Problem set is due Thursday, Feb. 21st in class.

1. Assume a straightforward linear regression model.

   (a) Suppose that we have a regression that meets the Gauss-Markov assumptions except that the disturbances have different variances and are correlated with each other. Is the OLS estimator unbiased anyway? Why or why not?

   (b) Suppose that the Gauss-Markov conditions hold, except that the covariance between the disturbances is positive. What is the true sampling variance of the OLS estimator in this case?

   (c) Regression programs in computer packages will assume that the observations are uncorrelated and will print out a different estimate of the sampling variance of $\hat{\boldsymbol{\beta}}$ for the question above. What is that estimate? Compared to the true sampling error you computed above, is this "nominal sampling variance" printed by the computer program always bigger, always smaller, or can't you say?

   (d) Now what if the covariance between the observations is σ^2? What is the sampling variance of $\hat{\boldsymbol{\beta}}$ in this case? Say in plain language what your result means.

2. Assume a simple regression model with one independent variable. Suppose that each disturbance is either $+0.5$ or -0.5, with probability 50% each. All the other Gauss-Markov assumptions hold. Further suppose that $\beta_0 = \beta_1 = 1$ and there are just three observations. The independent variable equals 1 in the first observation, 2 in the second, and 3 in the third.

   (a) Show that there are eight possible samples that could be drawn, each with three observations. List the three values apiece of x and y that occur in each possible sample.

   (b) Run all eight regression in **R**. List the value of $\hat{\beta}_1$ for each of the eight regressions. What is the variance of these eight numbers? (Note that these eight numbers are the entire population, not a sample, so divide by eight in computing the variance, not seven.) What do we call this variance?

   (c) What is the true σ^2? What is the true variance of the independent variables? Hence what is the true sampling variance of $\hat{\beta}_1$ according to the formula for OLS? How does it compare to your answer in (b)?

   (d) Look at the estimated sampling variance of $\hat{\beta}_1$ in each of your eight regressions. These eight numbers are each an estimate of the true sampling variance. How well these eight estimates do in estimating the true variance you found in (b)?

   (e) Now repeat 2(a)-2(d), but this time use disturbances that are either +1 or -1 with probability 50% each. Do the sampling variances change in the way that theory predicts?[1]

[1] Problems draw on Achen to a greater or lesser extent.