TEST #1

STA 4853 **Name:**____________________________________

March 6, 2024

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

Directions

- This exam is **closed book** and **closed notes**.
- Circle the **single best** answer for each multiple choice question. Your choice should be made clearly.
- Always **circle the correct response**. (Sometimes the question has an empty blank or a box, but this is **NOT** where the answer goes.)
- Each question is worth equal credit.
- There is no penalty for guessing.
- There are **33** multiple choice questions.
- The exam has **17** pages.

Problem 1. For a regression model with residuals $e_1, e_2, \ldots, e_n$, the Durbin-Watson statistic is equal to _________.

a) $\dfrac{\sum_{t=2}^{n}(e_t + e_{t-1})^2}{\sum_{t=1}^{n} e_t^2}$

b) $\dfrac{\sum_{t=1}^{n} e_t^2}{\sum_{t=2}^{n}(e_t + e_{t-1})^2}$

c)⋆ $\dfrac{\sum_{t=2}^{n}(e_t - e_{t-1})^2}{\sum_{t=1}^{n} e_t^2}$

d) $\dfrac{\sum_{t=1}^{n} e_t^2}{\sum_{t=2}^{n}(e_t - e_{t-1})^2}$

e) $\dfrac{\sum_{t=2}^{n}(e_t + e_{t-1})}{\sum_{t=1}^{n} e_t}$

f) $\dfrac{\sum_{t=1}^{n} e_t}{\sum_{t=2}^{n}(e_t + e_{t-1})}$

g) $\dfrac{\sum_{t=2}^{n}(e_t - e_{t-1})}{\sum_{t=1}^{n} e_t}$

h) $\dfrac{\sum_{t=1}^{n} e_t}{\sum_{t=2}^{n}(e_t - e_{t-1})}$

Problem 2. The values of X and Y are observed for each individual in a random sample of size n. If the sample correlation between X and Y is exactly $+1$ or -1, then _________.

a)⋆ there is a perfect linear relationship between X and Y in the sample

b) the residuals from the regression of Y on X will exhibit strong serial correlation

c) the regression of Y on X will have residuals with variance exactly equal to 1

d) the sample covariance between X and Y is exactly zero

e) the n points $(X_1, Y_1), (X_2, Y_2), \ldots, (X_n, Y_n)$ lie exactly on a straight line with slope equal to $+1$ or -1

f) the regression line fit to the data $(X_1, Y_1), (X_2, Y_2), \ldots, (X_n, Y_n)$ has a slope exactly equal to $+1$ or -1

g) the residuals from the regression of Y on X will have an autocorrelation exactly equal to $+1$ or -1

Problem 3. What is the correct interpretation of the P-value of 0.0136 in the table fragment below? The P-value of 0.0136 is the probability _________.

Parameter	Estimate	Std. Err.	t-Value	P-value
θ	-0.24630	0.09986	-2.47	0.0136

a) of rejecting the null hypothesis when the true value of θ is 0.24630

b) of rejecting the null hypothesis when the true value of θ is zero

c) of getting an estimate greater than -0.24630 when the true value of θ is zero

d) of getting a t-value greater than -2.47 by chance when the true value of θ is 0.24630

e) that the magnitude of the true value of θ is more than 1.96×0.09986

f) that the true value of θ is outside the interval $-0.24630 \pm 1.96 \times 0.09986$

g) of getting an estimate whose magnitude is at least 0.24630 by chance when the true value of θ is zero

h)⋆ of getting a t-value whose magnitude is at least 2.47 by chance when the true value of θ is zero

Problem 4. If the series $z_1, z_2, \ldots, z_n$ is a realization of a stationary process, then

$$\frac{\sum_{t=1}^{n-k}(z_t - \bar{z})(z_{t+k} - \bar{z})}{\sum_{t=1}^{n}(z_t - \bar{z})^2}$$

is an estimate of ________.

a) σ_z^2 **b)** σ_a^2 **c)** θ_k **d)** ϕ_k

e)⋆ ρ_k **f)** γ_k **g)** ϕ_{kk} **h)** μ_z

Problem 5. The autocorrelation function (ACF) of a stationary AR(1) process satisfies ________.

a) $\rho_k = 0$ for $k \geq 2$ **b)** $\rho_k = 0$ for $k \geq 1$

c)⋆ $\rho_k = \phi_1 \rho_{k-1}$ for $k > 0$ **d)** $\rho_{k-1} = \phi_1 \rho_k$ for $k > 0$

e) $\rho_k = \sigma_a^2 \sum_{i=0}^{q-k} \psi_i \psi_{i+k}$ **f)** $\rho_k = \sigma_a^2 \sum_{i=0}^{q-k} \psi_i \psi_{i-k}$

g) $\rho_k = \dfrac{\sigma_a^2}{1 - \phi_1^k}$ **h)** $\rho_k = \dfrac{\sigma_a^k}{1 - \phi_1^2}$

Problem 6. The following table contains computer output obtained by fitting the model

$$Y_t = \beta_0 + \beta_1 X_{1,t} + \beta_2 X_{2,t} + \varepsilon_t\,, \quad t = 1, 2, \ldots, n,$$

using standard regression software using ordinary least squares. The data are time ordered and the errors $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$ display strong serial correlation.

