

优达学城数据分析师纳米学位 P7

说明：[点击此处下载此文档的英文版本](#)。

最终项目说明

试验概述：免费试学筛选器

在进行此试验时，优达学城当前的主页上有两个选项：“开始免费试学”和“访问课程资料”。如果学生点击“开始免费试学”，系统将要求他们输入信用卡信息，然后他们将进入付费课程版本的免费试学。14 天后，将对他们自动收费，除非他们在此期限结束前取消试用。若学生点击“访问课程材料”，他们将能够观看视频和免费进行小测试，但是他们不会获得导师指导支持或验证证书，无法提交最终项目来获取反馈。

在此试验中，优达学城测试了一项变化，如果学生点击“开始免费试学”，系统会问他们有多少时间投入到这个课程中。如果学生表示每周 5 小时或更多，将按常规程序进行登录。如果他们表示一周不到 5 小时，将出现一条消息说明优达学城的课程通常需要更多的时间投入才能成功完成，并建议学生可免费访问课程资料。在这里，学生可选择继续进行免费试学，或免费访问课程资料。[这张截图](#)展示了试验概况。

我们的假设是这会为学生预先设定明确的期望，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。如果这个假设最后为真，优达学城将改进整体学生体验和提高导师为能够完成课程的学生提供支持的能力。

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。

度量选择

你会选择以下哪个度量来对此试验进行测量，为什么？对于选择的每个度量，请说明你是否会将它用作不变度量或评估度量。每个度量的实际显著性边界——即它在成为有意义的业务变化前必须观察的差异，在括号中给出。所有实际显著性边界作为绝对变化给出。

任何提及“唯一 cookie”的地方，其唯一性按天决定。（在两个不同日期进行访问的同一个 cookie 将计算两次。）用户 id 自动唯一，因为网站不允许同一个用户 id 参与两次。

- cookie 的数量：即访问课程概述页面的唯一 cookie 的数量。（ $d_{\text{最小}}=3000$ ）

- 用户 id 的数量：即参与免费试学的用户数量。（ $d_{\text{最小}} = 50$ ）
- 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）。（ $d_{\text{最小}} = 240$ ）
- 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（ $d_{\text{最小}} = 0.01$ ）
- 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。（ $d_{\text{最小}} = 0.01$ ）
- 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。（ $d_{\text{最小}} = 0.01$ ）
- 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。（ $d_{\text{最小}} = 0.0075$ ）

现在你还需决定你希望启动试验后获得的结果。你的任何一个评估度量的变化是否充足？你是否想在启动试验时看到多个度量同时全部移动或不移动？这个决定将为你在设计试验时的决策提供有用信息。

测量可变性

[此电子表格](#)包含了三个度量的基准值的粗略估计（再次说明这些数字是从优达学城的真实数字变化而来）。

作为你选择作为评估度量的每个度量，用分析法估计其标准偏差。你预期分析估计是准确的吗？也就是说，如果你有足够的时间，你会为哪个度量收集变异的经验估计？

规模

选择支持的样本数量

使用变异的分析估计，你总共需要收集多少网页浏览数（两个组全部）来为试验提供足够的数据支持？使用 $\alpha = 0.05$, $\beta = 0.2$ 。确保每个度量具有足够的数据支持。

选择持续时间和风险暴露

你将转移多少优达学城的流量到此试验（假设你没有其他试验需要同时执行）？这个变化的风险是否太大，导致你不想在所有流量上运行？

在你所选择的比例条件下，使用可变性的分析估计，运行此试验需要多长时间？如果答案是需要好几周，那么就有点不合理了，你便需要重新考虑能够缩短时间的方案。

分析

你要分析的数据在[这里](#)。此数据包含计算上述度量所需的原始信息，按天细分。请注意电子表格中共有两个表，一个用于试验组，一个用于对照组。

每一列代表：

- 网页浏览数：该天唯一 `cookie` 访问课程概述页面的次数。
- 点击次数：该天唯一 `cookie` 点击课程概述的次数。
- 注册量：该天参与免费试学的用户 `id` 的数量。
- 付费量：该天参加试用且在 14 天后继续参加并进行了付费的用户 `id` 的数量。（注意此列的日期为开始日期，即参加日期，而非付费日期。试学 14 天后开始收费。因此，参加量和付费量跟踪时间比其他列少 14 天。）

合理性检查

首先检查一下这两组之间的不变度量是否相等。如果不变度量是应在这 2 组间随机拆分的一个简单计数，你可以使用第 5 课中展示的二项分布检验。否则，您将需要使用第 1 课中的类似策略对比例差异构造置信区间，然后检查组值之间的差异是否在该置信水平内。

如果你的合理性检查失败，观察每日数据，看看是否可找出问题所在。

检查实践和统计显著性

接下来，对于你的评估度量，计算试验组和控制组之间差异的置信区间，然后检查每个度量是否具有统计和/或实践显著性。若置信区间不包括 0，则度量具有统计显著性（也就是说，你可以确定发生了变化），而如果置信区间不包括实际显著性边界，则其具有实践显著性（也就是说，你可以确定存在对业务重要的变化）。

如果你选择了多个评估度量，你将需要决定是否使用 **Bonferroni** 校正。在决定时，记住你发起此试验想要获得的结果。你拥有多个度量的事实是否使这些结果偶然发生的概率大于 α 水平 0.05？

符号检验

对于每个评估度量，使用每日细分数据执行符号检验。若符号检验的结果与差异的置信区间不符，看你能否找出原因。

建议

最后提供建议。你会启动/不启动这个试验吗？还是会进行更深入的探究，运行后续试验，或者还是应由自己主观去判断？如果你要进行更深入地探究，解释你会调查的领域。如果你要运行后续试验，简要描述这个试验。如果应由自己主观决定，解释有哪些与决定相关的因素。

后续试验：如何减少提前终止

如果你想减少受挫并想要提前终止课程的学生数量，你会尝试什么试验？简要描述你会做出的改变，你对变化影响的假设是什么，你想要测量什么度量，你想要使用什么转移单位。对每个选择提供解释。