

# A/B text 实例(过程)



王学习ty

关注

 0.056 2018.05.23 11:55:37 字数 2,227 阅读 149

## A/B 测试项目

### 实验概述：免费试用筛选器

在本次实验时，课程目前在课程概览页面上有两个选项：“开始免费试用”和“访问课程资料”。如果学生点击“开始免费试用”，他们将被要求输入他们的信用卡信息，然后他们将参加免费试用课程的付费版本。14天后，除非他们先取消，否则将自动收费。如果学生点击“访问课程资料”，他们将能够免费观看视频并参加测验，但他们不会获得辅导支持或经过验证的证书，他们也不会提交最终项目以获得反馈。

在实验中，测试了一个变化，如果学生点击“开始免费试用”，他们被问到他们有多少时间可以投入课程。如果学生每周表示5小时或更长时间，他们将像往常一样通过结帐流程进行。如果他们每周表示少于5小时，则会显示一条消息，指出Udacity课程通常需要更多的时间才能成功完成，并且建议学生可能希望免费访问课程材料。在这一点上，学生可以选择继续参加免费试用，或免费使用访问课程材料。

我们的假设是，这可能会为学生预先设定更明确的预期，从而减少因为时间不够而离开免费试用的失意学生的数量- 这不会显著减少学生继续超过免费试用并最终完成课程的数量。如果这个假设成立，可以改善整体学生体验，并提高教练支持有可能完成课程的学生们的能力。

转移单位是一个cookie，但如果学生参加了免费试用，那么他们将从该用户id开始跟踪。相同的用户ID不能参加免费试用两次。对于未注册的用户，即使他们在访问课程概览页面时已登录，他们的用户ID也不会 在实验中进行跟踪。

## 试验设计

### 指标选择

Number of cookies:That is, number of unique cookies to view the course overview page. (dmin=3000)

Number of user-ids:That is, number of users who enroll in the free trial. (dmin=50)

Number of clicks: That is, number of unique cookies to click the "Start free trial" button (which happens before the free trial screener is trigger). (dmin=240)

Click-through-probability:That is, number of unique cookies to click the "Start free trial" button divided by number of unique cookies to view the course overview page. (dmin=0.01)

Gross conversion: That is, number of user-ids to complete checkout and enroll in the free trial divided by number of unique cookies to click the "Start free trial" button. (dmin= 0.01)


Retention: That is, number of user-ids to remain enrolled past the 14-day boundary (and thus make at least one payment) divided by number of user-ids to complete checkout. (dmin=0.01)

Net conversion: That is, number of user-ids to remain enrolled past the 14-day boundary (and thus

30 Y  
PRO

Ultra-  
Negat

Trading is r



王学习ty

总资产 1 (约0.08元)

关注

### 推荐阅读

因长得漂亮嫁“国外首富”，丈夫为不打扰其睡觉，切除自身器官  
阅读 14,167

28岁的我也很焦虑  
阅读 3,599

表姐的强买强卖  
阅读 1,631

第二十五章太晨喜事  
阅读 2,749

我的小诗赚钱了  
阅读 24,807

写下你的评论...

 评论0  赞1 ...

Invariant Metrics: Number of cookies, Number of clicks, Click-through-probability

Evaluation Metrics: Gross conversion, Retention(剔除), Net conversion

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。

测量标准偏差

列出你的每个评估指标的标准偏差。

对于每个Evaluation Metrics，计算标准偏差SD。条件当Number of cookies = 5000 pageview时。  
(使用的不是实验数据，而是基准数据)  
根据基准数据， $SD = \sqrt{P \cdot (1 - P) / N}$

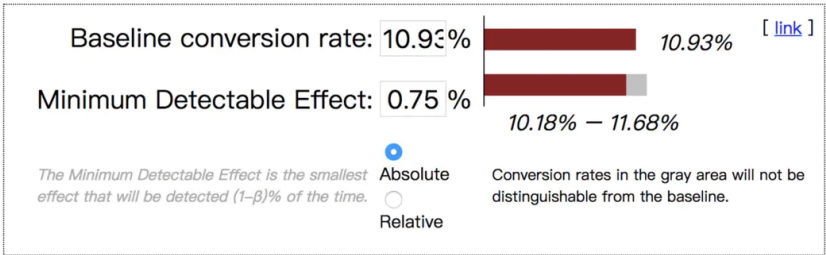
Unique cookies to view course overview page per day:	40000		5000
Unique cookies to click "Start free trial" per day:	3200		400
Enrollments per day:	660		82.5
Click-through-probability on "Start free trial":	0.08		
Probability of enrolling, given click:	0.20625	SD	0.02023
Probability of payment, given enroll:	0.53	SD	0.05495
Probability of payment, given click	0.1093125	SD	0.01560

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。

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。  
利用样本计算器，根据给定的 $\alpha = 0.05$ ,  $\beta = 0.2$ , dmin, BCR(基准数据)，算得是unit of analysis中分母的数量，eg.( net conversion, BCR=0.1093, dmin=0.01)  
no use of Bonferroni Correction, because the three evaluation metrics are likely covariant.  
得出 27411，需要除以点击概率0.08再乘以2 总样本数685275。由于Retention中unit of dividend跟 unit of analysis 不一致，sd异常大于另外两个指标，导致所需pageview过高在计算duration时会出现duration太长不适合实验，因此剔除此指标。总样本数为满足所有指标中最大的需求



Sample size:

27,411

per variation

Statistical power 1-β:  80% Percent of the time the minimum effect size will be detected, assuming it exists

Significance level α:  5% Percent of the time a difference will be detected, assuming one does NOT exist