		Col. #1	Col. #2	Col. #3	Col. #4
	Parameter	Estimate	Std. Error	t value	p-value
Row #1	(Intercept)	-9.4805	3.1025	-3.0558	0.0025
Row #2	X1	1.3384	0.0511	26.1843	0.0000
Row #3	X2	0.2477	0.1487	1.6652	0.0972

In situations like this, there is one row or column in this table where the numerical values have some validity (they are unbiased but likely to be inefficient), but the others could be completely wrong and way off. Which one of the rows or columns has some validity?

a) Row #1 **b)** Row #2 **c)** Row #3

d)⋆ Col. #1 **e)** Col. #2 **f)** Col. #3 **g)** Col. #4

Problem 7. A time series $z_1, z_2, \ldots, z_{113}, z_{114}$ of length $n = 114$ has some of its values displayed in the column labeled `z` in the table below. The remaining columns are obtained by "lagging" the column `z`. The sample correlation between the variables `z` and `zlag4` will be approximately equal to ________.

a)⋆ ρ_4 b) γ_4 c) ϕ_{44} d) θ_4

e) ϕ_4 f) 1 g) -1 h) 0

	z	zlag1	zlag2	zlag3	zlag4	zlag5	zlag6	zlag7	zlag8
1	5.59	.	.	.	.	.	.	.	.
2	5.77	5.59	.	.	.	.	.	.	.
3	6.37	5.77	5.59	.	.	.	.	.	.
4	6.77	6.37	5.77	5.59	.	.	.	.	.
5	7.30	6.77	6.37	5.77	5.59	.	.	.	.
6	7.94	7.30	6.77	6.37	5.77	5.59	.	.	.
7	8.28	7.94	7.30	6.77	6.37	5.77	5.59	.	.
8	8.69	8.28	7.94	7.30	6.77	6.37	5.77	5.59	.
9	8.51	8.69	8.28	7.94	7.30	6.77	6.37	5.77	5.59
10	7.85	8.51	8.69	8.28	7.94	7.30	6.77	6.37	5.77
11	6.26	7.85	8.51	8.69	8.28	7.94	7.30	6.77	6.37
12	4.58	6.26	7.85	8.51	8.69	8.28	7.94	7.30	6.77
(Rows 13 through 108 are omitted)									
109	6.18	6.27	7.34	7.98	8.18	7.80	7.03	5.99	5.43
110	6.50	6.18	6.27	7.34	7.98	8.18	7.80	7.03	5.99
111	6.91	6.50	6.18	6.27	7.34	7.98	8.18	7.80	7.03
112	7.37	6.91	6.50	6.18	6.27	7.34	7.98	8.18	7.80
113	7.88	7.37	6.91	6.50	6.18	6.27	7.34	7.98	8.18
114	8.13	7.88	7.37	6.91	6.50	6.18	6.27	7.34	7.98

Problem 8. If we generate a realization of a stationary AR(2) process with initial values $z_1 = z_2 = 0$, then, after an initial transient phase, the process converges to a long-run stationary behavior in which it varies about a fixed mean value equal to ________.

a)⋆ $\dfrac{C}{1-\phi_1-\phi_2}$ b) $\dfrac{C}{1-\phi_1}$ c) $\dfrac{C}{1-\phi_1^2}$ d) C

e) 0 f) $\dfrac{\sigma_a^2}{1-\phi_1^2}$ g) $\dfrac{\sigma_a^2}{1-\phi_1^2-\phi_2^2}$ h) $\sigma_a^2(\psi_0^2+\psi_1^2+\psi_2^2)$

Problem 9. Suppose $\{z_t\}$ is a stationary ARMA(p, q) process, and let $\tilde{z}_t$ be the mean-centered process. Then $E\tilde{z}_t =$ _________.

a) μ_z **b)** $\tilde{C}$ **c)** $\dfrac{\tilde{C}}{1 - \tilde{\phi}_1 - \tilde{\phi}_2 - \ldots - \tilde{\phi}_p}$ **d)** $\dfrac{\tilde{C}}{1 - \tilde{\theta}_1 - \tilde{\theta}_2 - \ldots - \tilde{\theta}_q}$

e)$\star$ 0 **f)** C **g)** $\dfrac{C}{1 - \phi_1 - \phi_2 - \ldots - \phi_p}$ **h)** $\dfrac{C}{1 - \theta_1 - \theta_2 - \ldots - \theta_q}$

Problem 10. An AR(2) process will be stationary if and only if all of the conditions _________ are true.

a)$\star$ $|\phi_2| < 1,\ \phi_2 + \phi_1 < 1,\ \phi_2 - \phi_1 < 1$ **b)** $|\phi_1| < 1,\ \phi_1 + \phi_2 < 1,\ \phi_1 - \phi_2 < 1$

c) $|\phi_2| > 1,\ \phi_2 + \phi_1 > 1,\ \phi_2 - \phi_1 > 1$ **d)** $|\phi_1| > 1,\ \phi_1 + \phi_2 > 1,\ \phi_1 - \phi_2 > 1$

e) $|\phi_2| < 1,\ |\phi_2 + \phi_1| < 1,\ |\phi_2 - \phi_1| < 1$ **f)** $|\phi_1| < 1,\ |\phi_1 + \phi_2| < 1,\ |\phi_1 - \phi_2| < 1$

g) $|\phi_2| > 1,\ |\phi_2 + \phi_1| > 1,\ |\phi_2 - \phi_1| > 1$ **h)** $|\phi_1| > 1,\ |\phi_1 + \phi_2| > 1,\ |\phi_1 - \phi_2| > 1$

Problem 11. Approximate 95% confidence intervals for a regression coefficient β_i use the value 1.96. If the regression assumptions are valid, an **exact** 95% confidence interval may be constructed by replacing 1.96 by _________

a) 2
b) 1.959964
c) 1.644854
d) 2.326348
e) a value obtained from the normal distribution
f) a value obtained from a chi-square distribution
g)$\star$ a value obtained from a t-distribution

Problem 12. For any stationary ARMA(p, q) process, z_t will be independent of **both** _________.

a) a_t and a_{t+2} **b)**$\star$ a_{t+1} and a_{t+3} **c)** a_t and a_{t-2}

d) a_{t-2} and a_{t+2} **e)** a_{t-1} and a_{t-3} **f)** a_{t-1} and a_{t+1}

Problem 13. Suppose $\{z_t\}$ is a stationary time series, and you are trying to predict z_t from the earlier values of the series by using a regression model:

$$z_t = \beta_0 + \beta_1 z_{t-1} + \beta_2 z_{t-2} + \cdots + \beta_m z_{t-m} + \varepsilon \,.$$

