Appendix D: Brief Review of Monte Carlo Simulation

The following is a brief review of the basics of a Monte Carlo simulation study.

When analytical arguments are intractable, a popular way to learn about the finite-sample properties of estimators is by Monte Carlo simulation. The objective of a simulation is to approximate the sampling distribution of an estimator by generating (via random deviate generation routines) some large number S independent data sets from a known situation and computing the estimator for each data set. The sample mean of the estimates over all S data sets is an estimate of the mean of the sampling distribution of the estimator; similarly, the standard deviation of the estimates over the S data sets is an estimate of the standard deviation of the sampling distribution (how good these quantities are at capturing the true features of the sampling distribution obviously depends on the size of S).

To carry out a simulation to evaluate to properties of several competing estimators for some parameter $\boldsymbol{\beta}$ ($p \times 1$) in a statistical model and for how well large sample approximations to the true sampling distribution work, the following are basic steps in a simulation.

- Generate S data sets from a scenario of interest. This scenario represents the true statistical model generating the data, with true value of $\boldsymbol{\beta}$ equal to some $\boldsymbol{\beta}_0$.
- For each data set, estimate $\boldsymbol{\beta}$ using each of the competing methods under consideration.
- Also obtain standard errors for the components of each estimator for $\boldsymbol{\beta}$ using the accompanying large sample theory approximation to the sampling distribution for that estimator.

Let $\beta_{k,0}$ be the kth component of the true value $\boldsymbol{\beta}_0$, $k = 1, \dots, p$. Let $\widehat{\boldsymbol{\beta}}$ be one of the estimators, and let $\widehat{\beta}_k$ be its k component, $k = 1, \dots, p$. Let $\widehat{\boldsymbol{\beta}}_s$ be the estimate obtained from the sth data set, $s = 1, \dots, S$, and let $\widehat{\beta}_{k,s}$ be its kth element, $k = 1, \dots, p$.

Then, for each estimator, do the following.

(i) If the estimators are consistent, we would hope that they would be approximately unbiased in finite samples. Thus, we would hope that the mean of the sampling distribution is close to the true value of $\boldsymbol{\beta}$, $\boldsymbol{\beta}_0$. with only minimal bias. To assess this based on the S observations from the sampling distribution, calculate the *Monte Carlo bias* for each component of an estimator $\widehat{\boldsymbol{\beta}}$, defined for the kth component as

$$S^{-1}\sum_{s=1}^{S}\widehat{\beta}_{k,s}-\beta_{0,k}.$$

It is standard to report this "raw" Monte Carlo bias. It is also standard to report this bias relative to to the true value (so report for each k

$$\frac{S^{-1}\sum_{s=1}^{S}\widehat{\beta}_{k,s}-\beta_{0,k}}{\beta_{0,k}},$$

which can of course be problematic when the true value is very close to 0) and relative to the Monte Carlo standard deviation (see (iii) below), so that the size of the bias relative to the variation in the estimator can be assessed.

(ii) To compare the precision of two competing estimators based on the S estimates of each, we could compare their sample variances, thus mimicking the idea of asymptotic relative efficiency. However, because the estimators may exhibit some finite sample *bias* for finite m in our case, it is standard instead routine to take this into account and compute the *Monte Carlo mean square error* (MSE) for each estimator. The estimated MSE based on the S estimates $\widehat{\boldsymbol{\beta}}_s$ for the kth component is defined as

$$S^{-1}\sum_{s=1}^{S}(\widehat{\beta}_{k,s}-\beta_{0,k})^2 = S^{-1}\sum_{s=1}^{S}(\widehat{\beta}_{k,s}-\bar{\beta}_k)^2+(\bar{\beta}_k-\beta_{0,k})^2,$$

where $\bar{\beta}_k$ is the sample average of the $\widehat{\beta}_{k,s}$. Note that MSE may thus be interpreted as sample variance over the S estimates plus observed bias, squared.

The ratio of estimated MSE values may be used as a measure of relative precision, similar to asymptotic relative efficiency. One ordinarily calculates MSE for each component k and then forms the ratio for each k. It is customary to put the MSE for the estimator that is thought to be *more efficient* in the numerator, so that a MSE ratio less than 1 reflects the relative inefficiency of the estimator whose MSE is in the denominator. It can of course be done either way as long as the user defines the MSE ratio so that it can be interpreted appropriately.

(iii) To assess how well the estimated standard errors approximate the true sampling variation, one can compare the sample standard deviation of each component of the S estimates $\widehat{\boldsymbol{\beta}}$, that is, the *Monte Carlo standard deviation*, to the average of the estimated standard errors for that component found using the large sample theory. If the theory is relevant, we would expect the sample standard deviation, an approximation to the true sampling variation, and the average of estimated standard errors, to be "close." Sometimes, the ratio of the two is formed for each component k to get a sense of this.

(iv) To assess further how well the approximate large sample sampling distribution approximates the true sampling distribution, for each estimator, one might calculate for each of the S data sets 95% Wald confidence intervals (using the usual critical value from the standard normal distribution of 1.96) for the true values of each component of β, and record the proportion of times that the intervals contain the true values. These proportions are *Monte Carlo coverage probabilities* – if the Wald intervals are reliable, we would expect them to be close to the nominal coverage probability of 0.95. If the Monte Carlo values are not close to 0.95, then using the large-sample approximation may be unreliable.