

## 1. 度量选择

- 不变度量: cookie的数量、点击次数、点进概率

以上三个指标都不会随着"变化"而改变。对照组和实验组中这三个指标都不应该发生显著的变化。

- 评估度量: 总转化率、净转化率


用户ID的数量不能作为评估度量。虽然用户ID数量发生在实验之后,会受到实验影响,可以作为一个评估度量,但是由于实验组和对照组的cookie数量不一定相同,也就是说两组中用户ID数量不同的原因除了我们希望的(受实验影响)之外,还有可能是因为cookie不同。因此我们选择了总转化率(归一化)代替用户ID数量。

留存率不能作为评估度量。因为在计算样本大小和持续时间时,发现若将留存率作为评估度量,所需要的样本数量太多,成本过高,所以将其排除。

- 期望得到的结果: 可以从评估度量中得到标准偏差

## 2. 测量标准偏差

- 根据二项分布公式  $SD = \sqrt{\frac{\hat{P}(1-\hat{P})}{N}}$ , 答案如下图SD所示

F5		=SQRT((B5*(1-B5))/400)			
A	B	C	D	E	F
Unique cookies to view page per day:	40000				
Unique cookies to click "Start free trial" per day:	3200				
Enrollments per day:	660				
Click-through-probability on "Start free trial":	0.08				SD
Probability of enrolling, given click:	0.20625	总转化率P	N=5000*(3200/40000)=400		0.020230604
Probability of payment, given enroll:	0.53	保存率P	N=5000*(3200/40000)*0.53		
Probability of payment, given click	0.1093125	净转换率P	N=5000*(3200/40000)=400		0.015601545

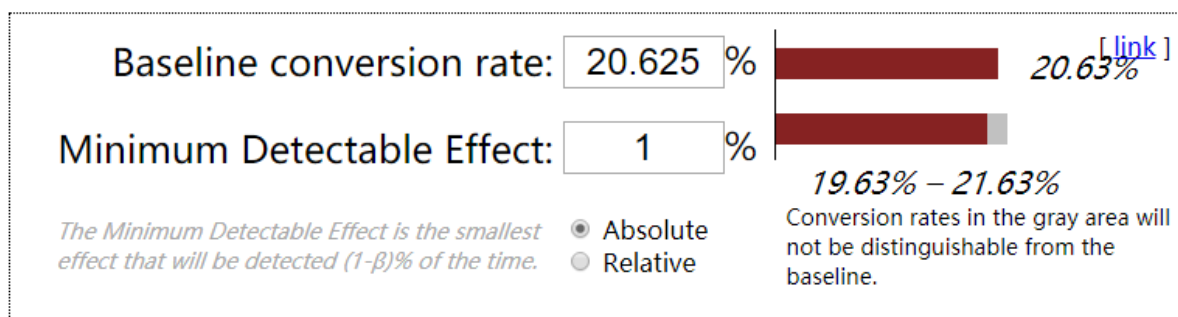
- 分析变异性是analytical variability, 经验变异性是empirical variability。二者相等的条件是“unit of analysis = unit of diversion”。因为unit of analysis是cookie, 而总转化率和净转换率的unit of diversion也是cookie(分母是cookie), 所以非分析估计和经验变异是类似的。

## 3. 规模

### 样本数量和支持

- 总转化率

Question: How many subjects are needed for an A/B test?



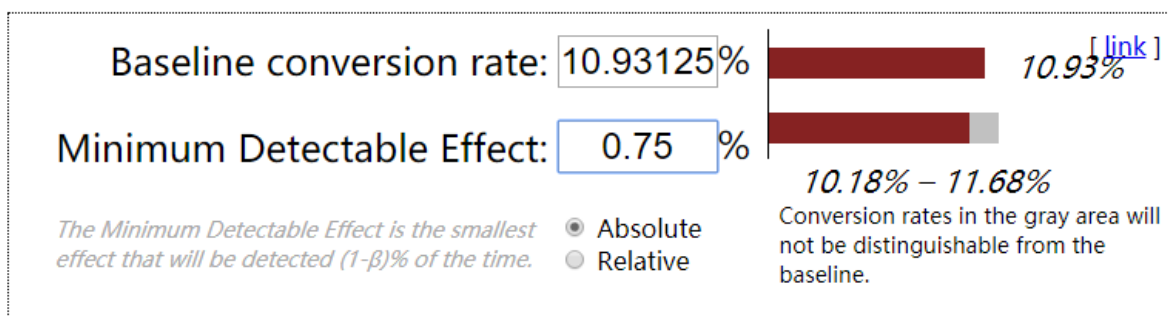
Sample size:  
**25,835**  
per variation

此处得到的是总转化率需要的cookie数, 转化为页面浏览量还需要如下计算  
乘以2表示实验组和对照组

$$SampleSize_{gross-conversion} = \frac{25835 \times 2}{3200/40000} = 645875$$

- 净转换率

Question: How many subjects are needed for an A/B test?



