Exam #1: Time Series Analysis

Fall Semester

Instructions. Follow all instructions below. If you have any questions, please ask me!

1. **Download the newest version** of `STA5362-functions.R` and source it.

2. This is an open-book, open-note exam. You may use materials from class notes, homework, lecture slides, R code, and the textbook. You may **NOT** use the internet, AI, or any other book, discuss the exam with anyone (other than Dr. Harvill) or use any other external source for the course. Ask me if you need clarification on whether a source is legitimate for this exam!

3. If you decide to hand-write your solutions (as opposed to using LaTeX to write them), that is fine. *However*, if I cannot read your handwriting, I will not grade the problem. So, your writing should be neat, legible, and dark enough that the scanned work is not too light to read. LaTeX to write them) that is fine. *However*, if I cannot read your handwriting, I will not grade the problem. So your writing should be neat and legible and dark enough that the scanned work is not too light to read.

4. You should turn in all problems in the order they're presented on this exam. Save your work in a PDF (Adobe) format and upload it to the `Exam 1` folder in our shared *Box* drive. If you have technical issues, please inform me as soon as possible so we can correct the problem or devise an alternative solution.

5. Partial credit can only be assigned if you show all your work. Don't leave steps out. Because of this, it is advisable, but not required, to include your *R* code for any problem requiring computations or graphs.

6. **ALL** graphs should be created using *R*.

7. This exam can be started any time on Friday, November 3rd. But it must be turned in by 11:59 PM that same day.

1. Consider the process $\{X_t,\ t \in \mathbb{Z}\}$, where

$$X_t = \alpha W_{t-2} + W_t.$$

where W_t is a white noise process with variance $\sigma_W^2 < \infty$. Prove that X_t is second-order stationary. Specifically, start with first principles (the definition) of second-order stationary and go through the proof of the result. Do **not** say something like, "On homework #4, we showed ... so." Do the proof like it is the first time you've ever done the proof. You can use properties of white noise without proving them.

2. Consider the two time series realizations in Figure 1. The data for each time plot can be downloaded from Canvas. The data the plot on the left is in the file `Exam1No2a`. For the plot on the right, the file is `Exam1No2b`. Both are available in plain text (files ending with `.txt`) and as an *R* data file (file ending in `.RData`). Both realizations are

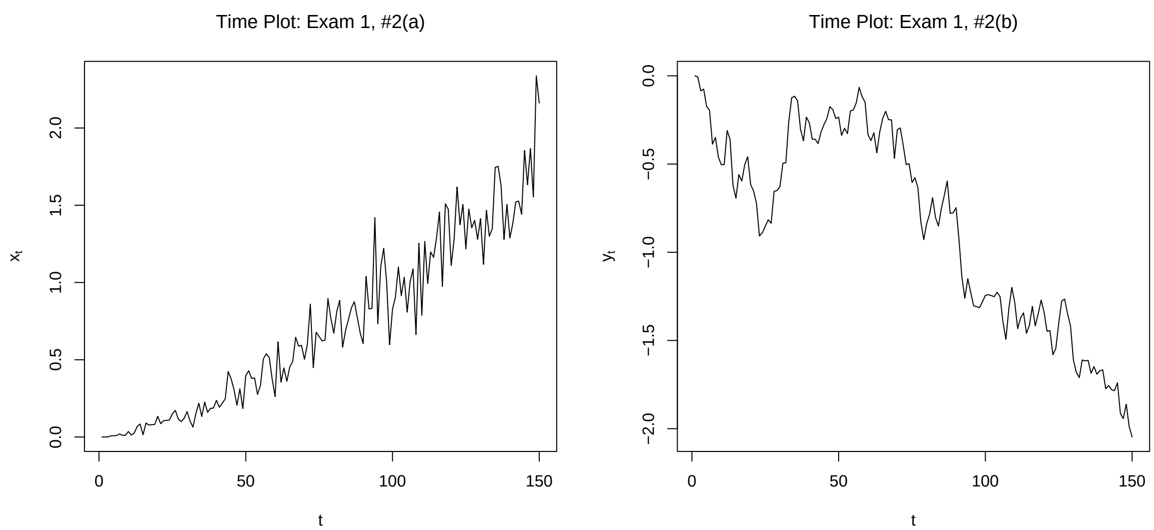

Figure 1: Time plots of series for Question #2.

