GROUND RULES:

- This exam contains two parts:
  - **Part 1.** Multiple Choice (50 questions, 1 point each)
  - **Part 2.** Problems/Short Answer (10 questions, 5 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 $\text{var}(e_t) = \sigma_e^2$.
- 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 of the following processes is **stationary**?
(a) An MA(1) process with $\theta = -1.4$
(b) $Y_t = 12.3 + 1.1Y_{t-1} + e_t$
(c) IMA(1,1)
(d) $Y_t = \beta_0 + \beta_1 t + e_t$

2. Which statement about an **AR(2) process** is always true?
(a) The process is invertible.
(b) The process is stationary.
(c) The theoretical ACF $\rho_k = 0$ for all $k > 2$.
(d) The theoretical PACF ϕ_{kk} decays exponentially or according to a sinusoidal pattern as k gets large.

3. Suppose that we have observations from an **MA(1) process** with $\theta = 0.9$. 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.

4. Suppose that we have observed the time series $Y_1, Y_2, ..., Y_t$. If the forecast error $e_t(l) = Y_{t+l} - \widehat{Y}_t(l)$ has **mean zero**, then we say that the MMSE forecast $\widehat{Y}_t(l)$ is
(a) stationary.
(b) unbiased.
(c) consistent.
(d) complementary.

5. What did we discover about the **method of moments** procedure when estimating parameters in ARIMA models?
(a) The procedure gives reliable results when the sample size $n > 100$.
(b) The procedure gives unbiased estimates.
(c) The procedure should not be used when models include AR components.
(d) The procedure should not be used when models include MA components.

6. What is an **AR characteristic polynomial**?
(a) A function that can be used to assess normality
(b) A function of that can be used to characterize invertibility properties.
(c) A function that can be used to characterize stationary properties.
(d) All of the above.

7. Which statement about **MMSE forecasts** in stationary ARMA models is true?
(a) If $\widehat{Y}_t(l)$ is the MMSE forecast of $\ln Y_{t+l}$, then $e^{\widehat{Y}_t(l)}$ is the MMSE forecast of Y_{t+l}.
(b) As the lead time l increases, $\widehat{Y}_t(l)$ will approach the process mean $E(Y_t) = \mu$.
(c) As the lead time l increases, $\text{var}[\widehat{Y}_t(l)]$ will approach the process variance $\text{var}(Y_t) = \gamma_0$.
(d) All of the above are true.

8. If $\{Y_t\}$ follows an **IMA(1,1) process**, then $\{\nabla Y_t\}$ follows a(n) ____________ process.
(a) ARI(1,1)
(b) MA(1)
(c) IMA(2,1)
(d) ARIMA(0,1,2)

Use the R output below to answer Questions 9 and 10.

```
> data.ar1.fit = arima(data,order=c(1,0,0),method='ML')
> data.ar1.fit
Coefficients:
         ar1  intercept
      0.4796   179.4921
s.e.  0.0565     0.4268
sigma^2 estimated as 6.495:  log likelihood = -126.24,  aic = 296.48
```

9. Concerning the output above, which statement is **true**?
(a) The AR(1) parameter estimate $\widehat{\phi}$ is statistically different from zero.
(b) This process has mean zero.
(c) The positive AIC means that an AR(1) model is not adequate.
(d) All of the above.

10. If you were going to **overfit** this model for diagnostic purposes, which two models would you fit?
(a) IMA(1,1) and ARMA(2,2)
(b) ARMA(1,1) and AR(2)
(c) ARMA(1,2) and ARI(1,1)
(d) IMA(1,1) and ARI(2,1)

11. We used R to generate a white noise process. We then calculated **first differences** of this white noise process. What would you expect the sample ACF r_k of the first differences to look like?
(a) Most of the r_k estimates should be close to zero, possibly with the exception of a small number of estimates which exceed the white noise bounds when k is larger.
(b) The r_k estimates will decay very, very slowly across k.
(c) The r_1 estimate should be close to -0.5 and all other r_k estimates, $k > 1$, should be small.
(d) It is impossible to tell unless we specify a distributional assumption for the white noise process (e.g., normality).