You are trying to choose m. If _________, then m is a reasonable choice.

a) If the sample ACF value r_m is **small**

b) If the p-value for testing $H_0 : \beta_m = 0$ is **large**

c) If the sample PACF $\hat{\phi}_{kk}$ is **large** for all lags k greater than m

d) If the sample ACF r_k is **large** for all lags k greater than m

e) If the sample PACF value $\hat{\phi}_{mm}$ is **small**

f) If the t-Value for testing $H_0 : \beta_m = 0$ is **small**

g)$\star$ If the sample PACF $\hat{\phi}_{kk}$ is **small** for all lags k greater than m

h) If the sample ACF r_k is **small** for all lags k greater than m

Problem 14. Suppose you fit an AR(2) model to a time series using PROC ARIMA. A piece of the estimation output you get is given below. This output contains an estimate of σ_a which is equal to _________.

a) 31.28354 **b)** 2.00188 **c)** 1.62627 **d)** 0.05263

e) -0.85162 **f)** 0.05332 **g)** 7.049993 **h)** 21.00325

i)$\star$ 4.58293 **j)** 595.2691 **k)** 603.0846

Maximum Likelihood Estimation

Parameter	Estimate	Standard Error	t Value	Approx Pr > |t|	Lag
MU	31.28354	2.00188	15.63	<.0001	0
AR1,1	1.62627	0.05263	30.90	<.0001	1
AR1,2	-0.85162	0.05332	-15.97	<.0001	2

Constant Estimate	7.049993
Variance Estimate	21.00325
Std Error Estimate	4.58293
AIC	595.2691
SBC	603.0846
Number of Residuals	100

Problem 15. Suppose that $\{z_t\}$ is a stationary process with mean μ_z and $\{a_t\}$ is a random shock sequence. The value of

$$E\left(7 + 2z_t + 3z_{t-1} + a_t + 4a_{t-2}\right)$$

is equal to ________.

a) $13\sigma_z^2 + 12$ **b)** 17 **c)** 12 **d)** $30\sigma_a^2$
e)$\star$ $7 + 5\mu_z$ **f)** $49 + 13\mu_z$ **g)** $12 + 5\mu_z$ **h)** $13\sigma_z^2 + 17\sigma_a^2$

Problem 16. Which **one** of the following processes is **stationary**?

a) $z_t = -1 + z_{t-1} + a_t - 0.4a_{t-1} - 0.3a_{t-2}$
b) $z_t = 2 - 1.2z_{t-1} + a_t$
c) $z_t = 0.8z_{t-1} + 0.6z_{t-2} + a_t$
d)$\star$ $z_t = 7 + 1.5z_{t-1} - 0.9z_{t-2} + a_t - 1.3a_{t-1}$
e) $z_t = 3 - 0.4z_{t-1} + 0.8z_{t-2} + a_t - 0.5a_{t-1}$

Problem 17. For a particular time series, the MINIC option in the IDENTIFY statement produced the table below. Based on this table, what would be a reasonable tentative model choice for this time series?

a) ARMA(4,4) **b)**$\star$ ARMA(4,2) **c)** ARMA(5,5) **d)** ARMA(5,1)
e) ARMA(5,3) **f)** ARMA(0,0) **g)** ARMA(0,1) **h)** ARMA(1,0)

Minimum Information Criterion						
Lags	**MA 0**	**MA 1**	**MA 2**	**MA 3**	**MA 4**	**MA 5**
AR 0	7.1511	7.0691	7.011455	6.944779	6.868762	6.837542
AR 1	5.186998	5.026385	5.032009	5.036951	5.038342	5.014086
AR 2	4.271171	4.259883	3.9342	3.799679	3.764675	3.624006
AR 3	4.272556	4.218942	3.844689	3.807723	3.723637	3.627763
AR 4	3.736673	3.577102	3.526859	3.537519	3.546234	3.554121
AR 5	3.704242	3.548928	3.536487	3.547038	3.555908	3.564622

Problem 18. Suppose we have collected data on a response variable Y and p covariates $X_1, X_2, \ldots, X_p$ for n observations (cases), and we plan to fit a linear regression of Y on $X_1, X_2, \ldots, X_p$. Consider the following list of statements:

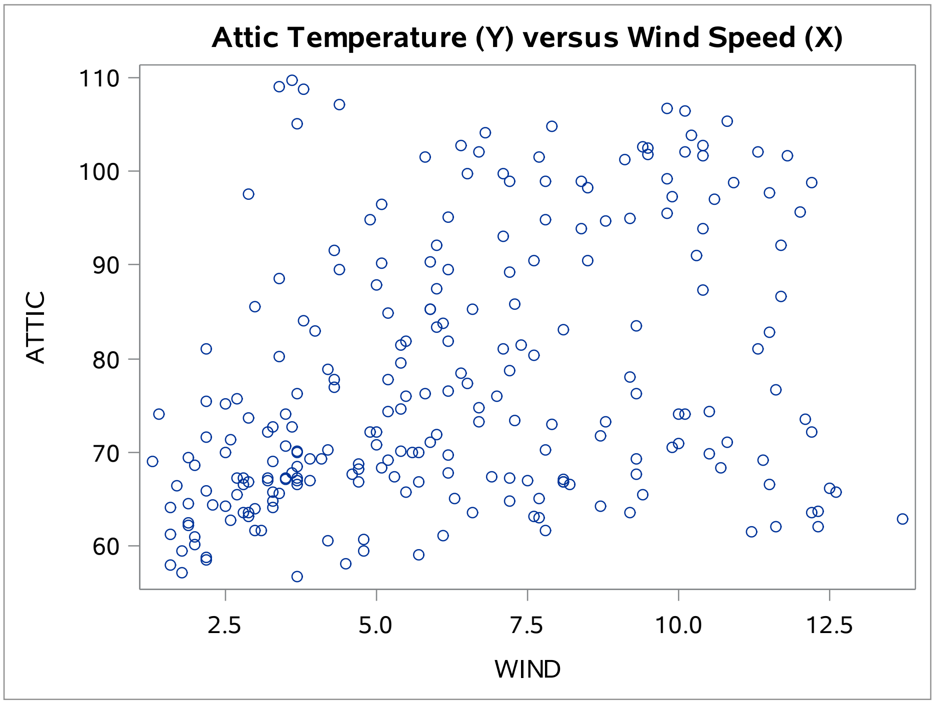
1. $Y_i = \beta_0 + \beta_1 X_{1,i} + \beta_2 X_{2,i} + \cdots \beta_p X_{p,i} + \varepsilon_i\,, \quad i = 1, \ldots, n.$
2. The errors $\varepsilon_1, \ldots, \varepsilon_n$ each have a $N(0, \sigma^2)$ distribution.
3. The errors $\varepsilon_1, \ldots, \varepsilon_n$ are independent of each other.
4. The errors $\varepsilon_1, \ldots, \varepsilon_n$ are independent of the covariates.
5. The Error Mean Square $\hat{\sigma}^2$ is an unbiased estimator of σ^2.
6. The parameter estimate $\hat{\beta}_i$ is an unbiased estimator of β_i.
7. When n is large enough, $\hat{\beta}_i \approx N(\beta_i, \mathrm{SE}(\hat{\beta}_i)^2)$.
8. None of the observations (cases) have a large value of the Studentized Residual (RStudent).
9. None of the observations have a large Leverage (H) value.
10. None of the observations have a large Cook's D.

Which of these statements are **assumptions** of the linear regression model?

a)$\star$ 1, 2, 3, 4 **b)** 5, 6, 7
c) 8, 9, 10 **d)** 1, 2, 3, 4 and 5, 6, 7
e) 1, 2, 3, 4 and 8, 9, 10 **f)** 5, 6, 7 and 8, 9, 10
g) all of them **h)** none of them

Problem 19. Attic temperature and wind speed have been observed at hourly intervals for 240 consecutive hours. The graph given on the next page is a scatter plot of the attic temperature (Y) versus the wind speed (X). What would you expect **solely on the basis of this scatter plot?**

a) The ACF of X will decay to zero slowly.
b) The ACF of X will decay to zero rapidly.
c)$\star$ In a regression of Y on X, the coefficient of X will be small and positive.
d) In a regression of Y on X, the coefficient of X will be large and negative.
e) The ACF of Y will decay to zero slowly.
f) The ACF of Y will decay to zero rapidly.
g) The time series plot of X will exhibit a strong daily (24 hour) pattern.
h) The PACF of Y will cutoff to zero.
i) The PACF of X will cutoff to zero.



Problem 20. Let $r_1, r_2, r_3, \ldots$ be the sample autocorrelations of a stationary time series of length n. The statistic

$$Q(m) = n(n+2)\sum_{k=1}^{m} \frac{r_k^2}{n-k}$$

is used to test _________.

a)$\star$ $H_0 : \rho_1 = \rho_2 = \ldots = \rho_m = 0$

b) $H_0 : \rho_m = 0$

c) $H_0 : \phi_{mm} = 0$

d) $H_0 : \theta_m = 0$

e) $H_0 : \phi_m = 0$

f) the normality of the residuals

g) whether the variance is constant

h) whether the ACF is constant

Problem 21. Under the appropriate null hypothesis, the statistic $Q(m)$ defined in the previous problem will have approximately a _________.

a) $N(0, 1)$ distribution

b) $N(\mu_z, \sigma_z^2)$ distribution

c) $N(0, \sigma_a^2)$ distribution

d)$\star$ χ_m^2 distribution

e) t-distribution with $n - m - 1$ degrees of freedom

f) $F_{m,n}$ distribution

Problem 22. A billion realizations of an AR(1) process $\{z_t\}$ with $C = \mathbf{20}$, $\phi_1 = \mathbf{0.7}$, $\sigma_a = \mathbf{6.0}$ were generated, each of length 100 with initial value $z_1 = \mathbf{0}$. Consider this collection of one billion time series to be like a population of one billion individuals. Each of the billion time series has a pair of values (z_{20}, z_{19}), so you can calculate the population correlation $\text{Corr}(z_{20}, z_{19})$ of these one billion pairs of values. Similarly, you can calculate $\text{Corr}(z_t, z_{t-1})$ for any time $t = 3, 4, 5, \ldots, 100$. If you were to plot these values of $\text{Corr}(z_t, z_{t-1})$ versus time t, you would see that the plotted values __________.

- **a)** converge to a limiting value of **zero** after an initial transient phase
- **b)**⋆ converge to a limiting value of **0.7** after an initial transient phase
- **c)** converge to a limiting value of **20** after an initial transient phase
- **d)** are all exactly equal to **zero** since the process is stationary
- **e)** are all exactly equal to **0.7** since the process is stationary
- **f)** are all exactly equal to **20** since the process is stationary
- **g)** vary randomly about a constant mean of **zero** after an initial transient phase
- **h)** vary randomly about a constant mean of **0.7** after an initial transient phase
- **i)** vary randomly about a constant mean of **20** after an initial transient phase

The following information applies to the next 2 problems:

SAS PROC ARIMA gives a plot of the sample ACF along with a band. For an MA(**4**) process, which choices make the following statement true:

The spike in the sample ACF at $\boxed{}$ will lie **inside** the band about $\boxed{}$ of the time.