from different models. For each, provide a discussion of whether or not the data seems to be from a stationary model. If you conclude the model is not stationary, explain why. If you conclude nonstationarity, explain which transformations you would make and why. Use *R* to make those transformations, create the plots of the transformed data and include them with your work. Can you conclude the transformed data is white? Include any additional graphs to support your conclusion.

3. Let the process $\{Y_t\}$ be given by

$$Y_t = a_0 + a_1 t + a_2 t^2 + W_t, \quad t = 1, 2, \ldots, n,$$

where W_t is white noise with variance $\sigma_W^2 < \infty$. Prove taking the first difference Y_t of Y_t two times removes any nonstationarity.

4. Consider the series V_t using

$$V_t = 10\cos\{2\pi(t-1)\omega_1\} + 2.5\cos\{2\pi(t-1)\omega_2\} + W_t, \quad t = 1, 2, \ldots, 200,$$

where $W_t \sim \mathrm{WN}(\sigma_W^2 = 1.5)$, $\omega_1 = 0.1$ and $\omega_2 = 0.25$.

(a) Use properties of models that are sinusoids to find the theoretical variance of the process without the white noise. Show your work. What is the variance of V_t when the white noise is included?

(b) Generate the skeleton (non-random) part of the process. Create time plot, and plots of the sample autocorrelation function and standardized periodogram. Describe the three plots. Numerically compute the variance of the data you generated and report that. Does it match the theoretical variance in part (a)?

(c) Add the white noise and repeat part (b). Explain the similarities and differences. f

5. The data set `co2` in (base) *R* contains monthly atmospheric concentrations of CO_2 expressed in parts per million (ppm) from January 1959 through December of 1997 over Mauna Loa, a volcano on the island of Hawaii. The data are maintained and updated by Scripps Institution of Oceanography. **WARNING:** The data `co2` is *not* the same thing as the data `CO2`.

(a) Create and describe a time plot of the data. Describe any features that indicate nonstationarity.

(b) Remove any noticeable trend using a polynomial regression, and save the residuals. Report any findings regarding the model.

(c) Create and describe a time plot of the residuals from the fit in part (b). Are there any remaining features that indicate nonstationarity?

(d) Use the periodogram to find important frequencies (no more than five) that may be present in the residuals and estimate the amplitudes corresponding to those frequencies.

(e) The *R* function `decompose` takes a time series, and decomposes it into a seasonal part, a trend, and any irregular components (random) using moving average smoother. In *R*, run the code below.

```
co2.d <- decompose(co2)
plot(co2.d)
co2.s <- co2.d$seasonal   # seasonal part of the series
co2.r <- co2.d$random     # random component of the series
```

Create and plot the periodogram for the seasonal part of the series. Use that to find important frequencies (as many as indicated) and estimate the amplitudes for those frequencies. How well does your answer here agree with your answer in part (d)? It's OK if they don't match. Discuss the similarities and differences in the results from the two methods.

(f) Finally, create a plot of the sample autocorrelation function of the random part of the results from part (e). What does the autocorrelation function indicate?

(g) Discuss your findings regarding the steps you took above regarding the carbon dioxide levels above Mauna Loa.

Note: The `co2` data set is a sneaky hard. I am not expecting perfection or ground-breaking analysis. I only want to see if you can use the stuff we've learned.

6. **Matching.** For Figure 2 on page 5, match the time plots on the left side to the autocorrelation functions.

7. **Matching.** For Figure 3 on page 6, match the time plots on the right side to the spectral density functions on the right side.

Figure 2 for Problem #6.

