Date: Tuesday, April 9, 2019

Plan for today

- Finish up multivariate experimentation
- Tie everything to business + examples + interviews

Note: When testing for interactions, you test the full model.

Example: If you have three factors, each w/ 2 levels, we will have cross-product terms like:

$$\beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \beta_4 x_{1i} x_{2i} x_{3i} + \beta_5 x_{1i} x_{2i} + \beta_6 x_{1i} x_{3i} + \beta_7 x_{2i} x_{3i} + \varepsilon_i$$

* If interaction terms are not significant (i.e. we fail to reject F-test of H_0: interaction coeffs = 0), then we re-run the regression w/ main effects only.

Interview question: We tested the main effects of ad frequency and type of session duration and estimated the model:

$$y_i = \hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \hat{\beta}_2 x_{i2} + \hat{\beta}_3 x_{i3} + \hat{\beta}_4 x_{i4}$$

x_{i1} = 7:1, x_{i2} = 4:1, x_{i3} = 1:1, x_{i4} = Video
Base level = {photo, none}

↳ What is the expected metric on {4:1, video}

Answer: $\hat{\beta}_0 + \hat{\beta}_2 + \hat{\beta}_4$

* We can next assess the significance of main effects over a factor by F-testing

  H_0: coeffs of factor = 0

  From this example: $H_0: \beta_1 = \beta_2 = \beta_3 = 0$

  test whether there is a main effect over frequency.

* Since we controlled the allocation of experimental units into conditions (ie through the regression coefficients quantify casual effects! This is not true for observational studies)

* This casual inference is the main advantage of experimentation over observational studies.

Interview question: Why run an A/B test?

Answer: we can assess the casual effect of changing a design factor on a metric that is important to the company!

Note: Running full model in R

```
lm(Time ~ Frequency * Type)
```

`*` → Run interaction effects

Frequency, Type: Factor variables w/ base levels "none" & "photo" respectively.

```
anova(reg) ⇒ output
```

The partial F-test for interaction and main effects across all factors.

* After running f-tests, we know which factors are important with respect to our response. All significant factors should be considered in future experiments for optimization.

* Optimization can then be done using pairwise t-tests or multi-armed bandits.

* Pairwise t-tests → compare metric (mean or proportion) across all pairs of conditions.

  ↳ use multiple comparisons.