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](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

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“选择不变度量”和“选择评估度量”小测试中使用的度量一样）

不变度量：

1、cookies的数量：即查看课程概述页面的唯一cookie的数量。

2、点击次数clicks：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click-through-probability：即点击“开始免费试用”按钮的唯一cookie数量除以查看课程概述页面的唯一cookie的数量所得的结果。

评估度量：

1、总转化率gross conversion：即完成登陆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。

2、净转化率net conversion：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。

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。

不变度量在免费试用屏幕触犯前需要保持不变，cookies的数量和点击次数clicks是不受影响的，需要保持不变，所以这两个为不变度量，而点进概率是cookies的数量与点击次数clicks的商，自然也保持不变。

评估度量是期望看到变化的度量，此试验中，我们期望减少因为没有足够时间而离开免费试学并因此受挫的学生数量，即总转化率，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，即净转化率。

用户 id 的数量：即报名参加免费试用的用户数量。不用此度量是因为它不直接与评估度量相关，虽然用户id数量增加，在一定程度上可能增加最后的净转化率，但却并不是充分条件。

留存率retention：即在 14 天期限后仍保持参加（并进行了至少一次支付）的用户 id 的数量除以完成登录的用户 id 的数量。留存率retention不用作评估度量是因为它的分析单位（用户id）与引流单位（clicks）不一致，导致试验所需的样本数量较大，从而试验持续时间过长。

此试验中，我们期望减少因为没有足够时间而离开免费试学并因此受挫的学生数量，即总转化率，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，即净转化率。当总转化率和净转化率都具有实际显著性时，可以实施这项改动。

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。）

总转化率gross conversion的标准偏差为sqrt(0.20625\*(1-0.20625)/（5000\*0.08）)=0.0202；同理可计算净转化率的标准偏差为sqrt(0.1093125\*(1-0.1093125)/（5000\*0.08）) =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。

总转化率和净转化率的分析估计与经验变异是类似的，因为他们的分析单位和分组单位都是cookies。

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。）

不用。首先根据总转化率，利用在线计算，将baseline conversion rate填入20.625%，minimum detectable effect设为1%，1-beta设为80%，alpha为5%，得出sample size为25835，即点击次数clicks为25835，根据点进概率得出网页数为25835/0.08=322937.5，由于控制组和试验组要保持网页访问数一致，所以开展试验所需的支持网页访问数为322937.5\*2=645875.同样，根据净转化率，将baseline conversion rate填入10.93125%，minimum detectable effect设为0.75%，1-beta设为80%，alpha为5%，得出sample size为27413，即点击次数clicks为27413，从而开展试验所需的支持网页访问数为27413/0.08\*2=685325.两者取较大值，即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。）

根据上一题，需要685325页面访问数，将70%的流量用于此试验，需要685325/0.7/40000=24.48，取整即25天。

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？

由于试验具有不确定性，如果全部页面用于试验容易产生风险，所以只用70%的流量来进行试验。此试验对优达学城有风险，如果试验结果不理想，也就是总转化率和净转化率并没有改善，会影响收益。

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）。

cookies，期望观察到的值是0.5，实际观测到的是345543/(345543+344660)=0.5006，SD=SQRT(0.5\*0.5/(345543+344660))=0.0006，1.96\*SD=0.0012，lower 0.5-0.0012=0.4988，upper 0.5+0.0012=0.5012，实际观测值在95%的置信区间内，通过合理性检查；同理，对于clicks，期望观察到的是0.5，实际观测到的是28378/（28378+28325）=0.50047，95%置信区间（0.4959，0.5041），实际观测值在95%置信区间内，通过合理性检查。

点击概率click through probability=clicks/cookies，p1=28378/345543=0.0821，se=sqrt(p1\*(1-p1)/n1=0.0005，95%置信区间为（p1-1.96\*se，p1+1.96\*se），即（0.0812，0.0830）,试验组p2=28325/344660=0.0822，在置信区间内，通过合理性检查。

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。）

对于总转化率，分别计算试验和对照组Enrollment总和X1=3423、X2=3785，及clicks总和N1=17260、N2=17294，P\_pool=（X1+X2）/（N1+N2）=0.2086，SE\_pool=sqrt(P\_pool\*(1-P\_pool)\*（1/N1+1/N2）)=0.0044,d\_hat=X1/N1-X2/N2=-0.02055,95%置信区间为（d\_hat-1.96\*SE\_pool, d\_hat+1.96\*SE\_pool），即（-0.0291，-0.0120），都小于0，因此总转化率具有统计显著性，|d\_hat|>2%，因此具有实际显著性。

同理，对于净转化率，可得P\_pool=0.1151，d\_hat=-0.0049，95%置信区间为（-0.0116，0.0018），包含0，因此净转化率不具有统计显著性，|d\_hat|<2%，因此净转化率也不具有实际显著性。

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。）

对于总转化率，分别计算试验和对照组每天的转化率，并相减，可得23个数据中有19个试验组小于控制组，利用在线计算器可得p\_value=0.0026<5%，具有统计显著性。

同理，对于净转化率，23个数据中有13个试验组小于控制组，得出p\_value=0.6776>5%，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，Bonferroni校正过于保守，因为多重比较时，假设一次检验显著性水平为5%，那么n次独立检测至少犯一次第一类错误的概率1-0.05^n会大于实际可接受的5%的水平，在Bonferroni校正时，单次检验的显著性水平被设定为总体显著水平的1/n,如果整体5%，10次独立检测1/n只有0.5%，显得过于保守。效应大小假设检验和符号检验之间不存在差异。

建议

提供建议并简要说明你的理由。

不建议发布。因为净转化率不具有统计显著性和实际显著性，符号检验的p值很大，也不显著。如果贸然发布，很可能会引起盈利水平的下降。

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。

由于净转化率并不显著，后续试验将对用户按地区进行细分，测试每个细分领域的净转化率情况，不变度量是cookies。假设发达地区的用户更易付费，欠发达地区付费率低。试验将测量净转化率，转移单位是cookies。将cookies作为转移单位，分析单位和转移单位一致，差异性较小。

参考网站：

在线样本容量计算器<http://www.evanmiller.org/ab-testing/sample-size.html>

符号检验在线计算<http://graphpad.com/quickcalcs/binomial2/>

优达学城

2016年9月