

优达学城数据分析师纳米学位

A/B 测试项目

试验设计

指标选择

由于过滤学生每周投入时间是在点击“开始免费试学”之后，因此进行试验后点击“开始免费试学”按钮的唯一 cookie 数量没有显著的变化，可以保证实验组和对照组总体上相同。但因为每周投入<5 小时的学生不会完成登陆并开始试学，所以完成登陆并参加免费试学的用户 id 的数量会明显的减少。14 天后仍参与课程的用户 id 的数量会有小幅度的减少。而如 cookie 的数量、用户点击次数由于没有被筛选，因此不会发生明显变化。

此外，点击概率是由 cookie 数量、点击次数生成的一个变量，那么预期点击概率也不会发生变化，因此将点进概率列为不变指标

选择不变指标

基于以上分析，cookie 的数量、用户点击次数以及点击概率作为不变指标。

选择评估指标

在发生变化的指标中，尽管用户 id 的数量预期会有减少，但由于用户在点击“开始免费试学”的时候，并不需要登陆，因此无法了解被每周学习时间过滤掉的 cookie 的比率变化。而总转化率可以衡量点击按钮后最终登陆并参加试学的用户比率。净转化率可以衡量点击后的用户有多少最终发生了付费。留存率可以由总转化率和净转化率计算得出，预期在试验后可能会有所上升。

因此选择总转化率以及净转化率以及留存率作为评估指标。

预期实验结果：

实验组与对照组的总转化率之差具有统计和实际显著性，且实验组显著低于对照组。同时两组净转化率没有显著差异。留存率在试验后预期会有所增加。

测量标准偏差

对于每个评估指标，说明你是否认为分析估计与经验变异是类似还是不同（如果不同，在时间允许的情况下将有必要进行经验估计）。简要说明每个情况的理由。

计算标准偏差

总转化率 SD = 0.0202

净转化率 SD = 0.0156

留存率 SD = 0.0549

由于总转化率和净转化率的分母都是 cookie 浏览量，引流单位就是分析单位，因此分析估计与经验变异是类似的。

对于留存率而言，引流单位与分析单位都是用户 id，因此分析估计与经验变异是类似的。

规模

样本数量和功效

不使用 Bonferroni 校正，所需网页访问数为 4741212。

持续时间和曝光比例

如果采用全部三个指标的话，实验所需的页面访问数为 4741212，在全部投入使用的环境下，需要 100 天以上。三个月以上的时间过长，用户的学习效应也会对试验存在影响。在不考虑留存率的情况下，所需网页访问数为 685325，考虑到页面的改变不会影响后台数据库，因此其风险并不是很大。为了不使实验持续天数过长，所以将全部流量投入使用。

试验分析

合理性检查

对于每个不变指标，对您在 95%置信区间下期望观察到的值、实际观察的值及指标是否通过合理性检查给出结论。

Cookies 数量

置信区间：下限 0.4988 上限 0.5011 观察值 0.5006 通过了完整性检验

点击数

置信区间：下限 0.4958 上限 0.5041 观察值 0.5004 通过了完整性检验

点击概率

置信区间：下限 -0.0013 上限 0.0013 观察值(diff value) 0.0001 通过了完整性检验

结果分析

效应大小检验

总转化率

置信区间 下限 -0.0291 上限 -0.0120，统计显著并且实际显著

净转化率

置信区间 下限 -0.0116 上限 0.0019，统计上不显著且实际也不显著

符号检验

总转化率

P-value 0.0026 具有统计显著性

净转化率

P-value 0.6776 不具有统计显著性

汇总

试验中未使用 Bonferroni 校正，由于 Bonferroni 校正主要适用于 n 次独立检验。但是试验中的总转化率和净转化率并非是独立的，而是相关联的，Bonferroni 校正过于保守，会导致计算出的所需网页访问数过多，因此不使用 Bonferroni 校正。

建议

总的来看，实验中的不变指标通过了完整性检验。从分析结果来看，实验组相比于控制组的总转化率显著降低，具有统计显著性和实际显著性。而净转化率两组的差值不具有统计和实际显著性，根据此处的计算结果(-0.0116, 0.0019)，也就是说有很大的概率净转化率会减少，并且有一定的概率净转化率的减少会超过实际显著性 0.0075。因此不能说明“降低的程度不大”。所以不建议启动。

跟进试验

在后续试验中，在假定不启动“免费试学筛选器”试验的情况下，设计为完成免费试学的同学发送优惠券的试验。

。试验假设优惠券对完成免费试学的用户有较大的激励作用，提高了学生付费继续学习的意愿，从而减少了受挫的学生数量。

选择用户 id 作为转移单位，因为试验需要将用户而非事件分配至控制组和对照组，同时试验发生在用户完成试学的时候。选择用户 id 可以唯一的代表参加试学的用户。没有登录试学的用户也会产生 Cookie，同时一个用户会产生多个 Cookie。因此 Cookie 不是合适的引流单位。

不变度量。Cookie 的数量、用户 id 的数量、点击次数、点进概率、总转化率。

实验选择净转化率和留存率作为评估度量。比较实验组和对照组的净转化率和留存率可以分别测量点击免费试学、完成登陆的学生中，最终完成试学并付费的学生数量是否减少。