

# 优达学城 A/B 测试项目

## 业务背景

优达学城当前的主页上有两个选项：“开始免费试学”和“访问课程资料”。如果学生点击“开始免费试学”，14 天后将对他们自动收费，除非他们在此期限结束前取消试用。若学生点击“访问课程材料”，他们将能够观看视频，但是他们不会获得导师指导或认证证书，也无法提交最终项目来获取反馈。在此试验中，优达学城测试了一项变化，如果学生点击“开始免费试学”，系统会问他们有多少时间投入到这个课程中。如果学生表示每周 5 小时或更多，将按常规程序进行登录。如果他们表示一周不到 5 小时，将出现一条消息说明优达学城的课程通常需要更多的时间投入才能成功完成，并建议学生可免费访问课程资料。

## 业务目标

1. 减少因为没有足够的时间而离开免费试学的学生数量。
2. 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。

## 指标含义

这里选择 cookie 作为分组单位，因为这样能保证登录前后看到的页面是一样的。

- cookie 的数量：即访问课程概述页面的唯一 cookie 的数量。（d 最小=3000）
- 用户 id 的数量：即参与免费试学的用户数量。（d 最小 =50）
- 点击次数：即点击“开始免费试学”按钮的唯一 cookie 的数量（在免费试学筛选器触发前发生）。（d 最小 =240）
- 点进概率：即点击“开始免费试学”按钮的唯一 cookie 的数量除以查看课程概述页的唯一 cookie 的数量所得的比率（d 最小=0.01）
- 总转化率：即完成登录并参加免费试学的用户 id 的数量除以点击“开始免费试学”按钮的唯一 cookie 的数量所得的比率，衡量的是从 cookie 转化为试学用户的比例。（d 最小 =0.01）
- 留存率：即在 14 天的期限过后仍参加课程（因此至少进行了一次付费）的用户 id 数量除以完成登录的用户 id 的数量，反映的是试学用户的流失情况。（d 最小 =0.01）
- 净转换率：即在 14 天的期限后仍参与课程的用户 id 的数量（因此至少进行了一次付费）除以点击了“开始免费试学”按钮的唯一 cookie 的数量所得的比率，衡量的是从 cookie 转化为付费用户的比例。（d 最小 =0.0075）

（注：d 最小指的是实际显著性边界，用于衡量指标变动是否达到业务预期）

## 试验设计

### 一、指标选择

#### 1. 不变指标：

- （1）cookie 数，因为 cookie 是分组单元，可随机分配给实验组和对照组，故各组的 cookie

数就应大致相同。

(2) 点击次数，因为只有点击了按钮才会弹出试验的内容，对是否点击按钮没有影响。

(3) 点进概率，因为用于计算该指标的 cookie 数和点击次数都是不变指标。

这里不选用户 ID 数，因为新增的变化会影响一些人报名参加免费试用的信心，从而影响用户 ID 数。不选总转化率、留存率、净转化率是因为它们是我们评估的指标。

## 2. 评估指标：

(1) 总转化率，因为试验的一个目标是减少没有足够时间却参加试学的学生数，而总转化率能反映这一情况。

(2) 留存率，减少因为没有足够时间而离开免费试学的学生数，便提高了留存率。

(3) 净转化率，实验的另一个目标是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，净转化率能反映这一情况。

这里不选用户 ID 数，因为用户是按照 cookie 随机分配到实验组或对照组的，且 cookie 数不完全相同，故采取比率形式即总转化率会更适合。

## 二、预期结果

1. 总转化率有明显降低

2. 留存率有明显提升

3. 净转化率不会出现明显降低

## 三、计算标准偏差（计算过程略）

评估指标的标准偏差

1. 总转化率：0.0202

2. 留存率：0.0549

3. 净转化率：0.0156

## 四、确定样本规模

通过样本容量计算器，在给定  $\alpha = 0.05$ ,  $\beta = 0.2$  的条件下，代入总转化率，留存率和净转化率的基准值及对应实际显著性水平，得到的结果只是各指标每组分析单元的数值，还需要转化为页面浏览量并乘以 2，计算出最高的页面浏览量为 4741212.12，这是由留存率计算得出的，即使每天 40000 的页面浏览量全部转入，也需要近 4 个月时间，实验时间过长，成本过高，故排除留存率作为评估指标，剩余两个指标当中计算出最大的页面浏览量为 685325，按每天 40000 的页面浏览量，一共只需要 18 天，时间合理，故选其作为实验所需页面浏览量。

