第八章　假设检验

§1　假设检验

统计推断的另一类重要问题是假设检验问题. 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，为了推断总体的某些未知特性，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. 我们要根据样本对所提出的假设作出是接受，还是拒绝的决策. 假设检验是作出这一决策的过程. 这里，先结合例子来说明假设检验的基本思想和做法.

例 1　某车间用一台包装机包装葡萄糖. 袋装糖的净重是一个随机变量，它服从正态分布. 当机器正常时，其均值为 0.5 kg，标准差为 0.015 kg. 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(kg)

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?

以 μ,σ 分别表示这一天袋装糖的净重总体 X 的均值和标准差. 由于长期实践表明标准差比较稳定，我们就设 $\sigma=0.015$. 于是 $X\sim N(\mu,0.015^2)$，这里 μ 未知. 问题是根据样本值来判断 $\mu=0.5$ 还是 $\mu\neq0.5$. 为此，我们提出两个相互对立的假设

$$H_0:\mu=\mu_0=0.5$$

和

$$H_1:\mu\neq\mu_0.$$

然后，我们给出一个合理的法则，根据这一法则，利用已知样本作出决策是接受假设 H_0(即拒绝假设 H_1)，还是拒绝假设 H_0(即接受假设 H_1). 如果作出的决策是接受 H_0，则认为 $\mu=\mu_0$，即认为机器工作是正常的，否则，则认为是不正常的.

由于要检验的假设涉及总体均值 μ，故首先想到是否可借助样本均值 $\overline{X}$ 这一统计量来进行判断. 我们知道，$\overline{X}$是 μ 的无偏估计，$\overline{X}$的观察值 $\overline{x}$ 的大小在一定程度上反映 μ 的大小. 因此，如果假设 H_0 为真，则观察值$\overline{x}$与 μ_0 的偏差 $|\overline{x}-\mu_0|$ 一般不应太大. 若 $|\overline{x}-\mu_0|$ 过分大，我们就怀疑假设 H_0 的正确性而拒绝 H_0，

并考虑到当 H_0 为真时 $\dfrac{\overline{X}-\mu_0}{\sigma/\sqrt{n}} \sim N(0,1)$. 而衡量 $|\overline{x}-\mu_0|$ 的大小可归结为衡量 $\dfrac{|\overline{x}-\mu_0|}{\sigma/\sqrt{n}}$ 的大小. 基于上面的想法，我们可适当选定一正数 k，使当观察值 $\overline{x}$ 满足 $\dfrac{|\overline{x}-\mu_0|}{\sigma/\sqrt{n}} \geqslant k$ 时就拒绝假设 H_0，反之，若 $\dfrac{|\overline{x}-\mu_0|}{\sigma/\sqrt{n}} < k$，就接受假设 H_0.

然而，由于作出决策的依据是一个样本，当实际上 H_0 为真时仍可能作出拒绝 H_0 的决策（这种可能性是无法消除的），这是一种错误，犯这种错误的概率记为

$$P\{\text{当 } H_0 \text{ 为真拒绝 } H_0\} \quad \text{或} \quad P_{\mu_0}\{\text{拒绝 } H_0\} \quad \text{或} \quad P_{\mu\in H_0}\{\text{拒绝 } H_0\}.$$

记号 $P_{\mu_0}\{\cdot\}$ 表示参数 μ 取 μ_0 时事件 $\{\cdot\}$ 的概率，$P_{\mu\in H_0}\{\cdot\}$ 表示 μ 取 H_0 规定的值时事件 $\{\cdot\}$ 的概率. 我们无法排除犯这类错误的可能性，因此自然希望将犯这类错误的概率控制在一定限度之内，即给出一个较小的数 α $(0<\alpha<1)$，使犯这类错误的概率不超过 α，即使得

$$P\{\text{当 } H_0 \text{ 为真拒绝 } H_0\} \leqslant \alpha. \tag{1.1}$$

为了确定常数 k，我们考虑统计量 $\dfrac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}$. 由于只允许犯这类错误的概率最大为 α，令(1.1)式右端取等号，即令

$$P\{\text{当 } H_0 \text{ 为真拒绝 } H_0\} = P_{\mu_0}\left\{\left|\frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}\right| \geqslant k\right\} = \alpha,$$

由于当 H_0 为真时，$Z=\dfrac{\overline{X}-\mu_0}{\sigma/\sqrt{n}} \sim N(0,1)$，由标准正态分布分位点的定义得（如图 8—1）.

$$k = z_{\alpha/2}.$$

因而，若 Z 的观察值满足

$$|z| = \left|\frac{\overline{x}-\mu_0}{\sigma/\sqrt{n}}\right| \geqslant k = z_{\alpha/2},$$

则拒绝 H_0，而若

$$|z| = \left|\frac{\overline{x}-\mu_0}{\sigma/\sqrt{n}}\right| < k = z_{\alpha/2},$$

则接受 H_0.

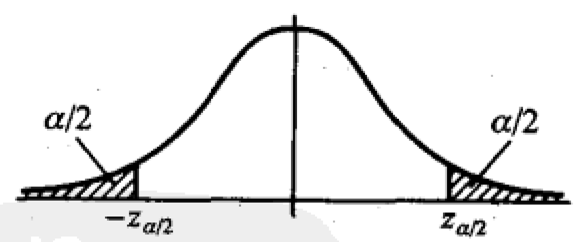

图 8—1

例如，在本例中取 $\alpha=0.05$，则有 $k=z_{0.05/2}=z_{0.025}=1.96$，又已知 $n=9$，$\sigma=0.015$，再由样本算得 $\overline{x}=0.511$，即有

$$\left|\frac{\overline{x}-\mu_0}{\sigma/\sqrt{n}}\right| = 2.2 > 1.96,$$

于是拒绝 H_0，认为这天包装机工作不正常. □

上例中所采用的检验法则是符合实际推断原理的. 因通常 α 总是取得较小，一般取 $\alpha=0.01, 0.05$. 因而若 H_0 为真，即当 $\mu=\mu_0$ 时，$\left\{\left|\dfrac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}\right|\geqslant z_{\alpha/2}\right\}$ 是一个小概率事件，根据实际推断原理，就可以认为，如果 H_0 为真，则由一次试验得到的观察值 $\overline{x}$，满足不等式 $\left|\dfrac{\overline{x}-\mu_0}{\sigma/\sqrt{n}}\right|\geqslant z_{\alpha/2}$ 几乎是不会发生的. 现在在一次观察中竟然出现了满足 $\left|\dfrac{\overline{x}-\mu_0}{\sigma/\sqrt{n}}\right|\geqslant z_{\alpha/2}$ 的 $\overline{x}$，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. 若出现的观察值 $\overline{x}$ 满足 $\left|\dfrac{\overline{x}-\mu_0}{\sigma/\sqrt{n}}\right|<z_{\alpha/2}$，此时没有理由拒绝假设 H_0，因此只能接受假设 H_0.

在上例的做法中，我们看到当样本容量固定时，选定 α 后，数 k 就可以确定，然后按照统计量 $Z=\dfrac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}$ 的观察值的绝对值 $|z|$ 大于等于 k 还是小于 k 来作出决策. 数 k 是检验上述假设的一个门槛值. 如果 $|z|=\left|\dfrac{\overline{x}-\mu_0}{\sigma/\sqrt{n}}\right|\geqslant k$，则称 $\overline{x}$ 与 μ_0 的差异是显著的，这时拒绝 H_0；反之，如果 $|z|=\left|\dfrac{\overline{x}-\mu_0}{\sigma/\sqrt{n}}\right|<k$，则称 $\overline{x}$ 与 μ_0 的差异是不显著的，这时接受 H_0. 数 α 称为**显著性水平**，上面关于 $\overline{x}$ 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.

统计量 $Z=\dfrac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}$ 称为**检验统计量**.

前面的检验问题通常叙述成：在显著性水平 α 下，检验假设

$$H_0:\mu=\mu_0,\quad H_1:\mu\neq\mu_0. \tag{1.2}$$

也常说成"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**（意指在原假设被拒绝后可供选择的假设）. 我们要进行的工作是，根据样本，按上述检验方法作出决策在 H_0 与 H_1 两者之间接受其一.

当检验统计量取某个区域 C 中的值时，我们拒绝原假设 H_0，则称区域 C 为**拒绝域**，拒绝域的边界点称为**临界点**. 如在上例中拒绝域为 $|z|\geqslant z_{\alpha/2}$，而 $z=-z_{\alpha/2}$，$z=z_{\alpha/2}$ 为临界点.

由于检验法则是根据样本作出的，总有可能作出错误的决策. 如上面所说的那样，在假设 H_0 实际上为真时，我们可能犯拒绝 H_0 的错误，称这类"弃真"的错误为第Ⅰ类错误. 又当 H_0 实际上不真时，我们也有可能接受 H_0. 称这类"取伪"的错误为第Ⅱ类错误. 犯第Ⅱ类错误的概率记为

$$P\{\text{当 } H_0 \text{ 不真接受 } H_0\} \quad \text{或} \quad P_{\mu\in H_1}\{\text{接受 } H_0\}.$$

为此，在确定检验法则时，我们应尽可能使犯两类错误的概率都较小. 但是，进一步讨论可知，一般来说，当样本容量固定时，若减少犯一类错误的概率，则犯另一类错误的概率往往增大. 若要使犯两类错误的概率都减小，除非增加样本容量. 在给定样本容量的情况下，一般来说，我们总是控制犯第Ⅰ类错误的概率，使它不大于 α. α 的大小视具体情况而定，通常 α 取 0.1，0.05，0.01，0.005 等值. 这种只对犯第Ⅰ类错误的概率加以控制，而不考虑犯第Ⅱ类错误的概率的检验，称为**显著性检验**.

形如(1.2)式中的备择假设 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**，而称形如(1.2)的假设检验为**双边假设检验**.

有时，我们只关心总体均值是否增大，例如，试验新工艺以提高材料的强度. 这时，所考虑的总体的均值应该越大越好. 如果我们能判断在新工艺下总体均值较以往正常生产的大，则可考虑采用新工艺. 此时，我们需要检验假设

$$H_0:\mu\leqslant\mu_0, \quad H_1:\mu>\mu_0. \tag{1.3}$$

形如(1.3)的假设检验，称为**右边检验**. 类似地，有时我们需要检验假设

$$H_0:\mu\geqslant\mu_0, \quad H_1:\mu<\mu_0. \tag{1.4}$$

形如(1.4)的假设检验，称为**左边检验**. 右边检验和左边检验统称为**单边检验**.

下面来讨论单边检验的拒绝域.

设总体 $X\sim N(\mu,\sigma^2)$，μ 未知、σ 为已知，$X_1,X_2,\cdots,X_n$ 是来自 X 的样本. 给定显著性水平 α. 我们来求检验问题(1.3)

$$H_0:\mu\leqslant\mu_0, \quad H_1:\mu>\mu_0$$

的拒绝域.

因 H_0 中的全部 μ 都比 H_1 中的 μ 要小，当 H_1 为真时，观察值 $\overline{x}$ 往往偏大，因此，拒绝域的形式为

$$\overline{x}\geqslant k \quad (k \text{ 是某一正常数}).$$

下面来确定常数 k，其做法与例 1 中的做法类似.

$$\begin{aligned} P\{\text{当 } H_0 \text{ 为真拒绝 } H_0\} &= P_{\mu\in H_0}\{\overline{X}\geqslant k\} \\ &= P_{\mu\leqslant\mu_0}\left\{\frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}\geqslant\frac{k-\mu_0}{\sigma/\sqrt{n}}\right\} \\ &\leqslant P_{\mu\leqslant\mu_0}\left\{\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}\geqslant\frac{k-\mu_0}{\sigma/\sqrt{n}}\right\} \end{aligned}$$

$\left(\text{上式不等号成立是由于 } \mu\leqslant\mu_0, \frac{\overline{X}-\mu}{\sigma/\sqrt{n}}\geqslant\frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}, \text{事件} \left\{\frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}\geqslant\frac{k-\mu_0}{\sigma/\sqrt{n}}\right\}\subset\left\{\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}\geqslant\frac{k-\mu_0}{\sigma/\sqrt{n}}\right\}\right)$. 要控制 $P\{\text{当 } H_0 \text{ 为真拒绝 } H_0\}\leqslant\alpha$，只需令

$$P_{\mu\leqslant\mu_0}\left\{\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}\geqslant\frac{k-\mu_0}{\sigma/\sqrt{n}}\right\}=\alpha. \tag{1.5}$$

由于$\dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}}\sim N(0,1)$，由(1.5)得到$\dfrac{k-\mu_0}{\sigma/\sqrt{n}}=z_\alpha$(如图 8-2)，$k=\mu_0+\dfrac{\sigma}{\sqrt{n}}z_\alpha$，即得检验问题(1.3)的拒绝域为

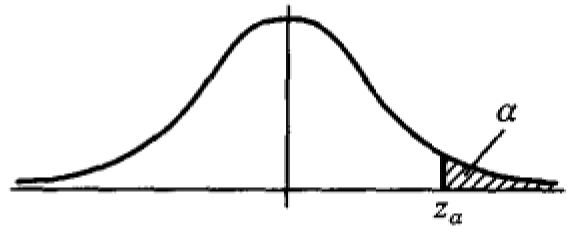

图 8-2

$$\overline{x}\geqslant\mu_0+\frac{\sigma}{\sqrt{n}}z_\alpha,$$

即
$$z=\frac{\overline{x}-\mu_0}{\sigma/\sqrt{n}}\geqslant z_\alpha. \tag{1.6}$$

类似地，可得左边检验问题(1.4)

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

的拒绝域为

$$z=\frac{\overline{x}-\mu_0}{\sigma/\sqrt{n}}\leqslant -z_\alpha. \tag{1.7}$$

例 2　公司从生产商购买牛奶。公司怀疑生产商在牛奶中掺水以谋利。通过测定牛奶的冰点，可以检验出牛奶是否掺水. 天然牛奶的冰点温度近似服从正态分布，均值 $\mu_0=-0.545$ ℃，标准差 $\sigma=0.008$ ℃. 牛奶掺水可使冰点温度升高而接近于水的冰点温度(0 ℃). 测得生产商提交的 5 批牛奶的冰点温度，其均值为$\overline{x}=-0.535$ ℃，问是否可以认为生产商在牛奶中掺了水？取 $\alpha=0.05$.

解　按题意需检验假设

$$H_0:\mu\leqslant\mu_0=-0.545\text{(即设牛奶未掺水)},$$
$$H_1:\mu>\mu_0\text{(即设牛奶已掺水)}$$

这是右边检验问题，其拒绝域如(1.6)式所示，即为

$$z=\frac{\overline{x}-\mu_0}{\sigma/\sqrt{n}}\geqslant z_{0.05}=1.645.$$

现在 $z=\dfrac{-0.535-(-0.545)}{0.008/\sqrt{5}}=2.7951>1.645$，$z$ 的值落在拒绝域中，所以我们在显著性水平 $\alpha=0.05$ 下拒绝 H_0，即认为牛奶商在牛奶中掺了水. □

综上所述，可得处理参数的假设检验问题的步骤如下：

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° 给定显著性水平 α 以及样本容量 n；

3° 确定检验统计量以及拒绝域的形式；

4° 按 $P\{$当 H_0 为真拒绝 $H_0\}\leqslant\alpha$ 求出拒绝域；

5° 取样，根据样本观察值作出决策，是接受 H_0 还是拒绝 H_0.

下面我们只讨论正态总体参数的假设检验问题.

§2　正态总体均值的假设检验

（一）单个总体 $N(\mu,\sigma^2)$ 均值 μ 的检验

1. σ^2 已知，关于 μ 的检验（Z 检验）

在§1中已讨论过正态总体 $N(\mu,\sigma^2)$ 当 σ^2 已知时关于 μ 的检验问题(1.2)，(1.3)，(1.4). 在这些检验问题中，我们都是利用统计量 $Z=\dfrac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}$ 来确定拒绝域的. 这种检验法常称为 Z **检验法**.

2. σ^2 未知，关于 μ 的检验（t 检验）

设总体 $X\sim N(\mu,\sigma^2)$，其中 μ,σ^2 未知，我们来求检验问题

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

的拒绝域（显著性水平为 α）.

设 $X_1,X_2,\cdots,X_n$ 是来自总体 X 的样本. 由于 σ^2 未知，现在不能利用 $\dfrac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}$ 来确定拒绝域了. 注意到 S^2 是 σ^2 的无偏估计，我们用 S 来代替 σ，采用

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

作为检验统计量. 当观察值 $|t|=\left|\dfrac{\overline{x}-\mu_0}{s/\sqrt{n}}\right|$ 过分大时就拒绝 H_0，拒绝域的形式为

$$|t|=\left|\frac{\overline{x}-\mu_0}{s/\sqrt{n}}\right|\geqslant k.$$

由第六章§3定理三知，当 H_0 为真时，$\dfrac{\overline{X}-\mu_0}{S/\sqrt{n}}\sim t(n-1)$，故由

$$P\{\text{当 } H_0 \text{ 为真拒绝 } H_0\}=P_{\mu_0}\left\{\left|\frac{\overline{X}-\mu_0}{S/\sqrt{n}}\right|\geqslant k\right\}=\alpha$$

得 $k=t_{\alpha/2}(n-1)$，即得拒绝域为

$$|t|=\left|\frac{\overline{x}-\mu_0}{s/\sqrt{n}}\right|\geqslant t_{\alpha/2}(n-1). \tag{2.1}$$

对于正态总体 $N(\mu,\sigma^2)$，当 σ^2 未知时，关于 μ 的单边检验的拒绝域在表 8.1 中给出.

上述利用 t 统计量得出的检验法称为 t **检验法**.

在实际中,正态总体的方差常为未知,所以我们常用 t 检验法来检验关于正态总体均值的检验问题.

