非参数检验

数据分析

单样本非参数检验、两独立样本非参数检验、多独立样本非参数检验、两配对样本非参数检验的的实验及结果分析。

李金哲

2020/4/21

非参数检验

【数据分析 】

一、 单样本非参数检验

1、总体分布卡方检验

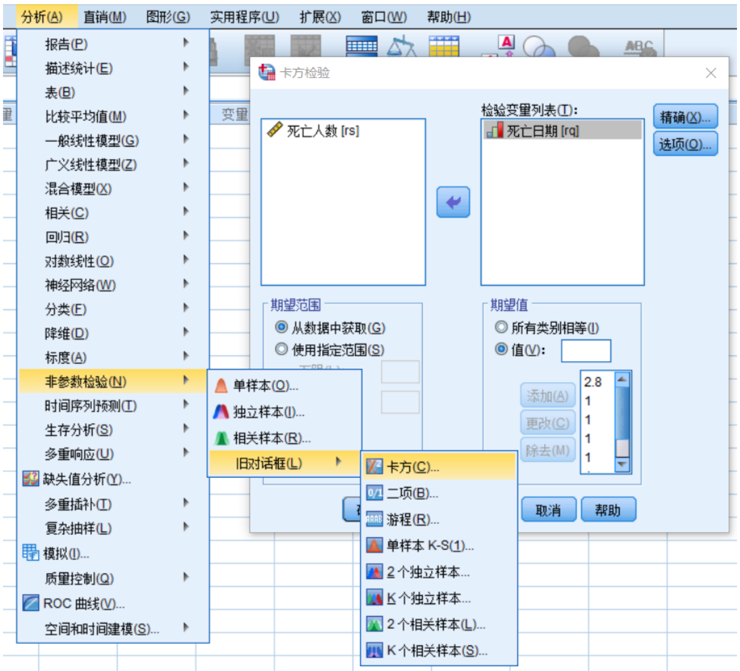
DATA：心脏病猝死.SAV

【分析过程】

SPSS 总体分布的卡方检验的基本操作步骤如下：

(1)选择菜单：【分析（A)】一【非参数检验（N)】一【旧对话框（L)】一【卡方（C)】

(2)选定待检验的变量到【检验变量列表（T)】框中。



【检验结果】

死亡日期	实测个案数	期望个案数	残差
1.00	55	53.5	1.5
2.00	23	19.1	3.9
3.00	18	19.1	-1.1
4.00	11	19.1	-8.1
5.00	26	19.1	6.9
6.00	20	19.1	.9
7.00	15	19.1	-4.1
总计	168		

【结果分析】

通过上表，共有 168 个观测数据。其中，观察到的一周内每天猝死人数分别为 55, 23, 18, 11, 26, 20, 15 人；但如果按照理论（周一到周日心脏病猝死人数比：2.8:1:1:1:1:1:1）计算，得到的预测值分别为：53.3, 19.1, 19.1, 19.1, 19.1, 19.1, 19.1。

观测值与预测值的差分别为：1.5, 3.9, -1.1, -8.1, 6.9, 0.9, -4.1。

检验统计

	死亡日期
卡方	7.757[a]
自由度	6
渐近显著性	.256

a. 0 个单元格 (0.0%) 的期望频率低于 5。期望的最低单元格频率为 19.1。

卡方统计量为 7.757，而对应的概率 P 值为 0.256

P>α，接受原假设，认为实际分布与理论分布无明显差异

即心脏病猝死人数与日期的关系基本为 2.8:1:1:1:1:1:1 的分布。

2、二项分布检验

DATA：产品合格率.SAV

【检验过程】

选择菜单：【分析（A)】一【非参数检验（N)】一【旧对话框（L)】一【二项（B)】

【检验结果】

二项检验

		类别	个案数	实测比例	检验比例	精确显著性（单尾）
是否合格	组 1	合格	19	.8	.9	.193[a]
	组 2	不合格	4	.2		
	总计		23	1.0		

