Template Format

This template can be used to organize your answers to the final project. Items that should be copied from your answers to the quizzes should be given in blue.

Experiment Design

Metric Choice

Invariant Metrics:

1. **Number of cookies :** Since change is introduced after a user clicks on "start free trial" button in control and experiment group. So, number of unique cookies to view the course overview page shouldn't vary much in control and experimental group.
2. **Click through probability:** Again, the user of udacity experiences change after he/she clicks on "start free trial" button, this metric should also be an invariant metric.

Evaluation Metrics:

1. **Gross conversion:** Since, the change is introduced after clicking the start free trial button i.e. a message flashes informing about the time needs of course in experiment group, it will lead to change in number of users who enroll in a course. Thus, this metric serves as an evaluation metric. Also, enrollment is compared on the number of unique id's to click start free trial button which is unaffected by the change, So, it will serve as a good evaluation metric.
2. **Net Conversion:** The prime motive of this study is to gauge the effect of experiment on conversions(i.e. Payment after enrollment, past 14 days). Thus, this metric is also an evaluation metric.

The hypothesis was that the introduced change might set clearer expectations for students upfront, thus reducing 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. If this hypothesis held true, Udacity could improve the overall student experience and improve coaches' capacity to support students who are likely to complete the course.

It is expected that gross conversion will decrease, since the pop up message will state the time requirements of course.However, the hypothesis assumes that net conversion will either improve or stay same.

Thus,gross conversion should decrease and become less than d(min) while net conversion should not decrease, i.e. stays in the confidence interval. But, also CI shouldn't be negative since it implies a negative effect on net conversion.

Remaining Metrics :

1. **Number of User ID's :** The metric is affected by introduction of change.So, it cannot be an invariant metric. Also, the metric isn't a good measure of change proposed in the hypothesis. Though, it can gauge the first part of the hypothesis but it isn't normalised. Thus, it could probably be used as an evaluation metric, but, gross conversion is more suitable option.

2. **Number of clicks:** Again, it won't serve as a good invariant metric since number of unique clicks on "Start free trial", as we are more interested to know how many students progress to second level of the funnel and in such cases when we have to gauge impact , we opt for probability. Thus, click through probability would be a better choice as an invariant metric. It can't be chosen as evaluation metric because it is unaffected by change.
3. **Retention:** Metric is affected by introduction of change.So, it cannot serve as an invariant metric. Retention's unit of analysis (user-ids) differs from the unit of diversion (cookies), it's also going to have a pretty high variance compared to the other metrics.Also, if I choose it as an evaluation metric, time required to run the experiment increases heavily. So, it cannot be used as an evaluation metric.

Measuring Standard Deviation

Standard deviations :
1. Gross Conversion : 0.0202
2. Net Conversion : 0.0156

Q- Whether analytic estimate would be comparable to empirical variability can be decided on the basis of unit of analysis and unit of diversion.
If, the units of analysis and diversion are different then analytical estimate and empirical variability would be different and vice versa.
Both, evaluation metrics have unit of diversion and analysis as "cookie". So,analytical estimate and empirical variability would be same.

Sizing

Number of Samples vs. Power
No, I will not use the Bonferroni correction during analysis phase.
Number of pageviews :685325
Duration vs. Exposure
Fraction of traffic I would divert to this experiment : 1
Days needed to run the experiment : 18.

Q- Give your reasoning for the fraction you chose to divert. How risky do you think this experiment would be for Udacity?

As suggested in the review, the experiment does not any sensitive data or chances of causing harm to someone, it is a case of minimal risk. Also, financial loss would be approximately same with full diversion or partial diversion of traffic.Probably, introduction of change might help students to avoid possible distress. Therefore, I would divert 100% traffic.

Experiment Analysis

Sanity Checks

For each of your invariant metrics, give the 95% confidence interval for the value you expect to observe, the actual observed value, and whether the metric passes your sanity check. (These should be the answers from the "Sanity Checks" quiz.)

