

## Public Health Sciences 33300: Applied Longitudinal Data Analysis - Winter 2024

Problem Set 1 - Due: Tuesday January 16, 2024

The data for this problem are from a study investigating treatment-related changes in symptomatology severity in a sample of schizophrenic patients. Subjects were assigned to one of four treatments: placebo, chlorpromazine, fluphenazine, and thioridazine, however, for this problem the three non-placebo drug groups have been combined into one group. Severity of schizophrenic symptomatology was assessed across time using the Inpatient Multidimensional Psychiatric Scale (IMPS) Item 79, "Severity of Illness," which was coded as: 1 = normal, not at all ill, 2 = borderline mentally ill, 3 = mildly ill, 4 = moderately ill, 5 = markedly ill, 6 = severely ill, or 7 = among the most extremely ill. Patients were sometimes classified by two psychiatric raters (in terms of the severity as measured by this scale) and when these raters differed an average of the two scores was used for that patient at that timepoint. The file SCHIZREP.DAT.TXT (at the Canvas site under Assignments) contains some of the data from this study. Specifically, in this file you'll find 1603 records from 437 patients with five fields of data:

field 1: Patient ID

field 2: IMPS79 (7-point measure of severity of illness)

field 3: Week - from 0 (baseline) to week 6 (most measurements were on weeks 0, 1, 3, & 6)

field 4: treatment group (0 = placebo 1 = drug)

field 5: sex (0 = female 1 = male) *ignore this variable*

For this problem set do the following:

1. Obtain the IMPS79 means, std deviations and sample sizes across the 7 timepoints. Comment on the values you obtain and what might be suggested for a statistical modeling of these data (ideas about time-related trends might help here).
2. For this question ignore treatment group, and just fit random-effects model(s) for the trend in IMPS79 scores across time. Using IMPS79 as your dependent variable, examine whether the overall trend across timepoints is linear or quadratic. Regarding the random effects, is there significant individual-level variation in the trends across time; specifically, is there significant individual-level variation in first, the linear, and then, the quadratic trend? Write down both the within-subjects and between-subjects components for your final model. Describe the meaning of the various model parameters and your conclusions regarding their statistical significance.
3. Now, investigate whether there is evidence for differential trend due to treatment group. Perform an analysis including the main effect of treatment and any treatment by time interactions that you see fit. Describe the significance of these additional model parameters and what they may suggest about the idea of better living through chemistry. Write down both the within-subjects and between-subjects components for your final model. Describe the meaning of the various model parameters. Obtain the means across time for the two treatment groups and compare these with the estimated means derived from your final model. How well does this model fit the observed means? Use Stata or graph paper (or some other program) to plot the observed means for both groups against the estimated means for weeks 0, 1, 3, & 6.