

	应用数理统计方法
第二章 比较总体大小特征的假设检验	
2.1 假设检验与假设检验方法	
2.2 比较一个或两个总体均值的参数检验	
2.3-2.4 比较两个总体大小的非参数检验	
2.5 比较多个总体大小的非参数检验	
应用举例	

假设检验问题	p.63
▪ 特征比较	
两总体或多总体大小比较	2.2 - 2.5
两总体或多总体离散程度比较	3.1 - 3.2
两总体分布特征比较	3.4
总体分布是否服从特定理论分布	3.3 - 3.4
▪ 影响因素	
方差分析及补充分析	4.1 - 4.4
▪ 变量关系	
相关分析	5.1 - 5.2
回归分析	5.3 - 5.4

2.2.1	
2.2.1 关于总体大小比较的假设检验	
2.2.3 比较一个总体均值与某一特定值的 t-检验	
2.2.4 比较两个独立总体均值的 t-检验	
2.2.5 比较两个对应总体均值的成对数据 t-检验	

总体均值与样本均值	p.84-85
▪ 从一个总体中抽样	
样本均值与总体均值	模拟 $\mu = 45.6, \bar{x} = 35.2 - 66.5$
个体越离散, 偏离远的可能越大	离散程度影响估计的不确定性
▪ 从两个总体中抽样	
样本与总体均值大小关系不同	模拟 $\mu_1 = 43.6, \mu_2 = 48.6$
不能直接比较样本均值	$\bar{x}_1 = 34.2 - 61.0; \bar{x}_2 = 39.5 - 56.5$
个体越离散, 偏离远的可能越大	离散程度影响比较结果

影响样本均值的因素	
▪ 影响两个样本均值大小关系的因素	
总体均值大小与个体离散程度	
▪ 均值大小	
总体均值差别越大, 样本均值差别也越大	
▪ 个体离散	
个体越离散, 样本均值对总体均值偏离远的概率越大	

为什么不能直接比较两个样本均值	
▪ 影响样本均值大小关系的因素	
总体均值大小	总体均值差别越大, 样本均值差别也越大
个体离散程度	个体越离散, 样本均值对总体均值偏离远的概率越大
▪ 举例	
若干黑白棋子, 棋子重量各异	
随机抽取黑白棋子各 20 粒, 称得黑白棋子平均重量分别为 3.2 g 和 3.4 g	
问: 黑白棋子重量有差别吗? 为什么?	
不能直接回答, 要了解棋子重量的离散程度	
如果两者标准差分别为 0.001 g 和 0.002 g 呢?	
如果两者标准差分别为 1.0 g 和 1.1 g 呢?	

假设检验 复习

p.64-65

▪ 基本思路

建立: 总体没有显著差别的原假设

计算: 相伴概率 p , 即原假设成立条件下得到实际观察结果及更极端结果的概率

– 更极端结果: 比实际观测结果更偏离原假设

判断: 相伴概率大小则拒绝原假设, 否则接受原假设

大小比较检验思路

p.84

▪ 比较两总体大小的假设检验

假设: H_0 两总体大小没有显著差别; H_1 两总体大小有显著差别

计算: 总体大小无显著差别时, 得到实际观察结果和差别更大结果的概率 p

判断: 若可能性太小, 则判定两总体大小有显著差异

▪ 参数方法

原假设成立条件下, 影响相伴概率的因素包括:

样本均值差距越小, p 越大

标准差越大, p 越大

即 是否显著取决于均值差/标准差相对关系 两个分布重叠程度

A_1-A_2 比 A_1-A_2 更倾向显著

B_1-B_2 比 A_1-A_2 更倾向显著

大小比较检验思路

p.84

▪ 例: 比较两总体大小的假设检验

假设: H_0 两总体大小没有显著差别; H_1 两总体大小有显著差别

计算: 总体大小无显著差别时, 得到实际观察结果和差别更大结果的概率 p

判断: 若可能性太小, 则判定两总体大小有显著差异

▪ 非参数方法

总体不是正态分布, 均值等参数无意义

利用个体的大小次序信息

秩数据

若两总体大小无显著差异, 共同排序时个体随机出现 如: 红蓝总体个体大小排序

如果观测精度高于秩数据, 可计算准确概率

大小比较检验方法选择

p.87

▪ 目的

确认关注大小特征

▪ 总体个数

确定总体个数

一个, 两个或多个, 取决于研究者的目的

▪ 总体分布特征与样本量大小

是否服从正态分布

样本量大小

参数方法还是非参数方法

影响不同非参数方法功效效率的主要因素

▪ 对立假设范畴及个体对应关系

能否排除一半可能?