12. True or False. The use of an **intercept term** θ_0 has the same effect in stationary and nonstationary ARIMA models.
(a) True
(b) False

13. The augmented **Dickey-Fuller** unit root test can be used to test for
(a) normality.
(b) independence.
(c) stationarity.
(d) invertibility.

14. An observed time series displays a clear upward linear trend. We fit a straight line regression model to remove this trend, and we notice that the residuals from the straight line fit are stationary in the mean level. **What should we do next?**
(a) Search for a stationary ARMA process to model the residuals.
(b) Perform a Shapiro-Wilk test.
(c) Calculate the first differences of the residuals and then consider fitting another regression model to them.
(d) Perform a t-test for the straight line slope estimate.

15. Excluding the intercept θ_0 and white noise variance σ_e^2, which model has the **largest** number of parameters?
(a) ARIMA$(1,1,1) \times (2,0,1)_{12}$
(b) ARMA(3,3)
(c) ARMA$(1,1) \times (1,2)_4$
(d) ARIMA(2,2,3)

16. In performing diagnostics for an ARMA(1,1) model fit, I see the following output in R:

```
> runs(rstandard(data.arma11.fit))
$pvalue
[1] 0.27
```

How do I **interpret** this output?
(a) The standardized residuals seem to be well modeled by a normal distribution.
(b) The standardized residuals are not well represented by a normal distribution.
(c) The standardized residuals appear to be independent.
(d) We should probably consider a model with either $p > 1$ or $q > 1$ (or both).

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

18. A 95 percent confidence interval for the **Box Cox transformation parameter** λ is $(0.77, 1.41)$. Which transformation is appropriate?
(a) Square root
(b) Square
(c) Log
(d) Identity (no transformation)

19. What is the **name** of the process defined by

$$(1 + 0.6B)(1 - B)Y_t = (1 - 0.9B)^2 e_t?$$

(a) ARIMA(1,1,2)
(b) ARIMA(2,1,1)
(c) ARIMA(1,2,1)
(d) ARIMA(2,0,1)

20. If an $\mathbf{AR}(1)_{12}$ model is the correct model for a data set, which model is also mathematically correct?
(a) AR(1)
(b) AR(12)
(c) AR(11)
(d) ARMA(1,12)

21. In Chapter 3, we discussed the use of regression to detrend a time series. Two models used to handle seasonal trends were the **cosine-trend model** and the **seasonal means model**. Which statement is true?
(a) With monthly data (so that the number of seasonal means is 12), the cosine trend model is more parsimonious.
(b) Standardized residuals from the cosine trend model fit will be normally distributed if the process is truly sinusoidal.
(c) Differencing should be used before fitting a seasonal means model.
(d) All of the above are true.

22. When we used least squares regression to fit the deterministic trend regression model $Y_t = \beta_0 + \beta_1 t + X_t$ in Chapter 3, the only assumption we needed for the least squares estimators $\widehat{\beta}_0$ and $\widehat{\beta}_1$ to be **unbiased** was that
(a) $\{X_t\}$ has constant variance.
(b) $\{X_t\}$ is a white noise process.
(c) $\{X_t\}$ has zero mean.
(d) $\{X_t\}$ is normally distributed.

23. If $\{\nabla Y_t\}$ is an **ARI(1,1) process**, then what is the correct model for $\{Y_t\}$?
(a) Random walk with drift
(b) AR(1)
(c) IMA(2,1)
(d) None of the above

24. For a stationary ARMA(p, q) process, when the sample size n is large, the **sample autocorrelations** r_k
(a) follow an MA(1) process.
(b) are approximately normal.
(c) are likely not to be statistically significant from zero.
(d) have variances which decrease as k gets larger.