例 1　某种元件的寿命 X(以 h 计)服从正态分布 $N(\mu,\sigma^2)$,μ,σ^2 均未知.现测得 16 只元件的寿命如下:

159	280	101	212	224	379	179	264
222	362	168	250	149	260	485	170

问是否有理由认为元件的平均寿命大于 225 h?

解　按题意需检验

$$H_0:\mu\leqslant\mu_0=225,\quad H_1:\mu>225.$$

取 $\alpha=0.05$.由表 8.1 知此检验问题的拒绝域为

$$t=\frac{\overline{x}-\mu_0}{s/\sqrt{n}}\geqslant t_\alpha(n-1).$$

现在 $n=16$,$t_{0.05}(15)=1.753\,1$.又算得 $\overline{x}=241.5$,$s=98.7259$,即有

$$t=\frac{\overline{x}-\mu_0}{s/\sqrt{n}}=0.668\,5<1.753\,1.$$

t 没有落在拒绝域中,故接受 H_0,即认为元件的平均寿命不大于 225 h.　□

(二) 两个正态总体均值差的检验(t 检验)

我们还可以用 t 检验法检验具有相同方差的两正态总体均值差的假设.设 $X_1,X_2,\cdots,X_{n_1}$ 是来自正态总体 $N(\mu_1,\sigma^2)$ 的样本,$Y_1,Y_2,\cdots,Y_{n_2}$ 是来自正态总体 $N(\mu_2,\sigma^2)$ 的样本,且设两样本独立.又分别记它们的样本均值为 $\overline{X},\overline{Y}$,记样本方差为 S_1^2,S_2^2.设 μ_1,μ_2,σ^2 均为未知,要特别引起注意的是,在这里假设两总体的方差是相等的.现在来求检验问题:

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

(δ 为已知常数)的拒绝域.取显著性水平为 α.

引用下述 t 统计量作为检验统计量:

$$t=\frac{(\overline{X}-\overline{Y})-\delta}{S_w\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}},$$

其中

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

当 H_0 为真时,由第六章§3 定理四知 $t\sim t(n_1+n_2-2)$.与单个总体的 t 检验法相仿,其拒绝域的形式为

$$\left|\frac{(\overline{x}-\overline{y})-\delta}{s_w\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}\right|\geqslant k.$$

由

$$P\{\text{当 } H_0 \text{ 为真拒绝 } H_0\} = P_{\mu_1-\mu_2=\delta}\left\{\left|\frac{(\overline{X}-\overline{Y})-\delta}{S_w\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}\right| \geqslant k\right\} = \alpha$$

可得 $k=t_{\alpha/2}(n_1+n_2-2)$. 于是得拒绝域为

$$|t| = \frac{|(\overline{x}-\overline{y})-\delta|}{s_w\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}} \geqslant t_{\alpha/2}(n_1+n_2-2). \tag{2.2}$$

关于均值差的两个单边检验问题的拒绝域在表 8.1 中给出. 常用的是 $\delta=0$ 的情况.

当两个正态总体的方差均为已知(不一定相等)时,我们可用 Z 检验法来检验两正态总体均值差的假设问题,见表 8.1.

例 2　用两种方法(A 和 B)测定冰自 -0.72 ℃转变为 0 ℃的水的融化热(以 cal/g 计). 测得以下的数据:

方法 A:　79.98　80.04　80.02　80.04　80.03　80.03

　　　　80.04　79.97　80.05　80.03　80.02　80.00　80.02

方法 B:　80.02　79.94　79.98　79.97　79.97　80.03　79.95　78.97

设这两个样本相互独立,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. 试检验假设(取显著性水平 $\alpha=0.05$)

$$H_0: \mu_1-\mu_2\leqslant 0,\ H_1: \mu_1-\mu_2>0.$$

解　分别画出对应于方法 A 和方法 B 的数据的箱线图,如图 8－3. 这两种方法所得的结果是有明显差异的,现在来检验上述假设.

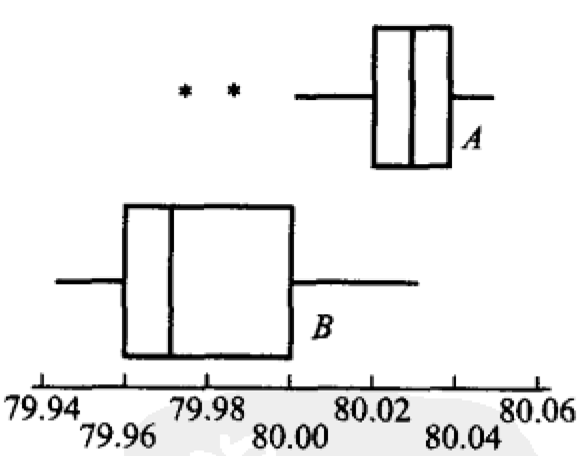

图 8－3

$n_1=13$, $\overline{x}_A=80.02$, $s_A^2=0.024^2$

$n_2=8$, $\overline{x}_B=79.98$, $s_B^2=0.03^2$

$$s_w^2=\frac{12\times s_A^2+7\times s_B^2}{19}=0.000\,717\,8.$$

$$t=\frac{\overline{x}_A-\overline{x}_B}{s_w\sqrt{1/13+1/8}}=3.33>t_{0.05}(13+8-2)=1.729\,1.$$

故拒绝 H_0,认为方法 A 比方法 B 测得的融化热要大. □

(三) 基于成对数据的检验(t 检验)

有时为了比较两种产品、两种仪器、两种方法等的差异,我们常在相同的条件下做对比试验,得到一批成对的观察值. 然后分析观察数据作出推断. 这种方

法常称为**逐对比较法**.

例 3　有两台光谱仪 I_x,I_y,用来测量材料中某种金属的含量,为鉴定它们的测量结果有无显著的差异,制备了 9 件试块(它们的成分、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块测量一次,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　本题中的数据是成对的,即对同一试块测出一对数据.我们看到一对与另一对之间的差异是由各种因素,如材料成分、金属含量、均匀性等因素引起的.由于各试块的特性有广泛的差别,就不能将仪器 I_x 对 9 个试块的测量结果(即表中第一行)看成是同分布随机变量的观察值.因而表中第一行不能看成是一个样本的样本值.同样,表中第二行也不能看成是一个样本的样本值.再者,对于每一对数据而言,它们是同一试块用不同仪器 I_x,I_y 测得的结果,因此,它们不是两个独立的随机变量的观察值.综上所述,我们不能用表 8.1 中第 4 栏的检验法来作检验.而同一对中两个数据的差异则可看成是仅由这两台仪器性能的差异所引起的,这样,局限于各对中两个数据来比较就能排除种种其他因素,而只考虑单独由仪器的性能所产生的影响.从而能比较这两台仪器的测量结果是否有显著的差异.

一般,设有 n 对相互独立的观察结果:$(X_1,Y_1),(X_2,Y_2),\cdots,(X_n,Y_n)$,令 $D_1=X_1-Y_1,D_2=X_2-Y_2,\cdots,D_n=X_n-Y_n$,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.今假设 $D_i\sim N(\mu_D,\sigma_D^2),i=1,2,\cdots,n$.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.我们需要基于这一样本检验假设:

(1)　$H_0:\mu_D=0$,　$H_1:\mu_D\neq 0$;

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

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

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 $\overline{d},s_D^2$,按表 8.1 第 2 栏中关于单个正态总体均值的 t 检验.知检验问题(1),(2),(3)的拒绝域分别为(显著性水平为 α):

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

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

$$t=\frac{\overline{d}}{s_D/\sqrt{n}}\leqslant -t_\alpha(n-1).$$

现在回过来讨论本例的检验问题．先作出同一试块分别由仪器 I_x, I_y 测得的结果之差，列于上表的第三行．按题意需检验假设

$$H_0:\mu_D=0,\quad H_1:\mu_D\neq 0.$$

现在 $n=9, t_{\alpha/2}(8)=t_{0.005}(8)=3.355\,4$ 即知拒绝域为

$$|t|=\left|\frac{\overline{d}}{s_D/\sqrt{n}}\right|\geqslant 3.3554.$$

由观察值得$\overline{d}=0.06, s_D=0.122\,7, |t|=\dfrac{0.06}{0.122\,7/\sqrt{9}}=1.467<3.3554$. 现$|t|$的值不落在拒绝域内，故接受 H_0，认为两台仪器的测量结果并无显著差异. □

例 4 做以下的实验以比较人对红光或绿光的反应时间(以 s 计). 实验在点亮红光或绿光的同时，启动计时器，要求受试者见到红光或绿光点亮时，就按下按钮，切断计时器，这就能测得反应时间. 测量的结果如下表：

红光(x)	0.30	0.23	0.41	0.53	0.24	0.36	0.38	0.51
绿光(y)	0.43	0.32	0.58	0.46	0.27	0.41	0.38	0.61
$d=x-y$	−0.13	−0.09	−0.17	0.07	−0.03	−0.05	0.00	−0.10

设 $D_i=X_i-Y_i\quad(i=1,2,\cdots,8)$是来自正态总体 $N(\mu_D,\sigma_D^2)$的样本，μ_D,σ_D^2均未知. 试检验假设(取显著性水平 $\alpha=0.05$)

$$H_0:\mu_D\geqslant 0,\quad H_1:\mu_D<0.$$

解 现在 $n=8, \overline{x}_d=-0.0625, s_d=0.0765$，而

$$\frac{\overline{x}_d}{s_d/\sqrt{8}}=-2.311<-t_{0.05}(7)=-1.894\,6,$$

故拒绝 H_0，认为 $\mu_D<0$，即认为人对红光的反应时间小于对绿光的反应时间，也就是人对红光的反应要比绿光快. □

§3 正态总体方差的假设检验

现在来讨论有关正态总体方差的假设检验问题. 以下分单个总体和两个总体的情况来讨论.

(一) 单个总体的情况

设总体 $X\sim N(\mu,\sigma^2)$，μ,σ^2 均未知，$X_1,X_2,\cdots,X_n$ 是来自 X 的样本. 要求检

验假设(显著性水平为 α)

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

σ_0^2 为已知常数.

由于 S^2 是 σ^2 的无偏估计,当 H_0 为真时,观察值 s^2 与 σ_0^2 的比值 $\frac{s^2}{\sigma_0^2}$ 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. 由第六章 §3 定理二知当 H_0 为真时

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

我们取

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

作为检验统计量,如上所说知道上述检验问题的拒绝域具有以下的形式:

$$\frac{(n-1)s^2}{\sigma_0^2}\leqslant k_1\quad\text{或}\quad\frac{(n-1)s^2}{\sigma_0^2}\geqslant k_2,$$

此处 k_1,k_2 的值由下式确定:

$$\begin{aligned}&P\{\text{当 } H_0 \text{ 为真拒绝 } H_0\}\\&=P_{\sigma_0^2}\left\{\left(\frac{(n-1)S^2}{\sigma_0^2}\leqslant k_1\right)\cup\left(\frac{(n-1)S^2}{\sigma_0^2}\geqslant k_2\right)\right\}=\alpha.\end{aligned}$$

为计算方便起见,习惯上取

$$P_{\sigma_0^2}\left\{\frac{(n-1)S^2}{\sigma_0^2}\leqslant k_1\right\}=\frac{\alpha}{2},P_{\sigma_0^2}\left\{\frac{(n-1)S^2}{\sigma_0^2}\geqslant k_2\right\}=\frac{\alpha}{2},$$

故得 $k_1=\chi^2_{1-\alpha/2}(n-1),k_2=\chi^2_{\alpha/2}(n-1)$. 于是得拒绝域为

$$\frac{(n-1)s^2}{\sigma_0^2}\leqslant\chi^2_{1-\alpha/2}(n-1)\quad\text{或}\quad\frac{(n-1)s^2}{\sigma_0^2}\geqslant\chi^2_{\alpha/2}(n-1)\text{①}.\tag{3.1}$$

下面来求单边检验问题(显著性水平为 α)

$$H_0:\sigma^2\leqslant\sigma_0^2,\ H_1:\sigma^2>\sigma_0^2\tag{3.2}$$

的拒绝域. 因 H_0 中的全部 σ^2 都比 H_1 中的 σ^2 要小,当 H_1 为真时,S^2 的观察值 s^2 往往偏大,因此拒绝域的形式为

$$s^2\geqslant k.$$

下面来确定常数 k.

$$\begin{aligned}P\{\text{当 } H_0 \text{ 为真拒绝 } H_0\}&=P_{\sigma^2\leqslant\sigma_0^2}\{S^2\geqslant k\}\\&=P_{\sigma^2\leqslant\sigma_0^2}\left\{\frac{(n-1)S^2}{\sigma_0^2}\geqslant\frac{(n-1)k}{\sigma_0^2}\right\}\end{aligned}$$

---

① 这里指的是 $\frac{(n-1)s^2}{\sigma_0^2}\leqslant\chi^2_{1-\alpha/2}(n-1)$ 与 $\frac{(n-1)s^2}{\sigma_0^2}\geqslant\chi^2_{\alpha/2}(n-1)$ 的并集.

$$\leqslant P_{\sigma^2\leqslant\sigma_0^2}\left\{\frac{(n-1)S^2}{\sigma^2}\geqslant\frac{(n-1)k}{\sigma_0^2}\right\}\ (\text{因为 } \sigma^2\leqslant\sigma_0^2).$$

要控制 $P\{$当 H_0 为真拒绝 $H_0\}\leqslant\alpha$，只需令

$$P_{\sigma^2\leqslant\sigma_0^2}\left\{\frac{(n-1)S^2}{\sigma^2}\geqslant\frac{(n-1)k}{\sigma_0^2}\right\}=\alpha. \tag{3.3}$$

因 $\frac{(n-1)S^2}{\sigma^2}\sim\chi^2(n-1)$，由(3.3)得 $\frac{(n-1)k}{\sigma_0^2}=\chi_\alpha^2(n-1)$（见图 8－4）. 于是 $k=\frac{\sigma_0^2}{n-1}\chi_\alpha^2(n-1)$，得检验问题(3.2)的拒绝域为 $s^2\geqslant\frac{\sigma_0^2}{n-1}\chi_\alpha^2(n-1)$，即

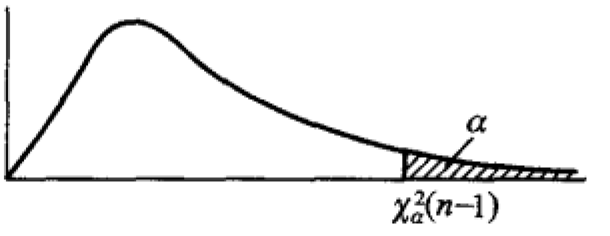

图 8－4

$$\chi^2=\frac{(n-1)s^2}{\sigma_0^2}\geqslant\chi_\alpha^2(n-1). \tag{3.4}$$

类似地，可得左边检验问题

$$H_0:\sigma^2\geqslant\sigma_0^2,\quad H_1:\sigma^2<\sigma_0^2$$

的拒绝域为

$$\chi^2=\frac{(n-1)s^2}{\sigma_0^2}\leqslant\chi_{1-\alpha}^2(n-1). \tag{3.5}$$

以上检验法称为 χ^2 **检验法**.

表 8－1　正态总体均值、方差的检验法(显著性水平为 α)

	原假设 H_0	检验统计量	备择假设 H_1	拒绝域
1	$\mu\leqslant\mu_0$ $\mu\geqslant\mu_0$ $\mu=\mu_0$ (σ^2 已知)	$Z=\frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}$	$\mu>\mu_0$ $\mu<\mu_0$ $\mu\neq\mu_0$	$z\geqslant z_\alpha$ $z\leqslant -z_\alpha$ $\lvert z\rvert\geqslant z_{\alpha/2}$
2	$\mu\leqslant\mu_0$ $\mu\geqslant\mu_0$ $\mu=\mu_0$ (σ^2 未知)	$t=\frac{\overline{X}-\mu_0}{S/\sqrt{n}}$	$\mu>\mu_0$ $\mu<\mu_0$ $\mu\neq\mu_0$	$t\geqslant t_\alpha(n-1)$ $t\leqslant -t_\alpha(n-1)$ $\lvert t\rvert\geqslant t_{\alpha/2}(n-1)$
3	$\mu_1-\mu_2\leqslant\delta$ $\mu_1-\mu_2\geqslant\delta$ $\mu_1-\mu_2=\delta$ (σ_1^2,σ_2^2 已知)	$Z=\frac{\overline{X}-\overline{Y}-\delta}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}$	$\mu_1-\mu_2>\delta$ $\mu_1-\mu_2<\delta$ $\mu_1-\mu_2\neq\delta$	$z\geqslant z_\alpha$ $z\leqslant -z_\alpha$ $\lvert z\rvert\geqslant z_{\alpha/2}$

