How do we calculate the metrics from the models?

Zane

2024-08-29

After we fit the models to the full data and to each of the subsampled datasets, we need to calculate the performance metrics for each of the models. Notably, we want to calculate the new metrics (AUC, p40, intercept) with and without a censoring correction in the fitted linear model, and we then want to figure out what the analogous metrics are for the "old" / current way of doing things.

- The magnitude is somewhat obvious. We want to calculate either the homologous GMT or titer increase, using only the homologous strain. So, we should get the intercept from an intercept-only linear model either with or without a censoring correction.
- The breadth is calculated as the average seroconversion rate. That is, we calculate the seroconversion rate for each individual and then average them. The seroconversion rate is not affected by the censoring correction. Any individual whose titer is recorded as 5 and is indicated as having seroconverted would have seroconverted even if the titer was actually lower. We can calculate the seroconversion rate from the real data instead of having to fit a separate model because we don't need the censoring correction.
- For the total strength of the response, we calculate the GMT across all strains and individuals. This is equivalent to the intercept of a linear model which uses all of the strains in an intercept-only model, so we can correct for censoring by including a censoring correction in the likelihood of that model.

As a justification for ignoring the censoring when we calculate the seroconversion, consider three individual who all have pre-titers below LoD and respective post-titers of 20, 40, and 80. If we record a pre-titer below the LoD as 5, we would respectively compute their seroconversion status as no (20/5 = 4 fold rise but they are not seroprotected), yes (40/5 = 8 fold rise and the posttiter is high enough), yes (80/5 = 16 fold rise and the posttiter is sufficient). However, we know that person's pre-titer is censored and can range from anywhere in the interval (0, 10). If that person's pretiter is infinitesimally close to 10, we would still calculate their seroconversion status as yes (40/10 = 4 fold rise and the posttiter is high enough). So we can ignore the censoring aspect when we calculate the fold change.

Calculating metrics from model fits

Once we fit the models to all of the subsampled datasets (simulated studies), we need to obtain the metrics.

New / Our metrics

To get the **new** set of metrics, we fit a multilevel model with varying effects for the intercept and slope to the data where the outcome is either log post-vaccination titer or titer increase, and the only predictor is the antigenic distance. When we run HMC via Stan, we then get k number of samples of each model parameter. In order to focus on the global mean, ignoring individual variability, we only consider the samples of the marginal model parameters, i.e. the population-level α and β intercept and slope. Since we have k samples, and our new metrics are functions of the slope and intercept only, we therefore get k samples of each metric for each subset.

Next, we have to condense the distribution of metric values for each subset into some summary statistic over the subsamples. Let $m_{s,k}$ be the kth sample of metric $m = g(\alpha, \beta)$