第三节 非参数假设检验

一、什么是非参数假设检验

目前对非参数假设检验尚没有一个严格的定义，它泛指参数假设检验以外的各种检验。目前较普遍的有以下几种解释：

1. 非参数假设检验不依赖于总体分布。因为参数假设检验除了大样本情况下进行的参数假设检验外，其余都是假定总体服从某一分布的参数检验，如 t 检验等都依赖于总体是正态分布的。

2. 非参数假设检验适用于较低的计量水准，如等级的、顺序的计量，如中位数检验，甚至还可以应用于列名的计量。

3. 非参数假设检验还包括对参数以外的各种检验，如随机变量是否服从某种规律或某种分布的拟合优度检验，数据是否随机的游程检验等。

因此非参数假设检验的内容十分广泛，本节也只是介绍一些在经济管理中最常用的检验。有些数据可以同时使用参数检验和非参数检验，由于非参数检验只应用于顺序等计量，没有充分利用信息，因此其效率就不如参数检验。

二、χ^2检验

在抽样分布一章已经介绍过什么是χ^2分布，它在非参数统计中得到了广泛的应用。

（一）分类数据的拟合优度检验

在实际应用统计方法时，常常先将收集到的数据进行分组，形成频数分布表，这是显示数据规律性的一种方法。人们为了掌握其规律性，往往还想进一步了解这一数据是否来自某一分布或与某一理论分布相一致的程度如何。χ^2 检验就可以用于这一类问题的研究。

利用 χ^2 分布进行拟合优度检验的方法是：

第一步：先将观测到的数据分类，假设分成 m 类，每类中的频数为：$v_1,v_2,\cdots,v_m$，或简记为 $v_i(i=1,2,\cdots,m)$。

第二步：根据观测结果似乎服从某一理论分布的规律，需要进一步检验。按照理论分布，各类的频数应为 $e_i=nP_i(i=1,2,\cdots,m)$，其中 P_i 为根据理论分布，观测发生在第 i 类的概率。

第三步：计算统计量

$$\chi^2=\sum_{i=1}^{m}\frac{(v_i-e_i)^2}{e_i} \tag{6.9}$$

如果理论分布的参数是予先给定的（已知的），则 χ^2 统计量服从自由度为 $m-1$ 的 χ^2 分布。若理论分布的参数是未知的，需要用样本观测值来估计时，χ^2 统计量服从自由度为 $m-\mathrm{r}-1$ 的 χ^2 分布，其中 r 为需要估计的参数的个数。

第四步：根据显著性水平 α 查 χ^2 分布表求相应的临界值 χ^2_α。$\chi^2>\chi^2_\alpha$ 时，拒绝原假设，说明样本观测并非来自该理论分布。

【例 6.10】 某百货公司的电器部下半年各月洗衣机的销售数量如下：

月份	七月	八月	九月	十月	十一月	十二月	合计
销售量（台）	27	18	15	24	36	30	150

该电器部经理想了解洗衣机的销售数量是否在各月是均匀分布的，也就是说各月中销售数量的差别可以归结为随机原因，这样可以为以后的进货提供依据。要求以 $\alpha=0.05$ 的显著性水平进行检验。

解：本例中的观测值以月为组，共分为 $m=6$ 组，每月的销售台数即为观测的频数 v_i，观测的总次数为 $n=150$。现欲检验是否服从（离散的）均匀分布，即每月的销售量是否为 $e_i=nP_i=\frac{150}{6}=25$ 台，其中 $P_i=\frac{1}{6}$，$i=1,2,\cdots,6$。为此，设

