TEST #1
STA 4853/5856 **Name:**____
March 6, 2013

Please read the following directions.
DO NOT TURN THE PAGE UNTIL INSTRUCTED TO DO SO

Directions

- This exam is **closed book** and **closed notes**.
- The problems are all multiple choice or fill in the blank.
- Circle a **single** answer for each multiple choice question. Your choice should be made clearly.
- There is no penalty for guessing.
- In one problem, you need to know a fact about the normal distribution: $z_{\alpha/2} = 1.96$ for $\alpha = .05$.
- The exam has **12** pages.
- Each multiple choice question and "fill in the blank" is worth equal credit. There are a total of 32 equally weighted items (since Problem 13 has two blanks to fill in).

Problem 1. You wish to apply the methods of ARIMA modeling to a time series. Suppose you observe that the variability of the series increases systematically with the level. What should you do?

a)⋆ Try a transformation

b) Try differencing at lag 1

c) Try differencing at the seasonal lag

d) Use a seasonal model

e) Try fitting a trend

f) None of the above

Problem 2. A time series has a stationary mean, but you observe that the variability increases steadily with time. Which of the following is likely to help?

a) Try a transformation

b) Try differencing at lag 1

c) Try differencing at the seasonal lag

d) Use a seasonal model

e) Try fitting a trend

f)⋆ None of the above

Problem 3. Suppose $\{z_t\}$ is an AR(2) process:

$$z_t = C + \phi_1 z_{t-1} + \phi_2 z_{t-2} + a_t$$

If you observe a realization $z_1, z_2, \ldots, z_n$ where n is very large (say, a million), and fit a regression model

$$z_t = \beta_0 + \beta_1 z_{t-1} + \beta_2 z_{t-2} + \beta_3 z_{t-3} + \varepsilon_t$$

to this data, then the estimate of β_3 will be approximately equal to ...

a) a_{t-3}	**b)** C	**c)** ϕ_1	**d)** ϕ_2	**e)** σ_a
f) z_3	**g)** a_3	**h)** z_{t-3}	**i)**⋆ 0	**j)** 1

Problem 4. Suppose $\{z_t\}$ is a stationary AR(2) process:

$$z_t = C + \phi_1 z_{t-1} + \phi_2 z_{t-2} + a_t$$

Which one of the following pairs of random variables are independent of each other?

a) z_t and z_{t+1}	**b)** z_t and a_t	**c)** z_t and a_{t-1}	**d)** z_t and z_{t-2}
e) z_t and a_{t-3}	**f)** z_t and z_{t-1}	**g)**⋆ z_t and a_{t+1}	**h)** z_t and z_{t-3}

Problem 5. In the blank provided, write an equation expressing a purely seasonal $\text{ARIMA}(1,0,1)_8$ model.

____

The answer is $z_t = C + \Phi_1 z_{t-8} + a_t - \Theta_1 a_{t-8}$. *The coefficient names are not important. Give full credit if they have the right kind of terms at the right lags. If they forget the constant but everything else is right, give full credit. The answer can also be written in backshift form as* $(1 - \Phi_1 B^8) z_t = C + (1 - \Theta_1 B^8) a_t$.

For problems 6 to 9, fill in the blanks with numerical answers.

These problems all use the fragment of output given below obtained from fitting a regression model $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \varepsilon$ for a response variable Y on two covariates X_1 and X_2.

Root MSE	10.00	R-Square	0.88
Dependent Mean	12.00	Adj R-Sq	0.87
Coeff Var	20.00		

Parameter Estimates					
		Parameter	Standard		
Variable	DF	Estimate	Error	t-value	Pr > $\|t\|$
Intercept	1	4.0	0.5	8.0	< .0001
X1	1	3.0	2.0	1.5	.1369
X2	1	1.0	2.5	0.4	.6900

Problem 6. Find a 95% confidence interval for β_1.

the confidence interval is ______________________________

Problem 7. Assuming the residuals are approximately normally distributed with mean zero, give a value A such that you would expect about 68% of the residuals to be between $-A$ and $+A$.

