第七章　假设检验

§7.1　假设检验的基本思想与概念

1. 假设

- 参数空间 $\Theta=\{\theta\}$ 的非空子集或有关参数 θ 的命题,称为统计假设,简称假设.
- 原假设,根据需要而设立的假设,常记为 $H_0:\theta\in\Theta_0$.
- 备择假设,在原假设被拒绝后而采用(接受)的假设,常记为 $H_1:\theta\in\Theta_1$.

注:要求 $\Theta_0\cap\Theta_1=\varnothing$,即原假设 H_0 与备择假设 H_1 不含公共参数.

如果假设 H_0(或 H_1)只含一个点,则称该假设是简单假设,否则称为复杂假设.

2. 检验　对原假设 $H_0:\theta\in\Theta_0$ 作出是否拒绝 H_0 的判断的法则称为检验法则,简称检验.检验有两个结果:

- "原假设不正确",称为拒绝原假设,或称检验显著.
- "原假设正确",称为接受(保留)原假设,或称检验不显著.

统计假设检验的着力点不在于说明原假设正确,而在于说明原假设不正确,因为用一个样本去说明原假设正确是根据不足的,而用一个样本推翻原假设却是合理的,因此统计学在讨论假设检验时着力点在于建立检验的拒绝域.

3. 检验问题

- 由原假设 H_0 和备择假设 H_1 组成一个需要作判断的问题称为检验问题.
- 参数假设检验问题,两个假设都是有关参数的命题组成的检验问题.
- 非参数假设检验问题,两个假设都是有关分布的命题组成的检验问题.

常用的参数假设检验问题有如下三个,其中 θ_0 是已知常数.

（1）$H_0: \theta \leqslant \theta_0 \quad \text{vs} \quad H_1: \theta > \theta_0$；

（2）$H_0: \theta \geqslant \theta_0 \quad \text{vs} \quad H_1: \theta < \theta_0$；

（3）$H_0: \theta = \theta_0 \quad \text{vs} \quad H_1: \theta \neq \theta_0$.

其中(1)与(2)又称单侧检验问题,因为一个假设位于另一个假设的一侧,(3)称为双侧检验问题,因为备择假设位于原假设的两侧.

4. 两类错误及其发生概率

• 原假设 H_0 正确,但被拒绝,这种判断错误称为犯第一类错误,其发生概率称为犯第一类错误的概率,或称拒真概率,常记为 α.

• 原假设 H_0 不真,但被接受,这种判断错误称为犯第二类错误,其发生概率称为犯第二类错误的概率,或称受伪概率,常记为 β.

5. 假设检验的基本步骤

（1）建立假设.根据要求建立原假设 H_0 和备择假设 H_1.

（2）选择检验统计量,给出拒绝域 W 的形式.

• 用于对原假设 H_0 作出判断的统计量称为检验统计量；

• 使原假设被拒绝的样本观察值所在区域称为拒绝域,常用 W 表示.

• 一个拒绝域 W 唯一确定一个检验法则,反之,一个检验法则唯一确定一个拒绝域 W.

（3）选择显著性水平 $\alpha(0<\alpha<1)$.

（4）给出拒绝域.由概率等式 $P(W)=\alpha$ 确定具体的拒绝域.

（5）作出判断.

• 当样本$(x_1, x_2, \cdots, x_n) \in W$,则拒绝 H_0,即接受 H_1.

• 当样本$(x_1, x_2, \cdots, x_n) \in \overline{W}$,则接受 H_0.

6. 势函数　设检验问题 $H_0: \theta \in \Theta_0 \quad \text{vs} \quad H_1: \theta \in \Theta_1$ 的拒绝域为 W,则样本观测值 $x_1, x_2, \cdots, x_n$ 落在拒绝域 W 内的概率称为该检验的势函数,记为

$$g(\theta) = P_\theta((x_1, x_2, \cdots, x_n) \in W), \quad \theta \in \Theta_0 \cup \Theta_1.$$

由势函数 $g(\theta)$ 容易得到犯两类错误的概率

$$g(\theta) = \begin{cases} \alpha(\theta), & \theta \in \Theta_0, \\ 1-\beta(\theta), & \theta \in \Theta_1. \end{cases}$$

7. 水平为 α 的检验　对检验问题 $H_0: \theta \in \Theta_0 \quad \text{vs} \quad H_1: \theta \in \Theta_1$,如果一

个检验犯第一类错误的概率 $\alpha(\theta)$ 不超过事先给定的显著性水平 α,即

$$\alpha(\theta)\leqslant\alpha,\theta\in\Theta_0,$$

则称该检验为水平为 α 的检验,或称显著性水平为 α 的显著性检验.

在实际使用中 α 不宜选得过小,α 过小会导致 β 过大,应在适当控制 α 中制约 β.最常用的选择是 $\alpha=0.05$,有时也选用 $\alpha=0.10$ 或 $\alpha=0.01$.