Problem 23. Choices for Box 1: (Circle the correct choice)

a) lag 1 **b)** lag 2 **c)** lag 3 **d)** lag 4 **e)**⋆ lag 5

Problem 24. Choices for Box 2: (Circle the correct choice)

a) 1% **b)** 2% **c)** 5% **d)** 10% **e)** 90% **f)**⋆ 95% **g)** 98% **h)** 99%

Problem 25. Suppose z_t is an AR(2) process:

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

What is the correlation between z_{t-2} and a_t?

a) 1 **b)** -1 **c)**⋆ 0 **d)** ϕ_2 **e)** ϕ_{22} **f)** ρ_2 **g)** ϕ_1^2 **h)** ϕ_2^2

Problem 26. Suppose you have a time series $z_1, z_2, \ldots, z_n$ which is a very long (i.e., n is very large) realization of a stationary ARMA process. You fit a regression of z_t on $z_{t-1}, z_{t-2}, z_{t-3}, z_{t-4}$ by creating a SAS data set containing `z` (the time series $\{z_t\}$) and the lagged variables `zlag1, zlag2, zlag3, zlag4` and then using SAS PROC REG to fit the regression model

$$\texttt{z} = \beta_0 + \beta_1\,\texttt{zlag1} + \beta_2\,\texttt{zlag2} + \beta_3\,\texttt{zlag3} + \beta_4\,\texttt{zlag4} + \varepsilon$$

and obtain the estimates $\hat\beta_1, \hat\beta_2, \hat\beta_3, \hat\beta_4$. After this, you use SAS PROC ARIMA to find the sample PACF $\hat\phi_{11}, \hat\phi_{22}, \ldots$ of the series $\{z_t\}$. When you compare $\hat\beta_1, \hat\beta_2, \hat\beta_3, \hat\beta_4$ and $\hat\phi_{11}, \hat\phi_{22}, \hat\phi_{33}, \hat\phi_{44}$, what should you expect to see? (Choose the **single** best response. Note that $x \approx y$ means "x and y are nearly equal".)

a) $\hat\beta_1 \approx \hat\phi_{11}$ **b)** $\hat\beta_2 \approx \hat\phi_{22}$ **c)** $\hat\beta_3 \approx \hat\phi_{33}$ **d)**$\star$ $\hat\beta_4 \approx \hat\phi_{44}$
e) All of (a) through (d) **f)** None of (a) through (d)

Problem 27. To simulate a realization from an AR(3) process, we need k starting values $z_1, z_2, \ldots, z_k$. What is the value of k?

a) 0 **b)** 1 **c)** 2 **d)**$\star$ 3 **e)** 4 **f)** 5 **g)** 6

The next three questions use the SAS output described below which is given on the following three pages.

A simulated data set has 103 observations and four variables, a response variable Y and three covariates X1, X2, X3. SAS PROC REG was used to fit a regression of Y on all three covariates X1, X2, X3. The output on the following pages contains: (1) a matrix of scatterplots for the covariates X1, X2, X3; (2) some of the plots produced by SAS PROC REG displaying the case diagnostics and other items; and (4) a printout of the first 24 observations giving the values of Y, X1, X2, X3, the predicted (fitted) values, residuals, Cook's D, Leverage, and RStudent.

Three unusual observations (i.e., rows) have been planted in this data set.

Problem 28. Two of the unusual observations are easily visible (i.e., they stick way out) in some of the pairwise scatterplots of X1, X2, X3, and these observations have been circled in those plots. What are these observations? (Select the correct pair of observation numbers.)

a) 4, 6 **b)** 4, 11 **c)** 4, 22 **d)** 4, 24 **e)** 6, 11
f)$\star$ 6, 22 **g)** 6, 24 **h)** 11, 22 **i)** 11, 24 **j)** 22, 24

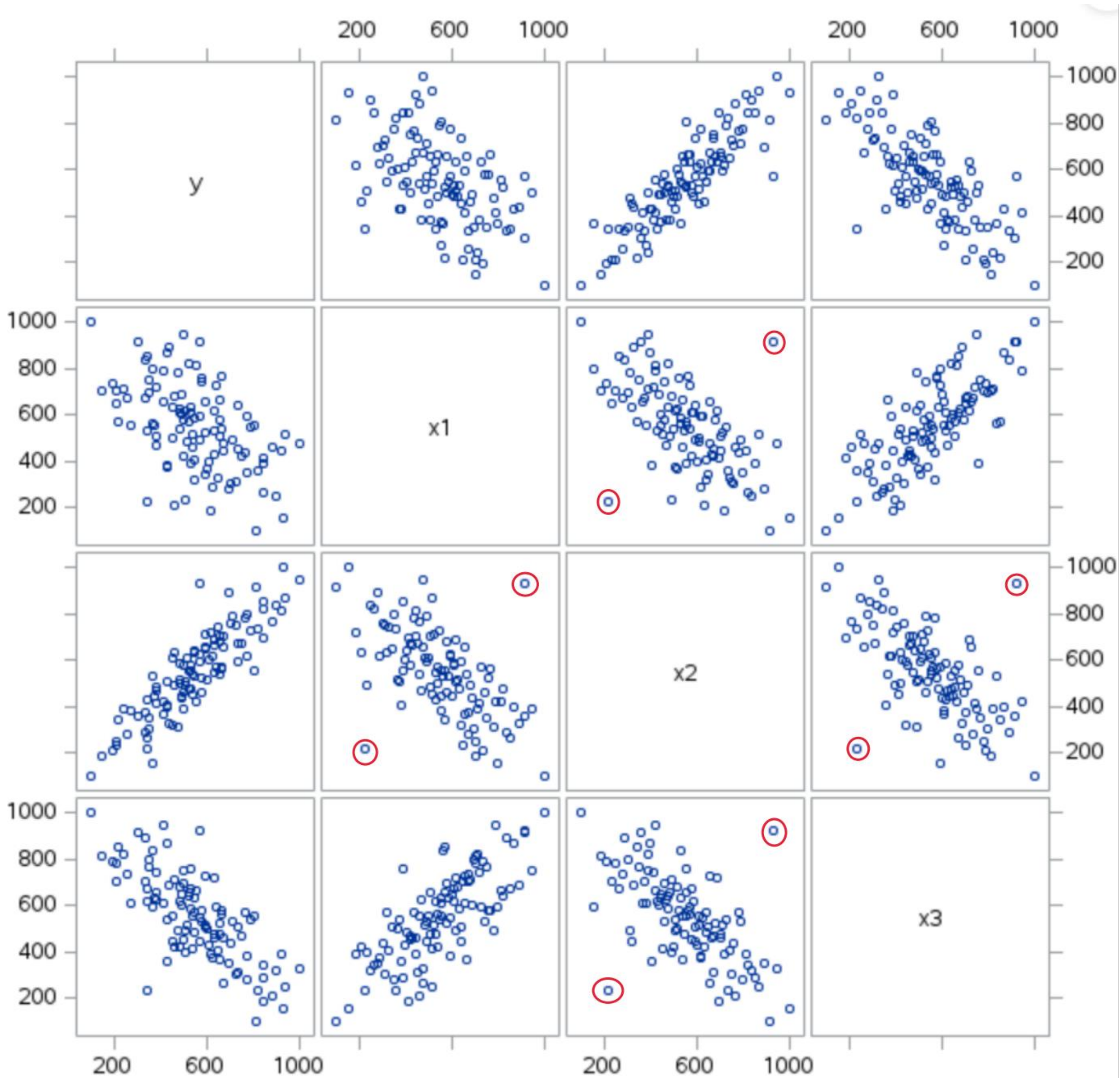
Problem 29. One of the unusual points has a much greater effect on the regression model (such as on the the estimated parameters and predicted values) than the other two. Which observation is this?

