

Brief solutions for Biostat 209 Repeated Measures Homework, Part 1

Q1. Needs to exhibit clustering and predictors specific to various levels of the dataset.

Q2. mixed gluc exer || pptid: or can use xtgee as well.

Gives a somewhat less stat significant result. The standard error is smaller, but so is the coefficient (by a fair amount). A likely explanation is that there is confounding in the cross-sectional (baseline) analysis that is partially being removed in the longitudinal analysis.

c) The simplest model that could be run would be

```
mixed gluc exer age bmi avgdrpwk || pptid:
```

but the three continuous predictors need to be checked for linearity.

d) There are lots of extreme outliers that do affect the results somewhat. Should remove them and look at the effect on the analysis. Note: xtgee won't generate residuals directly, but can generate predicted values, from which you can get residuals:

```
predict glucpreds  
gen glucresid=glucose - glucpreds  
scatter glucresid glucpreds
```

Q3. _cons, 97.37, is estimated mean glucose level for placebo group at visit 0. Visit, 1.225, is increase per visit for placebo group. 1.group is the difference between the groups at visit 0 (baseline). Group#c.visit, -0.811, is the difference in the increase per visit between treated and placebo groups. Yes, stat significant. No evidence for need for other than linear.