MSAN 631: Design and Analysis of Experiments

Lab 3: Google Analytics
Due: Thursday June 22nd, 2017

Problem Description:
In practice the data associated with web experiments is collected through an experimentation platform developed either in-house or by a third party. Examples of third party experimentation platforms include Optimizely, Google Analytics, Mixpanel and Wasabi, to name a few. In this lab you will be carrying out a web experiment using Google Analytics.

In MSAN 695 you have become proficient in using Google Analytics to collect and consume data on website user behavior. In this lab you will learn how Google Analytics may also be used to experiment with different variations of your web application. In self-selected groups you will be required to do the following:

- Choose a response variable that characterizes user behavior that you would like to optimize (i.e., average time on site, click through rate, bounce rate etc.).
- Choose one factor that you believe influences this response and identify 3 or more levels for that factor to experiment with.
- Develop a different version (variant) of your web application corresponding to each level of this factor. Note that the URL for each variant should be of the form www.example.com/v1, www.example.com/v2, etc.
- Create a multi-armed bandit experiment in Google Analytics that tracks your response variable across each variant of your web application. For instructions on building an experiment in Google Analytics see: https://support.google.com/analytics/answer/1745216?hl=en
- Run this experiment until the multi-armed bandit algorithm has declared a winner, or until the due date of the lab (whichever comes first).

Deliverable:
You will submit a report that includes the following:

- A description of your application and the purpose of the experiment. Be sure to indicate your response variable, factor, and factor levels.
- A description of how multi-armed bandit experiments are carried out and a discussion of the advantages and disadvantages of this approach relative to standard hypothesis testing (i.e., ANOVA F-tests and χ^2-tests). In this discussion be sure to use your sample size calculators from Lab 2 to determine what sample size would have been necessary in the classical hypothesis testing approach.

- A statement of which variant was deemed optimal by the end of the experiment. Please discuss any potential limitations or reservations you believe may be associated with this conclusion.

Grading:
The members of each group will receive a common grade out of 20. Groups will be graded on the thoughtfulness of the experiment and the reports will be graded on the technical correctness and overall quality of the written exposition.