

因果推断

2025/8/26

1 PSM

倾向评分匹配。将每个实验组的样本与具有相似特征的未接受处理的样本进行匹配，人为构建一个对照组。

选择协变量，常用逻辑回归计算每个用户划分到实验组的概率（即，倾向评分，PS）。

匹配：实验组用户和未被处理的用户的 PS 频次分布，看是否有足够的用户重叠。有足够的用户交叉重叠时，用 KNN 开始匹配，进行 1:1 匹配，将成功配对的用户从原有池子剔除。

判断匹配效果：用标准均差法 SMD 评估匹配后的用户质量。理想情况下接近零。

$$\text{SMD} = \frac{\bar{X}_{\text{treat}} - \bar{X}_{\text{control}}}{\sqrt{\frac{s_{\text{treat}}^2 + s_{\text{control}}^2}{2}}}$$

2 DiD

双重差分法：通过一定方法（PSM）找到与实验组相似的对照组人群