**AB TEST Report**

**Author: Rocha**

**Date: April 2017**

**Experiment Design**

**Metric Choice**

 **Number of cookies:**  Invariant Metrics. We should have a proper variable to record the information before the change happen. Among other metrics we chosen, we should get consistent because our unit of diversion is cookie.

 **Number of user-ids:** Not chosen. Although user-id is a good metric, we should get consist with other metric.

 **Number of clicks:**Invariant Metrics. This variable could let us measure the result and it consistent with our unit of diversion.

 **Click-through-probability:** Invariant Metrics. This metrics is just a derivative of our two other invariant variables.

 **Gross conversion:**Evaluation Metrics. This one is a measurement of our changes result in the “free trial Screenshot”.

 **Retention:**Evaluation Metrics. This one is a measurement of our changes result.

 **Net conversion:**Evaluation Metrics. It measures the result of whether our changes enhance retention rate without reduce the number of user a lot.

**Measuring Standard Deviation**

For the sheet of baseline is 40,000 unique cookies to view the page per day. And in our question it’s 5,000 cookies. First, we assume our two evaluation metrics follows binomial distribution. Then the standard deviation is . In 5000 samples, we get our number of enrollment is 82.5. So our **Retention** is 0.0549. The same as our **Net conversion**. In 5000 samples, the number of unique cookie click on “start free button” is 400. So our result is 0.0156.

**Sizing**

**Number of Samples vs. Power**

I don’t use the Bonferroni correction. is 5% , is 20%.

**Gross Conversion**: Baseline conversion rate is 20.625% and the minimum detectable effect is 1% .Our sample size needed is 25,835 / (320/40,000) = 645,875

**Retention**: Baseline conversion rate is 53% and the minimum detectable effect is 1% . Our sample size needed is 39,115 / (660/40,000) \* 2 = 4,741,212.

**Net Conversion**: Baseline conversion rate is 20.625% and the minimum detectable effect is 1% . Our sample size needed is 27,413 / (3200/40,000) \* 2 = 685,325

And my number of pageview is 4,741,212. Because of this number is so large, so we don’t choose the Retention as our evaluation metrics. So our sample size we used is 685,325.

**Duration vs. Exposure**

From above, our sample size is 685,325. Because divert the traffic doesn’t have moral effect so we can divert 100% of the traffic to our experiment group. Our daily number of pageviews per cookie is 40,000. So we collect our data just need 17 days.

**Experiment Analysis**

**Sanity Checks**

In the sanity check, we should check our invariant, namely number of cookies and number of click on “Start free trial”. In our check, equals 0.05

**number of cookies**: [0.4988, 0.5012] and our observation is 0.5006. So it’s pass our check.

**number of click on “Start free trial”:** [0.4959, 0.5041] and our observation is 0.5005. Pass.

**Result Analysis**

**Effect Size Tests**

**Gross Conversion**: [ -0.0291, -0.0120]. Statistical significant. Practical significant.

**Net Conversion**:[ -0.0116, 0.0019]. Both not significant.

**Sign Tests**

We assume our evaluation metrics follows binomial distribution.

**Gross Conversion:** We have 4/23 times, so p = 0.0026. That’s statistically significant.

**Net Conversion:** We have 10/23 times, so p = 0.6776. That’s not statistically significant.

**Summary**

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.

I don’t use the Bonferroni correction because if we use that we will need more time to collect our sample. There is a discrepancy in the gross conversion rate but we don't have it in net conversion rate. Because our goal is reducing the number of frustrated students who left the free trial because they didn't have enough time

**Recommendation**

So in the end, we can conclude that our test would reduce the click on “free start screenshot”, however it won’t show a significant increase in the net conversion. So we didn’t reach our goal, so we shouldn’t launch the experiment.

**Follow-Up Experiment**

In the future work, if we wanted to reduce the number of frustrated students who cancel early in the course, we can do things such as: 1. Have more specific instruction before they start the course. 2. Enhance the free trial courses, the same as improve the course’s quality. 3. Have stricter prerequisite to filter the students. We can use user-id as our unit of diversion. We could measure the same net conversion rate, use similar invariant like number of pagesviews, number of clicks on “free start trial.”

**Reference**

Discussion Forum

1V1 Mentor