# Laboratory #4 Questions and Discussion Part 2 – ARIMA Models

You have a lot of information and examples for ARIMA models already. I’ll share one more. This is from Dr. Robert Nau at Duke University. Dr. Nau is now a professor emeritus at Duke. He has put a lot of very valuable information online. The home page for Forecasting is at <https://people.duke.edu/~rnau/411home.htm>. There is a very good summary of how to choose the “right” model, i.e. a regression model, at ets() model, an ARIMA model or what??? This is a different question than choosing the most accurate model.

In addition to discussing how to select the “right” model, under ARIMA models it tells us what the nomenclature is that we get from the forecast package in R, That is, in our textbook in Section 9.9,

*ARIMA(p,d,q) are ARIMA models for non-seasonal data. “p” is the order of the autoregressive terms, “d” is the number of differences, and “q” is the order of the moving average, terms in the model.*

*ARIMA(p,d,q)(P,D,Q), where D, P, and Q are, respectively, the number of seasonal differences, seasonal autoregressive terms (lags of the differenced series at multiples of the seasonal period), and seasonal moving average terms (lags of the forecast errors at multiples of the seasonal period)* (Nau, 2020)*.*

Later, when you are finishing forecasting the Amtrak data the first time you’ll find that the forecast package in R produces an ARIMA model as

Series: Amtrak.ts

ARIMA(0,1,3)(2,1,1)[12]

Here we have p,d,q equal to 0,1,3 and “D” the number of seasonal differences are 1 (one), “P” the seasonal autoregressive terms (lags of the differenced series at multiples of the seasonal period) are 2 (two), and “Q” the seasonal moving average terms (lags of the forecast errors at multiples of the seasonal period) are 1 (one). The frequency of the data is 12 or monthly.

Keep in mind that AR(p) models are similar to linear regression models except… the predictors are the past values of the (time) series. For example, an AR(1) model, or an AR model of order 1, will have one term for past predictors. An AR(2) model, or AR model of order 2, will have two terms for past predictors. These equations are:

A MA(q) model uses values of past forecast errors. The general equation for an MA(q) model is:

You can actually use the arima.sim() command in the fpp3 package to simulate the AR(p) and MA(q) processes. That’s a really instructional thing to do!

Besides the order of the model built and its accuracy, output from the code also a table of coefficients for the non-seasonal and seasonal terms in the equations for the model. For example, coefficients for the non-seasonal terms are shown as: ar1, ar2, …, ma1, ma2, …. And coefficients for the seasonal terms are shown as: sar1, sar2, …, sma1, sma2, …. You are also given the standard error (se) as well as other information about the model.

Another area that can cause considerable frustration for students are unit root tests which include the ADF and KPSS tests. If you’ve had linear algebra you may recall that a “unit root” has a specific connotation. There is a short blog at <https://www.learnsignal.com/blog/unit-root/> that has a pretty good explanation of why a unit root is important for us. Essentially, a unit root test is a statistical test provides an indication of whether or not differencing is required for data to be stationary. The Augmented Dickey-Fuller test is probably the most popular test. It is essentially a t-test which has a null hypothesis that the data are non-stationary, i.e. there exists a unit root at some lags. The Kwiatkowski–Phillips–Schmidt–Shin test or KPSS test assumes a “trend-stationary” process and so its null hypothesis is that the data are stationary. Again, there are nuances to this. There are also more unit root tests. The reason that there are so many unit root tests is because they all have “low statistical power”. None of the unit root tests stands out as having more “power” than any of the others. Because they all differ, unit root tests can often produce conflicting results. Therefore, additional testing may be needed to conclusively determine if data are stationary or not. There is more information at <https://www.statisticshowto.com/unit-root/>.

Let’s jump into the questions for this lab. Beginning with Step 2 we started looking at ARIMA models. The following questions are based on the R script for Laboratory #4 and the results produced by this script for Laboratory #4 Step (or Part) 2.

