GROUND RULES:

- This exam contains two parts:

  - **Part 1.** Multiple Choice (40 questions, 1 point each)
  - **Part 2.** Problems/Short Answer (10 questions, 6 points each)

  The maximum number of points on this exam is 100.

- Print your name at the top of this page in the upper right hand corner.

- **IMPORTANT:** Although not always stated, it is understood that $\{e_t\}$ is a zero mean white noise process with $\text{var}(e_t) = \sigma_e^2$. In addition, I use B to denote the backshift operator and following standard abbreviations:

  - AR = autoregressive
  - MA = moving average
  - ACF = autocorrelation function
  - PACF = partial autocorrelation function
  - EACF = extended autocorrelation function
  - MMSE = minimum mean-squared error

- This is a closed-book and closed-notes exam. You may use a calculator if you wish. Cell phones are not allowed.

- Any discussion or otherwise inappropriate communication between examinees, as well as the appearance of any unnecessary material, will be dealt with severely.

- You have 3 hours to complete this exam. GOOD LUCK!

HONOR PLEDGE FOR THIS EXAM:

After you have finished the exam, please read the following statement and sign your name below it.

I promise that I did not discuss any aspect of this exam with anyone other than the instructor, that I neither gave nor received any unauthorized assistance on this exam, and that the work presented herein is entirely my own.

PART 1: MULTIPLE CHOICE. Circle the best answer. Make sure your answer is clearly marked. Ambiguous responses will be marked wrong.

1. Which time series model assumption are you testing when you perform a **runs test**?
(a) stationarity
(b) independence
(c) normality
(d) invertibility

2. Consider an invertible MA(2) process

$$Y_t = e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2}.$$

Which statement is **true**?
(a) Its PACF can decay exponentially or in a sinusoidal manner depending on the roots of the MA characteristic polynomial.
(b) It is always stationary.
(c) Its ACF is nonzero at lags $k = 1$ and $k = 2$ and is equal to zero when $k > 2$.
(d) All of the above.

3. Which function is best suited to determine the orders of a mixed **ARMA**(p, q) process?
(a) EACF
(b) PACF
(c) ACF
(d) a bivariate cross-correlation function

4. The **length** of a prediction interval for Y_{t+l} from fitting a nonstationary ARIMA(p, d, q) model generally
(a) increases as l increases.
(b) decreases as l increases.
(c) becomes constant for l sufficiently large.
(d) tends to zero as l increases.

5. In Homework 5, I had you examine recently-proposed goodness of fit (GOF) statistics by Fisher and Gallagher. What **primary advantage** do these new statistics have over the Ljung-Box GOF statistic we discussed in class?
(a) They are much easier to calculate.
(b) They are normally distributed in large samples (when H_0 is true).
(c) They correspond to tests that have a higher chance of rejecting an incorrect H_0 model (i.e., higher power).
(d) They incorporate decision making principles which unify the large-sample theory behind the ACF, PACF, and EACF.

6. True or False. When compared to Akaike's Information Criterion (AIC), the **Bayesian Information Criterion (BIC)** is more likely to favor overly-large models (i.e., models with more parameters).
(a) True
(b) False

7. In class, we used Monte Carlo simulation to examine the sampling distribution of $\widehat{\theta}$, the MA(1) model parameter estimate of θ using the **method of moments (MOM)**. What did we learn in that example?
(a) Monte Carlo simulation should not be used for MOM estimates when the sample size is small.
(b) We should use overfitting to determine if MA(1) MOM estimates are statistically different from zero.
(c) MOM estimates should not be used in ARIMA models with moving average components.
(d) MOM estimates are satisfactory in stationary models but not in nonstationary models.

8. I have calculated the ACF of an MA(1) process with $\theta = 0.5$. Which value of θ in the MA(1) family produces the **same ACF**?
(a) -0.5
(b) 0
(c) 1
(d) 2

9. If a **seasonal MA$(2)_4$** model is the correct model for a data set, which model is also mathematically correct?
(a) MA$(4)_2$
(b) ARMA(2,4)
(c) AR(8)
(d) MA(8)

10. Here is the R output from fitting an ARMA(1,1) model to a data set that we examined in class:

```
> arima(cows,order=c(1,0,1),method='CSS') # conditional least squares
Coefficients:
         ar1     ma1  intercept
      0.6625  0.6111    58.7013
s.e.  0.0616  0.0670     1.6192
```

What is the **estimate** of the autoregressive parameter ϕ?
(a) -0.6625
(b) -0.6111
(c) 0.6625
(d) 0.6111

11. True or False. If $\{\nabla Y_t\}$ is a **stationary** process, then $\{Y_t\}$ must be stationary.
(a) True
(b) False

12. Consider the seasonal $AR(1)_{12}$ process

$$Y_t = \Phi Y_{t-12} + e_t?$$

Which statement is **true**?
(a) This process is stationary when $-1/12 < \Phi < 1/12$.
(b) The ACF is nonzero at the seasonal lags only.
(c) The PACF decays exponentially across the seasonal lags.
(d) Adding a nonzero mean μ to this model will affect its stationarity properties at seasonal lags only.