这里不使用 Bonferroni 校正，因为 Bonferroni 校正在指标间不独立的情况下过于保守，而总转化率和净转化率具备一定相关性，故不采用。

## 五、确定试验持续时间和曝光比例

1. 试验天数：18 天

计算过程见“确定样本规模”。

2. 曝光比例：100%

因为该实验是以 cookie 分组，对于身份的辨识度很低，不用太在意安全和保密问题，同时所作变更只是为了让大家谨慎考虑是否参加，不具危害性，对于优达学城也不至于引起损伤名誉等问题，故风险性很低，可将流量全部转入。

试验分析

一、合理性检查（计算过程略）

	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"	0.0812	0.0830	0.0822	<input checked="" type="checkbox"/>
• Gross conversion				<input type="checkbox"/>
• Retention				<input type="checkbox"/>
• Net conversion				<input type="checkbox"/>

所有不变指标通过合理性检查，意味着实验组和对照组具备可比性。

二、结果分析

1.效应大小检验（计算过程略）

Did you use the Bonferroni correction? <input type="radio"/> Yes <input checked="" type="radio"/> No				
	Lower bound	Upper bound	Statistical significance	Practical significance
• Number of cookies			<input type="checkbox"/>	<input type="checkbox"/>
• Number of user-ids			<input type="checkbox"/>	<input type="checkbox"/>
• Number of clicks on "Start free trial"			<input type="checkbox"/>	<input type="checkbox"/>
• Click-through-probability on "Start free trial"			<input type="checkbox"/>	<input type="checkbox"/>
• Gross conversion	-0.0291	-0.0120	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
• Retention			<input type="checkbox"/>	<input type="checkbox"/>
• Net conversion	-0.0116	0.0019	<input type="checkbox"/>	<input type="checkbox"/>

可以看到在 95%的置信水平下，总转化率同时具备统计显著性和实际显著性，而净转化率都不具备。

2.符号检验（计算过程略）

Did you use the Bonferroni correction? <input type="radio"/> Yes <input checked="" type="radio"/> No		
	p-value	Statistical significance
• Number of cookies		<input type="checkbox"/>
• Number of user-ids		<input type="checkbox"/>
• Number of clicks on "Start free trial"		<input type="checkbox"/>
• Click-through-probability on "Start free trial"		<input type="checkbox"/>
• Gross conversion	0.0026	<input checked="" type="checkbox"/>
• Retention		<input type="checkbox"/>
• Net conversion	0.6776	<input type="checkbox"/>

因为符号检验的 P 值小于 0.05，故总转化率具有统计显著性，即效应大小检验和符号检验结果一致。这里没有使用 Bonferroni 校正，因为评估指标只有两个，不至于使总体的  $\alpha$  值太高，又因为这两个指标具有一定相关性，使用 Bonferroni 校正会导致结果较为保守。

## 结论建议

不建议全面实施实验，因为该举措虽然降低了总转化率，即时间不足却进行试学的学生比例，但净转化率的置信区间左边界-0.0116 超出了它的实际显著性左边界-0.0075，意味着有可能净转化率出现了我们无法接受的下降，此时可以通过增强功效来做进一步研究，手段有增加样本量、适度提升  $\alpha$  水平等。当然，也有必要和决策者商量是否存在战略及成本等考量因素后，再作进一步打算。

## 后续试验

如果预算充足，我会进行一个后续试验，在学员点击退出时询问是否需要与课程老师做一次交谈，有助于帮他们弄清楚无法继续的原因，假设此举会减少想要提前终止课程的学生数量。

这里仅确定试验基本要素：

一、转移单位：用户 ID，因为试验是针对已经注册登陆的用户，故需要 ID 进行跟踪来反映试验效果。

二、评估指标：留存率，即在 14 天的期限过后仍参加课程（因此至少进行了一次付费）的用户 id 数量除以完成登录的用户 id 的数量。因为它反映了学生是否在登陆后能坚持学习的情况，能很好地衡量试验目标。

三、不变指标：

1. 用户 ID 数，即参与免费试学的用户数量。因为用户 ID 为分组单元，可随机分配给实验组和对照组，故各组的 ID 数就应大致相同。

2. cookie 数，即访问课程概述页面的唯一 cookie 的数量，唯一性按天决定。因为用户是由 cookie 转化而来，用户大致相等，说明需要大致相等的 cookie 数，且试验内容是注册之后的事，不会影响该指标。

3. 点击提前终止课程按钮的次数，因为只有点击了按钮才会出现新的选项框。