8. 检验的 p 值 在一个假设检验问题中,利用样本观察值能够做出拒绝原假设的最小显著性水平称为该检验的 p 值,引入检验的 p 值的好处是:

(1) 它比较客观,避免了事先确定显著性水平;

(2) 由检验的 p 值与人们心目中的显著性水平 α 进行比较,

- 如果 $\alpha\geqslant p$,则在显著性水平 α 下拒绝 H_0.
- 如果 $\alpha<p$,则在显著性水平 α 下应保留 H_0.

(3) 检验的 p 值的计算是复杂的,会涉及各种抽样分布.如今统计软件都有计算 p 值的功能,因此这对使用者反而更方便,它不需要备用各种抽样分布的分位数表,而只需要观察计算机的输出的 p 值多少就可以做出判断.

§7.2 正态总体参数假设检验

1. 单个正态总体均值的假设检验

检验法	H_0	H_1	检验统计量	拒绝域	p 值
u 检验 σ 已知	$\mu\leqslant\mu_0$	$\mu>\mu_0$	$u=\dfrac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}$	$\{u\geqslant u_{1-\alpha}\}$	$1-\Phi(u_0)$
	$\mu\geqslant\mu_0$	$\mu<\mu_0$		$\{u\leqslant u_\alpha\}$	$\Phi(u_0)$
	$\mu=\mu_0$	$\mu\neq\mu_0$		$\{\lvert u\rvert\geqslant u_{1-\alpha/2}\}$	$2(1-\Phi(\lvert u_0\rvert))$
t 检验 σ 未知	$\mu\leqslant\mu_0$	$\mu>\mu_0$	$t=\dfrac{\bar{x}-\mu_0}{s/\sqrt{n}}$	$\{t\geqslant t_{1-\alpha}(n-1)\}$	$P(t\geqslant t_0)$
	$\mu\geqslant\mu_0$	$\mu<\mu_0$		$\{t\leqslant t_\alpha(n-1)\}$	$P(t\leqslant t_0)$
	$\mu=\mu_0$	$\mu\neq\mu_0$		$\{\lvert t\rvert\geqslant t_{1-\alpha/2}(n-1)\}$	$P(\lvert t\rvert\geqslant\lvert t_0\rvert)$

注 1:表中 $\bar{x}$ 和 s 分别为来自正态总体 $N(\mu,\sigma^2)$ 的样本 $x_1,x_2,\cdots,x_n$ 的均值和标准差,$u_0=\sqrt{n}(\bar{x}-\mu_0)/\sigma$,$t_0=\sqrt{n}(\bar{x}-\mu_0)/s$,$t$ 是服从 $t(n-1)$ 的随机变量.

注 2:检验问题 $H_0:\mu=\mu_0$ vs $H_1:\mu>\mu_0$ 与检验问题 $H_0:\mu\leqslant\mu_0$ vs $H_1:\mu>\mu_0$ 的拒绝域相同,p 值也一样.另一对单边假设检验也有同样结论.

2. 假设检验与置信区间的关系(以下 α 为显著性水平,$1-\alpha$ 为置信水平)

- 双侧检验问题:$H_0:\mu=\mu_0$ vs $H_1:\mu\neq\mu_0$ 的接受域 $\overline{W}$ 可定出正态均值 μ 的 $1-\alpha$ 置信区间.
- 单侧检验问题:$H_0:\mu\leqslant\mu_0$ vs $H_1:\mu>\mu_0$ 的接受域 $\overline{W}$ 可定出正态均值 μ 的 $1-\alpha$ 置信上限.
- 单侧检验问题:$H_0:\mu\geqslant\mu_0$ vs $H_1:\mu<\mu_0$ 的接受域 $\overline{W}$ 可定出正态均值 μ 的 $1-\alpha$ 置信下限.

