第七章 假设检验

简介

假设检验

- 参数检验
  - 数学期望
    - σ^2 已知 (U检验法)
    - σ^2 未知 (t 检验法)
  - 方差
    - χ^2检验法 (单正态总体)
    - F检验法 (双正态总体)
- 非参数检验
  - 分布检验
  - 独立性检验

下页 返回

第一节　假设检验的基本概念

- 一、假设检验的基本原理
- 二、假设检验的相关概念
- 三、假设检验的一般步骤
- 四、内容小结

一、假设检验的基本原理

在总体的分布函数完全未知或只知其形式、但不知其参数的情况下，为了推断总体的某些性质，提出某些关于总体的假设.

例如，提出总体服从泊松分布的假设；

又如，对于正态总体提出数学期望等于 μ_0 的假设等.

假设检验就是根据样本对所提出的假设作出判断：是接受，还是拒绝.

1. 基本原理

$0 < \alpha \leq 0.05$

小概率推断原理： 小概率事件(概率接近0的事件)，在一次试验中，实际上可认为不会发生.

2. 基本思想方法

采用概率性质的反证法： 先提出假设H_0，再根据一次抽样所得到的样本值进行计算. 若导致小概率事件发生，则否认假设H_0；否则，接受假设H_0.

下面结合实例来说明假设检验的基本思想.

例1 某厂有一批产品，共有200件，需检验合格才能出厂．按国家标准，次品率不得超过3%．今在其中随机地抽取10件，发现其中有2件次品，问：这批产品能否出厂?

分析： 从直观上分析，这批产品不能出厂．

因为抽样得到的次品率：$\dfrac{2}{10} > 3\%$

然而，由于样本的随机性，如何才能根据抽样结果判断总体(所有产品)的次品率是否≤3%?

解 用假设检验法，**步骤：**

1° 提出假设 H_0: $p \le 0.03$

其中 p为总体的次品率.

2° 设 $X_i = \begin{cases} 1, & 第i次抽取的产品是次品 \\ 0, & 否则 \end{cases}$

则 $X_i \sim B(1, p) \qquad (i = 1, 2, \cdots, 10)$

令 $Y = X_1 + X_2 + \cdots + X_n$

$=\{$ 抽取的10件产品中的次品数 $\}$

则 $Y \sim B(10, p)$

3° 在假设 H_0成立的条件下，计算

$$\mathrm{H}_0: \ p \le 0.03$$

$$Y \sim B(10, p)$$

$$\begin{aligned} f(p) &= P\{Y \ge 2;\ p\} \\ &= 1 - P\{Y < 2;\ p\} \\ &= 1 - [P\{Y = 0;\ p\} + P\{Y = 1;\ p\}] \\ &= 1 - [C_{10}^{0} p^{0} (1-p)^{10} + C_{10}^{1} p^{1} (1-p)^{9}] \\ &= 1 - [(1-p)^{10} + 10p(1-p)^{9}] \end{aligned}$$

$\because \quad \dfrac{\mathrm{d} f(p)}{\mathrm{d} p} = 90p(1-p)^{8} > 0$

$\therefore$ 当 $p \le 0.03$ 时，$f(p)$单调增加

$\therefore$ 当 $p \le 0.03$ 时，

$$f(p) = P\{Y \ge 2;\ p\} = 1 - [(1-p)^{10} + 10p(1-p)^9]$$

$$\le f(0.03) = P\{Y \ge 2;\ 0.03\} \approx 0.035 < \alpha = 0.05$$

从而 $P\{Y = 2;\ p\} < P\{Y \ge 2;\ p\} < \alpha = 0.05$

故 $\{ Y = 2 \}$ 是小概率事件.

4° 作判断

由于在假设 H_0 成立的条件下，$\{ Y = 2 \}$ 是小概率事件， 而实际情况是：小概率事件竟然在一次试验中发生了，这违背了小概率原理，

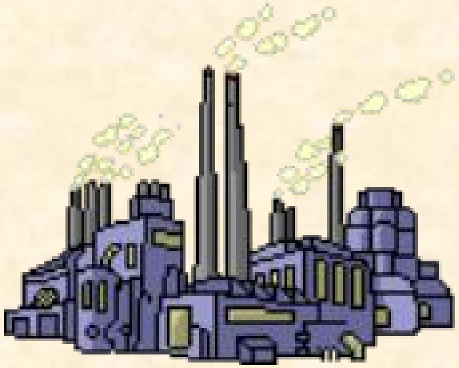
是不合理的，故应该否定原假设H_0，认为产品的次品率 $p > 3\%$.