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。

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？

因为此项实验风险很低，所以我想把流量分成两部分，一部分流量用于实验，另一部分流量用于对照。这样做的目的是为了确保实验结果的准确性，同时也能保证用户体验不受影响。

试验分析

合理性检查 (Sanity checks)

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。

此处用到实验数据，每列的含义是：

综合浏览量：当天查看课程概览页面的唯一用户数。

点击次数：当天点击课程概览页面的唯一用户数。

注册：当天免费试用中注册的用户数。

付款：在当天注册的用户数的数量，以保持注册，天，从而进行付款。（请注意，此栏的日期是开始日期，即注册日期，而不是付款日期，付款在，天后发生，因此，注册和付款的跟踪时间缩短了，天比其他栏。）

实验中不变指标是否合理，对于实验组和对照组，对照实验组随机分布，满足正态分布，可以用正态分布的期望和标准差来衡量。如果实验结果与期望值相差较大，说明实验可能存在异常。

对于实验组和对照组，计算期望值和标准差，然后比较实验结果与期望值的差异是否在合理范围内。如果差异过大，说明实验可能存在异常。

### Sanity checks

For each metric that you chose as an invariant metric, compute a 95% confidence interval for the value you expect to observe. Enter the upper and lower bounds, and the observed value, all to 4 decimal places. Check the box if the metric passes your sanity check.

	Lower bound	Upper bound	Observed	Passes
• Number of cookies	0.4988	0.5012	0.5006	<input checked="" type="checkbox"/>
• Number of user-ids				<input type="checkbox"/>
• Number of clicks on "Start free trial"	0.4959	0.5041	0.5005	<input checked="" type="checkbox"/>
• Click-through-probability on "Start free trial"	-0.0013	0.0013	0.0000	<input checked="" type="checkbox"/>
• Gross conversion				<input type="checkbox"/>
• Retention				<input type="checkbox"/>
• Net conversion				<input type="checkbox"/>

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。

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。只计算实验组的数据。

b" ertit tē"l. 1 1k" mēbēā t ū:qān Xeritē"l, 1, 1, 1,  
b" ertit tē"l kē b" ertit Xēāē"l, 1, 1. 2222 1222  
t ū:qān Xeritē"l kē b" ertit Xēāē"l, 1, 1 222. 222 2  
1 1tē"l 222222 1ē"l 22 2t tē"l. 1"1"l, 1, 1, 22  
k ū mē ā 1, 1, 1, 1  
t"1 Xēb/ē"l ōēl 1k 2 2tē"l 1tē"l 222222 1, 1, 1. 22  
n 1tēb/ē"l ōēl 1k 2 2tē"l 1tē"l 222222 1, 1, 1. 22  
gmrākrā gāpāāēl ē\Xēēl\lā ēgāpāāēl ē\X  
  
ŌXrib" ertit Xēāēlē"l, 1, 1, 1, 1. 2222 2222 gmrākrā gāpāāēl ē\Xēō" ēl\lā ēgāpāāēl ē\X

### 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 H 值以及结果是否具有统计显著性。

[网络计算器](#), ē"l, 1, 1, 1, 1 6Xl ōkX1X kōh XrāXē"l 1 2222ē\Xē  
ō1kēkō\1Xkē 1tē"lē"l, 1ēāē 22Xl, 1, 1, 1. 2222 2222!!gmrākrā gāpāāēl ē\X  
ŌXēkō\1Xkē, 1tē"lē"l, 1ēāē 22Xl, 11111tē"l 2222!!Ō" gmrākrā gāpāāēl ē\X

### 汇总


说明你是否使用了^" 2222ēā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。

Ō" ōkX" 22" 2222ēā" 2222ēl ēlēāē\Xē" 1 Xrāē ū tXl" 2222ēl ēlēāē\Xē" 2222ēāē 2222X  
m" 1" ēkXēl 22Xē  
ēXkōēkē 1 X1 2222 2222ēāē\Xēkē


### 建议


### 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。

 1 人点赞 >



 1 m ē ēkē



小礼物走一走，来简书关注我

赞赏支持

还没有人赞赏，支持一下

推荐阅读

更多精彩内容>



我从小就喜欢画画，长大后，我成了画家。我的画，不仅在国内，而且在国外也很受欢迎。我的画，不仅在国内，而且在国外也很受欢迎。我的画，不仅在国内，而且在国外也很受欢迎。



kxqelua 阅读 1 评论 0 赞 1

敬拜限定性原则与圣诞节

敬拜限定性原则与圣诞节，这是一个关于宗教自由和信仰自由的问题。在圣诞节期间，许多人都会在教堂里参加礼拜，这是他们的信仰自由。但是，对于那些不信仰基督教的人来说，他们是否也应该有参加礼拜的自由呢？这是一个值得探讨的问题。



kqaxs 阅读 1 评论 0 赞 1

一晃两三年，匆匆又夏天

就这样，我们上完了最后一天的课。眨眼间，已经初三了，三年转瞬即逝。我们还在还在彷徨懵懂中，中考渐渐逼近，我们没有时间了。



青柠酸 阅读 1 评论 0 赞 1

我是中国人

我听着故事 从万国朝贺到东亚病夫 我摸着母亲曲折的脉络 深刻而沉重 一个属于母亲的故事 每一次倾听，我都要将心脏安放在



风陌白 阅读 1 评论 0 赞 1

亲子日记 9篇

9月10日，星期天。昨天没做完作业，从早上起来吃完饭就开始做作业，中午才做完，不认识字，我一边读题目一边写。



爱笑的丫头 16 阅读 1 评论 0 赞 1

写下你的评论



评论



赞