3. 两个正态均值差的假设检验

检验	条件	H_0	H_1	检验统计量	拒绝域	p 值
u 检验	σ_1,σ_2 已知	$\mu_1\leqslant\mu_2$	$\mu_1>\mu_2$	$u=\dfrac{(\bar{x}-\bar{y})}{\sqrt{\dfrac{\sigma_1^2}{m}+\dfrac{\sigma_2^2}{n}}}$	$\{u\geqslant u_{1-\alpha}\}$	$1-\Phi(u_1)$
		$\mu_1\geqslant\mu_2$	$\mu_1<\mu_2$		$\{u\leqslant u_\alpha\}$	$\Phi(u_1)$
		$\mu_1=\mu_2$	$\mu_1\neq\mu_2$		$\{\lvert u\rvert\geqslant u_{1-\alpha/2}\}$	$2(1-\Phi(\lvert u_1\rvert))$
t 检验	$\sigma_1=\sigma_2$ 未知	$\mu_1\leqslant\mu_2$	$\mu_1>\mu_2$	$t=\dfrac{(\bar{x}-\bar{y})}{s_w\sqrt{\dfrac{1}{m}+\dfrac{1}{n}}}$	$\{t\geqslant t_{1-\alpha}(m+n-2)\}$	$P(T_1\geqslant t_1)$
		$\mu_1\geqslant\mu_2$	$\mu_1<\mu_2$		$\{t\leqslant t_\alpha(m+n-2)\}$	$P(T_1\leqslant t_1)$
		$\mu_1=\mu_2$	$\mu_1\neq\mu_2$		$\{\lvert t\rvert\geqslant t_{1-\alpha/2}(m+n-2)\}$	$P(\lvert T_1\rvert\geqslant\lvert t_1\rvert)$
大样本 u 检验	m,n 充分大	$\mu_1\leqslant\mu_2$	$\mu_1>\mu_2$	$u=\dfrac{(\bar{x}-\bar{y})}{\sqrt{\dfrac{s_x^2}{m}+\dfrac{s_y^2}{n}}}$	$\{u\geqslant u_{1-\alpha}\}$	$1-\Phi(u_2)$
		$\mu_1\geqslant\mu_2$	$\mu_1<\mu_2$		$\{u\leqslant u_\alpha\}$	$\Phi(u_2)$
		$\mu_1=\mu_2$	$\mu_1\neq\mu_2$		$\{\lvert u\rvert\geqslant u_{1-\alpha/2}\}$	$2(1-\Phi(\lvert u_2\rvert))$
近似 t 检验	m,n 不很大	$\mu_1\leqslant\mu_2$	$\mu_1>\mu_2$	$t=\dfrac{(\bar{x}-\bar{y})}{\sqrt{\dfrac{s_x^2}{m}+\dfrac{s_y^2}{n}}}$	$\{t\geqslant t_{1-\alpha}(l)\}$	$P(T_2\geqslant t_2)$
		$\mu_1\geqslant\mu_2$	$\mu_1<\mu_2$		$\{t\leqslant t_\alpha(l)\}$	$P(T_2\leqslant t_2)$
		$\mu_1=\mu_2$	$\mu_1\neq\mu_2$		$\{\lvert t\rvert\geqslant t_{1-\alpha/2}(l)\}$	$P(\lvert T_2\rvert\geqslant\lvert t_2\rvert)$

注:表中 $\bar{x}$ 和 s_x^2 分别为来自正态总体 $N(\mu_1,\sigma_1^2)$ 的样本 $x_1,x_2,\cdots,x_m$ 的均值和方差;$\bar{y}$ 和 s_y^2 分别为来自正态总体 $N(\mu_2,\sigma_2^2)$ 的样本 $y_1,y_2,\cdots,y_n$ 的均值和方差,$s_w^2=\dfrac{(m-1)s_x^2+(n-1)s_y^2}{m+n-2}$,$s_0^2=\dfrac{s_x^2}{m}+\dfrac{s_y^2}{n}$,$l=\dfrac{s_0^4}{\dfrac{s_x^4}{m^2(m-1)}+\dfrac{s_y^4}{n^2(n-1)}}$,$u_1=\dfrac{(\bar{x}-\bar{y})}{\sqrt{\dfrac{\sigma_1^2}{m}+\dfrac{\sigma_2^2}{n}}}$,$u_2=\dfrac{(\bar{x}-\bar{y})}{\sqrt{\dfrac{s_x^2}{m}+\dfrac{s_y^2}{n}}}$,$t_1=\dfrac{(\bar{x}-\bar{y})}{s_w\sqrt{\dfrac{1}{m}+\dfrac{1}{n}}}$,$t_2=\dfrac{(\bar{x}-\bar{y})}{\sqrt{\dfrac{s_x^2}{m}+\dfrac{s_y^2}{n}}}$,$T_1$ 是服从自由度是 $n+m-2$ 的 t 分布的随机变量,T_2 是服从自由度是 l 的 t 分布的随机变量.