数据是否有对应关系?

单侧检验还是双侧检验

一般检验还是成对数据检验

大小比较检验方法

p.84-87

▪ 大小比较假设检验方法

参数方法与非参数方法

一个, 两个或多个总体

多总体大小比较参数方法见第四章

取決于分布特征

两个以上总体可能涉及对应关系

方差分析不限于简单大小比较

总体	总体关系	参数方法	非参数方法举例
一个		t-检验, 正态检验	2.2.3
二个	非对应	t-检验	2.2.4
	对应	成对数据 t-检验	2.2.5
多个	非对应	方差分析	4
	对应	随机区组设计方差分析	4

一个或两个总体的大小比较

p.84-99

▪ 一个总体与已知值比较的参数方法

t-检验与正态检验 后者只是前者样本量趋向无穷的特例, 无实际用途

▪ 两个总体大小比较的参数方法

无对应关系的 t-检验与成对数据 t-检验

总体	总体关系	参数方法	非参数方法举例
一个		t-检验, 正态检验	2.2.3
二个	非对应	t-检验	2.2.4
	对应	成对数据 t-检验	2.2.5
多个	非对应	方差分析	4
	对应	随机区组设计方差分析	4

4 大小比较

2

2.2.3

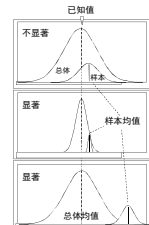
- 2.2.1 关于总体大小比较的假设检验
- 2.2.3 比较一个总体均值与某一特定值的 t-检验
- 2.2.4 比较两个独立总体均值的 t-检验
- 2.2.5 比较两个对应总体均值的成对数据 t-检验

总体均值与已知值比较

p.88-89

- 检验总体均值与某已知值是否有显著差异
相当于判断样本是否来自以该已知值为均值的总体
- 思路
大小: 样本均值距已知值越远, p 值越小
离散: 个体越集中, p 值越小

- 检验方法
t-检验 基于检验统计量 t
正态检验 无实际用途
从标准正态分布中抽样得到 t -分布
正态检验是 t -检验在样本量趋向无穷时的特例



总体均值与已知值比较的 t-检验

p.88-89

$$H_0: \mu = \mu_0$$

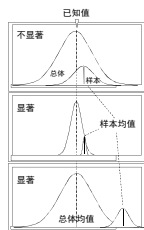
$$H_1: \mu \neq \mu_0 \quad (\mu < \mu_0, \mu > \mu_0) \quad \text{双侧与单侧}$$

$$t_s = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

分子为距离, 分母为标准误差

直接判断 $p < \alpha$

间接判断 $t_s > t_{\alpha/2}(n-1), t_s < -t_{\alpha/2}(n-1)$ 双侧或单侧



2.2.4

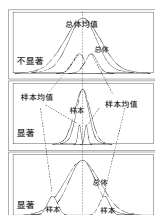
- 2.2.1 关于总体大小比较的假设检验
- 2.2.3 比较一个总体均值与某一特定值的 t-检验
- 2.2.4 比较两个独立总体均值的 t-检验
- 2.2.5 比较两个对应总体均值的成对数据 t-检验

两总体大小比较思路

p.92-94

- 检验两个总体均值是否有显著差异
相当于判断两个样本是否来自均值相同的总体
- 思路
大小: 两样本均值距离越远, p 越小
离散: 个体越集中, p 越小

- 检验方法
t-检验 基于检验统计量 t



两总体大小比较的 t-检验

p.92-94

$$H_0: \mu_1 = \mu_2$$

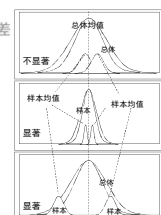
$$H_1: \mu_1 \neq \mu_2 \quad (\mu_1 < \mu_2, \mu_1 > \mu_2) \quad \text{双侧与单侧}$$

$$t_s = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}} \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

