

# 第13章 非参数检验

---

# 第13章 非参数检验——目录

---

- 利用二项分布对两类数据以及中位数进行推断
  - 第1节 符号检验
- 利用秩-顺序数据对两个或多个总体进行推断
  - 第2节 威尔科克森符号秩检验
  - 第3节 威尔科克森秩和检验
  - 第4节 克鲁斯卡尔-沃利斯检验
- 利用秩-顺序数据对两个变量计算等级相关系数
  - 第5节 斯皮尔曼秩相关系数

# 13.1 符号检验

---

**符号检验：**用于检验两个相关样本的观测值之差的正负号频次是否存在显著差异

- 多用途的非参数方法
- 对总体分布没有假定的要求
- 既能分析分类型数据，也能分析数量型数据

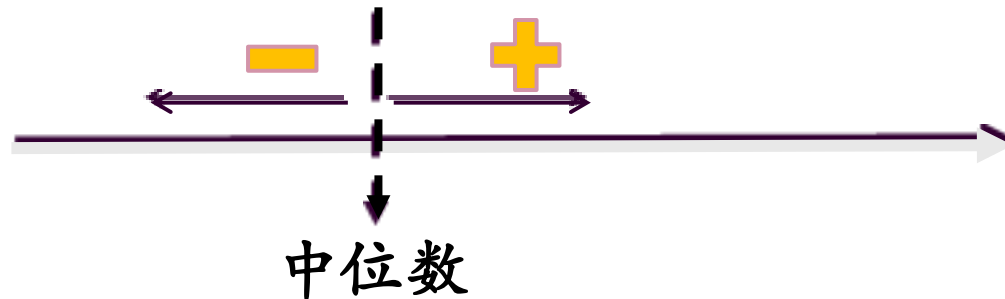
**符号检验应用：**

- 总体中位数假设检验（独立样本）
- 匹配样本的假设检验

# 13.1 符号检验

- 为了进行符号检验，从总体中随机抽取样本。将抽取的随机样本中大于中位数的值标记为加号，小于中位数的值标记为减号。如果一个值等于中位数，则在进一步分析中将其删除。
- 令 $p$ 表示加号的概率。若中位数实际值和假设值相等的假设成立，那么样本数据中每个观测值大于或小于假定值的概率都是0.5。

抽样分布： $p=0.5$ 的二项分布：



# 13.1 符号检验——总体中位数假设检验

**总体中位数假设检验：**考虑一个总体，其中没有任何一个数值恰好等于中位数，则中位数是中心趋势的度量，它将总体分为两个部分，其中50%的值大于中位数，另外50%的值小于中位数。

- 当总体是偏斜分布时，中位数是总体中心位置的最佳度量。
- 总体中位数假设检验的原假设和备择假设的一般形式如下：

$H_0$ : 总体中位数的实际值与假定值一致

$H_1$ : 总体中位数的实际值与假定值不一致

- 总体中位数的检验统计量如下：

$S^+$  : 观测值与假定的中位数的差值为正的个数

## 13.1 符号检验——总体中位数假设检验

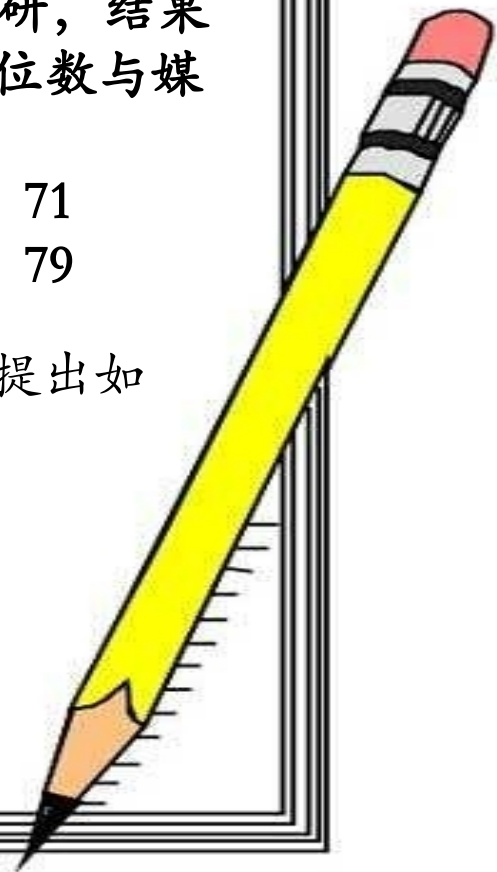
**例：**某地产商想要确认城市平均楼盘价格的中位数与媒体公布的7600元/m<sup>2</sup>是否有显著差异，故对某城市16座预出售的楼盘均价进行调研，结果如下所示（单位：百元/m<sup>2</sup>）。试检验该城市平均楼盘价格的中位数与媒体公布的7600元/m<sup>2</sup>是否有显著差异。

72	73	76	77	87	90	70	71
68	80	83	76	82	89	78	79

**解：**设该城市平均楼盘价格假定的中位数为7600元/m<sup>2</sup>，依题意提出如下假设：

$H_0$ ：该城市平均楼盘价格的中位数等于7600元/m<sup>2</sup>

$H_1$ ：该城市平均楼盘价格的中位数不等于7600元/m<sup>2</sup>



# 13.1 符号检验——总体中位数假设检验

SPSS输出检验结果如表13-1所示：

表13-1 楼盘均价的中位数符号检验

二项式检验

		类别	数字	观测到的比例	检验比例	精确显著性水平（双尾）
楼盘价格	组 1	<= 7600	7	.44	.50	.804
	组 2	> 7600	9	.56		
	总计		16	1.00		