Sample size:  
**27,413**  
per variation

此处得到的是净转换率需要的cookie数, 转化为页面浏览量还需要如下计算

$$SampleSize_{net-conversion} = \frac{27413 \times 2}{3200/40000} = 685325$$

$$\frac{685326}{40000} = 18$$

- 结果: 合理性检查均通过

## 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	0.0000566271	<input checked="" type="checkbox"/>
• Gross conversion				<input type="checkbox"/>
• Retention				<input type="checkbox"/>
• Net conversion				<input type="checkbox"/>

- ### ● 计算过程

A	B	C	D	E	F	G	H	I	J	K
Date	Pageviews	Clicks	Enrollments	Payments						
Tue, Nov 4	9427	743								
Wed, Nov 5	9633	808								
Thu, Nov 6	9842	831								
Fri, Nov 7	9272	767								
Sat, Nov 8	8969	760								
Sun, Nov 9	9697	850								
Mon, Nov 10	10445	851								
Tue, Nov 11	9931	831								
Wed, Nov 12	10042	802								
Thu, Nov 13	9721	829								
Fri, Nov 14	9304	770								
Sat, Nov 15	8668	724								
Sun, Nov 16	8988	710								
total	344660	28325								
subtotal	690203	56703	p----->	0.082154091						
SD	0.000601841	0.002099747	SE计算过程---->	0.075044796	5.79541E-06	4.37001E-07	0.000661061	<-----SE	Xexp/Nexp	0.082182441
margin	0.001179608	0.004115504					0.001295679	<-----margin	Xcont/Ncont	0.082125814
upper	0.5011179608	0.504115504							d	5.66271E-05
lower	0.498820392	0.495884496							upper	0.001295679
observed	0.500639667	0.500467347							lower	-0.001295679

5. 结果分析效应大小检验

- 结果

Effect Size Tests

For each of your evaluation metrics, compute a confidence interval around the difference.

Did you use the Bonferroni correction? ☐ Yes ☒ No

	Lower bound	Upper bound	Statistical significance	Practical significance
• Number of cookies	<input type="text"/>	<input type="text"/>	<input type="checkbox"/>	<input type="checkbox"/>
• Number of user-ids	<input type="text"/>	<input type="text"/>	<input type="checkbox"/>	<input type="checkbox"/>
• Number of clicks on "Start free trial"	<input type="text"/>	<input type="text"/>	<input type="checkbox"/>	<input type="checkbox"/>
• Click-through-probability on "Start free trial"	<input type="text"/>	<input type="text"/>	<input type="checkbox"/>	<input type="checkbox"/>
• Gross conversion	-0.029123358	-0.011986391	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
• Retention	<input type="text"/>	<input type="text"/>	<input type="checkbox"/>	<input type="checkbox"/>
• Net conversion	-0.011604624	0.001857179	<input type="checkbox"/>	<input type="checkbox"/>

- 计算过程

A	B	C	D	E	F	G
Date	Pageviews	Clicks	Enrollments	Payments		
Thu, Oct 23	8324	665	127	56		
Fri, Oct 24	9434	673	220	122		
Sat, Oct 25	8687	691	176	128		
Sun, Oct 26	8896	708	161	104		
Mon, Oct 27	9535	759	233	124		
Tue, Oct 28	9363	736	154	91		
Wed, Oct 29	9327	739	196	86		
Thu, Oct 30	9345	734	167	75		
Fri, Oct 31	8890	706	174	101		
Sat, Nov 1	8460	681	156	93		
Sun, Nov 2	8836	693	206	67		
Mon, Nov 3	9437	788	3785	<--Xcont	2033	<--Xcont
Tue, Nov 4	9420	781	17293	<--Ncont	17293	<--Ncont
Wed, Nov 5	9570	805	3423	<--Xexp	1945	<--Xexp
Thu, Nov 6	9921	830	17260	<--Nexp	17260	<--Nexp
Fri, Nov 7	9424	781	0.208607067	<--Ppool	0.115127485	<--Ppool
Sat, Nov 8	9010	756	0.004371675	<--SEpool	0.003434134	<--SEpool
Sun, Nov 9	9656	825	0.008568484	<--margin	0.006730902	<--margin
Mon, Nov 10	10419	874	-0.020554875	<--d	-0.004873723	<--d
Tue, Nov 11	9880	830	-0.029123358	<--lower	-0.011604624	<--lower
Wed, Nov 12	10134	801	-0.011986391	<--upper	0.001857179	<--upper
Thu, Nov 13	9717	814	总转化率		净转换率	

6. 符号检验

总转化率

- 结果

## Sign and binomial test

Number of "successes": 19

Number of trials (or subjects) per experiment: 23

Sign test. If the probability of "success" in each trial or subject is 0.500, then:

- The one-tail P value is 0.0013  
This is the chance of observing 19 or more successes in 23 trials.
- The two-tail P value is 0.0026  
This is the chance of observing either 19 or more successes, or 4 or fewer successes, in 23 trials.

