# Experiment Design

## Metric Choice

**Invariant Metrics**

* Number of cookies: overall population is same
* Click-through-probability
* Number of clicks: population invariance, to make sure there’s relatively same number of cookies in each groups

**Evaluation Metrics**

* Gross conversion: to check if the suggestion will reduce the number of people who wants to enroll the free trial
* Net conversion: check whether the change would influence the number of clients who pass the free trial but still enrolled

**Expected Results**

Reduced gross conversion, increased/no change of new conversion

## Measuring Standard Deviation

SD of Gross Conversion: 0.0202

SD of Net Conversion: 0.0156

## Sizing

### Number of Samples vs. Power

<https://www.evanmiller.org/ab-testing/sample-size.html>

Total Pageviews for Gross Conversion: 645875

Total Pageviews for Net Conversion: 685325

Did not use Bonferroni correlation, because it yielded too small of alpha values for each test, thus needed more pages. It was too conservative.

### Duration vs. Exposure

Choosing 90% and run the experiment for 20 days

Although there's no other experiments running at the same time, but it is usually not safe to put all (100%) of traffic into consideration, because 1)not too much days difference for 90% and 100%, 2)if put all traffic, cannot tell if other issues affect the results

# Experiment Analysis

## Sanity Checks

**Number of Clicks**

Expected value of clicks: 0.5

Observed value of clicks: 0.5005

The 95% Confidence Interval of Clicks: (0.4959, 0.5041)

**Number of Cookies**

Expected value of cookies: 0.5

Observed value of cookies: 0.5006

The 95% Confidence Interval of cookies: (0.4988, 0.5012)

**Click-through-probability**

Expected value of Click-through-probability: 0

Observed value of Click-through-probability: 0.0001

The 95% Confidence Interval of Click-through-probability: (-0.0013, 0.0013)

Passed the sanity check

## Result Analysis

### Effect Size Tests

At 95% CI level:

Gross Conversion is statistically significant and also practically significant

Net Conversion is neither statistically significant nor practically significant

### Sign Tests

According to Sign Test:

Gross Conversion is statistically significant (two tail p-value = 0.0347 < 0.05)

Net Conversion is statistically significant (two tail p-value = 0.0026 < 0.05)

### Summary

Yes, I used Bonferroni correction, because the two evaluation metrics are not independent

Under Bonferroni correction:

Gross Conversion: p-value = 0.0347 > 0.025 -> not statistically significant

Net Conversion: p-value = 0.0026 < 0.025 -> statistically significant

There is also a discrepancy between the effect size hypothesis tests and the sign tests:

GC: statistically significant and also practically significant in hypothesis tests but not statistically significant in sign tests

NC: neither statistically significant nor practically significant in hypothesis tests but not statistically significant in sign tests

Reasons:

1) Net Conversion of some control groups are high -> relatively low number of click cookies or low number of payments that day

2) NC of some experiment groups are low -> low number of payments

-> the overall invariant matrix matched but the day-by-day invariance metrics did not match with two groups

## Recommendation

**No Launch**

Because the overall hypothesis tests are not fully practically significant, and didn't match the sign tests results

# Follow-Up Experiment

#### **Experiment Overview**

Fully-refund after free trial within 7days, half-refund after free trial within 14 days

At the end of free trial, reminder the user they can still in 7 days with no reason fully refund if they cancel subscription

#### **Hypothesis**

The calcellation rate at the end of free trial will decrease with more people enroll in the course fully paid

#### **Unit of Diversion**

cookies

#### **Main Evaluation Metrics**

Retention: number of user-ids to remain enrolled past the 14-day boundary (and thus make at least one payment) divided by number of user-ids to complete checkout. (dmin=0.01)