续表

	原假设 H_0	检验统计量	备择假设 H_1	拒绝域
4	$\mu_1-\mu_2\leqslant\delta$ $\mu_1-\mu_2\geqslant\delta$ $\mu_1-\mu_2=\delta$ ($\sigma_1^2=\sigma_2^2=\sigma^2$ 未知)	$t=\dfrac{\overline{X}-\overline{Y}-\delta}{S_w\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}$ $S_w^2=\dfrac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}$	$\mu_1-\mu_2>\delta$ $\mu_1-\mu_2<\delta$ $\mu_1-\mu_2\neq\delta$	$t\geqslant t_\alpha(n_1+n_2-2)$ $t\leqslant -t_\alpha(n_1+n_2-2)$ $\lvert t\rvert\geqslant t_{\alpha/2}(n_1+n_2-2)$
5	$\sigma^2\leqslant\sigma_0^2$ $\sigma^2\geqslant\sigma_0^2$ $\sigma^2=\sigma_0^2$ (μ 未知)	$\chi^2=\dfrac{(n-1)S^2}{\sigma_0^2}$	$\sigma^2>\sigma_0^2$ $\sigma^2<\sigma_0^2$ $\sigma^2\neq\sigma_0^2$	$\chi^2\geqslant\chi_\alpha^2(n-1)$ $\chi^2\leqslant\chi_{1-\alpha}^2(n-1)$ $\chi^2\geqslant\chi_{\alpha/2}^2(n-1)$或 $\chi^2\leqslant\chi_{1-\alpha/2}^2(n-1)$
6	$\sigma_1^2\leqslant\sigma_2^2$ $\sigma_1^2\geqslant\sigma_2^2$ $\sigma_1^2=\sigma_2^2$ (μ_1,μ_2 未知)	$F=\dfrac{S_1^2}{S_2^2}$	$\sigma_1^2>\sigma_2^2$ $\sigma_1^2<\sigma_2^2$ $\sigma_1^2\neq\sigma_2^2$	$F\geqslant F_\alpha(n_1-1,n_2-1)$ $F\leqslant F_{1-\alpha}(n_1-1,n_2-1)$ $F\geqslant F_{\alpha/2}(n_1-1,n_2-1)$或 $F\leqslant F_{1-\alpha/2}(n_1-1,n_2-1)$
7	$\mu_D\leqslant 0$ $\mu_D\geqslant 0$ $\mu_D=0$ (成对数据)	$t=\dfrac{\overline{D}-0}{S_D/\sqrt{n}}$	$\mu_D>0$ $\mu_D<0$ $\mu_D\neq 0$	$t\geqslant t_\alpha(n-1)$ $t\leqslant -t_\alpha(n-1)$ $\lvert t\rvert\geqslant t_{\alpha/2}(n-1)$

例 1　某厂生产的某种型号的电池，其寿命(以 h 计)长期以来服从方差 $\sigma^2=5\ 000$ 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 $s^2=9\ 200$. 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　本题要求在水平 $\alpha=0.02$ 下检验假设

$$H_0:\sigma^2=5\ 000,\quad H_1:\sigma^2\neq 5\ 000.$$

现在 $n=26$, $\chi_{\alpha/2}^2(n-1)=\chi_{0.01}^2(25)=44.314$, $\chi_{1-\alpha/2}^2(25)=\chi_{0.99}^2(25)=11.524$. $\sigma_0^2=5\ 000$，由(3.1)式拒绝域为

$$\frac{(n-1)s^2}{\sigma_0^2}\geqslant 44.314\quad 或\quad \frac{(n-1)s^2}{\sigma_0^2}\leqslant 11.524.$$

由观察值 $s^2=9\ 200$ 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的变化. □

（二）两个总体的情况

设 $X_1,X_2,\cdots,X_{n_1}$ 是来自总体 $N(\mu_1,\sigma_1^2)$ 的样本，$Y_1,Y_2,\cdots,Y_{n_2}$ 是来自总体 $N(\mu_2,\sigma_2^2)$ 的样本，且两样本独立. 其样本方差分别为 S_1^2,S_2^2. 且设 $\mu_1,\mu_2,\sigma_1^2,\sigma_2^2$ 均为未知. 现在需要检验假设（显著性水平为 α）

$$H_0:\sigma_1^2\leqslant\sigma_2^2,\quad H_1:\sigma_1^2>\sigma_2^2. \tag{3.6}$$

当 H_0 为真时，$E(S_1^2)=\sigma_1^2\leqslant\sigma_2^2=E(S_2^2)$，当 H_1 为真时 $E(S_1^2)=\sigma_1^2>\sigma_2^2=E(S_2^2)$. 当 H_1 为真时，观察值 $\dfrac{S_1^2}{S_2^2}$ 有偏大的趋势，故拒绝域具有形式

$$\frac{s_1^2}{s_2^2}\geqslant k,$$

常数 k 确定如下：

$$\begin{aligned}P\{\text{当 } H_0 \text{ 为真拒绝 } H_0\}&=P_{\sigma_1^2\leqslant\sigma_2^2}\left\{\frac{S_1^2}{S_2^2}\geqslant k\right\}\\&\leqslant P_{\sigma_1^2\leqslant\sigma_2^2}\left\{\frac{S_1^2/S_2^2}{\sigma_1^2/\sigma_2^2}\geqslant k\right\}\ (\text{因为 } \sigma_1^2/\sigma_2^2\leqslant 1).\end{aligned}$$

要控制 $P\{$当 H_0 为真拒绝 $H_0\}\leqslant\alpha$，只需令

$$P_{\sigma_1^2\leqslant\sigma_2^2}\left\{\frac{S_1^2/S_2^2}{\sigma_1^2/\sigma_2^2}\geqslant k\right\}=\alpha. \tag{3.7}$$

由第六章 §3 定理四知 $\dfrac{S_1^2/S_2^2}{\sigma_1^2/\sigma_2^2}\sim F(n_1-1,n_2-1)$，由(3.7)式得 $k=F_\alpha(n_1-1,n_2-1)$. 即得检验问题(3.6)的拒绝域为

$$F=\frac{s_1^2}{s_2^2}\geqslant F_\alpha(n_1-1,n_2-1). \tag{3.8}$$

上述检验法称为 **F 检验法**. 关于 σ_1^2,σ_2^2 的另外两个检验问题的拒绝域在表 8.1 中给出.

例 2　设 §2 例 2 中的两个样本分别来自总体 $N(\mu_A,\sigma_A^2)$，$N(\mu_B,\sigma_B^2)$，且两样本独立. 试检验 $H_0:\sigma_A^2=\sigma_B^2$，$H_1:\sigma_A^2\neq\sigma_B^2$，以说明我们假设 $\sigma_A^2=\sigma_B^2$ 是合理的.（取显著性水平 $\alpha=0.01$）

解　此处 $n_1=13,n_2=8,\alpha=0.01$，拒绝域为

$$\frac{s_A^2}{s_B^2}\geqslant F_{0.005}(12,7)=8.18,$$

或

$$\frac{s_A^2}{s_B^2}\leqslant F_{0.995}(12,7)=\frac{1}{F_{0.005}(7,12)}=\frac{1}{5.52}=0.18.$$

现在 $s_A^2=0.024^2,s_B^2=0.03^2,s_A^2/s_B^2=0.64$，

$$0.18<0.64<8.18,$$

故接受 H_0，认为两总体方差相等. 两总体方差相等也称两总体具有**方差齐性**，这也表明 § 2 例 2 中假设两总体方差相等是合理的. □

§ 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有明显的联系，先考察置信区间与双边检验之间的对应关系. 设 $X_1,\cdots,X_n$ 是一个来自总体的样本，$x_1,\cdots,x_n$ 是相应的样本值，Θ 是参数 θ 的可能取值范围.

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

$$P_\theta\{\underline{\theta}(X_1,\cdots,X_n)<\theta<\overline{\theta}(X_1,\cdots,X_n)\}\geqslant 1-\alpha, \tag{4.1}$$

考虑显著性水平为 α 的双边检验

$$H_0:\theta=\theta_0,\ \ H_1:\theta\neq\theta_0. \tag{4.2}$$

由(4.1)式

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

即有

$$P_{\theta_0}\{(\theta_0\leqslant\underline{\theta}(X_1,\cdots,X_n))\cup(\theta_0\geqslant\overline{\theta}(X_1,\cdots,X_n))\}\leqslant\alpha.$$

按显著性水平为 α 的假设检验的拒绝域的定义，检验(4.2)的拒绝域为

$$\theta_0\leqslant\underline{\theta}(x_1,\cdots,x_n)\quad 或\quad \theta_0\geqslant\overline{\theta}(x_1,\cdots,x_n);$$

接受域为

$$\underline{\theta}(x_1,\cdots,x_n)<\theta_0<\overline{\theta}(x_1,\cdots,x_n).$$

这就是说，当我们要检验假设(4.2)时，先求出 θ 的置信水平为 $1-\alpha$ 的置信区间 $(\underline{\theta},\overline{\theta})$，然后考察区间 $(\underline{\theta},\overline{\theta})$ 是否包含 θ_0，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，对于任意 $\theta_0\in\Theta$，考虑显著性水平为 α 的假设检验问题

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

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

$$P_{\theta_0}\{\underline{\theta}(X_1,\cdots,X_n)<\theta_0<\overline{\theta}(X_1,\cdots,X_n)\}\geqslant 1-\alpha.$$

由 θ_0 的任意性，由上式知对于任意 $\theta\in\Theta$，有

$$P_\theta\{\underline{\theta}(X_1,\cdots,X_n)<\theta<\overline{\theta}(X_1,\cdots,X_n)\}\geqslant 1-\alpha.$$

因此 $(\underline{\theta}(X_1,\cdots,X_n),\overline{\theta}(X_1,\cdots,X_n))$ 是参数 θ 的一个置信水平为 $1-\alpha$ 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：$H_0:\theta=\theta_0$，$H_1:\theta\neq\theta_0$ 的接受域 $\underline{\theta}(x_1,\cdots,x_n)<\theta_0<\bar{\theta}(x_1,\cdots,x_n)$，那么，$(\underline{\theta}(X_1,\cdots,X_n),\bar{\theta}(X_1,\cdots,X_n))$ 就是 θ 的置信水平为 $1-\alpha$ 的置信区间.

还可验证，置信水平为 $1-\alpha$ 的单侧置信区间 $(-\infty,\bar{\theta}(X_1,\cdots,X_n))$ 与显著性水平为 α 的左边检验问题 $H_0:\theta\geqslant\theta_0$，$H_1:\theta<\theta_0$ 有类似的对应关系. 即若已求得单侧置信区间 $(-\infty,\bar{\theta}(X_1,\cdots,X_n))$，则当 $\theta_0\in(-\infty,\bar{\theta}(x_1,\cdots,x_n))$ 时接受 H_0，当 $\theta_0\notin(-\infty,\bar{\theta}(x_1,\cdots,x_n))$ 时拒绝 H_0. 反之，若已求得检验问题 $H_0:\theta\geqslant\theta_0$，$H_1:\theta<\theta_0$ 的接受域为 $-\infty<\theta_0\leqslant\bar{\theta}(x_1,\cdots,x_n)$，则可得 θ 的一个单侧置信区间 $(-\infty,\bar{\theta}(X_1,\cdots,X_n))$.

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 α 的右边检验问题 $H_0:\theta\leqslant\theta_0$，$H_1:\theta>\theta_0$ 也有类似的对应关系. 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 $H_0:\theta\leqslant\theta_0$，$H_1:\theta>\theta_0$ 的接受域为 $\underline{\theta}(x_1,\cdots,x_n)\leqslant\theta_0<\infty$，则可得 θ 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，$\alpha=0.05$，$n=16$，且由一样本算得 $\bar{x}=5.20$，于是得到参数 μ 的一个置信水平为 0.95 的置信区间

$$\left(\bar{x}-\frac{1}{\sqrt{16}}z_{0.025},\bar{x}+\frac{1}{\sqrt{16}}z_{0.025}\right)=(5.20-0.49,5.20+0.49)$$

$$=(4.71,5.69).$$

现在考虑检验问题 $H_0:\mu=5.5$，　$H_1:\mu\neq5.5$. 由于 $5.5\in(4.71,5.69)$，故接受 H_0. □

例 2 数据如上例. 试求右边检验问题 $H_0:\mu\leqslant\mu_0$，$H_1:\mu>\mu_0$ 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 $z=\dfrac{\bar{x}-\mu_0}{1/\sqrt{16}}\geqslant z_{0.05}$，或即 $\mu_0\leqslant4.79$. 于是检验问题的接受域为 $\mu_0>4.79$. 这样就得到 μ 的单侧置信区间 $(4.79,\infty)$，单侧置信下限 $\underline{\mu}=4.79$. □

§5 样本容量的选取

以上我们在进行假设检验时，总是根据问题的要求，预先给出显著性水平以控制犯第Ⅰ类错误的概率，而犯第Ⅱ类错误的概率则依赖于样本容量的选择. 在

一些实际问题中，我们除了希望控制犯第Ⅰ类错误的概率外，往往还希望控制犯第Ⅱ类错误的概率. 在这一节，我们将阐明如何选取样本的容量使得犯第Ⅱ类错误的概率控制在预先给定的限度之内. 为此，我们引入施行特征函数.

定义 若 C 是参数 θ 的某检验问题的一个检验法，

$$\beta(\theta)=P_\theta(\text{接受 } H_0) \tag{5.1}$$

称为检验法 C 的**施行特征函数**或 **OC 函数**，其图形称为 **OC 曲线**.

由定义知，若此检验法的显著性水平为 α，那么当真值 $\theta\in H_0$ 时，$\beta(\theta)$ 就是作出正确判断（即 H_0 为真时接受 H_0）的概率，故此时 $\beta(\theta)\geqslant 1-\alpha$；而当 $\theta\in H_1$ 时，则 $\beta(\theta)$ 就是犯第Ⅱ类错误的概率，而 $1-\beta(\theta)$ 是作出正确判断（即 H_0 为不真时拒绝 H_0）的概率. 函数 $1-\beta(\theta)$ 称为检验法 C 的**功效函数**. 当 $\theta^*\in H_1$ 时，值 $1-\beta(\theta^*)$ 称为检验法 C 在点 θ^* 的**功效**. 它表示当参数 θ 的真值为 θ^* 时，检验法 C 作出正确判断的概率.

本书只介绍正态总体均值的检验法的 OC 函数及其图形.

1. Z 检验法的 OC 函数

右边检验问题. $H_0:\mu\leqslant\mu_0$， $H_1:\mu>\mu_0$ 的 OC 函数是

$$\begin{aligned}\beta(\mu)&=P_\mu(\text{接受 } H_0)=P_\mu\left\{\frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}<z_\alpha\right\}\\&=P_\mu\left\{\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}<z_\alpha-\frac{\mu-\mu_0}{\sigma/\sqrt{n}}\right\}=\Phi(z_\alpha-\lambda),\end{aligned} \tag{5.2}$$

其中 $\lambda=\dfrac{\mu-\mu_0}{\sigma/\sqrt{n}}$. 其图形如图 8－5 所示. 此 OC 函数 $\beta(\mu)$ 有如下性质：

(1) 它是 $\lambda=\dfrac{\mu-\mu_0}{\sigma/\sqrt{n}}$ 的单调递减连续函数；

(2) $\lim\limits_{\mu\to\mu_0^+}\beta(\mu)=1-\alpha$， $\lim\limits_{\mu\to\infty}\beta(\mu)=0$.

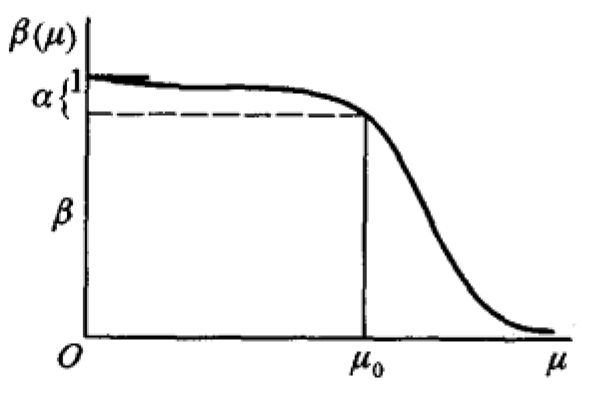

图 8－5

由 $\beta(\mu)$ 的连续性可知，当参数的真值 $\mu(\mu>\mu_0)$ 在 μ_0 附近时，检验法的功效很低，即 $\beta(\mu)$ 的值很大，亦即犯第Ⅱ类错误的概率很大. 因为 α 通常取得比较小，而不管 σ 多么小，n 多么大，只要 n 给定，总存在 μ_0 附近的点 $\mu(\mu>\mu_0)$ 使 $\beta(\mu)$ 几乎等于 $1-\alpha$.

这表明，无论样本容量 n 多么大，要想对所有 $\mu\in H_1$，即真值为 H_1 所规定的任一点时，控制犯第Ⅱ类错误的概率都很小是不可能的. 但是我们可以使用 OC 函数 $\beta(\mu)$ 以确定样本容量 n，使当真值 $\mu\geqslant\mu_0+\delta$ （$\delta>0$ 为取定的值）时，犯第Ⅱ类错误的概率不超过给定的 β. 这是由于 $\beta(\mu)$ 是 μ 的递减函数，故当 $\mu\geqslant\mu_0+\delta$

时有

$$\beta(\mu_0+\delta)\geqslant\beta(\mu).$$

于是只要 $\beta(\mu_0+\delta)=\Phi(z_\alpha-\sqrt{n}\delta/\sigma)\leqslant\beta$,亦即只要 n 满足

$$z_\alpha-\sqrt{n}\delta/\sigma\leqslant-z_\beta$$

即可.这就是说,只要

$$\sqrt{n}\geqslant\frac{(z_\alpha+z_\beta)\sigma}{\delta}, \tag{5.3}$$

就能使当 $\mu\in H_1$ 且 $\mu\geqslant\mu_0+\delta$ 时(即真值 $\mu\geqslant\mu_0+\delta$ 时)犯第Ⅱ类错误的概率不超过 β.

类似地,可得左边检验问题 $H_0:\mu\geqslant\mu_0,H_1:\mu<\mu_0$ 的 OC 函数为

$$\beta(\mu)=\Phi(z_\alpha+\lambda),\quad \lambda=\frac{\mu-\mu_0}{\sigma/\sqrt{n}}. \tag{5.4}$$

当真值 $\mu\geqslant\mu_0$ 时 $\beta(\mu)$ 为作出正确判断的概率;当真值 $\mu<\mu_0$ 时,$\beta(\mu)$ 给出犯第Ⅱ类错误的概率.只要样本容量 n 满足

$$\sqrt{n}\geqslant\frac{(z_\alpha+z_\beta)\sigma}{\delta}, \tag{5.5}$$

就能使当 $\mu\in H_1$ 且 $\mu\leqslant\mu_0-\delta(\delta>0$,为取定的值)时,犯第Ⅱ类错误的概率不超过给定的值 β.

