Introduction

The emergence and transmission (COVID-19) is a pressing, current global-health epidemic, that has rapidly impacted the entire world. The first case of COVID-19 in the United States was found on January 20, 2020 in Washington State. Since then, the outbreak has expanded to all 50 states. Currently, the U.S. has the highest number of coronavirus related cases and deaths globally, with over 500,000 cases and 20,000 deaths in less than 4 months.

As a response to the rapidly-intensifying nature of this pandemic, the majority of states in the US have adopted social distancing guidelines, restricting movements of residents, closing local businesses, and enacting personal protective equipment guidelines in an attempt to control community transmission of the virus and “flatten the curve”. As a result of social distancing, mobility of individuals has been heavily restricted.

We are interested in identifying predictors beyond social distancing guidelines that contribute to the reduction in the “Twitter Social Mobility Index” (SMI), defined by researchers at Johns Hopkins University and George Washington University. Identification of potentially significant predictors could be important which states can have better COVID-19 outcomes based on resident mobility, and potential steps that can be taken to help social distancing until we “flatten the curve”.

Methods

The SMI is a measure of social mobility gathered from geotagged.

The predictors we are considering are:

For our Bayesian modeling framework, we will be fitting a Bayesian linear with a Gaussian family multivariate regression model to assess different factors that can contribute to emergence and mortality to see how they differ across states. A multivariate model is needed to consider several predictors that could potentially contribute to the case occurrence and mortality rate as a result of novel coronavirus. We want to use the results of this data to try and make certain policy recommendations in order to “flatten the curve” and try and decrease the impact of COVID-19 on the US.

Our goal was to perform bayesian model selection on the full set of predictors plus some interaction effects, however, we ran into computational difficulties using the full set of predictors. Therefore we first performed model selection using AIC and then selected the top nineteen most significant predictors by their p-values. We used the Bayesian Model Averaging Library to perform model selection. We used a uniform model prior since we were not biased to one model over another and used Zellner’s g-priors as the prior distribution for our coefficients for the various linear models as g-priors benefit from computational efficiency and are scale invariant. We then used the chosen predictors to fit a Gaussian family multivariate regression model. We used a MCMC algorithm to sample the posterior distributions for the coefficients and we checked the trace plots to make sure our markov chains converged successfully before performing posterior analysis.

Results

Conclusions