Week 1 Assignments

It is important that you start your RStudio, with no library added. The default library should do the following work.

1. **Assume the timeseries S has observations as . ∈ means “a member of”. Assume S has *t* observations. (5 points)**

* **Any one of the observations.**
* Each is a member of S.
* **First, middle, and Last observations**.
* First observation is
* Middle observation

If t is even, then will have two middle values and

If t is odd, then middle observation will be

* Last observation is
* **First forecasted observation and kth forecasted observation.**
* First forecasted observation
* kth forecasted observation
* **Error of the kth forecasted observation. What Error of the kth forecasted observation is equal to?**
* We cannot find the error if we do not have actual value.
* We can forecast the kth observation using .
* If the actual value is then the error of the kth forecasted observation will be

=

* **If the forecast horizon is *h* then, list the forecasted observations.**
* If the forecast horizon is h, then the forecasted observations would be , , …….,

1. **Generate 1000 random number with normal distribution (10 Points)**

* random\_numbers = rnorm(1000)
  1. **Transform this dataset to time series using ts().**
* random\_numbers = ts(random\_numbers)
  1. **Plot the timeseries using plot()**
* plot(random\_numbers, main="Timeseries Data")

A graph showing a number of data

Description automatically generated with medium confidence

* 1. What are start, end and the frequency of this time series?

A white background with black text

Description automatically generated

* 1. Show the series is white noise with normal distribution.

Here the series with mean is approximately 0 and variance is approximately 1 and the series is normal distributed as shown below.

A number and a number

Description automatically generated with medium confidenceA graph of random numbers

Description automatically generated

* 1. Save your dataset as “white\_noise.csv”

**(Implemented In R)**

1. Open the “white\_noise.csv” file. (10 Points)
   1. Generate 500 observations which each observation represents the following autoregressive equation:

where t represents any time index between 1:500 and

is the observation value at time t

is the white noise value at the time t

* 1. Save your dataset as “autoregressive\_2.csv”

**(Implemented In R)**

1. Open the Open the “white\_noise.csv” file (5 Points)
   1. Generate 500 observation with the random walk feature
   2. Generate 500 observation with the random walk with drift of 0.3?
   3. Show both random walk and random walk with drift of part *a* and *b* in one plot.

**(Implemented in R)**

A graph showing a graph of a walking path

Description automatically generated with medium confidence

1. **If timeseries dataset follows a general sinusoidal waveform as then**

**What are the amplitude, periodical frequency, and the phase of the following equation (5 Points)**

* From the given equation
* Amplitude (A) =5
* periodical frequency ( =
* phase shift () = =

**Plot 500 observations of this equation (t, 1:500) with and without error**

* The first plot is without error and second with error.

A graph of a wave

Description automatically generated with medium confidence

Part 2 (10 Points)

1. Page 42-44, problems 1and 5

**Problem 1:**1. Impact of September 11 on Air Travel in the United States: The Research and Innovative Technology Administration’s Bureau of Transportation Statistics (BTS) conducted a study to evaluate the impact of the September 11, 2001, terrorist attack on U.S. transportation. The study report and the data can be found at www.bts.gov/publications/estimated\_impacts\_of\_9\_11\_on\_ us\_travel. The goal of the study was stated as follows: The purpose of this study is to provide a greater understanding of the passenger travel behavior patterns of persons making long distance trips before and after September 11. The report analyzes monthly passenger movement data be tween January 1990 and April 2004. Data on three monthly time series are given in the file Sept11Travel.xls for this period: (1) actual airline revenue passenger miles (Air), (2) rail passenger miles (Rail), and (3) vehicle miles traveled (Auto). In order to assess the impact of September 11, BTS took the following approach: Using data before September 11, it fore casted future data (under the assumption of no terrorist attack). Then, BTS compared the forecasted series with the actual data to assess the impact of the event. Plot each of the three pre-event time series (Air, Rail, Car).

1. **What time series components appear from the plot?**

From the plot of each of the three pre-event time series (Air, Rail, Car), several time series components may appear:

**Trend:** The trend component shows the long-term movement or direction of the data. It may reveal whether the passenger travel behavior patterns were increasing, decreasing, or stable over time before the September 11 event.

**Seasonality:** Seasonality refers to patterns that repeat at regular intervals, such as yearly cycles. Seasonal patterns might appear if there are recurring trends in passenger travel behavior related to specific times of the year.

A graph of a travel

Description automatically generated

1. **What type of trend appears? Change the scale of the series, add trend lines, and suppress seasonality to better visualize the trend pattern.**

**Trend for Three types of travel:**

**Air.RPM\_Thausends**: The trend in the airline revenue passenger miles (Air.RPM\_Thausends) appears to be generally increasing over time. There was a slight decrease around 2002, followed by a slight increase in 2004. Overall, the trend shows an upward movement with some fluctuations.

**Rail.PM:** For rail passenger miles (Rail.PM), the trend shows a downward movement from 1990 to 1996, indicating a decrease in rail passenger miles during that period. However, from 1998 to 2004, the trend appears to be slightly upward, suggesting an increase in rail passenger miles during this time frame.

**Vehicle Miles Traveled**: The trend in vehicle miles traveled seems to be consistently increasing over the entire period, indicating a general upward movement in vehicle miles traveled over time.

**A graph of a trend

Description automatically generated with medium confidence**

**5. Souvenir Sales: The file SouvenirSales.xls contains monthly sales for a souvenir shop at a beach resort town in Queensland, Australia, between 1995 and 2001.8 Back in 2001, the store wanted to use the data to forecast sales for the next 12 months (year 2002). They hired an analyst to generate forecasts. The analyst first partitioned the data into training and validation periods, with the validation period containing the last 12 months of data (year 2001). She then fit a regression model to sales, using the training period.**

**A. Create a well-formatted time plot of the data.**

A graph showing the number of sales

Description automatically generated**b. Change the scale of the x axis, y axis, or both to logarithmic scale in order to achieve a linear relationship. Select the time plot that seems most linear.**

A graph showing a number of sales

Description automatically generated

1. **Comparing the two-time plots, what can be said about the type of trend in the data?**

* The second plot (with logarithmic scale) shows a more linear relationship between Date and Sales compared to the first plot, it suggests that the sales data may exhibit exponential growth or decay over time.