4. 一个和二个正态总体方差的假设检验

检验法	H_0	H_1	检验统计量	拒绝域	p 值
χ^2 检验	$\sigma^2\leqslant\sigma_0^2$	$\sigma^2>\sigma_0^2$	$\chi^2=\frac{(n-1)s^2}{\sigma_0^2}$	$\chi^2\geqslant\chi^2_{1-\alpha}(n-1)$	$P(\chi^2\geqslant\chi_0^2)$
	$\sigma^2\geqslant\sigma_0^2$	$\sigma^2<\sigma_0^2$		$\chi^2\leqslant\chi^2_{\alpha}(n-1)$	$P(\chi^2\leqslant\chi_0^2)$
	$\sigma^2=\sigma_0^2$	$\sigma^2\neq\sigma_0^2$		$\chi^2\leqslant\chi^2_{\alpha/2}(n-1)$ 或 $\chi^2\geqslant\chi^2_{1-\alpha/2}(n-1)$	$2\min\{P(\chi^2\geqslant\chi_0^2),P(\chi^2\leqslant\chi_0^2)\}$
F 检验	$\sigma_1^2\leqslant\sigma_2^2$	$\sigma_1^2>\sigma_2^2$	$F=\frac{s_x^2}{s_y^2}$	$F\geqslant F_{1-\alpha}(m-1,n-1)$	$P(F\geqslant F_0)$
	$\sigma_1^2\geqslant\sigma_2^2$	$\sigma_1^2<\sigma_2^2$		$F\leqslant F_{\alpha}(m-1,n-1)$	$P(F\leqslant F_0)$
	$\sigma_1^2=\sigma_2^2$	$\sigma_1^2\neq\sigma_2^2$		$F\leqslant F_{\alpha/2}(m-1,n-1)$ 或 $F\geqslant F_{1-\alpha/2}(m-1,n-1)$	$2\min\{P(F\geqslant F_0),P(F\leqslant F_0)\}$

注:表中 s^2 为来自正态总体 $N(\mu,\sigma^2)$ 的样本 $x_1,x_2,\cdots,x_n$ 的方差.$\chi_0^2=(n-1)s^2/\sigma_0^2$ 是由样本计算得到的检验统计量的值.χ^2 是服从 $\chi^2(n-1)$ 分布的随机变量.s_x^2 为来自正态总体 $N(\mu,\sigma_1^2)$ 的样本 $x_1,x_2,\cdots,x_m$ 的方差;s_y^2 为来自正态总体 $N(\mu_2,\sigma_2^2)$ 的样本 $y_1,y_2,\cdots,y_n$ 的方差,$F_0=s_x^2/s_y^2$ 是由样本计算得到的检验统计量的值,F 是服从 $F(m-1,n-1)$ 分布的随机变量.

5. 成对数据检验

在对两个总体均值进行比较时,若数据是成对出现的,则应采用成对数据检验.

成对数据检验就是把两个总体均值的比较通过成对数据的差转变为对单个总体均值的检验,方法完全等同.

成对数据的获得事先要作周全的安排(即试验设计).在获得成对数据时不要发生“错位”,从而失去“成对数据”的信息.

§7.3 其他分布参数的假设检验

1. 指数分布均值 θ 的假设检验——χ^2 检验

H_0	H_1	检验统计量	拒绝域	p 值
$\theta\leqslant\theta_0$	$\theta>\theta_0$	$\chi^2=\frac{2n\bar{x}}{\theta_0}$	$W=\{\chi^2\geqslant\chi^2_{1-\alpha}(2n)\}$	$P(\chi^2\geqslant\chi_0^2)$
$\theta\geqslant\theta_0$	$\theta<\theta_0$		$W=\{\chi^2\leqslant\chi^2_{\alpha}(2n)\}$	$P(\chi^2\leqslant\chi_0^2)$
$\theta=\theta_0$	$\theta\neq\theta_0$		$W=\{\chi^2\leqslant\chi^2_{\alpha/2}(2n)$ 或 $\chi^2\geqslant\chi^2_{1-\alpha/2}(2n)\}$	$2\min\{P(\chi^2\geqslant\chi_0^2),P(\chi^2\leqslant\chi_0^2)\}$

注:$\chi_0^2=\frac{2n\bar{x}}{\theta_0}$为由样本计算得到的检验统计量的值,$\chi^2$ 是服从$\chi^2(2n)$分布的随机变量.

2. 比例 p 的假设检验

(1) 小样本方法:检验统计量为 n 次试验中成功次数 x.

H_0	H_1	p 值
$p\leqslant p_0$	$p>p_0$	$P_{p_0}(x\geqslant x_0)$
$p\geqslant p_0$	$p<p_0$	$P_{\theta_0}(x\leqslant x_0)$
$p=p_0$	$p\neq p_0$	$2\min\{P_{p_0}(x\leqslant x_0),P_{p_0}(x\geqslant x_0)\}$

注：x_0 为样本观测值，x 是服从 $b(n,p_0)$ 的随机变量.

（2）大样本方法：检验统计量为 $u=\dfrac{\sqrt{n}\,(\bar{x}-p_0)}{\sqrt{p_0(1-p_0)}}$.

H_0	H_1	拒绝域	p 值
$p\leqslant p_0$	$p>p_0$	$\{u\geqslant u_{1-\alpha}\}$	$1-\Phi(u_0)$
$p\geqslant p_0$	$p<p_0$	$\{u\leqslant u_{\alpha}\}$	$\Phi(u_0)$
$p=p_0$	$p\neq p_0$	$\{\lvert u\rvert\geqslant u_{1-\alpha/2}\}$	$2(1-\Phi(\lvert u_0\rvert))$

注：u_0 为由样本计算出的观测值.

