

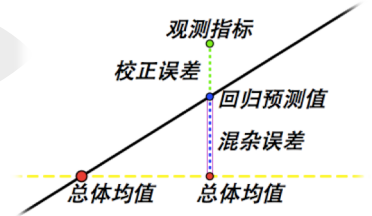
协方差分析

analysis of covariance ANOCVA

结合线性回归与方差分析，将未加或难以控制的混杂因素 X 看做协变量 (covariate)，建立应变量 Y 随协变量 X 变化的线性回归关系，并利用该回归关系消除混杂因素 X 对观测指标 Y 的影响，再对各组观测指标 Y 的修正均值 (adjusted means) 进行方差分析。其实质是从 Y 的总平方和中减去协变量 X 对 Y 的回归平方和，对残差平方和进行方差分析，以求更好地评价处理因素的效应。

应用条件：

- 各观察变量相互独立，服从正态分布，各样本所在总体具有方差齐性；
- 并且各总体存在应变量对协变量相同的线性回归关系，即各样本回归系数均有统计学意义并且相互差异不具有统计学意义。



【 k 组独立样本，每组有 n 对双变量观测值 (X, Y) ，其中 X 为协变量， Y 为应变量，比较各组应变量均值是否相等。】
($k=1, 2, \dots, i, \dots$; $n=1, 2, \dots, j, \dots$)

1. 求出 X, Y, XY 的误差变异；

$$SS_{eXX} = \sum_{i=1}^k \sum_{j=1}^n (X_{ij} - \bar{X}_i)^2; \quad SS_{eYY} = \sum_{i=1}^k \sum_{j=1}^n (Y_{ij} - \bar{Y}_i)^2; \quad SS_{eXY} = \sum_{i=1}^k \sum_{j=1}^n (X_{ij} - \bar{X}_i)(Y_{ij} - \bar{Y}_i)$$

2. 求出公共回归系数 (common regression coefficient)，以及误差项回归平方和与残差平方和，对回归关系进行检验；

$$b_c = \frac{SS_{eXY}}{SS_{eXX}}$$

$$SS_{e(\text{regression})} = \frac{SS_{eXY}^2}{SS_{eXX}}, \quad \nu_{e(\text{regression})} = 1; \quad SS_{e(\text{residual})} = SS_{eYY} - \frac{SS_{eXY}^2}{SS_{eXX}}, \quad \nu_{e(\text{residual})} = nk - 3$$

$$F = \frac{SS_{e(\text{regression})} / \nu_{e(\text{regression})}}{SS_{e(\text{residual})} / \nu_{e(\text{residual})}}$$

3. 求出 X, Y, XY 的总变异，并求出 Y 校正后的总变异；

$$SS_{tXX} = \sum_{i=1}^k \sum_{j=1}^n (X_{ij} - \bar{X})^2; \quad SS_{tYY} = \sum_{i=1}^k \sum_{j=1}^n (Y_{ij} - \bar{Y})^2; \quad SS_{tXY} = \sum_{i=1}^k \sum_{j=1}^n (X_{ij} - \bar{X})(Y_{ij} - \bar{Y})$$

$$SS_{t(\text{residual})} = SS_{tYY} - \frac{SS_{tXY}^2}{SS_{tXX}}$$

4. 求出校正后的处理组间变异，进行方差分析；

$$SS_{b(\text{adjusted})} = SS_{t(\text{residual})} - SS_{e(\text{residual})}; \quad \nu_{b(\text{adjusted})} = k - 1$$

$$F' = \frac{SS_{b(\text{adjusted})} / \nu_{b(\text{adjusted})}}{SS_{e(\text{residual})} / \nu_{e(\text{residual})}}$$

5. 建立校正回归方程，求各处理组的校正均值。

$$Y'_i = \bar{Y} - b_c(\bar{X}_i - \bar{X})$$