$\boldsymbol{A}$ = ____________________

Consider a case with values:

Y	X_1	X_2
5.0	2.0	3.0

Problem 8. What is the **predicted value** (also known as fitted value) for this case?

predicted value= ____________________

Problem 9. What is the **residual** for this case?

residual= ____________________

Problem 10. For an AR(p) process,

a) The theoretical **PACF** decays to zero, either exponentially or with a damped sine wave pattern or with both of these patterns.

b)$\star$ The theoretical **ACF** decays to zero, either exponentially or with a damped sine wave pattern or with both of these patterns.

c) The theoretical **ACF** has a cutoff to zero after lag p.

Problem 11. A stationary AR(3) process can be re-written as a ______ process.

a)$\star$ MA(∞) **b)** AR(2) **c)** AR(1) **d)** MA(1) **e)** MA(2) **f)** MA(3)

Problem 12. Suppose $EX = 3$ and $EY = 7$. What is the value of $E(2X + 3Y + 4)$? (Give a numerical answer.)

$\boldsymbol{E(2X + 3Y + 4) =}$ ______________________________

Problem 13. The ARMA process

$$z_t = \phi_1 z_{t-1} + \phi_2 z_{t-2} + a_t - \theta_1 a_{t-1}$$

can be written in backshift form as $\phi(B)z_t = \theta(B)a_t$. Write expressions for $\phi(B)$ and $\theta(B)$ in the blanks provided.

$\boldsymbol{\phi(B) =}$ ____

$\boldsymbol{\theta(B) =}$ ____

Problem 14. The ARMA process $\phi(B)z_t = \theta(B)a_t$ can also be written as $z_t = \psi(B)a_t$. What is $\psi(B)$?

a) $\phi(B)/\theta(B)$ **b)** $(1 - B)^d$ **c)** $1 - B$ **d)** $\theta(B) + \phi(B)$
e)$\star$ $\theta(B)/\phi(B)$ **f)** $\theta(B) - \phi(B)$ **g)** $(1 - B)\phi(B)$ **h)** $(1 - B)\theta(B)$

Problem 15. Suppose $\psi(B) = \sum_{k=0}^{\infty} \psi_k B^k$. The expression $\psi(B)a_t$ is equivalent to which of the following:

a) $a_t - \theta_1 a_{t-1} - \cdots - \theta_q a_{t-q}$
b) $a_t + \theta_1 a_{t-1} + \cdots + \theta_q a_{t-q}$
c) $\sum_{k=0}^{\infty} \psi_{t-k} a_t$
d) $\sum_{k=0}^{\infty} \psi_{t+k} a_t$
e)$\star$ $\sum_{k=0}^{\infty} \psi_k a_{t-k}$
f) $\sum_{k=0}^{\infty} \psi_k a_{t+k}$

Problem 16. What is the formula for SBC (Schwarz's Bayesian Criterion)? (In the choices below, L is the likelihood, k is the number of parameters, and n is the number of residuals, the length of the series minus any observations lost due to differencing.)

a) $-2\ln(L) + 2k$

b) $+2\ln(L) - 2k$

c) $2L - 2k$

d) $-2L - 2k$

e)⋆ $-2\ln(L) + k\ln(n)$

f) $+2\ln(L) - k\ln(n)$

Problem 17. The process generated by

$$z_t = 20.0 + 0.5z_{t-1} + a_t - 1.5\theta_1 a_{t-1} + 0.8a_{t-2}$$

is called a ______ process.

a) ARIMA(1,1,2) **b)** ARIMA(2,1,1) **c)**⋆ ARMA(1,2) **d)** ARMA(2,1) **e)** MA(2) **f)** AR(1)

Problem 18. An MA(5) process is stationary

a) if the roots of $\theta(B) = 0$ lie strictly outside the unit circle

b) if $|\theta_1| < 1$

c) if $|\theta_1| > 1$

d) if $|\theta_2| < 1$ and $\theta_2 \pm \theta_1 < 1$

e) if $|\theta_5| < 1$

f)⋆ always

g) never

Problem 19. Suppose you have fit four different non-seasonal models to a time series, and all models have acceptable residual diagnostics with all their parameters significantly different from zero. The SBC values for these models are as follows:

	Model	SBC
A	(0,0,1)	+109
B	(0,0,2)	+108
C	(2,0,0)	+107
D	(1,0,0)	+106
E	(1,0,1)	+105

Given only this information, which model is preferred?

a) A **b)** B **c)** C **d)** D **e)**⋆ E **f)** not clear

Problem 20. On another time series, suppose you have fit four different seasonal models, and all models have acceptable residual diagnostics with all their parameters significantly different from zero. The AIC values for these models are as follows:

	Model	AIC
A	$(0,0,2)(1,0,0)_{12}$	-100
B	$(1,0,0)(1,0,0)_{12}$	-101
C	$(1,1,0)(0,0,1)_{12}$	-102
D	$(1,0,0)(1,1,0)_{12}$	-104

Given only this information, which model is preferred?

a) A **b)** B **c)** C **d)** D **e)**$\star$ not clear

Problem 21. The model

$$(1-\phi_1 B-\phi_2 B^2-\phi_3 B^3)(1-\Phi_1 B^6)z_t = C + a_t$$

is a special case of a ______ model.

a)$\star$ AR(9) **b)** AR(18) **c)** AR(4) **d)** AR(6) **e)** AR(3)

f) MA(9) **g)** MA(18) **h)** MA(4) **i)** MA(6) **j)** MA(3)

Problem 22. For a stationary process $\{z_t\}$, let $\gamma_k = \mathrm{Cov}(z_t, z_{t-k})$ for $k = 0, \pm 1, \pm 2, \ldots$ Which of the following statements is always true?

a) $\mu_z = \dfrac{C}{1-\gamma_1-\gamma_2-\cdots-\gamma_p}$

b) $\gamma_k = \theta_1^k$

c) $\gamma_k = 0$ for $k > q$

d) $\gamma_k = \dfrac{\sigma_a^2}{1-\phi_k^2}$

e)$\star$ $\rho_k = \dfrac{\gamma_k}{\gamma_0}$

Problem 23. For any stationary ARMA(p,q) process, the value of ϕ_{kk} (the partial autocorrelation at lag k) can be computed as a function of ______

a) $\phi_1, \phi_2, \ldots, \phi_p$ **b)** $\theta_1, \theta_2, \ldots, \theta_q$ **c)**$\star$ $\rho_1, \rho_2, \ldots, \rho_k$ **d)** ρ_k **e)** θ_k **f)** ϕ_k

Problem 24. The Ljung-Box statistic $Q(m)$ is a test of the null hypothesis

a) $H_0 : \rho_m = 0$

b)$\star$ $H_0 : \rho_1 = \rho_2 = \cdots = \rho_m = 0$

c) $H_0 : \rho_1 = 0$

d) $H_0 : \rho_m = \rho_{m+1} = \rho_{m+2} = \cdots = 0$

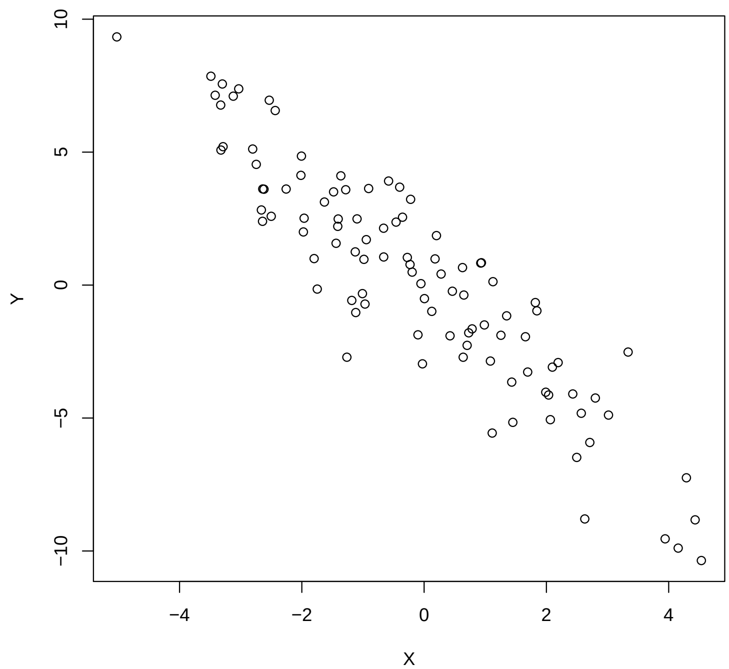
Problem 25. If the value of the Ljung-Box statistic $Q(m)$ is large, the associated p-value will be ______.