25. What is the main characteristic of an **AR(1) process** with parameter $\phi = 0.2$?
(a) The mean of the process is equal to 0.2.
(b) The variance of the process is equal to $(0.2)^2 = 0.04$.
(c) The autocorrelation function ρ_k exhibits a slow decay across lags.
(d) None of the above.

26. Consider the process

$$Y_t - Y_{t-1} = e_t - 0.5e_{t-1}.$$

How is this process written using **backshift notation**?
(a) $(1 - B)Y_t = (1 - 0.5B)e_t$
(b) $BY_t = (1 - 0.5B)e_t$
(c) $B(Y_t - Y_{t-1}) = 0.5Be_t$
(d) None of the above.

27. What is the **name** of the process identified in Question #26?
(a) IMA(2,2)
(b) IMA(1,1)
(c) ARI(1,1)
(d) None of the above.

28. In class, we proved that the autocorrelation function for a **zero-mean random walk** $Y_t = Y_{t-1} + e_t$ is equal to

$$\rho_k = \sqrt{\frac{t}{t+k}},$$

for $k = 1, 2, ...,$. Which of the following statements is true?
(a) This process is stationary.
(b) The variance of this process approaches $\gamma_0 = 1$.
(c) $\text{corr}(Y_1, Y_2)$ is larger than $\text{corr}(Y_{99}, Y_{100})$.
(d) None of the above.

29. Why is **invertbility** so important?
(a) If a process is not invertible, residuals can not be calculated for diagnostics purposes.
(b) If a process is not invertible, MMSE forecasts can not be determined uniquely.
(c) If a process is not invertible, we can not use the ACF, PACF, and EACF for model specification.
(d) All of the above.

30. We used the **Bonferroni criterion** for judging a standardized residual (from an ARIMA model fit) as an outlier. What essentially does this mean?
(a) We look at the mean of each residual and take the largest one as an outlier.
(b) Each residual is compared to $z_{0.025} \approx 1.96$, and all those beyond 1.96 (in absolute value) are declared outliers.
(c) We perform an intervention analysis and determine if the associated parameter estimate is significant.
(d) None of the above.

31. I have fit the following model:

$$(1-B)(1-B^4)(1-0.43B^4)Y_t = (1+0.22B)(1+0.88B^4)e_t.$$

Which model does this represent?
(a) $\text{ARIMA}(1,1,1) \times \text{ARIMA}(1,0,1)_4$
(b) $\text{ARIMA}(0,1,1) \times \text{ARIMA}(1,1,1)_4$
(c) $\text{ARIMA}(1,1,1) \times \text{ARIMA}(1,0,1)_4$
(d) $\text{ARIMA}(0,2,1) \times \text{ARIMA}(1,1,1)_4$

32. What statement about the **seasonal** $\mathbf{MA}(2)_{12}$ **model** is true?
(a) It is stationary.
(b) It is invertible.
(c) Its ACF ρ_k is nonzero when $k=1$ and $k=2$; otherwise, $\rho_k = 0$.
(d) Its PACF ϕ_{kk} is nonzero only when $k=12$ and $k=24$.

33. Which process is **not stationary**?
(a) White noise
(b) $(1-0.3B)Y_t = (1-B)e_t$
(c) The first difference of a ARIMA(1,1,1) process with $\phi = 1.5$ and $\theta = -0.5$
(d) An AR(2) process whose AR characteristic polynomial roots are $x = 0.8 \pm 0.9i$

34. If an ARIMA model is correctly specified and our estimates are reasonably close to the true parameters, then the residuals should behave **roughly** like an iid normal white noise process. Why do we say "roughly?"
(a) R can not calculate residuals at early lags, so it is impossible to tell.
(b) The mean of the residuals at early lags is not zero (as in a white noise process).
(c) The autocorrelations of residuals are slightly different than those of white noise at early lags.
(d) None of the above.

35. The length of a prediction interval for Y_{t+l} computed from fitting a **stationary** ARMA(p,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.

