A/B Test
Project 7: Data Analyst Nanodegree

Jorge Santamaría Cruzado

March 19, 2016

1 Experiment Design

1.1 Metric choice

Number of cookies INVARIANT This is the unit of diversion until the user logs in. It's a good invariant metric because not all user will be logged in when they load the page.

Number of user-ids I wont use this metric. It can be used as an evaluation metric though. The users that enroll on the free trial need to be logged in so it can be used to track them. However there are better choices for evaluation metrics below.

Number of clicks INVARIANT I will use this as an invariant metric to check the proportion of users of each group is the same. The experiment shouldn't affect this because changes occur after they click this button

Click-through-probability INVARIANT Same as above, it should be the same for both groups.

Gross conversion EVALUATION The number of logged in users that enroll divided by the number of cookies that clicked the button. This is the metric that's more likely to be affected by the experiment, probably the new pop up will "scare away" some users that don't have the required time to complete the courses.

Retention At the beginning I also choose Retention to be an evaluation metric. However it needed like 4M pageviews and the experiment would take a very long time, so I had to drop it.

Net conversion EVALUATION The number of users that make a payment divided by the number of users that clicked the button. I will also use this as an evaluation metric, it is logic to think that it gross conversion is lower, this will be lower too.

The initial hyposthesis consist in two parts:

1. It will reduce the number of students that left the free trial.

2. It won't significantly reduce the number of students that continue past the free trial and make a payment.

The first part of the hypotheses can be tracked by the Retention metric, while the second part can be tracked by the Net Conversion metric. However, Retention metric is not an option here for the reasons already shown. I will use Gross Conversion combined with Net Conversion, seeing a significant decrease on Gross Conversion while Net conversion keeps the same values will mean that the Retention has probably increased here.

1.2 Measuring Standard Deviation

I will calculate now the standard deviation of gross conversion and net conversion. They are binomial distributions so the formula will be $\sqrt{\frac{\hat{p}(1-\hat{p})}{N}}$

Gross conversion 0.0202

Net conversion 0.0156

For these two metrics, the unit of analysis is cookies while the unit of diversion is user-id. However, each user-id comes from a single cookie if the user was not logged in so we can use them as if they were the same, and the analytical estimate can be used here.

1.3 Sizing

I wont use the Bonferroni correction here. At the first iteration I needed like 4M pages using retention as an evaluation metric. This is way too much so I drooped it and now I only need 685325 pageviews.

All traffic will be diverted to the experiment so it will take only 18 days to complete.

The risk of the experiment is low. Maybe the popup will scare away some users, but it's not likely that they are too many, 5 hours per week is not that much so almost all people can afford to dedicate that amount of time to the courses.

2 Experiment Analysis

2.1 Sanity checks

For number of cookies, the observed value is 0.5006 and the confidence interval: [0.4988, 0.5012]

For number of clicks on Start free trial, the observed value is 0.5004 and the confidence interval is: [0.4959, 0.5041]

All sanity checks pass.

2.2 Result Analysis

2.2.1 Effect size

For gross conversion the confidence interval is [-0.0291, -0.012] this has both statistical significance and practical significance.

For net conversion the confidence interval is [-0.0116, 0.0019] this interval has no statistical or practical significance

2.2.2 Sign tests

For gross conversion the p-value is 0.0026, which have statistical significance.

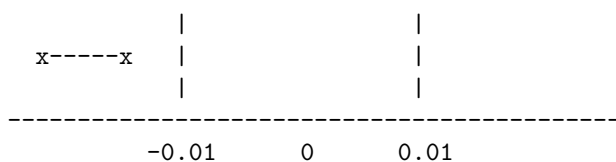For net conversion the p-value is 0.6776. It doesn't have statistical significance.

2.2.3 Summary

I haven't use Bonferroni correction because It looks too conservative. I want to avoid a type II error here, and be sure I reject the null if its false, even if I increase my chances of getting a type I error.
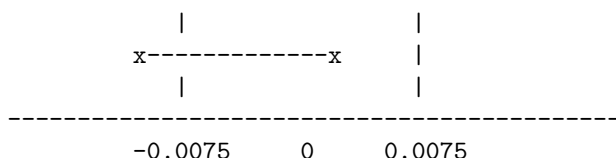
Looks like the only metric with statistical significance is gross conversion. That was expected because the popup showing probably scare some users who doesn't finish the enroll process.

2.3 Recommendation

For gross conversion, the expectations were that it will have a significant lower value. And that's what the results of the experiment shows, so the null can be rejected on the first part of the hypotheses.

```
                    |                   |
     x-----x        |                   |
                    |                   |
     ----------------------------------------------
              -0.01         0         0.01
```

For net conversion, the expectations were that it won't change, and this is what it looks like. However, it looks like it intersects with the practical significance boundary, so more data will be needed before a decision can be make.

```
                 |                   |
        x-------------x       |
                 |                   |
     ----------------------------------------------
            -0.0075      0       0.0075
```

The recomendation here is to gather more date because the net conversion results are not clear.

3 Follow-Up Experiment

To try to reduce early cancellations a simple experiment can be run. People always like to compete with each other, so it would be great to show some histogram with the time taken to complete the course by the rest of students so he can try to be as left on the histogram as he can.

Then also show a scoreboard with the fastest students.

The hypothesis is that this will increase student's engagement and more of them will finish the course, instead of procrastinating it away until cancellation.

I think this will keep the student working day by day on the course and it will reduce a lot of early cancellations (I think early cancellations happens most when you leave the course for a week, and then wait another day to come back, and so on, so this will probably reduce that, forcing students to put more time on the courses.)

The unit of diversion will be user-id here, so we can track each person. The evaluation metrics will be cancellations before finishing the course on each group, and it will be also interesting to use engagement measuring hours per week dedicated to the course. The invariant metric will be enrollments.