13. What is a **likelihood function**?
(a) A function that describes how extended autocorrelations are computed
(b) A function that is maximized to find model estimates
(c) A function that can determine the order of seasonal and nonseasonal differencing
(d) A function that can be used to test for a unit root

14. In class, we compared the art of **ARIMA model selection** to
(a) the unpredictable orbits of Jupiter's moons.
(b) an FBI agent attempting to find a person of interest who may or may not be guilty.
(c) a confused college professor stranded on a deserted island.
(d) the irregular mating habits of the Canadian stink beetle.

15. In an analysis, we have determined the following:

- The sample PACF for the series $\{Y_t\}$ has a slow decay.

- The mean and variance for the series $\{Y_t\}$ are constant over time.

- The first difference process $\{\nabla Y_t\}$ has a sample ACF with a very large spike at lag 1, a smaller spike at lag 2, and no spikes elsewhere.

Which model is **most consistent** with these observations?
(a) MA(1)
(b) ARI(2,1)
(c) AR(2)
(d) IMA(1,2)

16. Suppose that $\{Y_t\}$ is a **white noise process** with a sample size of $n = 100$. If we performed a simulation to study the sampling variation of r_1, the lag one sample autocorrelation, about 95 percent of our estimates r_1 would fall between
(a) -0.025 and 0.025
(b) -0.05 and 0.05
(c) -0.1 and 0.1
(d) -0.2 and 0.2

17. Consider the nonseasonal process defined by

$$(1 + 0.6B)(1 - B)Y_t = (1 - B + 0.25B^2)e_t.$$

This process is identified by which **ARIMA** model?
(a) ARIMA(1,1,2)
(b) ARIMA(2,1,1)
(c) ARIMA(2,2,1)
(d) ARIMA(2,1,2)

18. True or False. The process identified in Question 17 is **stationary**.
(a) True
(b) False

19. You have performed a **Ljung-Box test** to determine if an ARMA(1,1) model is suitable for a data set. The p-value for the test of

$$H_0 : \text{the ARMA}(1, 1) \text{ model is appropriate}$$
$$\text{versus}$$
$$H_1 : \text{the ARMA}(1, 1) \text{ model is not appropriate.}$$

is equal to 0.329. What should you do?
(a) Reject H_0 and look for a better model.
(b) Do not reject H_0 on the basis of this test.

20. In Chapter 3, we talked about using regression methods to detrend time series data that exhibited deterministic trends (e.g., linear, quadratic, seasonal, etc.). For all of the regression output in R to be applicable, we needed the regression model errors to have zero mean, constant variance, and be independent. What **additional assumption** was needed?
(a) invertibility
(b) normality
(c) unbiasedness
(d) strict stationarity

21. Consider an AR(2) model
$$(1 - \phi_1 B - \phi_2 B^2)Y_t = e_t.$$
True or False: If the AR(2) characteristic polynomial $\phi(x) = 1 - \phi_1 x - \phi_2 x^2$ has **imaginary roots**, then this model is not stationary.
(a) True
(b) False

22. In class, we examined the Lake Huron elevation data and decided that an AR(1) model was a good model for these data. Below, I display the estimated standard errors of the forecast error associated with the next 20 MMSE forecasts under the AR(1) model assumption:

```
> round(huron.ar1.predict$se,3)
Start = 2007 End = 2026
 [1] 0.704 0.927 1.063 1.152 1.214 1.258 1.289 1.311 1.328 1.340 1.349 1.355 1.360
[14] 1.363 1.366 1.367 1.369 1.370 1.371 1.371
```

What **quantity** do these estimated standard errors approach as the lead time l increases?
(a) the overall process mean
(b) the overall process variance
(c) the overall process standard deviation
(d) the overall process range

23. I have tentatively decided on an ARI(1,1) model for a process with a decreasing linear trend. I now want to use **overfitting**. Which two models should I fit?
(a) ARI(2,1) and ARIMA(1,1,1)
(b) ARI(1,2) and ARI(2,1)
(c) IMA(1,2) and IMA(2,1)
(d) ARI(1,2) and IMA(2,1)

24. What is the **difference** between strict stationarity and weak stationarity?
(a) Strict stationarity requires that the mean function and autocovariance function be free of time t. Weak stationarity does not.
(b) Strict stationarity is required to guarantee that MMSE forecasts are unbiased (in ARIMA models). These forecasts may not be unbiased under weak stationarity.
(c) Strict stationarity is a stronger form of stationarity that does not rely on large-sample theory.
(d) Strict stationarity refers to characteristics involving joint probability distributions. Weak stationarity refers only to conditions placed on the mean and autocovariance function.