双边检验问题 $H_0:\mu=\mu_0$, $H_1:\mu\neq\mu_0$ 的 OC 函数是

$$\begin{aligned}\beta(\mu)&=P_\mu(\text{接受 }H_0)=P_\mu\left\{-z_{\alpha/2}<\frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}<z_{\alpha/2}\right\}\\&=P_\mu\left\{-\lambda-z_{\alpha/2}<\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}<-\lambda+z_{\alpha/2}\right\}=\Phi(z_{\alpha/2}-\lambda)-\Phi(-z_{\alpha/2}-\lambda)\\&=\Phi(z_{\alpha/2}-\lambda)+\Phi(z_{\alpha/2}+\lambda)-1,\lambda=\frac{\mu-\mu_0}{\sigma/\sqrt{n}}.\end{aligned} \tag{5.6}$$

OC 曲线如图 8－6 所示.注意 $\beta(\mu)$ 是 $|\lambda|$ 的严格单调下降函数.

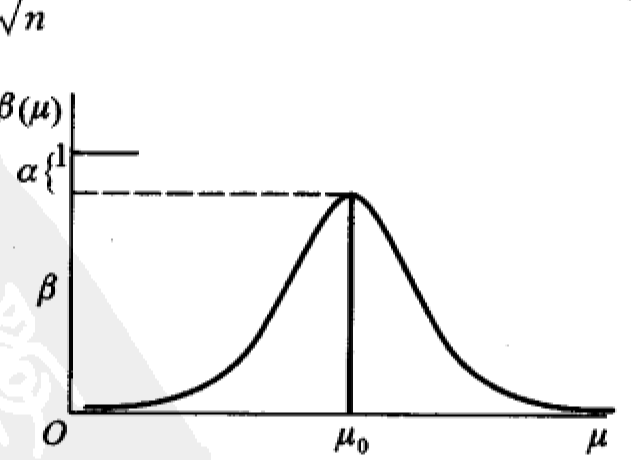

图 8－6

在双边检验问题中,若要求对 H_1 中满足 $|\mu-\mu_0|\geqslant\delta>0$ 的 μ 处的函数值 $\beta(\mu)\leqslant\beta$,则需解超越方程

$$\beta=\Phi(z_{\alpha/2}-\sqrt{n}\delta/\sigma)+\Phi(z_{\alpha/2}+\sqrt{n}\delta/\sigma)-1$$

才能确定 n.通常因 n 较大,故总可以认为 $z_{\alpha/2}+\sqrt{n}\delta/\sigma\geqslant4$,于是 $\Phi(z_{\alpha/2}+\sqrt{n}\delta/\sigma)\approx1$,故近似地有

$$\beta \approx \Phi(z_{\alpha/2}-\sqrt{n}\delta/\sigma).$$

由此知只要样本容量 n 满足

$$z_{\alpha/2}-\sqrt{n}\delta/\sigma \leqslant -z_\beta,$$

即只要 n 满足

$$\sqrt{n} \geqslant (z_{\alpha/2}+z_\beta)\frac{\sigma}{\delta}, \tag{5.7}$$

就能使当 $\mu \in H_1$ 且 $|\mu-\mu_0| \geqslant \delta$ ($\delta>0$，为取定的值)时，犯第Ⅱ类错误的概率不超过给定的值 β.

例 1(工业产品质量抽验方案)设有一大批产品，产品质量指标 $X \sim N(\mu,\sigma^2)$. 以 μ 小者为佳，厂方要求所确定的验收方案对高质量的产品($\mu \leqslant \mu_0$)能以高概率 $1-\alpha$ 为买方所接受. 买方则要求低质产品($\mu \geqslant \mu_0+\delta, \delta>0$)能以高概率 $1-\beta$ 被拒绝. α,β 由厂方与买方协商给出. 并采取一次抽样以确定该批产品是否为买方所接受. 问应怎样安排抽样方案. 已知 $\mu_0=120, \delta=20$，且由工厂长期经验知 $\sigma^2=900$. 又经商定 α,β 均取为 0.05.

解　检验问题可表达为

$$H_0: \mu \leqslant \mu_0, \quad H_1: \mu > \mu_0, \tag{5.8}$$

且要求当 $\mu \geqslant \mu_1+\delta$ 时能以 $1-\beta=0.95$ 的概率拒绝 H_0. 由 Z 检验，拒绝域为

$$\frac{\overline{x}-\mu_0}{\sigma/\sqrt{n}} \geqslant z_\alpha,$$

故 OC 函数为

$$\beta(\mu)=P_\mu\left\{\frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}<z_\alpha\right\}=P_\mu\left\{\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}<z_\alpha-\frac{\mu-\mu_0}{\sigma/\sqrt{n}}\right\}$$

$$=\Phi\left(z_\alpha-\frac{\mu-\mu_0}{\sigma/\sqrt{n}}\right). \tag{5.9}$$

现要求当 $\mu \geqslant \mu_0+\delta$ 时 $\beta(\mu) \leqslant \beta$. 因 $\beta(\mu)$ 是 μ 的递减函数，故只需 $\beta(\mu_0+\delta)=\beta$ 即可. 此时，由(5.9)式可得

$$\sqrt{n} \geqslant \frac{(z_\alpha+z_\beta)\sigma}{\delta}.$$

按给定的数据算得 $n \geqslant 24.35$，故取 $n=25$. 且当 $\overline{x}$ 满足 $\dfrac{\overline{x}-\mu_0}{\sigma/\sqrt{n}} \geqslant z_\alpha=z_{0.05}=1.645$ 时，即当 $\overline{x} \geqslant 129.87$ 时，买方就拒绝这批产品，而当 $\overline{x}<129.87$ 时，买方接受这批产品.　□

2. t 检验法的 OC 函数

右边检验问题 $H_0: \mu \leqslant \mu_0, H_1: \mu > \mu_0$ 的 t 检验法的 OC 函数是

$$\beta(\mu)=P_\mu(\text{接受 } H_0)=P_\mu\left\{\frac{\overline{X}-\mu_0}{S/\sqrt{n}}<t_\alpha(n-1)\right\}, \tag{5.10}$$

其中变量

$$\frac{\overline{X}-\mu_0}{S/\sqrt{n}}=\left(\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}+\lambda\right)\Big/\left(\frac{S}{\sigma}\right),\ \lambda=\frac{\mu-\mu_0}{\sigma/\sqrt{n}}. \tag{5.11}$$

我们称变量$\dfrac{\overline{X}-\mu_0}{S/\sqrt{n}}$服从非中心参数为$\lambda$、自由度为$n-1$的非中心$t$分布. 在$\lambda=0$时,它是通常的$t(n-1)$变量.

若给定α,β以及$\delta>0$,则可从书末附表7查得所需容量n,使得当$\mu\in H_1$且$\dfrac{\mu-\mu_0}{\sigma}\geqslant\delta$时犯第Ⅱ类错误的概率不超过$\beta$.

若给定α,β以及$\delta>0$,对于左边检验问题$H_0:\mu\geqslant\mu_0,H_1:\mu<\mu_0$的$t$检验法,也可从附表7查得所需容量$n$,使得当$\mu\in H_1$且$\dfrac{\mu-\mu_0}{\sigma}\leqslant-\delta$时犯第Ⅱ类错误的概率不超过$\beta$. 对于双边检验问题$H_0:\mu=\mu_0,H_1:\mu\neq\mu_0$的$t$检验法也可从附表7查得所需容量$n$,使得当$\mu\in H_1$,且$\dfrac{|\mu-\mu_0|}{\sigma}\geqslant\delta$时犯第Ⅱ类错误的概率不超过$\beta$.

例 2 考虑在显著性水平$\alpha=0.05$下进行t检验

$$H_0:\mu\leqslant 68,\quad H_1:\mu>68.$$

(1) 要求在H_1中$\mu\geqslant\mu_1=68+\sigma$时犯第Ⅱ类错误的概率不超过$\beta=0.05$. 求所需的样本容量.

(2) 若样本容量为$n=30$,问在H_1中$\mu=\mu_1=68+0.75\sigma$时犯第Ⅱ类错误的概率是多少?

解 (1) 此处$\alpha=\beta=0.05,\mu_0=68,\delta=\dfrac{\mu_1-\mu_0}{\sigma}=\dfrac{(68+\sigma)-68}{\sigma}=1$,查附表7得$n=13$.

(2) 现在$\alpha=0.05,n=30,\delta=\dfrac{\mu_1-\mu_0}{\sigma}=\dfrac{(68+0.75\sigma)-68}{\sigma}=0.75$,查附表7,得$\beta=0.01$. □

例 3 考虑在显著性水平$\alpha=0.05$下进行t检验

$$H_0:\mu=14,\quad H_1:\mu\neq 14.$$

要求在H_1中$\dfrac{|\mu-14|}{\sigma}\geqslant 0.4$时犯第Ⅱ类错误的概率不超过$\beta=0.1$,求所需样本容量.

解 此处$\alpha=0.05,\beta=0.1,\delta=0.4$,查附表7得$n=68$. □

在实际问题中,有时只给出α,β及$|\mu_1-\mu_0|$的值,而需要确定所需的样本容量n. 这时由于σ未知,不能确定$\delta=|\mu_1-\mu_0|/\sigma$的值,因而不能直接查表以确定

样本容量.此时可采用下述近似方法.先适当取一值 n_1,抽取容量为 n_1 的样本,根据这一样本计算 s^2 的值,以 s^2 作为 σ^2 的估计,算出 δ 的近似值.由 α,β,δ 的值查附表 7 定出样本的容量,记为 n_2.若 $n_1 \geqslant n_2$,则取 n_1 作为所求的容量,即取 $n=n_1$.否则,再抽 n_2-n_1 个独立观察值与原来抽得的观察值合并,重新计算 δ 的近似值.然后用 δ 的新近似值和 α,β 查附表 7,再次定出样本容量.记为 n_3.若 $n_2 \geqslant n_3$,则取 $n=n_2$,否则再按上法重复进行.一般,只需试少数几次就可得到所求的样本容量 n.

下面考虑两个正态总体均值差的 t 检验.

若两个正态总体 $N(\mu_1,\sigma_1^2)$,$N(\mu_2,\sigma_2^2)$ 中 $\sigma_1^2=\sigma_2^2=\sigma^2$ 而 σ^2 未知.在均值差 $\mu_1-\mu_2$ 的检验问题 $H_0:\mu_1-\mu_2=0$,$H_1:\mu_1-\mu_2\neq 0$ (或 $H_0:\mu_1-\mu_2\leqslant 0$,$H_1:\mu_1-\mu_2>0$ 或 $H_0:\mu_1-\mu_2\geqslant 0$,$H_1:\mu_1-\mu_2<0$)的 t 检验法中,当分别自两个总体取得的相互独立的样本其容量 $n_1=n_2=n$ 时,给定 α,β 以及 $\delta=|\mu_1-\mu_2|/\sigma$ 的值后可以查附表 8 得到所需样本容量,使当 $|\mu_1-\mu_2|/\sigma\geqslant\delta$ 时犯第Ⅱ类错误的概率小于或等于 β.当仅给出 α,β 以及 $|\mu_1-\mu_2|$ 的值时,可按类似于上面所说的方法处理.

例 4 需要比较两种汽车用的燃料的辛烷值,得数据:

燃料 A	81	84	79	76	82	83	84	80	79	82	81	79
燃料 B	76	74	78	79	80	79	82	76	81	79	82	78

燃料的辛烷值越高,燃料质量越好.因燃料 B 较燃料 A 价格便宜,因此,若两者辛烷值相同时,则使用燃料 B;但若含量的均值差 $\mu_A-\mu_B\geqslant 5$,则使用燃料 A.设两总体的分布均可认为是正态的,而两个样本相互独立.问应采用哪种燃料(取 $\alpha=0.01,\beta=0.01$)?

解 按题意需要在显著性水平 $\alpha=0.01$ 下检验假设

$$H_0:\mu_A-\mu_B\leqslant 0,\quad H_1:\mu_A-\mu_B>0,$$

并要求在 $\mu_A-\mu_B\geqslant 5$ 时,犯第Ⅱ类错误的概率不超过 $\beta=0.01$.

所取的样本容量为 $n_A=n_B=12$,且有 $\bar{x}_A=80.83$,$\bar{x}_B=78.67$,$s_A^2=5.61$,$s_B^2=6.06$.经显著性水平为 0.1 的 F 检验知,可认为两总体的方差相等,即有 $\sigma_A^2=\sigma_B^2$,记为 σ^2.因 $n_1=n_2$,取 $\hat{\sigma}^2=(s_A^2+s_B^2)/2=5.835$ 作为 σ^2 的点估计,取 $\hat{\sigma}=\sqrt{\hat{\sigma}^2}$,于是 $\delta=5/\hat{\sigma}=2.07$,查表,当 $\alpha=0.01,\beta=0.01,\delta=2.07$ 时 $n\geqslant 8$.现 $n=12$,故已近似地满足要求.而右边检验的拒绝域为

$$t=\frac{\bar{x}_A-\bar{x}_B}{s_w\sqrt{1/n_A+1/n_B}}\geqslant t_{0.01}(n_1+n_2-2)=2.5083.$$

由样本观察值算得 $t=2.19<2.5083$,故接受 H_0,即采用燃料 B. □

§6 分布拟合检验

上面介绍的各种检验法都是在总体分布形式为已知的前提下进行讨论的.

但在实际问题中，有时不能知道总体服从什么类型的分布，这时就需要根据样本来检验关于分布的假设. 本节介绍 χ^2 拟合检验法. 它可以用来检验总体是否具有某一个指定的分布或属于某一个分布族，还介绍专用于检验分布是否为正态的"偏度、峰度检验法".

(一)单个分布的 χ^2 拟合检验法

设总体 X 的分布未知，$x_1,x_2,\cdots,x_n$ 是来自 X 的样本值. 我们来检验假设

$$H_0: 总体 X 的分布函数为 F(x). \tag{6.1}$$

$$H_1: 总体 X 的分布函数不是 F(x)①$$

其中设 $F(x)$ 不含未知参数.(也常以分布律或概率密度代替 $F(x)$).

下面来定义检验统计量. 将在 H_0 下 X 可能取值的全体 Ω 分成互不相交的子集 $A_1,A_2,\cdots,A_k$，以 $f_i\ (i=1,2,\cdots,k)$ 记样本观察值 $x_1,x_2,\cdots,x_n$ 中落在 A_i 的个数，这表示事件 $A_i=\{X 的值落在子集 A_i 内\}$ 在 n 次独立试验中发生 f_i 次，于是在这 n 次试验中事件 A_i 发生的频率为 f_i/n. 另一方面，当 H_0 为真时，我们可以根据 H_0 中所假设的 X 的分布函数来计算事件 A_i 的概率，得到 $p_i=P(A_i),i=1,2,\cdots,k$. 频率 f_i/n 与概率 p_i 会有差异，但一般来说，当 H_0 为真，且试验的次数又甚多时，这种差异不应太大，因此 $\left(\frac{f_i}{n}-p_i\right)^2$ 不应太大. 我们采用形如

$$\sum_{i=1}^{k}C_i\left(\frac{f_i}{n}-p_i\right)^2 \tag{6.2}$$

的统计量来度量样本与 H_0 中所假设的分布的吻合程度，其中 $C_i\ (i=1,2,\cdots,k)$②为给定的常数. 皮尔逊证明，如果选取 $C_i=n/p_i\ (i=1,2,\cdots,k)$，则由(6.2)定义的统计量具有下述定理中所述的简单性质. 于是我们就采用

$$\chi^2=\sum_{i=1}^{k}\frac{n}{p_i}\left(\frac{f_i}{n}-p_i\right)^2=\sum_{i=1}^{n}\frac{f_i^2}{np_i}-n \tag{6.3}$$

作为检验统计量.

定理 若 n 充分大($n\geqslant 50$)，则当 H_0 为真时统计量(6.3)近似服从 $\chi^2(k-1)$分布.(证略)

据以上的讨论，当 H_0 为真时，(6.3)式中的 χ^2 不应太大，如 χ^2 过分大就拒绝 H_0，拒绝域的形式为

$$\chi^2\geqslant G\quad (G 为正常数).$$

对于给定的显著性水平 α，确定 G 使

---

① 在这里备择假设 H_1 可以不必写出.

② 在每一项前乘以 C_i，是为了能够适当选择 C_i，使得统计量(6.2)有一个理想的极限分布.

$$P\{当\ H_0\ 为真时拒绝\ H_0\}=P_{H_0}\{\chi^2\geqslant G\}=\alpha.$$

由上述定理得 $G=\chi_\alpha^2(k-1)$. 即当样本观察值使(6.3)式中的 χ^2 的值有

$$\chi^2\geqslant\chi_\alpha^2(k-1),$$

则在显著性水平 α 下拒绝 H_0;否则就接受 H_0. 这就是单个分布的 χ^2 **拟合检验法**.

χ^2 拟合检验法是基于上述定理得到的,所以使用时必须注意 n 不能小于 50. 另外 np_i 不能太小,应有 $np_i\geqslant 5$,否则应适当合并 A_i,以满足这个要求(见下例).

例 1 下表列出了某一地区在夏季的一个月中由 100 个气象站报告的雷暴雨的次数.

i	0	1	2	3	4	5	$\geqslant 6$
f_i	22	37	20	13	6	2	0
A_i	A_0	A_1	A_2	A_3	A_4	A_5	A_6

其中 f_i 是报告雷暴雨次数为 i 的气象站数. 试用 χ^2 拟合检验法检验雷暴雨的次数 X 是否服从均值 $\lambda=1$ 的泊松分布(取显著性水平 $\alpha=0.05$).

