

# 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:

- 1, Cookie 的数量: 即查看课程概述页面的唯一 cookie 的数量。
- 2, 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。
- 3, 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。

选择不不变指标的理由:

- 1, 查看课程概述页面的唯一 cookie 的数量是在实验之前的, 不会受到实验的影响, 因此试验组和控制组中查看课程概述页面的唯一 cookie 的数量可以是相同的, 即为不变指标。
- 2, user id 是在试验之后, 我们先看到了提示页面, 然后再选择是否注册, 这样试验组和控制组中注册 user id 的数量可能会受到试验的影响而发生不同。因此 user id 的数量不是不变度量。
- 3, 点击“开始免费试用”按钮的唯一 cookie 的数量是在免费试用屏幕触犯前发生即实验之前, 故不受实验的影响, 则点击“开始免费试用”按钮的唯一 cookie 的数量可以是相同的, 即为不变指标。
- 4, 点击“开始免费试用”按钮的唯一 cookie 与查看课程概述页面的唯一 cookie 的数量均在实验之前, 不会受到实验的影响, 因此点进概率可以是相同的, 即为不变指标。

评估指标及期望结果:

- 1, 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。期望结果: 实验组总转化率显著小于控制组。
- 2, 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。期望结果: 实验组净转化率与控制组没有显著差异。

选择评估指标的理由:

- 1, 用户 id 数量发生于试验之后, 会受到试验的影响, 由于实验组和对照组的 cookie 数量不一定相同, 故仅使用用户 id 数量不能很好的评估实验效果, 使用比例化的评估度量即总转化率能够很好的评估实验效果。故总转化率可以作为评估指标, 而用户 id 数量不作为评估指标。
- 2, 留存率所需的样本数量太多, 在 A/B 实验持续时间内, 我们无法采集到足够数量的样本来对留存率进行评估, 因此留存率不作为评估指标。
- 3, 实验的目的是为了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, 同时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。净转化率可以评估实验组和对照组中继续通过免费试学和最终完成课程的学生数量是否有显著差异, 故净转化率可作为评估指标。

#### 测量标准偏差

- 1, 总转化率: 0.0202
- 2, 净转化率: 0.0156

总转化率与净转化率的分析单元为 cookie, 分组单元也为 cookie, 故分析估计与经验变异是类似的。

#### 规模

##### 样本数量和功效

- 1, 分析阶段不使用 Bonferroni 校正
- 2, 实验正确设计所需的页面浏览量: 685325

##### 持续时间和曝光比例

- 1, 将 80% 的页面流量转入此试验, 以及鉴于此条件, 需要 22 天来运行试验。
- 2, 因为实验可能会影响继续通过免费试学并最终完成课程的学生数量, 故转移部分流量仅对部分用户开展实验。

##### 风险评估

- 1, 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另一个页面, 如果今后有需要学生

仍然可以进入免费试学、登录并可能继续完成课程，不会因此影响用户使用网站的习惯。

2，没有在页面展示上有过大的改动，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。

3，该试验没有关于数据库及后台的改变，不用担心数据的丢失及由于后台的失误导致网页崩溃用户无法访问网页等大问题。

4，此试验也不会对用户的个人信息安全造成风险，因为不论网页是否增加了提醒，用户在确认参加免费试学时都得输入信用卡信息，而很明显系统一定会保护用户的个人信息。

5，该试验同样也没有道德上的风险。

## 试验分析

### 合理性检查

1，Cookie 的数量：95%的置信区间（0.4988，0.5012），实际观察值：0.5006，指标通过合理性检查；

2，点击次数：95%的置信区间（0.4959，0.5041），实际观察值：0.5005，指标通过合理性检查；

3，点进概率：95%的置信区间（-0.0013，0.0013），实际观察值：0.0001，指标通过合理性检查。

### 结果分析

#### 效应大小检验

1，总转化率：95%的置信区间（-0.0291，-0.0120），具有统计和实际显著性；

2，净转化率：95%的置信区间（-0.0116，0.0019），不具有统计和实际显著性。

#### 符号检验

1，总转化率：P 值：0.0026，具有统计显著性；

2，净转化率：P 值：0.6776，不具有统计显著性。

### 汇总

本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，因此使用 Bonferroni 校正会使得试验结果过于保守，故不使用。

## 建议

总转化率具有统计和实际显著性，与预期一致；而净转化率的置信区间包含负值，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。因此我们无法说明“降低的程度不大”，所以不建议启动。

## 后续试验

调换“开始免费试学”按键和“访问课程资料”按键的位置和颜色。

我们的假设是学生访问课程概述页面的时候更容易看到“访问课程资料”按键，通过访问课程资料了解课程是否适合自己而决定是否要购买课程，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。同时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。如果假设成真，则该实验会改变点击“开始免费试学”按钮的唯一 cookie 数量，从而用户的 id 数量也会受到实验的影响，故以访问课程概述页面的唯一 cookie 数量为不变指标，以净转换率（即在 14 天的期限后仍参与课程的用户 id 的数量除以点击了“开始免费试学”按钮的唯一 cookie 的数量所得的比率。）为评估指标。如果净转化率具有统计和实际显著性，则说明该项实验内容可用于实施。

转移单位为 cookie，尽管学生参加的是免费试学，但在登录后他们的用户 id 便被跟踪。同一个用户 id 不能两次参加免费试学。对于不参加免费试学的用户，他们的用户 id 不会在试验中被跟踪，即使他们在访问课程概述页面时登录了网站。

### 参考资料：

优达学城论坛：设计 A/B 测试。