

第十章 非参数秩和检验

一句话抓住本质

非参数检验不在乎“数值多大”，
只关心：谁大、谁小、排第几。

一、为什么需要“非参数”？ (Why)

参数检验 (parametric test) 的前提

- 总体分布 **已知**
- 常见假设：**正态分布**
- 检验对象：**总体参数** (均值、方差)

✦ 典型代表：

- t 检验
 - 方差分析 (ANOVA)
-

现实问题是：

- 分布 **未知 / 明显非正态**
- 数据有偏态、重尾、离群值
- 数据本身就是“等级”“程度”

👉 参数检验的理论基础站不住

二、非参数检验在干什么？ (Core idea)

非参数检验不假设分布，
也不直接检验参数，
而是检验“位置是否不同”。

“位置”是什么意思？

不是精确的均值，而是：

- 整体偏大还是偏小
- 中位数是否不同
- 排名是否系统性偏移

👉 通过“秩次 (rank)”来比较。

三、秩和检验的基本操作（非常重要）

所有秩和检验，底层操作完全一致：

- 1 把所有观测值放在一起
- 2 从小到大排序
- 3 给每个值一个“名次（秩）”
- 4 比较不同组的秩之和 / 秩分布

🔥 你已经从此刻开始，
不再比较“数值”，而是比较“顺序”。

四、什么时候该用非参数？（三大典型场景）

① 总体分布未知或非正态

- 明显偏态
- 小样本 + 不确定分布
- 离群值多、极端值

② 有序或半定量资料（非常关键）

- 痊愈 / 显效 / 有效 / 无效
- 轻 / 中 / 重
- 优 / 良 / 中 / 差

✦ 这些数据：

- 有顺序
- 但“间距不等”
👉 不能当连续变量

③ 数据端点不确定

- “大于某值” “小于检测下限”
- 区间数据、删失数据

五、非参数检验的优缺点（必须正视）

优点

- 不依赖分布假设
- 对异常值不敏感
- 适用面非常广

代价（非常重要）

只利用了“秩次信息”，
丢掉了“数值大小信息”。

结果就是：

- 检验效能 (power) 较低
- 当数据满足参数检验条件时：
👉 非参数 不如参数检验敏感

✦ 所以选择原则是：

能用参数检验，绝不用非参数；
只有参数前提不成立，才退而求其次。

六、非参数 vs 参数：一句话对照（记住这个）

参数检验：

“我相信数据来自某个分布”

非参数检验：

“我不相信分布，只相信排序”

七、本章的“方法地图”（给你预期）

既然是秩和检验，后面你会看到的都是：

- 两独立样本秩和检验
(替代两独立样本 t 检验)
- 配对样本秩和检验
(替代配对 t 检验)
- 多样本秩和检验
(替代单因素 ANOVA)

✦ 逻辑结构和前几章一一对应。

八、考试级一句话总结（直接可写）

非参数检验是在总体分布类型未知或不满足参数检验条件时，通过对样本数据进行排序编秩并比较秩次分布，对总体分布位置进行假设检验的方法。由于其不依赖分布假设，适用范围广，但当数据满足参数检验条件时，其检验效能低于参数检验。

给你一个“方法选择”的终极锚点

分布清楚、条件满足 → 参数检验

分布存疑、数据有序 → 非参数检验

第一节 配对设计资料的符号秩和检验

Wilcoxon 符号秩和检验 (Signed-rank test)

一句话抓住本质

Wilcoxon 符号秩和检验问的不是“差多少”，
而是：

👉 正差和负差，谁在系统性地占上风？

一、它解决的是什么问题？ (Purpose)

这是一个**配对设计的非参数检验**，用于替代：

配对 t 检验

适用问题表述为：

- 同一对象 **处理前 vs 处理后**
- 同一对象 **方法 A vs 方法 B**
- 同一对象 **两种条件下的测量值**

🔴 检验目标不是均值，而是：

差值总体的中位数是否为 0

二、原假设在说什么？ (Hypothesis)

H_0 : 差值来自中位数为 0、分布对称的总体

翻译成成人话：

如果两种处理真的没差别，
那：

- 有正差
- 也有负差
- 大小和方向应该是对称的

三、核心直觉：为什么要“符号 + 秩”？ (Very important)