36. Under normality, what is a valid interpretation of the **partial autocorrelation** ϕ_{kk}?
(a) It measures the autocorrelation in the data $Y_1, Y_2, ..., Y_n$ after taking kth differences.
(b) It is the correlation between Y_t and Y_{t-k}, after removing the linear effects of the variables between Y_t and Y_{t-k}.
(c) It equals the variance of the large-sample distribution of r_k.
(d) It is the correlation of the first k residuals in an ARIMA model fit.

37. Recall that a model's **BIC** is given by

$$\text{BIC} = -2 \ln L + k \ln n,$$

where $\ln L$ is the natural logarithm of the maximized likelihood function and k is the number of parameters in the model. Which statement is true?
(a) The smaller the BIC, the better the model.
(b) When compared to the AIC, BIC offers a stiffer penalty for models with a large number of parameters.
(c) Both (a) and (b) are true.
(d) Neither (a) nor (b) are true.

38. True or False. In a stationary ARMA(p, q) model, **maximum likelihood estimators** of model parameters (i.e., the ϕ's and the θ's) are approximately normal in large samples.
(a) True
(b) False

39. In the acronymn "**ARIMA**," what does the "I" stand for?
(a) Independence
(b) Integrated
(c) Intraclass
(d) Irreversible

40. The first difference of a **stationary AR(1) process** can be expressed as
(a) a white noise process.
(b) an invertible MA(1) process.
(c) a nonstationary AR(2) process.
(d) a stationary ARMA(1,1) process.

41. You have an observed time series that has clear nonconstant variance and a sharp linear trend over time. **What should you do?**
(a) Display the ACF, PACF, and EACF of the observed time series to look for candidate models.
(b) Split the data set up into halves and then fit a linear regression model to each part.
(c) Try differencing the series first and then try a transformation to stabilize the variance.
(d) Try a variance-stabilizing transformation first and then use differencing to remove the trend.

42. During lecture, I recounted a true story where I had asked a fortune teller (in the French Quarters) to comment on the precision of her predictions about my future. In **what city** did this story take place?
(a) New Orleans
(b) Nome
(c) Nairobi
(d) Neverland

43. What was the **famous quote** we cited from Box?
(a) "Going to time series class is like going to church; many attend but few understand."
(b) "Modeling time series data is like running over hot coals; do it fast so that it hurts less."
(c) "Detrending time series data is for amateur statisticians."
(d) "All models are wrong; some are useful."

44. Suppose that $\{Y_t\}$ is a **white noise process** and that $n = 400$. 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

45. What is the **stationarity condition** for the seasonal $\text{AR}(1)_{12}$ given by

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

(a) $-12 < \Phi < 12$
(b) $-1 < \Phi < 1$
(c) $-1/12 < \Phi < 1/12$
(d) None of the above

46. What **technique** did we use in class to simulate the sampling distribution of sample autocorrelations and method of moments estimators?
(a) Monte Carlo
(b) Bootstrapping
(c) Jackknifing
(d) Backcasting

47. True or False. In a stationary ARMA(p, q) process, the **MMSE forecast** $\widehat{Y}_t(l)$ depends on the MA components only when $l \leq q$.
(a) True
(b) False

48. The **Yule-Walker equations** can be used to
(a) assess if standardized residuals from a stationary ARMA(p, q) fit are normally distributed.
(b) calculate "early" residuals for a stationary ARMA(p, q) fit.
(c) compute autocorrelations for a stationary ARMA(p, q) process.
(d) determine if a variance-stabilizing transformation is necessary.

49. In the notes, we discussed this quote:

> *"Simulation evidence suggests a preference for the maximum likelihood estimator for small or moderate sample sizes, especially if the moving average operator has a root close to the boundary of the invertibility region."*

In this quote, what is meant by the **phrase** "especially if the moving average operator has a root close to the boundary of the invertibility region?"
(a) the MA characteristic polynomial has roots that are close to zero.
(b) the roots of the MA characteristic polynomial that are all less than one in absolute value or modulus.
(c) the MA characteristic polynomial has at least one root which is complex.
(d) None of the above.