§7.4　似然比检验与分布拟合检验

1. 似然比检验

设 $x_1,x_2,\cdots,x_n$ 为来自密度函数为 $p(x;\theta)$，$\theta\in\Theta$ 的样本，则对检验问题 $H_0:\theta\in\Theta_0$　vs　$H_1:\theta\in\Theta_1=\Theta-\Theta_0$，可用似然比统计量 $\Lambda=\dfrac{\sup\limits_{\theta\in\Theta}p(x_1,x_2,\cdots,x_n;\theta)}{\sup\limits_{\theta\in\Theta_0}p(x_1,x_2,\cdots,x_n;\theta)}$作检验统计量，该检验称为似然比（Likelihood Ratio）检验，有时也称之为广义似然比检验.

检验统计量也可以写为 $\Lambda=\dfrac{p(x_1,x_2,\cdots,x_n;\hat{\theta})}{p(x_1,x_2,\cdots,x_n;\hat{\theta}_0)}$，其中$\hat{\theta}$表示在全参数空间 Θ 上 θ 的最大似然估计，$\hat{\theta}_0$ 表示在原假设成立时子参数空间 Θ_0 上 θ 的最大似然估计.拒绝域为 $W=\{\Lambda\geqslant c\}$，其中临界值 c 由 $P_\theta(\Lambda\geqslant c)\leqslant\alpha(\forall\theta\in\Theta_0)$确定.

似然比检验方法是产生检验统计量的另一种思路.该似然比检验统计量没有统一的精确分布形式，但其对数似然比的 2 倍——$2\ln\Lambda$，渐近服从

$\chi^2(k)$分布,其中 k 为 Λ 中独立参数个数.

2. 总体可以分成 k 类:$A_1,A_2,\cdots,A_k$ 时的分布拟合优度检验

- 原假设 $H_0:P(A_i)=p_i,\quad i=1,2,\cdots,k$,其中诸 $p_i\geqslant 0$,且 $\sum\limits_{i=1}^{k}p_i=1$.
- 数据:对总体作 n 次观察,k 个类各出现的频数分别为 $n_1,n_2,\cdots,n_k$,且 $\sum\limits_{i=1}^{k}n_i=n$.分两种情况给出检验统计量及其拒绝域:

(1) 诸 p_i 均已知,检验统计量 $\chi^2=\sum\limits_{i=1}^{k}\dfrac{(n_i-np_i)^2}{np_i}$,拒绝域为 $W=\{\chi^2\geqslant\chi^2_{1-\alpha}(k-1)\}$;

(2) 诸 p_i 不完全已知,检验统计量 $\chi^2=\sum\limits_{i=1}^{k}\dfrac{(n_i-n\hat{p}_i)^2}{n\hat{p}_i}$,拒绝域为 $\{\chi^2\geqslant\chi^2_{1-\alpha}(k-r-1)\}$,其中 r 为 $p_1,p_2,\cdots,p_k$ 中独立参数个数,$\hat{p}_i$ 为 p_i 的最大似然估计.

这个检验被皮尔逊(K. Pearson)称为 χ^2 拟合优度检验,$p=P(\chi^2\geqslant\chi^2_0)$ 被称为拟合优度,p 值愈大拟合优度愈好,p 值愈小拟合优度愈差,从而拒绝原假设 H_0.这就是检验的 p 值最原始的想法.

3. 列联表的独立性检验 $r\times c$ 的二维列联表:总体按两个属性 A 与 B 分类,A 有 r 个类:$A_1,A_2,\cdots,A_r$,B 有 c 个类:$B_1,B_2,\cdots,B_c$,共有 rc 个类,若进行 n 次试验,其中既属 A_i 又属 B_j 的结果有 n_{ij}个,按矩阵排列,就得 $r\times c$ 二维列联表.

A	B					行和
	B_1	$\cdots$	B_j	$\cdots$	B_c	
A_1	n_{11}	$\cdots$	n_{1j}	$\cdots$	n_{1c}	$n_{1\cdot}$
$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$
A_i	n_{i1}	$\cdots$	n_{ij}	$\cdots$	n_{ic}	$n_{i\cdot}$
$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$
A_r	n_{r1}	$\cdots$	n_{rj}	$\cdots$	n_{rc}	$n_{r\cdot}$
列和	$n_{\cdot 1}$	$\cdots$	$n_{\cdot j}$	$\cdots$	$n_{\cdot c}$	总和 n

- 原假设 $H_0:P(A_iB_j)=p_{ij}=p_{i\cdot}\,p_{\cdot j}=P(A_i)P(A_j)$,$i=1,2,\cdots,r,j=1,2,\cdots,c$,其意为:属性 A 与 B 相互独立.

● 在诸 p_{ij} 未知(常见)场合,检验统计量为 $\chi^2=\sum_{i=1}^{r}\sum_{j=1}^{c}\frac{(n_{ij}-n\hat{p}_{ij})^2}{n\hat{p}_{ij}}$,其中 $\hat{p}_{ij}=\hat{p}_{i\cdot}\hat{p}_{\cdot j}=\frac{n_{i\cdot}}{n}\cdot\frac{n_{\cdot j}}{n}$ 是 p_{ij} 的最大似然估计.