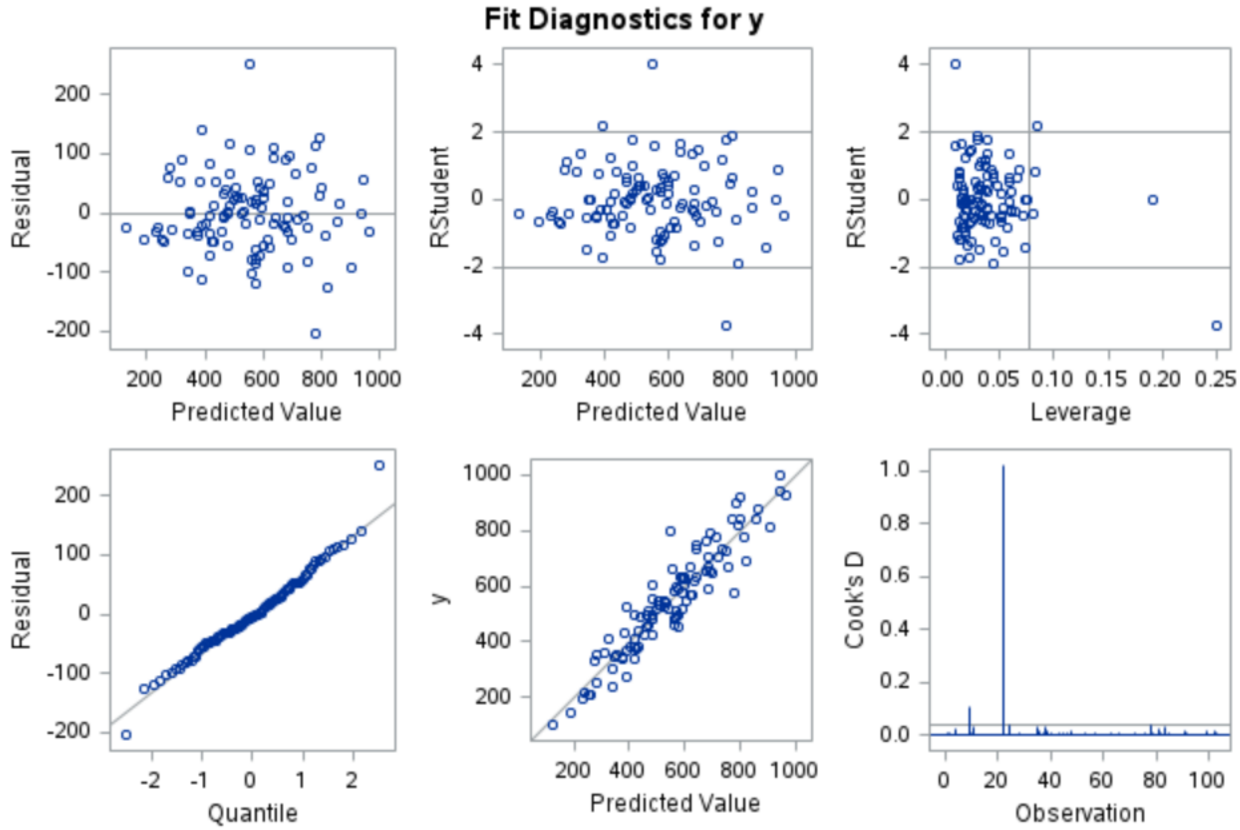
a) 4 **b)** 6 **c)** 8 **d)** 11 **e)** 14
f) 15 **g)** 17 **h)** 20 **i)**$\star$ 22 **j)** 24

Problem 30. Two of the three points have unusual response values. Which observations are they?

a) 4, 6 **b**) 4, 11 **c**) 4, 22 **d**) 4, 24 **e**) 6, 11

f) 6, 22 **g**) 6, 24 **h**)$\star$ 11, 22 **i**) 11, 24 **j**) 22, 24




Fit Diagnostics for y
Residual
Predicted Value
RStudent
Predicted Value
RStudent
Leverage
Residual
Quantile
y
Predicted Value
Cook's D
Observation