分子为距离, 分母包含方差

直接判断 $p < \alpha$

间接判断 $t_s > t_{\alpha/2}(n_1+n_2-2), t_s < -t_{\alpha/2}(n_1+n_2-2)$ 双侧或单侧



2.2.5

2.2.1 关于总体大小比较的假设检验

2.2.3 比较一个总体均值与某一特定值的 t-检验

2.2.4 比较两个独立总体均值的 t-检验

2.2.5 比较两个对应总体均值的成对数据 t-检验

数据的对应关系 复习

p.95

■ 对应关系的利用

用于大小比较,排除规律性个体差异的影响,提高检验功效

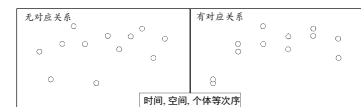
大小比较基于差异/波动相对大小 更多例子参见 2.2.5

■ 个体对应关系举例

时间 某河流上下游 BOD 差别: 按时间随机采样 vs. 按时间同步采样

空间 北京两年 PM2.5 差别: 两年分别随机布点 vs. 两年在相同点位

个体 下页



数据中个体的对应关系 复习

p.95

■ 个体对应关系

受特定因素影响,两个或多个总体中个体具有对应且相同的变化规律

有对应关系的两总体数据构成成对数据 不限于两个总体

这样的对应关系可体现在个体,时间,空间等方面

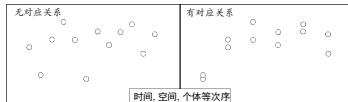
■ 对应总体中个体的数据构成

无对应关系: 总体差异 + 随机波动 个体间差异也是随机的

有对应关系: 总体差异 + 个体间差异 + 随机波动 两总体个体间差异一致

例: 蓝色和棕色两样本比较

■ 重要概念



成对数据举例 复习

■ 据称某药物能控制体重,征集志愿者测试药效

方案一 200 名志愿者,分两个组,分别服用药物和安慰剂 随机分组

方案二 100 名志愿者,分两阶段,分别服用药物和安慰剂 个体随机分阶段

■ 影响体重的因素

药物作用 可能的药物作用,关注的问题

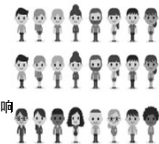
个体差异 志愿者个体间差异,方案二关注每人服药与否的差别,排除个体差异影响

随机效应 纯随机,所有个体自身的随机波动

■ 数据分析

方案一 相对个体差异和随机波动,药效是否显著

方案二 相对随机波动,药效是否显著,排除了个体差异影响



更多成对数据比较例子

■ 某连锁餐馆在北京和南京两地的日销售额有差别吗?

方案一 各随机抽取 30 个门店,随机抽取 20 天 日期随机

方案二 各随机抽取 30 个门店,同步抽取 20 天 日期同步,空间不能同步

■ 影响销售额的因素

两地差异 消费习惯,消费能力,开店密度,其它选择

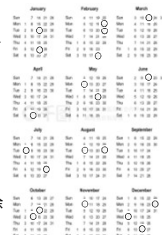
时间差异 日变化,周末,节假日等

随机效应 纯随机波动

■ 数据分析

方案一 相对日期差异与随机效应,两地是否有差别

方案二 相对随机效应,两地是否有差别,日期差异被排除



更多成对数据比较例子

■ 某市禁止机动车使用喇叭,禁令实施前后环境噪声有差别吗?

方案一 分别在禁令前后随机布设 100 个样点,进行一周噪声测定

方案二 随机布设 100 个样点,分别在禁令前后进行一周噪声测定

■ 影响机动车对环境噪声贡献的因素

样点差异 不同样点位置间差异

实施禁令 汽车使用喇叭的影响

随机效应 纯随机波动

■ 数据分析

方案一 相对样点差异与随机波动,禁令是否有效

方案二 相对随机波动,禁令是否有效,排除样点间差异



成对数据大小比较思路 参数方法

p.95-99

成对数据构成 – 绿色与蓝色两个样本

非成对数据: 总体差异 + 随机波动

成对数据: 总体差异 + 个体差异 + 随机波动

个体差异混入随机波动

对应个体差异一致

成对数据大小比较思路 参数方法

p.95-99

成对数据构成 – 绿色与蓝色两个样本

非成对数据: 总体差异 + 随机波动

成对数据: 总体差异 + 个体差异 + 随机波动

个体差异混入随机波动

对应个体差异一致

检验策略

普通比较方法会损失对应关系信息

策略: 成对数据检验

转换: 差值总体均值与 0 比较

蓝绿两总体: 3.1±0.94, 3.5±0.95

逐对差减排除个体差异 → 黑色总体

差减总体 0.40±0.13, 变异仅随机波动

成对数据大小比较 t-检验

p.95-99

假设

H0: $\mu_1 = \mu_2$

H1: $\mu_1 \neq \mu_2, \mu_1 < \mu_2, \mu_1 > \mu_2$ 双侧与单侧检验

计算

个体差 D_i

判断

一个总体与 0 比较的 t-检验 参见 2.2.3

比较 D_i 与 0 是否有显著差异

应用数理统计方法

第二章 比较总体大小特征的假设检验

2.1 假设检验与假设检验方法

2.2 比较一个或两个总体均值的参数检验

2.3-2.4 比较两个总体大小的非参数检验

2.5 比较多个总体大小的非参数检验

应用举例

两总体大小比较的非参数方法

p.100-102

方法举例

两个总体大小比较的非参数方法 次序及节号与教科书不完全相同

非对应数据 Mann-Whitney U 检验, 随机化检验

对应数据 Wilcoxon 秩和检验, 随机化检验

总体	总体关系	参数方法	非参数方法举例
一个		t-检验, 正态检验	2.2.3
二个	非对应	t-检验	2.2.4
	对应	成对数据 t-检验	2.2.5
多个	非对应	方差分析	4
	对应	随机区组设计方差分析	