Wilcoxon 的设计非常巧：

① 先看方向（符号）

- 正差：+
- 负差：-

👉 判断谁多、谁少

② 再看大小（秩）

- 不直接用差值
- 而是对 |差值| 排序、编秩

👉 判断谁更“有分量”

③ 最终看什么？

正秩和 vs 负秩和

- 若两处理等效：
 - 👉 两者绝对值应接近
 - 若有系统差异：
 - 👉 一边会明显“压倒”另一边
-

四、和“符号检验”的关键区别（很容易混）

你可以把它们对成一行：

符号检验：

只看正负，不看大小（更粗糙）

符号秩和检验（Wilcoxon）：

既看正负，又看大小（更有力）

👉 Wilcoxon 是符号检验的升级版

👉 在非参数检验里，优先用 Wilcoxon

五、Wilcoxon 的隐含前提（要说清楚）

虽然它是非参数的，但**不是完全无前提**：

- 配对差值来自 **对称分布**
- 不要求正态
- 但不能极端偏斜

✖ 如果差值分布严重不对称：

👉 连 Wilcoxon 也会变得不稳

六、它和配对 t 检验的关系（一句话打通）

配对 t 检验：

比较差值的“均值”

Wilcoxon 符号秩和检验：

比较差值的“中位数 / 排名结构”

当：

- 正态假设可信 → **t 检验更有力**
 - 正态假设存疑 → **Wilcoxon 更安全**
-

七、什么时候该第一反应用它？（考试 + 实战）

- 配对设计 ✓
- 连续或半定量数据 ✓
- 样本量不大 ✓
- 差值分布非正态 / 有异常值 ✓

👉 **Wilcoxon 符号秩和检验**

八、考试级一句话总结（直接可写）

Wilcoxon 符号秩和检验是一种用于配对设计资料的非参数检验方法，通过对配对差值进行排序并比较正负秩和的大小，检验差值总体分布的中位数是否为 0，适用于差值分布不满足正态假设的情况。

给你一个“不会用错”的终极锚点

配对 + 非正态 → Wilcoxon

配对 + 正态 → 配对 t

只要你在做题时先问这两个条件，这一节你就不会再犹豫。

【例 10-1】8 名男性志愿者参加一项睡眠剥夺对精神运动警觉性任务注意力的影响研究，分别在正常睡眠状况下以及睡眠剥夺状态下，测试其精神运动警觉性任务注意力脱漏次数，检测结果如表 10-1 所示。

表 10-1 8 名男性志愿者不同睡眠状态下精神运动警觉性任务注意力脱漏次数

志愿者编号 (1)	正常睡眠 (2)	睡眠剥夺 (3)	$d=(3)-(2)$ (4)	秩次 (5)
1	4	3	-1	-1
2	2	13	11	6
3	0	20	20	8
4	3	6	3	3
5	1	18	17	7
6	5	9	4	4
7	4	6	2	2
8	1	11	10	5
合计				$T_+=35, T_-=1$

(李康,贺佳主编, 2024, p. 101) - Please click Zotero - Refresh in Word/LibreOffice to update all fields.

(1) 建立检验假设,确定检验水准

$H_0: M_d = 0$, 即不同睡眠状态下注意力脱漏次数差值的总体中位数为零

$H_1: M_d \neq 0$, 即不同睡眠状态下注意力脱漏次数差值的总体中位数不为零

$\alpha = 0.05$

(2) 编秩次并求秩和统计量

首先求出各对数据的差值,见表 10-1 中的第(4)栏;然后编秩次,按照差值绝对值由小到大编秩,并按差值的正负给秩次加上正负号;若差值为“0”,舍去不计,总的对子数也要减去此对子数(记为 n);若差值的绝对值相等,取其平均秩次。最后,分别求正负秩次之和 T_+ 与 T_- (表 10-1),任取 T_+ 或 T_- 为检验统计量 T ,本例选取 $T = T_- = 1$ 。

(3) 确定 P 值,作出推断

当 $n \leq 50$ 时,根据 n 和 T 可查配对设计用的 T 界值表(附表 8),若检验统计量 T 值在上下界值范围内,则 P 值大于表上方对应的概率值;若 T 值在上下界值外,则 P 值小于表上方对应的概率值。本例 $n = 8$,查 T 界值表 $T_{0.05(8)} = 3 \sim 33$, $T = 1$ 不在 $3 \sim 33$ 范围内, $P < 0.05$,按 $\alpha = 0.05$ 水准,拒绝 H_0 ,即不同睡眠状态下注意力脱漏次数差异有统计学意义。