25. True or False. In an **MA(1) process**, MMSE forecasts for lead times $l = 2, 3, 4, ...$, will be identical.
(a) True
(b) False

26. I have a process $\{Y_t\}$. The **first difference** process $\{\nabla Y_t\}$ follows a MA(2) model. What is the appropriate model for $\{Y_t\}$?
(a) MA(1)
(b) ARI(2,1)
(c) IMA(1,2)
(d) ARIMA(0,2,2)

27. I have quarterly time series data $\{Y_t\}$ which display a clear linear trend over time plus a within-quarter seasonal pattern. There are no problems with nonconstant variance. I should consider using a **stationary process** to model
(a) $(1 - B)^4(1 - B^4)Y_t$
(b) $B^4(1 - B)Y_t$
(c) $(1 - B^4)(1 - B)Y_t$
(d) $(1 - B^4)Y_t$

28. Suppose that we have observations from an **AR(1) process** with $\phi = 0.9$. Which of the following is true?
(a) The scatterplot of Y_t versus Y_{t-1} will display a negative linear trend and the scatterplot of Y_t versus Y_{t-2} will display a negative linear trend.
(b) The scatterplot of Y_t versus Y_{t-1} will display a positive linear trend and the scatterplot of Y_t versus Y_{t-2} will display a positive linear trend.
(c) The scatterplot of Y_t versus Y_{t-1} will display a negative linear trend and the scatterplot of Y_t versus Y_{t-2} will display a random scatter of points.
(d) The scatterplot of Y_t versus Y_{t-1} will display a positive linear trend and the scatterplot of Y_t versus Y_{t-2} will display a random scatter of points.

29. Which of the following models would be the best model for **daily stock prices** which have constant variance over time?
(a) AR(1), with ϕ close to -1
(b) ARMA$(P = 1, Q = 1)_5$
(c) IMA(1,1)
(d) MA(5)

30. In Chapter 3, we discussed seasonal means and cosine trend regression models to detrend time series data that exhibited seasonality. When compared to seasonal means models, what **advantage** do the cosine trend models have?
(a) Cosine trend models produce unbiased estimates.
(b) Cosine trend models are easier to fit.
(c) Cosine trend models involve fewer parameters.
(d) Cosine trend models are more likely to produce residuals that have constant variance.

31. What statement best describes the **Principle of Parsimony**?
(a) Complex models do a better job of capturing inherent variability.
(b) Model selection procedures do not necessarily lead to the correct model.
(c) Simple models may be preferred because of their ease of interpretation.
(d) Diagnosing a model's fit is a key part of the data analysis phase.

32. In a nonseasonal AR(2) process with parameters ϕ_1 and ϕ_2, what **geometric figure** emerges if we plot the stationarity region in the ϕ_1-ϕ_2 plane?
(a) unit circle (i.e., a circle with radius 1, centered at the origin)
(b) square
(c) triangle
(d) trapezoid

33. Which **method of estimation** involves equating sample autocorrelations to population autocorrelations and solving the resulting set of equations for ARMA model parameters?
(a) Maximum likelihood
(b) Conditional least squares
(c) Method of moments
(d) Bootstrapping

34. When I was researching the motivation behind the **Augmented Dickey-Fuller ADF** unit root test, I found this passage in another textbook on time series (from a prominent author):

> *"Unfortunately, tests for unit roots generally have poor power, even for moderately sized samples...My view is that a formal test for a unit root can only ever be a small contribution to the important task of modelling (near) non-stationary behaviour."*

What is the author trying to tell us?
(a) Be careful about placing too much faith in the conclusion reached by the ADF test.
(b) Don't use an ADF test unless your sample size is very large.
(c) The ADF test will tell us to reject nonstationarity more often than it should.
(d) Only perform an ADF test if you can find a current R package that implements it reliably.

35. In introductory statistics courses, students are taught about the importance of **means** and **standard deviations** to measure "center" and "spread" of a distribution. Why aren't these concepts a primary focus in time series?
(a) In most time series models, means and standard deviations are biased estimates.
(b) Means and standard deviations are meaningful only for time series data sets that are nonstationary.
(c) The key feature in time series data is that observations over time are correlated; it is this correlation that we look to incorporate in our models.
(d) Means and standard deviations refer to probability distributions; these distributions are of less importance in time series applications.

36. You have a time series that displays **nonconstant variance** (it "fans out" over time) and also a pronounced **linear trend** in the mean over time. What should you do?
(a) Take one difference to remove the linear trend, and then take a second difference to remove the nonconstant variance.
(b) Difference the data first and then determine the suitable variance-stabilizing transformation.
(c) Try a log or square root transformation to remove the linear trend and then difference the resulting series.
(d) I would do something else.

37. Many of the model identification tools and statistical inference procedures we have discussed rely on what we called **asymptotic** theoretical results. What is a potential danger of applying these results in practice?
(a) These tools and procedures may not "work" if we apply them to transformed processes (e.g., square-root, log, etc.).
(b) These tools and procedures may give misleading results if the sample size is small.
(c) MMSE predictions might be valid for large lead times only.
(d) All of the above are potential dangers.

38. True or False. **Overdifferencing** can lead to fitting models that are not invertible.
(a) True
(b) False

39. Who said the **famous quote**, "All models are wrong; some are useful."?
(a) George Box
(b) George Harrison
(c) George Washington
(d) George Clooney

40. In an analysis, we have determined that

- The Dickey-Fuller unit root test for the series $\{Y_t\}$ does not reject a unit root.

