

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。](#)

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。

试验设计

度量选择

列出你将在此用作不变度量和评估度量的度量。（这些应与你在“选择不变量”和“选择评估度量”小测试中使用的度量一样）

对于每个度量，解释你为什么使用或不使用它作为不变度量和评估度量。此外，说明你期望从评估度量中获得什么试验结果。

不变度量：

- cookie 的数量：即访问课程概述页面的唯一 cookie 的数量。（d 最小=3000）
- 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）
- 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）

对于控制组和实验组这三个度量都发生在试验之前，不受试验的影响，因此是不变的。

评估度量：

- 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）
在实验组中，点击“开始免费试学”按钮的唯一 cookie 的数量不变，完成登录并参加免费试学的用户 id 的数量变少，所以期望总转化率在实验组比控制组数值变小。
- 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）
实验说明中提到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 数量可以看作不变。在实验组中，完成登录的用户 id 的数量变小，所以期望留存率在实验组比控制组数值变大。
- 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）
在实验组中，在 14 天的期限后仍参与课程的用户 id 的数量不变，点击了“开始免费试学”按钮的唯一 cookie 的数量不变，所以净转换率在实验组和控制组数值不变。

以上三个度量发生在试验之后，且受到试验影响，可以作为评估度量。

其他：

- 用户 id 的数量：即参与免费试学的用户数量。（d 最小 =50）

用户 id 的数量这个度量比较复杂。首先它发生于试验之后，会受到试验的影响，因此它是一个 ok 的评估度量。但是，由于实验组和对照组的 cookie 数量不一定相同，也就是说两组中用户 ID 数量不同可能是由于实验的影响，也可能是由于两组 cookie 的不同。所以使用用户 ID 数量的区别不能够很好的评估试验的效果。在一个比例化的评估度量（总转化率）存在的情况下，我们可以不选择用户 ID 的数量作为评估度量。

测量标准偏差

列出你的每个评估度量的标准偏差。（这些应是来自“计算标准偏差”小测试中的答案。）

对于每个评估度量，说明你是否认为分析估计与经验变异是类似的，或者你是否期望它们是不同的（如果是这样，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。

标准偏差：

总转换率： 0.020230604 分析分组 unit of analysis 总转换率分母，即点击“开始免费试学”按钮的唯一 cookie 的数量 与 unit of diversion 相同，分析变异性可能匹配经验变异性。

留存率： 0.054949012 分析分组 unit of analysis 留存率分母，即完成登录的用户 id 的数量与 unit of diversion 不同，理论上分析变异性无法匹配经验变异性。

净转换率： 0.015601545 分析分组 unit of analysis 净转换率分母，即点击“开始免费试学”按钮的唯一 cookie 的数量 与 unit of diversion 相同，分析变异性可能匹配经验变异性。

规模

样本数量和支持

说明你是否会在分析阶段使用 Bonferroni 校正，并给出你适当开展试验所需的支持网页访问数。（这些应是来自“计算网页访问数”小测试中的答案。）

不会使用 Bonferroni 矫正

需要网页访问数：

总转换率： 645875

留存率： 4741212.121

净转换率： 685325

持续时间和风险暴露

说明你会将哪一部分流量转入此试验，以及鉴于此条件，你需要多少天来运行试验。（这些应是来自“选择持续时间和风险暴露”小测试中的答案。）

说明你选择所转移流量部分的原因。你认为此试验对优达学城来说有多大风险？

留存率需要 4741212.121 网页访问数，约 118.5 天，所以选择通过总转换率和净转换率中较大的净转换率求出天数。

选取总流量的 80% 进入此实验，净转换率需要 685325 网页访问数，每天 40000×0.8 ，约 22 天完成。

我认为所选转移流量对优达学城没有风险。系统增加会问他们有多少时间投入到这个课程中，不会因此影响用户使用网站的习惯，而且在页面展示上没有太大变化，不会对用户造成情感上的冲击。技术上，该试验没有后台的改变，不用担心系统崩溃、用户无法访问等问题。另外，此试验也不涉及隐私，在道德上同样也没有风险。

试验分析

合理性检查

对于每个不变度量，对你期望观察到的值、实际观察的值及度量是否通过合理性检查给出 95% 置信区间（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猜测的最可能的原因。在所有合理性检查通过前，不要开始其他分析工作。

	下界	上界	观测值	是否 pass
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	0.0001	是

结果分析效应大小检验

对于每个评估度量，对试验和对照组之间的差异给出 95% 置信区间。说明每个度量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。）

	dmin	下界	上界	统计显著性	实际显著性
总转换率	0.01	-0.0291	-0.0120	有	有
净转换率	0.0075	-0.0116	0.0019	无	无

符号检验

对于每个评估度量，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。）

	p-value	统计显著性
总转换率	0.0026	有
净转换率	0.6776	无

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。

没有使用 **Bonferroni** 校正，总转换率和净转化率很相近，我认为并不是独立的，所以不需要校正。

建议

提供建议并简要说明你的理由。

我不建议启动本试验。

虽然总转化率具有统计和实际显著性，是我们希望看到的结果。但是净转化率的置信区间包含负数，这是我们不希望看到的，这也说明在投入人力和时间成本之后，进行该试验之后净转化率有可能会减少。

后续试验

将 14 天后对他们收费改为 20 天。 我的假设是延长收费期限会减少受挫并想要提前终止课程的学生数量。我依旧测试总转换率、留存率、净转换率使用 **cookie** 作为分流指标来区分点击“开始免费试用”的用户。

优达学城

2016 年 9 月