H_0：洗衣机销售量服从均匀分布；

H_1：并不服从均匀分布；

计算 χ^2 统计量的值：

$$\chi^2 = \sum_{i=1}^{6} \frac{(v_i - e_i)^2}{e_i}$$

$$= \frac{(27-25)^2}{25} + \frac{(18-25)^2}{25} + \frac{(15-25)^2}{25} + \frac{(24-25)^2}{25}$$

$$+ \frac{(36-25)^2}{25} + \frac{(30-25)^2}{25}$$

$$= 12$$

在本例的情况下，χ^2 统计量的自由度为 $m-1=6-1=5$。查表得知，$\chi^2_{0.05}(5)=11.07$。由于 $\chi^2=12>11.07$，所以拒绝 H_0，说明下半年各月的销售量与均匀分布有差别，这些差别尚不能完全归结为随机原因。

【例 6.11】 在高速公路收费站 100 分钟内观测到通过收费站的汽车共 190 辆，每分钟通过的汽车辆数分布如下表：

每分钟通过辆数	0	1	2	3	4 或更多
分钟数	10	26	35	24	5

用显著性水平 $\alpha=0.05$ 检验这些数据是否来自泊松分布。

解：设

H_0：汽车通过收费站的辆数服从泊松分布；

H_1：不服从泊松分布。

观测值分为 5 组，且有 $v_0=10, v_1=26, v_2=35, v_3=24, v_4=5$。

回忆泊松分布为

$$P(X=x) = \frac{e^{-\lambda}\lambda^x}{x!}, x=0,1,2,\cdots$$

其中 λ 为泊松分布的期望值，是未知的，需要用样本观测值来估计。由于 100 分钟内观测到 190 辆汽车，所以平均每分钟观测到 $\frac{190}{100}=1.9$ 辆汽车，故取 $\lambda=1.9$。据此，我们可以用参数 $\lambda=1.9$ 的泊松分布来计算每分钟内通过收费站的汽车为 0 辆、1 辆、2 辆、

3 辆、4 辆或更多的概率。例如，

$$P(X=0)=\frac{e^{-1.9}(1.9)^0}{0!}=0.1496$$

$$P(X=1)=\frac{e^{-1.9}(1.9)^1}{1!}=0.2842,$$

等等。各概率乘以观测总数 $n=100$，便得到理论频数 e_i，具体结果见下表：

每分钟通过辆数	0	1	2	3	4 或更多
观测频数 v_i	10	26	35	24	5
概率 P_i	0.149 6	0.284 2	0.270 0	0.171 0	0.125 2
理论频数 e_i	14.96	28.42	27.0	17.1	12.52

计算 χ^2 统计量：

$$\chi^2=\frac{(14.96-10)^2}{14.96}+\frac{(28.42-26)^2}{28.42}+\frac{(27.0-35)^2}{27.0}$$

$$+\frac{(17.1-24)^2}{17.1}+\frac{(12.52-5)^2}{12.52}=11.52$$

在本例中，χ^2 统计量的自由度为 $m-1-1=5-1-1=3$。临界值 $\chi^2_{0.05}(3)=7.815$，所以拒绝 H_0，说明每分钟通过收费站的汽车辆数不服从泊松分布。

在应用 χ^2 分布拟合优度检验时，应注意每一类中理论频数不宜过小，通常应不小于 5。如果出现理论频数太低，就应当与邻近的类进行合并。

（二）χ^2 分布的独立性检验

拟合优度检验是根据样本观察值与一个理论值进行比较来检验的，但是有些数值并不知道服从何种理论分布。因此在双变量的分布中，有时想了解两个变量是相依的还是独立的，χ^2 检验也可用于这样的检验，称作 χ^2 的独立性检验。这种检验建立在观察数据的交叉分类基础上，这种交叉分类的表称作列联表，如表

6.2 所示。

表 6.2　2×2 列联表

		j		合计
		$j=1$	$j=2$	
i	$i=1$	f_{11}	f_{12}	r_1
	$i=2$	f_{21}	f_{22}	r_2
合　计		c_1	c_2	n

