DATS 6313 – Time Series Analysis & Modeling

Instructor: Reza Jafari

Lab #1

Bradley Reardon

1/21/2022

1 – Abstract:

This lab pertains to using various visual and statistical methods for determining if a subset of data is stationary. The methods used throughout this lab are visualizing data using graphs and using both an ADF-test and KPPS-test. The results and observations of all three methods are compared to determine if they reinforce one another.

2 – Introduction:

This experiment was performed to increase understanding of stationary versus non-stationary data, and how to test for such using both visual and statistical approaches. Stationary data has a consistent mean, variance, and autocorrelation structure that do not change over time, meaning they metrics are constant throughout the subset of data being tested. As mentioned in the abstract, some methods for testing if a subset of data is stationary are visual cues, ADF-tests, and KPPS-tests.

3 – Method, Theory, and Procedures:

The first method we used was checking visual cues. To check for stationary data using a visual approach, we calculate the rolling mean and variance of a subset of data and plot it on a graph to see if the trend spikes at all or if it remains constant throughout. The rolling metrics are calculated by finding the difference between specific value and a value located at a designated interval from said specific value, and this is repeated for each sample in the subset of data.

The next method used was the ADF-test, which is a test to determine if a unit root is present in a time series sample, or also known as determining if the data is stationary. The ADF-test produces results that can be used inform the degree to which a null hypothesis can be rejected or failed to be rejected. If the null hypothesis is failed to be rejected, it suggests the time-series has a unit root and is non-stationary. If the null hypothesis is rejected, it suggests that the time series is stationary. The resulting p-value will help us determine whether to reject or not reject the null hypothesis, usually at a 5% threshold while working with a 95% confidence interval. Below the p-value threshold suggests we reject, and vice versa.

The final method used was the KPPS-test. Similarly, to the ADF-test, this is a statistical test used to determine if a time-series is stationary. One key difference between the two tests is that the null and alternative hypotheses are switched – with the KPPS test, null says the trend is stationary and the alternative says otherwise. Again, the p-value is used to determine whether to reject or fail to reject the null hypothesis and the criteria is the same as with the ADF-test.

4 – Answers to Lab Questions:

Fill in after checking with Divya

5 – Conclusion:

Whether a subset of data is stationary or not indicates if the mean, variance, and autocorrelation do not change over time. A few approaches for determining if your data is stationary include a visual check by plotting the rolling mean and variance, conducting an ADF-test, and conducting a KPPS-test. In this lab, we used all three approaches and cross referenced the results from each to see if our observations were consistent throughout. This lab involved data regarding quarterly sales for a small company over the period of 1981-2005. The features tested to determine if they were stationary are Sales, AdBudget, and GDP. I found some inconsistency amongst the features when performing the three tests, with the visual check seeming to show both Sales and GDP as non-stationary and AdBudget as stationary, the ADF-test following suit with the visual test observations, but the KPPS-test results showing all features as stationary. This proves that no single test is always representative of the actuality of stationary data, and that multiple tests should be used to cross-reference results with one another. One problem with the visual approach is that the observations are subjective and cannot always be interpreted correctly.