

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 比较两个独立总体大小的非参数检验 2.4 比较两个对应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 2.1.1

- 2.1.1 假设检验
- 2.1.2 假设检验思路与方法
- 2.1.3 假设检验的风险和功效
- 2.1.4 参数与非参数方法
- 2.1.5 假设检验方法选择

统计推断 复习 p.2-3

■ 统计推断

在概率论的基础上,根据对<u>样本</u>的观测判断<u>总体</u>的某些特征 包括参数估值和假设检验

■ 参数估值

描述总体统计特征,即根据<u>样本统计量对总体统计量</u>进行估计如根据样本均值估计总体均值

■ 假设检验

据样本观测结果,对针对总体的假设成立与否作出判断如 两总体大小是否相同,某因素是否有显著影响,两变量是否相关

随机变量的主要特征 复习

p.20

■ 随机变量的特征

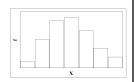
宏观规律的表现

• 统计学特征

分布 不同取值 (范围) 出现的概率

大小 分布中心在数轴上的位置,取值大小

离散 个体到分布中心的距离,聚集/分散程度



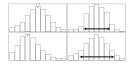
假设检验问题 大小与离散比较举例

- 个别方法之前已介绍 如 异常剔除,独立性检验
- 总体大小特征比较

湖泊 A 和 湖泊 B 中鱼的平均重量是否有差别? 两总体大小比较大、中、小三类湖泊湖水 pH 值是否有差别? 三总体大小比较

■ 总体离散特征比较

湖泊中草食性鱼与杂食性鱼的重量变幅是否一致? 两总体离散程度比较湖水 pH 值和湖水天然有机物的变异是否有差别? 两总体离散程度比较



假设检验问题 分布检验举例

■ 经验分布与理论分布比较

某鱼种大小是否服从正态分布, 抑或对数正态分布? 正态检验 湖泊丰水期水面面积是否服从正态分布?

某种鱼的雌雄比是否为1?

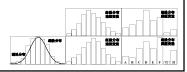
二项检验

■ 经验分布与经验分布比较

两种鱼的重量分布是否一致?

拟合度检验

两个湖泊中八个主要鱼种的比例是否有差异?



假设检验问题 影响因素检验举例

■ 単因素

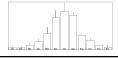
■ 多因素

汇流面积和深度不同湖泊水面面积有差别吗? 汇流面积与深度对面积的影响 不同水化学类型湖泊鱼的生长速率是否不同? 类型亚类对生长速率的影响

■ 方差分析

研究影响因素

与<u>回归分析</u>的关系



假设检验问题 变量关系研究举例

■ 两变量共变关系

■ 预测模型

能否/如何根据湖泊蓄水量估算鱼的总重量? 单自变量回归模型能否/如何根据湖泊多种参数预测鱼的总数? 多自变量回归模型

■ 相关与回归

目的和用法不同

<u>方差</u>,<u>相关</u>和<u>回归</u>分析的关系



假设检验内容安排

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 总体分布特征比较 3.3 - 3.4

■ 影响因素

方差分析及补充分析

■ 变量关系

相关分析 5.1-5.2 回归分析 5.3-5.4 2.1.2

2.1.1 假设检验

2.1.2 假设检验思路与方法

2.1.3 假设检验的风险和功效

2.1.4 参数与非参数方法

2.1.5 假设检验方法选择

假设检验 p.64-66

4.1 - 4.4

■ 假设检验

在概率论的基础上,对是否接受特定统计假设进行判断的手段

■ 方法

据<u>样本观测结果</u>,对针对<u>总体</u>的假设成立与否作出判断 仅在是与否两种可能性之间做出判断

假设检验的基本思路

p.64-65

■ 基本思路

建立: 总体<u>没有显著差别</u>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

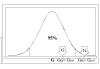
■ 更极端结果举例

异常 更极端情况包括比可疑值更偏离分布中心的所有取值

判断:相伴概率太小则拒绝原假设,否则接受原假设

独立 观察游程为 28 或 3, 更极端情况: <u>29-30</u> 或 <u>1-2</u>





三假设检验

假设检验思路举例 黑白棋子

袋中有棋子共 $\underline{200}$ 粒, \underline{k} 點 黑白各 $\underline{100}$ 粒, 混匀后随机取出 $\underline{20}$ 粒 若白色 18 粒, 黑色 2 粒, 相信黑白一样多的说法吗?判断有把握吗? 若 白色 11 粒, 黑色 9 粒, 相信黑白一样多的说法吗? 判断有把握吗?



