Room 2: Fountain

STAT 464/864 ∼ Fall 2024

Discrete Time Series Analysis | *Skye P. Griffith* ∼ Queen's University

Deadline for Original Problems: Friday, October 11th (midnight)

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

You need 4 Stars to complete this Room. Choose 4 Original Stars to submit.
The other 2 will be your Alternate Stars, which you may complete later to boost your score.
Alternate Stars may only replace Original Stars from the same "Room."

Graduate students must also complete 1 **comet.** Choose 1 Original Comet to submit.
The other will be your Alternate Comet. The same rules apply.

Formatting

Your submission **must** be a rendered Quarto document, otherwise it will not be graded.

On your computer, make a folder called `464_Room2` (or `864_Room2`, obviously.)
Create a Quarto file called `464_Room2.qmd` file in that folder.
Store all components of this Room in that folder, including datasets and image files.

- **R-code** must be typed and executed in "chunks".
- **Commentary** must be typed in-document.
- **Mathematical Proofs** may be typed, or hand-written and inserted into your Quarto document. The syntax to do so is: `![Caption](image.png){width=100%}`
- **Use headers** (hashtags) to organize your work. See lines 22-34 of W1_solutions.qmd.

Learning Outcomes

1. [Review] Trend Elimination using R; Evaluate whether residuals are white noise

2. Estimate AR(1) coefficient ϕ using linear regression in R,
   plot and interpret theoretical ACF defined by $\hat{\phi}$

3. Calculate theoretical ACVFs/ACFs of MA(1) and AR(2) processes,
   show these processes are stationary (or 1-correlated)

4. Understand how linear filters can pass polynomial trends without distortion.
   Prove and/or apply this phenomenon's constraints on filter coefficients.

Star 1: DOW Jones Index Galaxy

Data	Daily closing price of the Dow Jones Index (billion??? Dollars????)
Times Sampled	January 1, 2010 − December 31, 2015 (weird time intervals) for modelling purposes, treat Δt as being regular.

Data:

Download the dataset dow.csv from Github, and save it where? That Room 2 folder you just made. Make sure the data and the .qmd file are both in there. While you're at it, make sure the .png files for your handwritten Stars are in there too. Now LOAD the data!

```
dow <- read.csv("dow.csv", header = TRUE)
```

Plotting Etiquette

1. Line-plots. Not just points.
2. Create scientifically meaningful labels.
3. Use colour to distinguish different lines on the same plot.

Planet a) *Plotting*

Create a line plot of the time series. The x-axis is tricky. Place one tick at the first observed day of each year − this isn't always January 1st, as some days are missing from the dataset. Place a final tick at January 1st, 2016: the day after the data's last observation.

Comment: Is there any apparent trend or pattern in the data? What do you see?

Planet b) *Trend Elimination*

1. **Polynomial Regression [1 plot]:**
   Estimate the data's trend using polynomial regression with degree $p = 1$.
   Remember how to do that? Check out Workshop 2 if you don't.

   Create a line-plot displaying your regression line on top of the original time series.
   Don't forget your <u>plotting etiquette</u>.

2. **Residuals [2 plots]:**
   Compute the residuals $\{r_t\}$ and plot them against time.
   Next, create a separate plot displaying their sample ACF.

   Comment: Do you think the vector of residuals is a white noise process?
   How does the residual plot influence your response? What about the ACF plot?

Planet c)   *Autoregression*

1. **Estimating ϕ [1 plot]:**
   Plot a scatterplot of r_t versus r_{t-1} for $t = 2, \dots, N$.

   Fit the line $r_t = \phi\, r_{t-1}$ and plot it on top of the scatterplot.
   (you can use the R function `lm()` to do this, just don't include an intercept).

   Comment: What is the estimated AR coefficient $\hat{\phi}$?

2. **Comparing ACFs [1 plot]:**
   Plug your estimate of $\hat{\phi}$ into the formula for the theoretical ACF of an AR(1) process.
   Plot the resulting ACF as a line-and-point plot using the parameter `type = "o"`.

   On the same plot, add the sample ACF of the planet b) residuals $\{r_t\}$.
   Be sure to use colour to distinguish the two ACFs. Use `type = "o"` for both.

   Comment: How well did the structure of your estimated AR(1)-$\hat{\phi}$ process match
   that of the polynomial regression residuals $\{r_t\}$?

Star 2: Simulated Galaxy

Planet a)   *Simulation*

1. **Data:**
   Generate a time series $\{x_t\}_{t=1}^{300}$ realizing AR(1) model

$$X_t = 0.7 X_{t-1} + Z_t \qquad (\spadesuit)$$

$$Z_t \sim wn(\sigma^2 = 2)$$

   To simulate ($\spadesuit$), you can use the `specify()` function followed by `sim()`.

2. **Visualizing ($\spadesuit$) [2 plots]:**
   Create a plot of your simulated time series $\{x_t\}$.
   Create a separate plot displaying the sample ACF of this simulation.

Planet b)   *Autoregression*

1. **Estimating ϕ:**
   Estimate the AR coefficient ϕ by fitting the linear regression model

$$
\begin{bmatrix} x_2 \\ x_3 \\ \vdots \\ x_{300} \end{bmatrix} = \phi \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{299} \end{bmatrix} \qquad (\clubsuit)
$$

   You can use the `lm()` function to do this. Call the fitted coefficient $\hat{\phi}$.