Metric :Number of cookies
CI intervals: [0.4988,0.5011]
Observed Value: 0.5006

Metric :Click through probability on "Start free trial"
CI intervals: [-0.00129,0.00129]
Observed Value: -0.000056

Q- For any sanity check that did not pass, explain your best guess as to what went wrong based on the day-by-day data. Do not proceed to the rest of the analysis unless all sanity checks pass.
All sanity checks passed.

Result Analysis

Effect Size Tests

For each of your evaluation metrics, give a 95% confidence interval around the difference between the experiment and control groups. Indicate whether each metric is statistically and practically significant. (These should be the answers from the "Effect Size Tests" quiz.)

Metric :Gross Conversion
CI intervals: [-0.0291,-0.0120]
Practically and statistically significant.

Metric :Net Conversion
CI intervals: [-0.0116,0.0019]
Neither practically nor statistically significant

Sign Tests

For each of your evaluation metrics, do a sign test using the day-by-day data, and report the p-value of the sign test and whether the result is statistically significant. (These should be the answers from the "Sign Tests" quiz.)

Metric :Gross Conversion
P- value :0.0026

Summary
Q- State whether you used the Bonferroni correction, and explain why or why not. If there are any discrepancies between the effect size hypothesis tests and the sign tests, describe the discrepancy and why you think it arose.
No, I didn't use Bonferroni correction because in this experiment we want both of our metrics(highly correlated) to meet our expectations and in such case using Bonferroni correction isn't advisable.The Bonferroni correction is designed to avoid type I errors while in this case, we need all the metrics to meet criteria in order to launch i.e. we might reject some metrics by mistake, and that is a type II error. And, Bonferroni correction is useless in such cases.Benjamini & Yekutiely adjustment could be a probable option.
No, there wasn't any discrepancies between the effect size hypothesis tests and the sign tests. Gross conversion was statistically significant and same was reflected in sign test. Similarly, Net conversion was statistically insignificant and it was reflected correctly in sign test.

Recommendation
Q-Make a recommendation and briefly describe your reasoning.

Hypothesis for this experiment was that, this might set clearer expectations for students upfront, thus reducing 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. If this hypothesis held true, Udacity could improve the overall student experience and improve coaches' capacity to support students who are likely to complete the course.
Thus, the evaluation metrics(gross conversion and net conversion) should have reflected the same to adopt this change keeping the invariant metrics fairly stable over the controlled and experiment groups.
HO(1): Gross conversion stays same between control and experiment group.
HO(2): Net conversion stays same between control and experiment group.
HA(1): Gross conversion decreases towards negative side and lower than d(min) with CI of 95%

We expected, HA(1) to be true and HO(2) to hold. Results for gross conversion are good and initially net conversion seems to favour the change(being not significant statistically and practically)
But negative CI bounds indicate that the experiment has lead to decrease in net conversion(a negative effect) which isn't desirable. Also, the confidence interval of the net conversion includes the negative of the practical significance boundary(-0.0075). So, it might be possible

that net conversion might have went down by an amount that would negatively affect the business.

So, Recommendation is to drop this change.

Follow-Up Experiment

Q- Give a high-level description of the follow up experiment you would run, what your hypothesis would be, what metrics you would want to measure, what your unit of diversion would be, and your reasoning for these choices.

As a follow up experiment, I would like to flash a message as a user clicks on "Start free trial", saying, basic requirements (list of basic requirements for concerned course) with two options:
1. Learn basics - If user clicks on this link, it will lead to supporting courses to teach the basics.
2. Continue - It will lead to payment page.

Hypothesis: Showing the basic requirements of course will help students to understand that with their current knowledge, it could be easy or difficult for them to progress in the nanodegree. Thus, they can opt to learn basics first and enroll after that if they feel that they need to learn basics first rather than going for free trial and end up being frustrated.

Metrics : Similar to the given experiment.

No. of cookies and click through probability as invariant metrics

Gross conversion and Net conversion as evaluation metrics.

Cookies as unit of diversion and analysis.

I don't want to affect the traffic coming on udacity, so the basic requirements message will flash only after knowing that student is willing to enroll i.e. he/she clicks on the "Start free Trial" button. Thus, I chose the similar metrics .