50. I have monthly time series data $\{Y_t\}$ which display a clear quadratic trend over time plus a within-year seasonal component. There are no problems with nonconstant variance. I should consider using a **stationary process** to model
(a) $(1 - B)^{12}(1 - B^2)Y_t$
(b) $(1 - B)^{24}Y_t$
(c) $(1 - B)^2(1 - B^{12})Y_t$
(d) $(1 - B^{12})^2Y_t$

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 an MA$(1)_4$ process with mean μ, that is,

$$Y_t = \mu + e_t - \Theta e_{t-4},$$

where $\{e_t\}$ is a zero mean white noise process with $\text{var}(e_t) = \sigma_e^2$.
(a) Find $\mu_t = E(Y_t)$ and $\gamma_0 = \text{var}(Y_t)$.
(b) Show that $\{Y_t\}$ is (weakly) stationary.

2. Suppose that $\{e_t\}$ is zero mean white noise with $\text{var}(e_t) = \sigma_e^2$. Consider the model

$$Y_t = 0.5Y_{t-1} + e_t - 0.2e_{t-1} - 0.15e_{t-2}.$$

(a) Write this model using backshift notation.
(b) Determine whether this model is stationary and/or invertible.
(c) Identify this model as an ARIMA(p, d, q) process; that is, specify p, d, and q.

3. Suppose that $\{e_t\}$ is zero mean white noise with $\text{var}(e_t) = \sigma_e^2$. Consider the deterministic model

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

where $X_t = X_{t-1} + e_t - \theta e_{t-1}$.
(a) Derive an expression for ∇Y_t.
(b) What is the name of the process identified by $\{\nabla Y_t\}$? Is $\{\nabla Y_t\}$ stationary?

4. Explain how the sample ACF, PACF, and EACF can be used to specify the orders p and q of a stationary ARMA(p, q) process.

5. Suppose that $\{e_t\}$ is zero mean white noise with $\text{var}(e_t) = \sigma_e^2$. Consider the random walk with drift model

$$Y_t = \theta_0 + Y_{t-1} + e_t.$$

We have observed the data $Y_1, Y_2, ..., Y_t$.
(a) Show that the minimum mean squared error (MMSE) forecast of Y_{t+1} is equal to

