Hero X Outreach Experiment Power Analysis

# Purpose

Here I consider the necessary subject numbers to detect a significant difference in email click-through rates of Hero X outreach emails.

# Summary

We may need to extend the experiment over multiple email campaigns to have the numbers of subjects needed to detect treatment differences with confidence. Using historic data on click-throughs and a treatment effect that doubles the historic rate, 4000 subjects are needed to have an 80% detection rate. Previous campaigns typically reach 2714 inboxes, with 30% reaching more than 4000. I consider these estimates conservative, as the effect size is large and no covariates are considered.

# Statistical Framework

Assume a basic logit model is used to model the click-through rates in the experiment.

Assume the hypothesis test of interest is a difference in treatment and control.

Assume the level of the test is the 0.05 standard.

# Calibrating the Framework

To determine the number of subjects, the size of the treatment effect and error distribution need to be fixed. I assume the click-through rates of previous outreach emails are indicative of typical click-through rates.

Data on 10 previous *outreach* emails are in “Email Campaign Metrics for Harvard.xlsx”. I assume that the number of emails sent, net of bounced emails, are recipients. I consider the number of click-throughs to be positive responses.

|  |  |  |
| --- | --- | --- |
|  | Average | [Min, Max] |
| Click-through Probability | 0.0047 | [0.0011, 0.0102] |
| Click-through log-odds | -5.5707 | [-6.8363, -4.5756] |

Table 1: Summary Statistics of previous outreach click-throughs.



Figure 1: Logistic distribution fit to previous click-through rates.

Equation 1: Fitted distribution parameters.

Assume **doubling** of the average click-through rate in the treatment group and the average click-through rate in the control group.

|  |  |  |
| --- | --- | --- |
|  | Click-through Probability |  |
| Control Group | 0.0047 | -2.3215 |
| Treated Group | 0.0095 | -2.0184 |

Table 2: Assumed treatment effects.

# Detection Rate

Given the assumptions detailed above, the probability can be estimated by simulating datasets—essentially implementing a parametric bootstrap. Figure 2 shows the detection rate of doubling the historic click-through rate against the number of subjects. The detection rate is the proportion of simulated datasets of a given size for which the null hypothesis of no treatment effect is rejected with 95% confidence. Each dataset size is simulated 5000 times.

The analysis indicates that 4000 subjects (2000 per group) are needed to have an 80% chance of detecting a doubling of the click-through rate from historic levels.



Figure 2: Probability of detecting a doubling of historic click-through rates.

The previous outreach campaigns

|  |  |  |
| --- | --- | --- |
|  | Average | [Min, Max] |
| Emails Received | 2714 | [886, 4915] |

Table 3: Summary statistics of number of emails received in previous outreach campaigns.