所以，这批产品不能出厂.

例1 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 $\mu \neq 0.5$?

解 1° 提出两个对立假设

$$H_0 : \mu = \mu_0 = 0.5 \text{ 和 } H_1 : \mu \neq \mu_0 .$$

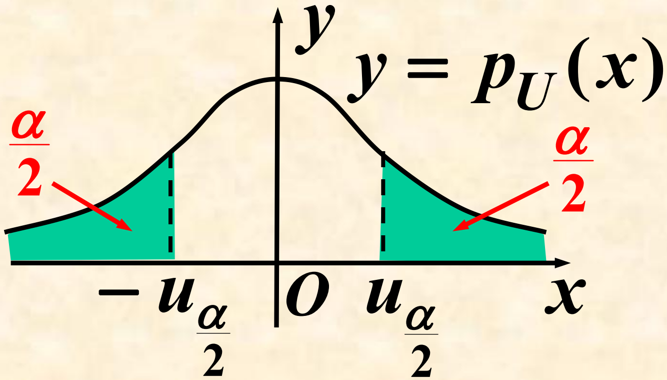
2° $\because$ $\overline{X}$ 是 μ 的无偏估计量,

$\therefore$ 若 H_0 为真, 则 $|\bar{x} - \mu_0|$ 不应太大,

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

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

$\because\quad P\{|U|>u_{\frac{\alpha}{2}}\}=\alpha$



当$\alpha>0$很小时，

$\{|U|>u_{\frac{\alpha}{2}}\}$是小概率事件

根据小概率原理，可以认为如果H_0为真，则由一次试验得到满足不等式 $|u|=\left|\dfrac{\overline{x}-\mu_0}{\sigma/\sqrt{n}}\right|\geq u_{\alpha/2}$

的观察值$\overline{x}$，几乎不会发生.

若在一次试验中,得到了满足不等式

$$|u| = \left|\frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}\right| \geq u_{\alpha/2}$$

的观察值 $\bar{x}$，则我们有理由怀疑原 来的假设H_0的正确性，因而拒绝 H_0.

若出现观察值 $\bar{x}$ 满足不等式

$$|u| = \left|\frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}\right| < u_{\alpha/2},$$

则没有理由拒绝假设H_0, 因而只能接受H_0.

当 $\dfrac{|\bar{x}-\mu_0|}{\sigma/\sqrt{n}} \geq u_{\alpha/2}$ 时，拒绝 H_0;

当 $\dfrac{|\bar{x}-\mu_0|}{\sigma/\sqrt{n}} < u_{\alpha/2}$ 时, 接受H_0.

如：若取定 $\alpha = 0.05$, 则 $u_{\alpha/2} = u_{0.025} = 1.96$,

3° 在假设 H_0成立的条件下，由样本计算

$$|u| = \frac{|\bar{x}-\mu_0|}{\sigma/\sqrt{n}} = 2.2 > u_{\alpha/2} = 1.96,$$

于是拒绝假设H_0, 认为包装机工作不正常.

二、假设检验的相关概念

1. 显著性水平

$$\alpha = P\{拒绝原假设H_0 \mid H_0为真\}$$

数 α 称为显著性水平.

如：对于例1,

当H_0: $\mu = 0.5$为真时， $U = \dfrac{\overline{X} - \mu_0}{\sigma / \sqrt{n}} \sim N(0,1)$,

$$P\{|U| > u_{\frac{\alpha}{2}} \mid H_0为真\} = \alpha$$

如果$|u|=\left|\dfrac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}\right|\geq u_{\frac{\alpha}{2}}$,则称 $\bar{x}$ 与μ_0的差异是显著的,

则我们拒绝 H_0,

反之,如果$|u|=\left|\dfrac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}\right|< u_{\frac{\alpha}{2}}$, 则称 $\bar{x}$ 与μ_0的差异是

不显著的,则我们接受 H_0,

上述关于 $\bar{x}$ 与 μ_0 有无显著差异的判断是 在显著性水平 α 之下作出的.