a) close to 1 **b)**⋆ close to 0 **c)** close to −1 **d)** close to 0.5

Problem 26. The scatterplot below depicts a sample of size 100 from a population with correlation ______.

a) −10 **b)** −2 **c)** −1 **d)**⋆ −0.9 **e)** −0.2

f) 0 **g)** +0.2 **h)** +0.9 **i)** +1 **j)** +2 **k)** +10



Problem 27. A list of stationary processes is given below. Which of these processes will be the most difficult to distinguish from a non-stationary process based on the sample ACF?

a) MA(1) with $\theta_1 = -0.5$

b) MA(1) with $\theta_1 = 0.9$

c) MA(1) with $\theta_1 = 0.5$

d) AR(1) with $\phi_1 = -0.5$

e)⋆ AR(1) with $\phi_1 = 0.9$

f) AR(1) with $\phi_1 = 0.5$

g) random shocks

Problem 28. Suppose $a_1, a_2, a_3, \ldots$ is a random shock sequence. What is the value of

$$E[(\psi_0 a_3 + \psi_1 a_2 + \psi_2 a_1)^2] \text{ ?}$$

a) $\psi_0^2 a_3^2 + \psi_1^2 a_2^2 + \psi_2^2 a_1^2$
b) $\sigma_a^2(\psi_0\psi_1 + \psi_1\psi_2)$
c)⋆ $\sigma_a^2(\psi_0^2 + \psi_1^2 + \psi_2^2)$
d) $\sigma_a(\psi_0 + \psi_1 + \psi_2)$
e) $\sigma_a(\psi_0\psi_1 + \psi_1\psi_2)$
f) $\sigma_z(\psi_0 + \psi_1 + \psi_2)$
g) $\sigma_z(\psi_0\psi_1 + \psi_1\psi_2)$

The following three pages (pages 9 to 11) give output (produced by the IDENTIFY statement in PROC ARIMA) for a series z_t and its first and second differences. Using this output, select a reasonable ARIMA(p, d, q) model for this series. Specify your answer in the next two questions.

Problem 29. What value of d should be used?

a) $d = 0$ **b)** $d = 1$ **c)**⋆ $d = 2$ **d)** $d = 3$

Problem 30. What values of p and q should be used?

a) $p = 1, q = 0$ **b)** $p = 2, q = 0$ **c)** $p = 3, q = 0$ **d)** $p = 0, q = 1$
e) $p = 1, q = 1$ **f)**⋆ $p = 0, q = 2$ **g)** $p = 0, q = 3$ **h)** $p = 2, q = 2$