1. You can use several packages to plot the ACF for the Amtrak data. Using the original Amtrak data processed only to change the Date variable from character format to date format. And, using any of the packages to compute and plot the ACF results in a pattern that indicates the data are \_\_\_\_\_\_\_\_\_\_ . (Note that some packages, such as fable, require additional processing the data into a tsibble data object to compute the ACF.)
   1. non-stationary
   2. periodic
   3. both a and b x
   4. neither a nor b
2. Decomposing the original data into its components reveals that the data are periodic.
   1. True x
   2. False
3. The data have a monotonically increasing trend.
   1. True
   2. False x
4. Deseasonalizing and detrending the data is intended to result in data that are \_\_\_\_\_\_\_\_\_\_\_\_\_\_ (by the observed magnitudes of the trend and seasonal components).
   1. non-stationary
   2. stationary (or quasi-stationary) x
   3. periodic
   4. cyclical
5. You only need to subtract (difference) the data once to subtract the seasonality and trend.
   1. True
   2. False x
6. You can always tell by the ACF and/or PACF that the data are stationary.
   1. True
   2. False x
7. Taking a moving average of data always results in data that are \_\_\_\_\_\_\_\_\_\_ , regardless of the period the data are averaged over.
   1. Non-stationary
   2. Stationary
   3. Non-periodic
   4. None of the above x
8. Taking a moving average of data is the same as building a MA(q) (moving average) model.
   1. True
   2. False x
9. Smoothing are data driven methods of forecasting that estimate time series components directly and are particularly useful in time series with components that change over time, i.e. a “data-driven method” rather than a “model-based” method.
   1. True x
   2. False
10. Centered moving averages cannot be used for forecasting, i.e. to make predictions, because the centered average depends on data that is not yet available.
    1. True x
    2. False
11. We have studied the Amtrak data for a while now. We know that there is a strong annual (seasonal) component in the Amtrak data. Therefore, we want to make the window width (w) at least 12 months or more long in order to preserve as much information as possible in the data.
    1. True
    2. False x
12. If you answered “False” to the last question, good for you! But why? Does a window of width 12 months or greater \_\_\_\_\_\_\_\_\_\_ .
    1. leave too much variance in the data
    2. cause the data to oscillate about some value
    3. smooths out too much variance causing lost information from the data x
    4. none of above choices
13. From the time plot of the twice-differenced Amtrak data we can see that there is still a trend and periodicity left in the data.
    1. True
    2. False x
14. Now that we have reloaded our data and added a couple more packages in R (tseries and urca), we will start building models; an AR(p) and an MA(q) model and ARIMA models. We start by considering whether the data are stationary or not using the ADF and KPSS tests. The results from the ADF and KPSS tests are initially \_\_\_\_\_\_\_\_\_\_ .
    1. Consistent that the data are stationary
    2. Consistent that the data are non-stationary
    3. Inconsistent, the ADF test indicates that the data are stationary x
    4. Inconsistent, the ADF test indicates that the data are non-stationary
15. A “unit root test” tells us \_\_\_\_\_\_\_\_\_\_ .
    1. that a unit root exists
    2. that the data are stationary or non-stationary
    3. that we need to do something, e.g. differencing, to make our data stationary
    4. all these choices x
16. What is the number of differences that the first unit root test (unitroot\_ndiffs(Amtrak.ts.91.16)) tells us we have to take for the Amtrak data to be stationary?
    1. 0
    2. 1 x
    3. 2
    4. 3
    5. 4
17. After taking 1 (one) difference the unit root tests tells us Amtrak data are \_\_\_\_\_\_\_\_\_\_ .
    1. Stationary
    2. Non-stationary
    3. Neither (the results of the unit root tests are inconsistent again) x
    4. None of the choices
18. What is the number of differences that the first unit root test (unitroot\_ndiffs(Amtrak\_seasdiff\_1)) tells us we have to take for the Amtrak data to be stationary?
    1. 0
    2. 1 x
    3. 2
    4. 3
    5. 4
19. Taking the difference of already once differenced data produces data that are \_\_\_\_\_\_\_\_\_\_ .
