The University of British Columbia – Okanagan

DATA315

Assignment 3

Due on Mar 16 by uploading the file to Canvas that contains your assignment. The data sets that will be used are posted separately under data on canvas.

Question 1

Calculate the first 10 lag sample autocorrelations for the time series given below. You may use the `acf` function. For each time series, discribe the sample acf pattern and pre-determine which model might be appropriate from what you have learned so far (White noise, MA(1) and AR(1))? If there are no models that you have learned about that can fit the data, your answer will be "so far, there are no models that fit". Explain each case clearly.

(a) (2 points) Data for electroless nickel concentrations in a chrome plating process is in the data set called `nickel`. You can use `source("nickel.R")`.

(b) (2 points) The time series of the number of lynx trapped in Hudson's Bay Company territory and later Canada in the years from 1821 through 1934 is in an object called `lynx`. This is built into R.

(c) (2 points) Global average temperatures are recorded in terms of number of Celsius degrees above a baseline temperature. The baseline temperature is the average temperature for the year 1990. You can use `source("Globaltemps.R")` to get this data set which is not set up as a time series object yet. You need to use the `ts` function (`temps <- ts(temps, start = 1880, end = 2016)`) to transform `Globaltemps` vector to a time series vector in order to run the `acf` function)

(d) (4 points) The daily closing pricess for DAX (German) stock index is in the first column of the built-in data `EuStockMarkets`. First plot this time series and its sample autocorrelation function. Apply the `diff()` function to the natural log of the data and plot the resulting time series and the sample autocorrelation function.

Question 2

(a) (4 points) Fit an MA(1) model to the `nickel` data. Use a hypothesis test to test whether the MA parameter equals 0. Write down the fitted model. What are the forecasts for the 2nd and 3rd observations after the end of this time series? (You can answer this without using any function to do the calculation.)

(b) (2 points) Use one of the portmanteau tests we have learned about to check if the residuals from the model above are adequately modelled as white noise. (Suppose we want to check the first 10 lags). Write down the details of this hypothesis test.

(c) (4 points) Suppose the 75th observation has been deleted by accident. Provide a forecast for the 75th observation, together with an estimate of the standard deviation of the forecast error. How far is the true value from your forecast, in terms of numbers of standard errors?

(d) (4 points) Fit an AR(1) model to the `nickel` data. What are the forecasts for the 2nd and 3rd observations after the end of this time series? (Use the `predict` function to calculate these). Check if the residuals are adequately modelled as white noise.

Question 3

(6 points) A time series follows the model $y_t = \mu + \phi(y_{t-1} - \mu) + \varepsilon_t$. The first few measurements are $\{3.2, 3.2, 2.2, 2.3, 1.8, 1.3, 2.2, 2.7\}$. Use the method of moments to estimate μ, ϕ and σ **manually**.

Question 4

(5 points) Consider the `longitudinalAcceleration` time series. Calculate the sample first lag autocorrelation using the `acf` function. Assuming an AR(1) process is appropriate, obtain theoretical values for the second, third, and fourth lag autocorrelations, using the previously obtained estimate of the first lag autocorrelation. Compare these values with the sample autocorrelations obtained using the `acf` function. Does an AR(1) process seem like a reasonable approximation for these data?

Question 5

Suppose $x_0 = 0$ and

$$x_t = .5x_{t-1}$$

for $t = 1, 2, \ldots$.

(a) (2 points) Find x_1 and x_2.

(b) (2 points) Find a formula for x_t in terms of t only.

(c) (1 point) Find $\lim_{t\to\infty} x_t$.

(d) (4 points) Repeat (a), (b) and (c) for the case where $x_0 = 1$.

(e) (4 points) Repeat (a), (b), (c) and (d) for the model as below

$$x_t = 1.5x_{t-1}$$

Question 6

Suppose $x_0 = 10$ and

$$x_t = 0.8x_{t-1}$$

for $t = 1, 2, \ldots$.

(a) (2 points) Find x_1, x_2, x_3, x_4.

(b) (2 points) Find a formula for x_t.

(c) (1 point) Find $\lim_{t\to\infty} x_t$.

(d) (3 points) Sketch a plot of x_t vs. x_{t-1} for $t = 1, 2, 3, \ldots, 10$.

(e) (1 point) Sketch a plot of x_t vs. x_{t-2} for $t = 2, 3, 4, \ldots, 10$.

(f) (8 points) Repeat (a), (b), (d), (e)for

$$x_t = 0.8x_{t-1} + z_t$$

where $z_1, \ldots, z_{10}$ take on values
-1.2, 0.2, -1.0, 0.5, 1.7, -0.5, -2.1, 1.0, 0.8, -0.1

(g) (2 points) Repeat(d), (e) for

$$x_t = z_t$$

(h) (8 points) Repeat(a), (b), (d), (e) for

$$x_t = z_t - 0.8z_{t-1}, \quad \text{t=2, 3, \ldots, 10}$$

Question 7

Suppose $x_0 = 2$ and $x_1 = 1$

$$x_t = 0.8x_{t-1} - 0.7x_{t-2}$$

for $t = 2, 3, 4, \ldots$.

(a) (2 points) Find x_2, x_3, x_4.

(b) (3 points) Sketch a plot of x_t vs. x_{t-1} for $t = 2, 3, 4, \ldots, 10$.

(c) (1 point) Sketch a plot of x_t vs. x_{t-2} for $t = 2, 3, 4, \ldots, 10$.

(d) (6 points) Repeat (a), (b), (c) parts for

$$x_t = 0.8x_{t-1} - 0.7x_{t-2} + z_t$$

where $z_2, \ldots, z_{11}$ take on values
-1.2, 0.2, -1.0, 0.5, 1.7, -0.5, -2.1, 1.0, 0.8, -0.1.

Question 8

Suppose $\{z_t\}$ is a sequence of independent normal random variables with mean 0 and variance 1 (NID(0,1)) for $t = 0, 1, 2, \ldots$. Suppose also that $x_0 = 0$ and

$$x_t = .8x_{t-1} + z_t$$

(a) (2 points) What is the distribution of x_1? Specifiy $E[x_1]$ amd $Var(x_1)$

(b) (2 points) What is the distribution of x_2? Specifiy $E[x_2]$ amd $Var(x_2)$

(c) (2 points) What is the distribution of x_3? Specifiy $E[x_3]$ amd $Var(x_3)$

(d) (2 points) If x_2 takes the value 3, make a point prediction for x_3.

(e) (2 points) What is the distribution of the error for your prediction?

Question 9

(a) (1 point) Write down the AR(1) model with mean 0 in matrix and backshift matrix form.

(b) (2 points) Can any AR(1) processes express in term of ϵ_t? Explain clearly.