两个非对应总体大小比较的非参数方法

p.102

推荐方法

Mann-Whitney U 检验 大样本量非正态分布总体, 样本量大小无确切边界

随机化检验 小样本量非正态分布总体, 数据精度最好优于秩数据

其它方法

Wilcoxon 秩和检验 与 Mann-Whitney U 检验无差别

Kolmogorov 单侧检验 仅限单侧

中位数检验 二分类类型数据, 功效很低

检验方法	数据类型	功效效率
Mann-Whitney U 检验	顺序量	随样本量增大趋向 95%
Wilcoxon 符号检验	顺序量	随样本量增大趋向 95%
Kolmogorov 单侧检验	离数量	随样本量减小趋向 96%
随机化检验	离数量	随样本量减小趋向 100%
中位数检验	类型量	从 n=6 的 95% 到大样本量的 63%

4 大小比较

5

两个对应总体大小比较的非参数方法

p.120

▪ 推荐方法

Wilcoxon 成对数据加符秩检验 大样本量非正态分布总体, 样本量无确切边界
成对数据随机化检验 小样本量非正态分布总体, 数据精度优于秩数据

▪ 其它方法

Walsh 检验 数据要求高于 Wilcoxon 检验
符号检验 二分类类型数据, 功效极低

检验方法	数据要求	功效效率
Wilcoxon成对数据加符秩检验	连续量或离散量	大样本趋向95%
Walsh 检验	顺序量, 个体差对称分布	大样本趋向95%
成对数据随机化检验	测量水平高	小样本趋向100%
符号检验	仅需了解数据对大小关系	63-95%

2.3.1

2.3.1 Mann-Whitney U 检验

2.3.2 Wilcoxon 成对数据加符秩检验

2.3.3 随机化检验

2.3.4 成对数据随机化检验

Mann-Whitney U 检验思路

p.102-104

▪ 方法

用途: 两个非对应非正态分布总体大小比较
适用: 大样本量数据 样本量增加, 功效效率趋向 95%

▪ 思路

两样本一起求秩, 若两总体大小无显著差别, 两样本的秩随机出现
计算任一样本的秩和 两样本秩和信息相同
若原假设成立, 秩和不应太大或太小 原假设成立条件下的小概率事件
实际观测结果与更极端情况 区分单侧与双侧

极端, 小概率 ↑ 秩和 = 1+2+...+ 26 = 351
秩和 = 1+2+3+5+8+...+53 = 693

极端, 小概率 ↓ 秩和 = 31+32+...+56 = 1130

Mann-Whitney U 检验方法

p.102-104

▪ 假设

H0: 两总体大小没有显著差异
H1: 两总体大小有显著差异 双侧检验
H1: 一总体显著大于另一总体 单侧检验

▪ 计算

基于秩和直接计算相伴概率 p 或检验统计量 Us

▪ 拒绝原假设条件

直接判断 $p < \alpha$
间接判断 $Us > U_{0.5\alpha[n_1, n_2]}$ 或 $Us > U_{\alpha[n_1, n_2]}$ 双侧或单侧

2.3.2

2.3.1 Mann-Whitney U 检验

2.3.2 Wilcoxon 成对数据加符秩检验

2.3.3 随机化检验

2.3.4 成对数据随机化检验

Wilcoxon 成对数据加符秩检验思路

p.120-122

▪ 方法

适用: 非正态分布成对数据大小比较
特点: 大样本连续量或离散量 样本量减小, 功效效率趋向 95.5%

▪ 思路

若大小无显著差异, 成对数据差别方向及大小均应随机出现
逐对计算数据差, 按差的绝对值求秩, 计算正负差的秩和 R+, R-
若原假设成立, 正负差秩和差别不应太小 同时利用差别方向和大小信息

