**Chapter 23 Summary**

The primary objective of this chapter was to introduce the method of bootstrap sampling, which is becoming popular in applied econometrics. The classical linear regression model, the workhorse of econometrics, is based on the critical assumption that the regression error terms are independently and identically distributed (IID) as a normal distribution with zero mean and constant variance.

With this assumption, it is easy to show that the OLS estimators of the regression parameters are best linear biased estimators (BLUE) of their population values. With the normality assumption, it was easy to establish that the OLS estimators are themselves normally distributed. As a result, it was easy to establish confidence intervals and test hypotheses regarding the population values of the estimators.

In many situations the assumption that the error terms are IID may not be tenable. In that case, one has to resort to the large sample, or asymptotic, theory to derive the statistical properties of the estimators, such as asymptotic normality. However, the results of the asymptotic theory are difficult to apply in finite samples.

In these situations, one can use bootstrap sampling to derive the sampling distributions of the OLS estimators, particularly the standard errors of the estimators. Without that, it is not possible to establish confidence intervals and test statistical hypothesis.

The key feature of bootstrap sampling is that we have a random sample from some population of interest and from that single sample we draw repeated samples, equal in size to the original sample, but with the important proviso that every time we draw a member of the sample we put it back before drawing another member from the original sample. This is called sampling with replacement. We can draw as many bootstrap samples as we want. In practice, we may have to draw 500 or more

bootstrap samples.

From the bootstrap samples, we can obtain the empirical frequency distributions (EFD) of the parameters of the model. From the EFD, we can estimate the mean value of the parameters and their standard errors and then engage in statistical inference – confidence intervals and hypothesis testing.

We considered four methods of establishing confidence intervals: (1) the normal approximation, (2) the percentile method, (3) the bias-corrected (BC) and accelerated bias-corrected (ABC) methods, and (4) the percentile *t* method. We discussed the pros and cons of each method and illustrated them with a simple example of 50 bootstrap samples from the standard normal distribution.

We also discussed the use of these methods in regression modeling. In particular, we estimated the wage regression model first introduced in Chapter 1, using 500 bootstrap samples, and showed how we interpret the results. Since the sample size in this example was fairly large (*n* = 1,289), the results of the bootstrap methods and the standard OLS method do not differ vastly. Of course, that may not be case in nonlinear parameter models, models with substantial number of outliers, models with heteroscedastic error terms, or models involving ratio variables.

Although a very useful empirical tool, we discussed the pros and cons of bootstrapping. Bootstrapping may not be appropriate if the original sample is not a random sample. Other situations where bootstrapping might fail are: (1) if the sample size is too small, (2) distributions with infinite moments, (3) estimating extreme values, and (4) survey sampling.

In parametric problems, like regression analysis, we have several diagnostic tools to check on the assumptions underlying the analysis. Since non-parametric sampling, such as bootstrapping, involves minimal assumptions, it may not be easy to assess the conditions in which bootstrapping may work or fail. However, some diagnostic tools have been developed for assessing non-parametric bootstrap. Efron (1992), for example, introduced the concept of a jackknife-after-bootstrap measure to assess the non-parametric bootstrap.

The basic idea behind jackknifing is to take repeated subsamples of the original sample, eliminating observations one at a time, and see the effect on the original bootstrap calculations. Of course, this is computationally very intensive, for we would be repeating bootstrap resampling n times, n being the original sample size.

Research on bootstrapping is evolving and the reader is urged to surf the Internet for the latest developments in the field.

Since the objective of this chapter was to give the bare essentials of bootstrapping, we have not touched on many aspects of this topic, such as bootstrap vs. asymptotic tests, the (statistical) power of bootstrap tests vs. exact tests, the power of bootstrap hypothesis tests, and the like. The interested reader may consult the references.