

Computer Lab Week 11

STAT221

Bootstrapping Quantiles

One of the major benefits of bootstrapping is that it does not require assumptions about the population distribution, or any statistical theory for the distribution of the statistic. So it's useful when the statistic has an unknown or complicated sampling distribution. Here we will explore the sampling distribution of various quantiles, for which there is no general result for their sampling distribution.

1. Simulate $n = 200$ observations from a standard normal distribution. Use the `quantile` function to estimate the 10%, 25%, 50%, 75% and 90% quantiles of the distribution, using this sample data. These are often referred to as the empirical quantiles. Also calculate the theoretical quantiles for the standard normal (for example by using `qnorm`).
2. Use bootstrapping to estimate a 90% confidence interval for each of these quantiles.
3. Produce a density histogram of each of the bootstrapped quantiles. Notice how the sample median is close to symmetric, but the others are less symmetric. In fact, as you go further into the tail the sampling distribution will get less and less symmetric. You may want to try some really extreme quantiles, e.g. 5% and 95%. Add vertical lines to represent the empirical and theoretical quantiles, along with the 90% confidence interval. Think about why the median might be more symmetric than the quantiles further out in the tails.
4. Repeat the above for $n = 200$ observation from a standard exponential distribution, again focussing on the behaviour of the sampling distributions for the 10% and 90% quantiles.