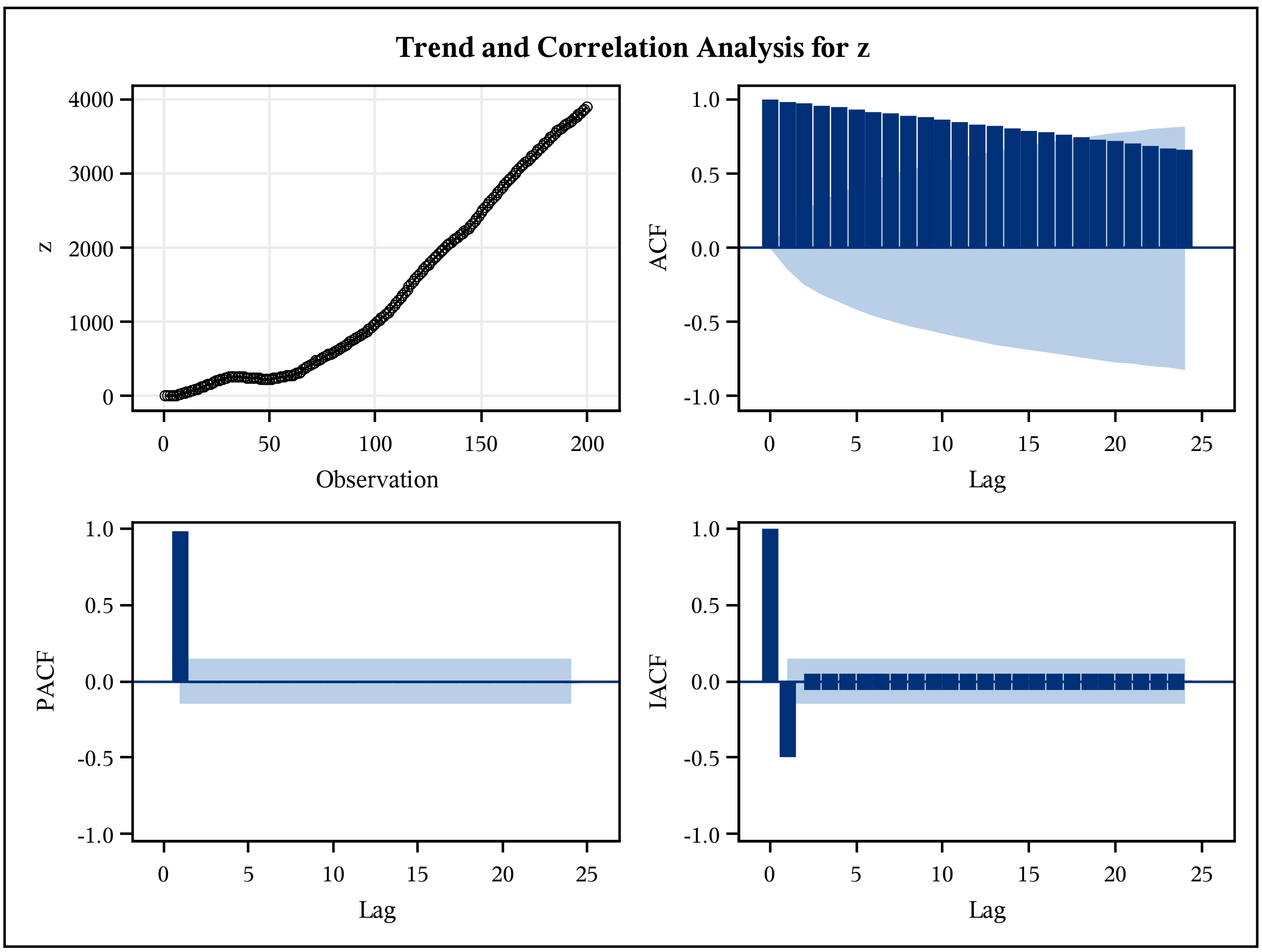
Problem 31. The last page of the exam (page 12) gives the sample ACF and PACF of a monthly time series $\{y_t\}$. The time series plot is not given but appears to be stationary. Select a plausible ARIMA$(p, d, q)(P, D, Q)_{12}$ model for this series.

a)⋆ $(1, 0, 0)(0, 0, 1)_{12}$
b) $(1, 0, 0)(1, 0, 0)_{12}$
c) $(0, 0, 1)(0, 0, 1)_{12}$
d) $(0, 0, 1)(1, 0, 0)_{12}$
e) $(1, 1, 0)(1, 0, 0)_{12}$
f) $(0, 0, 1)(0, 1, 1)_{12}$
g) $(1, 1, 0)(0, 1, 1)_{12}$

The ARIMA Procedure

Name of Variable = z	
Mean of Working Series	1420.536
Standard Deviation	1251.21
Number of Observations	200