样本一	12	22	11	34	74	24	54	32	43	19	31
样本二	11	17	21	30	80	43	27	33	25	26	21
差值	1	5	-10	4	-6	-19	27	-1	18	-7	10
绝对值秩	1.5	4	7.5	3	5	10	11	1.5	9	6	7.5
秩和差别太大则拒绝两总体没有差异的原假设 $R^+ = 30$ $R^- = 36$											

成对数据随机化检验方法

p.128-130

假设

H_0 : 两总体大小无显著差异 H_1 : 两总体大小有显著差异/一总体大

检验

求数据对差 D_i , 所有 D_i 组合按极端性顺序排序 ΣD_i 递降或递增

计算全部可能的组合数

$$C = 2^n$$

否定域为原假设成立条件下的极端组合 $J = \alpha C$, 单双侧检验分别在一侧或两侧

若观察结果在否定域中, 拒绝原假设 可计算精确概率 p

数据		D = A - B		值和差 极端次序	
A	5 11 12 16 22 $n_1=5$				
B	19 20 23 29 11 $n_2=5$				
D=A-B	-14 -9 -11 -13 -11				

D = A - B		值和差 极端次序	
A小的极端情况	-14 -9 -11 -13 -11	-59	1
	-14 -9 -11 -13 -11	-40	2
	-14 -9 -11 -13 -11	-36	3
	-14 -9 -11 -13 -11	-36	4
	-14 -9 -11 -13 -11	-32	5
	-14 -9 -11 -13 -11
B小的极端情况	-14 -9 -11 -13 -11	+59	32

随机化检验/成对数据随机化检验的否定域

p.111,130

与参数方法类比

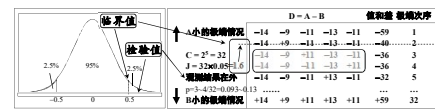
随机化/成对数据随机化检验 列举全部可能组合, 确定否定域 (计算相伴概率)


两总体大小比较 t-检验 计算检验统计量, 确定临界值 (计算相伴概率)

单侧与双侧检验

非参数检验 双侧与单侧检验分别将否定域放在两端或一端

参数检验 双侧与单侧检验分别考察分布两侧或一侧概率





应用数理统计方法

第二章 比较总体大小特征的假设检验

2.1 假设检验与假设检验方法

2.2 比较一个或两个总体均值的参数检验

2.3-2.4 比较两个总体大小的非参数检验

2.5 比较多个总体大小的非参数检验

应用举例

多总体大小比较的非参数方法

p.136-139

多总体大小比较的非参数方法

Kruskal-Wallis 检验 非对应总体

Friedman 检验 对应总体

总体	总体关系	参数方法	非参数方法举例
一个		t-检验, 正态检验	2.2.3
二个	非对应	t-检验	2.2.4
	对应	成对数据 t-检验	2.2.5
多个	非对应	方差分析	4
	对应	随机区组设计方差分析	4

多总体大小比较的非参数方法

p.139

推荐方法

Kruskal-Wallis 检验 无对应关系的多个非正态分布总体, 样本量可以不相等

Friedman 检验 有对应关系的多个非正态分布总体, 等样本量

其它方法

推广的中位数检验 类型量独立总体, 功效低

推广的 Mann-Whitney 检验 限于三个总体, 局限性

Jonckheere 检验 多总体

检验方法	研究对象	数据要求	功效效率
Kruskal-Wallis 检验	独立总体	顺序量	95.5%
推广的中位数检验	独立总体	类型量	低
Friedman 检验	对应总体	顺序量	接近参数方法
推广的 Mann-Whitney 检验	三总体	顺序量	
Jonckheere 检验	多总体	顺序量	

2.5.1

2.5.1 Kruskal-Wallis 检验

2.5.3 Friedman 秩方差分析

Kruskal-Wallis 检验思路

p.139-142

▪ 方法

多个非对应非正态总体大小比较 对应正态分布的单因子方差分析 参见第四章
与 Mann-Whitney U 类似 两总体推广到多总体, 不再有单双侧之分

▪ 思路

考察个体在所有样本中的出现次序 所有样本共同排序求秩
计算不同总体中个体的秩和 每个样本独立求秩和
若原假设成立, 秩和差别不应太大 类似 U-检验, 据秩和差别计算相伴概率