需要注意:当 $n > 50$ 时,无法查表,可利用秩和分布的近似正态法进行检验。已知在原假设 H_0 成立时,近似有

$$z = \frac{|T - n(n+1)/4| - 0.5}{\sqrt{n(n+1)(2n+1)/24}} \quad (10-1)$$

其中 0.5 为连续性校正数, z 近似服从标准正态分布。

当相同秩次较多时, z 值偏小,应采用校正公式,即

$$z_c = \frac{|T - n(n+1)/4| - 0.5}{\sqrt{\frac{n(n+1)(2n+1)}{24} - \frac{\sum(t_j^3 - t_j)}{48}}} \quad (10-2)$$

其中 t_j 为第 j 个相同秩次(即平均秩次)的个数,假定有 2 个秩次为 2.5, 4 个秩次为 8.5, 则 $t_1 = 2$, $t_2 = 4$, 故有

$$\sum(t_j^3 - t_j) = (t_1^3 - t_1) + (t_2^3 - t_2) = (2^3 - 2) + (4^3 - 4) = 66$$

(李康,贺佳主编, 2024, p. 102) - Please click Zotero - Refresh in Word/LibreOffice to update all fields.

【例 10-2】指导 28 名有轻度牙周疾病的成年人进行良好的口腔卫生保健,6 个月后,按照牙周情况好转高低程度分别给予 +3, +2, +1;牙周情况变差程度依次给予分数 -1, -2, -3;没有变化给予 0 分。数据如表 10-2 所示,试对此项干预的结果进行评价。

表 10-2 实行良好口腔卫生习惯 6 个月后牙周情况的变化程度

变化分数	人数	变化分数	人数
+3	4	-1	4
+2	5	-2	2
+1	6	-3	2
0	5		

- TES
- (1) 建立检验假设,确定检验水准

$H_0:M_d=0$,即前后变化分数的总体中位数为零

$H_1:M_d\neq 0$,即前后变化分数的总体中位数不为零

$\alpha=0.05$

(2) 计算 T 统计量

记变化分数的绝对值为 d ,编秩及计算正负秩和结果如表 10-3 所示。

(李康,贺佳主编, 2024, p. 102) - Please click Zotero - Refresh in Word/LibreOffice to update all fields.

表 10-3 实行良好口腔卫生习惯 6 个月后牙周情况的变化程度

d (1)	频数		秩次范围 (5)	平均秩次 (6)	负秩和 (7)=(2)×(6)	正秩和 (8)=(3)×(6)
	— (2)	+ (3)				
1	4	6	10	5.5	22	33
2	2	5	7	14	28	70
3	2	4	6	20.5	41	82
合计	8	15	23	—	$T_- = 91$	$T_+ = 185$

(3) 确定 P 值,作出推断

查 T 界值表(附表 8), $T_{0.05(23)} = 73 \sim 203$, $T = T_- = 91 > 73$, T 统计量值落在上下界值之间, $P > 0.05$, 按 $\alpha = 0.05$ 水准, 不拒绝 H_0 , 即对有轻度牙周疾病的成年人, 实行良好的口腔卫生 6 个月后, 尚不能说明此项干预对牙周改善有效果。

本例若用近似正态法计算, 由于上述资料相同秩次较多, 用校正公式(10-2)得出

$$\begin{aligned}
 z_c &= \frac{|T - n(n+1)/4| - 0.5}{\sqrt{\frac{n(n+1)(2n+1)}{24} - \frac{\sum(t^3 - t)}{48}}} \\
 &= \frac{|91 - 23(23+1)/4| - 0.5}{\sqrt{\frac{23 \times (23+1) \times (2 \times 23+1)}{24} - \frac{(10^3 - 10) + (7^3 - 7) + (6^3 - 6)}{48}}} \\
 &= 1.44
 \end{aligned}$$

查 z 界值表 $z_{0.05/2} = 1.96$, 本例中 $z_c < 1.96$, $P > 0.05$, 按 $\alpha = 0.05$ 的检验水准, 不拒绝 H_0 , 结论同前。

(李康, 贺佳主编, 2024, p. 103) - Please click Zotero - Refresh in Word/LibreOffice to update all fields.

