Estimating incrementality over time

Notation

- c : Cohort of shoppers; $c = 0$ for "control" cohort that do NOT experience Fenix, and $c = 1$ for the "treatment" cohort that experience Fenix products.
- G : Set of mutually exclusive and collectively exhaustive subgroups of shoppers (e.g., $G = \{$ Male in CA, Female in CA, Everyone else $\}$ would be a valid set of three subgroups)
- $\theta_{c,g}$: Conversion rate, defined as the proportion of unique sessions in cohort c and group g that result in an order being placed. The subscript for g may be omitted to mean the entire cohort c (e.g., θ_1 represents the conversion rate for all sessions in the treatment group and $\theta_{1,g}$ represents the conversion rate for sessions in treatment group that meet the criteria for group g)
- $\delta_g = \theta_{1,g} - \theta_{0,g}$: The causal effect of Fenix product on the conversion rate of group g

The subscript g may be omitted to mean the entire cohort (e.g., θ_1 represents the conversion rate for all sessions in the treatment group, and δ represents the average treatment effect over the entire population.)

Problem statement

Given at least one prior experimental measure of δ, and a new set of non-experimental observations under treatment ($c = 1$), we wish to extrapolate estimates of δ to the new observations.

Setup

- We have some prior experimental measure of $\widehat{\delta_g}$ for every $g \in G$.[1]
- Once Fenix is "live", without an explicit holdout, for some specific period, we observe
  - n_g : Total number of unique sessions in some group g
  - o_g : Total number of orders from sessions in group g
  - $\theta_{1,g} = \frac{o_g}{n_g}$ (i.e., every session is considered "treatment")
  - $N = \sum_{g \in G} n_g$: Total number of unique sessions

[1] Note, this requires that, for each unique session, we have group identifiers (e.g., "male/female", location info regarding the session) *prior* to the order.

Proposal (with concrete grouping data at session-level)

- We wish to know the counterfactual conversion rate:

$$\theta_0 = \sum_{g \in G} \left(\theta_{1,g} - \widehat{\delta_g}\right) \frac{n_g}{N}$$

- Assume the treatment effect on any given <u>*group*</u> does NOT change over time (e.g., the conversion rates for "Men in California" might change, but the *difference* of what conversion rate might have been with/without Fenix is always the same for that group)
- **What <u>*does*</u> change is the conversion rate itself and composition of shoppers (e.g., more "Men in California" are visiting, and the conversion rate for that group is dropping)**
- Then, given all the information we already have and observe, we can readily compute θ_0
- The main challenge is in collecting the data necessary for assigning sessions to a group *prior* or orders, both for the experiment and post-experiment.

[WIP] Proposal (WITHOUT concrete grouping data at session-level)

- Now assume that we *cannot* identify the group of a session prior to an order being placed
- This will have implications on how we measure $\widehat{\delta_g}$ to begin with; but let's start by assuming we have those available from a previous experiment, as before.
- Then, we no longer observe n_g, and only observe o_g
- As a result, we also do not observe the subgroup conversion rate $\theta_{1,g} = \frac{o_g}{n_g}$

Rewriting the expression for θ_0, with substitution, we can arrive at

$$\theta_0 = \sum_{g \in G} \left(\frac{o_g - n_g\hat{\delta}_g}{n_g}\right)\frac{n_g}{N}$$

$$\theta_0 = N^{-1}\sum_{g \in G} \left(o_g - n_g\hat{\delta}_g\right)$$

$$\theta_0 = N^{-1}\left(o - \sum_{g \in G} n_g\hat{\delta}_g\right)$$

In other words, the problem is reduced to estimating n_g: the number of sessions that should be assigned to each group.

Notes regarding implications on initial measurement

If we don't have session-level grouping data, we wouldn't be able to identify group-level conversion rate because we can't know the total number of sessions for each group. What we *can* know is the total number of *orders* for each group, from which we will have to extrapolate the proportion of each group in sessions in such a way that control and treatment agree. [TBD].