Computational Methods in Econometrics: Assignment 2

The purpose of this assignment is to study trends in temperature series. We can use our econometric and statistical knowledge together with the newly discussed bootstrap methods. Consider the following linear regression model

$$y_t = \beta_0 + \beta_1 t + \varepsilon_t, \qquad t = 1, \ldots, n. \tag{0.1}$$

The time series $\{y_t\}$ can represent different things. In our case, it represents the average annual temperature at the well-known KNMI weather station in De Bilt (NL) from 1900 to 2014.[1] In this model, the time series is decomposed into a linear trend with intercept β_0 and slope β_1 and an error process $\{\epsilon_t\}$. To test for the presence of a significant upward trend, we consider the pair of hypotheses H_0: $\beta_1 = 0$ vs. H_1: $\beta_1 \geq 0$. We use the t-statistic

$$T_n(Y) = \frac{\hat{\beta}_1}{\sqrt{\widehat{\mathrm{Var}(\hat{\beta}_1)}}},$$

where $\hat{\beta}_1$ is the OLS estimator of β_1 and $\widehat{\mathrm{Var}(\hat{\beta}_1)}$ denotes a consistent estimator of the variance of $\hat{\beta}_1$. The significance level is given by α.

We can make different assumptions about the error terms of our model. Assuming that $\varepsilon_n \sim \mathrm{IID}(0, \sigma_\varepsilon^2)$ would allow us to use critical values from the standard normal distribution. However, this assumption is very restrictive in practice and it is likely not satisfied in temperature data. The residuals from our trend estimation are likely to suffer from serial correlation and/or heteroskedasticity. Therefore, we wish to exploit inference methods that are robust to either serial correlation or heteroskedasticity (or maybe even both?).

We can simply estimate our linear trend model by OLS and obtain $\hat{\beta}_{OLS} = \beta + (X'X)^{-1}X'\varepsilon$ and residuals $\hat{\varepsilon}_{OLS} = y - X\hat{\beta}_{OLS}$, where the X-matrix consists of only two columns (can you derive them from the trend model?). In case of heteroskedasticity and autocorrelation, we can use the famous HAC standard errors proposed by Newey & West to be robust to both. *(You can use a package to obtain HAC standard errors.)*

You can find the data for this assignment in the *DeBilt1900to2014.csv* file on Canvas.[2] The first column gives the years for which the average has been computed from daily measurements. The second column presents the annual temperature averages for De Bilt. This is our first series of interest (the other two columns will be used later in this assignment).

[1] Find an article (in Dutch) about the weather station here.
[2] The data is publicly available at CBS.

Exercise 1.

a. Estimate equation (0.1) using the annual averages and perform a t-test using standard normal critical values and HAC standard errors.

b. Now, perform the test using the following bootstrap methods:

   (a) Nonparametric residual bootstrap (i.i.d. bootstrap)

   (b) Block bootstrap

   (c) Wild bootstrap: use the standard normal and the Rademacher distribution

   (d) Sieve bootstrap

   For each method, calculate the bootstrap p-value. *(For the sieve and the block bootstrap, it is encouraged to play around with different lag orders p and different block lengths ℓ. But you can also use a default value of $p = 5$ and $\ell = n/10$.)*

c. Report and shortly discuss your results.

Scientific evidence indicates that there is seasonal variation in the warming trends. Research has shown that winters tend to get warmer faster than summers. To investigate this, we use the third and fourth column of our data file. The third column contains the yearly averages calculated over the winter months and the fourth column uses only the summer months. Given that you have done the programming of all the bootstrap methods in the previous exercise, this next exercise should be a quick one.

Exercise 2.

a. Repeat the previous exercise using the winter and summer averages.

b. Briefly comment on your findings. Do the seasonal averages from De Bilt give further indication for faster warming rates in winter?

If we are not interested in the sign of the slope coefficients, but only care about whether they are significant, we could perform a two-sided version of the above test. However, instead of performing a t-test, we can also construct confidence intervals around the coefficient estimates. In practice, bootstrap methods are often used for the construction of confidence intervals due to their many desirable properties. As we have learned in this course, there are different ways to obtain confidence intervals using bootstrap methods.

Exercise 3.

a. For the yearly average temperatures and for each bootstrap method considered above, construct the following confidence intervals:

   (a) Equal-tailed percentile intervals

   (b) Equal-tailed percentile-t intervals

   (c) Symmetric percentile intervals

   (d) Symmetric percentile-t intervals

   Can you use the same bootstrap algorithms as for Exercises 1 and 2 or do you have to construct different procedures? Please explain your reasoning.

b. Comment on your findings. How do they compare to the previous findings?

How do we decide which way of constructing intervals performs best and which bootstrap method to use? This can vary depending on the underlying model and error structures. In this final exercise, you will perform a small Monte Carlo experiment to investigate the empirical coverage of the proposed confidence intervals using model (0.1).

Exercise 4.
Generate data from model (0.1), using $\beta_0 = 1$, $\beta_1 = 0.5$ and $n = 500$. With your group, decide on one or more specifications for your error terms (i.i.d., ARMA or AR model, variance depending on time). Based on this choice, decide on at least two bootstrap methods that you would like to compare. With the generated data, calculate at least two different types of bootstrap confidence intervals around the estimated slope coefficient for your bootstrap methods of choice. Check whether the true β_1 is contained in your confidence intervals. Repeat this N times and report the empirical coverages. Comment on your results.

GOOD LUCK!