Homework IV (2022)

1. Consider the following two random effects models for discrete data.

   (a) Consider the Poisson random effects model with a random intercept,

   $$\begin{aligned} \log E(Y_{it}|b_i, x_i) &= \beta_0 + \beta_1 x_i + b_i, \\ b &\sim N(0, \tau^2). \end{aligned}$$

   Derive the population averaged effect of a one unit change in $x_i$. How does this compare to the subject-specific effect, $\beta_1$?

   (b) Consider the logisitc random effects model with a random intercept,

   $$\begin{aligned} \text{logit } P(Y_{it} = 1|b_i, x_i) &= \beta_0 + \beta_1 x_i + b_i, \\ b &\sim N(0, \tau^2). \end{aligned}$$

   Derive an approximation to the population averaged effect of a one unit change in $x_i$. (Hint: Recall the cdf of a standard logistic distribution,

   $$F(x) = \left[1 + \exp(-\pi x/\sqrt{3})\right]^{-1}.$$

   Approximate this with

   $$\hat{F}(x) = \Phi(16x/15),$$

   where $\Phi$ is the cdf of a standard normal). How does this compare to the subject-specific effect, $\beta_1$? How does this comparison depend on $\tau^2$?

2. Fit a GEE to the simulated data from CTQ II (data in ctqII.data, information in ctqII.info.txt). Assume a separate mean for each week and a constant treatment effect. Is there a significant effect of treatment in the final week? What structure (if any) did you choose for the working correlation matrix? How do the inferences on $\beta$ (the regression coefficients) vary with other choices? Are the standard errors correct if the working correlation matrix is mis-specified? Explain.