Homework 3

(hard copy due Wednesday 6/11/19 by 4 PM in my mailbox in CH 317)

Please type your responses and include the relevant printouts with each answer. Answers should be about one paragraph for each question, and they should be written as if they were part of a research report or journal article. Include the relevant statistical values in the text and include the relevant portions of your output. **Your answers should be in your own words.**

For all of the problems below, allow as many iterations as necessary for convergence. For problems using HLM, report the robust standard error estimates and significance tests.

The data for the first few problems are from the same health study of middle aged and older adults, but there is a new version of the data set, `hrsgrowth3.sav`. In addition to the weight to height ratio, `bmi`, there are two variables based on a measure of self-reported physical activity, one variable is the response to "how often do you engage in vigorous physical activity" rated on a 1 "never" to 5 "everyday" scale measured at each time point (`vigact`) and the other is a person-level version, based on the average of the three measurements (`actave`). There are four time variables, `t`, `t2`, `ct`, and `ct2`. `t` is the uncentered time variable (0, 1, 2) and `ct` is the centered time variable (-1, 0, 1). `t2` and `ct2` are the squared versions of the corresponding time variables. Download the SPSS and/or the HLM data file from the data page: http://web.pdx.edu/~newsomj/data.htm.

1. In HW 2, you should have found a significant change in BMI over time (`t`) and that that change varied across individuals. Use SPSS, R, or HLM to investigate whether the change in BMI is dependent on the average level of vigorous activity (`actave`). Grand-mean center the activity variable. Allow the slopes for the time variable to vary. Report and interpret your results in terms of the research problem.

2. In SPSS, R, or HLM, test a new model to investigate whether there is still a significant change in BMI (using uncentered time `t`) after controlling for vigorous activity measured at each time point (`vigact`). Allow slopes for the time variable but not the activity variable to vary across individuals. Report and interpret your results in terms of the research problem. Add a brief statement at the end that distinguishes the analytic goals of the activity variables as used in this problem and the previous problem.

3. In SPSS, R, or HLM, test for whether there is quadratic change in BMI using the uncentered time variables (`t` and `t2`). Allow the linear slopes to vary but not the quadratic slopes. Obtain a plot with a quadratic or loess-type prediction line. Report and interpret your results in terms of the research problem.

4. In SPSS, R, or HLM, test for whether there is quadratic change in BMI using the centered time variables (`ct` and `ct2`). Allow the linear slopes to vary but not the quadratic slopes. Report and interpret your results in terms of the research problem. Add a statement that compares the results from this model to those of the prior model with uncentered time variables, being sure to explain the difference in interpretation of the linear effect in the two models.

Data for the following problems come from the same high school student neighborhood survey used in the class example. The outcome for the following models will be a binary alcohol use variable (`alc`), whether the student reported using alcohol over the last year or not (0=no use, 1=have used). There are three other variables in the data set: the student's perceived neighborhood support (`nhsup`), an average of several items asking about how much support the student perceives to be

available in his/her neighborhood (values between 0 and 3); the student's gender (`gender; 0`=male, 1=female), and dropout rate for the school (`dropout`), which is a level-2 variable.

Note: For these models, allow additional iterations if necessary. Use adaptive quadrature results if using HLM, using 7 quadrature points and 150 maximum number of iterations. Report the PQL variance components but fixed effects from adaptive quad. If using R, request `family = binomial` and use the Laplace estimates from the `glmer` function in the `lme4` package. Because SPSS has only PQL estimation for fixed effects, I would like you to use one of the other program for the following problems.

5. Using HLM or R, estimate a multilevel logistic model with alcohol use (`alc`) as the (binary) outcome and perceived support (`nhsup`) as the predictor. Use grand-mean centering for the support variable and allow its slope to be random. Interpret the fixed effects, including the odds ratio, and the random effects.

6. Using HLM or R, test a second multilevel logistic model with alc as the outcome using the following group-mean centered predictors: `nhsup`, `gender`, and `dropout` as predictors. Use grand-mean centering for all predictors. Allow the slope for `nhsup` to vary across groups but not `gender`. Do not include any cross-level interactions. Interpret the fixed effects, including the odds ratio, and the random effects.