Obs	y	x1	x2	x3	predict	resid	cookd	leverage	rstudent
1	145	706	186	809	188.814	-43.814	0.00601	0.05130	-0.66516
2	788	547	727	538	691.011	96.989	0.01303	0.02400	1.46421
3	672	447	706	458	681.454	-9.454	0.00008	0.01640	-0.14065
4	922	441	809	386	796.398	125.602	0.02740	0.02974	1.91623
5	521	566	552	658	500.065	20.935	0.00042	0.01669	0.31164
6	346	229	219	236	346.263	-0.263	0.00000	0.19064	-0.00432
7	566	591	626	547	624.195	-58.195	0.00282	0.01469	-0.86830
8	519	822	473	641	539.006	-20.006	0.00093	0.03920	-0.30127
9	528	393	557	759	389.479	138.521	0.10650	0.08460	2.18762
10	211	647	231	700	259.803	-48.803	0.00594	0.04168	-0.73756
11	802	554	552	554	550.477	251.523	0.03449	0.00972	4.02567
12	458	681	494	713	467.317	-9.317	0.00009	0.01815	-0.13875
13	365	561	528	835	386.791	-21.791	0.00184	0.06188	-0.33213
14	411	658	364	605	415.940	-4.940	0.00003	0.02142	-0.07368
15	607	399	600	430	597.620	9.380	0.00008	0.01695	0.13960
16	197	735	211	787	229.890	-32.890	0.00274	0.04221	-0.49646
17	379	504	484	652	428.931	-49.931	0.00330	0.02297	-0.74741
18	594	340	647	508	571.851	22.149	0.00087	0.03031	0.33205
19	546	586	533	636	504.106	41.894	0.00125	0.01261	0.62326
20	702	301	759	435	682.542	19.458	0.00074	0.03323	0.29210
21	617	186	720	393	633.846	-16.846	0.00095	0.05455	-0.25570
22	573	917	929	918	777.829	-204.829	1.02222	0.24958	-3.72769
23	843	386	850	290	858.895	-15.895	0.00059	0.03937	-0.23934
24	812	100	913	100	905.819	-93.819	0.04151	0.07362	-1.45357
25	303	913	356	913	338.302	-35.302	0.00402	0.05262	-0.53589

In the remaining problems, you are asked to use the SAS output in the following three pages to identify three time series (z1, z2, z3) as one of the following processes: RS (= Random Shocks), AR(1), AR(2), AR(3), MA(1), MA(2), MA(3), ARMA(1,1).

Problem 31. What is z1?

a) RS **b)** AR(1) **c)** AR(2) **d)** AR(3)

e) MA(1) **f)** MA(2) **g)** MA(3) **h)**$\star$ ARMA(1,1)

Problem 32. What is z2?

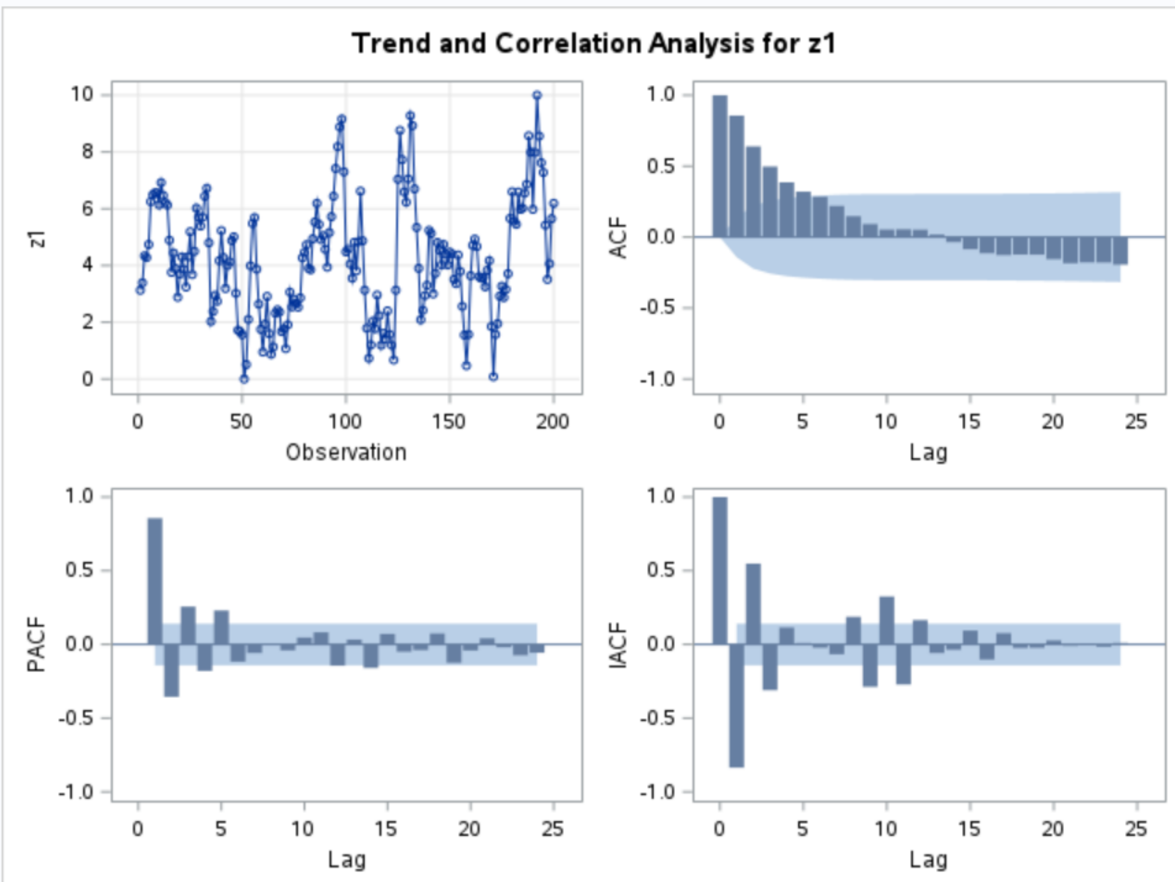
a) RS	**b)** AR(1)	**c)** AR(2)	**d)** AR(3)
e) MA(1)	**f)** MA(2)	**g)**$\star$ MA(3)	**h)** ARMA(1,1)

Problem 33. What is z3?