排序
求秩和

Mann-Whitney U

X₁₁ X₁₂ X₁₃ ... X_{1m}

X₂₁ X₂₂ X₂₃ ... X_{2m}

排序
求秩和

Kruskal-Wallis

X₁₁ X₁₂ X₁₃ ... X_{1m}

X₂₁ X₂₂ X₂₃ ... X_{2m}

X₃₁ X₃₂ X₃₃ ... X_{3m}

X₄₁ X₄₂ X₄₃ ... X_{4m}

Kruskal-Wallis 检验方法

p.139-142

▪ 假设

H₀: 多总体大小没有显著差异
H₁: 多总体大小有显著差异 无单双侧之分

▪ 计算

多个样本共同求秩, 每个样本独立求秩和 $\sum R_j$
计算相伴概率 p 或检验统计量 H

▪ 拒绝原假设的条件

直接判断 $p < \alpha$
间接判断 $H > \chi^2_{\alpha[k-1]}$

排序
求秩和

Kruskal-Wallis

X₁₁ X₁₂ X₁₃ ... X_{1m}

X₂₁ X₂₂ X₂₃ ... X_{2m}

X₃₁ X₃₂ X₃₃ ... X_{3m}

X₄₁ X₄₂ X₄₃ ... X_{4m}

2.5.3

2.5.1 Kruskal-Wallis 检验

2.5.3 Friedman 秩方差分析

Friedman 检验思路

p.149-150

▪ 方法

多个非正态分布对应总体大小比较 等样本量
与 Kruskal-Wallis 检验类似 基于排序和秩和, 无单侧之分

▪ 思路

按对应关系组分别排序 与 Kruskal-Wallis 不同
每个样本独立求秩和 与 Kruskal-Wallis 相同
若原假设成立, 秩和差别不应太大 据秩和差别计算相伴概率

排序
求秩和

Friedman

X₁₁ X₁₂ X₁₃ X₁₄ X₁₅

X₂₁ X₂₂ X₂₃ X₂₄ X₂₅

X₃₁ X₃₂ X₃₃ X₃₄ X₃₅

X₄₁ X₄₂ X₄₃ X₄₄ X₄₅

Friedman 检验方法

p.149-150

▪ 假设

H₀: 多总体大小没有显著差异
H₁: 多总体大小有显著差异 无单双侧之分

▪ 计算

按对应关系组求秩, 每个样本独立求秩和
计算相伴概率 p 或检验统计量 χ^2

▪ 拒绝原假设的条件

直接判断 $p < \alpha$
间接判断 $\chi^2 > \chi^2_{\alpha[k-1]}$

排序
求秩和

Friedman

X₁₁ X₁₂ X₁₃ X₁₄ X₁₅

X₂₁ X₂₂ X₂₃ X₂₄ X₂₅

X₃₁ X₃₂ X₃₃ X₃₄ X₃₅

X₄₁ X₄₂ X₄₃ X₄₄ X₄₅

两个多总体大小非参数比较方法对比

p.136-139

▪ 数据差异

Friedman 检验 无对应关系, 样本量可以不等
Kruskal-Wallis 检验 有对应关系, 等样本量

▪ 计算

排序: 全部样本数据一起排序, 或按对应关系独立排序
秩和: 计算每个样本的秩和 原假设成立条件下, 秩和不应太大或太小
类比: t-检验和成对数据-t 检验 直接比较或按对应关系差减排除个体间差异

排序
求秩和

Friedman

X₁₁ X₁₂ X₁₃ X₁₄ X₁₅

X₂₁ X₂₂ X₂₃ X₂₄ X₂₅

X₃₁ X₃₂ X₃₃ X₃₄ X₃₅

X₄₁ X₄₂ X₄₃ X₄₄ X₄₅

排序
求秩和

Kruskal-Wallis

X₁₁ X₁₂ X₁₃ ... X_{1m}

X₂₁ X₂₂ X₂₃ ... X_{2m}

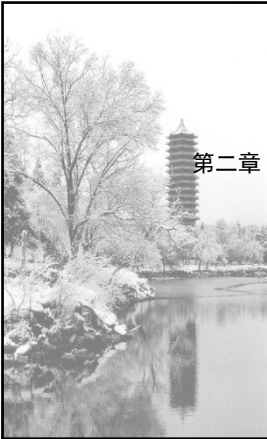
X₃₁ X₃₂ X₃₃ ... X_{3m}

X₄₁ X₄₂ X₄₃ ... X_{4m}

4 大小比较

9

大小比较假设检验方法汇总				p.84-87
<div>▪ 大小比较</div> <div>参数方法与非参数方法 取决于分布</div> <div>一个, 两个或多个总体 两个以上总体可能涉及对应关系</div> <div>多总体大小比较的方差分析见第四章 方差分析不限于简单大小比较</div>				
总体	总体关系	参数方法	非参数方法举例	
一个		t-检验, 正态检验	2.2.3	
二个	非对应	t-检验	2.2.4	Mann-Whitney U 检验 2.3.1 随机化检验 2.3.3
	对应	成对数据 t-检验	2.2.5	Wilcoxon 加秩秩检验 2.3.2 成对数据随机化检验 2.3.4
	非对应	方差分析	4	Kruskal Wallis 检验 2.5.1
	对应	随机区组设计方差分析		Friedman 秩方差分析 2.5.3

