优达学城数据分析师纳米学位

A/B 测试项目

说明：[点击此处下载此文档的英文版本](https://s3.cn-north-1.amazonaws.com.cn/static-documents/nd002/FinalProjectTemplate.pdf)。

模板格式

此模板可用于组织你最终项目的答案。应从你的答案中复制到小测试中的项用蓝色显示。

试验设计

指标选择

不变指标：

1. cookie的数量：即访问课程概述页面的唯一 ​cookie ​的数量（d 最小=3000）​。因为它不会受实验的影响，用户在被提问前就已经访问课程概述页面。
2. 点击次数：即点击“开始免费试学”按钮的唯一 ​cookie ​的数量（在免费试学筛选器触发前发生）（​d最小=240​）。因为它不受实验的影响。
3. 点进概率：即点击“开始免费试学”按钮的唯一 ​cookie ​的数量除以查看课程概述页的唯一 ​cookie ​的数量所得的比率​（​d最小=0.01​）。因为它也不受实验影响。

评估指标：

1. 总转化率：即完成登录并参加免费试学的用户 ​id ​的数量除以点击“开始免费试学” 按钮的唯一cookie ​的数量所得的比率（​d最小=0.01​）。因为它受实验影响，期望实验组中的参加免费试学的学院数量下降，即期望实验组总转化率比控制组的下降。
2. 留存率：即在 ​14 ​天的期限过后仍参加课程（因此至少进行了一次付费）的用户 ​id数量除以完成登录的用户 ​id ​的数量（​d最小=0.01​）。因为它受实验影响，期望实验组中留存率比控制组增多。
3. 净转换率：即在 ​14 ​天的期限后仍参与课程的用户 ​id ​的数量（因此至少进行了一次付费）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（​d最小=0.0075​）。因为它受实验影响，期望实验组的净转换率与控制组的无显著性不同，因为筛选器只会影响参加免费试学的人数，对付费行为影响不大。

用户id的数量是在实验之后被采集的度量，受到实验的影响；但是实验组和控制组的cookie不一定相同，导致用户id数量的不同可能会因为cookie的不同引起，所以不能将它作为评估指标。

测量标准偏差

通过标准偏差的计算公式，在有5000个cookie样本大小访问课程概述页面的情况下，计算测量标准偏差：

|  |  |  |
| --- | --- | --- |
|  |  | SE |
| Unique cookies to view page per day: | 40000 |  |
| Unique cookies to click "Start free trial" per day: | 3200 |  |
| Enrollments per day: | 660 |  |
| Click-through-probability on "Start free trial": | 0.08 |  |
| Probability of enrolling, given click: | 0.20625 | 0.0202 |
| Probability of payment, given enroll: | 0.53 | 0.0549 |
| Probability of payment, given click | 0.1093125 | 0.0156 |

总转化率和净转化率的分析估计和经验变异是类似的，因为这两个指标的分析单位和引流单位是唯一cookie；而留存率的分析估计与经验变异是不同的，因为其分析单位是用户id数量，但引流单位是cookie数量。

规模

样本数量和功效

在分析阶段我不使用Bonferroni校正，因为总转化率和净转化率并不独立，在这里使用Bonferroni校正会过于保守，它适用于多次独立实验。

使用alpha=0.05， beta=0.2，通过在线工具，计算出所需样本容量是：

1. 总转化率：2\*25835\*40000/3200=645875
2. 留存率：2\*39115\*40000/660=4741212
3. 净转化率：2\*27413\*40000/3200=685325

取最大的样本容量，所以总共需要收集的网页浏览数是4741212。

持续时间和暴光比例

为了减少实验时间，将会用100%的页面流量转入此实验，总共需要收集的页面浏览数是4741212，所以需要用4741212/40000=118.5303天，即大约需要118天来进行实验。可见时间太长了，不适合选取它作为评估指标。所以选取第二大的样本容量685325，其需要685325/40000=17.1131天，即大约需要18天来进行实验，这个时间比较合理。

实验风险很小：

1. 实验对页面展示的改动很小，仅仅是在学生点击“开始免费试学”后多弹出一个询问框，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动。
2. 实验不会阻止学生进行免费试学，不会影响用户的使用习惯。
3. 实验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。
4. 实验没有后台和数据库的改变，不会导致网站崩溃用户无法访问等大问题。
5. 实验没有道德风险。

试验分析

合理性检查

对于每个不变指标，对你在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。（这些应是来自“合理性检查”小测试中的答案）

对于任何未通过的合理性检查，根据每日数据解释你觉得最有可能的原因。**在所有合理性检查通过前，不要开始其他分析工作。**

结果分析

效应大小检验

对于每个评估指标，对试验和对照组之间的差异给出 95% 置信区间。说明每个指标是否具有统计和实际显著性。（这些应是来自“效应大小检验”小测试的答案。）

符号检验

对于每个评估指标，使用每日数据进行符号检验，然后报告符号检验的 p 值以及结果是否具有统计显著性。（这些应是“符号检验”小测试中的答案。）

汇总

说明你是否使用了 Bonferroni 校正，并解释原因。若效应大小假设检验和符号检验之间存在任何差异，描述差异并说明你认为导致差异的原因是什么。

建议

提供建议并简要说明你的理由。

后续试验

对你会开展的后续试验进行概括说明，你的假设会是什么，你将测量哪些指标，你的转移单位将是什么，以及做出这些选择的理由。

优达学城

2016年9月