第九章　假设检验

假设检验的概念

任何一个有关随机变量未知分布的假设称为统计假设或简称假设. 一个仅牵涉到随机变量分布中几个未知参数的假设称为参数假设. 这里所说的"假设"只是一个设想, 至于它是否成立, 在建立假设时并不知道, 还需进行考察.

对一个样本进行考察, 从而决定它能否合理地被认为与假设相符, 这一过程叫做假设检验. 判别参数假设的检验称为参数检验. 检验是一种决定规则, 通过一定的程序作出是与否的判断.

例1

抛掷一枚硬币100次, "正面"出现了40次, 问这枚硬币是否匀称?

若用ξ描述抛掷一枚硬币的试验, "ξ=1"及"ξ=0"分别表示"出现正面"和"出现反面", 上述问题就是要检验x是否服从p=1/2的0-1分布?

例2

从1975年的新生儿(女)中随机地抽取20个, 测得其平均体重为3160克, 样本标准差为300克. 而根据过去统计资料, 新生儿(女)平均本重为3140克. 问现在与过去的新生儿(女)体重有无显著差异(假设新生儿体重服从正态分布)?

若把所有1975年新生儿(女)体重体现为一个总体ξ, 问题就是判断$E\xi=3140$是否成立?

例3

在10个相同的地块上对甲,乙两种玉米进行对比试验, 得如下资料(单位:公斤)

甲	951	966	1008	1082	983
乙	730	864	742	774	990

从直观上看, 二者差异显著. 但是一方面由于抽样的随机性, 我们不能以个别值进行比较就得出结论; 另一方面直观的标准可能因人而异. 因此这实际上需要比较两个正态总体的期望值是否相等.

这种作为检验对象的假设称为待检假设, 通常用H_0表示,

例如, 例1的假设是

H_0: $\xi \sim B(1,0.5)$

例2的假设是

H_0: $E\xi=3140$

例3的假设是

H_0: $EX=EY$ (X与Y是两种玉米的产量期望值)

如何根据样本的信息来判断关于总体分布的某个设想是否成立, 也就是检验假设H_0成立与否的方法.

置信区间方法

用置信区间的方法进行检验, 基本思想是这样的: 首先设想H_0是真的成立; 然后考虑在H_0成立的条件下, 已经观测到的样本信息出现的概率. 如果这个概率很小, 这就表明一个概率很小的事件在一次试验中发生了. 而小概率原理认为, 概率很小的事件在一次试验中是几乎不可能发生的, 也就是说导出了一个违背小概率原理的不合理的现象. 这表明事先的设想H_0是不正确的, 因此拒绝原假设H_0. 否则, 不能拒绝H_0.

至于什么算是"概率很小", 在检验之前都事先指定. 比如概率为5%, 1%等, 一般记作α.

α是一个事先指定的小的正数, 称为显著性水平或检验水平.

两类错误

由于人们作出判断的依据是样本, 也就是由部分来推断整体, 因而假设检验不可能绝对准确, 它也可能犯错误. 其可能性的大小, 也是以统计规律性为依据的, 所可能犯的错误有两类.

第一类错误是: 原假设H_0符合实际情况, 而检验结果把它否定了, 这称为弃真错误.

第二类错误是: 原假设H_0不符合实际情况, 而检验结果把它肯定下来了, 这称为取伪错误.

一个正态总体的假设检验

设总体为$\xi \sim N(\mu,\sigma^2)$. 关于总体参数μ,σ^2的假设检验问题, 本节介绍下列四种:

(1) 已知方差σ^2, 检验假设$H_0:\mu=\mu_0$;

(2) 未知方差σ^2, 检验假设$H_0:\mu=\mu_0$;

(3) 未知期望μ, 检验假设$H_0:\sigma^2=\sigma_0^2$;

(4) 未知期望μ, 检验假设$H_0:\sigma^2\le\sigma_0^2$;

其中H_0中的σ_0^2, μ_0都是已知数.

例1 根据长期经验和资料的分析, 某砖瓦厂生产砖的"抗断强度"ξ服从正态分布, 方差$\sigma^2=1.21$. 从该厂产品中随机抽取6块, 测得抗断强度如下(单位: kg/cm^2):

32.56, 29.66, 31.64, 30.00, 31.87, 31.03

检验这批砖的平均抗断强度为32.50kg/cm^{2}是否成立($\alpha=0.05$)?

