Financial Econometrics
Prof Olivier Scaillet
TA Gaetan Bakalli

TP 8

Bootstrap and Asymptotic Confidence Intervals

This TP shows how the bootstrap can be used to generate confidence intervals.

The Delta Method

The delta method is a popular way to perform inference on nonlinear functions of model parameters. It is based on an asymptotic approximation and states the following result:

$$\text{if } \sqrt{T}(\hat{\theta}-\theta) \sim \mathcal{N}(0,\Sigma), \text{ then } \sqrt{T}\Big(g(\hat{\theta})-g(\theta)\Big) \sim \mathcal{N}\Big(0, \frac{\partial g}{\partial \theta}\Sigma\frac{\partial g'}{\partial \theta}\Big),$$

where g is a continuous function of the parameter vector θ. In practice, Σ is estimated by the estimated covariance matrix of $\hat{\theta}$ and $\frac{\partial g}{\partial \theta}$ is evaluated at $\hat{\theta}$.

You will also need the following property:

$$\text{if } \varphi \sim \mathcal{N}(\mu,\sigma^2), \text{ then:}$$
$$\sqrt{T}(\hat{\mu}-\mu) \sim \mathcal{N}(0,\sigma^2), \quad \sqrt{T}(\hat{\sigma}^2-\sigma^2) \sim \mathcal{N}(0,\sigma^4), \quad \text{and } \hat{\mu} \text{ and } \hat{\sigma} \text{ are orthogonal.}$$

1. Simulate $T = 100$ observations of the law $\mathcal{N}(1,16)$. (We will consider these observations as a series of monthly percentage returns on a stock.)
2. Estimate the model parameter $\theta = \begin{pmatrix} \mu \\ \sigma^2 \end{pmatrix}$ and compute the statistic $\widehat{SR} = \frac{\hat{\mu}}{\hat{\sigma}}$. Think of $\widehat{SR}$ as a Sharpe ratio, if we assume the risk-free rate is zero.
3. Using the delta method and the above property, form a 95% asymptotic confidence interval for SR. Hint: $\widehat{SR} = \frac{\hat{\mu}}{\sqrt{\hat{\sigma}^2}}$.

Bootstrap Confidence Interval

If the number of observations is small, it is often better to use bootstrap confidence intervals.

1. First, generate $B = 2000$ random samples of size T from the original data simulated in the previous section (see slide A.XIII-119). Use each of these samples to obtain new estimates of our statistic, $\widehat{SR}^{(b)}, b = 1,\ldots,B$.

In general, we do not necessarily know anything about the distribution of the new parameter estimates $\widehat{SR}^{(b)}$. But, although we may not have a perfect idea of the shape of this distribution, we can calculate quantiles $q^\star(\alpha)$, such that a fraction α of the bootstrap statistics are less than or equal to $q^\star(\alpha)$. For example, if we had 1000 bootstrap samples, the quantile $q^\star(0.05)$ would be the 50th largest observation. This means we can write:

$$\mathrm{P}\Big(q^\star(\alpha/2) \leq \widehat{SR}^{(b)} \leq q^\star(1-\alpha/2)\Big) = 1-\alpha.$$

Now suppose we want to look at the distribution of $\widehat{SR}^{(b)} - \widehat{SR}$. From the expression above, we can see that:

$$\mathrm{P}\Big(q^\star(\alpha/2) - \widehat{SR} \leq \widehat{SR}^{(b)} - \widehat{SR} \leq q^\star(1-\alpha/2) - \widehat{SR}\Big) = 1-\alpha.$$

In addition, we can argue that we can estimate the distribution of $\sqrt{T}(\widehat{SR} - SR)$ by the distribution of $\sqrt{T}(\widehat{SR}^{(b)} - \widehat{SR})$. This makes sense if you think of the analogy that $\widehat{SR}$ arose from sampling from a distribution with parameter θ, while $\widehat{SR}^{(b)}$ arose from sampling from a distribution with parameter $\hat{\theta}$. Hence:

$$\begin{aligned}\mathrm{P}\Big(q^\star(\alpha/2) - \widehat{SR} \leq \widehat{SR}^{(b)} - \widehat{SR} \leq q^\star(1-\alpha/2) - \widehat{SR}\Big) &= \\ \mathrm{P}\Big(q^\star(\alpha/2) - \widehat{SR} \leq \widehat{SR} - SR \leq q^\star(1-\alpha/2) - \widehat{SR}\Big) &= 1-\alpha.\end{aligned}$$

Rearranging:

$$\mathrm{P}\Big(2\widehat{SR} - q^\star(1-\alpha/2) \leq SR \leq 2\widehat{SR} - q^\star(\alpha/2)\Big) = 1-\alpha.$$

This means the bootstrap confidence interval for SR is:

$$\Big[2\widehat{SR} - q^\star(1-\alpha/2),\ 2\widehat{SR} - q^\star(\alpha/2)\Big].$$

2. Form a 95% bootstrap confidence interval for SR.

3. Compare the asymptotic and bootstrap confidence intervals for different values of T.

Dependent Data: The Block Bootstrap

The bootstrap procedure just presented, however, only works with independent data. If a dependency between neighboring observations is suspected, as may be the case with high-frequency stock data, our original bootstrap procedure will not work. To deal with this issue, a procedure known as the block bootstrap has been proposed in the literature. In the block bootstrap, groups (i.e., blocks) of consecutive observations are sampled instead of individual observations. In its simplest form, the data is divided into k non-overlapping blocks of length l, where $T = kl$. If k is large enough, then the generated samples should preserve most of the dependency between neighboring observations. Under a number of relatively mild conditions, it can be shown that this estimator is consistent, though its rate of convergence may not be as high as that for the iid bootstrap seen above. In this exercise, we will get a sense of the relative accuracy of the two types of bootstrap.

1. Generate again $B = 2000$ random samples of size T from the original data simulated in the first section. This time, however, do the sampling based on blocks of size 10, i.e., $k = 10$. Use each of the generated samples to obtain new estimates of our statistic SR, denoted by $\widehat{SR}^{(BB)}, b = 1, \ldots, B$.

2. Using a similar approach to the previous section, form a 95% bootstrap confidence interval for SR.

3. How does the new confidence interval obtained from block sampling compare with the confidence interval computed in the previous section? Is it wider? Repeat steps (1)-(2) several times to get a good idea of how both confidence intervals compare. Comment