ADA2: Homework Ch 07, ANCOVA, due Mar 28 Tursday

## ANCOVA model: Faculty political tolerances

A political scientist developed a questionnaire to determine political tolerance scores for a random sample of faculty members at her university. She wanted to compare mean scores adjusted for the age for each of the three categories: full professors (coded 1), associate professors (coded 2), and assistant professors (coded 3). The data are given below. Note the higher the score, the more tolerant the individual.

Below we will fit and interpret a model to assess the dependence of tolerance score on age and rank.

```
tolerate <- read.csv("http://statacumen.com/teach/ADA2/worksheet/ADA2_WS_11_tolerate.csv")
tolerate$rank <- factor(tolerate$rank)
# set 3="Ast" as baseline level
tolerate$rank <- relevel(tolerate$rank, "3")
str(tolerate)
```

```
'data.frame':   30 obs. of  3 variables:
 $ score: num  3.03 4.31 5.09 3.71 5.29 2.7 2.7 4.02 5.52 4.62 ...
 $ age  : int  65 47 49 41 40 61 52 45 41 39 ...
 $ rank : Factor w/ 3 levels "3","1","2": 2 2 2 2 2 2 2 2 2 2 ...
```

(1.) Below is a plot of tolerance against age, using rank as a plotting symbol. Describe what do you find from the plots, (for example, how tolerance score depends on age within ranks).

Tolerance score data
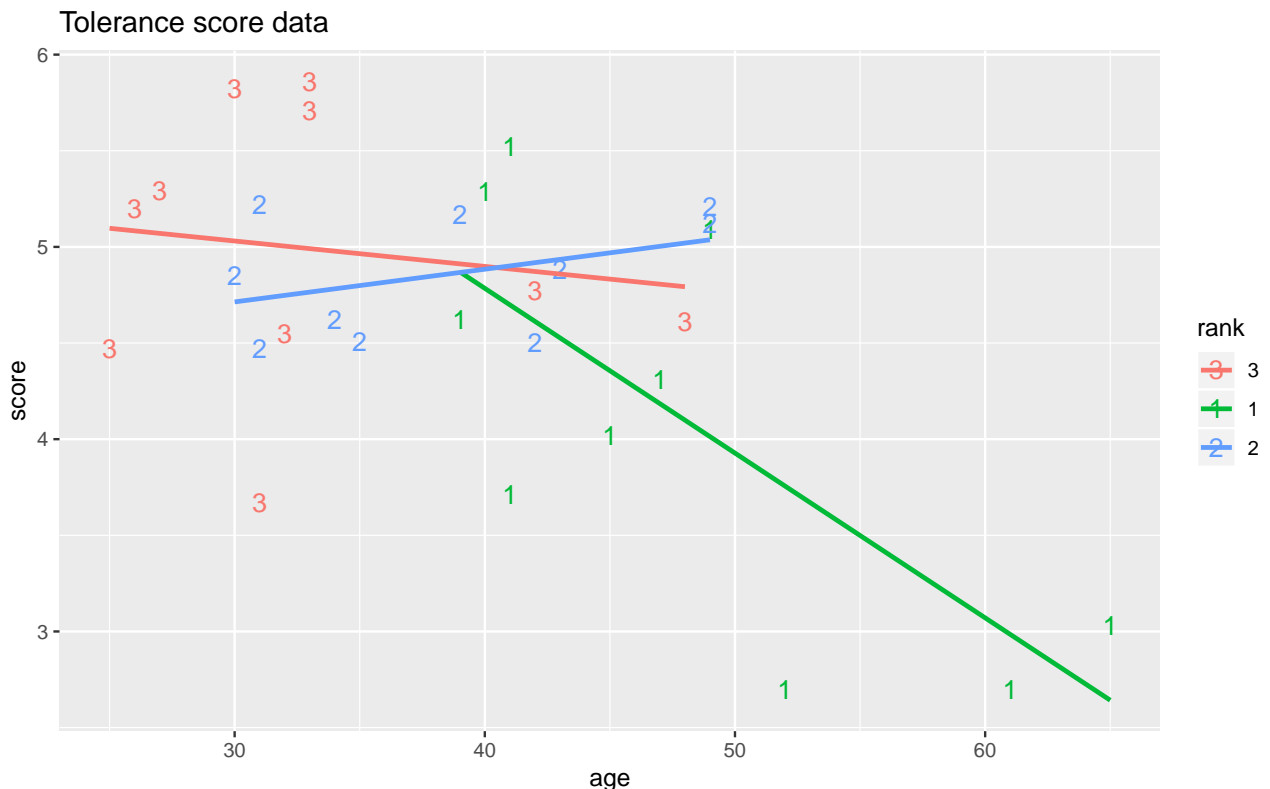
(2.) Create indicators for full and associate professors, so that assistant professors serve as the reference group. Write the full model (general model that allows each rank to have it's own intercept and slope), then the separate model for each rank using general notation. (please use β_i to illustrate instead of fitted coefficient)

(3.) Fit the general model, use the Wald test to perform pairwise comparisons for the regression line slope and intercept between ranks.

(4.) Test for equal slopes. If the hypothesis of equal slopes is plausible, fit the model of equal slopes and test whether intercepts are equal.

(5.) Reduce the model according to result from **(3)**, check assumptions. If assumptions are violated, do transformations, and check assumptions again.

(6.) Write the fitted model equation.

(7.) One feature to notice is that the observation **7** in the group of full professors appears to have an unusually low tolerance for his age (**2.70 52 1**). Now let's temporarily hold this observation out of the analysis, fit the general model. This observation has a fairly large impact on the estimated intercept and slope for the full professor regression line, but has no effect whatsoever on the estimated intercepts or slopes for the two other ranks. Why?

```
# exclude observation 7 from tolerate7 dataset
tolerate7 <- tolerate[-7,]
```