a. 备用假设指出第一个组中的个案比例 < .9。

【结果分析】

我们从上图可以得出 23 个样品中合格品为 19 个，不合格品为 4 个，合格品样本的实际比例 0. 8。检验合格品率是否显著低于 0. 9

由于是小样本，SPSS 自动计算精确概率，如果合格率为 0. 9, 那么 23 个样品中合格品个数小于等于 19 的概率为 0. 193

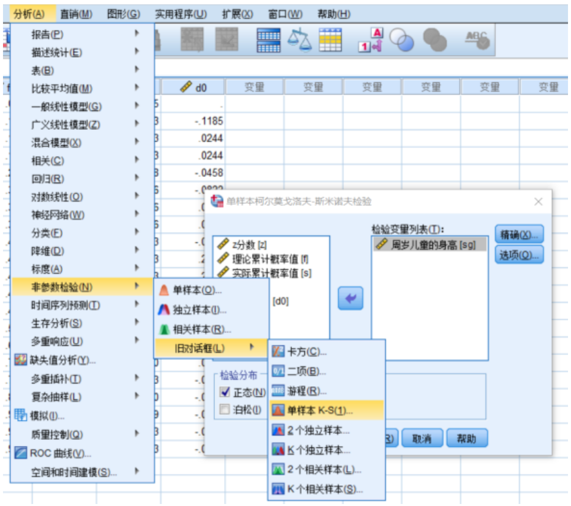
如果显著性水平 a 为 0. 05, 由于概率值大于显著性水平 a, 因此不应拒绝原假设，即没有充分理由认为合格品率显著低于 90%

3、单样本 K-S 检验

DATA：儿童身高.SAV

【分析过程】

选择菜单：【分析（A)】一【非参数检验（N)】一【旧对话框（L)】一【1-样本 K-S】



【检验结果】

单样本柯尔莫戈洛夫-斯米诺夫检验

		周岁儿童的身高
个案数		21
正态参数[a,b]	平均值	71.8571
	标准差	3.97851
最极端差值	绝对	.204
	正	.204
	负	-.119
检验统计		.204
渐近显著性（双尾）		.022[c]

a. 检验分布为正态分布。

b. 根据数据计算。

c. 里利氏显著性修正。

【结果分析】

数据的均值为 71. 8571, 标准差为 3. 97851.最大绝对差值为 0. 204,正差极值为 0. 204, 负差极值为－0. 119

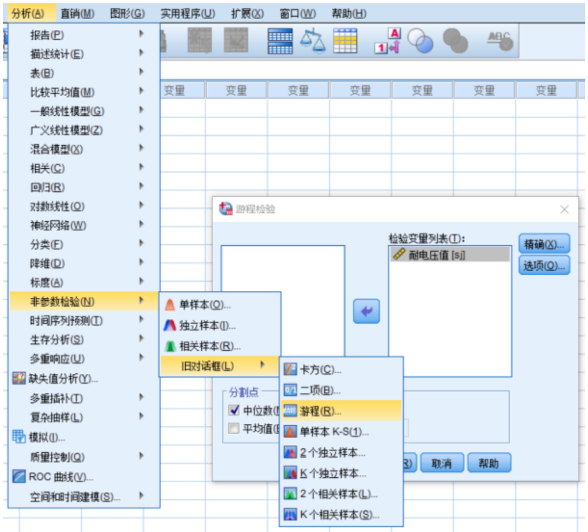
如果显著性水平为 0. 05, 由于概率 P-值大于显著性水平，因此不能拒绝原假设，接受周岁儿童身高的总体分布为正态分布的假设。

4、 变量值随机性检验

DATA： 电缆数据.SAV

【分析过程】

选择菜单：【分析（A)】—【非参数检验（N)】—【旧对话框（L)】—【游程】



【检验结果】

游程检验

	耐电压值
检验值[a]	204.55
个案数 < 检验值	10
个案数 >= 检验值	10
总个案数	20
游程数	13
Z	.689
渐近显著性（双尾）	.491