假设检验思路举例 黑白棋子

袋中有棋子共 200 粒, 据称黑白各 100 粒, 混匀后随机取出 20 粒 若白色 18 粒, 黑色 2 粒, 相信黑白一样多的说法吗?判断有把握吗? 若 白色 11 粒, 黑色 9 粒, 相信黑白一样多的说法吗? 判断有把握吗?

■ 猜测

前者黑白一样多的可能性 极小, 有很大把握 后者黑白一样多的可能性 比前者大的多,有把握吗?



假设检验思路举例 黑白棋子

■ 问题

袋中有棋子共 200 粒, 据称黑白各 100 粒, 混匀后随机取出 20 粒 若白色18粒,黑色2粒,相信黑白一样多的说法吗?判断有把握吗? 若 白色 11 粒, 黑色 9 粒, 相信黑白一样多的说法吗? 判断有把握吗?

■ 检验

假设 袋中黑白棋子相等 无差别的原假设 计算 白色≥18 或≥11 的概率 若原假设成立的前提 判断 概率太小则判定黑白不同 拒绝原假设

两种判断的可靠性完全不同 见 2.1.2 两类风险

否则 接受原假设

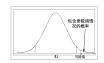


假设检验思路举例 异常值剔除

p.52

■ 问题

从正态分布总体中采样,判断是否有异常值 可疑值为最大和最小值 如果可疑值出现在分布的尾端, 会是异常吗? 远离均值, 属于总体概率小 如果可疑值出现在分布的中部,会是异常吗?接近均值,属于概率概率大



假设检验思路举例 异常值剔除

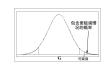
p.52

■ 问题

从正态分布总体中采样,判断是否有异常值 可疑值为最大和最小值 如果可疑值出现在分布的尾端、会是异常吗? 远离均值、属于总体概率小 如果可疑值出现在分布的中部, 会是异常吗? 接近均值, 属于概率概率大

■ 猜测

前者是异常的可能性 很大, 有很大把握 后者是异常的可能性 较小,仍可能不属于这个总体,有把握吗?



假设检验思路举例 异常值剔除

■ 问题

从正态分布总体中采样,判断是否有异常值 可疑值为最大和最小值 如果可疑值出现在分布的尾端、会是异常吗? 远离均值、属于总体概率小 如果可疑值出现在分布的中部,会是异常吗? 接近均值,属于概率概率大

■ 检验

假设可疑值属于总体,不是异常 无差别的原假设 计算可疑值及更偏值的出现概率 如果原假设成立: 可疑值属于该总体

判断概率太小则判定为异常 拒绝原假设 否则

接受原假设

两种判断的可靠性完全不同 参见2.1.2两类风险

包含更极端情况的概率

假设检验思路举例 独立性检验

p.49

从连续变量总体中抽样 n=30, 判断采样是否有独立性 是否受相邻个体影响? 将连续数据转换为二分类类型数据 大于中位数计正,否则计负,计算游程数 如果游程数为 28,3 或 14, 样本有独立性吗?

假设检验思路举例 独立性检验

p.49

从连续变量总体中抽样 n=30, 判断采样是否有独立性 是否受相邻个体影响? 将连续数据转换为二分类类型数据 大于中位数计正,否则计负,计算游程数 如果游程数为 28,3 或 14,样本有独立性吗?

■ 猜测

前两种情况有独立性的可能性 极小, 有很大把握 后一种情况有独立性的可能性 较高, 有把握吗?

假设检验思路举例 独立性检验

p.49

■ 问题

从连续变量总体中抽样 n=30, 判断采样是否有独立性 是否受相邻个体影响? 将连续数据转换为二分类类型数据 大于中位数计正, 否则计负, 计算游程数 如果游程数为 28,3 或 14, 样本有独立性吗?

