第八章　假设检验

第一节　假设检验的基本概念

一、假设检验的背景及提法

1. 假设检验的背景

例 1 某洗衣粉厂用自动包装机进行包装，正常情况下包装量 $X \sim N(500,9)$（单位：克），现随机抽取 25 袋洗衣粉，测得平均重量 $\bar{x} = 501.5$ 克，假定方差不变，问可否认为平均包装量 μ 仍为 500 克？

例 2 设有某车间生产的甲、乙两批同型号的产品，其次品率分别为 p_1 和 p_2，其中 p_1 和 p_2 均未知．现从甲批产品中任取36件产品，发现有2件次品；再从乙批产品中任取50件产品，发现有3件次品，问是否有 $p_1 < p_2$？

例 3 将一枚骰子抛120次，统计出各点数出现的次数如下

点数	1	2	3	4	5	6
出现的次数	21	28	19	24	16	12

问这枚骰子的六个面是否均匀？

假设检验问题是非常丰富的．按检验的内容，假设检验可分为参数检验和非参数检验．

如果总体 X 的分布类型已知，检验只涉及到其中的某些参数，这类假设检验称为参数检验．

例 1，例 2 为参数检验．例 3 为非参数检验．

在参数检验问题中，又会出现单总体和多总体情形．

例 1 为单总体情形；例 2 为双总体情形．

假设检验不同于参数估计．参数估计是想了解总体 X 中未知参数 θ 的取值大约是多少，从而进行点估计等等．而假设检验并不想知道未知参数 θ 的取值，只是判断未知参数 θ 是否满足某种关系．

如例 1 中，检验的问题是接受 $\mu = 500$，还是 $\mu \neq 500$，如果是 $\mu \neq 500$，那么此时 μ 取值多少并不是重点关心的问题．

2. 假设的提法

称检验问题中相互对立的两个方面命题为**假设或统计假设**．并将其中一个命题称为**原假设**或**零假设**，记为 H_0；另一个命题称为**备择假设**，记为 H_1．因此检验问题常简记为 (H_0, H_1)．

注 1： 原假设中的“原”可理解为“原本有的”，具有“保持不变”（带有“$=$”，“$\leq$”或“$\geq$”） 的特征.

备择假设指抛弃原假设后可供选择的假设，具有“发生变化”（含有“$\neq$”，“$<$”或“$>$”）的特征.

二、假设检验的思想和方法

1. 假设检验中的反证法思想

反证法思想（注意：不是指严格的反证法）：

先假定 H_0 成立，然后根据统计分析的思想和方法，进行推理和演算，如果推理和演算的结果中有“矛盾”的现象出现，就“主动地”拒绝 H_0，接受 H_1；如果其结果中没有“矛盾”的现象出现，就不能拒绝 H_0，因此只好“被动地”接受 H_0，拒绝 H_1 .

2. 假设检验的基本原理

小概率原理：在正常情况下，小概率事件在一次抽样中是几乎不可能发生的.

如果在一次抽样中，某小概率事件 A 发生了，应属于“不正常现象”，即“‘矛盾’的现象”出现了. 在检验问题(H_0, H_1)中，就会认为对总体所做的原假设H_0不正确，从而拒绝H_0，接受H_1.

3. 假设检验的两类错误

根据假设检验的基本原理知道，在假定 H_0 成立的情况下，选择统计量 $g(X_1, X_2, \cdots, X_n)$，并由其分布确定一个小概率事件 A.

当经过抽样得到样本值 $(x_1, x_2, \cdots, x_n)$ 时，计算统计量 $g(X_1, X_2, \cdots, X_n)$ 的观察值 $g(x_1, x_2, \cdots, x_n)$，根据其结果，决定小概率事件 A 是否发生，并依此对检验问题 (H_0, H_1) 作出判断.

如果小概率事件 A 发生，则拒绝 H_0．因此，导致小概率事件 A 发生的 $g(X_1, X_2, \cdots, X_n)$ 的全体取值范围称为 H_0 **的拒绝域**，记为 W．

如果小概率事件 A 不发生，则接受 H_0．同理，导致小概率事件 A 不发生的 $g(X_1, X_2, \cdots, X_n)$ 的全体取值范围称为 H_0 **的接受域**.

由于样本具有随机性（小概率事件有时也会发生），在一次抽样中，A 可能发生，A 也可能不发生．因此，检验结果与真实情况之间就有四种情形：

(1) 真实情况 H_0 成立，且检验结果接受 H_0，拒绝 H_1；

(2) 真实情况 H_0 成立，而检验结果拒绝 H_0，接受 H_1；

(3) 真实情况 H_1 成立，而检验结果接受 H_0，拒绝 H_1；

(4) 真实情况 H_1 成立，且检验结果拒绝 H_0，接受 H_1．

其中(1)和(4)中的检验结果与真实情况完全吻合，表明理论判断正确．但(2)和(3)中两者不一致，表明理论判断有误，这就是假设检验的两类错误．

称真实情况 H_0 成立，而检验结果拒绝 H_0 为**第一类错误**或**弃真错误**；称真实情况 H_1 成立，而检验结果接受 H_0 为**第二类错误**或**存伪错误**.

