Dr. Özlem İLK DAĞ

Fall 2023-2024

**STAT 545- LONGITUDINAL DATA ANALYSIS**

**HOMEWORK 2**

**Due 9 January 2024, Tuesday, 15:30**

**You should work on your homework on your own. Please feel free to get help from me, but not from anyone else. Let me know if my wording in the questions is not clear.**

**Your answers should be ≤ 10 pages under a standard format (normal margins, Times new roman font=12 …)**

**I will post the answers to this homework on the same day. Therefore, no late homeworks will be allowed.**

**1.** For each of the following parts, state which covariance structure is the most appropriate.

**a)**  where α and λ are two unknown parameters, tj and tl are two follow-up points.

**b)** 

**c)** An estimated correlation matrix from data:



2. Please use the demo dataset called nepali in faraway package in R. Use ?nepali in R to get information about the dataset.

**a)** Investigate the missing data pattern in nepali. Report the most important findings.

**b)** Impute the missing cases with a method in the mice package in R. Report which method you used; your codes; and the first 10 rows of the imputed data. If you are imputing more than one set, you can randomly pick one of them.

**c)** Fit a marginal model to the imputed data (use any one of the imputed sets) to analyze the relation between weight and some possible covariates (but do not use the height). Which covariance structure did you use and why? Write down the fitted model and interpret the coefficient estimates in a few sentences.

**d)** Fit a transition model to the imputed data (use any one of the imputed sets) to analyze the relation between weight and some possible covariates (but do not use the height). Write down the fitted model and interpret the coefficient estimates in a few sentences.

**e)** Fit a random effects model to the imputed data (use any one of the imputed sets) to analyze the relation between weight and some possible covariates (but do not use the height). Write down the fitted model and interpret the coefficient estimates in a few sentences.

**f)** Try a machine learning method on this dataset that takes weight as a response (do not use height as a predictor) and report your codes, oputputs and findings. This method could be any ML method we covered in class, or something new that you found in literature/web. Just be sure it is an appropriate method in terms of type of response (eg it accounts for longitudinal aspect and it is appropriate for continous?, normally distributed? response).