36-402/608 Homework #2 due 10:30AM 1/28

1. Serial correlation simulation (25 points)

   There are many ways to violate the uncorrelated errors assumption in regression. One of these is serial correlation, and one form of serial correlation is the "AR1" or autoregresive-1 correlation structure. In this structure, there is a parameter, say ar, such that the correlation between any two observations separated, say, t time units apart, equals ar^t. This parameter must be in $[0, 1]$, where 0 corresponds to uncorrelated error. We will learn more about this structure in a future week, but for now you need understand that bigger values of ar indicate a greater degree of assumption violation.

   Examine and load the function `ARcorSim()` from the file ARcorSim.R. Use this function along with `summary()` to calculate the power for nsim=1000, n=25, and $\beta_1 = 0$ over the set of "ar" values of 0, 0.2, 0.4, and 0.8 for both $\beta_0 = 0$ and $\beta_0 = 0.2$. Turn in a plot with power on the y-axis, "ar" on the x-axis, and separate curves for the two intercept values. Use `text()` to label the curves. Also turn in a brief statement of what you can learn about the robustness of regression analysis for estimation of the intercept under violation of the independent errors assumption through serial ar1 correlation.

2. Power study (25 points)

   A proposed experiment is designed to estimate the slope of the relationship between the explanatory variable "amount of fertilizer" and the outcome "plant growth". A reliable estimate of the error variance is $\sigma^2 = 64$. A reliable estimate of the intercept is $\beta_0 = 120$. (These estimates come from previous research reports.) The minimum meaningful slope is $\beta_1 = 0.5$. A total of 30 plants will be used, and we need to choose between two ways of selecting what fertilizer levels to use for these 30 plants. Using 1000 simulations each, calculate and turn in the power for

   (a) half the plants at each of two fertilizer levels of 0 and and 9 vs.

   (b) 3 plants each at fertilizer levels 0 through 9 at each whole number.

   Turn in your R code, too.

   Bonus question: What advantage is there to using all ten levels of fertilizer?

continued on other side

3. Arm strength study (25 points)

   A study was designed to see if people who identify as "ambidexterous" have stronger left or right arms. The apparatus can only compare arms, and cannot make a quantitative measurement for either arm. The results are that 20 subjects were stronger on the right, 5 were equal, and 27 were stronger on the left. Compute a p-value for the null hypothesis that either arm is equally likely to be stronger. Show your work.

4. Do Sleuth problem 28 on page 107 using Darwin.dat. (25 points)