    1. Stationary x
    2. Non-stationary
    3. Neither (the results of the unit root tests are inconsistent again)
    4. None of the choices
20. Consider the time plot of the twice-differenced Amtrak data. Just based on that time plot does it look like \_\_\_\_\_\_\_\_\_\_ .
    1. there is no more trend in the data
    2. there is no more periodicity in the data
    3. there are no more discernible patterns in the data
    4. all of these choices x
21. The first model we will build uses the **twice-differenced Amtrak data** and the auto.arima() command from the forecast package. Using that command the model produced is an \_\_\_\_\_\_\_\_\_\_ model.
    1. AR(p)
    2. MA(q)
    3. ARMA(p,q)
    4. ARIMA(p,d,q) x
22. The output for the ARIMA() model that the code automatically builds includes the coefficients for the AR(5) model, the coefficients for the MA(4) model as well as values for the various tested accuracies. (Note that this is because the model built is an ARIMA(5,0,4) model.)
    1. True x
    2. False
23. The computed value of AICc for the auto built ARIMA(5,0,4) model is \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ . (Note that we haven’t really been using “best practices,” or rather I haven’t been when I’ve been coding these scripts. Best practices is to use the set.seed() command in order to conduct reproducible research. I have not used that. The result is that when I ran the scripts and parts of them a number of times I came up with different answers. That should not happen but can when you don’t use “best practices”. Again, caveat emptor… I haven’t entered the set.seed() command in the scripts so if you do get different or unexpected answers, I suggest that you **close RStudio, re-execute RStudio and run the script(s) again**.)
    1. 2880.97 x
    2. -5519.17
    3. not computed for this model
    4. none of these choices
24. The auto.arima() command in the forecast package starts building the model based on the \_\_\_\_\_\_\_\_\_\_\_ first (the default unless otherwise specified) to find starting values and then continues using the \_\_\_\_\_\_\_\_\_\_ . Wrong
    1. maximum likelihood, conditional sum of squares x
    2. conditional sum of squares, maximum likelihood xx
    3. sum of squares, sum of squares
    4. maximum likelihood, maximum likelihood
25. In the fable package, the ARIMA() command uses \_\_\_\_\_\_\_\_\_\_\_ to estimate an ARIMA() model.
    1. conditional sum of squares
    2. Likelihood sum of squares
    3. ordinary least squares
    4. maximum likelihood x
26. Codes used to produce essentially the same result but which use different basis for computation, e.g. the difference between starting with the conditional-sum-of-squares then maximum likelihood versus starting and continuing with maximum likelihood, can and often will produce somewhat different answers.
    1. True x
    2. False
27. When a model is built using the **original** Amtrak data and the auto.arima() command a simpler model is built, i.e. an ARIMA(0, 1, 3)(2,1,1)[12] model.
    1. True x
    2. False
28. The computed AICc for this model, the ARIMA(0,1,3)(2,1,1)[12] model is \_\_\_\_\_\_\_\_\_\_ . AIC=-5519.56 AICc=-5519.17 BIC=-5493.63 xx Wrong
29. Using this model and producing a 5-year forecast using the forecast() command yields a forecast that appears to replicate the periodicity well but is very flat, almost like a seasonal naïve forecast would be.
    1. True x
    2. False
30. Thinking about the previous question, the reason the flat forecast is almost like a seasonal naïve forecast is because a seasonal naïve forecast \_\_\_\_\_\_\_\_\_\_ .
    1. uses the prior period values for all values of the forecast
    2. uses the prior period error value for all values of the forecast
    3. uses the value from the same period in the prior season for all values of the forecast x
    4. none of these choices
31. Using the ARIMA(0, 1, 3) model built by the auto.arima() command, the ARIMA() command in fable does build an ARIMA(0, 1, 3) model. The computed AICc for the ARIMA(0,1,3) model built in fable is \_\_\_\_\_\_\_\_\_\_ . This compares well to the computed model from the forecast package. 2770
32. The 5-year forecast produced by the fable package does appear to have a slightly downward slope as the just prior periods of the Amtrak data does and has the same scale as the forecast from the forecast package.
    1. True x
    2. False