- The ACF for the series $\{Y_t\}$ has a very, very slow decay.

- The PACF for the differences $\{\nabla Y_t\}$ has significant spikes at lags 1 and 2 (and is negligible at higher lags).

Which model is **most consistent** with these observations?
(a) IMA(1,1)
(b) ARI(2,1)
(c) ARIMA(2,2,2)
(d) IMA(2,2)

PART 2: PROBLEMS/SHORT ANSWER. Show all of your work and explain all of your reasoning to receive full credit.

1. Suppose that $\{Y_t\}$ is a seasonal $\mathrm{MA}(1) \times \mathrm{MA}(1)_{12}$ process, that is,

$$
\begin{aligned}
Y_t &= (1 - \theta B)(1 - \Theta B^{12})e_t \\
&= e_t - \theta e_{t-1} - \Theta e_{t-12} + \theta\Theta e_{t-13},
\end{aligned}
$$

where $\{e_t\}$ is a zero mean white noise process with variance $\mathrm{var}(e_t) = \sigma_e^2$.
(a) Derive expressions for $\mu_t = E(Y_t)$ and $\gamma_0 = \mathrm{var}(Y_t)$.
(b) Derive the autocovariance function, that is, calculate $\gamma_k = \mathrm{cov}(Y_t, Y_{t-k})$, for $k = 1, 2, \ldots$.

2. Consider the following models:

$$\text{Model A:} \quad (1 - 0.3B + 1.1B^2)Y_t = (1 - 0.5B)e_t$$
$$\text{Model B:} \quad (1 - 0.8B)Y_t = (1 - 1.6B + 0.64B^2)e_t$$

(a) Characterize these models as models in the $\text{ARMA}(p, q)$ family, that is, identify p and q.
(b) Determine if each model corresponds to a stationary process or not.
(c) Write each model above without using backshift notation.

3. Give an example of a process $\{Y_t\}$ that satisfies the following (if we discussed a "named" example in class or elsewhere, just name the process.). Treat each criteria separately, that is, give an example of a process that satisfies (i), one that satisfies (ii), and so on.

 (i) A stationary process that has nonzero autocorrelation at lag 4 only.

 (ii) A nonstationary process with a constant mean.

 (iii) A process with constant mean but variance that increases with time.

 (iv) A nonstationary process that exhibits two types of nonstationarity, namely, an increasing linear trend over time and also a change in seasonal behavior.

 (v) A stationary process that has a lag 1 partial autocorrelation of -0.9 and lag 2 partial autocorrelation equal to zero.

 (vi) A nonstationary process whose first seasonal differences $\{\nabla_s Y_t\}$ are stationary.

4. Consider the time series model

$$Y_t = \theta_0 + \phi^2 Y_{t-1} + e_t,$$

where θ_0 and ϕ are fixed parameters and $\{e_t\}$ is a white noise process with mean zero and variance σ_e^2.

**(a)** Recall that $\widehat{Y}_t(1)$ denotes the $l = 1$ step ahead MMSE forecast. Show that

$$\widehat{Y}_t(1) = \theta_0 + \phi^2 Y_t.$$

(b) Recall that $e_t(1)$ denotes the $l = 1$ step ahead forecast error. Show that

$$\mathrm{var}[e_t(1)] = \sigma_e^2.$$

(c) This process is stationary when $0 < \phi^2 < 1$. When stationarity holds, what will the MMSE forecasts $\widehat{Y}_t(l)$ get close to when the lead time l is very large?

