Test #1

STAT 878

Spring 2024

Complete the problems below. Make sure to fully explain all answers and show your work to receive full credit. Use R for all calculations and plots.

1. (31 total points) Consider the following ARMA(1,1) model:

(1 – ϕB)xt = (1 + θB)wt

where μ = 0. Complete the following parts.

* 1. (12 points) Rewrite the model as an infinite order moving average model so that the only parameters are ϕ and θ in the final model given. To save time, you may use “…” in your answer for ψ4 and higher index values. Please take advantage of using the ϕ, θ, and ψ symbols given here for your derivation.
	2. (12 points) We defined a linear process in the notes as

 with  and wt ~ independent N(0,)

for a stationary process. It can be shown that  for h ≥ 0 and γ(h) = γ(-h). Using this expression for the autocovariance, find γ(0) for the ARMA(1,1) model.

* 1. (7 points) Suppose ϕ = -0.4 and θ = 0.8. Plot the ACF for lags 1, …, 20. Rather than coding your answer for b), I recommend using the R function that directly calculates the ACF for any ARMA model!

1. (31 total points) The National Oceanic and Atmospheric Administration (NOAA) monitors the amount of ice cover on the Great Lakes each year. This is done for a number of reasons, such as tracking the effects of global warming and examining the availability of shipping lanes. The file Superior.csv (available from the graded web page) provides the maximum ice cover percentage per year for Lake Superior (one of the lakes in the Great Lakes). Below is how I read the data into R and examine the first six observations.

> superior.ice <- read.csv(file = "Superior.csv")

> head(superior.ice)

 year ice

1 1973 69.8

2 1974 73.8

3 1975 64.9

4 1976 49.9

5 1977 96.0

6 1978 92.5

x <- superior.ice$ice

For example, the maximum ice cover percentage in 1973 was 69.8%. Using these data, complete the following.

* 1. (8 points) Plot the maximum ice cover percentage vs. time. Is there evidence against stationarity based upon this plot? Explain.
	2. (9 points) Plot the estimated ACF and PACF. Describe the dependence that exists in the data by referencing the plots.

* 1. (7 points) Estimate and state an ARIMA(1,0,1) model for these data. Make sure to include the estimate of α in your answer!
	2. (7 points) What method functions in R are available for objects created by arima()?
1. (24 total points) Determine an appropriate ARMA model for each set of estimated ACF and PACF plots below. Each ACF and PACF plot starts at lag 1 (not 0). Justify all of your answers using the appropriate terminology!
	1.





* 1.



* 1.



1. (14 total points) Answer the questions below.
	1. (7 points) Suppose n observations are simulated from an ARIMA(2,0,0) model. The estimated ACF and the true (actual) ACF are plotted, and there are some differences in their appearance (e.g., an estimated autocorrelation at lag h is positive and a true autocorrelation at lag h is negative). What is the most likely reason for these differences? What could one change in the data simulation to make these ACFs become closer in agreement?
	2. (7 points) What is the main potential benefit of first differencing?