第二节 两独立样本比较的秩和检验

Wilcoxon 秩和检验 (Rank sum test / Mann-Whitney U)

一句话抓住本质

Wilcoxon 秩和检验不比较“均值差多少”,

而是问:

👉 一组的数据, 是否系统性地“更大 / 更小”?

一、什么时候该用它？ (Purpose)

这是**两独立样本**的非参数检验，用来替代：

两独立样本 t 检验

适用条件非常清晰：

- 两组 **相互独立**
- 数据为连续型或有序等级资料
- 正态性 / 方差齐性 **不可信**

✦ 检验目标不是均值，而是：

两个总体的**分布位置**是否不同
(通常可理解为中位数是否不同)

二、原假设在说什么？ (Hypothesis)

H_0 : 两总体的分布位置相同

直觉翻译：

如果两组没有差别，
那它们的数据混在一起排序后，
秩次应该是“均匀交错”的。

三、秩和检验的核心操作 (Very important)

不管你用查表法还是正态近似，**底层逻辑完全一样**：

- 1 把两组样本 **合并**
- 2 从小到大 **排序编秩**
- 3 计算 **样本量较小那一组** 的秩和 T

🔴 关键直觉:

若 H_0 成立,
小样本组的秩和
应接近一个“理论平均值”。

四、统计量 T 在干什么? (直觉解释)

理论平均秩和

$$E(T) = \frac{n_1(N+1)}{2}, N = n_1 + n_2$$

这代表什么?

如果两组完全混合、无差别,
那 n_1 个秩的平均位置,
就应该落在整个秩序的正中间。

判断逻辑 (一句话版)

- $T \approx E(T)$
 - 👉 差异可由随机解释
 - T 明显偏大或偏小
 - 👉 某一组整体偏大 / 偏小
 - 👉 拒绝 H_0
-

五、为什么要用正态近似? (方法选择)

当:

- n_1 或 n_2 较大
- 超出秩和表范围

👉 T 的分布可以近似为正态

于是构造统计量：

$$z = \frac{|T - \frac{n_1(N+1)}{2}| - 0.5}{\sqrt{\frac{n_1 n_2 (N+1)}{12}}}$$

✦ 三个你要“看懂、不背”的点：

- 分子：
 - 👉 实际秩和 - 理论秩和
 - -0.5：
 - 👉 连续性校正
 - 分母：
 - 👉 抽样误差尺度
-

六、为什么要对“相同秩次”校正？（非常容易忽略）

当数据是：

- 等级资料
- 或大量重复值

👉 秩的离散程度会被低估

于是要用校正公式：

$$z_c = \frac{|T - \frac{n_1(N+1)}{2}| - 0.5}{\sqrt{\frac{n_1 n_2}{12N(N-1)} (N^3 - N - \sum(t_j^3 - t_j))}}$$

✦ 直觉一句话：

重复值越多，
随机波动越小，

必须修正标准误，
否则会“虚假显著”。

七、它和两独立样本 t 检验的关系（一眼打通）

t 检验：

比较均值，利用全部数值信息

Wilcoxon 秩和检验：

比较秩次，只利用相对大小

👉 当正态 + 方差齐性成立：

t 检验更有力

👉 当条件存疑：

Wilcoxon 更稳健

八、什么时候“第一反应”就该选它？（考试信号）

- “两组独立样本”
- “数据明显偏态 / 有离群值”
- “等级资料”
- “样本量不大”

👉 **Wilcoxon 秩和检验**

九、考试级一句话总结（直接可写）

Wilcoxon 秩和检验是一种用于比较两独立样本总体分布位置是否存在差异的非参数方法，通过对合并样本排序并比较秩和大小，在正态假设不成立时可替代两独立样本 t 检验，其统计量在样本量较大时可采用正态近似进行检验。

给你一个“不会选错方法”的终极锚点

- 独立 + 非正态 → Wilcoxon 秩和
- 独立 + 正态 → 两独立样本 t

只要你在题目里先确认“独立 or 配对”，再看正态性是否可信，这一节就不会再混乱。

【例 10-3】某地检测了 23 名成年人的尿镉含量(表 10-4),问男性女性的尿镉含量是否不同?

