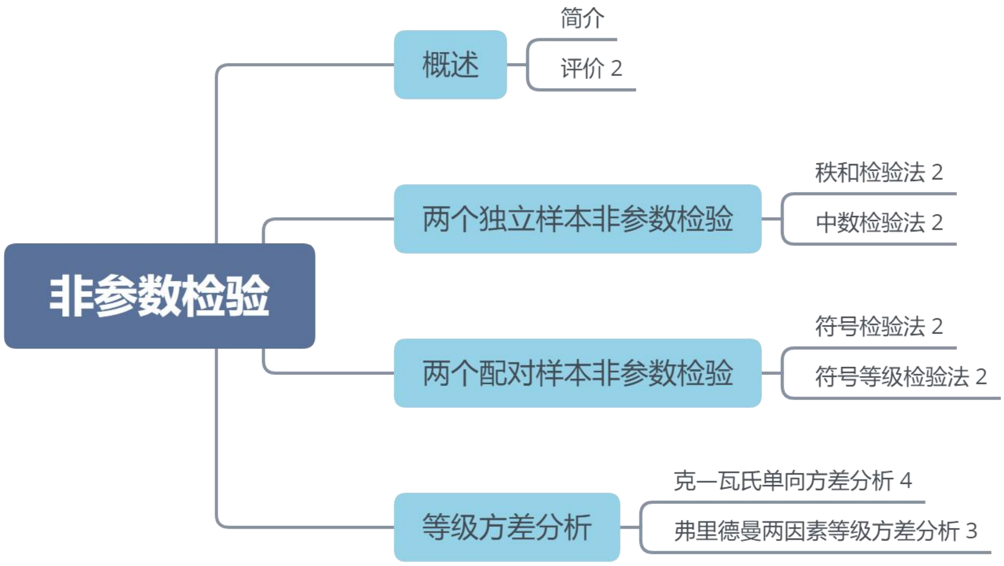
第十一章　非参数检验



一、概述

（一）简介

在实践中，对研究的总体可能知之不多，有时对参数检验中的诸多要求和假定难以满足，这时就应该使用非参数检验。如：斯皮尔曼等级相关、χ^2 检验。

（二）评价

1.优点

（1）不需要严格的前提假设；

（2）特别适合顺序变量；

（3）特别适合小样本，计算很快。

2.缺点

（1）未能利用数据的全部信息；

（2）不能处理交互作用。

二、两个独立样本的非参数检验

（一）秩和检验法

又称“维尔克松两样本检验法”“曼-特尼维尔克松秩和检验”“曼-特尼 U 检验”。

对应参数检验中“独立样本的 t 检验”，且当两独立样本均为顺序变量时也可以使用。其中，秩代表数据经过从小到大排序后的等级，秩和表示等级之和。

1.小样本（$n1 \le 10$, $n2 \le 10$）

（1）将两样本数据混合，从小到大排序，求秩次;

（2）对容量较小的样本求秩和，记为 T；

（3）查表，若 $T \le T1$ 或 $T \ge T2$，则两样本差异显著；若 $T1 < T < T2$，则差异不显著。

2.大样本（$n1>10$，$n2>10$）

T 值分布服从近似正态分布，使用近似 Z 检验：

$$\sigma_T=\sqrt{\frac{n_1n_2(n_1+n_2+1)}{12}}\quad \mu_T=\frac{n_1(n_1+n_2+1)}{2}\qquad Z=\frac{T-\mu_T}{\sigma_T}$$

（二）中数检验法

对应着参数检验中的“两独立样本平均数之差的 t 检验”，用中数作为集中趋势的量度，与秩和法的适用条件基本相同。

（1）将两个样本数据混合，从小到大排列，求混合排列的中数。

（2）分别找出每一样本中大于混合中数及小于混合中数的数据个数，列成四格表对四格表进行卡方检验。

$$\chi^2=N(\sum\frac{f_{oi}^2}{f_{xi}f_{yi}}-1)\qquad\qquad \chi^2=\frac{N(AD-BC)^2}{(A+B)(A+C)(B+D)(C+D)}$$

三、两个配对样本的非参数检验

（一）符号检验法

用于检验两个配对样本分布的差异。

对应参数检验中“配对样本差异显著性 t 检验”，也是以中数作为集中趋势的量度。

1.小样本（对子数 N≤25）

（1）对于每对数据之差，不计大小，只记符号。正号的个数记为 n_+，负号的个数记为 n_-， 0 不记在内。$N=n_++n_-$，$r=\min$（n_+，n_-）。

（2）查符号检验表，若 r 大于表中临界值，则差异不显著，即接受虚无假设（与一般参数检验相反，要特别注意）。可以看出与 n_+与 n_-差得越多，越显著。

2.大样本（对子数 N>25）

n_+和 n_-符合二项分布，N>25 时服从正态分布，故可用 Z 检验。

$$\mu=np=\frac{N}{2}，\quad \sigma=\sqrt{Npq}=\frac{\sqrt{N}}{2}\qquad Z=\frac{r-\mu}{\sigma}$$

为更接近正态分布，常用校正公式：$Z=\dfrac{(r+0.5)-\mu}{\sigma}$

（二）符号等级检验法

又称“符号秩和检验”“维尔克松检验法”“维尔克松 T 检验”，精确度比符号检验法高，在考虑差值的符号的同时还考虑差值大小。

对应参数检验中“配对样本差异显著性 t 检验”。

1.小样本（对子数 N≤25）

（1）一对数据之差按绝对值从小到大排列（0 不参加）；

（2）在等级前添正负号；

（3）求出带正号的等级和（T_+）与带负号的等级和（T_-），取两者之中较小的记作 T，T=min（T_+ ，T_-）；

（4）查表检验，若大于表中临界值，则差异不显著（与一般参数检验相反，要特别注意）。

2.大样本（对子数 N>25）

N>25 时，T 分布接近正态，使用 Z 检验：

$$\mu_T = \frac{N(N+1)}{4}, \quad \sigma_T = \sqrt{\frac{N(N+1)(2N+1)}{24}} \qquad Z = \frac{T-\mu_T}{\sigma_T}$$

四、等级方差分析

当方差分析三大前提不满足时使用。

（一）克-瓦氏单向方差分析

又称“克-瓦氏 H 检验”，对应“完全随机设计”的方差分析。

1.将所有数据混合，排序；

2.求出各个处理内的秩和，记为 R_i；

3.计算 H 值：

$$H = \frac{12}{N(N+1)}\sum_{1}^{K}\frac{R_i^2}{n_i} - 3(N+1)$$

n=某处理组中样本容量，
K=处理数
Ri 为每种处理中的等级和

4.查 H 表，大于临界值说明差异显著。

（二）弗里德曼两因素等级方差分析

对应“随机区组设计”的方差分析。

（1）将每一区组的 K 个数据（K 为实验处理数）从小到大排出等级；

（2）求出每种实验处理 n 个数据等级和，以 Ri 表示;

（3）代人公式

$$\chi_r^2 = \frac{12}{nK(K+1)}\sum R_i^2 - 3n(K+1)$$

n=区组个数，K=处理数
Ri 为每种处理中的等级和

	独立样本均值		相关样本均值	
	参数检验	非参数检验	参数检验	非参数检验
2个组	两独立样本t检验	秩和检验法	两个配对样本分布的差异	符号检验法
	两独立样本之差的t检验	中数检验法	配对样本差异显著性t检验	符号等级检验法
3个及3个以上组	方差分析的完全随机设计	克-瓦氏单向方差分析	随机区组设计	弗里德曼两因素等级方差分析