# Design Process and Calculations

## Final Project Baseline Values

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

Probability of enrolling, given click: 0.20625

Probability of payment, given enroll: 0.53

Probability of payment, given click: 0.1093125

# Experiment Design

### Metric Choice

**Invariant Metrics:**

1. Number of cookies: Number of unique cookies to view the course overview page
2. Number of clicks: Number of unique cookies to click the "Start free trial" button
3. Click-through-probability: number of unique cookies to click the "Start free trial" button divided by number of unique cookies to view the course overview page

**Evaluation Metrics:**

1. Gross Conversion:  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

2. Net Conversion: Number of user-ids to remain enrolled past the 14-day boundary (and thus make at least one payment) divided by the number of unique cookies to click the "Start free trial" button

For each metric, explain both why you did or did not use it as an invariant metric and why you did or did not use it as an evaluation metric. Also, state what results you will look for in your evaluation metrics in order to launch the experiment.

### Measuring Standard Deviation

Sample size of 5000 cookies visiting the course overview page:

$$Gross\ Conversion = \frac{Number\ of\ user-ids\ to\ enroll}{Number\ of\ unique\ cookies\ to\ click\ "Start\ free\ trial"}$$

$$Net\ Conversion = \frac{Number\ of\ user-ids\ to\ make\ a\ payment}{Number\ of\ unique\ cookies\ to\ click\ "Start\ free\ trial"}$$

$$S.D.(Gross\ Conversion) = \sqrt{\frac{0.020625(1-0.20625)}{3200/8}} = 0.0202$$

$$S.D.(Net\ Conversion) = \sqrt{\frac{0.109325(1-0.109325)}{3200/8}} = 0.0156$$

Gross Conversion: 0.0202
Net Conversion: 0.0156
(*Retention was chosen initially, but the number of pageviews required to measure retention were too many to measure over a reasonable length of time for the experiment. So retention was removed as an evaluation metric.*)
For each of your evaluation metrics, indicate whether you think the analytic estimate would be comparable to the the empirical variability, or whether you expect them to be different (in which case it might be worth doing an empirical estimate if there is time). Briefly give your reasoning in each case.

## Sizing

### Number of Samples vs. Power

Number of pageviews needed (using alpha=0.05 and beta=0.2): 685325
Bonferroni correction was not used since it can be over-conservative and only two evaluation metrics were chosen.

### Duration vs. Exposure

**Fraction of Udacity's traffic to divert to this experiment: 0.5**
**Given this, the number of days Udacity will need to run the experiment: 35**

$$Number\ of\ days = \frac{685325}{20000} = 34.266 \cong 35$$

## Experiment Analysis

### Sanity Checks

95% confidence interval for invariant metrics and sanity checks based on observed values

| | Lower Bound | Upper Bound | Observed | Passes |
|---|---|---|---|---|
| Number of cookies | 0.4988 | 0.5012 | 0.5006 | Yes |
| Number of clicks on "Start free trial" | 0.4958 | 0.5041 | 0.5004 | Yes |
| Click-through-probability on "Start free trial" | -0.0013 | 0.0013 | 0.00006 | Yes |

Number of Cookies:

$$Standard\ Error = \sqrt{\frac{0.5 \times 0.5}{345543 + 344660}} = 0.000602$$

$m = z \times Standard\ Error = 1.96 \times 0.000602 = 0.00118$
95% Confidence Interval = (0.5-m, 0.5+m) = (0.4988, 0.5012)

Number of clicks on "Start free trial":

$$Standard\ Error = \sqrt{\frac{0.5 \times 0.5}{28378 + 28325}} = 0.00209$$

$m = z \times Standard\ Error = 1.96 \times 0.00209 = 0.00411$
95% Confidence Interval = (0.5-m, 0.5+m) = (0.4958, 0.5041)

Click-through-probability on "Start free trial":

$$\hat{p}_{pool} = \frac{X_{control} + X_{exp.}}{N_{control} + N_{exp.}} = \frac{28378 + 28325}{345543 + 344660} = 0.08215$$

$$SE_{pool} = \sqrt{0.08215 \times (1 - 0.08215) \times (\frac{1}{345543} + \frac{1}{344660})} = 0.00066$$

$m = z \times Standard\ Error = 1.96 \times 0.00066 = 0.001296$
95% Confidence Interval: (-0.0013, 0.0013)

**Since all the sanity checks pass, we can proceed with the rest of the analysis.**

## Result Analysis

### Effect Size Tests

95% Confidence interval around the difference between the experiment and control groups for Evaluation metrics.

| | Lower Bound | Upper Bound | Statistical Significance | Practical Significance |
|---|---|---|---|---|
| **Gross Conversion** | -0.0291 | -0.0119 | Yes | Yes |
| **Net conversion** | -0.0116 | 0.0018 | No | No |

**Gross Conversion:**

$$\hat{p}_{exp} = \frac{Enrollments}{Clicks} = \frac{3423}{17260} = 0.1983$$

$$\hat{p}_{control} = \frac{Enrollments}{Clicks} = \frac{3785}{17293} = 0.2188$$

$\hat{d} = \hat{p}_{exp} - \hat{p}_{control} = -0.02055$

$$\hat{p}_{pool} = \frac{X_{control} + X_{exp.}}{N_{control} + N_{exp.}} = \frac{3785 + 3423}{17293 + 17260} = 0.20861$$

$$SE_{pool} = \sqrt{0.20861 \times (1 - 0.20861) \times (\frac{1}{17293} + \frac{1}{17260})} = 0.004372$$

$m = z \times Standard\ Error = 1.96 \times 0.004372 = 0.008568$
95% Confidence Interval: (-0.0291, -0.0116)

**Net Conversion:**

$$\hat{p}_{exp} = \frac{Enrollments}{Clicks} = \frac{1945}{17260} = 0.11268$$

$$\hat{p}_{control} = \frac{Enrollments}{Clicks} = \frac{2033}{17293} = 0.11756$$

$\hat{d} = \hat{p}_{exp} - \hat{p}_{control} = -0.00487$

$$\hat{p}_{pool} = \frac{X_{control} + X_{exp.}}{N_{control} + N_{exp.}} = \frac{2033 + 1945}{17293 + 17260} = 0.11512$$

$$SE_{pool} = \sqrt{0.11512 \times (1 - 0.11512) \times (\frac{1}{17293} + \frac{1}{17260})} = 0.003434$$

$m = z \times Standard\ Error = 1.96 \times 0.003434 = 0.00673$
95% Confidence Interval: (-0.0116, 0.0018)

**Sign Tests**

| | **p-value** | **Statistical Significance** |
|---|---|---|
| **Gross Conversion** | 0.0026 | Yes |
| **Net conversion** | 0.6776 | No |

**Summary**

**Recommendation**

I will recommend not to launch the change a do a follow-up experiment since only of the evaluation metrics showed a statistical significant change.