Autocorrelation Check for White Noise									
To Lag	Chi-Square	DF	Pr > ChiSq	Autocorrelations					
6	1120.89	6	<.0001	0.987	0.973	0.960	0.946	0.933	0.919
12	2084.34	12	<.0001	0.905	0.891	0.877	0.863	0.849	0.834
18	2891.39	18	<.0001	0.820	0.806	0.791	0.776	0.762	0.747
24	3548.84	24	<.0001	0.732	0.718	0.703	0.688	0.673	0.659

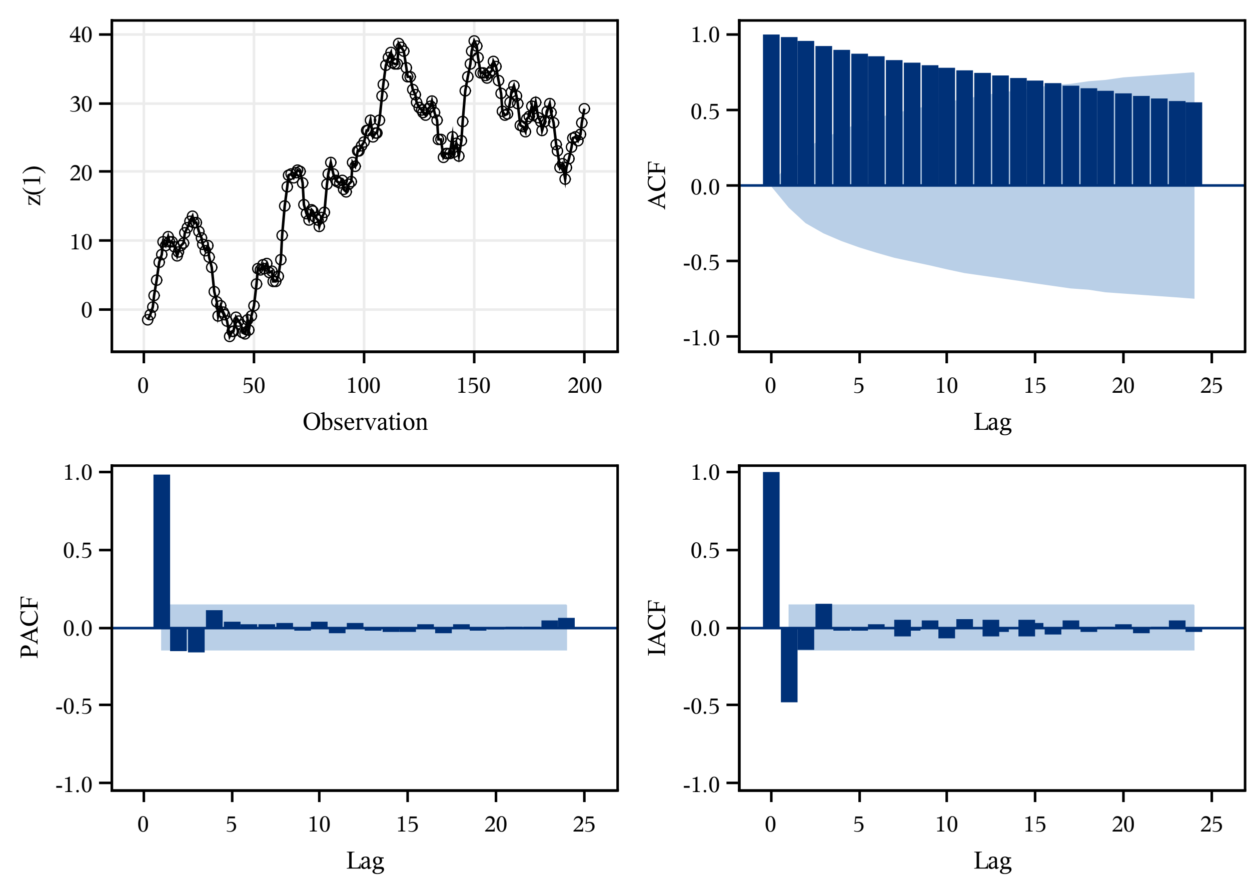


The ARIMA Procedure

Name of Variable = z	
Period(s) of Differencing	1
Mean of Working Series	19.6022
Standard Deviation	11.7548
Number of Observations	199
Observation(s) eliminated by differencing	1