a. 中位数

[数据集 4] E:\数据分析\data\chap7\电缆数据.sav

【结果分析】

通过上表，

检验值（中位数）为204.55，且共有20个观测样本。

小于和大于等于检验值的样本量各位10

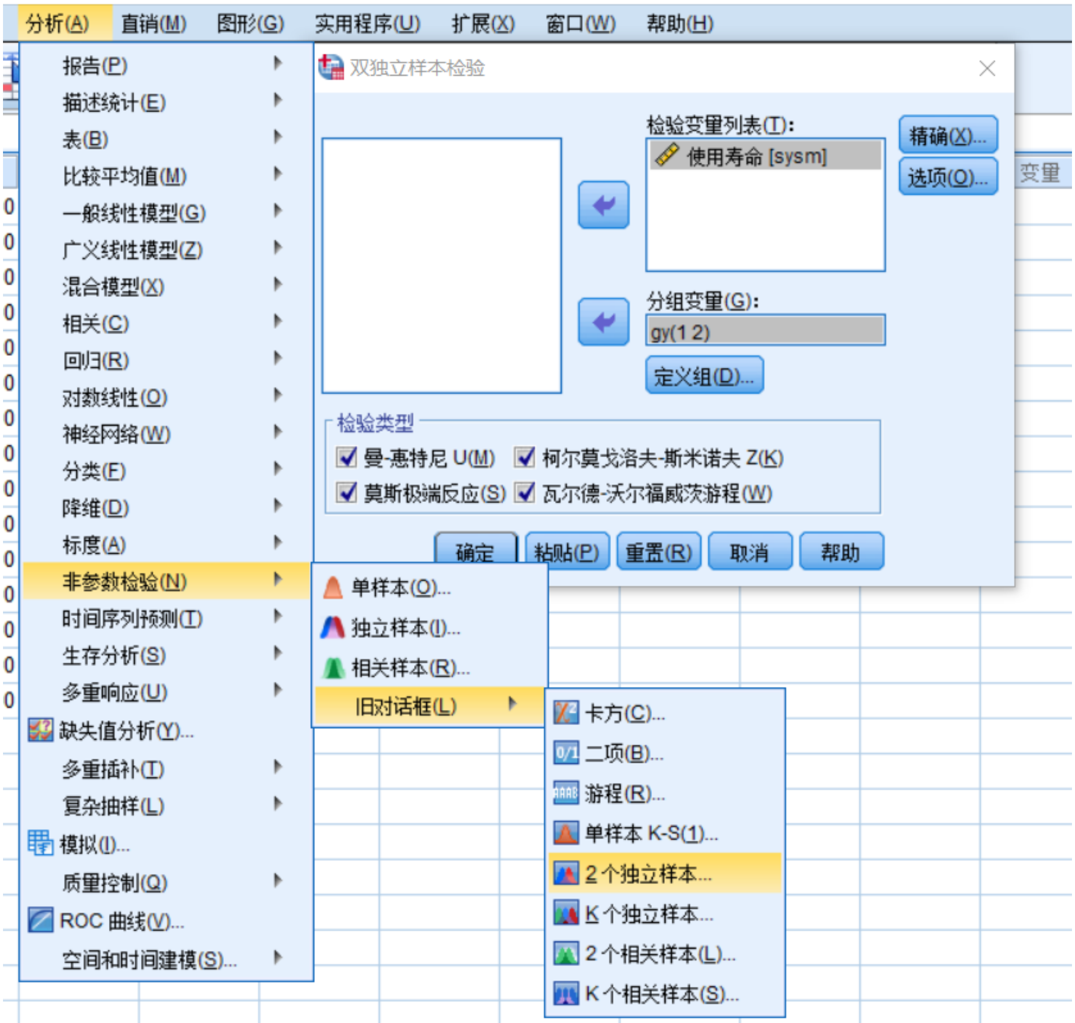
游程数为13，连续性校正的检验统计量的值为0.689

对应的概率P值为0.491，P>α，接受原假设，认为该设备在这段时间内的工作是基本正常的

二、 两独立样本非参数检验

DATA：使用寿命.SAV

【分析过程】



1、 两独立样本的曼-惠特尼 U 检验

【检验结果】

秩

	使用工艺	个案数	秩平均值	秩的总和
使用寿命	甲种工艺	7	11.43	80.00
	乙种工艺	8	5.00	40.00
	总计	15		

检验统计[a]

	使用寿命
曼-惠特尼 U	4.000
威尔科克森 W	40.000
Z	-2.777
渐近显著性（双尾）	.005
精确显著性 [2*(单尾显著性)]	.004[b]

a. 分组变量：使用工艺

b. 未针对绑定值进行修正。

【结果分析】

（1）从甲、乙两种工艺中分别抽取了 7 个和 8 个样本，两个秩和分别为 80 和40

（2）W 统计量应取乙种工艺的秩和 40；U,Z 统计量分别为 4，-2.777

（3）此次抽样属于小样本，因此采用 U 统计量的精确概率 0.005，P<α，拒绝原假设

认为甲、乙两种工艺下产品使用寿命的分布存在显著差异。

2、两独立样本的 K-S 检验

【检验结果】

频率

	使用工艺	个案数
使用寿命	甲种工艺	7
	乙种工艺	8
	总计	15

检验统计[a]