表 10-4 不同性别成年人的尿镉含量 单位:μg/L

男性		女性		男性		女性	
尿镉含量	秩次	尿镉含量	秩次	尿镉含量	秩次	尿镉含量	秩次
2.05	1	3.50	5	7.71	15	8.23	17
2.23	2	3.60	6	7.71	15	9.54	18
2.45	3	3.78	7	13.56	19	14.00	20
2.47	4	4.20	9.5			16.00	21
4.10	8	6.64	12			32.06	22
4.20	9.5	7.00	13			45.07	23
4.45	11	7.71	15	$n_1=10$	$T_1=87.5$	$n_2=13$	$T_2=188.5$

(李康,贺佳主编, 2024, p. 103) - Please click Zotero - Refresh in Word/LibreOffice to update all fields.

(1) 建立检验假设,确定检验水准
 H_0 :男性女性尿镉含量的总体分布相同
 H_1 :男性女性尿镉含量不同
 $\alpha=0.05$

(2) 确定秩和检验统计量 T

首先编秩号,即将两样本 23 个数据由小到大统一编秩,结果见表 10-4。排序时若有相同数据,取平均秩次。本例中两组有相同的观察值 4.20 与 7.71,原秩次分别应为 9,10 与 14,15,16,所以取平均秩次分别为 9.5 和 15。进而,将两组数据的秩次分别求和,若两组例数相同,则任取一组的秩和作为统计量;若两组例数不同,则以例数较小者对应的秩和作为统计量。本例中两组例数分别 10 和 13,取较小者为 $n_1=10, T=T_1=87.5$ 。

(3) 确定 P 值,作出推断结论

当 $n_1 \leq 10, n_2 - n_1 \leq 10$ 时,查两样本比较的 T 界值表(附表 9),先从表的左侧查 n_1 (两样本量较小者),本例为 10;再从表上方找到两样本量的差 ($n_2 - n_1$),本例 $n_2 - n_1 = 3$,两者交叉处即为 T 的临界值。将检验统计量 T 值与 T 的临界值作比较,如果 T 在界值范围内,则 P 值大于表上方的概率值;若 T 等于界值或在界值范围外,则 P 值等于或小于表上方的概率值。本例 T 的双侧临界值范围为 88~152,检验统计量 T 值为 87.5,超出范围, $P < 0.05$,按照 $\alpha=0.05$ 检验水准,拒绝 H_0 ,两组的秩和差别有统计学意义,说明不同性别成年人的尿镉含量不同,女性尿镉含量相对较高。

(李康,贺佳主编, 2024, p. 104) - Please click Zotero - Refresh in Word/LibreOffice to update all fields.

【例 10-4】44 名健康人与 24 名慢性气管炎病人痰液嗜酸性粒细胞数的测量结果如表 10-5,问健康人与慢性气管炎病人痰液嗜酸性粒细胞数有无差别?

表 10-5 两组人痰嗜酸性粒细胞的秩和计算

嗜酸性粒细胞数 (1)	例数		统一编秩		例数较小组的秩和 (6)=(3)×(5)
	健康人 (2)	病人 (3)	秩次范围 (4)	平均秩次 (5)	
—	5	11	1~16	8.5	93.5
+	18	10	17~44	30.5	305.0
++	16	3	45~63	54	162.0
+++	5	0	64~68	66	0.0
合计	44	24	—	—	$T_1=560.5$

(李康,贺佳主编, 2024, p. 104) - Please click Zotero - Refresh in Word/LibreOffice to update all fields.

(1) 建立检验假设,确定检验水准

H_0 :健康人与慢性气管炎病人痰液嗜酸性粒细胞数的总体分布相同

H_1 :健康人与慢性气管炎病人痰液嗜酸性粒细胞数不同

$\alpha = 0.05$

(2) 计算检验统计量

表 10-5 中第(4)栏按第(2)与(3)栏数据统一编秩号,第(5)栏为各等级的平均秩次,第(6)栏则是较小样本的秩和,本例中 $T = T_1 = 560.5$,将其代入公式(10-4)得出

$$\begin{aligned} z_c &= \frac{|T_1 - n_1(N+1)/2| - 0.5}{\sqrt{\frac{n_1 \cdot n_2}{12N(N-1)}(N^3 - N - \sum(t^3 - t))}} \\ &= \frac{|560.5 - 24 \times (68+1)/2| - 0.5}{\sqrt{\frac{24 \times 44}{12 \times 68 \times (68-1)} \times [68^3 - 68 - (16^3 - 16 + 28^3 - 28 + 19^3 - 19 + 5^3 - 5)]}} \\ &= 3.62 \end{aligned}$$