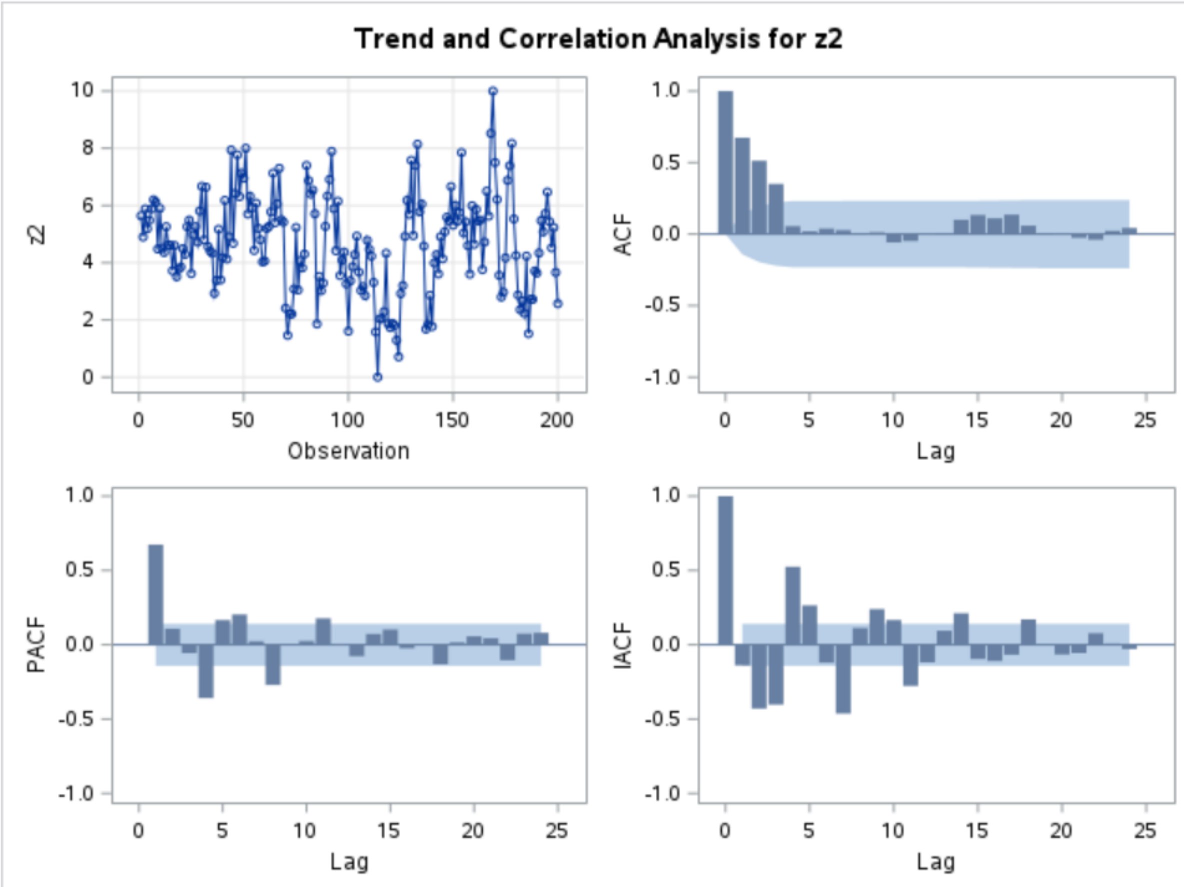
a) RS	**b)** AR(1)	**c)**$\star$ AR(2)	**d)** AR(3)
e) MA(1)	**f)** MA(2)	**g)** MA(3)	**h)** ARMA(1,1)

Autocorrelation Check for White Noise

To Lag	Chi-Square	DF	Pr > ChiSq	Autocorrelations					
6	352.18	6	<.0001	0.857	0.639	0.498	0.386	0.321	0.285
12	370.48	12	<.0001	0.220	0.147	0.092	0.053	0.055	0.052
18	382.13	18	<.0001	0.018	-0.035	-0.085	-0.114	-0.127	-0.123
24	421.77	24	<.0001	-0.123	-0.155	-0.184	-0.178	-0.179	-0.194



Autocorrelation Check for White Noise									
To Lag	Chi-Square	DF	Pr > ChiSq	Autocorrelations					
6	172.25	6	<.0001	0.674	0.513	0.351	0.055	0.019	0.038
12	173.67	12	<.0001	0.028	0.007	0.013	-0.058	-0.048	-0.000
18	187.30	18	<.0001	0.007	0.100	0.133	0.111	0.136	0.059
24	188.34	24	<.0001	0.006	0.002	-0.026	-0.039	0.023	0.043



Autocorrelation Check for White Noise									
To Lag	Chi-Square	DF	Pr > ChiSq	Autocorrelations					
6	349.58	6	<.0001	0.607	0.678	0.522	0.486	0.450	0.407
12	474.43	12	<.0001	0.352	0.369	0.298	0.315	0.267	0.266
18	607.08	18	<.0001	0.273	0.330	0.330	0.342	0.323	0.303
24	665.40	24	<.0001	0.238	0.262	0.224	0.203	0.164	0.122