		使用寿命
最极端差值	绝对	.732
	正	.732
	负	.000
柯尔莫戈洛夫-斯米诺夫 Z		1.415
渐近显著性（双尾）		.037

a. 分组变量：使用工艺

【结果分析】

通过以上两表看出，甲、乙两种工艺下产品使用寿命的累计概率的最大绝对差为 0.732，正差极值为 0.732，负差极值为 0.000。

Z 值为 1.145，概率 P 值为 0.037，P<α，拒绝原假设，认为甲、乙两种工艺下的产品使用寿命的分布存在显著差异。

3、两独立样本的游程检验

检验统计[a,b]

		游程数	Z	精确显著性（单尾）
使用寿命	精确游程数	6[c]	-1.059	.149

a. 瓦尔德-沃尔福威茨检验

b. 分组变量：使用工艺

c. 未遇到任何组内绑定值。

【结果分析】

通过上表可以看出，甲、乙两种工艺下产品使用寿命秩的游程数为 6，因此得出 Z 统计量观测值为-1.059，对应的概率 P 值为0.149，P>α，接受原假设

故认为甲、乙两种工艺下的产品使用寿命的分布不存在显著差异。

4、两独立样本的极端反应检验

【检验结果】

检验统计[a,b]

		使用寿命
实测控制组范围		10
	显著性（单尾）	.084
剪除后控制组跨度		6
	显著性（单尾）	.100
在两端剪除了离群值		1

a. 莫斯检验

b. 分组变量：使用工艺

【结果分析】

通过上表可以看出，观察控制组跨度、已修正控制组跨度分别为 10 和 6。两种情况下的概率分别为 0.084 和 0.100。

假设显著性水平α为 0.05，两个概率 P 值均大于α，无论是否剔除极端值，都接受原假设，认为甲、乙两种工艺下的产品使用寿命的分布不存在显著差异。

两独立样本非参数检验总结

Mann-Whitney U 检验法

检验两个样本的总体在某些位置上是否相同，其基于对平均秩的分析实现推断。

检验思路：首先对两个样本合并并按升序排列得出每个数据的秩，然后对这两个样本求平均秩并计算第一组样本的平均秩优于第二组样本的每个秩的格式 N1 和第二组样本的每个秩优于第一组样本的每个秩的个数 N2.如果平均秩和 N1、N2 之间的差距过大，则认为两个样本来自于不同的总体

Kolmogorov-Smirnov Z 双样本检验法

首先对两个样本合并并按升序排列得出每个数据的秩，然后得出样本秩的累积频率与样本点的累积频率的差值序列并计算该序列的 K-S Z 统计量，如果该统计量的伴随概率 P 值小于显著性水平则认为两个样本的总体分布具有显著性差异

Moses 极限反应检验法

将一个样本作为实验样本而另一个样本作为控制样本，将两个样本合并按升序排列得出每个数据的秩并计算控制组的跨度（即控制组样本中最大秩和最小秩之间包含的样本个数）。然后，忽略取值极高和极低的各 5%数据后计算截头跨度，如果跨度和截头跨度相差较大，则认为两个样本存在极限反应其总体分布具有显著性差异