This is the end of the questions about the Amtrak data. To be thorough you should be able to split the Amtrak data into a training set and a test or validation set; and report the differences in error for each of these sets of data. By doing this you are using the model built with the training set to generate a forecast that you can check against real data using the test or validation set. You will do more of this sort of testing in ANA 540.

The remainder of the script is intended to be skills-building for you to become more adept and comfortable with running time series analyses and forecasting with real world data. These examples are taken from your textbook. Remember that I’ve said if you’ve followed the textbook and run the examples, especially in the embedded videos you will be well prepared for the final exam. I do not expect you to have all this at your fingertips for the final exam so have highlighted the correct answers for you in this section.

1. From Chapter 9 (of our textbook) Exercise 3a, find the differencing and the Box-Cox transformation that produces stationary data. 1 0.1571804
2. Does taking the log of GDP linearize the data?
   1. Yes
   2. No
3. From Chapter 9, Exercise 3b, find the differencing and the Box-Cox transformation that produces stationary data. Note that this will take both a seasonal and first differencing. 2 -0.005076712
4. The Tasmania Takings data are \_\_\_\_\_\_\_\_\_\_ .
   1. Additive
   2. Multiplicative
   3. Additive and Multiplicative
   4. None of these choices
5. From Chapter 9, Exercise 3c, find the differencing and the Box-Cox transformation that produces stationary data. Note that this will also take both a seasonal and first differencing. 2 -0.2444328
6. The sales data that are shown are \_\_\_\_\_\_\_\_\_\_ /
   1. Additive
   2. Multiplicative
   3. Additive and Multiplicative
   4. None of these choices
7. From Chapter 9, Exercise 8a, for the U.S. GDP find a suitable Box-Cox transformation for the data. 0.2819443
8. From Chapter 9, Exercise 8b, fit an ARIMA() model to the transformed data. The ARIMA() model is:
   1. ARIMA(1,1,0) with drift
   2. ARIMA(1,1,0)
   3. ARIMA(0,1,1)
   4. ARIMA(1,1,1)
9. For the last question about the ARIMA() model if you selected the answer “with drift” you were correct. The drift is important to note because it adds \_\_\_\_\_\_\_\_\_\_ to the equation for the model.
   1. an additional term of past values
   2. an additional term for the error of past values
   3. a constant
   4. nothing
10. From Chapter 9, Exercise 8d, choose which of the models you built for part 8c that is best model in terms of AICc.
    1. ARIMA(0,1,0)
    2. ARIMA(2,1,3)
    3. ARIMA(2,1,1)
    4. ARIMA(1,1,0)
11. A forecast using your fitted ARIMA() model (Exercise 8e) compared to an ets() model (Exercise 8f) shows \_\_\_\_\_\_\_\_\_\_.
    1. roughly the same magnitude growth in GDP but narrower confidence intervals for the ARIMA() model
    2. roughly the same magnitude growth in GDP but wider confidence intervals for the ARIMA() model
    3. there is no difference in the forecasts
    4. none of these choices

Now, we’ll go back to the Amtrak data and continue questions about the ACF plot, the PACF plot, the residuals, and the models we can build or are automatically built for us.

1. Now going back to the Amtrak data, first let’s anchor what we do now to the results we produced before. A plot of the Amtrak data from 1991 through 2016 shows \_\_\_\_\_\_\_\_\_\_ .
   1. the data are not linearly increasing or decreasing
   2. the data are periodic, probably with multiple periods
   3. the periodicity in the data appears to be additive (if it were multiplicative the peaks and troughs would get bigger over time)
   4. all these choices x
2. The ACF plot is the same as was produced before and shows \_\_\_\_\_\_\_\_\_\_ .
   1. data that is multiplicative in nature
   2. data with a lot of periodicity in it x
   3. data that is additive in nature
   4. none of these choices
3. A histogram of the Amtrak data shows that it is nearly \_\_\_\_\_\_\_\_\_ in distribution.
   1. right skewed
   2. left skewed
   3. normal x
   4. none of these choices
4. Let’s let the code from the forecast package automatically build a model for the Amtrak data again, i.e. use the auto.arima() command. The ARIMA model that is built is a \_\_\_\_\_\_\_\_\_\_ model.
   1. (0,0,1)(0,1,1)[12]
   2. (0,1,1)(0,1,1)]12]
   3. (0,1,2)(0,1,2)]12]
   4. (0,1,3)(2,1,1)[12] x
5. The textbook talks about “innovation residuals” in several places. Section 5.3 describes “innovation residuals” as the residuals on a transformed scale if a transformation has been applied to the data. Before we generated the figure with a plot of these “innovation residuals” we had the code automatically build an ARIMA model. The model the code built is an ARIMA \_\_\_\_\_\_\_\_\_\_ model. (As discussed at the beginning of this document, keep in mind that the automatic modeling done by auto.arima in the forecast package starts from a different point than the automatic modeling done by ARIMA() in the fable package does. The models built by these two different packages might not be the same!)