检验结果 \ 真实情况	接受 H_0	接受 H_1（拒绝 H_0）
H_0 成立	判断正确	第一类错误（弃真错误）
H_1 成立（H_0 不成立）	第二类错误（存伪错误）	判断正确

记犯第一类错误，即弃真错误的概率为α，犯第二类错误，即存伪错误的概率为β.

理论上已经证明，当样本容量n无限增大时，可以同时降低α和β，而这在实际问题中是不可能做到的．但当样本容量n取某固定值时，α和β会出现此消彼长现象.

因此，在控制α和β时，要选择一个先后次序．目前比较流行的做法是采用“优先固定或限制第一类错误概率α的原则”，并在此基础上，降低犯第二类错误的概率.

在优先固定或限制第一类错误概率 α 后，可以通过构造“好的”统计量或统计方法，降低犯第二类错误的概率 β .

4. 显著性检验

根据假设检验的基本原理，先假定 H_0 成立，然后选择统计量 $g(X_1, X_2, \cdots, X_n)$，并由样本值 $(x_1, x_2, \cdots, x_n)$ 得到统计量的值 $g(x_1, x_2, \cdots, x_n)$.

如果 $g(x_1, x_2, \cdots, x_n)$ 反映了抽样结果与总体情况有显著的差异，则表明总体的状况已经发生了质的变化，这时就应该拒绝 H_0. 而拒绝 H_0 后，就有可能犯第一类错误.

因此，统计中，称**犯第一类错误的概率 α 为显著性水平**.

如果假设检验问题(H_0, H_1)为

$$H_0: \quad \theta = \theta_0, \quad H_1: \quad \theta \neq \theta_0,$$

就称之为**双侧（边）检验**.

如果假设检验问题(H_0, H_1)为

$$H_0: \quad \theta \geq \theta_0, \quad H_1: \quad \theta < \theta_0,$$

或

$$H_0: \quad \theta \leq \theta_0, \quad H_1: \quad \theta > \theta_0,$$

就称之为**单侧（边）检验**.

5. 假设检验的四个步骤

第一步： 根据给定的问题，建立假设检验问题 (H_0, H_1)；

第二步： 根据假设检验问题 (H_0, H_1) 及条件，选择检验统计量 $g(X_1, X_2, \cdots, X_n)$．当 H_0 为成立时，确定该统计量 $g(X_1, X_2, \cdots, X_n)$ 的分布；

第三步： 根据显著性水平 α，确定临界值和原假设 H_0 的拒绝域 W；

第四步： 通过样本值 $(x_1, x_2, \cdots, x_n)$，计算统计量 $g(X_1, X_2, \cdots, X_n)$ 的值 $g(x_1, x_2, \cdots, x_n)$． 若 $g(x_1, x_2, \cdots, x_n) \in W$，则拒绝 H_0，否则接受 H_0．

例 4 某食品厂生产的罐头重量 $X \sim N(\mu, 4)$（单位：克），在正常情况下，$\mu = 500$．现任意抽取了16听罐头，测得其平均重量为 $\bar{x} = 502$ 克．在显著性水平 $\alpha = 0.05$ 下，问可否认为现在仍有 $\mu = 500$？

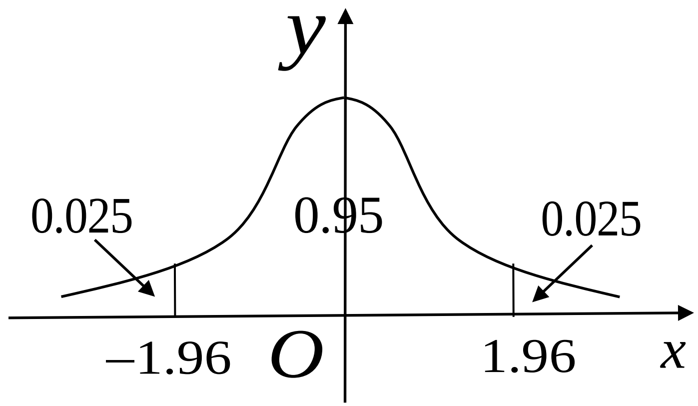
解 (1)假设检验问题为 H_0：$\mu = 500$，H_1：$\mu \neq 500$．

(2)选择统计量以及分布为 $U = \dfrac{\overline{X} - \mu}{\sigma/\sqrt{n}} \overset{\text{在}H_0\text{成立下}}{=} \dfrac{\overline{X} - 500}{\sigma/\sqrt{n}} \sim N(0,1)$．

(3)由于该检验为双侧检验，且 $\alpha = 0.05$，查表得临界值为

$U_{\frac{\alpha}{2}} = U_{0.025} = 1.96$ 和 $U_{1-\frac{\alpha}{2}} = U_{0.975} = -U_{0.025} = -1.96$，所以 H_0 的拒绝域 $W = \{U \le -1.96 \text{ 或 } U \ge 1.96\} = \{|U| \ge 1.96\}$．

（续解）



(4)又$\sigma=2, n=16, \bar{x}=502$，计算得统计量的观测值为

$$U_0=\frac{502-500}{2/\sqrt{16}}=4\in W,$$

所以拒绝H_0，即不可认为现在有$\mu=500$.