

# 优达学城数据分析师纳米学位

## A/B 测试项目

### 试验设计

#### 指标选择

对于每个指标，解释你为什么使用或不使用它作为不变指标或评估指标。此外，说明你期望从评估指标中获得什么样的试验结果。

在试验中我们选取了两组度量：不变度量和评估度量。其中不变度量包括：

1. Cookie 的数量：查看课程概述页面的唯一 cookie 的数量。
2. 点击次数：点击“开始免费试用”按钮的唯一 cookie 的数量。
3. 点击概率：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。

Cookie 的数量：因為我們修改的點擊「開始免費試學」後的子頁面，對應的影響也是其後的子頁面的數據，不會影響首頁的數據，所以對於首頁的 Cookie 數量不會改變。

因此期望結果為：不變

点击次数：同 Cookie，對於首頁的「開始免費試學」點擊次數也不會改變。

因此期望結果為：不變

点击概率：因為分子點擊「開始免費試學」的次數以及分母 Cookie 的數量都不變。

因此期望結果為：不變

评估度量包括：

1. 总转化率：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。
2. 留存率：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。
3. 净转化率：在 14 天期限结束后仍然参加的用户 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。

总转化率： 因為修改了點擊「開始免費試學」按鈕後的介面，同時會根據用戶所設的時間長短，走不同的路徑，對於大於五小時的用戶，就直接走登入填寫信用卡的流程路徑，反之則走完全免費的路徑。這個環節相較原分案，會導致登入用戶 ID 數減少，同時也因為點擊開始「開始免費試學」按鈕的用戶不變，分母不便而分子減少。

因此期望結果為：變小

留存率： 因為實驗期望「同时不会在很大程度上减少继续通过 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」，在分子不變的情況下，分母變小。

因此期望結果為：變大

净转化率: 因為實驗期望「同时不会在很大程度上减少继续通过 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」，且點擊開始「開始免費試學」按鈕的用戶不變，分子分母都不變。

因此期望結果為：不變

不選擇用戶 id 的理由:

用戶 id 的登入發生在實驗之後，會受到實驗的影響，因此它可以是評估度量指標。但是，由於實驗組和對照組的 cookie 數量不一定相同，也就是兩組中用戶 id 數量不同可能是由於實驗的影響，也可能是由於兩組 cookie 的不同。所以使用用戶 id 數量的區別不能夠很好的評估實驗的效果。因此，在使用一個比例化的評估度量(總轉化率)的情況下，我們可以  
~ 不選擇用戶 id 數量作為評估度量指標。

## 测量标准偏差

对于每个评估指标，说明你是否认为分析估计与经验变异是类似还是不同（如果不同，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：

$$P = 0.20625$$

$$\text{Unique cookies to click "Start free trial" per day} / \text{Unique cookies to view page per day}$$

$$= 3200 / 40000 = 0.08$$

$$N = 5000 * 0.08 = 400$$

$$\text{Std dev} = \sqrt{0.20625 * (1 - 0.20625) / 400} = 0.0202$$

总转化率 = 完成登录并报名参加免费试用的用户 id 的数量 / 点击“开始免费试用”按钮的唯一 cookie 的数量，所以作為分母的 cookie 即是轉移單位與分析單位，故分析估計與經驗估計變異相似。

留存率：

$$P = 0.53$$

$$N = 5000 * 0.20625 = 82.5$$

$$\text{Std dev} = \sqrt{0.53 \cdot (1-0.53) / 82.5} = 0.0549$$

留存率 = 即在 14 天的期限过后仍参加课程的用户 id 数量 / 完成登录的用户 id 的数量。分母是 id 數量，分析單位和轉移單位不相似，所以分析估計與經驗變異不相似

净转化率:

$$P = 0.1093125$$

$$N = 5000 \cdot 0.08 = 400$$

$$\text{Std dev} = \sqrt{0.1093125 \cdot (1-0.1093125) / 400} = 0.0156$$

净转化率 = 即在 14 天的期限后仍参与课程的用户 id 的数量 / 点击了“开始免费试学”按钮的唯一 cookie 的数量。作為分母的 cookie 數量既是分析單位和轉移單位，所以分析估計與經驗變異相似

## 规模

样本数量和功效

说明你是否会在分析阶段使用 Bonferroni 校正，并给出实验正确设计所需的页面浏览量。（这些应是来自“计算页面浏览量”小测试中的答案。）

不使用 Bonferroni 校正

因為  $\alpha = 0.05$ ， $\beta = 0.2$ ，根據在線計算器所得以下：

1. 总转化率：

$$\text{Baseline conversion rate} = \text{Probability of enrolling, given click} = 0.20625 = 20.625\%$$

$$\text{Minimum Detectable Effect} = 1\%$$

$$\text{樣本數量} = 25835$$

$$\text{總瀏覽量} = (25835 / 0.08) \cdot 2 = 645875$$

2. 留存率：

$$\text{Baseline conversion rate} = \text{Probability of payment, given enroll} = 0.53 = 53\%$$

$$\text{Minimum Detectable Effect} = 1\%$$

$$\text{樣本數量} = 39115$$

$$\text{總瀏覽量} = (39115 / 0.20625 / 0.08) \cdot 2 = 4741212$$

3. 净转化率:

$$\text{Baseline conversion rate} = \text{Probability of payment, given click} = 0.1093125 = 10.93125\%$$

$$\text{Minimum Detectable Effect} = 0.75\%$$

$$\text{樣本數量} = 27413$$

$$\text{總瀏覽量} = (27413 / 0.08) \cdot 2 = 685325$$

我們取最大的留存率的總瀏覽量作為樣本總數：4741212

$$\text{日流量} = 4000$$

以 80%流量來說實驗天數需要 148 天，但該實驗天數太長，我們需要修改樣本總數  
取淨轉化率的總瀏覽量 = 685325，如此以 80%流量來說實驗天數需要約 22 天，此數較為合理。