5. As part of a plan to lose weight and to get in better shape, I weighed myself each morning from July 24 through November 9. My weight data are below (recorded in lb). There are $n = 109$ daily measurements.
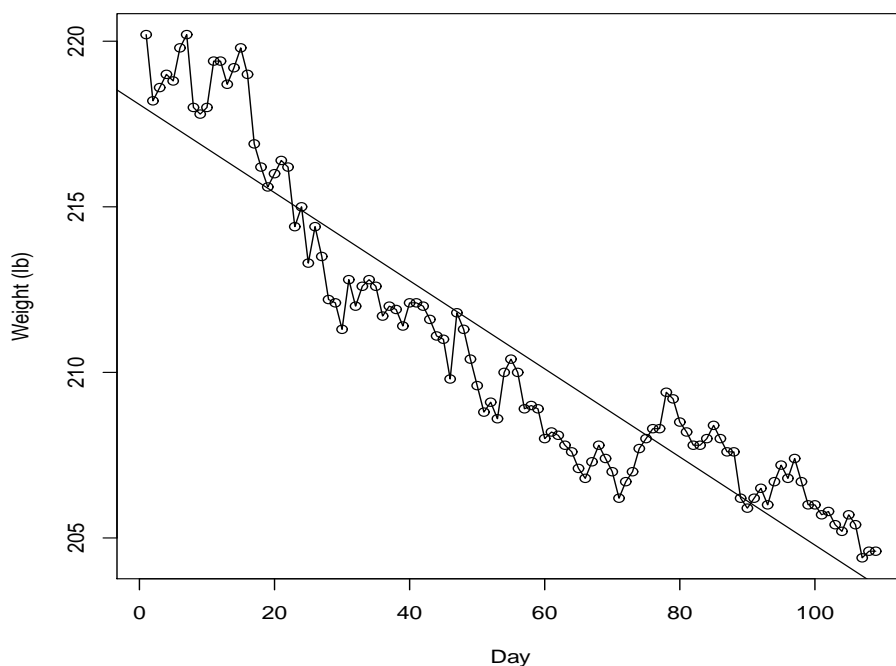
As you can see, I used a simple linear regression model to detrend the data; here is the R output from the straight line fit:

```
> summary(fit)
Call: lm(formula = weight ~ time(weight))

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  218.088277   0.298875  729.70   <2e-16 ***
time(weight)  -0.133031   0.004717  -28.20   <2e-16 ***

Residual standard error: 1.549 on 107 degrees of freedom
Multiple R-squared: 0.8814,     Adjusted R-squared: 0.8803
F-statistic: 795.5 on 1 and 107 DF,  p-value: < 2.2e-16
```
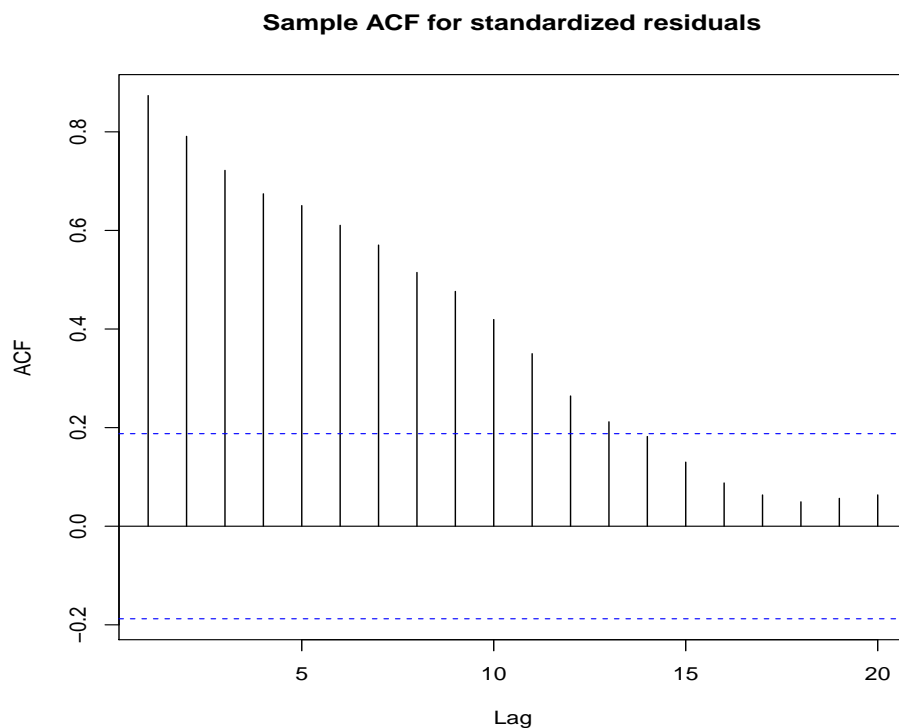
(a) On the next page, I display the sample ACF for the standardized residuals $\{\widehat{X}_t^*\}$ from the straight line fit, that is, the fit of the deterministic trend model

$$Y_t = \beta_0 + \beta_1 t + X_t,$$

where $E(X_t) = 0$. Does it look like the standardized residuals from the fit resemble a white noise process? Why or why not?

Sample ACF for standardized residuals

(b) Calculate the estimated MMSE forecast for Day 130. Based on the model fit (see output on the last page), characterize how MMSE forecasts will behave as you forecast my weight further out in the future.

(c) Instead of using a deterministic trend model, I also considered choosing a model from the ARIMA(p, d, q) family. **When I plotted the sample ACF of the first differences,** all the sample autocorrelations were very close to zero (well within the margin of error bands under the white noise assumption). Based on this information alone, what model in the ARIMA(p, d, q) family would you consider to model my daily weights?

