

Untitled

Graham Casey Gibson

8/20/2020

Reference: https://bbolker.github.io/mixedmodels-misc/notes/corr_braindump.html

Suppose we evaluate two models, a candidate model and a baseline model, across a series of time points, targets, and geographies. Specifically, suppose we are interested in evaluating the question of whether MechBayes (MB) is better than the COVID-Hub baseline model (BL) across 13 weeks and 51 geographies (50 states plus DC) and 4 targets (1-4 week ahead) according to MAE. The data-frame we are handed is a

Note that the size of the dataframe is less than $51412 * 2 = 4896$ due to right truncation of the 2-4 week ahead targets for weeks 9+.

This produces a boxplot of scores by time-zero as follows

In order to answer the question of whether there is a stisticaly significant difference between models in terms of MAE, we model $\log(\text{MAE})$ as follows

$$MAE_r = \beta_0 + \beta_1 * \text{model} + \beta_2 * 2 \text{ week ahead} + \beta_3 * 3 \text{ week ahead} + \beta_4 * 4 \text{ week ahead} \quad (1)$$

$$+ \beta_5 * 2 \text{ week ahead} * \text{model} + \beta_6 * 3 \text{ week ahead} * \text{model} \quad (2)$$

$$+ \beta_7 * 4 \text{ week ahead} * \text{model} + \Sigma \quad (3)$$

where