应用数理统计方法	
	第二章 比较总体大小特征的假设检验
	2.1 假设检验与假设检验方法
	2.2 比较一个或两个总体均值的参数检验
	2.3 比较两个总体大小的非参数检验
	2.5 比较多个总体大小的非参数检验
应用举例	

应用实例 鱼油治疗干眼症效果评估

DEAMSRG, NEJM 2018

▪ 研究问题

问题: 通常认为鱼油对眼睛有益, 主要有效成份是 Ω -3

方法: 募集 535 名干眼症患者参与研究, 分两组开展一年双盲实验

分别服用 Ω -3 补充剂和橄榄油安慰剂, 测定眼表疾病指数 OSDI 变化

统计: 两个独立总体大小比较 t-检验

假设: $H_0 \mu_1 = \mu_2$ $H_1 \mu_1 < \mu_2$

▪ 结果与讨论

结果: 不能拒绝原假设, 未发现 Ω -3 补充剂有显著作用

讨论: 成对数据设计更合理, 补充剂和安慰剂使用次序随机

CONCLUSIONS

Among patients with dry eye disease, those who were randomly assigned to receive supplements containing 3000 mg of n-3 fatty acids for 12 months did not have significantly better outcomes than those who were assigned to receive placebo. (Funded

An abstract graphic consisting of a rectangular frame. Inside the frame, at the top, is the text 'DEAMSRG, NEJM 2018' in a bold, sans-serif font. Below this text is a small, stylized illustration of a person's head and shoulders, facing forward. The rest of the frame is filled with a dense, repeating pattern of small, light-colored text, which appears to be a snippet of the article's abstract or a list of references, rendered in a very small font size.

应用实例 智能手机对焦虑感的影响

Gao et al., Peer J., 2016

▪ 研究问题

问题: 焦虑是否与智能手机使用行为有关

方法: 调查了 127 名用户, 据焦虑指数将人群分为高、低两组

统计: 用 Mann-Whitney U-检验比较两组人群使用行为差异

针对行为特征 电话、短信、拍照、健康程序 等使用频数

假设: H_0 无显著差异 H_1 有显著差异 不同特征分别比较

▪ 结果与讨论

结果: 高焦虑组 电话、短信、拍照少, 使用健康类程序多

讨论: 方法: 样本量

方法: 对比行为而非焦虑指数

方法: 其它分析方法, 如相关?



应用实例 某化合物在两种材料中的扩散速率比较

Frauenhofer et al., JCA, 2021

▪ 研究问题

问题: 不同温度条件下, 化合物在两种材料中的扩散系数是否相同

方法: 不同温度条件下做对比实验 扩散系数不服从正态分布

统计: 用 Wilcoxon 加秩秩检验 逐对比较

假设: H_0 两种材料无差别, H_1 两种材料有差别 不同温度实验分别检验

▪ 结果与讨论

结果: 在各种温度条件下均显著

讨论: 双因子?

扩散速率与温度回归比较 见第五章

p-values from the Wilcoxon signed-rank tests (paired data) comparing various physical quantities measured with HMW-PDMS and LMW-PDMS.

	Temperature (°C)			
Compared experimental quantities	60	70	80	90
Specific retention volumes	0.031	0.031	0.031	0.036
Molar enthalpies of sorption	0.058			
Interaction parameters	0.036	0.063	0.073	0.106
Diffusion coefficients	0.031	0.031	0.031	0.031
Activation energy of diffusion	0.036			

应用实例 废水中污染物的去除率评估

p.112

▪ 研究问题

问题: 特定设施在去除废水中某污染物的效果 出水浓度是否显著低于进水

方法: 测定结果, 进水 19, 20, 24, 29 mg/L 出水 0, 11, 12, 16, 22 mg/L

统计: 小样本量, 选择 随机化检验, 单侧检验

假设: H_0 进出水无差异, H_1 出水低于进水

▪ 结果与讨论

计算: 组合数 $C = 126$, 否定域 $J = 6.3$

结果: 观测结果第三极端 $p = 0.024$

拒绝原假设

讨论: 虽然功效效率不低, 但功效不高

进水取值	出水取值	值和差	极端性次序
0, 11, 12, 16, 19	20, 22, 24, 29	37	1
0, 11, 12, 16, 20	19, 22, 24, 29	35	2
0, 11, 12, 16, 22	19, 20, 24, 29	31	3
0, 11, 12, 19, 20	16, 22, 24, 29	29	4
0, 11, 12, 19, 22	16, 20, 24, 29	25	5
0, 11, 12, 20, 22	16, 19, 24, 29	21	6
0, 11, 12, 20, 24	16, 19, 22, 29	17	7