■ 检验

假设 采样有独立性,不受其它个体影响 无差别的原假设

计算 游程数 ≥28, ≤3 或 ≤14 的概率 如果原假设成立: 有独立性

判断 概率太小则判定不具独立性

拒绝原假设,判断可靠性高

否则

接受原假设,判断可靠性低

假设检验思路 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 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 判断: 相伴概率太小则拒绝原假设, 否则接受原假设

■ 换个说法

如果面对的是一个<u>小概率</u>事件 原假设成立条件下,p值太小,

则判定该事件不应该发生 原假设不成立

检验假设 p.64-65

■ 统计假设

用统计语言表达, 可检验是否成立的情景 基于概率计算 针对总体

用符号或文字表述

符号为总体参数 μ , σ , γ , ρ , β 等,均值,标准差,偏度-峰态系数,相关系数,斜率

■ 原假设与对立假设

待判断的两种可能

Ho 原假设 检验的直接对象

Hi 对立假设 拒绝原假设情况下接受的结果

假设的表达

p.66-69

以下例子中出现的若干具体方法及符号,内容之后陆续介绍,这里关注表达 与算术符号不同,≠和=分别代表有无显著差异

■ 用符号表达

 $H_0 \mu_1 = \mu_2 H_1 \mu_1 \neq \mu_2$ 两总体大小是否有显著差异, μ 为<u>总体</u>均值 H_0 ρ = 0 H_1 ρ ≠ 0 两变量是否显著相关, ρ 为 <u>总体</u>相关系数

■ 用文字叙述

Ho 两总体大小无显著差异; Hi 两总体大小有显著差异 Ho 总体服从正态分布; HI 总体不服从正态分布

检验假设举例

- 日常生活中的判断 排队选择,路径选择,食物选择...
- 黑白棋子 Ho 黑白棋子个数相同; Hi 黑白棋子个数不同 黑白棋子个数是否相同
- 异常值剔除 Ho可疑值不是异常; Hi 可疑值是异常

可疑值是否来自该总体

■ 独立性检验

Ho 个体具有独立性; Hi 个体不具独立性 与独

与独立总体无差别

检验假设举例

■ 主要特征比较

两类湖泊湖水 pH 是否一致? $H_0 \ \mu_1 = \mu_2$ $H_1 \ \mu_1 \neq \mu_2$ 总体均值 湖水硬度和 pH 变<u>异程度</u>相同吗? $H_0 \ V_{P_1} = V_{P_2}$ $H_1 \ V_{P_1} \neq V_{P_2}$ 总体变异系数 湖泊面积是否为正态分布? $H_0 \ \gamma_1 = 0, \gamma_2 = 0, \ H_1 \ \gamma_1 \neq 0, \gamma_2 \neq 0$ 总体偏峰系数

■ 影响因素研究

■ 变量关系研究

湖泊径流量与面积有关吗? $H_0\,\rho=0\,,H_1\,\rho\neq0$ 总体相关系数能否据湖泊蓄水量预测鱼产量? $Y=a+bX,\,H_0\,\beta=0\,,H_1\,\beta\neq0$ 总体斜率

原假设与对立假设

p.64-65

■ 两者关系

包括所有可能

相互矛盾

对立假设远比原假设包含更多可能性,两者不能互换

■ 举例

黑白棋子 相等或不等 H_0 100:100 vs. H_1 0:200 - 99:101, 101:99 - 200:0 大小比较 相同或不同 H_0 μ_1 = μ_2 差别为 0, H_1 μ_1 \neq μ_2 差别为 0 以外所有可能值

■ 例外

大小比较 $H_0 \mu_1 = \mu_2$ 差别为 0 $H_1 \mu_1 = \mu_2 + C$ 差别为 C

假设检验思路 复习

p.64-65

■ 基本思路

建立: 总体没有显著差别的原假设

<u>计算:相伴概率</u> 及 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 判断: 相伴概率太小则拒绝原假设, 否则接受原假设

描述统计量与检验统计量

p.67 总

■ 描述统计量

用于参数估值

根据样本获得样本参数,据此估计总体参数 点估计与区间估计如:算术均值,标准差,几何均值等及其置信区间

■ 检验统计量

用于假设检验

如 异常值剔除 Grubbs 检验中的 \underline{G} , t-检验中的 \underline{t} , 方差分析中的 \underline{F} 等