Wald-Wolfowitz 游程检验法

检验两个样本是否被随机地赋秩。

检验思路：首先对两个样本合并并按升序排列，然后对样本标志值序列求游程，如果得到的游程数较小则认为两个样本的总体分布具有显著性差异

三、　　多独立样本非参数检验

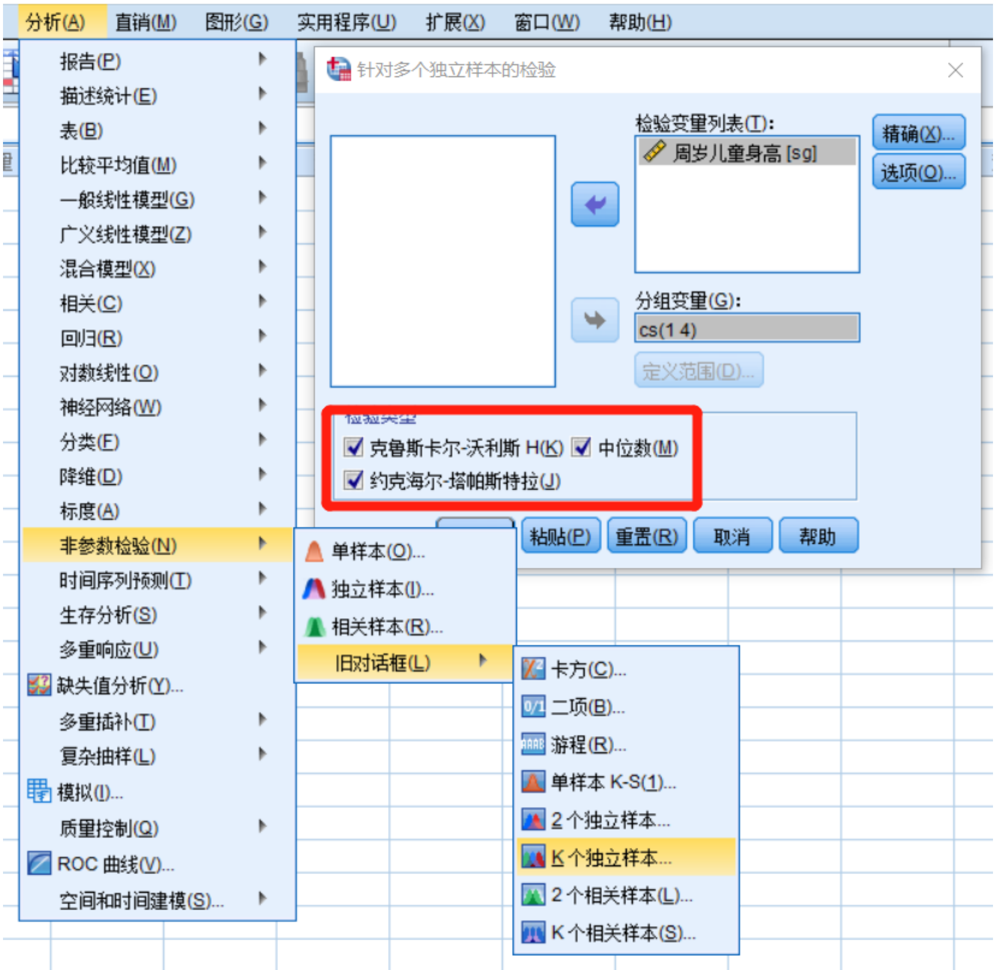
DATA：多城市儿童身高.SAV

简述

多独立样本的非参数检验是通过分析多组独立样本数据，推断样本来自的多个总体的中位数或分布是否存在显著差异。多独立样本是指按独立抽样方式获得的多组样本。

SPSS 提供的多独立样本的非参数检验方法主要包括中位数检验、Kruskal-Wm 检验、Jonckheere-Terpstra 检验。

【分析过程】



【检验结果】

1、 多独立样本的中位数检验：

频率

		城市标志 北京	上海	成都	广州
周岁儿童身高	> 中位数	4	0	5	0
	<= 中位数	1	5	0	5

检验统计[a]

	周岁儿童身高
个案数	20
中位数	74.0000
卡方	16.768[b]
自由度	3
渐近显著性	.001

a. 分组变量：城市标志

b. 8 个单元格 (100.0%) 的期望频率低于 5。期望的最低单元格频率为 2.3。

2、 多独立样本的 Kruskal-Wallis 检验：

秩

	城市标志	个案数	秩平均值
周岁儿童身高	北京	5	14.40
	上海	5	8.20
	成都	5	15.80
	广州	5	3.60
	总计	20	

检验统计[a,b]

	周岁儿童身高
卡方	13.900
自由度	3
渐近显著性	.003

a. 克鲁斯卡尔-沃利斯检验

b. 分组变量：城市标志

结果分析

通过左边两个表中数据可以得出：四组数据的中位数为 74，计算出的卡方统计量 16.768，对应的概率 P 值为 0.001，P<α，拒绝原假设，认为四个城市周岁儿童身高的分布存在显著差异。

注：应当注意的是卡方检验对期望频数小于 5 的单元格数是有一定限制的。

由于该例违背了限制，SPSS 出了提示信息，因此该检验只能作为一种参考。

由左侧表格可以看出，四个城市共 20 个周岁儿童，其身高的平均秩分别为 14.40，8.20，15.80，3.60，K-W 统计量为 13.9，概率 P 值为 0.003，P<α，拒绝原假设，认为四个城市周岁儿童身高的平均秩存在显著差异，且总体分布也存在显著差异。

3、 多独立样本的 Jonckheere-Terpstra 检验:

约克海尔-塔帕斯特拉检验[a]

	周岁儿童身高
城市标志 中的级别数	4
个案数	20
实测 J-T 统计	45.500
平均值 J-T 统计	75.000
J-T 统计的标准差	14.764
标准 J-T 统计	-1.998
渐近显著性（双尾）	.046

a. 分组变量：城市标志

通过上表可以看出，观测的 J-T 值为 45.500，所有 J-T 值的平均值为 75.000，标准差为 14.764，观测的 J-T 值的标准化值为-1.998,小于平均值且差距较明显。

J-T 统计量的概率 P 为 0.046，如果显著性水平 α 为 0.05，由于概率 P 值小于显著性水平 α， 因此应拒绝原假设，认为四个城市周岁儿童身高的分布存在显著差异。

综上，通过多独立样本的非参数检验（ 中位数检验法、 Kruskal-Wallis 检验法、Jonckheere-Terpstra 检验法），一致认为四个城市周岁儿童身高的分布存在显著差异。

四、两配对样本非参数检验

DATA：统计学学习.SAV

【检验步骤】

选择菜单：【分析（A)】—【非参数检验（N)】—【旧对话框（L)】—【两个相关样本】

1、两配对样本的 McNemar 检验:

【检验结果】

学习前的认识 & 学习后的认识

学习前的认识	学习后的认识：不重要	学习后的认识：重要
不重要	3	4
重要	2	3

检验统计[a]

	学习前的认识 & 学习后的认识
个案数	12
精确显著性（双尾）	.688[b]

a. 麦克尼马尔检验

b. 使用了二项分布。

【结果分析】

通过上面两张表可知，此次共抽去了 12 名同学，其中有 4 人将认为统计学“不重要”变成“重要”，有 2 人将认为统计学“重要”变成“不重要”。

双侧的二项分布累计概率为 0.687（单侧为 0.3435），P> α，接受原假设，认为学习前后学生对统计学重要性的认识没有发生显著变化。

2、两配对样本的符号检验:

频率

		个案数
学习后的认识 - 学习前的认识	负差值[a]	2
	正差值[b]	4
	绑定值[c]	6
	总计	12

a. 学习后的认识 < 学习前的认识

b. 学习后的认识 > 学习前的认识

c. 学习后的认识 = 学习前的认识

检验统计[a]

	学习后的认识 - 学习前的认识
精确显著性（双尾）	.688[b]

a. 符号检验

b. 使用了二项分布。

符号检验分析:

可以看出，在抽样的 12 名同学中，将“重要”变为“不重要”的有 2 人，将“不重要”变为“重要”的有 4 人，6 人保持不变。

双侧的二项分布累计概率为 0.687，P> α，接受原假设

认为学习前后学生对统计学重要性的认识没有发生显著变化。

3、两配对样本的 Wilcoxon 符号秩检验:

秩

		个案数	秩平均值	秩的总和
学习后的认识 - 学习前的认识	负秩	2[a]	3.50	7.00
	正秩	4[b]	3.50	14.00
	绑定值	6[c]		
	总计	12		

a. 学习后的认识 < 学习前的认识

b. 学习后的认识 > 学习前的认识

c. 学习后的认识 = 学习前的认识

Wilcoxon 符号秩检验分析:

此次样本数为 12，负号秩总和为 7，正号秩总和为 14

检验统计[a]

	学习后的认识 - 学习前的认识
Z	-.816[b]
渐近显著性（双尾）	.414

a. 威尔科克森符号秩检验

b. 基于负秩。

Wilcoxon 符号秩检验分析：

Z 检验统计量为-0.816（其中，n=6，W=6），对应的概率P值为0.414，P>α，接受原假设

认为学习前后学生对统计学重要性的认识没有发生显著变化。

综合以上检验，可以得出：两配对样本的非参数检验（McNemar 检验法、符号检验法、Wilcoxon 符号秩检验法）

学习前后学生对统计学重要性的认识没有发生显著变化的结论。

参数检验和非参数检验的区别

1、参数检验是针对参数做的假设；

非参数检验是针对总体分布情况做的假设，这是区分的一个重要特征；

2、参数检验要利用到总体的信息（总体的分布、总体的一些参数特征，如方差），以总体分布和样本信息对总体参数做出推断；

非参数检验不需要利用总体信息，以样本信息对总体分布做出推断；

3、正态分布用参数检验，非正态分布用非参数检验。