应用实例 英文课程效果评估

p.112

研究问题

问题: 考察某英文课程效果

方法: 分别在授课前后对 8 名学生进行测试 难度相似

课程前得分 63, 42, 74, 37, 51, 43, 80, 82

课程后得分 82, 69, 73, 43, 58, 56, 76, 85 得分差 -19, -27, 1, -6, -7, -13, 4, -3

统计: 成对数据随机化检验, 单侧

假设: H_0 不变, H_1 上升

结果与讨论

计算: $C = 28 = 256$, $J = 256 \times 0.05 = 13$

结果: 否定域 13 种组合, 极端性排序 6, $p = 0.025$

结论: 拒绝原假设, 课程有效

抽样方式	值和 极端性 差 次序
-19 -27 -1 -6 -7 -13 -4 -3	-80 1
-19 -27 +1 -6 -7 -13 -4 -3	-78 2
-19 -27 -1 -6 -7 -13 -4 +3	-74 3
-19 -27 -1 -6 -7 -13 +4 -3	-72 4.5
-19 -27 +1 -6 -7 -13 -4 +3	-72 4.5
-19 -27 +1 -6 -7 -13 +4 -3	-70 6
-19 -27 -1 +6 -7 -13 -4 -3	-68 7
-19 -27 -1 -6 +7 -13 -4 -3	-66 9
-19 -27 +1 +6 -7 -13 -4 -3	-66 9
-19 -27 -1 -6 -7 -13 +4 -3	-66 9
-19 -27 +1 -6 +7 -13 -4 -3	-64 11.5
-19 -27 +1 -6 -7 -13 +4 +3	-64 11.5
-19 -27 -1 +6 -7 -13 -4 -3	-62 13

应用实例 不同环境中微塑料含量比较

Li et al., Water Res. 2021

研究问题

问题: 不同环境介质中微塑料含量差别 浓度 个/ m^3

方法: 收集全国 709 样点微塑料数据 淡水, 海水, 淡水沉积物, 海洋沉积物, 表土

统计: 用 Kruskal-Wallis 检验 比较五个总体差异

用 Wilcoxon 加符秩检验 比较淡水与海水差异

假设: H_0 无显著差异, H_1 有显著差异

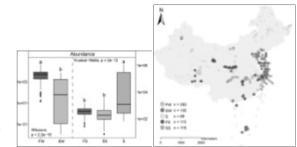
结果与讨论

结果: 不同环境介质有显著差别

海水与淡水有显著差别

讨论: 淡/海水样品不成对

不能用 Wilcoxon sign-ranks test



应用实例 野生动物、鱼和家畜消费量比较

Rao et al, Environ. Manag. 2010

研究问题

问题: 狩猎, 生计和保护关系研究

方法: 调查缅甸野生动物保护区 242 户野生动物、鱼和家畜食用量 kg/户-月

研究不同家庭三类食物消费量是否有差别, 种植季和狩猎季分别检验

方法: 有对应关系的多总体非参数检验 Friedman 检验 比较三个对应总体

假设: H_0 无显著差别; H_1 有显著差别

结果与讨论

结果: 两季均有显著差异 $p < 0.0001$

结论: 不同家庭三类食物消费量不同

Table 3 Consumption patterns of wild meat, fish and livestock (kg/household/month) for 242 households in 13 villages during the planting and hunting seasons

Variable	Minimum	Maximum	Mean	SD
Planting Season				
Wild meat	0.000	25.000	2.142	4.305
Fish	0.000	23.830	2.467	3.268
Livestock	0.000	11.170	1.464	1.616
Hunting Season				
Wild meat	0.000	25.830	2.008	3.612
Fish	0.000	25.000	2.822	3.752
Livestock	0.000	9.750	2.154	1.809

Friedman's test: $P < 0.0001$

要点

概念

样本均值与总体均值

正态分布与非正态分布

大小比较方法

选择依据 目的, 总体个数, 分布, 样本量, 对立假设, 个体对应关系

参数方法 - 样本均值差与离散程度 t-检验

非参数方法 - 数据排序 U, 随机化, Wilcoxon, Fruskal-Wallis, Friedman

特殊条件

重要方法 成对数据, 单侧检验

谢谢