表 6.2 是最简单的 2 行 2 列的列联表，它可以扩展到 $r\times c$ 列联表，即 r 行 c 列的列联表。也就是说把所有的观察结果的频数分布在各格之中，f_{11} 就代表行的第 1 类和列的第 1 类所出现的实际频数，依次类推。那么相应于 f_{11} 的理论值应如何计算呢？因为 f_{11} 位于第 1 行，整个样本量为 n，落入第 1 行的概率根据样本估计应该是$\frac{r_1}{n}$，f_{11} 又同时位于第 1 列，落入第 1 列的概率根据样本计算应该是$\frac{c_1}{n}$，根据概率论的原理，如果行和列的变量是独立的，那么落入第 1 行和第 1 列的概率应该是$\left(\frac{r_1}{n}\right)\left(\frac{c_1}{n}\right)$，由于样本量为 n，则落入第 1 行第 1 列的理论频数应该是：

$$e_{11} = n\left(\frac{r_1}{n}\right)\left(\frac{c_1}{n}\right) = \frac{r_1 c_1}{n}$$

由此可以推广到

$$e_{ij} = n\left(\frac{r_i}{n}\right)\left(\frac{c_j}{n}\right) = \frac{r_i c_j}{n}.$$

在独立性检验中的 χ^2 统计量为

$$\chi^2 = \sum_{i=1}^{r}\sum_{j=1}^{c}\frac{(f_{ij}-e_{ij})^2}{e_{ij}}$$

其自由度为 $(r-1)(c-1)$。

【例 6.12】　某副食品商店欲研究顾客的性别与购物金额大小之间是有关系，还是没有关系（意味着相互独立）。在该商店内

随机调查了 548 位顾客，按金额大小和性别进行分类，取得如下数据(见表 6.3)：

表 6.3　顾客的性别与购买金额列联表(括号内是理论频数 e_{ij})

金额 / 性别	10 元以下	10—50 元	50 元以上	合　计
男性	40(50.29)	90.(99.64)	130(110.07)	260
女性	66(55.71)	120(110.36)	102(121.93)	288
合计	106	210	232	548

要求用 $\alpha=0.05$ 的显著性水平检验顾客的性别和购买金额是否独立。

解：

H_0：购物的金额大小与性别无关(独立)；

H_1：购物的金额大小与性别有关。

计算列联表各格的理论值：

$$e_{ij}=\frac{r_i c_j}{n}$$

$$e_{11}=\frac{(106)(260)}{548}=50.29 \quad e_{12}=\frac{(210)(260)}{548}=99.64$$

$$e_{13}=\frac{(232)(260)}{548}=110.07 \quad e_{21}=\frac{(106)(288)}{548}=55.71$$

$$e_{22}=\frac{(210)(288)}{548}=110.36 \quad e_{23}=\frac{(232)(288)}{548}=121.93$$

并列入列联表各格的括号内。计算 χ^2 值

$$\chi^2=\frac{(40-50.29)^2}{50.29}+\frac{(90-99.64)^2}{99.64}+\frac{(130-110.07)^2}{110.07}$$

$$+\frac{(66-55.71)^2}{55.71}+\frac{(120-110.36)^2}{110.36}+\frac{(102-121.93)^2}{121.93}$$

$$=2.105+0.933+3.609+1.901+0.842+3.258$$

$$=12.648$$

2×3 列联表的自由度为 $(r-1)(c-1)=2$，当 $\alpha=0.05$ 时的

$\chi_{\alpha=0.05}(2)=5.991$

$\chi^2>\chi_{\alpha=0.05}(2)$拒绝 H_0,接受 H_1,即购物的金额大小与性别有关。

2×2 列联表的 χ^2 值计算还可以简化,为了说明方便,将列联表每格的数字用字母表示(见表 6.4)。

表 6.4　2×2 列联表用字母表示

行 \ 列		1	2	合　计
行	1	a	b	$a+b$
	2	c	d	$c+d$
合　计		$a+c$	$b+d$	n

$$\chi^2=\frac{n(ad-bc)^2}{(a+c)(b+d)(a+b)(c+d)} \tag{6.10}$$

【例 6.13】　某市场调研机构,调查某种光盘的购买者和性别之间是否有关系取得如下数据:

	想买	不想买	合计
男	32(26)	118(124)	150
女	20(26)	130(124)	150
合计	52	248	300

令 $\alpha=0.05$,用 χ^2 独立性检验推断购买某种光盘与性别是否有关?