通过符号检验，与效应大小检验结果一致

two-tail P value = 0.0026 < 0.05

- 计算过程

对照组	实验组	总转化率是否降低
0.195050946	0.153061224	1
0.188703466	0.147770701	1
0.183718372	0.164027149	1
0.186602871	0.166868198	1
0.19474313	0.168269231	1
0.167679222	0.163705584	1
0.195187166	0.162820513	1
0.174050633	0.144171779	1
0.189580318	0.172166428	1
0.191637631	0.177906977	1
0.226066897	0.165509259	1
0.193317422	0.15980025	1
0.190977444	0.190031153	1
0.326894502	0.278335725	1
0.254703329	0.189835575	1
0.22740113	0.220779221	1
0.306982872	0.276264591	1
0.20923913	0.220108696	0
0.265223275	0.27647868	0
0.227520436	0.284340659	0
0.246458924	0.252077562	0
0.22907489	0.204316547	1
0.297258297	0.251381215	1

## 净转换率

- 结果

## Sign and binomial test

Number of "successes": 13

Number of trials (or subjects) per experiment: 23

Sign test. If the probability of "success" in each trial or subject is 0.500, then:

- The one-tail P value is 0.3388  
This is the chance of observing 13 or more successes in 23 trials.
- The two-tail P value is 0.6776  
This is the chance of observing either 13 or more successes, or 10 or fewer successes, in 23 trials.

未通过符号检验，与效应大小检验结果一致

two-tail P value = 0.6776 > 0.05

- 计算过程

对照组	实验组	净转换率是否降低
0.1018923	0.0495627	1
0.0898588	0.1159236	0
0.1045105	0.0893665	1
0.1255981	0.1112455	1
0.0764636	0.1129808	0
0.0996355	0.0774112	1
0.1016043	0.0564103	1
0.1107595	0.095092	1
0.0868307	0.1104735	0
0.1126597	0.1139535	0
0.1211073	0.0821759	1
0.1097852	0.0873908	1
0.0842105	0.105919	0
0.1812779	0.1348637	1
0.1852388	0.1210762	1
0.1468927	0.1457431	1
0.1633729	0.154345	1
0.1236413	0.1630435	0
0.1163735	0.1320495	0
0.1021798	0.092033	1
0.1430595	0.1703601	0
0.1365639	0.1438849	0
0.0966811	0.1422652	0

## 7. 汇总

没有使用Bonferroni校正。因为Bonferroni校正主要适用于n次独立检验,但是本次试验中的总转化率和净转换率并不是独立的,它们是相互关联的,因此使用Bonferroni校正会使得实验结果过于保守,所以不使用Bonferroni校正。

## 8. 建议

通过结果分析得知:

- 总转化率统计性显著,实际性显著,并共同通过符号检验,说明此种"改变"会显著减少因为没有足够时间而离开免费试学的学生数量。
- 净转换率统计性不显著,实际性不显著,未通过符号检验,说明此种"改变"不会显著降低14天期限后仍参加课程的用户数量。

我会启动这个试验。因为通过以上分析,这种"改变"显著减少了因为没有足够的时间而离开免费试学,并因此受挫的学生数量,同时不会在很大程度上(不显著)减少继续通过免费试学和最终完成课程的学生数量。

## 9. 后续试验

如果要减少受挫并想要提前终止课程的学生数量,我会尝试:

实验时,主页上两个按钮。“开始学习”按钮和“免费视频”按钮。

- 点击"开始学习"按钮即让学生先付一定量的定金然后才能注册登录并享受完整功能(导师服务、论坛答疑、观看视频)。如果学生坚持继续学习14天后并付费,所缴纳的定金将作为折扣,即

最后缴纳费用 = 课程原价 - 定金  $\times 2$  以此作为对学生的鼓励。如果学生中途退出, 所缴纳的定金将不会返还给学生。

- 点击“**免费视频**”，学生则可以观看免费视频，但不能使用其他服务。

假设是学生会因为缴过定金而对此课程更加重视, 即使学习中受挫也会因为自己缴纳过的定金和最后的奖励而继续坚持学习，最终完成付费得到奖励。如果这个假设为真，优达学城将改进整体学生的体验和提高导师为能够完成课程的学生提供支持的能力。

本实验的分组单元是cookie。因为在学生决定是否缴纳定金之前，是针对所有的cookie，而不是只针对于注册用户。学生点击开始学习按钮后，将进行注册登录缴纳定金，用户id开始被跟踪。对于点击免费视频的学生，他们的用户id不会在实验中被追踪。

度量：

- 不变度量

cookie的数量:即访问课程概述页面的唯一cookie数量

- 评估度量

总转化率:即学生登录并缴纳定金的id数量除以点击了"开始学习"按钮的唯一cookie的数量的比率。

净转换率:即学生在14天后仍缴纳课程费用并享受折扣的用户id除以点击了"开始学习"按钮的唯一cookie的数量的比率。