6. The Olympia Paper Company makes absorbent paper towels. Below I depict the weekly paper towels sales (measured in 10,000s rolls) for the most recent 124 weeks.
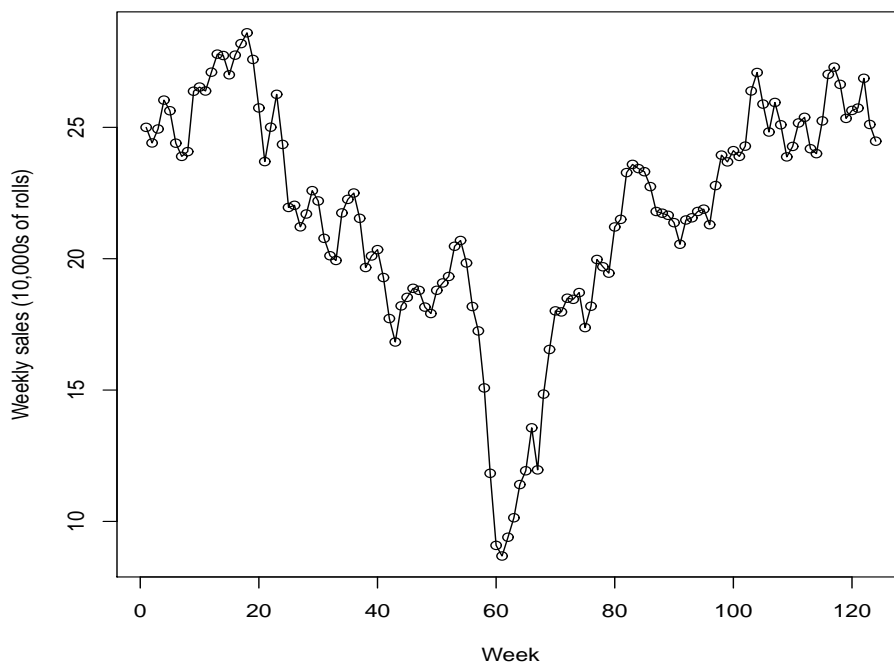
(a) I used R to investigate whether this series exhibited constant variance. A 95 percent confidence interval for Box Cox transformation parameter λ was calculated to be $(0.9, 1.5)$. Do you think a variance stabilizing transformation is necessary? Explain.

(b) I used R to perform an augmented Dickey-Fuller test using the `adf.test` function. Here is the output:

```
> ar(diff(papertowel))

Coefficients:
      1         2
 0.3403   -0.1662


> # Perform ADF test (use k=2 previous command)
> adf.test(papertowel, alternative = "stationary", k=2)

Augmented Dickey-Fuller Test

Dickey-Fuller = -1.6311, Lag order = 2, p-value = 0.7298
alternative hypothesis: stationary
```

Interpret the results from performing this test.

(c) Based on your answer in part (b), what you would expect the sample ACF to look like?

Use the back of this page for your answers.

7. A chemical company produces Chemical Product XB-77-5. In order to develop a control scheme for its production process, chemists need to develop a forecasting model that will give reliable predictions. Here are the data from the most recent 99 days.
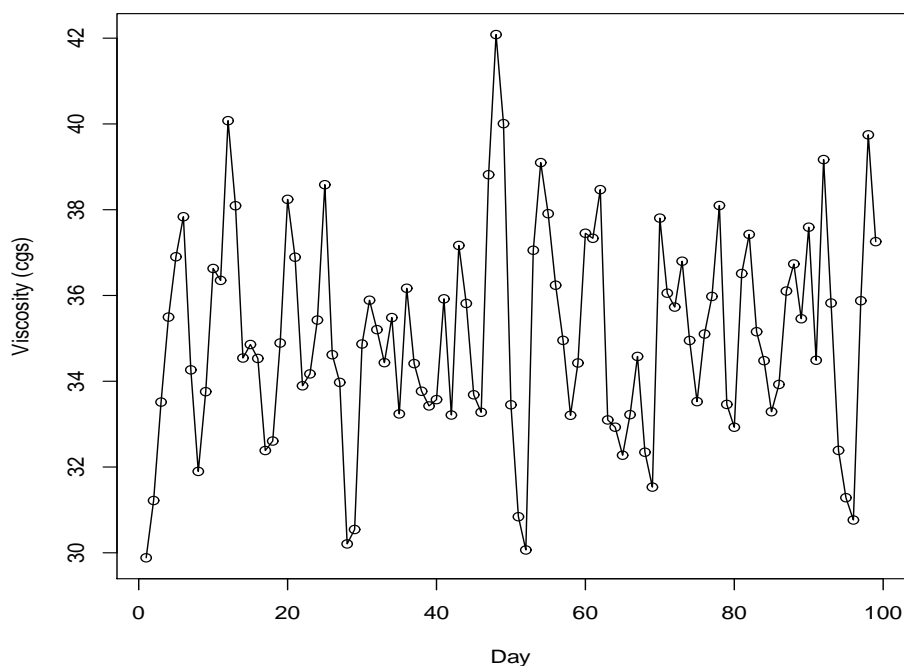
It was judged not necessary to transform the data in any way. The sample PACF revealed statistically significant sample partial autocorrelations at the first three lags (although barely).