解 按题意需检验假设

$$H_0:\quad P\{X=i\}=\frac{\lambda^i e^{-\lambda}}{i!}=\frac{e^{-1}}{i!},\quad i=0,1,\cdots.$$

在 H_0 下 X 所有可能取的值为 $\Omega=\{0,1,2,\cdots\}$,将 Ω 分成如表所示的两两不相交的子集 $A_0,A_1,\cdots,A_6$,则有 $P\{X=i\}$ 为

$$p_i=P\{X=i\}=\frac{e^{-1}}{i!},i=0,1,\cdots,5.$$

例如

$$p_0=P\{X=0\}=e^{-1}=0.367\ 88,$$

$$p_3=P\{X=3\}=\frac{e^{-1}}{3!}=0.061\ 31,$$

$$p_6=P\{X\geqslant 6\}=1-\sum_{i=0}^{5}p_i=0.059.$$

$$n=100.$$

表 8－2　例 1 的 χ^2 拟合检验计算表

A_i	f_i	p_i	np_i	$f_i^2/(np_i)$
$A_0:\{X=0\}$	22	e^{-1}	36.788	13.16
$A_1:\{X=1\}$	37	e^{-1}	36.788	37.21
$A_2:\{X=2\}$	20	$e^{-1}/2$	18.394	21.75
$A_3:\{X=3\}$	13	$e^{-1}/6$	6.131	
$A_4:\{X=4\}$	6	$e^{-1}/24$	1.533	
$A_5:\{X=5\}$	2	$e^{-1}/120$	0.307　}8.03	54.92
$A_6:\{X\geqslant 6\}$	0	$1-\sum_{i=0}^{5}p_i$	0.059	

$\sum=127.04$

计算结果如表 8－2 所示，其中有些 $np_i<5$ 的组予以适当合并，使得每组均有 $np_i\geqslant5$，如表中第 4 列花括号所示. 并组后 $k=4$，χ^2 的自由度为 $k-1=4-1=3$. $\chi^2_{0.05}(k-1)=\chi^2_{0.05}(3)=7.815$. 现在 $\chi^2=127.04-100=27.04>7.815$，故在显著性水平 0.05 下拒绝 H_0，认为样本不是来自均值 $\lambda=1$ 的泊松分布. □

例 2 在研究牛的毛色与牛角的有无，这样两对性状分离现象时，用黑色无角牛与红色有角牛杂交，子二代出现黑色无角牛 192 头，黑色有角牛 78 头，红色无角牛 72 头，红色有角牛 18 头，共 360 头，问这两对性状是否符合孟德尔遗传规律中 9∶3∶3∶1的遗传比例?

解 现将题中的数据列表如下

序号	1	2	3	4
种类	黑色无角	黑色有角	红色无角	红色有角
数量	192	78	72	18
A_i	A_1	A_2	A_3	A_4

以 X 记各种牛的序号，按题意需检验各类牛的头数符合比例 9∶3∶3∶1，即 (9/16)∶(3/16)∶(3/16)∶(1/16). 需检验假设：H_0：X 的分布律为

X	1	2	3	4
p_k	9/16	3/16	3/16	1/16

取显著性水平为 0.1. 所需计算列在表 8－3 中($n=360$).

表 8－3 例 2 的 χ^2 检验计算表

A_i	f_i	p_i	np_i	$f_i^2/(np_i)$
A_1	192	9/16	$360\times9/16=202.5$	$192^2/202.5=182.04$
A_2	78	3/16	$360\times3/16=67.5$	$78^2/67.5=90.13$
A_3	72	3/16	$360\times3/16=67.5$	$72^2/67.5=76.8$
A_4	18	1/16	$360\times1/16=22.5$	$18^2/22.5=14.4$
				$\sum=363.37$

现在 $\chi^2=363.37-360=3.37$，$k=4$，$\chi^2_{0.1}(4-1)=6.251>3.37$，故接受 H_0，认为两性状符合孟德尔遗传规律中 9∶3∶3∶1的遗传比例. □

(二) 分布族的 χ^2 拟合检验

在(一)中要检验的原假设是 H_0：总体 X 的分布函数是 $F(x)$，其中 $F(x)$是已知的，这种情况是不多的. 我们经常遇到的所需检验的原假设是

$$H_0\text{：总体 } X \text{ 的分布函数是 } F(x;\theta_1,\cdots,\theta_r), \tag{6.4}$$

其中 F 的形式已知，而 $\boldsymbol{\theta}=(\theta_1,\theta_2,\cdots,\theta_r)$是未知参数，它们在某一个范围取值. 在 $F(x;\theta_1,\theta_2,\cdots,\theta_r)$中当参数 $\theta_1,\theta_2,\cdots,\theta_r$ 取不同的值时，就得到不同的分布，因而 $F(x;\theta_1,\theta_2,\cdots,\theta_r)$代表一族分布. (6.4)中的 H_0 表示总体 X 的分布属于分

布族 $F(x;\theta_1,\theta_2,\cdots,\theta_r)$. 采用类似(一)中的方法来定义检验统计量，将在 H_0 下 X 可能取值的全体 Ω 分成 k $(k>r+1)$ 个互不相交的子集 $A_1,A_2,\cdots,A_k$，以 f_i $(i=1,2,\cdots,k)$ 记样本观察值 $x_1,x_2,\cdots,x_n$ 落在 A_i 的个数，则事件 $A_i=\{X$ 的值落在 A_i 内$\}$ 的频率为 f_i/n. 另一方面，当 H_0 为真时，由 H_0 所假设的分布函数来计算 $P(A_i)$，得到 $P(A_i)=p_i(\theta_1,\theta_2,\cdots,\theta_r)=p_i(\boldsymbol{\theta})=p_i$. 此时，需先利用样本求出未知参数的最大似然估计(在 H_0 下)，以估计值作为参数值，求出 p_i 的估计值 $\hat{p}_i=\hat{P}(A_i)$，在(6.3)式中以 $\hat{p}_i$ 代替 p_i，取

$$\chi^2=\sum_{i=1}^{k}\frac{f_i^2}{n\hat{p}_i}-n \tag{6.5}$$

作为检验假设 H_0 的统计量. 可以证明，在某些条件下，在 H_0 为真时近似地有

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

与在(一)中一样可得假设检验问题(6.4)的拒绝域为

$$\chi^2\geqslant\chi_\alpha^2(k-r-1), \tag{6.6}$$

α 为显著性水平. 以上就是用来检验分布族的 χ^2 拟合检验法.

例 3　在一实验中，每隔一定时间观察一次由某种铀所放射的到达计数器上的 α 粒子数 X，共观察了 100 次，得结果如下表：

表 8—4　铀放射的到达计数器上的 α 粒子数的实验记录

i	0	1	2	3	4	5	6	7	8	9	10	11	$\geqslant 12$
f_i	1	5	16	17	26	11	9	9	2	1	2	1	0
A_i	A_0	A_1	A_2	A_3	A_4	A_5	A_6	A_7	A_8	A_9	A_{10}	A_{11}	A_{12}

其中 f_i 是观察到有 i 个 α 粒子的次数. 从理论上考虑知 X 应服从泊松分布

$$P\{X=i\}=\frac{\lambda^i e^{-\lambda}}{i!},\quad i=0,1,2,\cdots. \tag{6.7}$$

问(6.7)式是否符合实际(取 $\alpha=0.05$)？即在显著性水平 0.05 下检验假设

H_0：总体 X 服从泊松分布

$$P\{X=i\}=\frac{\lambda^i e^{-\lambda}}{i!},\quad i=0,1,2,\cdots.$$

解　因在 H_0 中参数 λ 未具体给出，所以先估计 λ. 由最大似然估计法得 $\hat{\lambda}=\overline{x}=4.2$. 在 H_0 假设下，即在 X 服从泊松分布的假设下，X 所有可能取的值为 $\Omega=\{0,1,2,\cdots\}$，将 Ω 分成如表 8—4 所示的两两不相交的子集 $A_0,A_1,\cdots,A_{12}$. 则 $P\{X=i\}$ 有估计

$$\hat{p}_i=\hat{P}\{X=i\}=\frac{4.2^i e^{-4.2}}{i!},\quad i=0,1,\cdots.$$

例如　　$\hat{p}_0=\hat{P}\{X=0\}=e^{-4.2}=0.015,$

$$\hat{p}_3=\hat{P}\{X=3\}=\frac{4.2^3 e^{-4.2}}{3!}=0.185,$$

$$\hat{p}_{12}=\hat{P}\{X\geqslant 12\}=1-\sum_{i=0}^{11}\hat{p}_i=0.002.$$

表 8—5　例 3 的 χ^2 拟合检验计算表

A_i	f_i	$\hat{p}_i$	$n\hat{p}_i$	$f_i^2/n\hat{p}_i$
A_0	1 } 6	0.015 } 0.078	1.5 } 7.8	4.615
A_1	5	0.063	6.3	
A_2	16	0.132	13.2	19.394
A_3	17	0.185	18.5	15.622
A_4	26	0.194	19.4	34.845
A_5	11	0.163	16.3	7.423
A_6	9	0.114	11.4	7.105
A_7	9	0.069	6.9	11.739
A_8	2	0.036	3.6	
A_9	1	0.017	1.7	
A_{10}	2 } 6	0.007 } 0.065	0.7 } 6.5	5.538
A_{11}	1	0.003	0.3	
A_{12}	0	0.002	0.2	

$\sum=106.281$

计算结果如表 8—5 所示，其中有些 $n\hat{p}_i<5$ 的组予以适当合并，使得每组均有 $n\hat{p}_i\geqslant 5$，如表中第四列花括号所示. 此处，并组后 $k=8$，但因在计算概率时，估计了一个参数 λ，故 $r=1$，χ^2 的自由度为 $8-1-1=6$. $\chi^2_{0.05}(k-r-1)=\chi^2_{0.05}(6)=12.592$，现在 $\chi^2=106.281-100=6.281<12.592$，故在显著性水平 0.05 下接受 H_0. 即认为样本来自泊松分布总体. 也就是说认为理论上的结论是符合实际的. □

注意　本题答案是“接受 H_0，认为总体 X 的分布属于泊松分布族，即认为 $X\sim\pi(\lambda)$”，亦即“认为必有某一个参数 λ_0，$X\sim\pi(\lambda_0)$”，而不能将答案误写成“X 服从以 $\lambda=4.2$ 为参数的泊松分布”.

例 4　自 1965 年 1 月 1 日至 1971 年 2 月 9 日共 2 231 天中，全世界记录到里氏震级 4 级和 4 级以上地震计 162 次，统计如下：

相继两次地震间隔天数 x	0～4	5～9	10～14	15～19	20～24	25～29	30～34	35～39	≥40
出现的频数	50	31	26	17	10	8	6	6	8①

① 这里 8 个数值是 40,43,44,49,58,60,81,109.

试检验相继两次地震间隔的天数 X 服从指数分布($\alpha=0.05$).

解 按题意需检验假设

H_0:X 的概率密度为

$$f(x)=\begin{cases}\dfrac{1}{\theta}e^{-x/\theta}, & x>0,\\ 0, & x\leqslant 0.\end{cases} \tag{6.8}$$

在这里,H_0 中的参数 θ 未给出,先由最大似然估计法求得 θ 的估计为 $\hat{\theta}=\overline{x}=\dfrac{2231}{162}=13.77$. 在 H_0 下,X 可能取值的全体 Ω 为区间 $[0,\infty)$. 将区间 $[0,\infty)$ 分为 $k=9$ 个互不重叠的小区间:$A_1=[0,4.5]$,$A_2=(4.5,9.5]$,$\cdots$,$A_9=(39.5,\infty)$,如表 8-6 第二列所示. 若 H_0 为真,X 的分布函数的估计为

$$\hat{F}(x)=\begin{cases}1-e^{-x/13.77}, & x>0,\\ 0, & x\leqslant 0.\end{cases}$$

由上式可得概率 $p_i=P(A_i)$ 的估计:

$$\hat{p}_i=\hat{P}(A_i)=\hat{P}\{a_i<X\leqslant a_{i+1}\}=\hat{F}(a_{i+1})-\hat{F}(a_i).$$

例如

$$\hat{p}_2=\hat{P}(A_2)=\hat{P}\{4.5<X\leqslant 9.5\}=\hat{F}(9.5)-\hat{F}(4.5)=0.2196,$$

而

$$\hat{p}_9=\hat{P}(A_9)=1-\sum_{i=1}^{8}\hat{F}(A_i)=0.0568.$$

将计算结果列表如下.

表 8-6 例 4 的 χ^2 检验计算表

A_i	f_i	$\hat{p}_i$	$n\hat{p}_i$		$f_i^2/(n\hat{p}_i)$
$A_1:0\leqslant x\leqslant 4.5$	50	0.2788	45.1656		55.3519
$A_2:4.5<x\leqslant 9.5$	31	0.2196	35.5752		27.0132
$A_3:9.5<x\leqslant 14.5$	26	0.1527	24.7374		27.3270
$A_4:14.5<x\leqslant 19.5$	17	0.1062	17.2044		16.7980
$A_5:19.5<x\leqslant 24.5$	10	0.0739	11.9718		8.3530
$A_6:24.5<x\leqslant 29.5$	8	0.0514	8.3268		7.6860
$A_7:29.5<x\leqslant 34.5$	6	0.0358	5.7996		6.2073
$A_8:34.5<x\leqslant 39.5$	6	0.0248	4.0176	13.2192	14.8269
$A_9:39.5<x<\infty$	8	0.0568	9.2016		

$\sum=163.5633$

现在 $\chi^2=163.5633-162=1.5633$,因为 $\chi^2_{0.05}(k-r-1)=\chi^2_{0.05}(8-1-1)=\chi^2_{0.05}(6)=12.592>1.5633$,故在显著性水平 0.05 下接受 H_0,认为 X 服从指

数分布. □

例 5 对于第六章§2 例 1 中的数据,试检验它们是否来自正态总体 X(取显著性水平 $\alpha=0.1$).

解 需检验假设

H_0:X 的概率密度为

$$f(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}},\quad -\infty<x<\infty.$$

因在 H_0 中未给出 μ,σ^2 的数值.需先估计 μ,σ^2.由最大似然估计法得 μ,σ^2 的估计值为 $\hat{\mu}=143.8,\hat{\mu}^2=(6.0)^2$.我们将在 H_0 下 X 可能取值的区间$(-\infty,\infty)$分为 7 个小区间,并取事件 A_i 如表 8—7 中第一列所示.若 H_0 为真 X 的概率密度的估计为

$$\hat{f}(x)=\frac{1}{\sqrt{2\pi}\times 6}e^{-\frac{(x-143.8)^2}{2\times 6^2}},-\infty<x<\infty.$$

按上式并查标准正态分布的分布函数表即可得概率 $P(A_i)$的估计.例如

$$\begin{aligned}\hat{p}_2=\hat{P}(A_2)&=\hat{P}\{129.5<X\leqslant 134.5\}\\&=\Phi\left(\frac{134.5-143.8}{6}\right)-\Phi\left(\frac{129.5-143.8}{6}\right)\\&=\Phi(-1.55)-\Phi(-2.38)=0.0519.\end{aligned}$$

将计算结果列表如下:

表 8—7 例 5 的 χ^2 检验计算表

A_i	f_i	$\hat{p}_i$	$n\hat{p}_i$	$f_i^2/n\hat{p}_i$
$A_1:x\leqslant 129.5$	1 } 5	0.0087 }	0.73 } 5.09	4.91
$A_2:129.5<x\leqslant 134.5$	4	0.0519	4.36	
$A_3:134.5<x\leqslant 139.5$	10	0.1752	14.72	6.79
$A_4:139.5<x\leqslant 144.5$	33	0.3120	26.21	41.55
$A_5:144.5<x\leqslant 149.5$	24	0.2811	23.61	24.40
$A_6:149.5<x\leqslant 154.5$	9 } 12	0.1336 }	11.22 } 14.37	10.02
$A_7:154.5<x<\infty$	3	0.0375	3.15	

$\sum=87.67$

现在 $\chi^2=87.67-84=3.67$,因为 $\chi^2_{0.1}(k-r-1)=\chi^2_{0.1}(5-2-1)=\chi^2_{0.1}(2)=4.605>3.67$,故在水平 0.1 下接受 H_0,即认为数据来自正态分布总体. □

(三)偏度、峰度检验

根据第五章关于中心极限定理的论述知道,正态分布随机变量是较广泛地

存在的，因此，当研究一连续型总体时，人们往往先考察它是否服从正态分布．上面介绍的 χ^2 拟合检验法虽然是检验总体分布的较一般的方法，但用它来检验总体的正态性时，犯第Ⅱ类错误的概率往往较大．为此，统计学家们对检验正态总体的种种方法进行了比较，根据奥野忠一等人在 20 世纪 70 年代进行的大量模拟计算的结果，认为正态性检验方法中，总的来说，以“偏度、峰度检验法”及“夏皮罗－威尔克法”较为有效．在这里我们仅介绍偏度、峰度检验法．

随机变量 X 的偏度和峰度指的是 X 的标准化变量 $[X-E(X)]/\sqrt{D(X)}$ 的三阶矩和四阶矩：

$$\nu_1=E\left[\left(\frac{X-E(X)}{\sqrt{D(X)}}\right)^3\right]=\frac{E[(X-E(X))^3]}{(D(X))^{3/2}},$$

$$\nu_2=E\left[\left(\frac{X-E(X)}{\sqrt{D(X)}}\right)^4\right]=\frac{E[(X-E(X))^4]}{(D(X))^2}.$$

当随机变量 X 服从正态分布时，$\nu_1=0$ 且 $\nu_2=3$．

设 $X_1,X_2,\cdots,X_n$ 是来自总体 X 的样本，则 ν_1,ν_2 的矩估计量分别是

$$G_1=B_3/B_2^{3/2},\quad G_2=B_4/B_2^2,$$

其中 $B_k\ (k=2,3,4)$ 是样本 k 阶中心矩，并分别称 G_1,G_2 为样本偏度和样本峰度．

若总体 X 为正态变量，则可证当 n 充分大时，近似地有

$$G_1\sim N\left(0,\frac{6(n-2)}{(n+1)(n+3)}\right), \tag{6.9}$$

$$G_2\sim N\left(3-\frac{6}{n+1},\frac{24n(n-2)(n-3)}{(n+1)^2(n+3)(n+5)}\right). \tag{6.10}$$