Autocorrelation Check for White Noise									
To Lag	**Chi-Square**	**DF**	**Pr > ChiSq**	**Autocorrelations**					
6	1030.80	6	<.0001	0.981	0.957	0.927	0.900	0.875	0.853
12	1819.30	12	<.0001	0.832	0.814	0.796	0.780	0.763	0.747
18	2435.18	18	<.0001	0.730	0.713	0.695	0.677	0.659	0.642
24	2898.56	24	<.0001	0.624	0.607	0.591	0.575	0.562	0.551

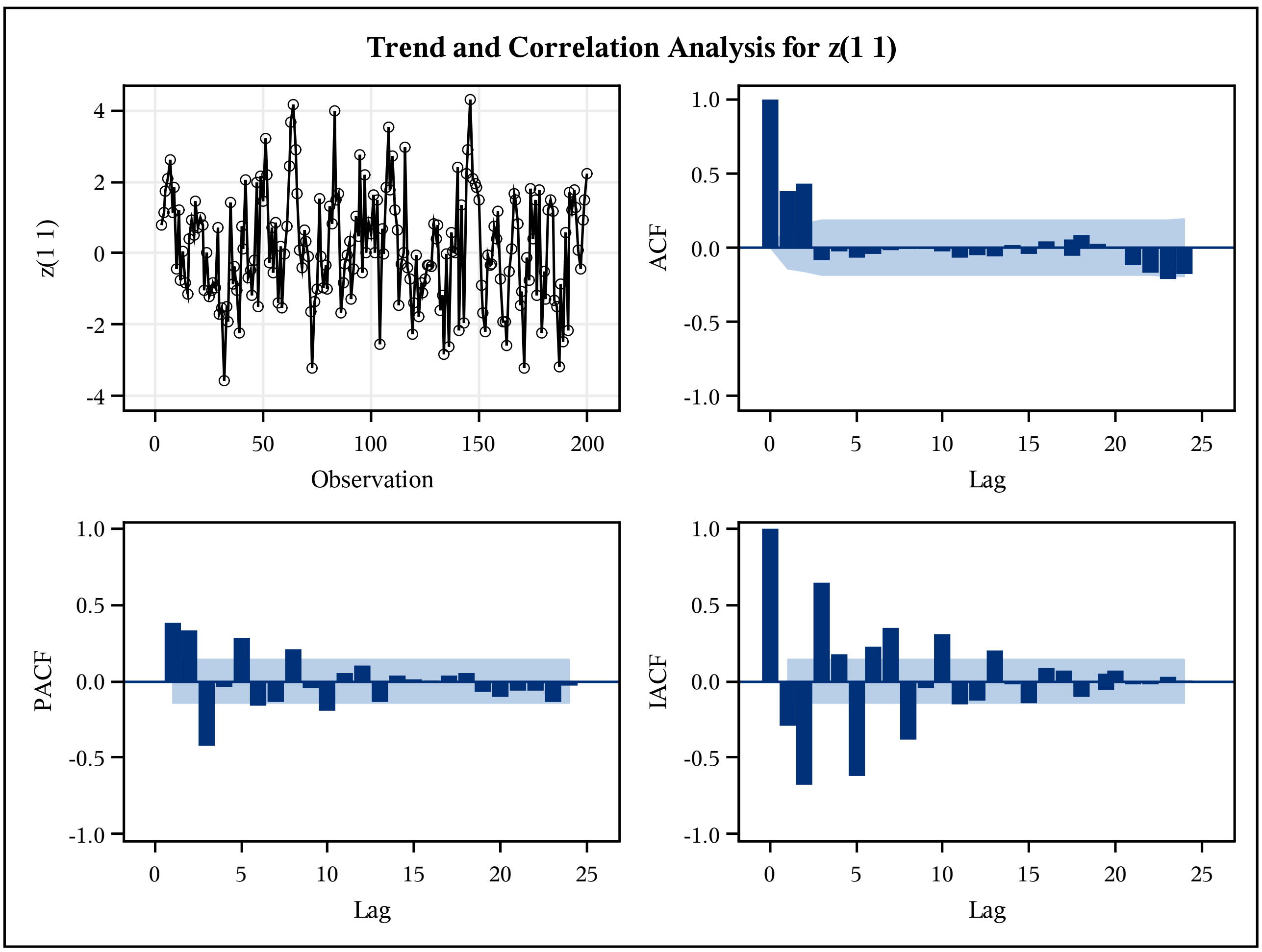
Trend and Correlation Analysis for z(1)



