

Meeting Notes

LATEX 2 ε

March 2, 2020

Contents

1		2
2		5
2.1	Overview	5
2.2	Bayesian MR	5

Minutes 1

Status Update and Next Steps Discussion

Those present Daniel Cizin, David Knowles, Stephen Malina

Date January 29, 2020

We discussed a few topics:

1. Getting standard error estimates from NN predictions.
2. Inverse variance weighted methods for conducting meta-analyses.
3. How to do in-silico mutation such that we don't condition on the exposure and also get effects we can combine.

Summarizing our decisions in order of dependency rather than chronologically:

1. At least for now, we're going to try and get results from the "multiple DAG instantiations with mutations as IVs" framework.
2. High-level our strategy will be:
 - randomly sample sequences;
 - do saturation mutagenesis for each, getting a standard error for each mutation's prediction;
 - run Egger (or similar, e.g; RAPS) on the result for each sequence, treating the difference between the binding probability for the mutated and wild type sequences as an effect size; and
 - combine the causal estimates for each sequence together using a meta-analysis protocol.
3. For getting standard errors for individual predictions, we'll use Yarin Gal's method of repeated dropout.

4. We hope that using a robust regression method will allow us to get an estimate of how much violation of exclusion-restriction there is in our model.

Given this, our follow-ups are:

Task: Verify that with classification, we don't care about the noise variables in the prediction variance formula. (Stephen)

Task: Figure out and implement the saturation mutagenesis with uncertainty estimates logic on top of Kipoi. (Stephen)

Task: Do some research into how likely it seems that sequence will influence accessibility directly. (Daniel)

Task: Try and understand the difference between Egger regression and RAPS. (Daniel)

Task: Keep thinking about the idea of using cell type, potentially in a deep IV framework, as an alternative IV.

Task: Keep thinking about whether there's something interesting in investigating the covariance between predictions from saturation mutagenesis (see figure 2.2.2).

On the next page, I put the two pictures I took of the whiteboard during our meeting.

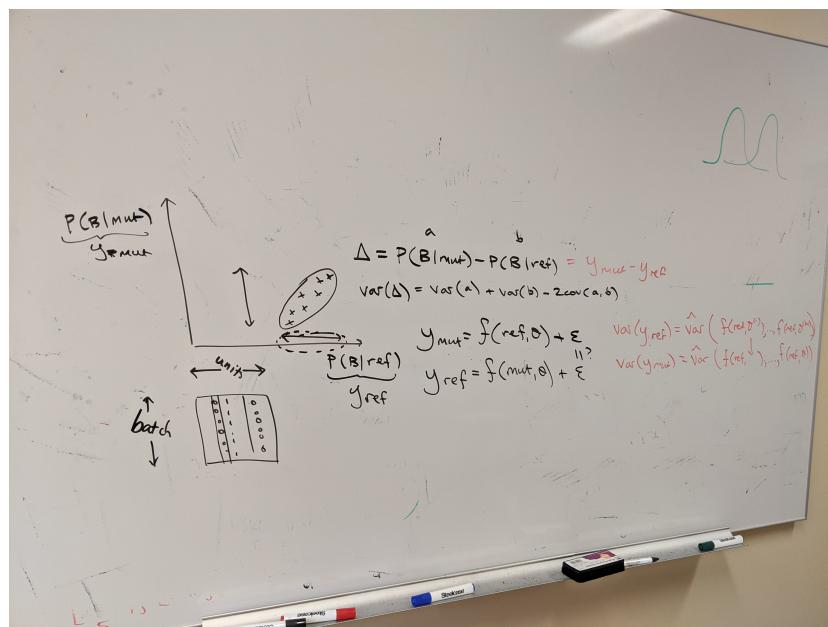


Figure 1.1: Shows two things: that we should take the std error of the diff between mut and wild type probabilities and that we want to use the same dropout mask for mut and wild type predictions so that we can estimate the covariance between the two.

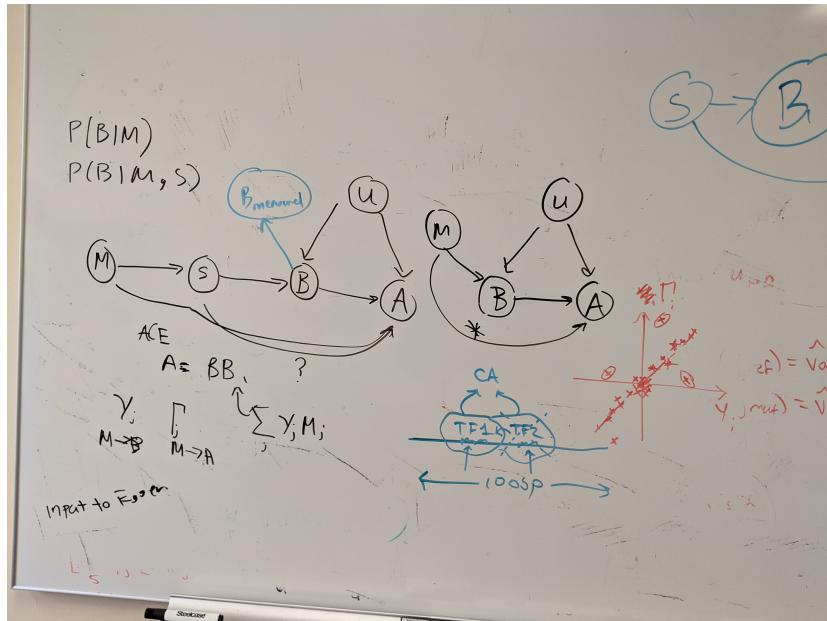


Figure 1.2: Captures two notable things. First, that we mostly came to the conclusion that because of potential exclusion restriction violations, it makes to analyze the causal effect separately for each sequence and then aggregate. Second, that there may be something interesting in the covariance between different mutation predictions for the same sequence.