从表13-1中可以看出楼盘价格小于等于中位数的有7个，大于中位数的有9个。SPSS给出的精确双尾概率为 $P = 0.804 > \alpha = 0.05$ ，所以不能拒绝原假设。也就是说该城市楼盘价格的实际中位数与7600元/m<sup>2</sup>有显著差异，媒体公布的数据不准确。



## 13.1 符号检验——匹配样本的假设检验

**匹配样本：**可以是同一研究对象分别给于两种不同处理的效果比较的观察值；或者同一研究对象处理前后的效果比较的观察值。

- 匹配样本的符号检验要求两个样本的观测值之差服从对称分布
- 匹配样本假设检验的原假设和备择假设的一般形式如下：

$H_0$ : 两个总体的分布相同

$H_1$ : 两个总体的分布不相同

- 匹配样本的检验统计量如下：

$N^+$  : 两个匹配样本数据对的差值为正的个数



## 13.1 符号检验——匹配样本的假设检验

**例：**为了研究放松（如听音乐）对于入睡时间的影响，随机选择了11名志愿者，分别记录他们未进行放松时的入睡时间及放松后的入睡时间，数据如下表13-2。检验该放松方法对睡眠时间的影响是否有显著差异。 $(\alpha = 0.05)$

表13-2 经历放松前后两种方式的入睡时常（单位：min）

编号	1	2	3	4	5	6	7	8	9	10	11
放松前	23	15	17	18	19	30	22	14	13	28	21
放松后	18	10	17	14	15	24	20	18	7	22	18

**解：**这里有两个匹配总体。为了检验放松对入睡时长的影响是否存在显著差异，提出如下假设：

$H_0$ ：两个总体分布相同

$H_1$ ：两个总体分布不同

# 13.1 符号检验——匹配样本的假设检验

SPSS输出检验结果如表13-3和表13-4所示：

表13-3频率

		数字
放松前的入睡时长 - 放松后的入睡时长	负差分 <sup>a</sup>	1
	正差分 <sup>b</sup>	9
	结 <sup>c</sup>	1
	总计	11

- a. 放松前的入睡时长 < 放松后的入睡时长
- b. 放松前的入睡时长 > 放松后的入睡时长
- c. 放松前的入睡时长 = 放松后的入睡时长

从表13-3中，可以看出有9个志愿者放松前的入睡时长高于放松后的入睡时长，有1个志愿者放松后的入睡时长高于放松前的入睡时长，有1个志愿者的入睡时长在放松前后没有变化，样本容量为10。



# 13.1 符号检验——匹配样本的假设检验

SPSS输出检验结果如表13-3和表13-4所示：

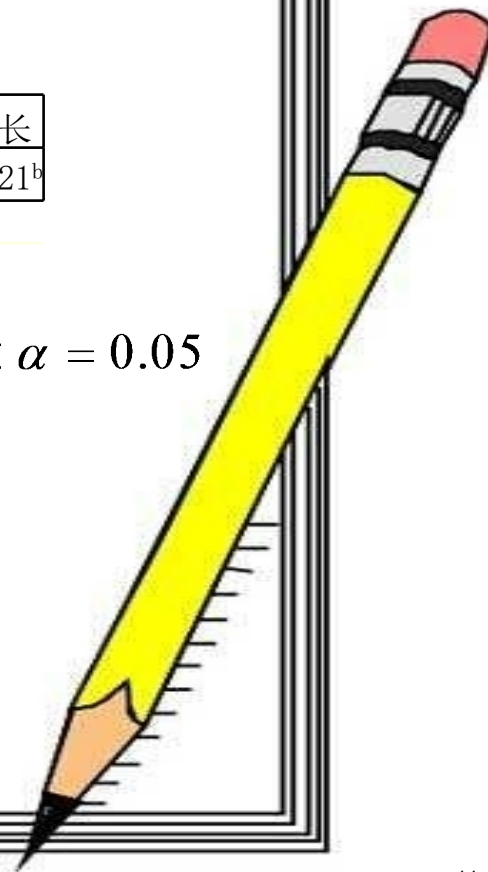
表13-4 检验统计<sup>a</sup>

	放松前的入睡时长 - 放松后的入睡时长
精确显著性水平（双尾）	.021 <sup>b</sup>

a. 符号检验

b. 使用了二项分布。

从表13-4中可以看出，SPSS给出的精确双尾概率为  $P = 0.021 < \alpha = 0.05$  所以拒绝原假设，认为放松对入睡时长有显著差异。



## 13.2 威尔科克森符号秩检验

**威尔科克森符号秩检验：**检验两个匹配总体的位置（中位数）是否存在显著差异

- 匹配样本观测值之差具有对称分布
- 可以同时考虑样本差异大小和差异方向上的信息。

秩 (rank)：一组数据按照从小到大顺序排列以后，每个观测值所在的位置。

威尔科克森符号秩检验的原假设和备择假设的一般形式如下：

$H_0$ : 两个总体的分布相同

$H_1$ : 两个总体的分布不相同

匹配样本的检验统计量计算步骤如下：

- 设  $X, Y$  是具有对称分布的两个匹配连续总体。从两个总体中分别抽取两个独立的随机样本分别为  $(x_1, x_2, \dots, x_n), (y_1, y_2, \dots, y_n)$ ；
- 组成数据对  $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ ；
- 计算  $\Delta_i = x_i - y_i, i = 1, 2, \dots, n$ ；

以上内容仅为本文档的试下载部分，为可阅读页数的一半内容。如要下载或阅读全文，请访问：  
<https://d.book118.com/025010022114011132>