● 对给定显著性水平 $\alpha(0\leqslant\alpha<1)$ 在 n 较大场合,该检验的拒绝域为 $W=\{\chi^2\geqslant\chi^2_{1-\alpha}((r-1)(c-1))\}$.

§7.5 正态性检验

1. 正态性概率纸检验

具体步骤如下:

● 首先将样本观察值按从小到大的次序排列:$x_{(1)}\leqslant x_{(2)}\leqslant\cdots\leqslant x_{(n)}$.

● 对每一个 i,计算修正频率 $\hat{F}_i=(i-0.375)/(n+0.25)$,$i=1,2,\cdots,n$,将 $\hat{F}_i$ 看作概率 $F(x_{(i)})$ 的估计.

● 将点 $(x_{(i)},\hat{F}_i)$,$i=1,2,\cdots,n$ 逐一点在正态概率纸上.

● 判断:若诸点在一条直线附近,则认为该样本来自正态总体;

若诸点明显不在一条直线附近,则认为该样本不是来自正态分布总体.

● 如果从正态概率纸上确认总体是非正态分布时,可从如下变换

$$y=\ln x,\qquad y=1/x,\quad y=\sqrt{x}$$

中选一个数据作变换,然后用 $(y_{(i)},\hat{F}_i)$ 再描点,判断变换后的数据是否来自正态分布,若 $\ln x\sim N(\mu,\sigma^2)$,则 $x\sim LN(\mu,\sigma^2)$,若 $\frac{1}{x}\sim N(\mu,\sigma^2)$,则 $x\sim IN(\mu,\sigma^2)$(倒正态分布),若 $\sqrt{x}\sim N(\mu,\sigma^2)$,则 χ 服从非中心 χ^2 分布(更一般的 χ^2 分布).

2. 夏皮罗-威尔克(Shapiro - Wilk)检验

具体步骤如下:

● 首先将观察值按从小到大的次序排列:$x_{(1)}\leqslant x_{(2)}\leqslant\cdots\leqslant x_{(n)}$.

● 从附表 6 中查得对应 n 的系数 $a_1,a_2,\cdots,a_n$,其中 $a_{n+1-i}=-a_i$,$i=1,2,\cdots,[n/2]$.

• 计算检验统计量 $W=\dfrac{\left[\sum\limits_{i=1}^{[n/2]} a_i\left(x_{(n+1-i)}-x_{(i)}\right)\right]^2}{\sum\limits_{i=1}^{n}\left(x_{(i)}-\bar{x}\right)^2}$.

• 拒绝域 $W=\{W\leqslant W_\alpha\}$,其中 W_α 可查附表 7.

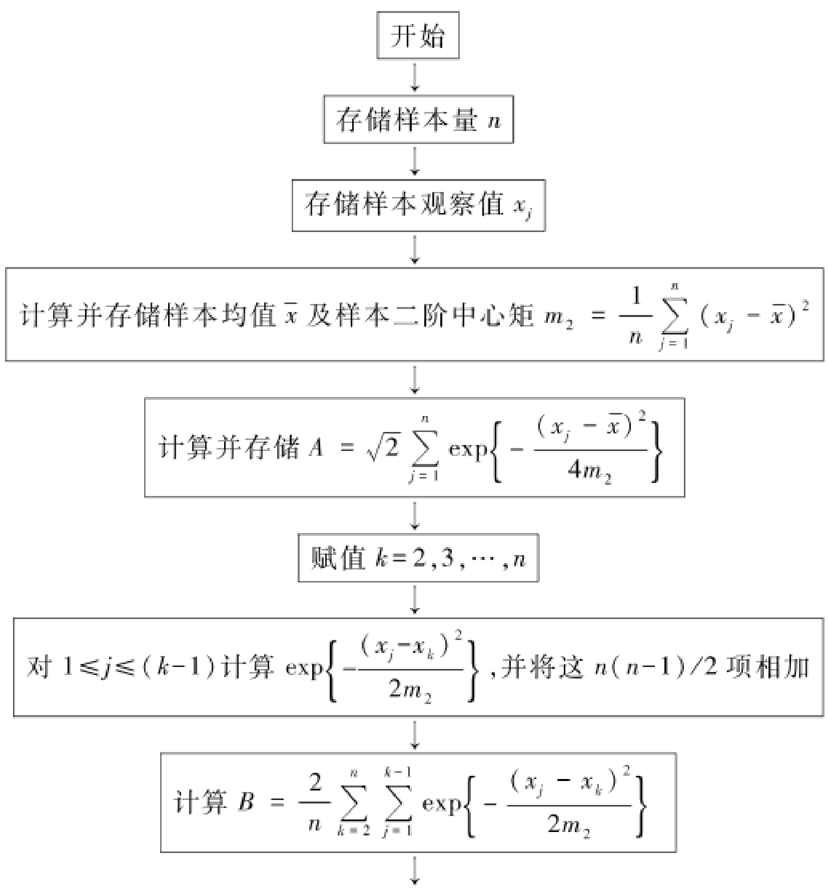
注:国家标准 GB/T4882—2001 中规定:样本量 $n\geqslant 8$,因为在 $n<8$ 时,对偏离正态分布的检验不太有效.

