

Project P7: Design an A/B Test

Peter Eisenschmidt

April 13, 2017

1 Experiment Design

1.1 Metric Choice

The following metrics have been selected as **Invariant Metrics**:

Number of cookies	The number of unique cookies to visit the page should not be affected by the experiment, as someone visiting the course overview page has not seen the changes yet.
Number of clicks	The same applies to the number of clicks on the "Start Free Trial" button; there should not be any impact of the experiment on this metric
Click-through-probability	As the CTR is defined as the number of unique cookies to click the "Start free trial" button divided by the number of unique cookies to view the course overview page (both of which are invariant metrics), the click-through-probability is also an invariant metric

These three metrics are not impacted by the experiment and hence one can expect similar distributions between control and experiment groups.

The following metrics have been selected as **Evaluation Metrics**:

Gross conversion	Being defined as the number of user-ids to complete checkout and enroll in the free trial divided by the number of unique cookies to click the "Start free trial" button, one would expect a lower gross conversion for the experiment as for the control group. The goal of the tested change is to reduce the number of frustrated students, so you could expect that students that are likely to drop out with the current design are filtered out early and do not complete the checkout.
Retention	Similarly, you would expect an increased retention as a result of the experiment, as the number of students that complete the checkout should reduce. At the same time, the number of students to make at least one payment should remain the same.

Net conversion Net conversion is the combination of the two previously mentioned metrics. It is expected that net conversion remains the same for both control and experiment group, as the number of students to remain enrolled past the 14-day boundary as well as the number of unique cookies to click the "Start Free Trial" Button should remain the same.

For each of these metrics, a practical significance boundary d_{min} is defined. This indicates the minimum difference that needs to be observed between control and experiment group in order to determine whether the change is meaningful or not. This is important for the decision to whether or not launch the change.

For the above given evaluations metrics, the practical significance boundaries are $d_{min} = .01$ (for gross conversion and retention) and $d_{min} = .0075$ (for net conversion)

1.2 Measuring Standard Deviation

The analytical estimate of the standard deviation can be calculated as follows:

$$\sigma = \sqrt{\frac{p(1-p)}{N}} \quad (1)$$

where the probabilities are given in the baseline values:

- Probability of enrolling, given click (Gross Conversion): $p = .20625$
- Probability of payment, given enroll (Retention): $p = .5300$
- Probability of payment, given click (Net Conversion): $p = 0.1093125$

Given that the sample size to visit the course overview page is 5000 cookies, the number of units of analysis for each metric can be calculated as follows. For gross conversion, it is given by:

$$N = \frac{PageViews \times Cookies_{ClickFreeTrial}}{Cookies_{ViewPagePerDay}} = \frac{5000 \times 3200}{40000} = 400 \quad (2)$$

For retention it can be calculated as:

$$N = \frac{PageViews \times Enrollments}{Cookies_{ViewPagePerDay}} = \frac{5000 \times 660}{40000} = 82.5 \quad (3)$$

For net conversion, it is the same as gross conversion:

$$N = \frac{PageViews \times Cookies_{ClickFreeTrial}}{Cookies_{ViewPagePerDay}} = \frac{5000 \times 3200}{40000} = 400 \quad (4)$$

This results in the following standard deviations:

- Gross Conversion: $\sigma = .0202$
- Retention: $\sigma = .0549$
- Net Conversion: $\sigma = .0156$

1.3 Sizing

Number of Samples vs. Power

Duration vs. Exposure

2 Experiment Analysis

2.1 Sanity Checks

2.2 Result Analysis

2.3 Recommendation

3 Follow-Up Experiment