设 $X_1,X_2,\cdots,X_n$ 是来自总体 X 的样本，现在来检验假设

$$H_0: X \text{ 为正态总体}.$$

记

$$\sigma_1=\sqrt{\frac{6(n-2)}{(n+1)(n+3)}},\quad \sigma_2=\sqrt{\frac{24n(n-2)(n-3)}{(n+1)^2(n+3)(n+5)}},$$

$\mu_2=3-\dfrac{6}{n+1}$，$U_1=G_1/\sigma_1$，$U_2=(G_2-\mu_2)/\sigma_2$．当 H_0 为真且 n 充分大时，近似地有

$$U_1\sim N(0,1),\quad U_2\sim N(0,1).$$

由第六章§3知样本偏度 G_1、样本峰度 G_2 分别依概率收敛于总体偏度 ν_1 和总体峰度 ν_2．因此当 H_0 为真且 n 充分大时，一般来说，G_1 与 $\nu_1=0$ 的偏离不应太大，而 G_2 与 $\nu_2=3$ 的偏离不应太大．故从直观来看当 $|U_1|$ 的观察值 $|u_1|$ 或 $|U_2|$ 的观察值 $|u_2|$ 过大时就拒绝 H_0．取显著性水平为 α，H_0 的拒绝域为

$$|u_1|\geqslant k_1 \quad 或 \quad |u_2|\geqslant k_2, \tag{6.11}$$

其中 k_1,k_2 由以下两式确定：

$$P_{H_0}\{|U_1|\geqslant k_1\}=\frac{\alpha}{2}, \quad P_{H_0}\{|U_2|\geqslant k_2\}=\frac{\alpha}{2}.$$

这里记号 $P_{H_0}\{\cdot\}$ 表示当 H_0 为真时事件 $\{\cdot\}$ 的概率，即有 $k_1=z_{\alpha/4}, k_2=z_{\alpha/4}$. 于是得拒绝域为

$$|u_1|\geqslant z_{\alpha/4} \quad 或 \quad |u_2|\geqslant z_{\alpha/4}. \tag{6.12}$$

下面来验证当 n 充分大时上述检验法近似地满足显著性水平为 α 的要求. 事实上当 n 充分大时有

$$\begin{aligned}&P\{当 H_0 为真拒绝 H_0\}\\&=P_{H_0}\{(|U_1|\geqslant z_{\alpha/4})\cup(|U_2|\geqslant z_{\alpha/4})\}\\&\leqslant P_{H_0}\{|U_1|\geqslant z_{\alpha/4}\}+P_{H_0}\{|U_2|\geqslant z_{\alpha/4}\}=\frac{\alpha}{2}+\frac{\alpha}{2}=\alpha.\end{aligned}$$

例 6 试用偏度、峰度检验法检验例 5 中的数据是否来自正态总体(取 $\alpha=0.1$).

解 现在来检验假设

$$H_0: 数据来自正态总体.$$

这里 $\alpha=0.1, n=84, \sigma_1=\sqrt{\dfrac{6(n-2)}{(n+1)(n+3)}}=0.257\,9$,

$$\sigma_2=\sqrt{\frac{24n(n-2)(n-3)}{(n+1)^2(n+3)(n+5)}}=0.489\,2,$$

$\mu_2=3-\dfrac{6}{n+1}=2.929\,4$. 下面来计算样本中心矩 B_2,B_3,B_4，计算时可利用以下关系式：

$$B_2=A_2-A_1^2, B_3=A_3-3A_2A_1+2A_1^3,$$
$$B_4=A_4-4A_3A_1+6A_2A_1^2-3A_1^4,$$

其中 $A_k=\dfrac{1}{n}\sum\limits_{i=1}^{n}X_i^k \quad (k=1,2,3,4)$ 为 k 阶样本矩. 经计算得 $A_k \quad (k=1,2,3,4)$, $B_k \quad (k=2,3,4)$ 的观察值分别为

$$A_1=143.773\,8, \quad A_2=20\,706.13, \quad A_3=2\,987\,099,$$
$$A_4=4.316\,426\times10^8, \quad B_2=35.224\,6, \quad B_3=-28.5, \quad B_4=3\,840.$$

样本偏度和样本峰度的观察值分别为

$$g_1=-0.136\,3, \quad g_2=3.094\,8.$$

而 $z_{\alpha/4}=z_{0.025}=1.96$. 由(6.11)式，拒绝域为

$$|u_1|=|g_1/\sigma_1|\geqslant 1.96 \quad 或 \quad |u_2|=|g_2-\mu_2|/\sigma_2\geqslant 1.96.$$

现算得 $|u_1|=0.528\,5<1.96$, $|u_2|=0.338\,1<1.96$，故接受 H_0，认为数据来自

正态分布的总体. □

上述检验法称为**偏度、峰度检验法**. 使用这一检验法时样本容量以大于 100 为宜.

§7 秩和检验

本节介绍一种有效的、且使用方便的检验方法——秩和检验法.

设有两个连续型总体,它们的概率密度函数分别为 $f_1(x)$,$f_2(x)$,均为未知,但已知

$$f_1(x)=f_2(x-a),\quad a \text{ 为未知常数}, \tag{7.1}$$

即 f_1 与 f_2 至多只差一平移. 我们要检验下述各项假设

$$H_0: a=0,\ H_1: a<0. \tag{7.2}$$

$$H_0: a=0,\ H_1: a>0. \tag{7.3}$$

$$H_0: a=0,\ H_1: a\neq 0. \tag{7.4}$$

特别,若两总体的均值存在,分别记作 μ_1,μ_2,则由于 f_1,f_2 至多只差一平移,故有

$$\mu_2=\mu_1-a.$$

此时,上述各项假设分别等价于

$$H_0: \mu_1=\mu_2,\ H_1: \mu_1<\mu_2. \tag{7.2$'$}$$

$$H_0: \mu_1=\mu_2,\ H_1: \mu_1>\mu_2. \tag{7.3$'$}$$

$$H_0: \mu_1=\mu_2,\ H_1: \mu_1\neq\mu_2. \tag{7.4$'$}$$

现在来介绍威尔柯克斯(Frank Wilcoxon)提出的秩和检验法以检验上述假设. 为此,先引入秩的概念.

秩 设 X 为一总体,将一容量为 n 的样本观察值按自小到大的次序编号排列成

$$x_{(1)}<x_{(2)}<\cdots<x_{(n)}, \tag{7.5}$$

称 $x_{(i)}$ 的足标 i 为 $x_{(i)}$ 的秩,$i=1,2,\cdots,n$.

现设自 1,2 两总体分别抽取容量为 n_1,n_2 的样本,且设两样本独立. 这里总假定 $n_1\leqslant n_2$. 我们将这 n_1+n_2 个观察值放在一起,按自小到大的次序排列,求出每个观察值的秩,然后将属于第 1 个总体的样本观察值的秩相加,其和记为 R_1,称为第 1 样本的秩和. 其余观察值的秩的总和记作 R_2,称为第 2 样本的秩和. 显然 R_1,R_2 是离散型的随机变量,且有

$$R_1+R_2=\frac{1}{2}(n_1+n_2)(n_1+n_2+1). \tag{7.6}$$

所以,R_1,R_2 中的一个确定后另一个随之而定. 这样,我们只要考虑统计量 R_1

即可.

现在来解决双边检验问题(7.4). 对此,先作直观分析. 当 H_0 为真时,即有 $f_1(x)=f_2(x)$,这时两个独立样本实际上是来自同一个总体. 因而第 1 个样本中诸元素的秩应该随机地、分散地在自然数 $1\sim n_1+n_2$ 中取值,一般来说不应过分集中取较小的或较大的值. 考虑到

$$\frac{1}{2}n_1(n_1+1)\leqslant R_1\leqslant\frac{1}{2}n_1(n_1+2n_2+1),$$

即知当 H_0 为真时秩和 R_1 一般来说不应取太靠近上述不等式两端的值. 因而,当 R_1 的观察值 r_1 过分大或过分小时,我们都拒绝 H_0.

据以上分析,对于双边检验(7.4),在给定显著性水平 α 下,H_0 的拒绝域为

$$r_1\leqslant C_U\left(\frac{\alpha}{2}\right) \quad \text{或} \quad r_1\geqslant C_L\left(\frac{\alpha}{2}\right),$$

其中临界点 $C_U\left(\frac{\alpha}{2}\right)$是满足 $P_{a=0}\left\{R_1\leqslant C_U\left(\frac{\alpha}{2}\right)\right\}\leqslant\frac{\alpha}{2}$的最大整数,而 $C_L\left(\frac{\alpha}{2}\right)$是满足 $P_{a=0}\left\{R_1\geqslant C_L\left(\frac{\alpha}{2}\right)\right\}\leqslant\frac{\alpha}{2}$的最小整数. 而犯第 I 类错误的概率为

$$P_{a=0}\left\{R_1\leqslant C_U\left(\frac{\alpha}{2}\right)\right\}+P_{a=0}\left\{R_1\geqslant C_L\left(\frac{\alpha}{2}\right)\right\}\leqslant\frac{\alpha}{2}+\frac{\alpha}{2}=\alpha.$$

如果知道 R_1 的分布,则临界点 $C_U\left(\frac{\alpha}{2}\right)$,$C_L\left(\frac{\alpha}{2}\right)$是不难求得的. 下面以 $n_1=3,n_2=4$ 的情况为例来说明求临界点的方法.

当 $n_1=3,n_2=4$ 时,第 1 个样本中各观察值的秩的不同取法共有$\binom{3+4}{3}=35$ 种. 现将这 35 种情况列表如下.

秩	R_1	秩	R_1	秩	R_1	秩	R_1	秩	R_1
123	6	136	10	167	14	247	13	356	14
124	7	137	11	234	9	256	13	357	15
125	8	145	10	235	10	257	14	367	16
126	9	146	11	236	11	267	15	456	15
127	10	147	12	237	12	345	12	457	16
134	8	156	12	245	11	346	13	467	17
135	9	157	13	246	12	347	14	567	18

由于这 35 种情况的出现是等可能的,由上表容易求得 R_1 的分布律和分布函数如下:

R_1	6	7	8	9	10	11	12
$P\{R_1=r_1\}$	1/35	1/35	2/35	3/35	4/35	4/35	5/35
$P\{R_1\leqslant r_1\}$	1/35	2/35	4/35	7/35	11/35	15/35	20/35
R_1	13	14	15	16	17	18	
$P\{R_1=r_1\}$	4/35	4/35	3/35	2/35	1/35	1/35	
$P\{R_1\leqslant r_1\}$	24/35	28/35	31/35	33/35	34/35	1	

于是，对于不同的 α 值，容易写出检验问题(7.4)的临界点和拒绝域. 例如，给定 $\alpha=0.2$. 由上表知

$$P_{a=0}\{R_1\leqslant 7\}=2/35<0.1=\frac{\alpha}{2},$$

$$P_{a=0}\{R_1\geqslant 17\}=2/35<0.1=\frac{\alpha}{2}.$$

即有 $C_U(0.1)=7, C_L(0.1)=17$. 故当 $n_1=3, n_2=4$，在显著性水平 0.2 下检验问题(7.4)的拒绝域为

$$r_1\leqslant 7 \quad 或 \quad r_1\geqslant 17.$$

此时，犯第Ⅰ类错误的概率为

$$P_{a=0}\{R_1\leqslant 7\}+P_{a=0}\{R_1\geqslant 17\}=2/35+2/35=0.114.$$

类似地可得左边检验(7.2)的拒绝域为(显著性水平为 α)

$$r_1\leqslant C_U(\alpha),$$

此处，临界点 $C_U(\alpha)$是满足 $P_{a=0}\{R_1\leqslant C_U(\alpha)\}\leqslant\alpha$ 的最大整数. 而右边检验问题的拒绝域为(显著性水平为 α)

$$r_1\geqslant C_L(\alpha),$$

此处，临界点 $C_L(\alpha)$是满足 $P_{a=0}\{R_1\geqslant C_L(\alpha)\}\leqslant\alpha$ 的最小整数.

例如，若给定 $\alpha=0.1$，抽取的样本容量为 $n_1=3, n_2=4$，则由上表知检验问题(7.3)的拒绝域为

$$r_1\geqslant 17.$$

此时犯第Ⅰ类错误的概率为 $2/35<0.1$.

书末附表 9 中列出了 n_1 和 n_2 自 2 到 10 为止的 n_1, n_2 的各种组合的临界点，以及相应的犯第Ⅰ类错误的概率.

例 1　为查明某种血清是否会抑制白血病，选取患白血病已到晚期的老鼠 9 只，其中有 5 只接受这种治疗，另 4 只则不做这种治疗. 设两样本相互独立. 从试验开始时计算，其存活时间(以月计)如下：

不作治疗	1.9	0.5	0.9	2.1	
接受治疗	3.1	5.3	1.4	4.6	2.8

设治疗与否的存活时间的概率密度至多只差一个平移. 取 $\alpha=0.05$,问这种血清对白血病是否有抑制作用?

解 本题需检验接受治疗的老鼠的存活期是否有增长. 分别以 μ_1,μ_2 表示不作治疗和接受治疗的老鼠的存活时间总体的均值. 需要检验的假设是

$$H_0:\mu_1=\mu_2,\quad H_1:\mu_1<\mu_2.$$

这里,$n_1=4,n_2=5,\alpha=0.05$. 先计算对应于 $n_1=4$ 的一组观察值的秩和. 将两组数据放在一起按自小到大的次序排列. 对来自第 1 个总体($n_1=4$)的数据下面加一表示之. 即有

数据	$\underline{0.5}$	$\underline{0.9}$	1.4	$\underline{1.9}$	$\underline{2.1}$	2.8	3.1	4.6	5.3
秩	1	2	3	4	5	6	7	8	9

所以 R_1 的观察值为 $r_1=1+2+4+5=12$. 查附表 9 知$C_U(0.05)=12$,即拒绝域为 $r_1\leqslant 12$. 而现在 $r_1=12$,故拒绝 H_0,即认为这种血清对白血病有抑制作用. □

可以证明,当 H_0 为真时(即 $a=0$ 时)

$$\begin{aligned}\mu_{R_1}&=E(R_1)=\frac{n_1(n_1+n_2+1)}{2},\\ \sigma_{R_1}^2&=D(R_1)=\frac{n_1n_2(n_1+n_2+1)}{12}.\end{aligned}\tag{7.7}$$

而当 $n_1,n_2\geqslant 10$,当 H_0 为真时,近似地有

$$R_1\sim N(\mu_{R_1},\sigma_{R_1}^2).\tag{7.8}$$

因此,当 $n_1,n_2\geqslant 10$ 时我们可以采用

$$Z=\frac{R_1-\mu_{R_1}}{\sigma_{R_1}}$$

作为检验统计量. 在显著性水平 α 下双边检验、右边检验、左边检验的近似拒绝域分别为

$$|z|\geqslant z_{\alpha/2},\quad z\geqslant z_\alpha,\quad z\leqslant -z_\alpha.$$

这里 z 是 Z 的观察值 .

例 2 某商店为了确定向公司 A 或公司 B 购买某种商品,将 A,B 公司以往各次进货的次品率进行比较,数据如下,设两样本独立. 问两公司的商品的质量有无显著差异. 设两公司的商品的次品率的密度至多只差一个平移,取显著性水平 $\alpha=0.05$.

A	7.0	3.5	9.6	8.1	6.2	5.1	10.4	4.0	2.0	10.5			
B	5.7	3.2	4.2	11.0	9.7	6.9	3.6	4.8	5.6	8.4	10.1	5.5	12.3

解 分别以 μ_A,μ_B 记公司 A,B 的商品次品率总体的均值. 所需检验的假设是

$$H_0:\mu_A=\mu_B,\quad H_1:\mu_A\neq\mu_B.$$

先将数据按自小到大的次序排列,得到对应于 $n_1=10$ 的样本的秩和为

$$r_1=1+3+5+8+12+14+15+17+20+21=116.$$

又当 H_0 为真时,

$$E(R_1)=\frac{1}{2}n_1(n_1+n_2+1)=\frac{1}{2}\times 10(10+13+1)=120,$$

$$D(R_1)=\frac{1}{12}n_1n_2(n_1+n_2+1)=260.$$

故知当 H_0 为真时近似地有

$$R_1\sim N(120,260).$$

拒绝域为

$$\frac{|r_1-120|}{\sqrt{260}}\geqslant z_{0.025}=1.96.$$

现在 R_1 的观察值为 $r_1=116$,得 $|r_1-120|/\sqrt{260}=0.25<1.96$,故接受 H_0,认为两个公司商品的质量无显著差异. □

在实际问题中(7.5)式中会出现某些观察值相等的情况,对于这种观察值的秩定义为足标的平均值. 例如,若抽得的样本按次序排成 0,1,1,1,2,3,3,则三个 1 的秩均为(2+3+4)/3=3,两个 3 的秩均为(6+7)/2=6.5.

将两个样本 $n_1+n_2=n$ 个元素按自小到大的次序排列,若出现 k 个秩相同的组,设其中有 t_i 个数的秩为 $a_i,i=1,2,\cdots,k,a_1<\cdots<a_k$,则当 H_0 为真时 R_1 的均值仍为 $\mu_{R_1}=n_1(n_1+n_2+1)/2$,而 R_1 的方差修正为

$$\sigma_{R_1}^2=\frac{n_1n_2\left[n(n^2-1)-\sum_{i=1}^{k}t_i(t_i^2-1)\right]}{12n(n-1)}.\tag{7.9}$$

当 k 不大时,附表 9 仍能使用,但表载值为近似值. 又当 $n_1,n_2\geqslant 10$,H_0 为真,且 k 不大时,近似地有

$$R_1\sim N(\mu_{R_1},\sigma_{R_1}^2),$$

其中 $\mu_{R_1}=n_1(n_1+n_2+1)/2$,而 $\sigma_{R_1}^2$ 由(7.9)式确定. 这时我们就采用

$$Z=\frac{R_1-\mu_{R_1}}{\sigma_{R_1}}$$

作为检验统计量来检验假设检验问题(7.2)′～(7.4)′.

