Homework 3

CSE3102: Applied Probability for Computer Science

Due date: 2020/10/20

A. Background

The increase in annual world crude oil production from 1880 to 1973 follows a pattern of exponential growth. In order to fit a linear model to these data, the oil production variable must be transformed by taking the natural log. A scatterplot of the log of oil production vs. year follows a straight line from 1880 to 1973. Political turmoil in the oil-producing regions of the mideast affected patterns of oil production after 1973.

- Methods: Scatterplot , Time series , Transformation
- Data file: "OilProduction.txt" from the i-class
- Output: Scatterplot of oil production vs. year

B. Problems

1. Draw graphs using GNU plot. (30pts)

    a. Download "OilProduction.txt" available on the i-class

    b. Comment out unnecessary parts (Note: Do not delete) and insert any comments if needed

    d. Draw graphs with the outputs in Line, Point (scatterplot), line+point, dots, Boxes, Impulses)

2. Data transform (50pts)

    a. Define a model to fit the oil production data. You should choose a model and select model parameters, manually or automatically. Visualize your model together with the scattered points in a plot chart.

    b. In order to understand the data better, write a code to transform the data as the way you need (possibly exponential transform with your estimated parameter) using a programming language (C/C++, JAVA, Matlab, etc.) at your convenience.

    b. Draw a scatterplot graph of the transformed data using GNU plot

3. Provide your own analysis and discussion about the oil production. (20pts)

What to submit: A report with (1) the answers from the problems and (2) the source code from the problem 2.