

STAT/BIOST 571: Homework 5

To be handed in on Friday, February 22. Where solutions require use of R, summarize your findings in a written answer. For each question, write up your solution on your own, using **full sentences**.

IMPORTANT NOTE: Please include your annotated code in an appendix, to show what you did.

Problem 1: Sandwich and bootstrap standard error estimates (10 points)

As on slide 2.76, fit the model

$$EY_{ij} = \beta_0 + \beta_1(\text{Age}_{ij} - 8) + \beta_2\text{Gender}_i + \beta_3(\text{Age}_{ij} - 8) \times \text{Gender}_i$$

to the dental data by using REML, but use a homoscedastic covariance models with no correlation.

- (a) Calculate sandwich-based standard error estimates for $\hat{\beta}_3$ that account for clustering by subject. Write your own code for this, using matrix algebra.
- (b) Calculate bootstrap standard error estimates for $\hat{\beta}_3$ by resampling clusters. Describe the results of some basic diagnostics you can do to provide confidence that bootstrap intervals are valid for this dataset and that you have simulated a sufficient number of draws to accurately approximate true bootstrap intervals?
- (c) Calculate bootstrap standard error estimates for $\hat{\beta}_3$ based on resampling observations without regard to cluster and resampling both clusters and observations within clusters.
- (d) Discuss any differences between your sandwich standard error estimates and the three versions of bootstrap standard errors.