1. Experiment Design

Sr. No	Invariant Metrics	Evaluation Metrics	Metrics not used at all
1	Number of cookies	Gross Conversion	Number of user-ids
2	Number of Clicks	Net Conversion	Retention
3	Click through Probability (CTP)		

Invariant Metrics

The change we made for this experiment is in the enrolling stage i.e. between 'Start Free Trail' and actual enrollment. As such, we expect the metrics after the change is invoked or is presented to be affected and not those before.

The following metrics (***number of cookies, number of clicks and CTP)*** measure the stats before the supposed change is presented to the user. As such, those metrics should remain relatively unchanged before or after the experiment is done. The experiment should in no way affect those metrics.

Number of cookies is just measuring the page views. The experiment should not affect the number of visitors to the site and so is considered as invariant metric.

Number of Clicks is just measuring the number of unique users that click 'Start free Trail' button. Again, the experiment should not affect this metric.

CTP is just a measure of the probability that a click for pageview. This is just a combination of the other two metrics and thus should remain invariant.

Evaluation Metrics

Evaluation metrics are those metrics that we expect to change or be affected as a result of the experiment. In other words, these are the metrics we intended to use to measure the success of the experiment.

The general hypothesis of the experiment is '*to reduce the number of frustrated students who left the free trial because they didn't have enough time—without significantly reducing the number of students to continue past the free trial and eventually complete the course*.'

Gross Conversion metrics is used to measure the first part of the hypothesis i.e. gross conversion is the measure of the number of those who actually enrolled in the free trail to those who clicked the 'start free trail'. Thus, with this experiment we are expecting the number of those students who enroll in the free trail to be less than those that are not exposed to the experiment (control group). By reducing the enrollees, the expectation is the highest percentage of those enrollees would untimely go beyond the train and start the first paid course thus reducing frustrated students.

However, by reducing the actual numbers of students that enroll in the free trail we should not significantly affect the number of students that go pass the free trail and start a paid course. In other words the experiment should try to affect only the would be frustrated students i.e. students with few hours to spare to the course. The metrics used to measure this phenomenon is Net conversion. It measures the number of students that start a paid class to the number of visits to the 'start free trail'. The expectation is that this metric shouldn't be affected significantly (the second part of the hypothesis).

Metrics not used at all

Number of user Ids- this metrics is already included in the Gross conversion. Number of user Ids on its own doesn't give much information but when combined with the number of clicks i.e. gross conversion, it can tell us the effect of the experiment. Thus, the information is redundant and thus not used.

Retention- this metrics was initially used as one of the evaluation metrics. However, the use of this metrics resulted in lengthy experiment i.e. the size requirement become too big. Thus, I decide to drop it. Besides, we are able to measure our hypothesis with the above evaluation metrics and so there was no need for this metrics.

Expected Results

As mentioned above, the expectation is

- Compared to the control group, the gross conversion in the experiment group should be significantly less.
- There net conversion in the experiment group should not be significantly less than those in the control group.

According to the hypothesis, I would launch the change under these circumstances.

2. Measuring Standard Deviation

Gross Conversion: **0.0202**

Net Conversion: **0.0156**

I expect the analytical and empirical results to be similar. In our case we have used the formulas for normal distribution and expect the empirical results to not be significantly different from those.

3. Sizing

I expect the analytical and empirical results to be similar. In our case, are using the same unit for analysis and diversion- Cookies. As such, we expect the results obtained through empirical and analytical result to be similar.

Number of Samples vs Power

No of page views: **685325** (No Benferroni correction)

Duration vs Exposer

Fraction of Traffic Exposed: **100%**

Length of Experiment: **18**

The experiment has very low risk level. In my case I have chosen to divert 100% of my traffic for this experiment and complete the experiment in short period of time.

Experimental Analysis

Number of cookies: [**0.4988, 0.5012**], observed: **0.5006,** Sanity check: **Pass**

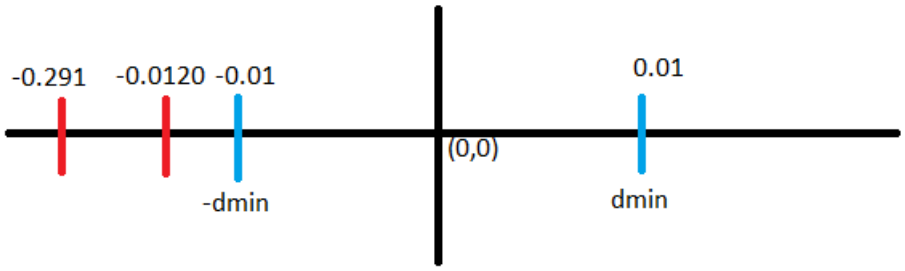
Number of clicks on 'Start Free Trail':[**0.4959, 0.5041**], Observed: **0.5005**, Sanity Check: **Pass**