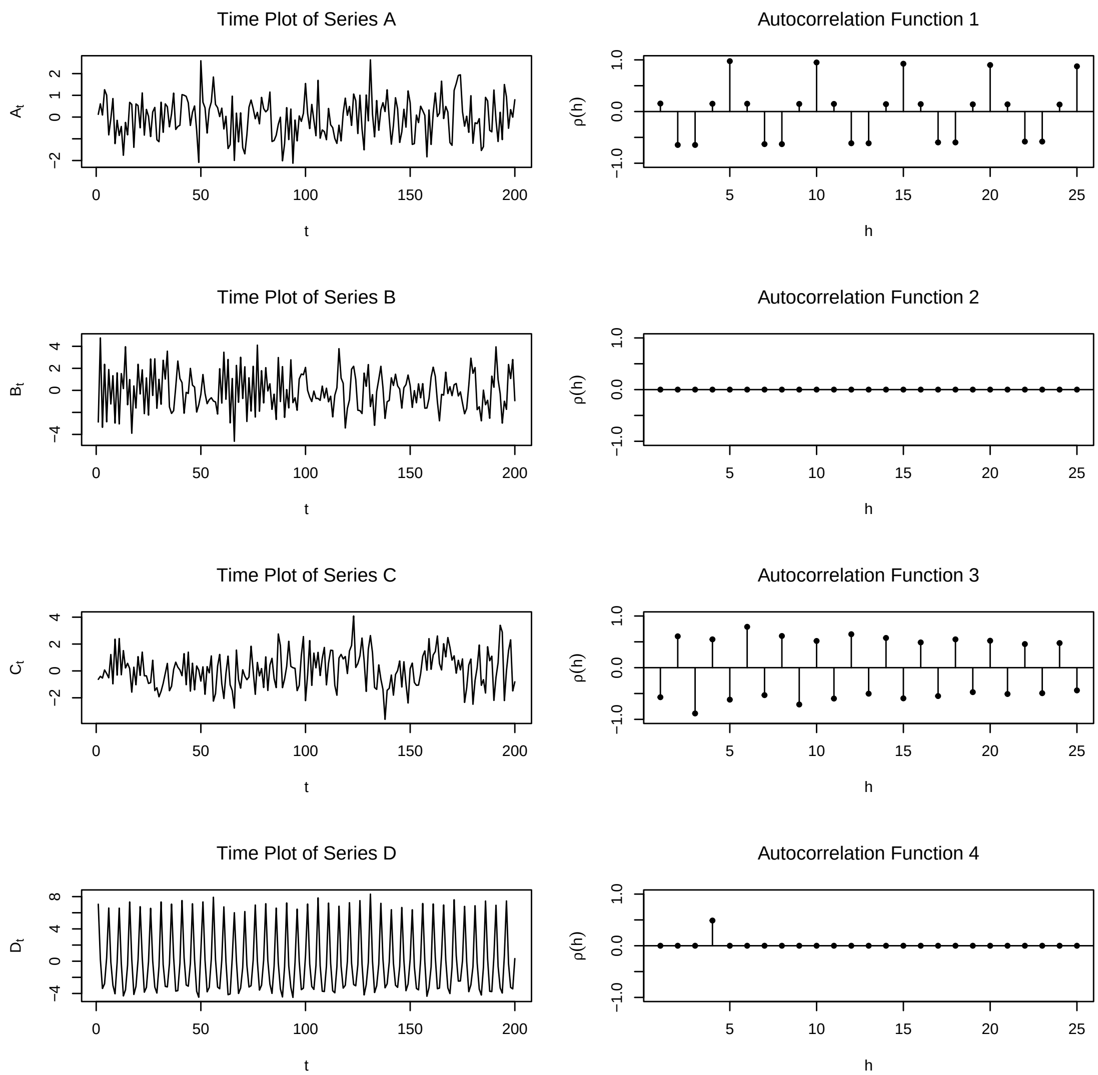

Figure 2: Time plots (left) and autocorrelation functions (right) for Problem #6.

Figure 3 for Problem #7.