3. EP 检验 EP 检验统计量定义为

$$T_{\mathrm{EP}}=1+\frac{n}{\sqrt{3}}+\frac{2}{n}\sum_{i=2}^{n}\sum_{j=1}^{i-1}\exp\left\{\frac{-\left(x_j-x_i\right)^2}{2s_*^2}\right\}-\sqrt{2}\sum_{i=1}^{n}\exp\left\{\frac{-\left(x_i-\bar{x}\right)^2}{4s_*^2}\right\},$$

其中 $\bar{x}$,s_*^2 就是前述的样本均值和(除以 n 的)样本方差.该统计量通常需要编程计算,其拒绝域为 $\{T_{\mathrm{EP}}\geqslant T_{1-\alpha,\mathrm{EP}}(n)\}$,$T_{1-\alpha,\mathrm{EP}}(n)$ 可查附表 11,当 $n>200$ 时,统计量 T_{EP} 的分位数可以用 $n=200$ 时的分位数代替.对小于 200 而不在表内的 n,可采用线性插值的方法得到近似的分位数.

注意:样本观察值的次序是随机的,但一经选定后在计算 T_{EP} 中必须保持不变.计算 T_{EP} 的程序框图 7.4 如下所示(引自国标 GB/T4882—2001).



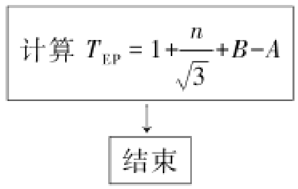


图 7.1　计算检验统计量 T_{EP} 的框图

§7.6　非参数检验

1. 游程检验

在怀疑数据可能不符合随机选取的原则时可对数据进行游程检验.

设 $x_1,x_2,\cdots,x_n$ 为依时间顺序连续得到的一组样本观测值序列.要考察如下原假设是否成立.

H_0:样本观察值序列符合随机抽取的原则.

为此记样本中位数为 m_e,把序列中小于 m_e 的那些 x_i 换成 0;大于或等于 m_e 的那些 x_i 换成 1.设序列中 0 和 1 的个数分别为 n_1 和 n_2,设 R 表示样本序列的总游程数,则检验的拒绝域为 $\{R\leqslant c_1\}\cup\{R\geqslant c_2\}$,其中临界值 c_1 和 c_2 根据 H_0 为真时 R 的分布确定,可查表.

当 n_1,n_2 都不太大时,游程检验用 p 值进行更加方便,记 R 的观测值为 R_0,则 $p=2\min\{P(R\leqslant R_0),P(R\geqslant R_0)\}$,其中概率可用下述两式进行计算.

$$P(R=2k)=\frac{2\binom{n_1-1}{k-1}\binom{n_2-1}{k-1}}{\binom{n_1+n_2}{n_1}},\quad k=1,2,\cdots,\left[\frac{n}{2}\right],$$

$$P(R=2k+1)=\frac{\binom{n_1-1}{k-1}\binom{n_2-1}{k}+\binom{n_1-1}{k}\binom{n_2-1}{k-1}}{\binom{n_1+n_2}{n_1}},\quad k=1,2,\cdots,\left[\frac{n-1}{2}\right].$$

对比较大的 n_1,n_2,可以使用渐近分布.对给定的显著性水平 α,两个临界值可近似取为

$$c_1=\left[\frac{2n_1n_2}{n_1+n_2}\left(1+\frac{u_{\alpha/2}}{\sqrt{n_1+n_2}}\right)\right],c_2=\left[\frac{2n_1n_2}{n_1+n_2}\left(1+\frac{u_{1-\alpha/2}}{\sqrt{n_1+n_2}}\right)\right]+1.$$

研究表明，当 n_1,n_2 都大于 20 时，上式近似效果是足够好的.

游程检验还可以用于检验两个总体是否有相同分布.

2. 符号检验

符号检验主要用来对总体 p 分位数 x_p 进行检验.