例 3 两位化验员各自读得某种液体黏度如下：

化验员 A	82	73	91	84	77	98	81	79	87	85	
化验员 B	80	76	92	86	74	96	83	79	80	75	79

设数据可以认为分别来自仅均值可能有差异的两个总体的样本. 试在 $\alpha=0.05$ 下，检验假设

$$H_0:\mu_1=\mu_2,\quad H_1:\mu_1>\mu_2,$$

其中 μ_1,μ_2 分别为两总体的均值.

解 将两个样本的元素混合，按自小到大次序排列. 并求出各个元素的秩如下：

数据	$\underline{73}$	74	75	76	$\underline{77}$	$\underline{79}$	79	79	80	80	$\underline{81}$
秩	1	2	3	4	5	7	7	7	9.5	9.5	11

数据	$\underline{82}$	83	$\underline{84}$	$\underline{85}$	86	$\underline{87}$	$\underline{91}$	92	96	$\underline{98}$
秩	12	13	14	15	16	17	18	19	20	21

现在 $n_1=10,n_2=11,n=21,\mu_{R_1}=10\times22/2=110,k=2,\ \sum_{i=1}^{2}t_i(t_i^2-1)=3\times(9-1)+2\times(4-1)=30$，按(7.9)式得 $\sigma_{R_1}^2=201$. 当 H_0 为真时近似地有

$$R_1\sim N(110,201).$$

拒绝域为

$$\frac{r_1-110}{\sqrt{201}}\geqslant z_{0.05}=1.645.$$

现在 R_1 的观察值为 $r_1=121$，得 $(r_1-110)/\sqrt{201}=0.776<1.645$. 故接受 H_0，认为两位化验员所测得的数据无显著差异. □

§8 假设检验问题的 p 值检验法

以上讨论的假设检验方法称为**临界值法**. 本节介绍另一种被称为 p 值检验法的检验方法. 先从一个例题讲起.

例 1 设总体 $X\sim N(\mu,\sigma^2)$，μ 未知，$\sigma^2=100$，现有样本 $x_1,x_2,\cdots,x_{52}$，算得 $\overline{x}=62.75$. 现在来检验假设

$$H_0:\mu=\mu_0=60,\qquad H_1:\mu>60.$$

采用 Z 检验法，检验统计量为

$$Z=\frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}.$$

以数据代入，得 Z 的观察值为

$$z_0=\frac{62.75-60}{10/\sqrt{52}}=1.983.$$

概率

$$P\{Z\geqslant z_0\}=P\{Z\geqslant 1.983\}=1-\Phi(1.983)=0.0238.$$

此即为图 8－7 中标准正态曲线下位于 z_0 右边的尾部面积.

此概率称为 Z 检验法的右边检验的 p 值. 记为

$$P\{Z\geqslant z_0\}=p\text{ 值}(=0.0237).$$

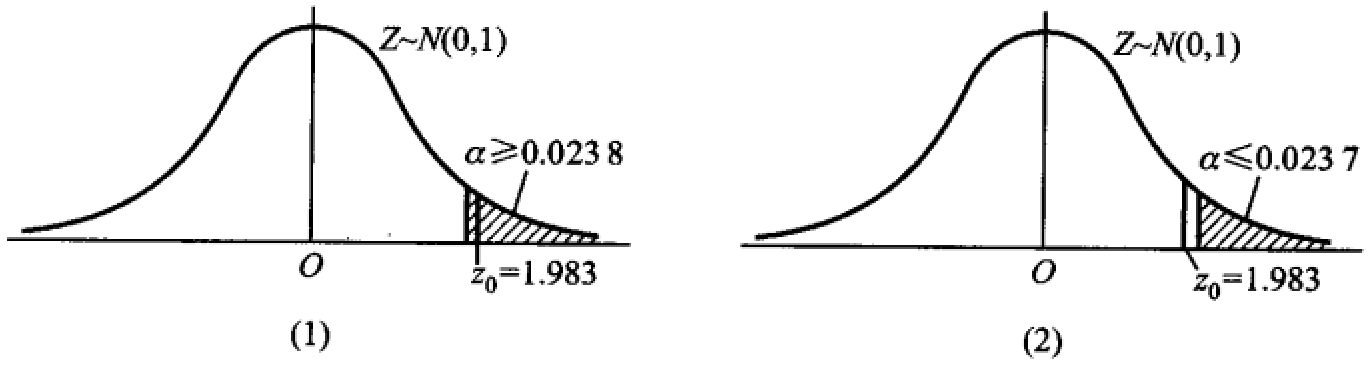

图 8－7

若显著性水平 $\alpha\geqslant p=0.0238$，则对应的临界值 $z_\alpha\leqslant 1.983$，这表示观察值 $z_0=1.983$ 落在拒绝域内(图 8－7(1))，因而拒绝 H_0；又若显著性水平 $\alpha<p=0.0238$，则对应的临界值 $z_\alpha>1.983$，这表示观察值 $z_0=1.983$ 不落在拒绝域内(图 8－7(2))，因而接受 H_0.

据此，p 值$=P\{Z\geqslant z_0\}=0.0238$ 是原假设 H_0 可被拒绝的最小显著性水平.

一般，p 值的定义是：

定义　假设检验问题的 p 值(probability value)是由检验统计量的样本观察值得出的原假设可被拒绝的最小显著性水平.

任一检验问题的 p 值可以根据检验统计量的样本观察值以及检验统计量在 H_0 下一个特定的参数值(一般是 H_0 与 H_1 所规定的参数的分界点)对应的分布求出. 例如在正态总体 $N(\mu,\sigma^2)$ 均值的检验中，当 σ 未知时，可采用检验统计量 $t=\dfrac{\overline{X}-\mu_0}{S/\sqrt{n}}$，在以下三个检验问题中，当 $\mu=\mu_0$ 时 $t\sim t(n-1)$. 如果由样本求得统计量 t 的观察值为 t_0，那么在检验问题

$H_0:\mu\leqslant\mu_0,H_1:\mu>\mu_0$ 中，

p 值$=P_{\mu_0}\{t\geqslant t_0\}=t_0$ 右侧尾部面积，如图(8－8(1))；

$H_0:\mu\geqslant\mu_0,H_1:\mu<\mu_0$ 中，

p 值 $=P_{\mu_0}\{t\leqslant t_0\}=t_0$ 左侧尾部面积，如图(8－8(2))

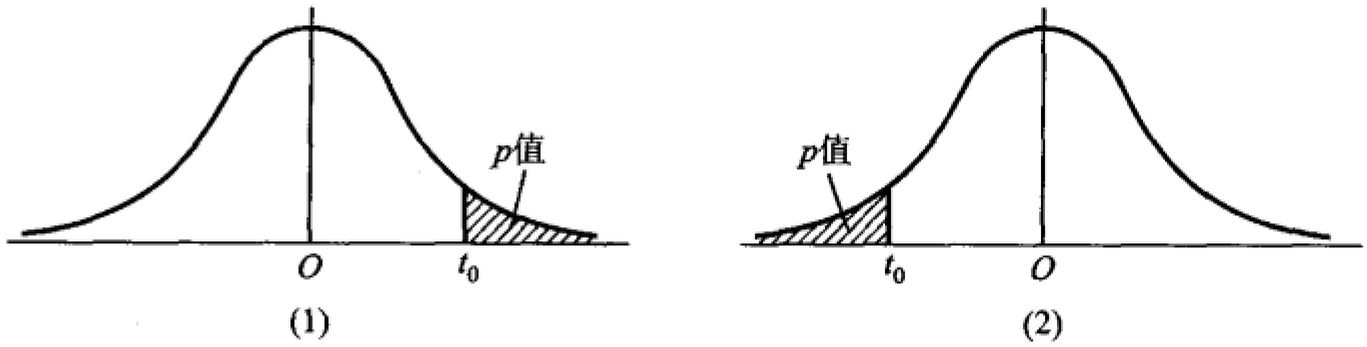

图 8－8

$H_0:\mu=\mu_0, H_1:\mu\neq\mu_0$ 中，

(i) 当 $t_0>0$ 时

$$p\text{ 值}=P_{\mu_0}\{|t|\geqslant t_0\}=P_{\mu_0}\{(t\leqslant -t_0)\cup(t\geqslant t_0)\}$$
$$=2\times(t_0\text{ 右侧尾部面积})\text{（如图 8－9(1)）}.$$

(ii) 当 $t_0<0$ 时

$$p\text{ 值}=P_{\mu_0}\{|t|\geqslant -t_0\}=P_{\mu_0}\{(t\leqslant t_0)\cup(t\geqslant -t_0)\}$$
$$=2\times(t_0\text{ 左侧尾部面积})\text{（如图 8－9(2)）}.$$

综合(i)(ii)，p 值 $=2\times$(由 t_0 界定的尾部面积).

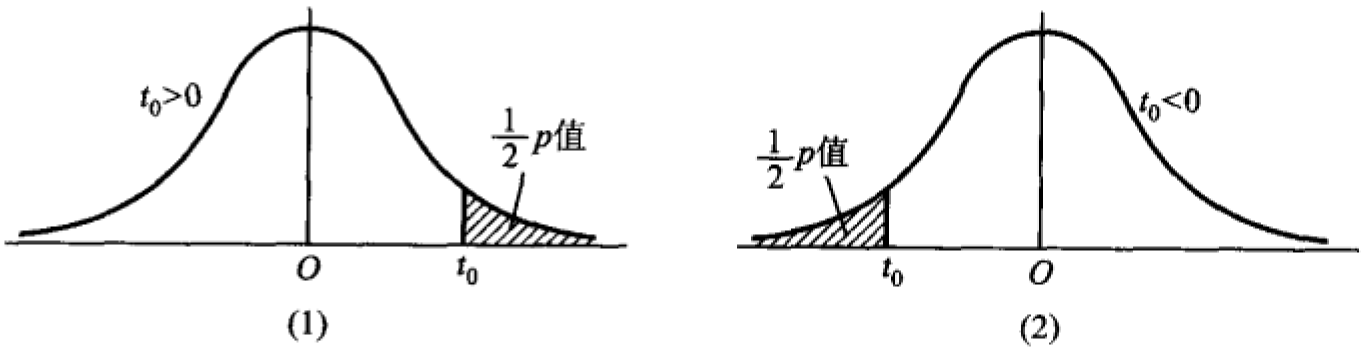

图 8－9

上述各图中的曲线均为 $t(n-1)$ 分布的概率密度曲线.

在现代计算机统计软件中，一般都给出检验问题的 p 值.

按 p 值的定义，对于任意指定的显著性水平 α，就有

(1) 若 p 值 $\leqslant\alpha$，则在显著性水平 α 下拒绝 H_0.

(2) 若 p 值 $>\alpha$，则在显著性水平 α 下接受 H_0.

有了这两条结论就能方便地确定 H_0 的拒绝域. 这种利用 p 值来确定检验拒绝域的方法，称为 p **值检验法**.

用临界值法来确定 H_0 的拒绝域时，例如当取 $\alpha=0.05$ 时知道要拒绝 H_0，再取 $\alpha=0.01$ 也要拒绝 H_0，但不能知道将 α 再降低一些是否也要拒绝 H_0. 而 p 值法给出了拒绝 H_0 的最小显著性水平. 因此 p 值法比临界值法给出了有关拒

绝域的更多的信息.

例 2　用 p 值法检验本章 § 1 例 2 的检验问题

$$H_0:\mu\leqslant\mu_0=-0.545,\quad H_1:\mu>\mu_0\quad \alpha=0.05.$$

解　用 Z 检验法，现在检验统计量 $Z=\dfrac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}$ 的观察值为

$$z_0=\frac{-0.535-(-0.545)}{0.008/\sqrt{5}}=2.795\ 5.$$

$$p\text{ 值}=P\{Z\geqslant 2.795\ 5\}=1-\Phi(2.795\ 5)=0.002\ 6.$$

p 值 $<\alpha=0.05$，故拒绝 H_0. □

例 3　用 p 检验法检验本章 § 2 例 1 的检验问题

$$H_0:\mu\leqslant\mu_0=225,\quad H_1:\mu>225,\quad \alpha=0.05.$$

解　用 t 检验法，现在检验统计量 $t=\dfrac{\overline{X}-\mu_0}{S/\sqrt{n}}$ 的观察值为

$$t_0=\frac{241.5-225}{98.725\ 9/\sqrt{16}}=0.668\ 5,$$

由计算机算得

$$p\text{ 值}=P\{t\geqslant 0.668\ 5\}=0.257\ 0,$$

p 值 $>\alpha=0.05$，故接受 H_0. □

例 4　用 p 值检验法检验本章 § 3 例 1 中检验问题

$$H_0:\sigma^2=5\ 000, H_1:\sigma^2\neq 5\ 000, \alpha=0.02.$$

解　用 χ^2 检验法. 现在检验统计量 $\chi^2=\dfrac{(n-1)S^2}{\sigma_0^2}$ 的观察值为

$$\chi^2=\frac{25\times 9\ 200}{5\ 000}=46,$$

由计算机得

$$p\text{ 值}=2\times P\{\chi^2\geqslant 46\}=0.012\ 8,$$

p 值 $<\alpha=0.02$，故拒绝 H_0. □

p 值表示反对原假设 H_0 的依据的强度，p 值越小，反对 H_0 的依据越强、越充分（譬如对于某个检验问题的检验统计量的观察值的 p 值 $=0.000\ 9$，p 值如此的小，以至于几乎不可能在 H_0 为真时出现目前的观察值，这说明拒绝 H_0 的理由很强，我们就拒绝 H_0）.

一般，若 p 值 $\leqslant 0.01$，称推断拒绝 H_0 的依据很强或称检验是高度显著的；若 $0.01<p$ 值 $\leqslant 0.05$ 称推断拒绝 H_0 的依据是强的或称检验是显著的；若 $0.05<p$ 值 $\leqslant 0.1$ 称推断拒绝 H_0 的理由是弱的，检验是不显著的；若 p 值 >0.1 一般来说没有理由拒绝 H_0. 基于 p 值，研究者可以使用任意希望的显著性水平来

作计算. 在杂志上或在一些技术报告中,许多研究者在讲述假设检验的结果时,常不明显地论及显著性水平以及临界值,代之以简单地引用假设检验的 p 值,利用或让读者利用它来评价反对原假设的依据的强度,作出推断.

小结

统计推断就是由样本来推断总体,它包括两个基本问题:统计估计和假设检验. 上一章讲述了参数估计,本章讨论假设检验问题. 有关总体分布的未知参数或未知分布形式的种种论断叫统计假设,人们要根据样本所提供的信息对所考虑的假设作出接受或拒绝的决策. 假设检验就是作出这一决策的过程.

一般,人们总是对原假设 H_0 作出接受或拒绝的决策. 由于作出判断原假设 H_0 是否为真的依据是一个样本,由于样本的随机性,当 H_0 为真时,检验统计量的观察值也会落入拒绝域,致使我们作出拒绝 H_0 的错误决策;而当 H_0 为不真时,检验统计量的观察值也会未落入拒绝域,致使我们作出接受 H_0 的错误决策.

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

我们使用"接受假设"或"拒绝假设"这样的术语. 接受一个假设并不意味着确信它是真的,它只意味着决定采取某种行动(例如 A);拒绝一个假设也不意味着它是假的,这也仅仅是作出采取另一种不同的行动(例如 B). 不论哪种情况,都存在作出错误选择的可能性.

当样本容量 n 固定时,减小犯第Ⅰ类错误的概率,就会增大犯第Ⅱ类错误的概率,反之亦然. 我们的做法是控制犯第Ⅰ类错误的概率,使

$$P\{当\ H_0\ 为真拒绝\ H_0\} \leqslant \alpha,$$

其中 $0<\alpha<1$ 是给定的小的数. α 称为检验的显著性水平. 这种只对犯第Ⅰ类错误的概率加以控制而不考虑犯第Ⅱ类错误的概率的检验称为显著性检验.

在进行显著性检验时,犯第Ⅰ类错误的概率是由我们控制的. α 取得小,则概率 $P\{当\ H_0$ 为真拒绝 $H_0\}$ 就小,这保证了当 H_0 为真时错误地拒绝 H_0 的可能性很小. 这意味着 H_0 是受到保护的,也表明 H_0、H_1 的地位不是对等的. 于是,在一对对立假设中,选哪一个作为 H_0 需要小心. 例如,考虑某种药品是否为真,这里可能犯两种错误:(1)将假药误作为真药,则冒着伤害病人的健康甚至生命的风险;(2)将真药误作为假药,则冒着造成经济损失的风险. 显然,犯错误(1)比犯错误(2)的后果严重,因此,我们选取"H_0:药品为假,H_1:药品为真",即是使得犯第Ⅰ类错误"当药品为假时错判药品为真"的概率$\leqslant\alpha$. 就是说,选择 H_0、H_1 使得两类错误中后果严重的错误成为第Ⅰ类错误. 这是选择 H_0、H_1 的一个原则.

如果在两类错误中,没有一类错误的后果严重更需要避免时,常常取 H_0 为维持现状,即

取 H_0 为“无效益”、“无改进”、“无价值”等等．例如，取

$$H_0:\text{新技术未提高效益},\quad H_1:\text{新技术提高效益}.$$

实际上，我们感兴趣的是 H_1“提高效益”，但对采用新技术应持慎重态度．选取 H_0 为“新技术未提高效益”，一旦 H_0 被拒绝了，表示有较强的理由去采用新技术．

在实际问题中，情况比较复杂，如何选取 H_0、H_1 只能在实践中积累经验，根据实际情况去判断了．

注意，拒绝域的形式是由 H_1 确定的．

我们还介绍了置信区间与假设检验的关系．知道了置信区间就能容易判明是否接受原假设；反之，知道了检验的接受域就得到了相应的置信区间．

■ **重要术语及主题**

原假设　备择假设　检验统计量　单边检验　双边检验　显著性水平　拒绝域　显著性检验　一个正态总体的参数的检验　两个正态总体均值差、方差比的检验　成对数据的检验　χ^2 分布拟合检验　偏度、峰度检验　秩和检验

习题

1. 某批矿砂的 5 个样品中的镍含量，经测定为(%)

3.25　3.27　3.24　3.26　3.24

设测定值总体服从正态分布，但参数均未知，问在 $\alpha=0.01$ 下能否接受假设：这批矿砂的镍含量的均值为 3.25.