2. **Visualizing ($\clubsuit$) [1 plot]:**
   Create a scatterplot of the response against the predictor for model ($\clubsuit$).
   Add the regression line defined by $\hat{\phi}$.

Planet c)   *Autocorrelation*

1. **Compare ACFs [1 plot]:**
   Plug your estimate of $\hat{\phi}$ into the formula for the theoretical ACF of an AR(1) process.
   Plot the resulting ACF as a line-and-point plot using the parameter `type = "o"`.

   On the same plot, add the sample ACF you computed on planet a).
   Finally, add the theoretical ACF for the true ($\spadesuit$) model ($\phi = 0.7$)

   Be sure to use different colours to distinguish the three ACFs.
   Use `type = "o"` for the estimated ACFs, and `type = "h"` for the true ACF.

   Comment: Which estimate $-$ the sample ACF or the ACF derived from $\hat{\phi}$ $-$ is more accurate to the true ACF? Why do you think that is?

Star 3: (MA) Moving Average Galaxy

For $t \in \mathbb{Z}$, define the following time series:

$$Z_t \sim wn(\sigma^2)$$

$$Y_t = \mu Z_t + Z_{t-1}$$

$$X_t = \frac{1}{2q+1} \sum_{k=-q}^{q} Y_{t-k}$$

where $q \geq 1$ is a fixed positive integer. Find the ACF of $\{X_t : t \in \mathbb{Z}\}$.

Star 4: (AR) Autoregressive Galaxy

For $t \in \mathbb{Z}$, define the following time series:

$$Z_t \sim wn(\sigma^2 = 1)$$
$$Z_t' \sim wn(\sigma^2 = 1)$$

$$X_t = 0.5 X_{t-1} + Z_t$$
$$Y_t = X_t + Z_t'$$

Let $\text{Cov}[Z_s Z_t'] = 0 \quad \forall\, s, t \in \mathbb{Z}$.

Planet a)

Show that $\{Y_t\}$ is weakly stationary and find its ACVF.

Planet b)

Show that the time series $W_t = Y_t - 0.5 Y_{t-1}$ is weakly stationary and 1-correlated. (That is, $\gamma_W(h) = 0$ for $|h| > 1$, where $\gamma_W(h)$ is the ACVF of $\{W_t\}$).

Star 5: Theoretical Linear Filter Galaxy

Show that a linear filter $\{b_j\}$ passes an arbitrary polynomial of degree p without distortion,

$$\text{i.e.,} \quad q_t = \sum_{j=-\infty}^{\infty} b_j q_{t-j}$$

for all kth-degree polynomials $q_t = \alpha_0 + \alpha_1 t + \cdots + \alpha_p t^p$ (where $\{\alpha_i\} \subset \mathbb{R}$) if and only if

$$1) \quad \sum_{j \in \mathbb{Z}} b_j = 1$$

$$2) \quad \sum_{j \in \mathbb{Z}} j^k b_j = 0 \quad \text{for } k = 1, \ldots, p.$$

Hint:
With $\{q_t\}_{t \in \mathbb{Z}}$ denoting the polynomial of degree p, use the binomial theorem to write the filtered version of $\{q_t\}$ as

$$\sum_{j \in \mathbb{Z}} b_j q_{t-j} = q_t \sum_{j \in \mathbb{Z}} b_j + \sum_{k=1}^{p} g_k(t) \sum_{j \in \mathbb{Z}} j^k b_j$$

for some functions $g_k(t)$, $k = 1, \ldots, p$, that depend only on t.

Star 6: Applied Linear Filter Galaxy

Find a linear filter of the form $1 + \alpha B + \beta B^2 + \gamma B^3$ (i.e., find α, β, and γ) that passes linear trends without distortion AND eliminates arbitrary seasonal components of period 2.

You may use the results of Star 5 for conditions to pass a linear trend.

Comet 1: Filter Swap Galaxy [GRAD Problem]

Suppose that $\{X_t\}$ is the MA(1) process

$$X_t = Z_t + \theta Z_{t-1}, \qquad \{Z_t\} \sim wn(0, \sigma^2),$$

where $|\theta| > 1$. Define a new process

$$W_t = \sum_{j=0}^{\infty} (-\theta)^{-j} X_{t-j}$$

and show that $\{W_t\} \sim \text{WN}(0, \sigma_W^2)$. Express σ_W^2 in terms of θ and σ^2 and show that $\{X_t\}$ can be represented in terms of $\{W_t\}$ as follows,

$$X_t = W_t + \frac{1}{\theta} W_{t-1}.$$

Comet 2: Causal Galaxy [GRAD Problem]

Let $\{X_t\}_{t \in \mathbb{Z}}$ denote the unique stationary solution of the autoregressive equations

$$X_t = \phi X_{t-1} + Z_t,$$
$$\text{where} \quad Z_t \sim wn(\sigma^2),$$
$$\text{and} \quad |\phi| > 1.$$

Then Xt is given by the expression (2.2.11) in the text.

Define the new sequence

$$W_t = X_t - \frac{1}{\phi} X_{t-1}.$$

1. Show that $\{W_t\}$ is a white noise process with mean zero and variance σ_W^2.
2. Express σ_W^2 in terms of σ^2 and ϕ.

This shows that $\{X_t\}$ is the (unique stationary) solution of the causal AR equations

$$X_t = \frac{1}{\phi} X_{t-1} + W_t \qquad t = 0, \pm 1, \ldots$$