用于计算相伴概率p 统计软件直接给出计算结果

少数情况可直接计算 如新冠阳性概率计算,黑白棋子判断,随机化检验 参见 2.3.4

总体分布与检验统计量分布

p.67

■ 总体中个体的经验分布

随机变量中个体出现规律 总体特征之一

决定检验方法选择 参数方法与非参数方法

■ 检验统计量的理论分布

用于假设检验,定义检验值与p的关系 相伴概率计算的基础

如 Grubbs 检验中的标准正态分布

t-检验中的 t-分布

方差分析中的 F-分布

三假设检验

5

取值与概率的对应关系

p.71

■ 概率密度函数 取值区间与概率有对应关系

■ 相伴概率

原假设成立条件下,得到实际观察结果和更极端结果的概率如: b 1,000 次硬币,950 次正,更极端情况 $951,952,\ldots,1,000$ 假设: 500:500 如: 异常值剔除 Grubbs 法中,可疑值及更偏离均值的取值 假设: Grubhar 因为 <math> Grubhar 假设: Grubhar 假设: Grubhar 假设: Grubhar 份。 Grubhar 代码: Grubhar 假设: Grubhar Grubhar



假设检验思路 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 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

<u>判断</u>: 相伴概率<u>太小则拒绝</u>原假设, 否则接受原假设

显著性水平 *p.71*

■ 定义

拒绝客观上正确原假设的最大概率,又称拒绝水平,记为 α 主观选择,<u>在检验前确定</u>以确保检验的客观性 拒绝原假设:表达为在 α 水平下<u>显著</u>即有显著差异 一般取值:0.01, 0.05, 0.10

■ 与相伴概率的关系 2 为实际概率

α 是拒绝原假设的概率边界

假设检验的判断

p.71-72

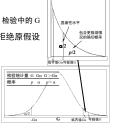
■ 直接判断

计算相伴概率, 若 $p < \alpha$, 则拒绝原假设可了解拒绝原假设情况下的确切错误率

■ 间接判断

若不能计算p值,可计算检验统计量 如 Grubbs 检验中的 G 查临界值,若统计量大于临界值 如 $G \ge G\alpha$,则拒绝原假设此时仅知道错误率小于 α , \overline{AC} 解确切值

■ 直接判断与间接判断的对应关系 计算值与相伴概率对应 G~p 临界值与显著性水平对应 Gα~α



双侧检验与单侧检验

p.68

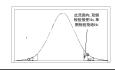
■ 双侧检验

如果有差异,可能出现在两个方向上,α <u>均分</u>在分布的两端 如 黑棋多或白棋多,高异常或低异常,游程太大或太小

■ 单侧检验

若有差异,只可能出现在一个方向上 前提是可根据已知事实排除另一半可能 对立假设包含的可能为双侧检验的一半 α 在分布的一端 如 黑棋只可能比白旗多,只可能高异常,只可能出现极小游程

在较少可能性中选择, <u>拒绝</u> H_0 的可能性增加



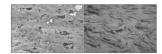
双侧与单侧检验问题举例

■ 双侧检验

大型湖泊和小型湖泊湖水硬度是否有差别?差别可以是双向的 杂食性鱼和草食性鱼大小是否有差异? 若有差别,两种鱼都可能偏大

■ 单侧检验

流域降水量增加是否导致湖泊面积增加? 降水增加,湖泊面积不可能减少 鱼体内农药含量与湖水农药含量是否相关?不可能负相关



假设检验步骤

p.73

■ 基本思路

建立: 总体<u>没有显著差别</u>的原假设

计算: 相性概率 p 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 判断: 相伴概率太小则拒绝原假设

判断: 怕件概率太小则<u>拒绝</u>原假设

■ 步骤

直接判断和间接判断 差别在于能否获得确切的相伴概率









确定显著 性水平α 计算 检验统计量

it it

界值比较

2.1.3

- 2.1.1 假设检验
- 2.1.2 假设检验思路与方法
- 2.1.3 假设检验的风险和功效
- 2.1.4 参数与非参数方法
- 2.1.5 假设检验方法选择

统计推断的可能错误

p.70

■ 参数估值的可能错误

根据<u>样本</u>参数估计<u>总体</u>参数

区间估计: 将点估计的可能错误表达为置信区间

■ 假设检验的可能错误

检验假设基于<u>样本</u>,针对<u>总体</u>作出判断

判断结果可能不对 不能穷尽所有个体

接受和拒绝原假设的错误率不同 原假设和对立假设包含的可能性不同

结论的把握不同

假设检验思路举例 黑白棋子

■ 问题

袋中有棋子共 $\underline{200}$ 粒, <u>据称</u>黑白各 $\underline{100}$ 粒, 混匀后随机取出 $\underline{20}$ 粒 若 白色 $\underline{18}$ 粒, 黑色 $\underline{2}$ 粒, 相信黑白一样多的说法吗?判断<u>有把握</u>吗?若 白色 $\underline{11}$ 粒, 黑色 $\underline{9}$ 粒, 相信黑白一样多的说法吗?判断<u>有把握</u>吗?

