**Homework 5 Due on Monday by 1:10 pm, 4/08, 2019**

The data are from a longitudinal study to identify salient parent and adolescent psychosocial factors related to emotional distress in adolescents, assessed using the Global Severity Index (GSI). A total of 409 adolescents whose parents had HIV were included, and their GSI measurements were obtained over 6 years (mean = 63.5 months). (Reference of the study**:** B. Bursch et. al.(2008). Psychosocial Predictors of Somatic Symptoms in Adolescents of Parents with HIV: a Six-Year Longitudinal Study. *AIDS Care*, 20, 667-676.)

The data “**hw5p1\_data**” is uploaded on blackboard. The data has been cleaned so all the missing values have been removed. The following variables are included in the data.

* ID: adolescent’s ID number
* Parent\_ID: parent’s id number
* GSI: Global severity index (outcome, continuous)
* True\_month: actual time since baseline (months)
* Gender: adolescent’s gender
* Hispanic: whether adolescent is Hispanic (not Hispanic vs. hispanic)
* Treatment: whether family was in the treatment =1 or control =0 group
* Season: 3 categories separated into months (spring=3-6, summer=7-10 ,winter=11-2)
* Parent\_base\_age: parent’s baseline age
* Parent\_died: whether parent died during the study(Yes/No)
* Parent\_gender: adolescent’s parent’s gender (male/female)
* Parent\_drug\_status: Parental hard drug use during the last 3 months:
* nonuser: parent never used hard drugs during the study
* nonusing-user: parent used hard drugs during the study, but not in the last 3 months
* using-user: parent used hard drugs during the last 3 months
* Parent\_alcohol: whether parent drank alcohol in the last 3 month (no=0, yes=1)
* Parent\_marijuana: whether parent smoked marijuana in the last 3 months (no=0,yes=1)
* Parent\_diagnosis: 3 level parent self-reported baseline illness level (asymptomatic, symptomatic, diagnosed\_with\_AIDS)

Perform a logarithm transformation with a base of 2 on the outcome variable GSI as follows:

**data** one; set hw5p1\_data; logGSI=log2(GSI + **1**/**53**); **run**;

The following variables are **time-dependent**: GSI, True\_month, Season, Parent\_died, Parent\_drug\_status, Parent\_alcohol, and Parent\_marijuana.

**Questions:**

1. Use the model building strategies to build a model for these data. Due to the presence of multiple time-dependent variables, assess the random slopes for significant variables only. If the test for random slope is not significant then it is necessary to keep the slope fixed. When comparing two models, if the likelihood ratio test is not significant then choose the simpler model.
2. From the final model, use the type 3 analysis to find significant effects and interpret them.