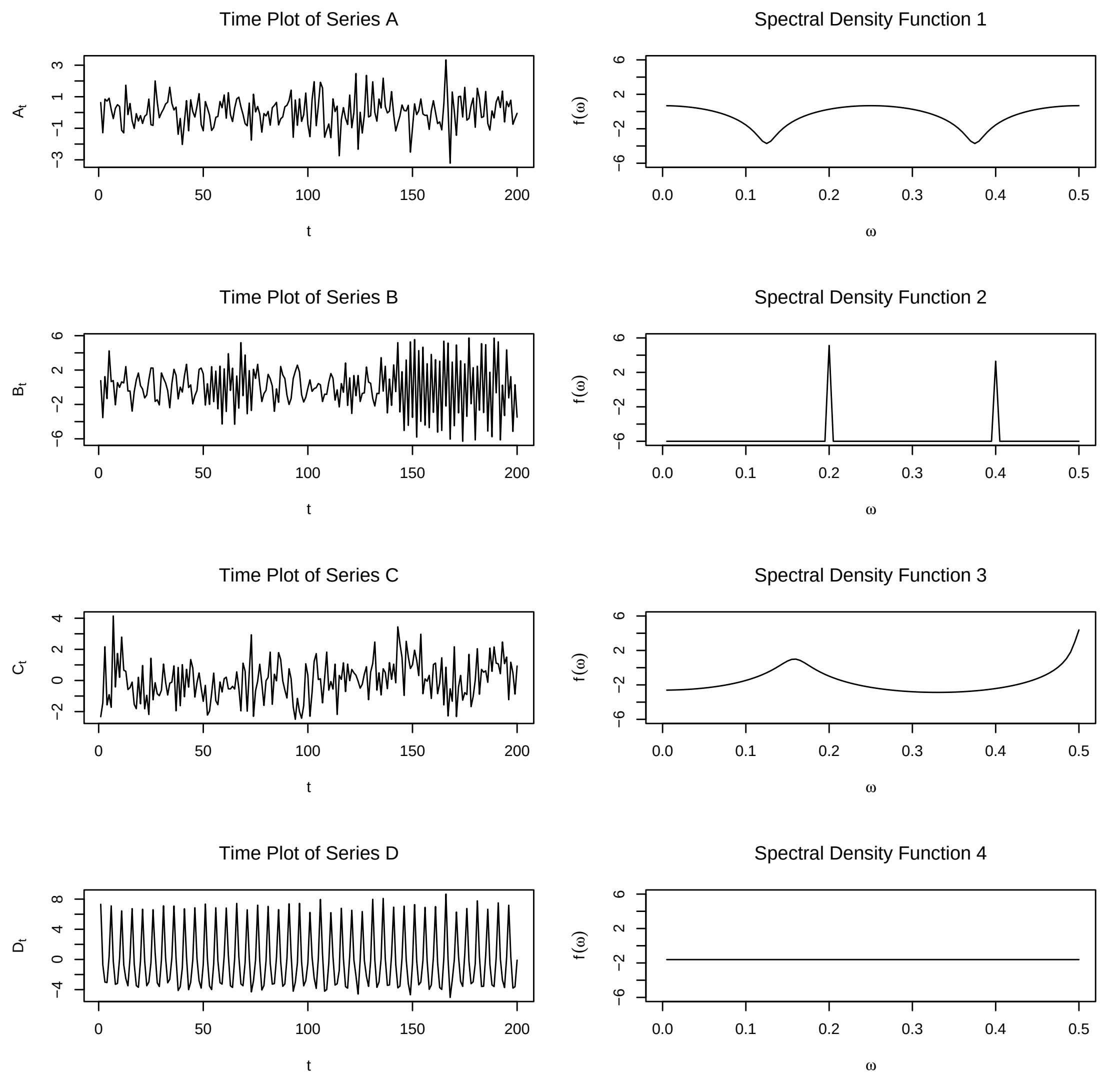

Figure 3: Time plots (left) and spectral density functions (right) for Problem #7.