(a) I have fit these seven models: AR(1), AR(2), AR(3), MA(1), ARMA(1,1), ARMA(2,1), and ARMA(3,1). Based on the output below, which model would you choose and why?

```
> # AR(1)
> arima(viscosity,order=c(1,0,0),method='ML')

Coefficients:
         ar1   intercept
      0.4451    35.0379
s.e.  0.0921     0.3942

sigma^2 estimated as 4.81:  log likelihood = -218.33,  aic = 440.66

> # AR(2)
> arima(viscosity,order=c(2,0,0),method='ML')

Coefficients:
        ar1      ar2   intercept
      0.618   -0.4315    35.0501
s.e.  0.092    0.0941     0.2466

sigma^2 estimated as 3.959:  log likelihood = -208.89,  aic = 423.78
```

```
> # AR(3)
> arima(viscosity,order=c(3,0,0),method='ML')

Coefficients:
         ar1      ar2      ar3  intercept
      0.5074  -0.2774  -0.2624    35.0600
s.e.  0.0982   0.1083   0.1003     0.1887

sigma^2 estimated as 3.695:  log likelihood = -205.6,  aic = 419.2

> # MA(1)
> arima(viscosity,order=c(0,0,1),method='ML')

Coefficients:
        ma1  intercept
      0.521    35.0445
s.e.  0.074     0.3218

sigma^2 estimated as 4.462:  log likelihood = -214.66,  aic = 433.32

> # ARMA(1,1)
> arima(viscosity,order=c(1,0,1),method='ML')

Coefficients:
         ar1     ma1  intercept
      0.1695  0.4167    35.0387
s.e.  0.1544  0.1283     0.3583

sigma^2 estimated as 4.411:  log likelihood = -214.11,  aic = 434.23

> # ARMA(2,1)
> arima(viscosity,order=c(2,0,1),method='ML')

Coefficients:
         ar1      ar2      ma1  intercept
      0.9727  -0.5997  -0.4537    35.0683
s.e.  0.1397   0.0864   0.1624     0.1714

sigma^2 estimated as 3.726:  log likelihood = -206,  aic = 420

> # ARMA(3,1)
> arima(viscosity,order=c(3,0,1),method='ML')

Coefficients:
         ar1      ar2      ar3      ma1  intercept
      0.6368  -0.3562  -0.2078  -0.1400    35.0628
s.e.  0.3207   0.2219   0.1731   0.3233     0.1809

sigma^2 estimated as 3.688:  log likelihood = -205.5,  aic = 421.01
```

(b) Are there any models in this list that you wish I had fit? If so, say which one(s) and tell me why. If not, tell me why you are satisfied. Use the back of this page if you need extra space.

8. This is a continuation of Problem 7. For the viscosity data, I chose what I felt was the "best" model from those available (I called the Oracle to ask her, but she did not answer). Based on the fit of my model, I then produced the next 21 MMSE forecasts for Days 100-120. These are depicted below; I cut off the Days 1-20 observations for aesthetic reasons.
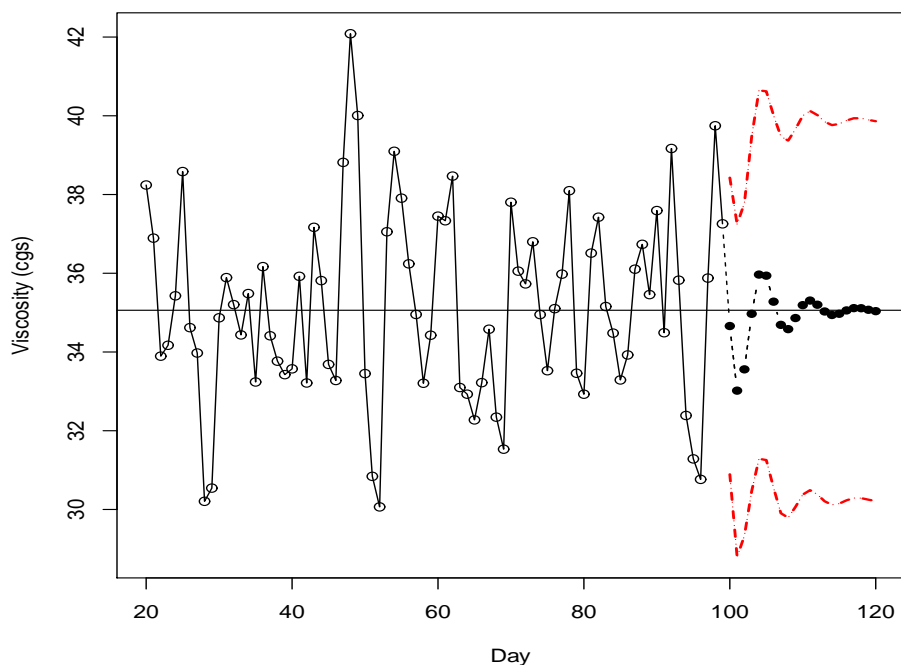

(a) My repeated "Oracle" references throughout the semester come from the 1999 hit movie "The Matrix." What is my intended meaning behind these references? Discuss this in terms of model selection/identification.

(b) A 95 percent prediction interval for Day 100 (the 1-step ahead interval) is $(30.8, 38.4)$. Interpret this interval in words. What key assumption is needed for this interval to be valid?

(c) I won't tell you what model I fit in Problem 7, but you should be able to tell something by looking at the pattern in the MMSE forecasts (for Days 100-120). Characterize this pattern and then tell me what type of model would produce this pattern in the forecasts.