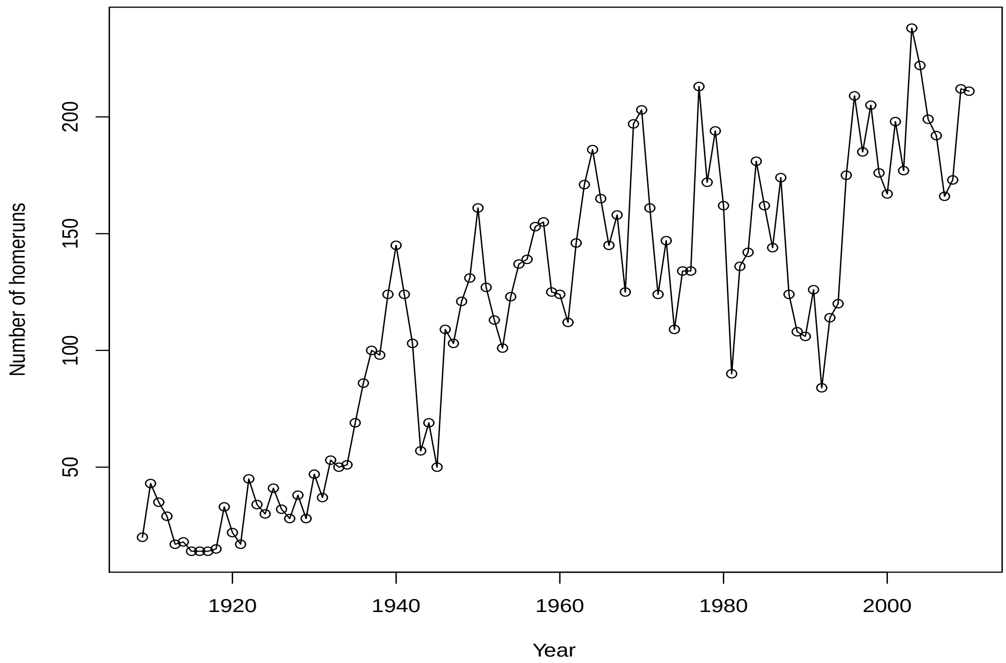
$$\widehat{Y}_t(1) = \theta_0 + Y_t.$$

(b) Show that the MMSE forecast of Y_{t+l} satisfies

$$\widehat{Y}_t(l) = \theta_0 + \widehat{Y}_t(l-1),$$

for all $l > 1$.

6. In class, we looked at the number of homeruns hit by the Boston Red Sox each year during 1909-2010. Denote this process by $\{Y_t\}$.



I have used R to fit the model

$$(1 - B)Z_t = e_t - \theta e_{t-1},$$

where $Z_t = \sqrt{Y_t}$. Here is the output:

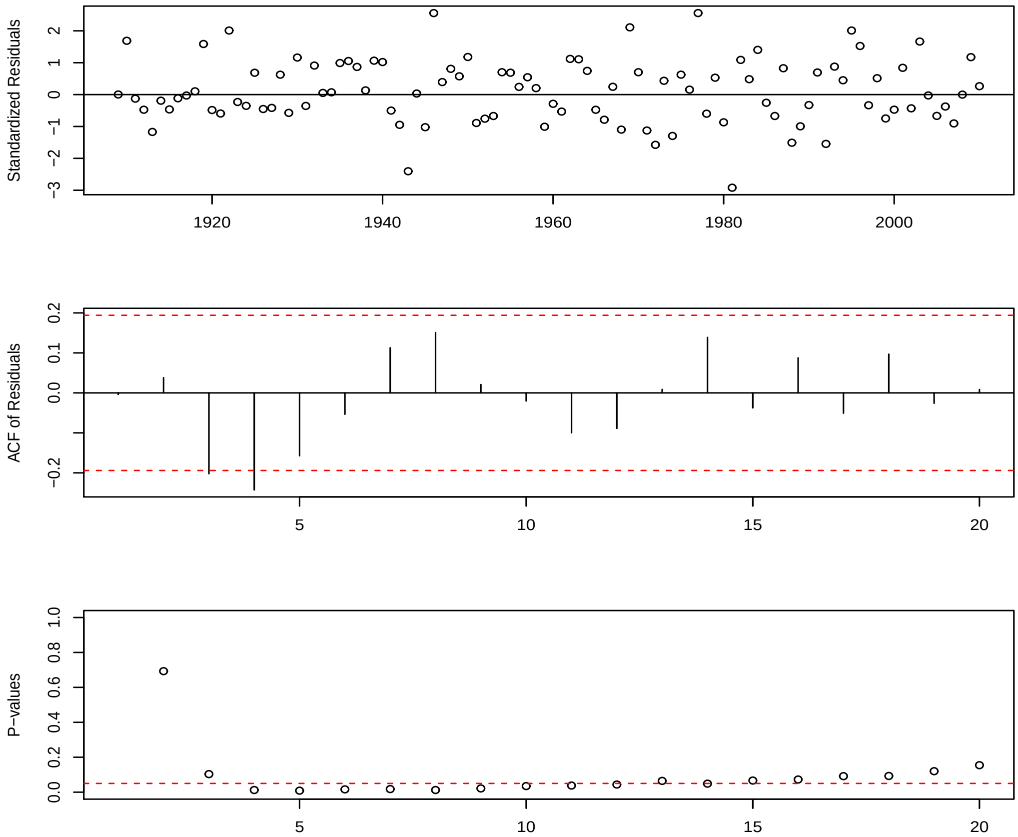
```
> arima(sqrt(homeruns),order=c(0,1,1),method='ML') # maximum likelihood
Coefficients:
          ma1
      -0.2488
s.e.   0.1072
sigma^2 estimated as 1.437:  log likelihood = -161.64,  aic = 325.28
```

Answer the questions below; the next page can be used for your responses.