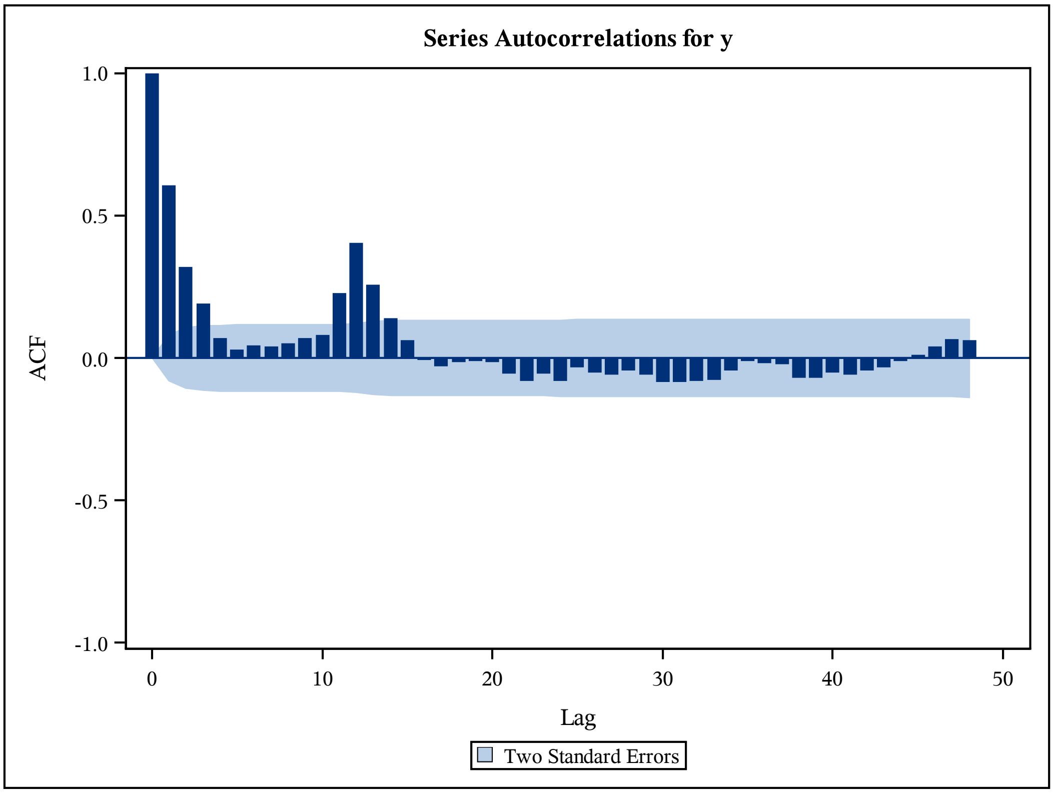
The ARIMA Procedure

Name of Variable = z	
Period(s) of Differencing	1,1
Mean of Working Series	0.155902
Standard Deviation	1.551766
Number of Observations	198
Observation(s) eliminated by differencing	2

Autocorrelation Check for White Noise									
To Lag	**Chi-Square**	**DF**	**Pr > ChiSq**	**Autocorrelations**					
6	69.21	6	<.0001	0.379	0.431	-0.087	-0.029	-0.064	-0.041
12	70.80	12	<.0001	-0.017	-0.004	-0.010	-0.021	-0.066	-0.048
18	73.73	18	<.0001	-0.057	0.015	-0.039	0.040	-0.003	0.083
24	100.61	24	<.0001	0.025	-0.012	-0.115	-0.166	-0.215	-0.177



The ARIMA Procedure



The ARIMA Procedure