■ 检验

假设 袋中黑白棋子相等 <u>无差别</u>的原假设 计算 白色 ≥18 或 ≥11 的概率 若原假设成立的前提

判断 概率太小则判定黑白不同 拒绝原假设 否则 接受原假设

两种判断的可靠性完全不同 见 2.1.2 两类风险



假设检验的不确定性

p.70

■ 两类错误

原假设<u>客观上</u>有两种可能: 正确或错误 检验者不了解的事实 主观判断也有两种可能: 接受或拒绝 都可能错, 但错误率不同 四种可能, 两种是错的

第 I 类错误: <u>拒绝</u>客观上<u>正确</u>的原假设 错误率 α , 主观设定 第 II 类错误: 接受客观上错误的原假设 错误率 β , 客观事实

■ 举例

黑白棋子 接受与拒绝的风险 异常值剔除 接受与拒绝的风险

接受原假设 拒绝原假设 假设客观上正确 正确: 1-α 错误 1: α 假设客观上错误 错误 11: β 正确: 1-β

第Ⅱ类错误率

p.70

■ 特占

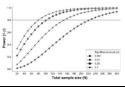
不可计算 已知分布的检验统计量是基于原假设构建的 不可控制 控制了 α

远高于 α 对立假设包含的可能性远多于原假设

■ 第 II 类错误率的理论表征

可构建假设总体, 模拟 β 随 α 和 n 变化的一般趋势 方法选择的依据

不能用于实际计算



假设检验判断的不确定性

p.70

■ 判断的不确定性

任何判断都没有绝对把握由于不能穷尽全部个体

拒绝和接受原假设的不确定性不同 由于两种假设包含的可能性完全不同

例: 从 2 种或 256 种颜色中猜

拒绝原假设 不确定性低,可控,p或 $\leq \alpha$ 接受原假设 不确定性高,不可控,未知, β

■ 假设检验的结论

拒绝原假设,可以有明确的结论 已知,且可控的错误率p或 $\leq \alpha$ 接受原假设,不能下明确的结论 未知,且很高的错误率 β

实例 药物疗效大小比较

- 关于假设检验结论的例子 强调结论,具体方法在后续各节中介绍
- 研究

问题 对比两种药物治疗某疾病的效果 方法 募集两组各 100 名患者, 分别用两种药物治疗 数据 采集两组人群的疗效

■ 结果与结论

假设 $H_0 \mu_1 = \mu_2; H_1 \mu_1 \neq \mu_2$ 两种药物疗效无/有差别结果 p > 0.05,接受原假设

结论 没有观察到两种药物的差别



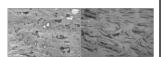
应用实例 鱼重与径流量间关系

■ 研究

问题 湖泊中鱼的重量与径流量有关吗? 方法 随机抽50个湖泊,每个湖采集50条鱼,测量湖泊年均径流量和鱼重

■ 结果与结论

假设 $H_0 \rho = 0; H_1 \rho \neq 0$ 总体相关系数 $\rho \to 0$ 则不相关,否则相关 结果 $\rho = 0.15, >0.05$ 接受原假设,重量与径流量不相关 结论 <u>看不到</u>鱼的重量与径流量之间有相关关系 无明确结论



实例 湖泊流量的正态检验

■ 研究

问题 某区域湖泊径流量是否服从正态分布 使用参数方法的前提方法 随机抽取 100 个湖泊, 测量湖泊年均径流量

■ 结果与结论

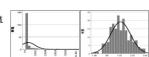
假设 $H_0 \gamma_1 = 0, \gamma_2 = 0; H_1 \gamma_1 \neq 0, \gamma_2 \neq 0$ 偏度系数 γ_1 和峰态系数 γ_2 结果 $p_1 < 0.05, p_2 < 0.05,$ 结果 $\underline{L$\underline{Z}$}$, 拒绝原假设, 分布偏离正态

