不同设计和资料类型的秩和检验

作者：张倩 来源：爱科学

本文原地址：https://www.iikx.com/news/statistics/5965.html

本文仅供学习交流之用，版权归原作者所有，请勿用于商业用途！

秩和检验的背景

秩和检验方法最早是由维尔克松提出，叫维尔克松两样本检验法。后来曼—惠特尼将其应用到两样本容量不等(n1 ≠
n2)的情况，因而又称为曼—惠特尼U检验。这种方法主要用于比较两个独立样本的差异。

提出问题

在实践中我们常常遇到以下一些资料，如需比较患者和正常人的血铁蛋白、血铅值、不同药物的溶解时间、实验鼠发癌后的生存日数、护理效果评分等，这类资料有如下特点：

(1)资料的总体分布类型未知;

(2)资料分布类型已知，但不符合正态分布;

(3)某些变量可能无法精确测量。

对于此类资料，除了进行变量变换或t ’ 检验外，可采用非参数统计方法。

参数统计与非参数统计的区别：

参数统计：即总体分布类型已知，用样本指标对总体参数进行推断或作假设检验的统计分析方法。

非参数统计：即不考虑总体分布类型是否已知，不比较总体参数，只比较总体分布的位置是否相同的统计方法。

下面我们将介绍非参数统计中一种常用的检验方法--秩和检验，其中“ 秩 ”又称等级、即按数据大小排定的次序号。上述次序号的和称“ 秩和 ”，秩和检验就是用秩和作为统计量进行假设检验的方法。

不同设计和资料类型的秩和检验

1. 配对比较的资料：

对配对比较的资料应采用符合秩和检验(Sighed rank test)，其基本思想是：若检验假设成立，则差值的总体分布应是对称的，故正负秩和相差不应悬殊。检验的基本步骤为：

(1)建立假设;

H0：差值的总体中位数为0;

h3：差值的总体中位数不为0;检验水准为0.05。

(2)算出各对值的代数差;

(3)根据差值的绝对值大小编秩;

(4)将秩次冠以正负号，计算正、负秩和;

(5)用不为“ 0 ”的对子数n及T(任取T+或T-)查检验界值表得到P值作出判断。

应注意的是当n>25时，可用正态近似法计算u值进行u检验，当相同秩次较多时u值需进行校正。

2. 两样本成组比较：

两样本成组资料的比较应用 Wilcoxon秩和检验，其基本思想是：若检验假设成立，则两组的秩和不应相差太大。其基本步骤是：

(1)建立假设;

H0：比较两组的总体分布相同;

h3：比较两组的总体分布位置不同;检验水准为 0.05。

(2)两组混合编秩;

(3)求样本数最小组的秩和作为检验统计量T;

(4)以样本含量较小组的个体数 n1、两组样本含量之差 n2-n1 及T值查检验界值表;

(5)根据P值作出统计结论。

同样应注意的是，当样本含量较大时，应用正态近似法作u检验;当相同秩次较多时，应用校正公式计算u值。

3. 多个样本比较：

多个样本比较的秩和检验可用 Kruskal-Wallis 法，其基本步骤为：

(1)建立假设;

H0：比较各组总体分布相同;

h3：比较各组总体分布位置不同或不全相同;检验水准为 0.05。

(2)多组混合编秩;

(3)计算各组秩和 Ri;

(4)利用 Ri 计算出检验统计量 H;

(5)查H界值表或利用卡方值确定概率大小。

应注意的是当相同秩次较多时，应计算校正Hc

4. 按等级分组资料或频数表资料：

这类资料的特点是无原始值，只知其所在组段，故应用该组段秩次的平均值作为其秩次，在此基础上计算秩和并进行假设检验，其步骤与两组或多组比较秩和检验相同。需注意的是由于样本含量较多，相同秩次也较多，应用校正后的u值和H值。

说到最后

1. 多个样本两两比较的秩和检验

同样的，多个样本组比较的秩和检验，如拒绝H0，只说明比较各组的总体分布位置不同或不全相同，应在此基础上进行两两比较，常用 Nemenyi 法。

2. 秩和检验的优缺点

秩和检验的优点是(1)不受总体分布限制，适用面广;(2)适用于等级资料及两端无缺定值的资料;(3)易于理解，易于计算。缺点是符合参数检验的资料，用秩和检验，则不能充分利用信息，检验效能低。

3. 应用中的注意事项：

(1)注意应用条件;

(2)编秩时相同值要取平均秩次;

(3)相同秩次较多时，统计量要校正。

更多 统计方法 请访问 https://www.iikx.com/news/statistics/

本文版权归原作者所有，请勿用于商业用途，爱科学iikx.com转发