解:H_0:购买与性别无关,H_1:购买与性别有关。现采用两种方法计算 χ^2 值。

$$\chi^2=\sum\frac{(f-e)^2}{e}=\frac{(32-26)^2}{26}+\frac{(118-124)^2}{124}$$

$$+\frac{(20-26)^2}{26}+\frac{(130-124)^2}{124}$$

$$=1.3846+0.2903+1.3846+0.2903$$

$$=3.3498$$

$$\chi^2 = \frac{n(ad - bc)^2}{(a+c)(b+d)(a+b)(c+d)}$$

$$= \frac{300[(32)(130)-(118)(20)]^2}{(52)(248)(150)(150)}$$

$$= 3.349\ 8$$

两种方法的计算结果相同，当 $\alpha = 0.05$ $\chi^2_{\alpha=0.05}(1) = 3.841$，$\chi^2 <$ 3.841，接受 H_0，说明买该光盘与性别的关系不显著。

三、秩和检验

秩和检验又称等级和检验。因为参数中的均值检验在小样本时通常要求总体服从正态分布，当总体不符合正态分布时可以把它转换成等级然后检验。这一类检验统称为秩和检验。

（一）曼－惠特尼 U 检验

曼－惠特尼 U 检验是由 H.B.曼和D.R.惠特尼于 1947 年提出来的。它假设两个样本分别来自两个总体，目的是检验这两个总体的均值是否有显著的差别。它的具体做法如下：

第一步：先把两组数据混合在一起，按照大小顺序编排等级。如最小的数据等级为 1，第二小的数据等级为 2，依次下去。若有两个数据相等，且它们在按大小顺序编排好的数列里是第 m 和第 $m+1$ 个数据，则它们的等级（也称作秩）都是 $\frac{m+(m+1)}{2} = \frac{2m+1}{2}$。同理，若有 3 个数据相等，且它们在按大小顺序编排好的数列里是第 m，第 $m+1$ 和第 $m+2$ 位数据，则它们的等级都是 $\frac{3m+3}{3} = m+1$。

假设第一个样本的样本量为 n_1，第二个样本的样本量为 n_2，则共有 $n_1 + n_2 = n$ 个观测值，因此最大的数据的等级应该是 n（若最大的数据只有一个的话）。

第二步：分别求两个样本的等级和。设第一个样本的等级和

为 W_1，第二个样本的等级和为 W_2，则有 $W_1+W_2=\frac{n(n+1)}{2}$。

第三步：计算曼－惠特尼 U 检验统计量

$$U_1=n_1n_2+\frac{n_1(n_1+1)}{2}-W_1$$

$$U_2=n_1n_2+\frac{n_2(n_2+1)}{2}-W_2$$

从 U_1 和 U_2 中选择较小者并称其为 U。

第四步：作出判断。

对于 n_1 和 n_2 都比较小的情形，可以查附表 6 得到临界值 U_α，在 $U<U_\alpha$ 时拒绝 $H_0:\mu_1=\mu_2$。

在原假设为真的情况下，可以证明随机变量 U 的均值和方差分别为

$$E(U)=\frac{n_1n_2}{2},$$

$$D(U)=\frac{n_1n_2(n_1+n_2+1)}{12}$$

并且当 n_1 和 n_2 都不小于 10 时，随机变量

$$Z=\frac{U-E(U)}{\sqrt{D(U)}} \tag{6.11}$$

近似地服从标准正态分布。

设第一个总体的均值为 μ_1，第二个总体的均值为 μ_2，则对于

$$H_0:\mu_1\leqslant\mu_2 \quad H_1:\mu_1>\mu_2$$

如果 $Z<-z_\alpha$，则拒绝 H_0；

对于

$$H_0:\mu_1\geqslant\mu_2 \quad H_1:\mu_1<\mu_2$$

如果 $Z>z_\alpha$，则拒绝 H_0；

对于

$$H_0:\mu_1=\mu_2 \quad H_1:\mu_1\neq\mu_2$$

如果 $|Z|>z_{\alpha/2}$，则拒绝 H_0。

【例 6.14】 为了比较两个小学贯彻素质教育的情况，现从甲

学校抽 15 名学生,乙学校抽 25 名学生,按素质教育的要求进行测试并评分,按评分高低顺序排队并编上等级,其结果如下:

甲学校			乙学校				
3	17	30	1	8	16	23	33
6	22	32	2	9	18	27	35
10	24	34	4	11	19	28	36
12	25	38	5	13	20	29	37
14	26	40	7	15	21	31	39

$n_1 = 15, W_1 = 333$　　　　$n_2 = 25,\ W_2 = 487$

要求以显著性水平 $\alpha = 0.05$ 检验两学校的素质教育有没有差别。

解:我们假设两个学校的素质教育除了平均水平以外在其他方面没有差异。我们需要检验

H_0:两校素质教育水平无差异,

H_1:两校素质教育水平有差异。

计算 U 值:

$$U = 15 \times 25 + \frac{(15)(16)}{2} - 333 = 162$$

U 的均值和标准差分别为

$$E(U) = \frac{n_1 n_2}{2} = \frac{(15)(25)}{2} = 187.5$$

$$\sqrt{D(U)} = \sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}} = 35.8$$

因此,$Z = \dfrac{U - E(U)}{\sqrt{D(U)}} = \dfrac{162 - 187.5}{35.8} = -0.71$。

因为 $|Z| = 0.71 < z_{\frac{0.05}{2}} = 1.96$,所以我们不能拒绝 H_0,说明两个学校素质教育的水平没有显著性的差异。

(二) 威尔科克森带符号的秩检验

威尔科克森带符号的秩检验是由 F. 威尔科克森于 1945 年提

出来的。它是在成对观测数据的符号检验基础上发展起来的，比传统的单独用正负号的检验更加有效。它适用于前面介绍的 t 检验中的成对比较，但 t 检验要求比较的两个总体必须是正态分布，从而成对数据之差 d_i 也服从正态分布。而威尔科克森带符号的秩检验则没有这种限制，它只要求数据之差所服从的分布是对称分布。它的目的是检验成对观测的数据之差是否来自均值为 0 的总体（或产生数据的两个总体是否具有相同的均值）。

威尔科克森带符号的秩检验的具体做法是：

第一步：求出成对观测数据的差 d_i，并将 d_i 的绝对值 $|d_i|$ 按大小顺序编上等级，最小的为 1，其次为 2 等等。当两个绝对值相等时就用相应等级的平均数代替（和曼－惠特尼 U 检验一样）。

第二步：编码等级后再恢复其正负号，并将正号的等级与负号的等级分别相加，用 T_+ 代表正号的等级之和，T_- 代表负号的等级之和。威尔科克森检验统计量 T 为 T_+ 和 T_- 之中较小的一个。

第三步：作出判断。

对于小样本，根据显著性水平 α 查数表 7，得到临界值 T_α。若 $T < T_\alpha$，则拒绝 H_0；

对于大样本（观测不少于 20 对），可以证明统计量 T 的均值和方差分别为

$$E(T) = \frac{n(n+1)}{4}$$

$$D(T) = \frac{n(n+1)(2n+1)}{24}$$

其中 n 为成对观测的个数，并且

$$Z = \frac{T - E(T)}{\sqrt{D(T)}}$$

近似地服从标准正态分布。因此，对于单侧检验，若 $Z < -z_\alpha$，则拒绝 H_0；对于双侧检验，若 $Z < -z_{\frac{\alpha}{2}}$，则拒绝 H_0。