持续时间和曝光比例並说明该试验的风险性:

说明你会将多少百分比的页面流量转入此试验 ,以及鉴于此条件 ,你需要多少天来运行试验。  
( 这些应是来自 “选择持续时间和曝光” 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？

取淨轉化率的總瀏覽量 = 685325 ，如此以 80%流量來說實驗天數需要約 22 天，此數較為合理。

该试验的风险性:

1. 因為不論網頁是否增加了提醒頁面，用戶在確認參加免費試學後都得輸入信用卡信息，而系統也一定會保護用戶的個人信息，因此，此實驗部會對用戶的個人信息安全造成風險
2. 該實驗沒有對後台或數據庫的更改，所以不需擔心數據丟失或因後台失誤導致網頁崩潰，使用戶無法訪問網頁等問題
3. 即使學生每周學不到 5 小時，他們也只被頁面的變更提醒引導到另外一個頁面，如果今後有需要的學生仍然可以進入免費試學、登入並繼續完成課程的，因此不會影響用戶使用網站的習慣。
4. 此實驗沒有在頁面展示上有過大的變更，部會對用戶產生情感上的衝擊，用戶也不需要花時間去適應頁面，因此不會對受試者造成生理或心理傷害
5. 此實驗沒有道德上的風險

## 试验分析

### 合理性检查

对于每个不变指标，对你在 95%置信区间下期望观察到的值、实际观察的值及指标是否通过合理性检查给出结论。( 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觉得最有可能的原因。**在所有合理性检查通过前，不要开始其他分析工作。**

	下限	上限	觀察值	是否通過
Cookie 數量	0.4988	0.5012	0.5006	通過
点击次数	0.4959	0.5041	0.5005	通過
点进概率	0.0812	0.0830	0.0822	通過

### 结果分析

#### 效应大小检验

对于每个评估指标，对试验和对照组之间的差异给出 95% 置信区间。说明每个指标是否具

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。)

	下限	上限	統計顯著性	實際顯著性
總轉化率	-0.0291	-0.0120	具備，因置信區間不包含 0	具備
淨轉化率	-0.0116	0.0019	不具備，因置信區間包含 0	不具備

### 符号检验

对于每个评估指标，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 值	統計顯著性
總轉化率	0.0026	具備，因小於 $\alpha = 0.05$
淨轉化率	0.6776	不具備，因大於 $\alpha = 0.05$

### 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。

不使用 Bonferroni 校正，因總轉化率與淨轉化率不是互相獨立，可能使結果過於保守。效應量檢驗與符號檢驗得到的結果一致，符合預期。

## 建议

提供建议并简要说明你的理由。

從實驗結果可以看到，總轉化率有 95% 的信心落在 (-0.0291, -0.0120) 之間，說明總轉化率降低，變化符合預期，且具有統計與實際顯著性。測試的修改起到了作用，新的提示使學生更加慎重地考慮自己是否有足夠的時間投入到課程的學習中，而這些沒有足夠時間的學生，顯然也不會成為付費用戶。這達到了減少因為沒有足夠的時間而離開免費試學、並因此受挫的學生數量的目的。

而淨轉化率有 95% 的信心落在 (-0.0116, 0.0019) 之間，說明淨轉化率降低，並不符合「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」的預期。

換言之，實驗可能“在很大程度上減少繼續通過免費試學和最終完成課程的學生數量”。因此不建議展開實驗。

## 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，你将测量哪些指标，你的转移单位将是什么，以及做出这些选择的理由。

為了減少因受挫而提前終止課程的學生數量，我認為改善錯誤提示可以提高部分用戶的信心，進而減少因為挫敗而流失的用戶數量。因此我想展開的後續實驗是在用戶練習出錯時改善錯誤提示，例如依據用戶出錯部分提出針對性的提示，或在進行項目練習時提供相關視頻鏈結，引導用戶重新複習相關內容再嘗試回答問題。

不变度量:

1. 免費試用的用戶 id 數量
2. 瀏覽課程頁面的 cookie 數量
3. 開始免費試用的點擊概率

上述三個不變度量因是在課程開始後發生，因此不產生影響。

评估度量:

1. 留存率  
定義為 “至少進行一次付費的用戶 id 數量 / 所有參加課程的用戶 id 數量”
2. 終止概率  
定義為 “提前終止課程的用戶 id 數量 / 所有參加課程的用戶 id 數量”

(轉移)分組單位:

1. 用戶 id  
後續試驗主要追蹤的是參與課程的用戶 id，需保持用戶體驗一致。

參考來源：

優達學城論壇 A / B 測試 區

<https://discussions.youdaxue.com/t/topic/54702>

<https://discussions.youdaxue.com/t/topic/49928/3>

<https://discussions.youdaxue.com/t/topic/42144/36>

知乎 : <https://zhuanlan.zhihu.com/chizaohuigengxin>

Github : <https://github.com/leesper/become-data-analyst>

优达学城

2017 年 3 月