

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。

试验设计

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。

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。

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)

不变度量: Number of cookies 、 Number of clicks 、 Click-through-probability

评估度量: Gross conversion 、 Retention 、 Net conversion

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。

此次实验, 控制组 and 对照组的区别在于是否使用了免费试学的筛选功能, Number of cookies、Number of clicks 两个变量在筛选器触发之前发生, 因此在两组实验中是不变的度量, Click-through-probability 是利用 Number of clicks/ Number of cookies 计算而得, 也是不变的变量, 因此上述三个度量可以作为不变度量。

Number of user-ids 是筛选器触发之后发生的度量, 因此是一个变化的度量, $Gross\ conversion = \frac{Number\ of\ user-ids}{Number\ of\ clicks}$, $Retention = \frac{Number\ of\ user-ids(14\ 天之后的用户)}{Number\ of\ user-ids}$, $Net\ conversion = \frac{Number\ of\ user-ids(14\ 天之后的用户)}{Number\ of\ clicks}$, 从中可以看出这三个度量都是变化的度量, 可以作为评估度量, 但是 Number of user-ids 不可以作为评估度量, 因为其无法反映出整体实验的情况。

我期望从评估度量中获得使用筛选器后, 总转化率能够保持不变或者提高, 净转化率能够提高。

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。)

Gross conversion: 0.0202

Retention: 0.0549

Net conversion: 0.0156

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。

Gross conversion 和 Net conversion 的分析估计与经验变异是类似的，因为它们的分析单元与分组单元都是 cookie；而 Retention 的分析估计与经验变异不是类似的，因为它的分析单元是用户 id，分组单元是 cookie。

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。）

分析阶段不会使用 Bonferroni 校正，因为评估变量是相互联系的，并非独立。评估变量选择了 Gross conversion，Net conversion，因 Retention 样本量太大，选择放弃，最终计算样本数量为 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。）

我会将全部流量转入此试验，需要 18 天来运行试验。

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？

因为所需要的流量非常大，所以需要将所有流量转入到此试验中，同时减少运行天数。此试验的风险可能会降低参与优达学城免费试学的数量，即影响生源率。

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）。

Number of cookies: lower bound 0.4988 upper bound 0.5012 observed 0.50064 pass

Number of clicks: lower bound 0.4959 upper bound 0.5041 observed 0.5005 pass

Click-through-probability: lower bound 0.0812 upper bound 0.0830 observed 0.0822 pass

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。

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。）

Gross conversion: lower bound -0.0291 upper bound -0.01198

0 不在 95% 置信区间内，因此具有统计显著性，同时此区间不包含 dmin，具有实际显

著性;

Net conversion: lower bound -0.0116 upper bound 0.0019

同理, 该评估度量不具有统计和实际显著性。

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。)

Gross conversion: p -value=0.0026 具有统计显著性;

Net conversion: p -value=0.6776 不具有统计显著性。

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, 因为 Bonferroni 校正主要适用于 n 次独立检验, 但本试验中 Gross conversion 和 Net conversion 并未独立的, 而是相关联的, 使用 Bonferroni 校正会使试验结果过于保守。

建议

提供建议并简要说明你的理由。

我不会启动这个试验, 因为在上述试验中我们可以看出通过设计免费试学筛选器后, Gross conversion 95%置信区间为 $(-0.0291, -0.01198)$, 具有统计显著性和实际显著性, 说明有 95% 的概率会出现总转化率下降的情况, 对优达学城来说是高风险的。同时 Net conversion 95%置信区间为 $(-0.0116, 0.0019)$, 不具有统计显著性和实际显著性, 也就是有可能会降低提前终止课程的数量, 也有可能不会, 因此我们不能得出该试验能够在提高通过免费试用和最终完成课程的数量方面发挥一定的作用。

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, 你将测量哪些度量, 你的转移单位将是什么, 以及做出这些选择的理由。

我会尝试在改进英文翻译方面进行试验, 优达学城的英文翻译是以外国人的说话方式进行翻译的, 在有限的时间内不能快速理解一句话的含义, 最终会花费很多时间在理解上面, 同时名词也没有实现标准化, 和在书本上学到的知识不一致, 也会导致理解起来有难度。

我们的假设是通过改进英文翻译会提高学生的快速理解, 从而减少受挫的频率, 进而很大程度上会减少提前终止学习的数量。

我将测量: 用户 id 数量、cookie 数量 (点击免费试学)、总转化率、留存率、净转换率。这 5 个度量中前三个是不变度量, 后面两个是评估度量。

转移单位是: 用户 id, 因为我们只需要关注免费试学阶段到正式参加学习阶段的学员数量, 不需要全部数据, 只需要部分数据。