【例 6.15】 为比较两种轮胎的平均使用里程，在 6 辆汽车的

后轮分别用两种不同牌号的轮胎，直到用坏后加以记录里程，取得的数据如表6.5。

表6.5　两种轮胎在同一车上试验数据

汽车	轮胎 A（英里）	轮胎 B（英里）	d_i	等级
1	20 000	19 000	1 000	1
2	24 600	23 000	1 600	2
3	32 500	37 000	−4 500	−3
4	36 000	30 100	6 500	4
5	37 000	25 500	11 700	5
6	23 000	39 500	−16 500	−6
			平均	$T_-=9$　$T_+=12$

要求 $\alpha=0.05$，检验两种轮胎的平均行驶里程是否有显著差别。

解：H_0：两种轮胎的平均使用里程无显著差别，

　　H_1：两种轮胎的平均使用里程有差别。

将成对的差列于表6.5的第4列（d_i），根据 d_i 的绝对值由小到大顺序编号，然后恢复正负号，再将不同符号的等级分别相加，见表6.5最后一列。计算得到正负号的等级和，$T_-=9$，$T_+=12$，用较小的 T 与临界值 T_α 相比较，由附表7得到：对于 $\alpha=0.05$ 的双侧检验，$n=6$ 时 $T_{0.025}=1$，T 已超过临界值，因此不能推翻 H_0，可认为两种轮胎的行驶里程无显著差别。

【例6.16】　某饮料商用两种不同的配料方法推出了两种新的饮料，现抽取了20个消费者，让其分别品尝两种饮料并加以评分，从不喜欢到喜欢，评分由1～10，其评分结果如下：

品尝者	1	2	3	4	5	6	7	8	9	10
A饮料	10	8	6	8	7	5	1	3	9	7
B饮料	6	5	2	2	4	6	4	5	9	8

品尝者	11	12	13	14	15	16	17	18	19	20
A饮料	4	5	8	6	8	7	4	8	9	3
B饮料	2	2	1	3	2	6	1	2	5	3

要求以 $\alpha=0.05$ 的显著性水平检验对两种饮料的评分是否有显著差别。

解:应用威尔科克森带符号的成对检验,将评分之差变换为等级,再恢复正负号,其计算过程见计算表6.6。将评分相同的样本加以剔除,因此样本量就由20变为18。

表6.6 例6.16的计算表

品尝者	A饮料评分	B饮料评分	d_i (A_i-B_i)	转换等级	正等级	负等级
1	10	6	4	13	13	
2	8	5	3	8.5	8.5	
3	6	2	4	13	13	
4	8	2	6	16	16	
5	7	4	3	8.5	8.5	
6	5	6	−1	2		2
7	1	4	−3	8.5		8.5
8	3	5	−2	4.5		4.5
9	9	9	0	——		
10	7	8	−1	2		2
11	4	2	2	4.5	4.5	
12	5	2	3	8.5	8.5	
13	8	1	7	18	18	
14	6	3	3	8.5	8.5	
15	8	2	6	16	16	
16	7	6	1	2	2	
17	4	1	3	8.5	8.5	
18	8	2	6	16	16	
19	9	5	4	13	13	
20	3	3	0	——		

$T_+=154$　$T_-=17$

最后得到 $T_+ = 154$, $T_- = 17$,取其中较小的 $T_- = 17$ 来检验,在大样本的情况下 T 近似正态分布

$$E(T) = \frac{n(n+1)}{4} = \frac{(18)(19)}{4} = 85.5$$

$$V(T) = \frac{n(n+1)(2n+1)}{24} = \frac{(18)(19)(37)}{24} = 527.25$$

$$Z = \frac{T - E(T)}{\sqrt{V(T)}} = \frac{17 - 85.5}{\sqrt{527.25}} = -2.98$$

当 $\alpha = 0.05$ 时双侧检验 $z_{\alpha/2} = \pm 1.96$, $Z < -1.96$,因此拒绝 H_0 接受 H_1,说明两种饮料的评分有差别。