(3) 确定 P 值,做出推断结论

由于 $z_c = 3.62 > z_{0.05/2} = 1.96$, $P < 0.05$,则按照 $\alpha = 0.05$ 检验水准,拒绝 H_0 ,两组的差别有统计学意义,认为健康人与慢性气管炎病人痰液嗜酸性粒细胞数不同。

(李康,贺佳主编, 2024, p. 105) - Please click Zotero - Refresh in Word/LibreOffice to update all fields.

第三节 多个独立样本比较的秩和检验

Kruskal-Wallis 秩和检验 (K-W / H 检验)

一句话抓住本质

Kruskal-Wallis 检验问的不是“均值是否不同”,

而是:

👉 哪一组的数据, 整体更靠前 / 更靠后?

一、它解决什么问题? (Purpose)

当你面对的是：

- ≥ 3 组独立样本
- 数据是 计量资料或有序资料
- ANOVA 的条件不成立（非正态、方差不齐、离群值多）

👉 Kruskal-Wallis 检验就是首选

它在非参数体系中的地位，相当于：

单因素方差分析（one-way ANOVA）的替代方案

二、原假设在说什么？（Hypothesis）

H_0 : 各组总体的分布位置相同

直觉翻译：

如果各组真的没有差别，
那把所有数据放在一起排序后，
各组的秩次分布应该是相似的。

三、核心直觉：它和 Wilcoxon 秩和有什么关系？

你可以把它看成一句话：

K-W = Wilcoxon 秩和的“多组推广版”

逻辑完全一致：

- 不比较数值本身
- 比较 秩次结构
- 看哪一组的秩和系统性偏大 / 偏小

四、K-W 检验在“暗中”做了什么？（Very important）

底层操作一模一样：

- 1 将所有组的观测值 **合并**
- 2 对全部数据 **排序编秩**
- 3 计算每一组的 **秩和 / 平均秩**
- 4 比较各组秩和之间的差异程度

✦ 如果 H_0 成立：

各组平均秩应接近

✦ 如果某组整体偏大 / 偏小：

秩和会明显偏离 → 拒绝 H_0

五、它到底检验的是什么？（别说错）

K-W 检验检验的是：

“总体分布位置是否相同”，
而不是严格意义上的“中位数相同”。

但在多数常见场景下：

- 若分布形态相近
 - 👉 可以 **近似理解为中位数差异检验**
-

六、什么时候一定要考虑它？（考试信号）

- “三组及以上独立样本”
- “数据明显偏态”
- “等级 / 评分资料”

- “样本量不大或极不均衡”
- “方差齐性不满足”

👉 Kruskal–Wallis

七、它解决不了什么？（非常重要）

和 ANOVA 一样：

K-W 只能告诉你：
至少有一组不同。

它不能告诉你：

- 哪几组不同
- 差异发生在哪

👉 拒绝 H_0 之后，
仍需做多重比较（秩和的两两比较）

（逻辑与 ANOVA + 多重比较完全平行）

八、K-W 与 ANOVA 的关系（一句话打通）

ANOVA:
比较均值，利用数值大小

K-W:
比较秩次，利用相对顺序

👉 条件满足 → **ANOVA 更有力**

👉 条件不满足 → **K-W 更稳健**

九、考试级一句话总结（直接可写）

Kruskal-Wallis 秩和检验是一种用于比较多个独立样本总体分布位置是否存在差异的非参数方法，通过对合并样本排序并比较各组秩和大小，在不满足方差分析条件时可替代单因素方差分析。

给你一个“方法选择”的终极锚点

独立样本 ≥ 3 组：

正态 + 方差齐 \rightarrow **ANOVA**

非正态 / 有序 \rightarrow **K-W**

只要你在做题时**先判断“独立 + 组数”**，再看**分布条件**，这一节你就不会再犹豫。

【例 10-5】为研究霍乱菌苗不同途径的免疫效果,对不同途径免疫 21 天后血清抗体滴度水平进行了测定,检测结果见表 10-6 中(1)~(3)栏,问各组间的血清抗体滴度水平之间是否存在差异?