2. 检验统计量

用于检验假设的统计量，称为检验统计量.

如：对于例2,

$$\text{统计量 } U=\frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}} \quad \text{— 检验统计量.}$$

3. 原假设与备择假设

假设检验问题通常叙述为：在显著性水平 α下,

检验假设 $H_0:\mu=\mu_0, H_1:\mu\neq\mu_0.$

或称为"在显著性水平 α下,针对 H_1检验 H_0".

H_0称为原假设或零假设, H_1 称为备择假设.

4. 拒绝域与临界点

拒绝域 W_1：拒绝原假设 H_0 的所有样本值 $(x_1, x_2, \cdots, x_n)$ 所组成的集合.

$W_1 \leftrightarrow W_1'$：拒绝原假设 H_0 的检验统计量的取值范围.

临界点(值)：拒绝域的边界点(处的检验统计量的值).

如: 在前面例2中,

拒绝域为 $|u| \geq u_{\alpha/2}$,

临界值为 $u = -u_{\alpha/2}$, $u = u_{\alpha/2}$.

5. 两类错误及记号

假设检验的依据是: 小概率事件在一次试验中很难发生, 但“很难发生”不等于“不发生”, 因而假设检验所作出的结论有可能是错误的. 这种错误有两类:

(1) 当原假设H_0为真, 观察值却落入拒绝域, 而作出了拒绝H_0的判断, 称为**第一类错误**, 又叫**弃真错误**, 这类错误是“以真为假”. 犯第一类错误的概率是显著性水平α.

$$\alpha = P\{拒绝原假设H_0 \mid H_0为真\}$$

(2) 当原假设H_0不真, 而观察值却落入接受域, 而作出了接受H_0的判断, 称为**第二类错误**, 又叫**取伪错误**, 这类错误是“以假为真”.

犯第二类错误的概率记为

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

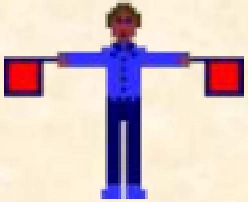
注 1° 当样本容量 n 一定时, 若减少犯第一类错误的概率, 则犯第二类错误的概率往往增大.

2° 若要使犯两类错误的概率都减小, 除非增加样本容量.

6. **显著性检验**

只对犯第一类错误的概率加以控制, 而不考虑犯第二类错误的概率的检验, 称为**显著性检验**.

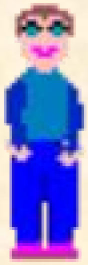
7. **双侧备择假设与双侧假设检验**



在 $H_0: \mu = \mu_0$ 和 $H_1: \mu \neq \mu_0$ 中, 备择假设 H_1 表示 μ 可能大于 μ_0, 也可能小于 μ_0, 称为双边备择假设, 形如 $H_0: \mu = \mu_0$, $H_1: \mu \neq \mu_0$ 的假设检验称为双边假设检验.

目录 上页 下页 返回 结束

8. 单侧检验(右侧检验与左侧检验)



形如 $H_0: \mu = \mu_0, H_1: \mu > \mu_0$ 的假设检验称为右侧检验.

形如 $H_0: \mu = \mu_0, H_1: \mu < \mu_0$ 的假设检验称为左侧检验.

右侧检验与左侧检验统称为**单侧检验**.

三、假设检验的一般步骤

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. 选择适当的检验统计量 ,在H_0成立的条件下,确定它的概率分布;

3. 给定显著性水平 α,确定拒绝域 W_1;

4. 根据样本观察值计算统 计量的值;

5. 根据统计量值是否落入拒绝域W_1中, 作出拒绝或者接受H_0的判断.

四、内容小结

假设检验的基本原理、相关概念和一般步骤.

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

典型例题

例3 设$(X_1, X_2, \cdots, X_n)$是来自正态总体$N(\mu,100)$的一个样本，要检验 $H_0: \mu = 0\ (H_1: \mu \neq 0)$，在下列两种情况下，分别确定常数$d$，使得以$W_1$为拒绝域的检验犯第一类错误的 概率为0.05 .