“Innovation residuals” are also discussed in Sections 5.4, 8.6, 9.9, and 10.2.

* 1. (1,0,1)(2,0,2)[12]
  2. (1,1,1)2,1,2)[12] x
  3. (0,1,3)(2,1,1)[12]
  4. none of these choices

1. In the notation used for the ARIMA models, the coefficients are given in a table for the non-seasonal components as the ar(p), ma(q), with differencing. This means that we have an AR(\_\_\_\_) and an MA(\_\_\_\_) model with \_\_\_\_\_ difference(s).
   1. 0,0,0
   2. 1,0,0
   3. 1,1,0
   4. 1,1,1 x
2. In the notation used for the ARIMA models, the coefficients are given in a table for the seasonal components as the ar(P), MA(Q), with differencing. This means that we have a seasonal component of AR(\_\_\_\_) and an MA(\_\_\_\_) model with \_\_\_\_\_ seasonal difference(s).
   1. 1,1,1
   2. 1,1,2
   3. 1,2,2
   4. 2,1,2 x
3. The frequency of the data, as given in the (output) model description is [\_\_\_\_].
   1. 1
   2. 2
   3. 12 x
   4. 0
4. Because the code applied one differencing to the data, residuals became “innovation residuals” that were plotted on the “transformed” scale. (Hint: in this case, a difference will not change the units of the scale.) Wrong
   1. True xx
   2. False x
5. The script provided for you follows this order in the analysis of the Amtrak data, prior to conducting cross-validation.
   * the original Amtrak data, its ACF, and the histogram of the original Amtrak data are plotted
   * an ARIMA model is fit to the original Amtrak data using the auto.arima() command from the forecast package
   * an ARIMA model is fit to the original Amtrak data using the ARIMA() command from the fable package
   * information from the model built from the fable package is plotted as the “Innovation residuals,” i.e. the residuals after differencing the data, the ACF of the residuals, and the histogram of the residuals
   * the original Amtrak data is plotted again, including the data, its ACF, and its PACF
   * models are built using the fable package including an ARIMA(0,1,0)(1,0,0) model and an automatically built model
   1. True x
   2. False
6. In order to conduct cross-validation the data must be split into a training set and a test set. The Amtrak data were split into a training set prior to 2009 and a test set after 2009 through 2016. (Hint to illustrate this, the training set is plotted in red and the training set in black.)
   1. True x
   2. False
7. The type of model built next is a(an) \_\_\_\_\_\_\_\_\_\_ .
   1. Naïve
   2. SNaive x
   3. ARIMA
   4. ets()
8. The Innovation residuals for this model appear to be \_\_\_\_\_\_\_\_\_\_ .
   1. “iid” (independent and identically distributed (about zero))
   2. normally distributed x
   3. perhaps some residual periodicity per the ACF
   4. all of these choices x
9. After completing the (validation) forecast, the test set has less accuracy or more error than the training set.
   1. True x
   2. False

***I have to do some more work on cross-validation for the Amtrak data. I’ll get a short message to you on that as soon as I can. In order to get all the rest of this to you quicker I’ll send this out now.***

1. The last chart produced by the script shows the linear model (lm()), the ARIMA() model, and the ets() model. These models are compared by the mean average error for twelve months. Based on this chart which model is the most accurate?
   1. lm()
   2. ARIMA()
   3. ets() x
   4. none of these choices

This is the end of Laboratory #4 Step (or Phase) 2.