

数据=(原始)分数
 总体: N
 样本: n M s
 描述总体/样本的数值: 参数/统计量
 相关法&实验F(控制组&实验组)&非实验F
 (准-非SYF)自变量: 至少2个值. 性别男/女
 称名(房间号)/顺序(学历/中杯)-条形图
 等距(相对零点. 10)/等比(绝对零点. 亮度)量
 表-直方图(精确上下限)&多边形图
 集中趋势: 均值&中数(偏态&极端&不确定值&
 空端&顺序)&众数(称名&离散&形状)
 正偏: 众<中<均
 变异性: 全距(精确界限)&标准差&方差(离均差
 平方和SS的均值)
 自由度: 样本中可以自由变化的分数组个数
 |概率质量函数PMFy[0, 1]离散. 取值为概率
 |概率密度函数PDFy[0, inf)面积为概率
 |l0 <- seq(55, 145, by = 0.1)
 |f_l0 <- dnorm(l0, mean = 100, sd = 15)
 |plot(l0, f_l0, type = 'l')
 |累计分布函数CDF. a>b则ya>=yb
 |对于二项分布, 只要p不为0或1, 那么只要
 |充分大, 就可以用正态分布进行近似
 |p: 当H0为真时, 观察到当前样本数据或比
 前样本更极端数据的概率
 |对任意给定的显著性水平 ($\alpha > 0$), 只要
 |使样本容量n充分大且效应量(effect size
 |不为0, 就能检验出显著差异
 |标准正态分布拐点: $z=1/-1$
 |z/t表选择较大的值: 由于在p值表中找不到

样本均值的分布: 所有来自总体的样本量为n的随机样本的样本均值的分布/集合
中心极限定理: 对于任何均值为 μ , 标准差为 σ 的总体, 样本大小为n的样本均值的分布的均值(μ的期望值)为 μ , 标准差(μ的标准误, σ/\sqrt{n})和 σ 的平均差异)为 σ/\sqrt{n} , 且当n趋近于inf时接近正态分布(总体正态/ $n >= 30$)
大数法则: 样本量n越大, \bar{M} 越可能接近 μ

假设检验: 使用样本数据来评估一个关于总体参数的假设

虚无假设 H_0 : 自变量(处理)对因变量BL没有作用
水平: H_0 为真时极不可能的样本结果的概率值(犯第一类错误-在 H_0 为真时拒绝 H_0 的概率),决定拒绝域. 小的 水平意味犯DYLCLW概率小, 但假设检验需要更多来自研究结果的证据, 处理效应需要较大

第二类错误：没有拒绝错误的虚无假设-处理效应存在时假设检验不能识别它。概率为
增加样本容量可同时减小

在拒绝域: 拒绝虚无假设, 备择假设在一定的概率下正确, 不是必然正确

不能证实虚无假设或得到不存在处理效应的结论，只是认为没有充分的证据证明存在处理效应

结果显著：当 H_0 为真，极不可能的结果发生，即
结果足以拒绝 H_0
—检验前提 α 的值不被处理器改变。随机样本。对

Z检验前提： μ 的值不被处理改变&随机样本&独立观察&正态样本分布
效应：对处理效应大小的具体测量Cohen's d

效应: 对处理效应大小的具体测量Cohen's
0.2小效应&0.5中等效应&0.8大效应
检验力: 检验正确拒绝一个错误的H₀的能力

检验力：检验正确拒绝一个错误的 H_0 的能力
检验确定一个处理效应真实存在的概率
随效应增加，样本量增加，水平增加，双侧

随效应增加，样本量增加，
检验变单侧，检验力增加

估计标准误：未知时真实标准误的估计
t统计量：未知时被用来检验关于总体未知

统计量：未知时被用未检验关于总体未知参数的假设，与 z 相似但分母为估计标准误差分子分母都变化，变异性更大（ z 只分子变）只需要 H_0 和一个来自未知总体的样本前提：样本的值相互独立&（差异量）总体正态方差小，样本量大，则 t 大，效应大&易显著&置信区间窄（做出更精确的估计）

解释的变异百分比
样本量不影响效应大小Cohen's d完全无，小
置信区间：一个以样本统计量为中心的范围内
的值。样本统计量应接近其对应的总体参数，故
可合理认为参数值落在置信区间内

无偏估计：统计量期望等于总体参数
z分数相同不代表排名相同
 $\text{Var}(X) = E(X - E(X))^2$ 计算的是总体方差而非样本方差