四、等级相关系数及其检验

(一)测定两组等级变量之间的相关系数

测定两组变量之间是否存在相关以及相关程度的常用测量是相关系数,这将在本书的第七章作介绍,但它主要是测量数值之间的相关。在对相关系数进行检验时,又涉及到对变量服从正态分布等假设。但有些现象是难以用数字确切计量的,如才智高低、色泽深浅、艺术水平等,怎样才能测量它们之间是否有关系及其密切程度呢?也要用到本节的非参数方法。斯皮尔曼提出了一种建立在等级基础上的相关系数,通常就称作斯皮尔曼等级相关系数。

在两组数据分别没有重复观测值的情况下,斯皮尔曼等级相关系数的公式为:

$$r_{SP} = 1 - \frac{6\Sigma d_i^2}{n(n^2 - 1)} \tag{6.12}$$

其中 d_i 表示两组数据的等级之差,n 为样本量。若数据中没有相同观测值,则斯皮尔曼等级相关系数就是两组数据的等级的相关系数,其计算公式参见本书第七章。

【例 6.17】 有一家公司招聘打字员,采用口试与实际操作两种考核方式。现有 6 个申请人的口试与实际操作的评分记录如下:

申　请　人	1	2	3	4	5	6
口试成绩	B^-	A^-	B^+	A	C	B
操作成绩	28	38	47	56	29	32

要求测定这些申请人实际操作成绩与口试成绩之间是否存在关系及关系的密切程度。

解：首先需要将口试成绩与操作成绩变换为等级，然后计算等级相关系数，见计算表 6.7。

表 6.7　例 6.17 的计算表

申请人	口试成绩	操作成绩	转换成等级			Σd_i^2
			口试	操作	d_i	
1	B^-	28	5	6	−1	1
2	A^-	38	2	3	−1	1
3	B^+	47	3	2	+1	1
4	A	56	1	1	0	0
5	C	29	6	5	+1	1
6	B	32	4	4	0	0
						4

应用斯皮尔曼等级相关系数

$$r_{SP}=1-\frac{6\Sigma d_i^2}{n(n^2-1)}=1-\frac{6(4)}{6(36^2-1)}=1-\frac{24}{210}=0.885\ 7$$

说明口试与操作成绩之间存在着相关，其相关程度为 0.885 7。和一般的相关系数一样，当斯皮尔曼等级相关系数等于 1 时表示完全正相关，当为 0 时为完全不相关，当等于 −1 时为完全的负相关。

（二）等级相关系数的检验

和其他的推断一样，当以样本的数据来说明总体时，由于样本

带有随机性,在小样本时有时看来相关,但总体之间则不一定相关。因此也同样有一个假设检验的问题:

H_0:研究的总体之间无相关(即 $\rho=0$),

H_1:研究的总体之间有相关(即 $\rho\neq0$)。

检验的样本估计量为样本的相关系数 r_{SP},在小样本的情况下通常临界值的 r 已编制成表可以直接查阅(见附表8),但在大样本的情况下可以通过变换

$$t = r_{SP}\sqrt{\frac{n-2}{1-r_{SP}^2}} \tag{6.13}$$

服从 $t(n-2)$的 t 分布,采用 t 检验。

【例6.18】 某公司在招聘雇员时需要通过口试和笔试两种形式,由于口试要单独进行,比较费时费力,如果二者相关程度很高,就可以采用笔试一种形式,如果总体不存在相关,则仍需采用两种考试方式。因此首先用 $n=62$ 的大样本进行口试和笔试两种形式,并计算等级相关系数 $r_{SP}=0.2$,要求用 $\alpha=0.05$ 的显著性水平确定总体是否相关。

解: $H_0:\rho=0$ $H_1:\rho\neq0$ (单侧检验)

利用公式(6.13) $t=0.2\sqrt{\dfrac{62-2}{1-0.2^2}}=1.58$

t 分布的自由度60已接近正态分布 $t_{0.05}=1.671$,$t<t_\alpha$ 落入接受域,接受 H_0,不能肯定存在相关。

习 题

6.1 H_0:该批药物合格。

要求(1) 写出备择假设;