解 设H_0:μ=32.50. 如果H_0正确, 则样本($X_1,\ldots,X_6$)来自正态总体$N(32.50,\ 1.1^2)$, 令

$$U=\frac{\bar{X}-32.50}{1.1/\sqrt{6}}\sim N(0,1)$$

对给定的$\alpha=0.05$, 查表得$u_\alpha=1.96$

使$P(|U|>u_\alpha)=\alpha$, 即$\Phi_0(u_\alpha)=1-\dfrac{\alpha}{2}$

计算求出U的样本值为

$$|u|=\left|\frac{31.13-32.50}{1.1/\sqrt{6}}\right|\approx 3.05>1.96$$

最后可以下结论否定H_0, 即不能认为这批产品的平均抗断强度是32.50kg/cm^2.

方差已知对期望值μ的检验步骤:

(1) 提出待检假设H_0:$\mu=\mu_0$(μ_0已知);

(2) 选取样本($X_1,\ldots,X_n$)的统计量, 如H_0成立,则

$$U=\frac{\overline{X}-\mu_0}{\sigma_0/\sqrt{n}}\sim N(0,1) \quad (\sigma_0\text{为已知})$$

(3) 根据检验水平α, 查表确定临界值u_α, 使

$P(|U|>u_\alpha)=\alpha$, 即$\Phi_0(u_\alpha)=1-\alpha/2$.

(4) 根据样本观察值计算统计量U的值u并与临界值u_α比较

(5) 若$|u|>u_\alpha$则否定H_0, 否则接受H_0.

例2 假定某厂生产一种钢索, 它的断裂强度ξ(kg/m^2)服从正态分布$N(\mu,40^2)$. 从中选取一个容量为9的样本, 得$\overline{X}$=780kg/m^2. 能否据此样本认为这批钢索的断裂强度为800kg/cm^2 (α=0.05)

解 首先建立$H_0:\mu=800.$如H_0成立则

$$U=(\overline{X}-800)\sqrt{9}/40\sim N(0,1),$$

因$\alpha=0.05$则$u_\alpha=1.96$

$$|u|=\left|\frac{780-800}{40/3}\right|=1.5<1.96$$

可以接受H_0, 认为断裂强度为800kg/cm^2

例3

从1975年的新生儿(女)中随机地抽取20个, 测得其平均体重为3160克, 样本标准差为300克. 而根据过去统计资料, 新生儿(女)平均本重为3140克. 问现在与过去的新生儿(女)体重有无显著差异(假设新生儿体重服从正态分布)?(α=0.01)

若把所有1975年新生儿(女)体重体现为一个正态总体$N(\mu,\sigma^2)$, 问题就是判断

$\mu=E\xi=3140$是否成立?

解 待检假设H_0:μ=3140. 由于σ^2未知, 自然想到用S^2代表σ^2. 则如果H_0成立, 则

$$T=\frac{\overline{X}-3140}{S/\sqrt{20}}\sim t(19)$$

由给定的检验水平$\alpha=0.01$,查表得$t_{0.01}(19)=2.861$,

即 $$P\left\{\left|\frac{\overline{X}-3140}{S/\sqrt{20}}\right|>2.861\right\}=0.01$$

而$$|t|=\frac{3160-3140}{300/\sqrt{20}}\approx 0.298<2.861$$

因此接收H_0,认为现在的新生儿体重与过去没有明显差异.

方差未知对期望值μ的检验步骤:

(1) 提出待检假设H_0:$\mu=\mu_0$(μ_0已知);

(2) 选取样本($X_1,\ldots,X_n$)的统计量, 如H_0成立,则

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

(3) 根据检验水平α, 查表确定临界值t_α, 使

$P(|T|>t_\alpha)=\alpha$;

(4) 根据样本观察值计算统计量T的值t并与临界值t_α比较;

(5) 若$|t|>t_\alpha$则否定H_0, 否则接受H_0.

例4 某炼铁厂的铁水含碳量ξ在正常情况下服从正态分布. 现对操作工艺进行了某些改进, 从中抽取5炉铁水测得含碳量数据如下:

4.412, 4.052, 4.357, 4.287, 4.683

据此是否可以认为新工艺炼出的铁水含碳量的方差仍为0.108^2(α=0.05).

