Room 1: Terrace

STAT 464/864 - Fall 2024

Discrete Time Series Analysis | Queen's University

Deadline for Original Problems: Friday, September 27th (midnight)

Galaxy Grading (Modified from *Specifications Grading*, Linda Nilson, 2014)

Instructions

This *Room* (assignment) has 6 *Stars* (problems). All are worth 5 points.

Choose any 4 to complete and submit. These are your **Original Stars.**
The remaining 2 are your **Alternate Stars.** You do not have to submit these.

Once you receive your grade for this *Room*, you can try for a higher score: Alternate Stars may be attempted *after* the Room's due date and, if favourable to your grade, may replace up to two of your Original Stars. All Alternate Stars are "due" December 2nd.

Graduate students must complete 4 stars, and also 1 **comet.** Yeah, there are 2 comets.
Choose 1 Original Comet to submit — the other will be your Alternate Comet.

Fine Print

1. Alternate Stars may only replace Original Stars **from the same homework "Room."** This ensures that your improved grade reflects an improved understanding of the learning outcomes specific to that Room.
2. Only 2 Alternate Stars permitted per Room. (Ex: If you submit 3/4 Original Stars by the Room's due date, you can't use the missing Original Star as a 3rd Alternate Star.)
3. Comets (gradate problems) can be treated as Stars, but the converse is not true.

Learning Outcomes

1. Estimate trend and seasonal components of an observed time series, using R
2. Identify strong stationarity, weak stationarity, and non-stationarity
3. Determine the ACVF and/or ACF of a stationary series
4. **[GRAD]** Derive limits (as $N \to \infty$) of sample ACFs for deterministic time-functions.

Formatting

Your submission **must** be a rendered Quarto document, otherwise it will not be graded. If you are having troubles rendering, please come to my office hours! I will help you!

R-based responses must use code chunks to execute code and display results. All written discussions must be typed in-document (for example: if you're asked to comment on the implications of a certain result, or to compare the performance of different models). See Workshop 1 for how to create a code chunk.

But you don't have to type out your *mathematical* proofs. You may write math formulae by hand, save the image as a **.png** in your **working directory** (the folder containing your **.qmd** file) and insert it into your document. The syntax to insert an image is:

`![Caption goes here](image.png){width=100%}`

"But Skye," you say, "isn't this just extra work? Since we upload it all to Crowdmark as individual pages, anyway?"

"Yes." I say. "But it helps ensure all your submissions are right-side-up, in order, etc. AND it gives a buffer of white space for the TA to leave comments. AND, it combines all the pieces of your submission and puts them in one place. AND this entire process is a good professional skill to make a habit of - it's built into RStudio because it's standard practice." Mic drop.

Beyond all this, there are a few basic expectations:

1. Don't print a bunch of unnecessary code. **Look** at the rendered pdf document. Is it 30 pages long? Are there hundreds of lines printing data values for no reason? In general, don't print out data. If you really want, you can give a preview of the data by using the command: `head(my_data)`.
2. Label your plots. Give each plot a relevant main title. The `x` axis will usually be time, in this course, but you should also indicate units and/or dates.
3. Organize your work using headers (hashtags). Example of syntax and output below.

   # Problem X

   #### Part y)
   Blablabla your answer here

Problem X

Part y)

Blablabla your answer here

Star 1: Carbon Dioxide Galaxy [R Problem]

Data	Global average marine surface CO_2 (parts per billion)
Times Sampled	(Monthly) January, 1980 – May, 2024
Source	NOAA (National Oceanic & Atmospheric Administration)

Load the dataset `co2.txt` into Rstudio. I recommend the code

```
# you can call the data variable [dat] whatever you want, of course
dat <- read.table("co2.txt", header = TRUE)
```

Part a) *Plotting*

Create a line plot of the time series, demonstrating the evolution of average marine surface CO_2 levels over time. **Comment** (as text in your Quarto doc) on any apparent trend and/or periodicity in the time series.

Plotting Etiquette

For full marks, hit the following criteria:

1. **Make it a line-plot.**
   This helps highlight the actual trajectory of the points.
   (You can see this in action if you try plotting a point-plot of the same data, and then compare the two plots. Which one better shows the trends we might be interested in? Not for marks, just food for thought.)
2. **Create an appropriate x-axis.**
   The time-units should be in *months,* however, the tick marks on the x-axis should correspond to *years.* (One tick every 12 observations.)
3. **Use scientifically meaningful labels:**
   Main title. x-axis. y-axis. Don't forget units!
4. **Don't panic:**
   Look, I know I'm really particular about the way I do things. But for YOU: all I care about are *learning outcomes.* As long as your plot is able to serve its scientific purpose (which does require conveying details such as units and point trajectories), you've completed the mission. Beyond this, feel free to take risks, write your own functions (unless I say otherwise), use shortcuts, etc.