\bar{X} - 2\bar{Y}}{\sqrt{\frac{\sigma_1^2}{n} + \frac{4\sigma_2^2}{m}}}$$

其拒绝域为

$$\{Z \geq z_\alpha\}$$

3. 双正态总体方差的检验--F检验

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

取检验统计量 $F = S_1^2 / S_2^2$

拒绝域: $F \geq F_{n_1-1, n_2-1}(\alpha/2) \quad \text{或} \quad F \leq F_{n_1-1, n_2-1}(1-\alpha/2)$

8.3单侧假设检验

设总体 $X \sim N(\mu, \sigma^2)$, μ, σ^2 未知, $X_1, X_2, \dots, X_n$ 是来自 X 的样本, 给定显著性水平 α (显著水平)

1. σ^2 已知

取统计量: $U = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}}$ —— u 检验法

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

拒绝域: $|u| \geq u_{\alpha/2}$

b) $H_0: \mu \leq \mu_0, H_1: \mu > \mu_0$

拒绝域: $u \geq u_{\varepsilon}$

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

拒绝域: $u \leq -u_{\varepsilon}$

2. σ^2 未知

取统计量

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

—— t 检验法

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

拒绝域: $|t| \geq t_{\alpha/2}(n-1)$

b) $H_0: \mu \leq \mu_0, H_1: \mu > \mu_0$

拒绝域: $t \geq t_{\alpha}(n-1)$

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

拒绝域: $t \leq -t_{\alpha}(n-1)$

3). σ^2 的检验问题

取检验统计量: $\chi^2 = \frac{(n-1)S^2}{\sigma_0^2}$ — χ^2 检验法

(1) 双边检验:

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

拒绝域:

$$\chi^2 \leq \chi_{n-1}^2(1-\alpha/2) \\ \text{或} \chi^2 \geq \chi_{n-1}^2(\alpha/2)$$

(2) 右边检验:

$$H_0: \sigma^2 \leq \sigma_0^2, H_1: \sigma^2 > \sigma_0^2$$

拒绝域:

$$\chi^2 \geq \chi_{n-1}^2(\alpha)$$

(3) 左边检验:

$$H_0: \sigma^2 \geq \sigma_0^2, H_1: \sigma^2 < \sigma_0^2$$

拒绝域:

$$\chi^2 \leq \chi_{n-1}^2(1-\alpha)$$

2. 两个总体 $N(\mu_1, \sigma_1^2)$ 与 $N(\mu_2, \sigma_2^2)$ 的情形

1) 两总体均值差的检验($\sigma_1^2 = \sigma_2^2$ 未知)

取检验统计量
$$t = \frac{(\bar{X} - \bar{Y}) - \delta}{S_{\omega} \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \quad \text{—— } t\text{检验法}$$

检验假设 $H_0: \mu_1 - \mu_2 = \delta, \quad H_1: \mu_1 - \mu_2 \neq \delta$ (双边) 拒绝域:

$$|t| \geq t_{n_1+n_2-2}(\alpha/2)$$

$$H_0: \mu_1 - \mu_2 \leq \delta, H_1: \mu_1 - \mu_2 > \delta \quad (\text{右边})$$

$$t \geq t_{n_1+n_2-2}(\alpha)$$

$$H_0: \mu_1 - \mu_2 \geq \delta, H_1: \mu_1 - \mu_2 < \delta \quad (\text{左边})$$

$$t \leq -t_{n_1+n_2-2}(\alpha)$$

3. 双正态总体方差的检验--F检验

检验

$$\begin{array}{ll} H_0 : \sigma_1^2 = \sigma_2^2, & H_1 : \sigma_1^2 \neq \sigma_2^2 \quad (\text{双边}) \\ H_0 : \sigma_1^2 \leq \sigma_2^2, & H_1 : \sigma_1^2 > \sigma_2^2 \quad (\text{右边}) \\ H_0 : \sigma_1^2 \geq \sigma_2^2, & H_1 : \sigma_1^2 < \sigma_2^2 \quad (\text{左边}) \end{array}$$

取检验统计量 $F = S_1^2 / S_2^2$

拒绝域: $F \geq F_{n_1-1, n_2-1}(\alpha/2)$ 或 $F \leq F_{n_1-1, n_2-1}(1-\alpha/2)$

$$F \geq F_{n_1-1, n_2-1}(\alpha) \quad (\text{右边})$$

$$F \leq F_{n_1-1, n_2-1}(1-\alpha) \quad (\text{左边})$$

8.4非正态总体的大样本及小样本检验

例 某厂有一大批产品需要出厂，其次品率 p 未知，按规定标准，次品率 p 不超过某个已知数 p_0 方可出厂. 今在其中任意选取 n 件产品进行检验，发现有次品 m 件，问这批产品能否出厂？假定 n 比较大.

解： 总体 X 的分布是 $(0-1)$ 分布 $B(1,p)$

要检验假设 $H_0: p \leq p_0, \quad H_1: p > p_0$

总体均值 $E(X)=p$ ，想到利用样本均值 $\bar{X} = m/n$ 来检验，取

$$U = \frac{\bar{X} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{m - np_0}{\sqrt{np_0(1-p_0)}}$$

因为 $\bar{x}$ 是 p 的无偏估计量，所以， H_0 为真时， u 应偏小，当 H_1 为真时， u 应偏大，故拒绝域的形式为

$$u = \frac{m - np_0}{\sqrt{np_0(1-p_0)}} \geq k \quad (k \text{ 待定})$$

根据中心极限定理，当 n 很大时， U 近似地服从标准正态分布 $N(0,1)$ ，有

$$P(U \geq u_\alpha) \approx \alpha$$

拒绝域：
$$u = \frac{m - np_0}{\sqrt{np_0(1-p_0)}} \geq u_\alpha$$

或：
$$m \geq np_0 + u_\alpha \sqrt{np_0(1-p_0)}$$

例 某厂家声称他们生产的电子元件平均寿命不低于2000h, 现随机抽取10个这种电子元件, 测得寿命 (单位: 小时) : 717, 1996, 2532, 4213, 1574, 5437, 925, 3654, 6273, 2150, 在检验水平 $\alpha=0.05$ 下能否相信厂家的说法?