CTP: [**-0.0013, 0.0013**], observed: **0.0000,** Sanity check, **Pass**

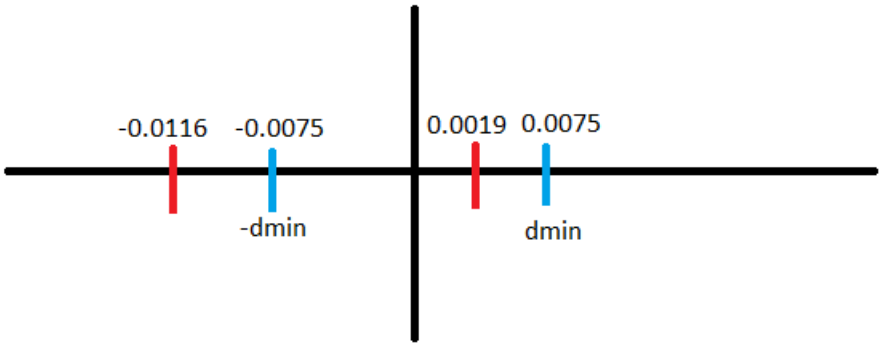
4. Result

Effect size tests

Gross Conversion:[**-0.0291, -0.0120**], Statistical Significance: **YES** Practical Significance: **YES**



Net Conversion:[**-0.0116, 0.0019**], *Statistical Significance: **NO** * Practical Significance: **NO**



For net conversion, the result is not statistically or practically significantly. However, looking at the lower bound of the confidence interval, it is '*possible*' that this could affect us significantly in practical sense i.e. the reduction in the number of students who actually complete the course be reduced significantly though less likely.

Sign Test

Gross Conversion: **0.0026,** Statistical Significance: **YES**

Net Conversion: **0.6776,** Statistical Significance: **NO**

Summary

In our experiment we didn't use Benferroni correction because we need all the metrics to meet our expectation for the change to be launched. Gross Conversion need to be statistically and practically significant while at the same time Net conversion is statistically and practically insignificant. I would have used Benferroni correction if it a case where the change would be launched if any of the two metrics meet our expectation.

5. Recommendation

I would repeat the experiment with more power. The experiment met our first hypothesis i.e. reduced significantly the number of students who enroll in the free trail, as indicated by Gross Conversion. At the same time it did not reduced significantly the number of students going beyond the 14 day trail and starting a paid course, as indicated by Net Conversion. However, there was a catch on this one- the confidence level has shown us it is probable that this experiment might reduce the number of students who start paid course though less likely. Hence, to be confident on this result and to avoid unintended consequence, we need to repeat the experiment with more power.

If further experiment doesn't meet our second hypothesis i.e. if there is likelihood that our net conversion might be affected significantly, then we don't have to launch the experiment.

6. Follow up experiment

One of the main reasons students enroll in such kind of programs is either to advance their career or embark on a new career path. On the other hand, one of the reasons for quitting such a program is a financial issue. At the moment, Udacity is addressing both issues by providing a financial incentive for early completion and by guaranteeing a job offer for some of the programs.

The job guarantee is a good incentive to complete the program however, it is only provided to just few of the programs. For instance, Data Analysis is not one of them. If not a job offer, what about guaranteeing an online or live meet-up/interview with recruiters from various companies? I think giving the enrollee a chance to pitch his/her skills directly to a recruiter is a big deal and could be an incentive

for the student to complete the program. This would be great for those courses that Audacity cannot guarantee job placement/offer.

Here is how I would go about this experiment:

Description of the Experiment

To do this experiment, we are going to divide the test group in to Control and experiment group. Once the students enroll in to the program, those under the experiment group will be presented with a screen message that indicates a guarantee for meet-up/online with recruiters from several companies up on successful completion. This should not just be a screen notice that shows up only once. The student needs to me constantly motivated and so, this incentive should be there as part of his homepage.

Hypothesis

The expectation is that we will see a higher completion rate of those students that were presented with the incentive than the other group. In other words, we expect the incentive to reduce early cancelations.

Metrics

Invariant metrics: User ID

There shouldn't be a significant difference between the number of students who enroll in the control group and experiment group. The experiment takes place only after the student has enrolled and so enrollment rate should not be affected.

In our case User ID is the metrics used to measure the enrollment rate. It will be our invariant metrics and will be used for sanity check.

Evaluation Metrics: Completion probability

As indicated in the hypothesis, we expect more students to complete the course if they are presented with the incentive. We will use Completion probability as the metrics to evaluate this hypothesis/expectation.

- **User ID:** Number of unique students that enrolled in the program
- **Completion probability:** That is, number of user-IDs to complete the program divided by the number of user IDs to enroll in the program

Unit of Diversion

User ID is used as a unit of diversion. This would make the unit of diversion similar to the unit of analysis and thus reducing variability. This will ultimately decrease the sizing requirement to meet the confidence interval. This will end up helping us in completing the experiment in shorter period of time.