Minutes 2

Bayesian Multi-modal MR discussion

Those present David Knowles, Stephen Malina, Brielin Brown

Date February 26, 2020

List of topics

2.1	Overview	5
2.2	Bayesian MR	5
2.2.1	Bi-directional MR issue	5
2.2.2	Mixture model	6

2.1 Overview

We discussed a few topics:

1. Mimicking the zero modal pleiotropy MR method using a mixture of Gaussians model
2. Next steps for Deep MR

I've included my pictures of the whiteboard from our discussion at the end of these notes.

2.2 Bayesian MR

2.2.1 Bi-directional MR issue

Brielin is interested in the ‘Bayesian’ approach to bi-directional MR because he’s discovered an issue with using significance as a proxy for picking instruments in

bi-directional MR. Selecting instruments based on significance leads to problems when we have different sample sizes for different pairings of the instrument and phenotypes we're looking at. E.g., say I is an instrument for A which actually causes B , but I only have a few samples of i, a and many samples of i, b . As a result, it's possible that I'll guess that I is an instrument for B . If I do this, then my Wald estimator for $B \rightarrow A$, $\hat{\beta}_{B \rightarrow A}$ will be $\frac{1}{\beta}$ where β is the true causal effect of A on B .

2.2.2 Mixture model

(This describes the model depicted in 2.2.2.) We want to parameterize a (eventually, mixture of) Gaussian by a radius (ρ) and angle parameter (θ) to approximate the non-Bayesian zero modal pleiotropy model. David described a hierarchical model as an approach for doing so. The model assumes that we have computed a set of effect sizes, $(\hat{\beta}_{iA}, \hat{\beta}_{iB}), \dots$ and standard errors $\hat{\sigma}_{iA}, \hat{\sigma}_{iB}$ where A, B correspond to the $Z \rightarrow X$ and $Z \rightarrow Y$ relationships respectively.

To describe the model, we start with the assumption that effect sizes for each variant (indexed by i) are jointly normally distributed,

$$\begin{bmatrix} \hat{\beta}_{iA} \\ \hat{\beta}_{iB} \end{bmatrix} | \theta, \rho \sim \text{Normal} \left(\rho \begin{bmatrix} \cos \theta \\ \sin \theta \end{bmatrix}, \text{diag} \left([\sigma_{iA} \quad \sigma_{iB}]^\top \right) \right).$$

Since ρ is a nuisance parameter in this model, we can integrate it out. Letting

$$\mathbf{t} = \begin{bmatrix} \cos \theta_i \\ \sin \theta_i \end{bmatrix},$$

we have

$$\begin{bmatrix} \hat{\beta}_{iA} \\ \hat{\beta}_{iB} \end{bmatrix} | \theta \sim \text{Normal} \left(0, \mathbf{t} \mathbf{t}^\top + \text{diag} \left([\sigma_{iA} \quad \sigma_{iB}]^\top \right) \right).$$

Next we need to put priors on our parameters. For our angles, $\theta_1, \dots, \theta_n$, we have

$$\theta \sim \text{VM}(\mu, \kappa)$$

where VM denotes the Von Mises distribution over the circle.

Then, our inferred distribution over θ s should reflect SNP-specific causal effects and our distribution over μ should reflect the 'overall' causal effect.

This covers the basics of the unimodal version of the model. Obviously, there's still a bunch of details on how to fit this thing, but I'll leave those for when I'm actually doing the work.

The other thing I wanted to note is what the c_i terms in figure 2.2.2 represent. The novel use-case for this model relates to Brielin's stuff about bi-directional MR I discussed above. Solving the bi-directional MR problem requires dealing with a bi-modal distribution of effect sizes, so the c_i s and the discrete distribution over them reflect that.

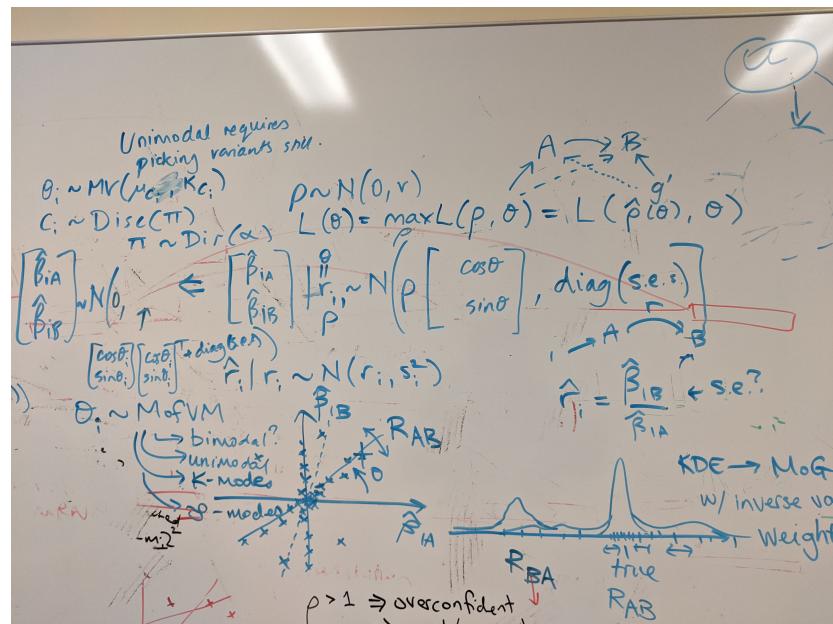


Figure 2.1: Mixture of Gaussians model for unimodal MR.