结果 若对原始数据做对数变换,检验结果均不显著

结论 不是正态分布 此结论可靠

结论 是对数正态分布 此结论不可靠

讨论 现实和妥协,误判的风险



假设检验的功效

p.74

■ 假设检验的功效

拒绝客观上错误原假设的可能性 $1-\beta$ 对比第二类错误率 β 不能就具体问题计算功效 不能计算 β

■応田

可模拟获得非参数方法功效随 α 和 n 的变化趋势根据样本量大小选择方法依据



影响检验功效的因素

p.74-75

■ 方法差别

参数方法的功效总是不低于相应的非参数方法 同类问题,如两总体大小比较同类检验的非参数方法功效不同,与<u>样本量</u>有关 如 U-检验和随机化检验

■ 其他因素

样本量大小 样本量越大,代表性越好,置信区间越小,功效越高显著性水平 α 越高,拒绝客观上错误原假设机会越大,但不应为提高功效改变 α 对立假设 包含可能性越小,功效越高,如:

H₀ μ = 10, H₁ μ ≠ 10 双侧检验

H₀ μ = 10, H₁ μ < 10 <u>单侧检验</u>, 排除一半可能

 $H_0 \mu = 10$, $H_1 \mu = 30$ <u>已知</u>对立假设成立条件下的检验统计量<u>分布</u>

课堂练习

■ 举例

举一个单侧假设检验问题的例子 最好是非专业问题

■ 要求

就近随意分组, 每组讨论一个例子 可以针对随机变量的任何特征

要求 简述问题

给出原假设和对立假设 假定的 p 值, 结论

字大些

2.1.4

- 2.1.1 假设检验
- 2.1.2 假设检验思路与方法
- 2.1.3 假设检验的风险和功效
- 2.1.4 参数与非参数方法
- 2.1.5 假设检验方法选择

总体分布与参数估值 复习

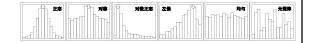
p.36-39

■ 总体大小表征

正态分布总体,用算术均值表征 代表集中趋势 非正态分布,若可做正态变换,仍可用算术均值 变换与逆变换 非正态分布且不能或不便变换,用中位数 算术均值不能代表集中趋势

■ 总体其它特征表征

离散, 偏度, 峰态, 相关, 截距, 斜率等计算均以算术均值为基础 总体参数 $\mu,\sigma,\sigma^2,\gamma_1,\gamma_2,\rho,\beta,\alpha,...$



总体分布与特征比较

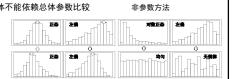
p.76-78

■ 总体大小比较

正态分布总体,可用基于均值的方法比较大小 比较集中趋势可变换为正态分布的总体,可用类似方法比较大小 数据变换非正态分布且不能或不便变换,不能依赖均值比较 非参数方法

■ 总体其它特征比较

正态分布总体可用基于均值的方法比较 非正态分布总体不能依赖总体参数比较 $\underline{\mu},\,\sigma,\,\sigma^2,\,\gamma_1,\gamma_2,\,\rho,\,\beta,\,\alpha,\,...$



假设检验的参数与非参数方法

p.76

■ 参数方法

基于参数的经典方法 类似参数估值,基于以均值为基础的参数 如 异常值的 Grubbs 检验

■ 非参数方法

不依赖参数的无分布方法 基于秩数据或精确概率计算

如 个体独立性波动游程检验

如 黑白棋子判断



参数与非参数检验方法对比

p.76

■ 参数与非参数方法的差别 本质差别: 是否基于总体参数 其它差别: 与总体参数有关的差别 附表

■ 方法选择

正态分布总体: 优先选择功效高的参数方法 不绝对,如大样本量 U-检验 非正态分布总体: 正态变换后用参数方法,或直接用非参数方法 类似参数估值

	参数	非参数
检验方法	经典方法	无分布方法
总体参数	基于总体参数	不依赖总体参数
变量类型	连续量,高散量	顺序量, 类型量
信息利用	充分	仅利用顺序信息
检验功效	相对高	相对低
样本要求	尚	低
异常干扰	强	46

