

# 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

##### 不变指标:

Cookie 的数量: 即查看课程概述页面的唯一 cookie 的数量。

点击次数: 即点击“开始免费试用”按钮(在免费试用屏幕触犯前发生)的唯一 cookie 的数量。

点进概率: 即点击“开始免费试用”按钮的唯一 cookie 除以查看课程概述页面的唯一 cookie 的数量所得的结果。

##### 评估指标:

总转化率: 即完成登录并报名参加免费试用的用户 id 的数量除以点击“开始免费试用”按钮的唯一 cookie 的数量所得的结果。

净转化率: 即在 14 天期限结束后仍然参加(并至少进行了一次支付)的用户 id 的数量除以点击“开始免费试用”按钮的唯一 cookie 的数量所得的结果。

##### 原因解释:

1. Cookie 的数量和点击次数是在试验之前就被采集的度量, 而点进概率是由 cookie 的数量和点击次数相除所得, 所以这三个指标作为不变指标。
2. 用户 id 的数量是在试验之后被采集的度量, 但是实验组和对照组的 cookie 不一定相同, 导致用户 id 数量的不同可能会因为 cookie 不同引起, 所以不将它作为评估指标。
3. 留存率在试验之后被采集, 且可以验证假设结果, 理论上可以作为评估指标, 但这个评估量需要的页面浏览量过大, 导致试验持续时间过长(具体数值后续说明), 所以也不将它作为评估指标。
4. 总体转化率和净转化率, 在试验之后被采集, 可以验证假设结果, 且需要的页面浏览量合适, 将其作为评估指标。

##### 期望的试验结果:

试验组总转化率显著小于控制组, 且试验组净转化率与控制组无显著性不同。

### 测量标准偏差

总转化率: 0.0202

净转化率: 0.0156

两个评估指标的分析变异与经验变异相同, 因为试验的引流单位是唯一 cookie, 两个评估指标的分析单位也是唯一 cookie, 分析单位与引流单位相同。

### 规模

#### 样本数量和功效

在分析阶段不会使用 Bonferroni 校正, 因为总转化率和净转化率并不独立, 而 Bonferroni 校正用于多次独立试验犯第一类错误概率的矫正, 在这里使用 Bonferroni 校正过于保守。

试验需要的页面浏览量为 685325(如果选取留存率为评估指标则需要页面浏览量 4741213)。

### 试验风险

此试验的风险很小。

1. 试验仅仅是在学生点击“开始免费试学”后多弹出一个询问框，对流程的改动很小，不会对用户产生感情上的冲击，用户也不需要花长时间去适应页面的改变。
2. 此试验与原状态相比，没有记录额外的个人信息，不会对用户个人信息安全造成风险。
3. 此试验没有后台及数据库的改变，不会导致网站崩溃用户无法访问等大问题。
4. 即使学生每周学不到五小时，他们只是被页面的变更提醒引导到了另外的一个页面，如果今后有需要学生仍然可以进入免费试学、登陆并可能完成继续课程的，不会因此影响用户使用网站的习惯
5. 此试验没有道德风险。

### 持续时间和曝光比例

由于试验风险很小，而试验需要的页面浏览量较大，因此选择 100% 的页面流量转入此试验，需要 18 天来运行试验（如果选取留存率为评估指标则需要 119 天来运行试验，时间太长，所以不选取它为评估指标）。

### 试验分析

#### 合理性检查

各不变指标的 95%置信区间、实际观察值和是否通过合理性检查情况如下

指标	置信区间下限	置信区间上限	观察值	是否合理
Cookie 的数量	0.4988	0.5012	0.5006	是
点击次数	0.4959	0.5041	0.5005	是
点进概率	-0.0013	0.0013	0.00005663	是

所有不变指标都通过合理性检查，可以开始其他分析工作。

### 结果分析

#### 效应大小检验

各评估指标的 95%置信区间、实际观察值和是否通过合理性检查情况如下

指标	置信区间下限	置信区间上限	统计显著性	实际显著性
总转化率	-0.0291	-0.0120	是	是
净转化率	-0.0116	0.0019	否	否

#### 符号检验

对每个评估指标使用每日数据进行符号检验的情况如下

指标	p 值	统计显著性
总转化率	0.0026	是
净转化率	0.6776	否

#### 汇总

分析中未使用 Bonferroni 校正，因为 Bonferroni 校正是用于多次独立试验，而总转化率和净转化率相互关联，并不独立。

效应大小假设检验和符号检验的结果一致。

## 建议

从试验结果上看，实验组的总转化率实际显著性小于控制组。

实验组和控制组的净转化率差，其 95% 置信区间为 (0.0116, 0.0019)，而  $d_{\min}=0.0075$ ，即无法排除实验组的净转化率实际显著小于控制组，无法证明假设“不会在很大程度上减少继续通过免费试学和最终完成课程的学生数量”，所以不建议启动试验。

## 后续试验

### 试验及假设

试验中，当用户点击取消试用后，在确认弹框内增加导师指导支持的链接。

假设部分学生想离开免费试学的原因是没有获得足够的辅导，增加此链接可以引导学生获得导师指导支持，从而减少因为没有足够的时间而离开免费试学、并因此受挫的学生数量。

### 转移单位

转移单位为参加免费试学的用户 ID。

此试验对参加了免费试学的学生进行处理，参加免费试学的每名学生都有一个用户 ID，而且每个用户 ID 对应唯一的一名学生，确保每名参加试验的学生只被分到实验组或控制组的一组中。

### 不变度量

1. Cookie 的数量：即查看课程概述页面的唯一 cookie 的数量。
2. 用户 id 的数量：即报名参加免费试用的用户数量。
3. 点击次数：即点击“开始免费试用”按钮（在免费试用屏幕触犯前发生）的唯一 cookie 的数量。
4. 点击退出用户 id 数量：点击过退出免费试学按钮（无论最终是否退出）的用户 id 数量。
5. 点进概率：即点击“开始免费试用”按钮的唯一 cookie 除以查看课程概述页面的唯一 cookie 的数量所得的结果。
6. 总转化率：即完成登录并报名参加免费试用的用户 id 的数量除以点击“开始免费试用”按钮的唯一 cookie 的数量所得的结果。

其中，Cookie 的数量，用户 id 的数量，点击次数和点击退出用户 id 数量是在试验之前就被采集的度量，而点进概率是由 cookie 的数量和点击次数两个不变量相除所得，而总转化率是由用户 id 的数量和点击次数两个不变量相除所得，所以这 6 个度量作为不变度量。

### 评估度量

实际离开率：退出试学的用户 ID 数量除以点击过退出免费试学按钮的用户 ID 数量。

实际退出试学的用户 ID 数量会受到试验影响，但是不能排除实验组和对照组中点击过退出免费试学按钮用户数量不同的影响，因此采用比例化的度量。