

A/B 测试项目报告

试验设计：

I. 指标选择：

A. 不变指标：

1. Cookie 的数量：即查看课程概述页面的唯一 cookie 的数量
 - a. 选择理由：本项目的转移单位是 Cookie，且查看课程概述页面这一事件发生在点击“免费试用”事件之前，不会产生影响，所以须保持 Cookie 数量不变。
 - b. 期望结果：不变
2. 点击次数：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。
 - a. 选择理由：本项目测试的是点击“开始免费试用”后的页面变化产生的影响，不会对点击次数本身产生影响，所以须保持点击次数不变。
 - b. 期望结果：不变
3. 点进概率：即点击“开始免费试用”按钮的唯一 cookie 除以查看课程概述页面的唯一 cookie 的数量所得的结果。
 - a. 选择理由：因为本项目测试的是点击“开始免费试用”后的页面变化产生的影响，如果点进概率变大或变小，都会对后续检验的评估指标产生很大影响，且点进这一事件发生在试验之前，其概率不受试验影响，所以续保持点进概率不变。
 - b. 期望结果：不变

B. 评估指标：

1. 总转化率：即完成登录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。
 - a. 选择理由：因为询问用户每周是否有 5 小时学习时间可能会对注册免费试用的用户人数产生影响，有些用户可能没有 5 小时时间而选择不注册免费试用，所以总转化率可以是一个评估指标。
 - b. 期望结果：减小
2. 留存率：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。
 - a. 选择理由：因为询问用户每周是否有 5 小时学习时间可能会对注册免费试用的用户人数产生影响，而对付费用户人数也有可能产生影响，所以留存率可以是一个评估指标。
 - b. 期望结果：增大
3. 净转化率：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。
 - a. 选择理由：因为点击次数是不变指标而付费用户人数可能是个变量，且希望试验不会减少付费用户人数，所以净转化率可以是一个评估指标。
 - b. 期望结果：不变

C. 无关指标：

1. 用户 id 的数量：注册“免费试用”的用户数量
2. 理由：注册“免费试用”在试验发生之后，会受到试验的影响，用户 id 的数量

可以作为一个评估度量，但是对照组和试验组的 cookie 数量可能会不同，这可能会使两组用户 id 的数量产生差异，从而可能对最终试验效果评估的准确性产生影响。而且已经存在一个更稳健的比例化评估度量——总转化率，用户 id 的数量就不适应作为评估度量。

II. 测量标准偏差:

- A. 总转化率: 0.0202
- B. 留存率: 0.0549
- C. 净转化率: 0.0156
- D. 总转化率和净转化率的转移单位和分析单位都是 Cookie 数量，所以分析估计与经验变异类似。但是留存率的转移单位是 Cookie 数量，分析单位却是 id 数量，分析估计与经验变异会不同，需要进一步进行经验估计。

III. 规模:

在分析阶段不使用 Bonferroni 校正，因为三个评估指标之间不是相互独立，可能存在关联，如果使用会使结果过于保守。

A. 样本数量和功效:

- 1. 总转化率: 页面浏览量: $25835 \times 2 / 0.08 = 645875$
- 2. 留存率: 页面浏览量: $39115 \times 2 / 0.0165 = 4741212$
- 3. 净转化率: 页面浏览量: $27413 \times 2 / 0.08 = 685325$ ✓

因为最大值 4741212 过大，试验持续时间将会过长，不适宜作为本试验的样本数量，所以舍弃留存率这个指标，选择 685325 作为试验所需页面浏览量。

B. 持续时间和曝光比例:

- 1. 曝光比例: 100%
 - a. 本试验只是询问用户每周学习时间，用户会被页面变更提醒引导到另一个页面，不会影响用户后续注册免费试学和继续学习课程，也不会影响用户使用网站的习惯。
 - b. 没有对页面展示有过大的改动，不会对用户的情感产生过大的波动，也不需要用户花费时间去适应改动后的页面。
 - c. 不涉及数据库和后台的变动，不影响网站的安全性，不会造成数据的丢失和后台的失误导致网页崩溃无法访问。
 - d. 用户输入的学习时间不是敏感数据，不会对用户个人信息安全造成风险，不论是否进行此试验，用户在确认参加免费试学时都得输入信用卡信息，且系统会保护用户的个人信息安全。
 - e. 没有道德上的风险，用户面临的风险也很小。
 - f. 所以本试验的风险较小，可以不必控制试验规模。
- 2. 持续时间: 18 天

试验分析:

I. 合理性检查:

A. Cookie 的数量:

- 1. 95%置信区间: (0.4988, 0.5012)
- 2. 观察值: 0.5006
- 3. 因为观察值在置信区间内，所以通过完整性检查

B. 点击次数:

- 1. 95%置信区间: (0.4959, 0.5041)
- 2. 观察值: 0.5005

3. 因为观察值在置信区间内, 所以**通过完整性检查**
- C. 点进概率:
 1. 95%置信区间: (-0.001296, 0.001296)
 2. 观察值: 0.00005663
 3. 因为观察值在置信区间内, 所以**通过完整性检查**
- II. 结果分析:
 - A. 效应大小检验:
 1. 总转化率: 95%置信区间: (-0.0291, -0.0120)
结论: 置信区间不包含 0, 且不包含实际性显著边界 0.01, 所以总转化率**具有统计显著性和实际显著性**。
 2. 净转化率: 95%置信区间: (-0.01160, 0.001857)
结论: 置信区间包含 0 且包含实际性显著边界 0.0075 所以净转化率**不具有统计显著性, 也不具有实际显著性**。
 - B. 符号检验:
 1. 总转化率: p 值: 0.0026
结论: p 值小于显著性水平 $\alpha=0.05$, 所以总转化率**具有统计显著性**。
 2. 净转化率: p 值: 0.6776
结论: p 值大于显著性水平 $\alpha=0.05$, 所以净转化率**不具有统计显著性**。
 - C. 汇总:
分析过程未使用 Bonferroni 校正, 因为评估指标之间不是相互独立的, 具有关联性, 如果使用会使结果过于保守。

建议:

由结果分析可以初步看出, 总转化率具有统计显著性和实际显著性, 效应为负, 说明在试验中, 注册免费试学的用户人数减少了, 与期望结果相符; 但是净转化率不具有统计显著性和实际显著性, 且效应有正有负, 负效应大于正效应, 说明净转化率有较大的概率减小, 与期望结果不符。因此不建议启动这个试验。

后续试验:

为了减少受挫而提前终止课程的用户数量, 建议在用户注册“免费试用”后的课程页面上加上“预约咨询导师/助教”的按钮。

假设: 导师/助教会解答用户对课程的各种疑惑, 并增强用户的学习信心, 那么将会有更多的用户继续留下来学习课程。

转移单位: 用户 id

理由: 本试验是在用户注册课程之后进行, 每个 id 只能注册一次课程, 注册之后 id 便会被追踪, 以用户 id 作为转移单位可以更好地获得每个用户的完整课程体验。

度量:

不变指标:

1. 用户 id 的数量: 注册“免费试用”的用户数量
2. 理由: 注册“免费试用”在试验进行之前, 试验不会对注册用户数量造成影响, 应该保持不变。

评估指标:

1. 留存率: 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。
2. 理由: 本试验会影响在 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, 留存率可以很好地检验本试验的效应。