(1) $n = 1, W_1 = \{x_1 \mid |x_1| > d\}$;

(2) $n = 25, W_1 = \{(x_1, \cdots, x_{25}) \mid |\bar{x}| > d\}$其中$\bar{x} = \dfrac{1}{25}\sum_{i=1}^{25} x_i$.

解 (1) $n = 1$时，若H_0成立， 则 $\dfrac{X_1}{10} \sim N(0,1)$,

$$P(X_1 \in W_1) = P(|X_1| > d)$$

$$= P\left(\left|\frac{X_1}{10}\right| > \frac{d}{10}\right) = \Phi\left(-\frac{d}{10}\right) - \Phi\left(\frac{d}{10}\right)$$

$$= 2\left(1 - \Phi\left(\frac{d}{10}\right)\right) = 0.05,$$

$$\Phi\left(\frac{d}{10}\right) = 0.975, \qquad \frac{d}{10} = 1.96, \qquad d = 19.6;$$

(2) $n=25$时, 若H_0成立, 则 $\sqrt{25}\dfrac{\overline{X}}{10}\sim N(0,1)$,

$$P\left(\left(X_1,\cdots X_{25}\right)\in W_1\right)=P\left(\left|\overline{X}\right|>d\right)$$

$$=P\left(\left|\frac{\overline{X}}{2}\right|>\frac{d}{2}\right)=\Phi\left(-\frac{d}{2}\right)-\Phi\left(\frac{d}{2}\right)$$

$$=2\left(1-\Phi\left(\frac{d}{2}\right)\right)=0.05,$$

$$\Phi\left(\frac{d}{2}\right)=0.975,\qquad \frac{d}{2}=1.96,\qquad d=3.92\,.$$

例4 设$(X_1, X_2, \cdots, X_n)$是来自正态总体$N(\mu, 9)$的一个样本,其中μ为未知参数,检验$H_0: \mu = \mu_0$ $(H_1: \mu \neq \mu_0)$,拒绝域$W_1 = \{(x_1, \cdots, x_n) \mid |\bar{x} - \mu_0| \geq C\}$,

(1) 确定常数C,使显著性水平为0.05;

(2) 在固定样本容量$n = 25$的情况下,分析犯两类错误的概率α和β之间的关系.

解 (1) 若H_0成立,则$\dfrac{\sqrt{n}(\bar{X} - \mu_0)}{3} \sim N(0,1)$,

$$P((X_1, \cdots X_n) \in W_1) = P(|\bar{X} - \mu_0| \geq C)$$

$$= P\left(\frac{\sqrt{n}\left|\overline{X} - \mu_0\right|}{3} \geq \frac{\sqrt{n}C}{3}\right) = 2\left(1 - \Phi\left(\frac{\sqrt{n}C}{3}\right)\right) = 0.05,$$

$$\Phi\left(\frac{\sqrt{n}C}{3}\right) = 0.975, \qquad \frac{\sqrt{n}C}{3} = 1.96, \qquad C = \frac{5.88}{\sqrt{n}};$$

(2) $n = 25$时, 若H_0成立,

$$\alpha = P((X_1, \cdots X_n) \in W_1)$$

$$= 2\left(1 - \Phi\left(\frac{\sqrt{n}C}{3}\right)\right) = 2\left(1 - \Phi\left(\frac{5C}{3}\right)\right)$$

若H_0不成立,

$$\beta = P((X_1,\cdots X_n) \in W_1) = P(|\overline{X} - \mu_0| \le C)$$

$$= P(-C + \mu_0 \le \overline{X} \le C + \mu_0)$$

$$= P\left(\frac{5}{3}(-C + \mu_0 - \mu) \le \frac{5}{3}(\overline{X} - \mu) \le \frac{5}{3}(C + \mu_0 - \mu)\right)$$

$$= \Phi\left(\frac{5}{3}(C + \mu_0 - \mu)\right) - \Phi\left(\frac{5}{3}(-C + \mu_0 - \mu)\right)$$

当α 较大时,C较小,因此β 较小;

当C 较大时,β 较大,因而α 较小.