2. 如果一个矩形的宽度 w 与长度 l 的比 $w/l=\frac{1}{2}(\sqrt{5}-1)\approx 0.618$，这样的矩形称为黄金矩形．这种尺寸的矩形使人们看上去有良好的感觉．现代的建筑构件(如窗架)、工艺品(如图片镜框)，甚至司机的执照、商业的信用卡等常常都是采用黄金矩形．下面列出某工艺品工厂随机取的 20 个矩形的宽度与长度的比值：

0.693	0.749	0.654	0.670	0.662	0.672	0.615	0.606	0.690	0.628
0.668	0.611	0.606	0.609	0.601	0.553	0.570	0.844	0.576	0.933

设这一工厂生产的矩形的宽度与长度的比值总体服从正态分布，其均值为 μ，方差为 σ^2，μ,σ^2 均未知．试检验假设(取 $\alpha=0.05$)

$$H_0:\mu=0.618,\quad H_1:\mu\neq 0.618.$$

3. 要求一种元件平均使用寿命不得低于 1 000 h，生产者从一批这种元件中随机抽取 25 件，测得其寿命的平均值为 950 h．已知该种元件寿命服从标准差为 $\sigma=100$ h 的正态分布．试在显著性水平 $\alpha=0.05$ 下判断这批元件是否合格？设总体均值为 μ，μ 未知．即需检验假设 $H_0:\mu\geqslant 1\,000$，$H_1:\mu<1\,000$.

4. 下面列出的是某工厂随机选取的 20 只部件的装配时间(min)：

9.8	10.4	10.6	9.6	9.7	9.9	10.9	11.1	9.6	10.2
10.3	9.6	9.9	11.2	10.6	9.8	10.5	10.1	10.5	9.7

设装配时间的总体服从正态分布 $N(\mu,\sigma^2)$，μ,σ^2 均未知．是否可以认为装配时间的均值显著大于 10 (取 $\alpha=0.05$)？

5. 按规定，100 g 罐头番茄汁中的平均维生素 C 含量不得少于 21 mg/g. 现从工厂的产品中抽取 17 个罐头，其 100 g 番茄汁中，测得维生素 C 含量(mg/g)记录如下：

16　25　21　20　23　21　19　15　13　23　17　20　29　18　22　16　22

设维生素含量服从正态分布 $N(\mu,\sigma^2)$，μ,σ^2 均未知，问这批罐头是否符合要求(取显著性水平 $\alpha=0.05$).

6. 下表分别给出两位文学家马克·吐温(Mark Twain)的 8 篇小品文以及斯诺特格拉斯(Snodgrass)的 10 篇小品文中由 3 个字母组成的单字的比例.

马克·吐温	0.225	0.262	0.217	0.240	0.230	0.229	0.235	0.217		
斯诺特格拉斯	0.209	0.205	0.196	0.210	0.202	0.207	0.224	0.223	0.220	0.201

设两组数据分别来自正态总体，且两总体方差相等，但参数均未知. 两样本相互独立. 问两位作家所写的小品文中包含由 3 个字母组成的单字的比例是否有显著的差异(取 $\alpha=0.05$)?

7. 在 20 世纪 70 年代后期人们发现，在酿造啤酒时，在麦芽干燥过程中形成致癌物质亚硝基二甲胺(NDMA). 到了 20 世纪 80 年代初期开发了一种新的麦芽干燥过程. 下面给出分别在新老两种过程中形成的 NDMA 含量(以 10 亿份中的份数计)：

老过程	6	4	5	5	6	5	5	6	4	6	7	4
新过程	2	1	2	2	1	0	3	2	1	0	1	3

设两样本分别来自正态总体，且两总体的方差相等，但参数均未知. 两样本独立. 分别以 μ_1，μ_2 记对应于老、新过程的总体的均值，试检验假设($\alpha=0.05$)

$$H_0:\mu_1-\mu_2\leqslant 2,\quad H_1:\mu_1-\mu_2>2.$$

8. 随机地选了 8 个人，分别测量了他们在早晨起床时和晚上就寝时的身高(cm)，得到以下的数据.

序号	1	2	3	4	5	6	7	8
早上(x_i)	172	168	180	181	160	163	165	177
晚上(y_i)	172	167	177	179	159	161	166	175

设各对数据的差 $D_i=X_i-Y_i\ (i=1,2,\cdots,8)$ 是来自正态总体 $N(\mu_D,\sigma_D^2)$ 的样本，μ_D,σ_D^2 均未知. 问是否可以认为早晨的身高比晚上的身高要高(取 $\alpha=0.05$)?

9. 为了比较用来做鞋子后跟的两种材料的质量，选取了 15 名男子(他们的生活条件各不相同)，每人穿一双新鞋，其中一只是以材料 A 做后跟，另一只以材料 B 做后跟，其厚度均为 10 mm. 过了一个月再测量厚度，得到数据如下：

男子	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
材料 $A(x_i)$	6.6	7.0	8.3	8.2	5.2	9.3	7.9	8.5	7.8	7.5	6.1	8.9	6.1	9.4	9.1
材料 $B(y_i)$	7.4	5.4	8.8	8.0	6.8	9.1	6.3	7.5	7.0	6.5	4.4	7.7	4.2	9.4	9.1

设 $D_i=X_i-Y_i\quad(i=1,2,\cdots,15)$是来自正态总体 $N(\mu_D,\sigma_D^2)$的样本，μ_D,σ_D^2 均未知.问是否可以认为以材料 A 制成的后跟比材料 B 的耐穿(取 $\alpha=0.05$)?

10. 为了试验两种不同的某谷物的种子的优劣，选取了10块土质不同的土地，并将每块土地分为面积相同的两部分，分别种植这两种种子.设在每块土地的两部分人工管理等条件完全一样.下面给出各块土地上的单位面积产量：

土地编号 i	1	2	3	4	5	6	7	8	9	10
种子 $A(x_i)$	23	35	29	42	39	29	37	34	35	28
种子 $B(y_i)$	26	39	35	40	38	24	36	27	41	27

设 $D_i=X_i-Y_i\quad(i=1,2,\cdots,10)$是来自正态总体 $N(\mu_D,\sigma_D^2)$的样本，μ_D,σ_D^2 均未知.问以这两种种子种植的谷物的产量是否有显著的差异(取 $\alpha=0.05$)?

11. 一种混杂的小麦品种，株高的标准差为 $\sigma_0=14$ cm，经提纯后随机抽取10株，它们的株高(以 cm 计)为

$$90\quad 105\quad 101\quad 95\quad 100\quad 100\quad 101\quad 105\quad 93\quad 97$$

考察提纯后群体是否比原群体整齐？取显著性水平 $\alpha=0.01$，并设小麦株高服从 $N(\mu,\sigma^2)$.

12. 某种导线，要求其电阻的标准差不得超过 0.005 Ω，今在生产的一批导线中取样品9根，测得 $s=0.007$ Ω，设总体为正态分布，参数均未知.问在显著性水平 $\alpha=0.05$ 下能否认为这批导线的标准差显著地偏大?

13. 在第2题中记总体的标准差为 σ，试检验假设(取 $\alpha=0.05$)

$$H_0:\sigma^2=0.11^2,\quad H_1:\sigma^2\neq 0.11^2.$$

14. 测定某种溶液中的水分，它的10个测定值给出 $s=0.037\%$，设测定值总体为正态分布，σ^2 为总体方差，σ^2 未知.试在显著性水平 $\alpha=0.05$ 下检验假设

$$H_0:\sigma\geqslant 0.04\%,\quad H_1:\sigma<0.04\%.$$

15. 在第6题中分别记两个总体的方差为 σ_1^2 和 σ_2^2.试检验假设(取 $\alpha=0.05$)

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

以说明在第6题中我们假设 $\sigma_1^2=\sigma_2^2$ 是合理的.

16. 在第7题中分别记两个总体的方差为 σ_1^2 和 σ_2^2.试检验假设(取 $\alpha=0.05$)

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

以说明在第7题中我们假设 $\sigma_1^2=\sigma_2^2$ 是合理的.

17. 两种小麦品种从播种到抽穗所需的天数如下：

x	101	100	99	99	98	100	98	99	99	99
y	100	98	100	99	98	99	98	98	99	100

设两样本依次来自正态总体 $N(\mu_1,\sigma_1^2)$，$N(\mu_2,\sigma_2^2)$，$\mu_i,\sigma_i\quad(i=1,2)$均未知，两样本相互独立.

(1) 试检验假设 $H_0:\sigma_1^2=\sigma_2^2$，$H_1:\sigma_1^2\neq\sigma_2^2$ (取 $\alpha=0.05$).

(2) 若能接受 H_0，接着检验假设 $H_0':\mu_1=\mu_2$，$H_1':\mu_1\neq\mu_2$ (取 $\alpha=0.05$).

18. 用一种叫"混乱指标"的尺度去衡量工程师的英语文章的可理解性，对混乱指标的打分越低表示可理解性越高．分别随机选取 13 篇刊载在工程杂志上的论文，以及 10 篇未出版的学术报告，对它们的打分列于下表：

工程杂志上的论文(数据Ⅰ)	未出版的学术报告(数据Ⅱ)
1.79　1.75　1.67　1.65	2.39　2.51　2.86
1.87　1.74　1.94	2.56　2.29　2.49
1.62　2.06　1.33	2.36　2.58
1.96　1.69　1.70	2.62　2.41

设数据Ⅰ，Ⅱ分别来自正态总体 $N(\mu_1,\sigma_1^2)$，$N(\mu_2,\sigma_2^2)$，$\mu_1,\mu_2,\sigma_1^2,\sigma_2^2$ 均未知，两样本独立．

(1) 试检验假设 $H_0:\sigma_1^2=\sigma_2^2$，$H_1:\sigma_1^2\neq\sigma_2^2$ (取 $\alpha=0.1$)．

(2) 若能接受 H_0，接着检验假设 $H_0':\mu_1=\mu_2$，$H_1':\mu_1\neq\mu_2$(取 $\alpha=0.1$)．

19. 有两台机器生产金属部件．分别在两台机器所生产的部件中各取一容量 $n_1=60$，$n_2=40$ 的样本，测得部件重量(以 kg 计)的样本方差分别为 $s_1^2=15.46$，$s_2^2=9.66$．设两样本相互独立．两总体分别服从 $N(\mu_1,\sigma_1^2)$，$N(\mu_2,\sigma_2^2)$分布．μ_i,σ_i^2 $(i=1,2)$均未知．试在显著性水平 $\alpha=0.05$ 下检验假设

$$H_0:\sigma_1^2\leqslant\sigma_2^2,\quad H_1:\sigma_1^2>\sigma_2^2.$$

20. 设需要对某一正态总体的均值进行假设检验

$$H_0:\mu\geqslant 15,\quad H_1:\mu<15.$$

已知 $\sigma^2=2.5$．取 $\alpha=0.05$．若要求当 H_1 中的 $\mu\leqslant 13$ 时犯第Ⅱ类错误的概率不超过 $\beta=0.05$，求所需的样本容量．

21. 电池在货架上滞留的时间不能太长．下面给出某商店随机选取的 8 只电池的货架滞留时间(以天计)：

$$108\quad 124\quad 124\quad 106\quad 138\quad 163\quad 159\quad 134$$

设数据来自正态总体 $N(\mu,\sigma^2)$，μ,σ^2 未知．

(1) 试检验假设 $H_0:\mu\leqslant 125$，$H_1:\mu>125$，取 $\alpha=0.05$．

(2) 若要求在上述 H_1 中$(\mu-125)/\sigma\geqslant 1.4$时，犯第Ⅱ类错误的概率不超过 $\beta=0.1$，求所需的样本容量．

22. 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

$$H_0:\mu_1\leqslant 2\mu_2,\quad H_1:\mu_1>2\mu_2.$$

此处 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_1^2,σ_2^2．现分别在两总体中取一样本 $X_1,X_2,\cdots,X_{n_1}$ 和 $Y_1,Y_2,\cdots,Y_{n_2}$，设两个样本独立．试给出上述假设 H_0 的拒绝域，取显著性水平为 α．

23. 检查了一本书的 100 页，记录各页中印刷错误的个数，其结果为

错误个数 f_i	0	1	2	3	4	5	6	$\geqslant 7$
含 f_i 个错误的页数	36	40	19	2	0	2	1	0

问能否认为一页的印刷错误的个数服从泊松分布(取 $\alpha=0.05$).

24. 在一批灯泡中抽取 300 只作寿命试验,其结果如下:

寿命 t(h)	$0\leqslant t\leqslant 100$	$100<t\leqslant 200$	$200<t\leqslant 300$	$t>300$
灯泡数	121	78	43	58

取 $\alpha=0.05$,试检验假设

H_0:灯泡寿命服从指数分布

$$f(t)=\begin{cases}0.005e^{-0.005t}, & t\geqslant 0,\\ 0, & t<0.\end{cases}$$

25. 下面给出了随机选取的某大学一年级学生(200 名)一次数学考试的成绩.

(1) 画出数据的直方图.

(2) 试取 $\alpha=0.1$ 检验数据来自正态总体 $N(60,15^2)$.

分数 x	$20\leqslant x\leqslant 30$	$30<x\leqslant 40$	$40<x\leqslant 50$	$50<x\leqslant 60$
学生数	5	15	30	51
分数 x	$60<x\leqslant 70$	$70<x\leqslant 80$	$80<x\leqslant 90$	$90<x\leqslant 100$
学生数	60	23	10	6

26. 袋中装有 8 只球,其中红球数未知.在其中任取 3 只,记录红球的只数 X,然后放回,再任取 3 只,记录红球的只数,然后放回.如此重复进行了 112 次,其结果如下:

x	0	1	2	3
次数	1	31	55	25

试取 $\alpha=0.05$ 检验假设

H_0:X 服从超几何分布

$$P\{X=k\}=\binom{5}{k}\binom{3}{3-k}\Big/\binom{8}{3},\quad k=0,1,2,3.$$

即检验假设 H_0:红球的只数为 5.

27. 一农场 10 年前在一鱼塘中按比例 20∶15∶40∶25 投放了四种鱼:鲑鱼、鲈鱼、竹夹鱼和鲇鱼的鱼苗,现在在鱼塘里获得一样本如下:

序号	1	2	3	4	
种类	鲑鱼	鲈鱼	竹夹鱼	鲇鱼	
数量(条)	132	100	200	168	$\sum=600$

试取 $\alpha=0.05$,检验各类鱼数量的比例较 10 年前是否有显著的改变.

28. 某种鸟在起飞前,双足齐跳的次数 X 服从几何分布,其分布律为

$$P\{X=x\}=p^{x-1}(1-p),\quad x=1,2,\cdots.$$

今获得一样本如下:

x	1	2	3	4	5	6	7	8	9	10	11	12	$\geqslant 13$
观察到 x 的次数	48	31	20	9	6	5	4	2	1	1	2	1	0

（1）求 p 的最大似然估计值.

（2）取 $\alpha=0.05$，检验假设：H_0：数据来自总体 $P\{X=x\}=p^{x-1}(1-p)$，$x=1,2,\cdots$.

29. 分别抽查了两球队部分队员行李的重量(kg)为：

1 队	34	39	41	28	33	
2 队	36	40	35	31	39	36

设两样本独立且 1，2 两队队员行李重量总体的概率密度至多差一个平移. 记两总体的均值分别为 μ_1，μ_2，且 μ_1，μ_2 均未知. 试检验假设：$H_0:\mu_1=\mu_2$，$H_1:\mu_1<\mu_2$ （取 $\alpha=0.05$）.

30. 下面给出两种型号的计算器充电以后所能使用的时间(h)：

型号 A	5.5	5.6	6.3	4.6	5.3	5.0	6.2	5.8	5.1	5.2	5.9	
型号 B	3.8	4.3	4.2	4.0	4.9	4.5	5.2	4.8	4.5	3.9	3.7	4.6

设两样本独立且数据所属的两总体的概率密度至多差一个平移，试问能否认为型号 A 的计算器平均使用时间比型号 B 来得长（$\alpha=0.01$）?

31. 下面给出两个工人五天生产同一种产品每天生产的件数：

工人 A	49	52	53	47	50
工人 B	56	48	58	46	55

设两样本独立且数据所属的两总体的概率密度至多差一个平移. 问能否认为工人 A、工人 B 平均每天完成的件数没有显著差异（$\alpha=0.1$）?

32. （1）设总体服从 $N(\mu,100)$，μ 未知，现有样本：$n=16$，$\overline{x}=13.5$，试检验假设 $H_0:\mu\leqslant 10$，$H_1:\mu>10$，(i) 取 $\alpha=0.05$，(ii) 取 $\alpha=0.10$，(iii) H_0 可被拒绝的最小显著性水平.

（2）考察生长在老鼠身上的肿块的大小. 以 X 表示在老鼠身上生长了 15 天的肿块的直径（以 mm 计），设 $X\sim N(\mu,\sigma^2)$，μ,σ^2 均未知. 今随机地取 9 只老鼠（在它们身上的肿块都长了 15 天），测得 $\overline{x}=4.3$，$s=1.2$，试取 $\alpha=0.05$，用 p 值检验法检验假设 $H_0:\mu=4.0$，$H_1:\mu\neq 4.0$，求出 p 值.

（3）用 p 值检验法检验 §2 例 4 的检验问题.

（4）用 p 值检验法检验第 27 题中的检验问题.