Assignment 3

Plots of Repeated Measures Longitudinal Data

This homework is intended to develop skills in representing repeated measures data graphically (as well as a small lesson on the permutation test). Download the relevant data set and then complete the following tasks and answer the questions.

Strength Data

Trial data consisting of 3 treatment groups: No training (tx=1), weight training with light weights and high repetition (tx=2), and weight training with heavy weights and low repetition (tx=3). Subjects were followed for 7 weeks and a measure of muscle strength was recorded each week. The questions of interest are

- Does weight training have any impact on strength?
- Is there a difference between tx 2 and 3?
- Which training program works quickest to increase strength?

The assignment is to address these using graphical methods and a statistical approach. Tip: A great deal of Stata syntax for graphics is covered in the Chapter 2 lecture slides and recent lab materials.

1. On one plot, plot separate line plots of all subjects' trajectories of strength versus week (relevant commands: *xtline, graph combine*).
2. Fit a linear regression by subject, and repeat 1, replacing the strength with predicted strength based on these regressions (relevant commands: *xtline, graph combine,* along with code below for running regression by id).
3. Optional: Consider whether there are other plots that might be useful, and include one more if you want to. (For example, box plots of the distribution of strength by both week and tx (relevant command *graph box* – note you can have more than one *over(var)* statement as option). Or feel free to choose another)
4. Write a short paragraph describing what the plots suggest about the bulleted questions above.
5. Reduce the data for each subject to 1 number – the slope of the change in strength estimated in each person separately (can use program below). One way you could test the treatment effect of tx=3 versus tx=2 on this outcome (ignoring tx=1) is to use a standard two-sample t-test.

6. To further explore the relationship between treatment and group, try fitting a simple cross sectional model to this data as discussed in Chapter 3. Look the model with both naïve and robust standard errors. What happens if you expand the model to include a longitudinal term?

Regression by id

```
** Initialize variables used in regbyid
capture drop predy
gen predy = .
capture drop slope
gen slope = .
capture drop intercept
gen intercept = .
** Program to fit regression by subject
capture program drop regbyid
program define regbyid, byable(recall)
syntax [varlist] [if] [in]
marksample touse
capture matrix drop beta
regress `varlist' if `touse'
matrix beta = get(_b)
replace slope = beta[1,1] if `touse'
replace intercept = beta[1,2] if `touse'
capture drop predt
predict predt
replace predy = predt if `touse'
end
** Sort by id
sort id
** Get Slope by subject
quietly by id: regbyid y time
```