非参数方法的功效效率

p.77

■ 非参数方法的检验功效 1 - β

同一问题,一般仅有唯一的参数方法,但可能有不止一个非参数方法 非参数方法的检验功效不可能高于相应的参数方法 信息利用不充分 对同类问题,不同非参数检验方法的检验功效不同 如两个总体大小比较

■ 功效效率

非参数方法对应于参数方法的相对功效 $\leq 100\%$ 相同功效,非参数方法需较大 n 计算: 达到同等功效的参数与非参数方法样本量比 $100~\mathrm{Np/Nn}$

举例 实现相同功效,参数方法 Np=20,非参数方法 Nn=25

非参数方法功效效率 = $100\% \times 20/25 = 80\%$

选择方法的依据,不能直接计算 因为 β 未知

只可模拟获得一般性趋势 如 n 趋向无穷时, U-检验功效效率趋向 95%

2.1.5

- 2.1.1 假设检验
- 2.1.2 假设检验思路与方法
- 2.1.3 假设检验的风险和功效
- 2.1.4 参数与非参数方法
- 2.1.5 假设检验方法选择

假设检验方法的前提

p.77-78

■ 基本要求

采样随机性 所有统计方法的前提,不可检验

个体独立性 所有统计方法的前提, 波动游程检验功效很低, <u>第二类错误率</u>很高通过可靠的实验设计和采样方法保证

■ 参数方法的要求

分布正态性 所有参数方法的前提,有时可用正态变换获得 方差同质性 方差分析的不同水平,回归分析的不同自变量,一般 n 太小,很难检验 方差加和性 方差分析的不同因子,方差分析的基础,可用来考察交互作用

数据中个体的对应关系

p.95

■ 个体对应关系

受特定因素影响,两个或多个总体中<u>个体</u>具有<u>对应</u>的共同变化规律 有对应关系的两总体数据构成<u>成对数据</u> 不限于两个总体 这样的对应关系可体现在<u>个体</u>、时间、空间等方面

■ 对应总体的数据构成

数据的对应关系

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 大小比较是基于<u>差异/波动</u>相对大小

在成对数据大小比较方法中介绍更多实例 参见 2.2.5, 2.4, 2.5.3

■ 个体对应关系举例

个体 100 只服药小鼠和 100 只未服药小鼠对比; 100只小鼠服药前后的差别时间 比较某年某河流上下游 BOD, 按时间随机采样或同步采样空间 比较两年各随机抽 100 个城市的 PM2.5; 随机抽 100 个城市比较两年 PM2.5

三假设检验

10

成对数据举例

■ 据称某药物能控制体重,征集志愿者测试药效 方案- 200 名志愿者,分两个组,分别服用药物和安慰剂 分组随机 方案二 100 名志愿者,分两阶段,分别服用药物和安慰剂 两阶段随机

■ 影响体重的因素

药物作用 药物可能的作用,关注的问题

个体差异 志愿者个体差异,<u>方案二</u>考察个体用药前后关系,可排除个体差异影响

随机效应 纯随机,各组各阶段所有个体之间

■ 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

TITITIII

检验方法选取原则

p.79-80

■ 检验目的

总体 根据研究目的确定总体及总体个数 特征, 变量 大小, 离散, 分布

■ 检验功效

原则上选择功效高的方法,功效差别不大时例外 如选择大样本量 U-检验

■ 其他因素

数据特点: 分布特征, 个体对应关系 如非参数检验, 成对数据比较

方法假设: 对立假设覆盖的可能性 单侧槽

样本量: 影响非参数方法的功效 如 U-检验和随机化检验的选择 保守性: 保守性/严格性, 相当于改变 α 如异常值 Grubbs 检验和 ι 检验

其它: 等价方法, 无概率意义的简便方法 如拟合度卡方检验和 G-检验

方法汇总表 p. 81-83 表 2-4 ■ 检验目的 秤本独立性检验 特征比较 大小, 离散, 分布 平方和相乘性检验 Patra 核酸 * Chauvernet 核酸 * Grubbe 核膜 c-核腺 Dixon 核腺 影响因素 方差分析 变量关系 相关,回归 总体均似与特定值比较 正高松級。 1-检验 其他检验 如预处理等 四独立总体大小比较 ■ 其他因素 两相关总体大小比较 成对数据 = 检验 总体个数 1个,2个或多个 个体关系 是否有对应关系 假设范畴 单侧或双侧 其它 如样本量大小 ■ 参数与非参数方法