(a) Why do you think I used the square-root transformation? Why do you think I used a nonstationary model?
(b) Based on the model that I fit (judged to be "reasonable" during the model specification phase), what do you think the sample ACF of $\{Z_t\}$ looked like? the sample ACF of $\{\nabla Z_t\}$?
(c) Write an approximate 95 percent confidence interval for θ based on the model fit. Interpret the interval.

This page is for your responses to **Question 6**.

7. I have displayed below the `tsdiag` output from fitting the model in Question 6 to the Boston Red Sox homerun data.



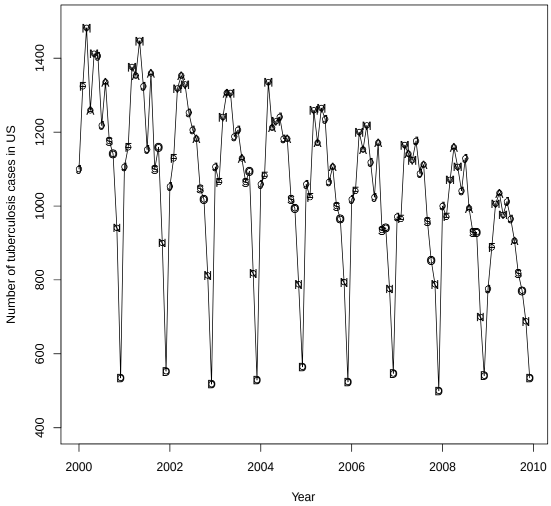
I have also performed the Shapiro-Wilk and runs tests for the standardized residuals; see the R output below:

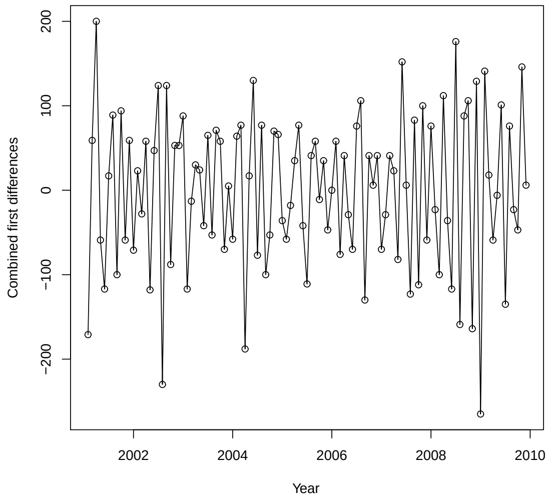
```
> shapiro.test(rstandard(homerun.fit)
W = 0.9884, p-value = 0.5256

> runs(rstandard(homerun.fit))
$pvalue
[1] 0.378
```

Based on the information available, what do you think of the adequacy of the model fit in Question 6? Use the **back of this page** if you need extra space.

8. Recall the TB data from our midterm; these data are the number of TB cases (per month) in the United States from January 2000 to December 2009. Denote this process by $\{Y_t\}$. On the midterm, you used regression methods to detrend the data. On this problem, we will use a SARIMA modeling approach. In the figure below, the TB data are on the left, and the combined first differenced data $(1-B)(1-B^{12})Y_t$ are on the right.





In Question 9, we will fit this model to the data:

$$(1-B)(1-B^{12})Y_t = (1-\theta B)(1-\Theta B^{12})e_t.$$

I have arrived at this (candidate) model from examining the sample ACF and sample PACF of the combined first differences $(1-B)(1-B^{12})Y_t$. Before we get to the model fit, answer the questions below. Use the **back of this page** if you need extra space.

(a) Why did I work with the combined first differences $(1-B)(1-B^{12})Y_t$?
(b) Assuming that the model above is a reasonable choice, what would you expect the sample ACF and sample PACF of the combined first differences $(1-B)(1-B^{12})Y_t$ to look like?
(c) The model above is a member of the $\text{ARIMA}(p,d,q)\times\text{ARIMA}(P,D,Q)_s$ class. Identify the values of p, d, q, P, D, Q, and s.