表 10-6 不同途径免疫 21 天后血清抗体滴度的分布与秩和计算

抗体滴度 (1)	气雾组		皮下注射组 (4)	合计 (5)	平均秩次 (6)	秩和		
	80 (2)	100 (3)				80 (7)	100 (8)	皮下 (9)
1 : 10	2	4	2	8	4.5	9	18	9
1 : 20	15	7	1	31	20	300	140	20
1 : 40	10	12	13	66	49	490	588	637
1 : 80	5	7	9	87	77	385	539	693
1 : 160	1	2	5	95	91.5	91.5	183	457.5
1 : 320	—	—	1	96	96	—	—	96
合计	33	32	31	—		1 275.5	1 468	1 912.5

(1) 建立检验假设,确定检验水准

H_0 :三组血清抗体滴度水平的总体分布相同

H_1 :三组血清抗体滴度水平的总体分布位置不全相同

$\alpha = 0.05$

(李康,贺佳主编, 2024, p. 105) - Please click Zotero - Refresh in Word/LibreOffice to update all fields.

(2) 计算检验统计量 H

首先将各组数据统一按从小到大顺序编秩,如有相等数值则取平均秩次;然后分别计算各组的秩和 R_i ;最后计算检验统计量 H ,即

$$H = \frac{12}{N(N+1)} \sum \frac{R_i^2}{n_i} - 3(N+1) \quad (10-5)$$

其中 $N = n_1 + n_2 + \cdots + n_g$ 为各组例数之和。本例 $N = 96, R_1 = 1\,275.5, R_2 = 1\,468, R_3 = 1\,912.5$, 由此得到

$$\begin{aligned} H &= \frac{12}{N(N+1)} \sum \frac{R_i^2}{n_i} - 3(N+1) \\ &= \frac{12}{96(96+1)} \left(\frac{1\,275.5^2}{33} + \frac{1\,468^2}{32} + \frac{1\,912.5^2}{31} \right) - 3(96+1) = 11.36 \end{aligned}$$

本例相同秩次较多,使用校正公式更加准确,即

$$H_c = \frac{H}{1 - \frac{\sum(t_j^3 - t_j)}{N^3 - N}} \quad (10-6)$$

其中 t_j 为第 j 个相同秩次(即平均秩次)的个数。本例有

$$H_c = \frac{11.36}{1 - \frac{8^3 - 8 + 23^3 - 23 + 35^3 - 35 + 21^3 - 21 + 8^3 - 8}{96^3 - 96}} = 12.27$$

(3) 确定 P 值,作出推断结论

当组数 $k=3$ 且每组例数 $n_i \leq 5$ 时,可查 H 界值表(附表 10)得到 P 值;当 $k > 3$ 或 $k=3$ 且最小样本例数 $n_i > 5$ 时, H 近似地服从自由度为 $\nu = k-1$ 的 χ^2 分布,可查 χ^2 界值表得到 P 值。

本例 $k=3$ 且最小样本例数 $n_i > 5$,查 χ^2 界值表, $\nu = k-1 = 3-1 = 2, \chi_{0.01,2}^2 = 9.21, H_c = 12.27 > \chi_{0.01,2}^2$, $P < 0.01$,按照 $\alpha = 0.05$ 检验水准,拒绝 H_0 ,三组血清抗体滴度水平的差别具有统计学意义。

用 Kruskal-Wallis 秩和检验当推断拒绝 H_0 ,接受 H_1 时,只能得出各总体分布不全相同的结论,但不能说明任两个总体分布不同。若要对每两个总体分布做出有无不同的推断,需要作组间的多重比较,具体比较方法请参考相关书籍。

需要注意:非参数检验方法的主要优点是放宽了 t 检验和方差分析的正态分布条件,如果满足参数检验的条件却使用非参数方法会降低检验效能。另外,由于两样本的秩和检验对总体分布的形状差别不敏感,如对总体均数相同、方差不等的正态分布,不能对其分布的形状进行推断,故备择假设 H_1 不能为总体分布不同,而只能写为总体分布的位置不同,其含义是两组数据的大小不同。

(李康,贺佳主编, 2024, p. 106) - Please click Zotero - Refresh in Word/LibreOffice to update all fields.

[BIBLIOGRAPHY] Please click Zotero - Refresh in Word/LibreOffice to update all fields