Ilana’s work plan, summer 2019

Goal: have proposal ready by December

June 3-14: Ilana in Scotland

When back:

* Set up weekly meetings with Deb and Brandy
* Make a progress tracking chart if you want

Tasks for summer – split time between these three things so you don’t get burnt out

1. Writing
2. Math proofs
3. Computing (Simulations)

Writing tasks – write every day!

* Literature review
  + Look into pattern mixture model, need one sentence describing why it’s different from what we are doing
* Pick journals for three papers – can use paper from target journal to base our paper formatting off of
  + Unreasonable journals: anals of statistics, biometrics, biometrika, journal of the ASA, journal of the royal statistical society series B
  + Reasonable journals: statistics in medicine, jabes, jcgs
* Start on proposal:
  + Put blank tables in results. For discussion, say what questions you will answer.
  + Simulation paper will look a lot like the Johnson paper published in Statistics and Medicine
    - Will have same first four subsections of “simulation” section, can use the same metrics
  + For quasilex paper:
    - Title
    - Authors
    - Abstract
    - Introduction
      * Motivating example
      * Literature review
      * Why we care
      * What we are not doing
      * Sections of paper
    - Methods
      * Notation and Definitions
        + Missingness process
        + Deletion classes
      * Preliminary results
        + Estimation in deletion classes
        + Weighted estimates

Unbiased

Consistent

Distribution theory

Test statistic

Distribution under the null

Simulations to check everything (proposed)

Example (proposed) – Brandy will help come up with

Computing tasks

* Brandy will show Ilana where to store files and how to document
* Make data generator that can generate correlated outcomes for different sample sizes, percentages of missingness, correlations between outcomes (sigma), and betas.
  + Generate Gaussian data using missingness mechanism developed by Qaqish
    - Brandy has code for simple random missingness – use this to get started. Be aware of freeing matrices to free up memory.
    - Kevin Josey paper describes mechanism, as well as Qaqish paper in dropbox folder
  + Start with full Y, sick the gremlin on it, then generate X (or else may run into less than full rank X)
  + Continuous Y
* Pick betas for testing under the alternative
  + Search algorithm
  + Pick so complete data has power of 0.2, 0.5, 0.8
* Run GEE and mixed models on simulated data
  + Tabulate Type I error, convergence, biasedness of estimates, and power

Math proofy tasks

* Finish Bonferroni correction proof
* Think about unifying notation as you go – you can introduce new operators (like Johnson did) to unify things

Things to keep in mind:

* We can talk to Katerina or Harry Smith about CPG data
* F31 grant – due December 8
  + <https://grants.nih.gov/grants/guide/pa-files/pa-18-671.html>
  + <https://grants.nih.gov/grants/how-to-apply-application-guide/due-dates-and-submission-policies/due-dates.htm>
* Ilana has funding from BIOS department to use on a conference
  + Talk to Ivana Yang about an epigenetics conference
  + Or, teaching conferences (SACNAS)
* Throughout summer: use Maya as needed. Could do team programming or writing with her. Could put her in charge of running code to generate data.