9. As promised, I used R to fit the model stated in Question 8 (using maximum likelihood). Here is the output:

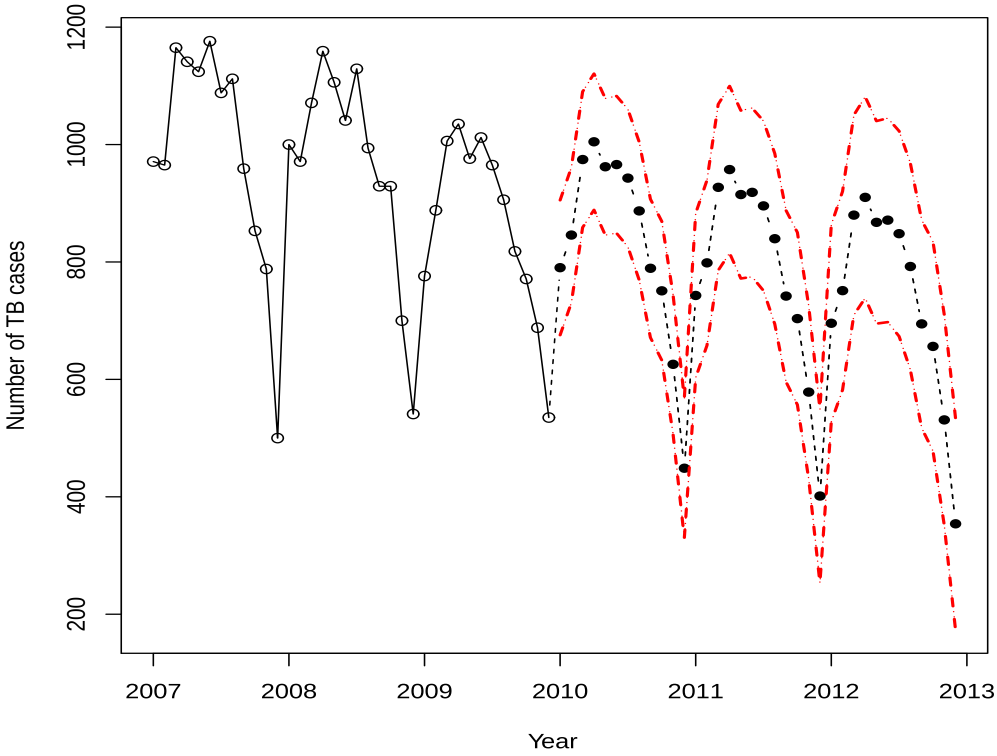
```
> tb.fit
Coefficients:
          ma1     sma1
      -0.9182  -0.4406
s.e.   0.0482   0.1098
sigma^2 estimated as 3436:  log likelihood = -589.81,  aic = 1183.63
```

Therefore, the fitted model is given by

$$(1-B)(1-B^{12})Y_t = (1-0.9182B)(1-0.4406B^{12})e_t.$$

(a) Rewrite this fitted model without using backshift notation and then simplify as much as possible. I want an equation with only Y_t on the left hand side and the rest of the model equation on the right hand side.
(b) Are the estimates of θ and Θ statistically different from 0? Explain.
(c) Why doesn't R display an "intercept term" in the output?

10. The model I fit to the TB data in Question 9 has been declared a very good model (after doing thorough residual diagnostics and overfitting). In the figure below, I have displayed the MMSE forecasts (with 95 percent prediction limits) for the next 3 years (Jan 2010 through Dec 2012).



(a) A 95 percent prediction interval for the TB count in January 2012 is (527.7, 863.5). Interpret this interval.
(b) Which assumption on the white noise terms $\{e_t\}$ is crucial for the prediction interval in part (a) to be valid?
(c) Careful inspection reveals that the prediction limits in the figure above tend to widen as the lead time increases. Why is this true?