解 建立待检假设H_0:σ2=0.108^2; 在H_0成立时, 样本来自总体$N(\mu,0.108^2)$, 这时

$$\chi^2 = \frac{(n-1)S^2}{0.108^2} \sim \chi^2(n-1)$$

对于给定的检验水平α=0.05, 可查表确定临界值χ_a^2及χ_b^2,使

$$P\left(\frac{(n-1)S^2}{0.108^2}<\chi_a^2\right)=P\left(\frac{(n-1)S^2}{0.108^2}>\chi_b^2\right)=\frac{\alpha}{2}$$

查表得$\chi_a^2=\chi_{0.975}^2(4)=0.484$,

$\chi_b^2=\chi_{0.025}^2(4)=11.1$,具体计算统计量$\chi^2$的值:

$$\chi^2=\frac{4\times 0.228^2}{0.108^2}\approx 17.827>11.1$$

因而应拒绝H_0, 即方差不能认为是0.108^2

未知期望对正态总体方差的假设检验步骤:

(1) 建立待检假设H_0:$\sigma^2=\sigma_0^2$;

(2) 如H_0成立, 则

$$\chi^2=\frac{(n-1)S^2}{\sigma_0^2}\sim\chi^2(n-1)$$

(3) 由给定的检验水平α查表求χ_a^2,χ_b^2满足:

$$P\left(\frac{(n-1)S^2}{\sigma_0^2}>\chi_b^2\right)=P\left(\frac{(n-1)S^2}{\sigma_0^2}<\chi_a^2\right)=\frac{\alpha}{2}$$

(4) 计算χ^2的值与χ_a^2,χ_b^2比较;

(5) 若$\chi^2>\chi_b^2$或$\chi^2<\chi_a^2$拒绝H_0否则接受H_0;

例5 机器包装食盐, 假设每袋盐的净重服从正态分布, 规定每袋标准重量为500g, 标准差不超过10g. 某天开工后, 为检查其机器工作是否正常, 从装好的食盐中随机取9袋, 测其净重(单位:g)为

497,507,510,475,484,488,524,491,515

问这天包装机工作是否正常(α=0.05)?

解 设ξ为一袋食盐的净重, 则$\xi \sim N(\mu,\sigma^2)$. 今需假设$H_0:\mu=500$以及$H_0':\sigma^2\leq 100$是否成立?

$H_0:\mu=500$:查表求临界值$t_{0.05}(9-1)=2.306$, 计算得

$$\bar{x} = 499, s = 16.03$$

如H_0成立, 则有

$$T=\frac{\overline{X}-500}{S/\sqrt{9}}\sim t(9-1)$$

具体算得

$$|t|=\left|\frac{499-500}{16.03/\sqrt{9}}\right|\approx 0.187<2.306$$

故接受H_0,即认为平均每袋食盐净重为$500g$

机器没有系统误差.

当$\sigma^2=10^2$时, 必有

$$\chi^2=\frac{(9-1)S^2}{10^2}\sim\chi^2(9-1)$$

对给定的α, 查表求χ_a^2使

$$P\{\chi^2\geq\chi_a^2\}=\alpha=0.05$$

查得$\chi_{0.05}^2(8)=15.5$

再算得$\chi^2=\dfrac{8\times16.03^2}{10^2}\approx20.56>15.5$

由此拒绝H_0', 即方差过大, 该天打包机工作不正常

顺便指出, 在检验假设H_0:$\sigma^2=\sigma_0^2$时, 有两个临界值χ_a^2和χ_b^2, 即当$(n-1)S^2/\sigma_0^2<\chi_a^2$或$(n-1)S^2/\sigma_0^2>\chi_b^2$时, 都应拒绝$H_0$. 这是基于若$H_0$成立, 即$\sigma^2=\sigma_0^2$, 则$S^2/\sigma_0^2$的值不应太大或太小. 而在检验假设$H_0'$:$\sigma^2\le\sigma_0^2$时, 考虑到若$H_0'$成立, 即$\sigma^2\le\sigma_0^2$, 则$S^2/\sigma_0^2$不应太大, 但较小是合理的, 因此拒绝域只取$(n-1)S^2/\sigma_0^2>\chi_a^2$.