Some Remarks on Standard Errors

Suppose you have data on (Y_i, X_i) where you believe the data are iid but n is small. You want to regress Y_i on X_i and you want to quantify the sampling uncertainty associated with your coefficients. What do you do?

Proposals:

(a) `reg Y X`
(b) `reg Y X, robust`
(c) `reg Y X, hc2`
(d) `reg Y X, hc3`
(e) "the bootstrap"

Proposal (a) assumes homoskedasticity, which there isn't really any good reason to presume holds, although it might be interesting as a reference point. Proposal (b) estimates the heteroskedasticity, but the resulting variance estimator is finite-sample biased even under the optimistic scenario that the errors are homoskedastic. Proposal (c) tends to work pretty well, as does (d), which might even be too conservative. Generally, all four of these methods have variance estimators of the form

$$\widehat{V} = (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\widehat{\Sigma}\mathbf{X} (\mathbf{X}'\mathbf{X})^{-1} \tag{1}$$

where $\widehat{\Sigma}$ is an estimate of the variance matrix of the fitted residuals. The first four proposals' definitions of $\widehat{\Sigma}$ are

(a) $\widehat{\Sigma} = \widehat{\sigma}^2 I$, in which case $\widehat{V} = \widehat{\sigma}^2 (\mathbf{X}'\mathbf{X})^{-1}$
(b) $\widehat{\Sigma} = \text{diag}\left\{\widehat{\varepsilon}_i^2\right\}$
(c) $\widehat{\Sigma} = \text{diag}\left\{\widehat{\varepsilon}_i^2/(1-\widehat{h}_i)\right\}$, where $\widehat{h}_i = X_i' (\mathbf{X}'\mathbf{X})^{-1} X_i$ is the "leverage" of the ith observation
(d) $\widehat{\Sigma} = \text{diag}\left\{\widehat{\varepsilon}_i^2/(1-\widehat{h}_i)^2\right\}$.

The final proposal is related to the jackknife, which is similar to, but different from, the bootstrap.

What is meant by "the bootstrap"? There are two big choices in deciding how to bootstrap: (1) What kind of resampling scheme you want to employ, and (2) what is to be computed for each resample. Regarding (1), you might resample from the data (this is called the pairs bootstrap, aka the nonparametric bootstrap), or you might resample from groups of the data (this is sometimes called the cluster bootstrap), or you might resample from another distribution and distort the moments corresponding to your model.

Regarding (2), it is often said that it is better to compute an "asymptotically pivotal statistic" for each resample. This is not usually what "regression runners" do, although maybe they should.

Discuss:

1. Compute standard deviation of estimates across bootstrap samples. Presumes 2 parameters of sampling distribution are all that are necessary. Did you believe in normality when you decided to bootstrap? Presentation problems.
2. Compute percentiles of estimates and report a confidence region? Typically report 2.5% and 97.5% quantities. This works really badly unless the distribution is symmetric.
3. Usually best to report the following confidence interval:

$$LHS = \widehat{\theta} - s(\widehat{\theta})q^*(1-\alpha/2) \tag{2}$$
$$RHS = \widehat{\theta} - s(\widehat{\theta})q^*(\alpha/2) \tag{3}$$

where $q^*(\cdot)$ is the quantile function for $(\widehat{\theta}^* - \widehat{\theta})/s(\widehat{\theta}^*)$