设 $x_1,x_2,\cdots,x_n$ 是来自总体 X 的样本，欲检验 $H_0:x_p\leqslant x_0 \quad \text{vs} \quad H_1:x_p>x_0$，其中 x_0 是某个给定常数，记

$$y_i=\begin{cases}1, & x_i>x_0,\\ 0, & x_i\leqslant x_0,\end{cases}$$

符号检验统计量为 $S^+=\sum_{i=1}^{n}y_i$，拒绝域为

$$W=\{S^+\geqslant c\},c=\inf_k\left\{k \mid \sum_{i=k}^{n}\binom{n}{i}p_0^i(1-p_0)^{n-i}\leqslant\alpha\right\}.$$

对符号检验，使用检验的 p 值较为简便.记 S_0^+ 为符号统计量的观测值，则检验的 p 值为

$$p=P(S^+\geqslant S_0^+)=\sum_{i=S_0^+}^{n}b(i;n,1-p),$$

其中 $b(i;n,1-p)=\binom{n}{i}(1-p)^ip^{n-i}$表示二项分布的概率函数.

三种假设下的符号检验：

H_0	H_1	拒绝域形式	检验的 p 值
$x_p\leqslant x_0$	$x_p>x_0$	$W_{\mathrm{I}}=\{S^+\geqslant c\}$	$\sum_{i=S_0^+}^{n}b(i;n,1-p)$
$x_p\geqslant x_0$	$x_p<x_0$	$W_{\mathrm{II}}=\{S^+\leqslant c\}$	$\sum_{i=0}^{S_0^+}b(i;n,1-p)$
$x_p=x_0$	$x_p\neq x_0$	$W_{\mathrm{III}}=\{S^+\leqslant c_1$ 或 $S^+\geqslant c_2\}$	$2\min\left\{\sum_{i=0}^{S_0^+}b(i;n,1-p),\sum_{i=S_0^+}^{n}b(i;n,1-p),0.5\right\}$

符号检验还可用于成对数据的比较.

3. 符号秩和检验

秩检验主要用于对对称分布的分布中心进行检验.

设连续总体关于某个参数 θ 对称，其分布函数记为 $F(x-\theta)$，要检验的假设为：

$$H_0: \theta = 0 \quad \text{vs} \quad H_1: \theta \neq 0.$$

设 $x_1, x_2, \cdots, x_n$ 是样本，记 R_i 为 $|x_i|$ 在$(|x_1|, |x_2|, \cdots, |x_n|)$中的秩，符号秩和统计量 $W^+ = \sum_{i=1}^{n} R_i I(x_i > 0)$，拒绝域为 $\{W^+ \leqslant W^+_{\alpha/2}(n)\} \cup \{W^+ \geqslant W^+_{1-\alpha/2}(n)\}$.

附表 13 给出了 $n \leqslant 50$ 时满足条件 $P(W^+ \leqslant W^+_\alpha(n)) \leqslant \alpha$ 的 $W^+_\alpha(n)$，至于满足条件 $P(W^+ \geqslant W^+_{1-\alpha}(n)) \leqslant \alpha$ 的 $W^+_{1-\alpha}(n)$，它等于$\frac{1}{2}n(n+1) - W^+_\alpha(n)$.

当 $n>50$ 时可采用正态近似$\dfrac{W^+ - \dfrac{n(n+1)}{4}}{\sqrt{\dfrac{n(n+1)(2n+1)}{24}}} \xrightarrow{L} N(0,1)$计算检验临界值.

4. 秩和检验

秩和检验常用来对两个总体的位置进行比较.

设有两个总体，分布函数为分别为 $F(x-\theta_1)$ 和 $F(x-\theta_2)$，$x_1, x_2, \cdots, x_m$ 和$y_1, y_2, \cdots, y_n$ 分别为其样本，对合样本 $x_1, x_2, \cdots, x_m, y_1, y_2, \cdots, y_n$ 进行排序，对应的秩为 $R = (Q_1, Q_2, \cdots, Q_m, R_1, R_2, \cdots, R_n)$. 在比较 θ_1, θ_2 的大小时，采用如下 Wilcoxon 秩和统计量 $W = \sum_{i=1}^{n} R_i$，假设与拒绝域分别为

Ⅰ $H_0: \theta_1 \leqslant \theta_2 \quad \text{vs} \quad H_1: \theta_1 > \theta_2, W_{\mathrm{I}} = \{W \leqslant W_\alpha(m,n)\}$；

Ⅱ $H_0: \theta_1 \geqslant \theta_2 \quad \text{vs} \quad H_1: \theta_1 < \theta_2, W_{\mathrm{II}} = \{W \geqslant W_{1-\alpha}(m,n)\}$；

Ⅲ $H_0: \theta_1 = \theta_2 \quad \text{vs} \quad H_1: \theta_1 \neq \theta_2, W_{\mathrm{III}} = \{W \leqslant W_{\alpha/2}(m,n)$ 或 $W \geqslant W_{1-\alpha/2}(m,n)\}$.

在样本容量较大时，可以用大样本近似公式 $W^* = \dfrac{W - \dfrac{n(m+n+1)}{2}}{\sqrt{\dfrac{mn(m+n+1)}{12}}} \sim N(0,1)$作上述检验. 在 m 与 n 都大于等于 20 时，此种近似效果很好.