

优达学城数据分析师纳米学位

A/B 测试项目

一、试验设计

1. 指标选择

- Cookie 的数量，不变指标。Cookie 是该试验的最佳引流单元。因此，实验组与对照组应该没有显著差异
- 用户 ID 的数量，不作为指标。尽管用户 id 的数量发生于试验之后，受试验影响。但是，实验组和对照组的 cookie 数量不一定相同，也就是说两组用户 ID 数量变化可能是由 cookie 数量不同引起的。所以，用户 ID 数量不是一个很好的评估指标。我们可以使用总转化率，作为评估指标。
- 点击次数，不变指标。试验的变化部分是点击“开始免费试用”按钮后的流程。因此，以唯一 cookie 为计数的点击次数不会有显著性变化。
- 点进概率，不变指标。分子是点击次数，分母是查看课程概述页面的唯一 cookie 的数量，都是不受试验影响的指标。因此，点进概率也不会有显著性变化
- 总转化率，评价指标。分子是完成登录并报名参加免费试用的用户 id 数，重点在报名参加，该数量可能会受到试验影响，由试验产生变化。分母是不变指标。因此，总转化率可以作为评价试验效果的指标。
- 留存率，不作为评价指标。留存率本身是可以作为评估指标的，但是在计算页面浏览量以及持续时间时，所需浏览量过大，持续时间太长，因此放弃使用留存率作为评估指标。
- 净转化率，评价指标。试验组与对照组的变化可能会引起分子的变化，分母是不变指标。因此，可以作为评价试验效果的指标。

期望：减少时间不足学生注册量的同时，不减少最终的付费学生数量。即，总转化率降低，净转化率保持不变。

2. 测量标准偏差

总转化率：SE = 0.0202

净转化率：SE = 0.0156

引流单元为 cookie，分析单元同样为 cookie。因此，分析变异性与经验变异性

应该比较接近。

二、规模

1. 样本数量和功效

不会使用 Bonferroni 校正，净转化率与总转化率分母相同，不独立。
所需页面浏览量：685325

2. 持续时间和曝光比例

50%的流量转入此实验。经计算，需要 35 天运行试验

3. 说明你选择所转移流量部分的原因。你认为此试验对于优达学城来说有多大风险？

选择转移部分流量是因为该试验存在降低优达学城的付费用户量的风险，通过转移部分流量，降低风险带来的影响。该试验不涉及道德问题，Cookie 并不是敏感数据，用户 id 是优达学城内部的数据，且只用作计数。对于不知情的参与者来说，试验风险是小于最低风险的。对于优达学城来说，试验有较大可能造成付费用户数量的降低。

三、试验分析

1. 合理性检查

95%的置信区间下：

	下限	上限	观察值	通过
Cookie 的数量	0.4988	0.5012	0.5006	是
“开始免费试用”的点击数	0.4959	0.5041	0.5005	是
“开始免费试用”的点入概率	-0.0013	0.0013	0.0001	是

四、结果分析

1. 效应大小检验

	上限	下限	统计显著性	实际显著性
总转化率	-0.0291	-0.0120	存在	存在
净转化率	-0.0116	0.0019	不存在	不存在

2. 符号检验

$\text{Alpha} = 0.05$

	P 值	统计显著性
总转化率	0.0026	存在
净转化率	0.6776	不存在

3. 汇总

不使用 Bonferroni 校正，总转化率与净转化率分母相同，不是独立指标。

五、建议

总转化率降低，说明时间不足学生的注册量会显著减少。但没有足够证据说明净转化率显著降低。从净转化率效应大小的置信区间来看，很大一部分都是负数。说明有较大的可能性会降低净转化率。这与我们的期望不符，所以不建议实施试验。

六、后续试验

试验：

在学生点击取消免费试用按钮后，出现一则帮助消息，并鼓励学生坚持完成课程。在这里学生可以放弃取消，坚持完成课程，也可以忽略消息，放弃退出。

假设：

会减少受挫并想要提前终止课程的学生数量

测量指标：

留存率，14 天后仍参加课程的用户 id 数量除以完成登录并参加免费试学的用户 ID 数量

转移单位：

完成登录并参加免费试学的用户 id

理由：

能参与该实验的用户一定是完成登录并且已经参加免费试学的用户，所以引流单元选择用户 id。试验弹出的帮助信息，可以鼓励甚至帮助用户坚持完成课程，所以 14 天后仍参加的用户 id 数量可能增加。因此选择留存率作为测量指标。