解: 总体 X 的分布是 $Exp(\lambda)$, λ 未知, 要检验

要检验假设 $H_0: \lambda \leq \lambda_0, H_1: \lambda > \lambda_0$

由于 $2n\lambda\bar{X} \sim \chi^2(2n)$, 因此拒绝域 $\chi^2 \leq \chi_{1-\alpha}^2(2n)$

计算得 $\chi^2 = 29.471 > \chi_{0.95}^2(20) = 10.851$

故接受原假设, 即认为电子元件的平均寿命超过2000h

8.5成对数据的假设检验

例 某培训班声称经过培训，学员掷实心球的成绩能够提高1米以上. 为检验他们的说法是否符合实际，随机选9人，记录他们的培训前后成绩为

培训前成绩 x_i	9.63	7.61	6.28	8.32	5.40	5.82	6.89	8.17	5.78
培训前成绩 y_i	10.24	8.75	7.05	8.91	6.51	6.58	8.02	9.08	6.32
差值 $x_i - y_i$	0.61	1.14	0.77	0.59	1.11	0.76	1.13	0.91	0.54

设差值数据来自正态总体，在检验水平0.05下，是否可以认为改培训班的说法是对的？

解：总体 X 的分布是 $Z \sim N(\mu, \sigma^2)$ ，要检验

$$H_0: \mu \geq 1, \quad H_1: \mu < 1$$

取

$$T = \frac{\bar{Z} - 1}{S_z / \sqrt{9}} \sim t(8)$$

拒绝域 $t \leq -t_{0.05}(8)$

计算得 $\bar{z} = 0.84, s_z = 0.242, t = -1.98 < -t_{0.05}(8) = -1.8595$

故拒绝原假设，即认为培训后平均成绩提高不到1m

8.6 分布假设检验

定理8.1 设一个随机试验的 r 个结果 $A_1, A_2, \dots, A_r$ 构成互斥完备事件群，在一次试验中它们发生的概率分别为 $p_1, p_2, \dots, p_r$ ，其中 $p_i > 0 (i=1, 2, \dots, r)$ ，且 $\sum_{i=1}^r p_i = 1$ ，以 m_i 表示在 n 次独立重复试验中 A_i 发生的次数，那么，当 $n \rightarrow \infty$ 时，随机变量

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

的分布收敛于自由度为 $r-1$ 的 χ^2 分布

这个定理称为**皮尔逊定理**

χ^2 拟合检验法

1. 理论分布完全已知的情形

设 $(X_1, X_2, \dots, X_n)$ 是来自总体 X 的样本, $F(x)$ 是一个完全已知的分布函数, 今要利用此样本假设

H_0 : X 的分布函数为 $F(x)$, H_1 : X 的分布函数不是 $F(x)$

χ^2 拟合检验法的基本思想和步骤:

(1) 把总体 X 的一切可能值的集合 Ω 划分成有限个, 比如 r 个互相不相交的子集 $\Omega_1, \Omega_2, \dots, \Omega_r$, 记 $A_i = \{X \in \Omega_i\}$ ($i=1, 2, \dots, r$), 则 $A_1, A_2, \dots, A_r$ 构成互斥完备事件群

(2) 在原假设 H_0 为真的前提下, 计算事件 A_i 的概率 $P(A_i)=p_i$ ($i=1,2,\dots,r$), 显然 $\sum_{i=1}^r p_i = 1$ 计算, 不妨设 $p_i > 0$ (否则可适当改变划分), 从而可得, 在对总体 X 的 n 次独立观察中, 事件 A_i 发生的理论频率为 np_i ($i=1,2,\dots,r$)

(3) 统计出样本值 $(x_1, x_2, \dots, x_n)$ 中落入 Ω_i 的频数, 即事件 A_i 发生的实际频数 m_i ($i=1,2,\dots,r$)

(4) 考虑统计量 $\chi^2 = \sum_{i=1}^r \frac{(m_i - np_i)^2}{np_i}$, χ^2 是样本的实际频数 m_i 对理论频数 np_i 偏差 $m_i - np_i$ 的加权平分和, 它值的大小刻画理论分布函数 $F(x)$ 与样本值 $(x_1, x_2, \dots, x_n)$ 的拟合程度, 因此 χ^2 可以作为检验统计量

2. 理论分布含有未知参数的情形

H_0 : X 的分布函数为 $F(x; \theta_1, \theta_2, \dots, \theta_l)$

H_1 : X 的分布函数不是 $F(x; \theta_1, \theta_2, \dots, \theta_l)$

先利用样本 $(X_1, X_2, \dots, X_n)$ 求出 $\theta_1, \theta_2, \dots, \theta_l$ 出的最大似然估计值 $\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_l$ 求, 那么 $F(x; \hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_l)$ 就是完全已知的部分函数, 再按照前面的步骤进行. 这里第(2)步中计算出估计值 $\hat{p}_i$, 统计量为

$$\chi^2 = \sum_{i=1}^r \frac{(m_i - n\hat{p}_i)^2}{n\hat{p}_i}$$

那么, 当 $n \rightarrow \infty$ 时, 其极限分布为自由度 $r-l-1$ 的 χ^2 分布