Stat 443 - Fall 2017

Assignment #1

(due Tuesday, October 3, at 1:00 pm)

1. In this problem you are asked to generate and plot 5 paths for each of the three processes described below (the notation is the same as in class). You may assume that the number of time steps for each process is $N = 200$.

(i) Random walk with $Z_1 \sim N(0, \sigma^2 = 0.2)$ and $X_0 = 0$.

(ii) MA(1) process with $Z_1 \sim N(0, \sigma^2 = 0.2)$ and $\theta = 0.7$. To start your iteration, for each path you may use

$$X_1 = Z_1, \quad \text{that is} \quad Z_0 = 0.$$

(iii) Process $Y_1, Y_2, \ldots, Y_{200}$ defined as

$$Y_t = 10\sqrt{t} + X_t, \ t = 1, 2, \ldots, 200,$$

where X_t, $t = 1, 2, \ldots$, follow an AR(1) process with $\phi = -0.5$, $\sigma^2 = 0.2$, and $X_0 = 0$.

For each process present:

(1) - a single graph that depicts all five simulated paths,

(2) - the R code that you used to simulate the paths,

(3) - a brief comment (2-3 sentences) as to whether or not the paths show "stationary" behavior over time (when plotting a path use lines connecting the generated points; this will help you see better the shape of the path).

2. Suppose that random variables X and Y are defined in the following way:

$$\begin{aligned} X &= Z + \epsilon_1 \\ Y &= Z + \epsilon_2 \end{aligned}$$

where $Z \sim N(0, \sigma^2)$, $\sigma > 0$, $\epsilon_1 \sim N(0, 1)$, and $\epsilon_2 \sim N(0, 1)$. Assume also that Z is independent of ϵ_1 and ϵ_2, and

$$\mathrm{Corr}(\epsilon_1, \epsilon_2) = \rho,$$

where $\rho \in [-1, 1]$.

(i) Find the correlation coefficient between X and Y.

(ii) Can you find values of ρ and σ for which X and Y are perfectly correlated? Briefly explain your answer.

(iii) Can you find values of ρ and σ for which X and Y are independent? Briefly explain your answer.

3. The density of a random variable Y conditioned on $X = x$ and the marginal density f_X of the variable X are given respectively by

$$f_{Y|X}(y|x) = xe^{-xy}, \ y > 0, \qquad \text{and} \ \ f_X(x) = 1 \ \ \text{if } x \in (0, 1), \ \text{and } 0 \text{ otherwise.}$$

Find the best nonlinear predictor of Y in terms of X, and calculate the corresponding minimum mean-square error for your predictor.

4. Obtain the historical data **hist_y.txt**, which consists of 200 consecutive observations $Y_1, \ldots, Y_{200}$ from a constant mean model

$$Y_t = \beta_0 + \epsilon_t, \qquad \text{where } \epsilon_t \sim \text{ OLSA.}$$

Let $D_n = \{Y_1, \ldots, Y_n\}$, where $n = 1, 2, \ldots, 200$. We know that if D_n is the only data available to us then the least squares estimate of β_0 is given by $\beta_0^*(n) = \bar{Y}_n$, the sample mean.

(i) Plot $\beta_0^*(n)$ vs. n for $n = 1, 2, \ldots, 200$. To what value does $\beta_0^*(n)$ appear to be converging? What value does this represent in the constant mean model?

(ii) Provide a plot of $\beta_0^*(n)$ vs. n superimposed on a plot of Y_n vs. n. Compare the variability of the curves. At what point does the updating of $\beta_0^*(n)$ appear to become almost negligible here?
Hint. You may create the vector $\beta_0^*(n)$ by using the following loop:

```
beta0 <- 0*c(1:200)
for (i in c(1:200)) {
  beta0[i] <- mean(y[c(1:i)])
}
```