(2) 说明第一类错误是什么;

(3) 说明第二类错误是什么。

6.2 某生产线在正常生产条件下有2%的产品是次品,一个检验员每小时抽取5个产品作检验,规则是发现一个次品就拒绝。

(1) 建立原假设与备择假设;

(2) 什么是检验统计量;

(3) 什么是接受域和拒绝域,并分别计算接受和拒绝的概率。

6.3 某厂生产的肥皂厚度服从正态分布,平均为 1.27 厘米,现抽取 10 块作样本,平均为 1.3 厘米,标准差为 0.1,要求 $\alpha=0.05$ 的显著性水平检验生产中厚度是正常的假设。

6.4 两种不同型号电缆的拉力强度检验,各抽 $n_1=n_2=100$ 小段做试验,取得如下结果

型号Ⅰ	型号Ⅱ
$\bar{x}_1=1\ 925$	$\bar{x}_2=1\ 905$
$S_1=40$	$S_2=30$

要求检验型号Ⅰ的强度是否显著高于型号Ⅱ($\alpha=0.05$)。

6.5 为调整产业结构需将某工厂并入另一厂,该厂领导认为有 60% 以上的人会赞同,现需征求职工意见加以证实,随机抽取 200 名职工调查,其中有 110 人表示赞成合并,用 $\alpha=0.05$ 检验原假设 $H_0: p\geqslant 60\%$。若已知该厂共有 1 000 人,其结论有何变化?

6.6 某调查机构研究某种商品在城市和乡村受欢迎的比例是否有差别,在城市和乡村各抽 100 人作样本,在城市有 32 人表示喜欢,在农村有 24 人表示喜欢,能否证明城市中喜欢该商品的比例高于农村($\alpha=0.05$)?

6.7 某药厂生产减肥药,想了解该减肥药对不同的年龄是否有差别,于是抽取 40 岁以上的样本 12 人,经试验平均减肥 4.04 公斤,标准差 1.86 公斤;40 岁以下的样本 12 人,平均减肥 5.13 公斤,标准差 1.72 公斤。用 $\alpha=0.05$ 及 $\alpha=0.01$ 检验 40 岁上下不同年龄的减肥效果是否有差别。

6.8 某一城市要通过一项地方法规,拟先进行民意测验。在下属的四个区各抽 100 人,调查结果如下:

	一区	二区	三区	四区
同意	38	43	78	53
不同意	62	57	22	47

从调查结果分别用 $\alpha=0.1$, $\alpha=0.05$, $\alpha=0.01$ 显著性检验四个区同意的比例是否一致。

6.9 为了研究豌豆经杂交后的实际结果与遗传学的理论是否一致，根据理论上四种类型的豌豆和实际的结果的如下数据：

类型	理论值	实际值
光滑绿色	321	303
皱皮绿色	105	109
光滑黄色	105	98
皱皮黄色	36	57

分别以 $\alpha=0.01$ 和 $\alpha=0.05$ 检验实际结果与理论是否一致。

6.10 某化工厂对铁管涂上两种不同涂料后的耐腐蚀程度作比较，各有6个样本，其使用时间如下：

涂料 A	42	37	61	74	55	57
涂料 B	39	43	43	52	52	59

用 $\alpha=0.05$，使用曼-惠特尼 U 检验检验两种涂料的耐腐蚀程度有否显著差别。

6.11 某工厂欲检验装配线上男工和女工的机械技能有何差别，抽选了9个男工和5个女工进行测试评分，取得数据如下表，用 $\alpha=0.05$ 检验男工和女工有否显著差别。

男工	1 400	1 500	570	700	1 000	700	1 220	1 050	500
女工	1 300	1 100	680	1 250	790				

6.12 随机抽选了10家公司，比较1994和1995两年科技投入占GNP的百分比。

1994	11	24	19	20	22	25	17	23	13	16
1995	18	26	17	22	19	31	24	20	14	20

用 $\alpha=0.05$ 检验1995年的比例有否增加。