

Project 7 A/B Testing

Kosuke Murakami

Experiment Design

Metric Choice

I used Number of cookies and Click-through-probability as invariant metrics. Evaluation metrics are Gross conversion and Net conversion.

Number of cookies: since unit of diversion is cookie, this should be invariant metric.

Number of user-ids and Number of clicks: these metrics are might be good invariant metrics but since these are not unit of diversion, these might not be randomly assigned to control and experiment group.

Click-through-probability: I chose this metric as invariant metric because showing new message does not affect Click-through-rate.

Gross conversion: I chose this one as evaluation metric since the unit of analysis and the unit of conversion is the same. In addition, gross conversion can measure whether showing message is effective for reducing frustrated students who do not have much time to continue paying course.

Retention: The unit of analysis and the unit of diversion are different. It is not suitable for evaluation metric.

Net conversion: The unit of analysis and the unit of diversion is same. Furthermore, it is helpful to confirm showing new message do not significantly reduce the number of students who pass free trial and continue to take paying courses.

The results I am looking for to launch the experiment are that Gross conversion indicates statistically and practically significance and net conversion do not suggest any difference between control and experiment group.

Measuring Standard Deviation

Standard deviations for gross conversion and net conversion are 0.0202 and 0.0156 respectively.

Both of evaluation metrics have the same unit of diversion and unit of analysis, so I expect the analytic estimate would be comparable to the empirical variability.

Sizing

Number of Samples vs. Power

I did not use Bonferroni correction, require 685275 pages to conduct an analysis.

Duration vs. Exposure

Fraction is 0.5. The experiments need 40 days.

Running on all traffic seems like risky since if something would happen in showing new message, all users could get in trouble. Therefore I chose to divert 0.5 of all traffic to the experiment.

Experiment Analysis

Sanity Checks

Number of cookies

Lower bound: 0.4988

Upper bound: 0.5012

Observed: 0.5006

Sanity checks:Pass

Click-through-probability on "Start free trial"

Lower bound: -0.0012

Upper bound: 0.0013

Observed: 0.0000

Sanity checks:Pass

Result Analysis

Effect Size Tests

Gross conversion

Lower bound: -0.0291

Upper bound: -0.0120

Statistically and practically significance

Net conversion

Lower bound: -0.0116

Upper bound: 0.0018

Not statistically significance

Not practically significance

Sign Tests

Gross conversion

p-value: 0.0026

Statistically significance

Net conversion

p-value: 0.6776

Statistically not significance

Summary

First of all, I expected only gross conversion to be statistically and practically significance and did not expect net conversion to be statistically significance. If I use bonferroni correction, the result would be too conservative, so I might not confirm the result of statistically difference in gross conversion. Second, evaluation metrics are independent each other and there are only two metrics. These facts make me consider not using bonferroni correction.

There are no discrepancies between effect size tests and sign tests.

Recommendation

Showing new message statistically and practically reduce the number of students who enroll in free trial courses without reducing the number of students who stay in paying courses after trial courses. Therefore, based on my A/B testing, we should change to showing new message when a student starts free trial course.

Follow-Up Experiment

In my A/B testing, the number of students who completed courses were not considered.

Therefore I would like to measure whether students successfully study more than 5 hours per week during free trial weeks. Hypothesis is that showing the new message significantly reduce the number of students who frustrated with commitment to classes.

I need to make study time metric: That is, the number of students who study more than 10 hours(5 hours multiplied by 2 weeks) during free trial divided by number of unique cookies to click the "Start free trial" button. Unit of diversion is cookie because I can track the number of page viewers.