9. The monthly retail sales in US sporting goods and bicycle stores are presented below from January 2000 through December 2012 (there are $n = 144$ observations).
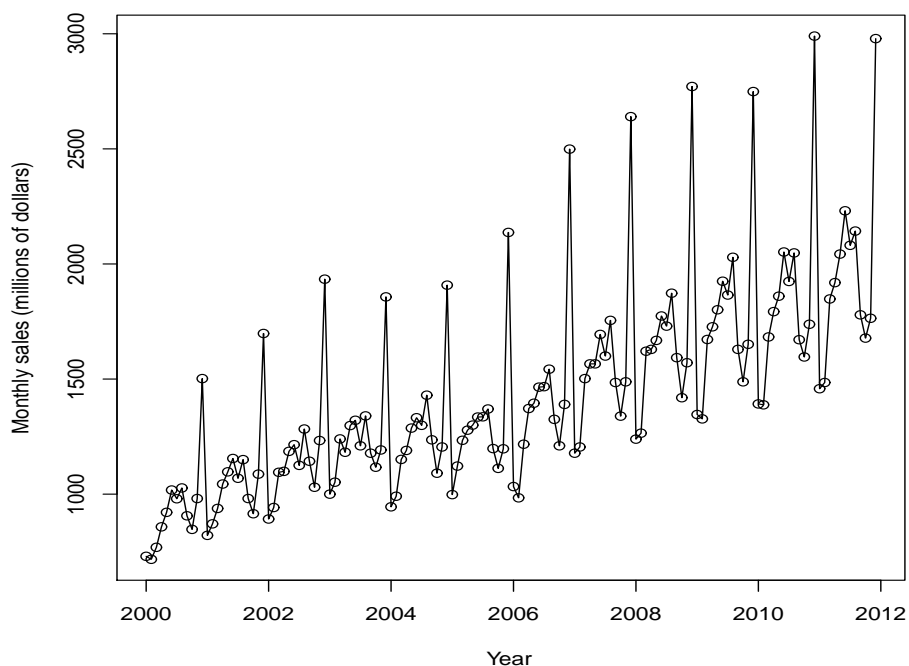

On the next page, I have displayed the sample ACF and sample PACF for the values of

$$(1 - B) \log Y_t = \log Y_t - \log Y_{t-1},$$

the first differences of the log-transformed data.

(a) Why do you think I used a log-transformation here? Why did I use the first (nonseasonal) differences? Answer each question precisely.

(b) Based on the sample ACF and sample PACF on the next page, what (seasonal) SARIMA model would you consider using for $\log Y_t$? That is, in the $\text{ARIMA}(p, d, q) \times \text{ARIMA}(P, D, Q)_s$ family of models, identify the values of p, d, q, P, D, Q, and s. **You must explain/defend your selections!**

Sample ACF: 1st differences of log-transformed data   **Sample PACF: 1st differences of log-transformed data**
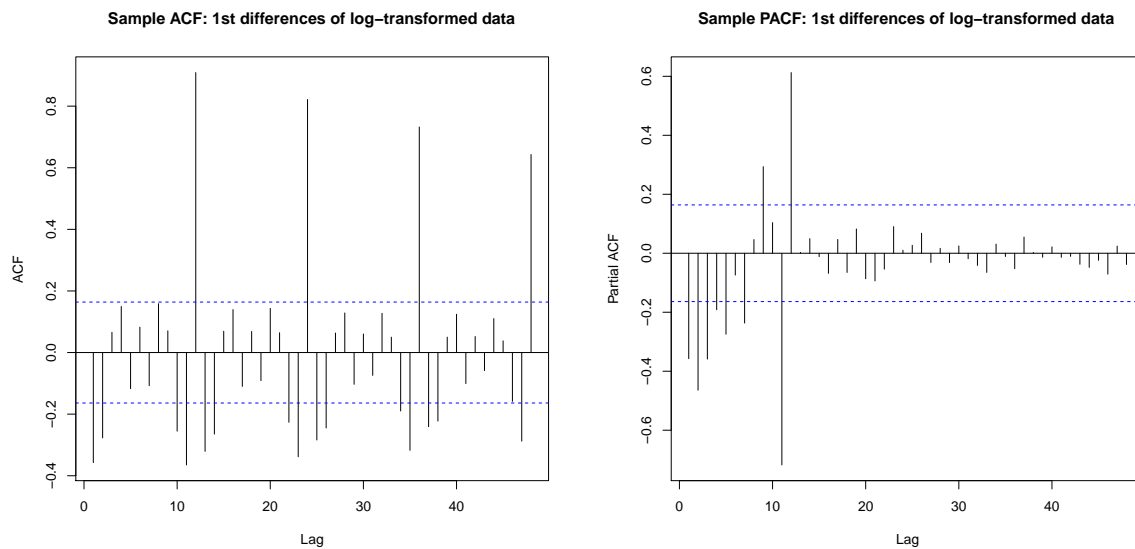
10. This is a continuation of Problem 9. For the sporting sales data, I selected what I felt was the best model from the SARIMA family and then performed model diagnostics for it. The `tsdiag` output from R is on the next page.

(a) Interpret each panel in the `tsdiag` output. How satisfied are you with the model that I fit?

(b) What effect does the log-transformation have if I want to calculate MMSE forecasts? Comment briefly.