TEST #2

STA 4853 **Name:**______________________________________

May 2, 2019

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

Directions

- This exam is **closed book** and **closed notes**.
- There are **34** multiple choice questions.
- 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.)
- There is no penalty for guessing.
- The exam has **14** pages.
- Each question is worth equal credit.

Problem 1. Suppose X, Y, Z are random variables, a, b, c are constants, and $\widehat{X} = E(X \mid \mathcal{I})$, $\widehat{Y} = E(Y \mid \mathcal{I})$, $\widehat{Z} = E(Z \mid \mathcal{I})$ for some information $\mathcal{I}$. If $X = aY + bZ + c$, then $\widehat{X} =$ ________.

a) $aY + bZ + c$ **b)** $\widehat{a}\,Y + \widehat{b}\,Z$ **c)** $a^2\widehat{Y} + b^2\widehat{Z}$ **d)** $\widehat{a^2}\,Y + \widehat{b^2}\,Z$

e) $a^2Y + b^2Z$ **f)** $a^2Y + b^2Z + c^2$ **g)**⋆ $a\widehat{Y} + b\widehat{Z} + c$ **h)** $\widehat{a}\,Y + \widehat{b}\,Z + \widehat{c}$

Problem 2. Suppose you are trying to find a good time series model for a series x_t. You first decide to use a log transformation $y_t = \log(x_t)$, and then you decide to difference the transformed series two times, obtaining the series $z_t = \nabla^2 y_t$. You then observe that the IACF of z_t is slowly decaying. This suggests ________.

a) trying a different transformation
b)⋆ differencing only once
c) differencing three times
d) not transforming the series at all
e) dropping the constant from the model
f) adding a seasonal term

Problem 3. Suppose you have a time series z_t with monthly data. You fit an ARIMA(2, 1, 0) model to this series. In this model, you retain the constant term C which is significant. For this series, if you use this model to compute forecasts for 36 months into the future, you expect these forecasts to ________.

a) exactly lie on a straight line with a nonzero slope
b) exactly lie on a straight line with a slope of zero
c) converge to a repeating seasonal pattern
d) exactly follow a repeating seasonal pattern
e) converge to a repeating seasonal pattern added to a straight line with a nonzero slope
f)⋆ converge to a straight line with a nonzero slope
g) converge to a straight line with a slope of zero

Problem 4. Suppose we have a transfer function model for explaining a response series $\{Y_t\}$ in terms of an input series $\{X_t\}$. If the noise process in this model is an ARIMA$(p, 1, q)$ process without a constant term, then the transfer function model can be written as ________.

a)⋆ $Y_t = C + v(B)X_t + \dfrac{\theta(B)}{\phi(B)(1-B)}a_t$ **b)** $Y_t = C + v(B)X_t + \dfrac{\theta(B)(1-B)}{\phi(B)}a_t$

c) $Y_t = C + v(B)X_t + \dfrac{\theta(B)}{\phi(B)}a_t$ **d)** $(1-B)Y_t = C + v(B)X_t + \dfrac{\theta(B)}{\phi(B)}a_t$

e) $(1-B)Y_t = C + v(B)X_t + \dfrac{\theta(B)(1-B)}{\phi(B)}a_t$ **f)** $Y_t = C + v(B)(1-B)X_t + \dfrac{\theta(B)(1-B)}{\phi(B)}a_t$

Problem 5. An AR(2)$_{12}$ or ARIMA$(2,0,0)_{12}$ is a purely seasonal model. It may be written as ________.

a) $z_t = C + \Phi_1 z_{t-1} + \Phi_2 z_{t-12} + a_t$

b) $z_t = C + \Phi_1 z_{t-12} + \Phi_2 z_{t-13} + a_t$

c) $z_t = C + \Phi_1 z_{t-12} + \Phi_2 z_{t-24} + a_{t-24}$

d) $z_t = C + \Phi_1 z_{t-1} + \Phi_2 z_{t-12} + a_{t-12}$

e) $z_t = C + \Phi_1 z_{t-12} + \Phi_2 z_{t-13} + a_{t-12}$

f)$\star$ $z_t = C + \Phi_1 z_{t-12} + \Phi_2 z_{t-24} + a_t$

Problem 6. A simple approach for identifying a linear transfer function model

$$Y_t = C + v(B)X_t + N_t$$

consists of following the steps listed below in a certain order. What is the correct order?

1. Study the ACF/PACF of the residuals.
2. Use the p-values for the estimates $\hat{v}_i$ to decide which values of v_i are likely to be non-zero.
3. Identify a plausible ARMA model for the noise process.
4. Choose a relatively large value of h.
5. Fit a multiple regression of Y_t on $X_t, X_{t-1}, \ldots, X_{t-h}$ assuming ARMA errors.
6. Fit a multiple regression of Y_t on $X_t, X_{t-1}, \ldots, X_{t-h}$.

a) 3, 6, 2, 4, 1, 5

b)$\star$ 4, 6, 1, 3, 5, 2

c) 3, 4, 6, 5, 2, 1

d) 4, 3, 5, 6, 2, 1

e) 1, 2, 4, 6, 5, 3

f) 6, 5, 4, 3, 1, 2

Problem 7. If z_t is an ARIMA$(2,3,1)(1,2,3)_9$ process, then $w_t =$ ________ will be an ARMA$(2,1)(1,3)_9$ process.

a) $(1-B^2)(1-B^3)^9 z_t$

b) $(1-3B-2B^9)z_t$

c) $(1-3B-2B^2)^9 z_t$

d) $(1-B)^2(1-B^9)^3 z_t$

e) $(1-B^3)(1-B^2)^9 z_t$

f)$\star$ $(1-B)^3(1-B^9)^2 z_t$

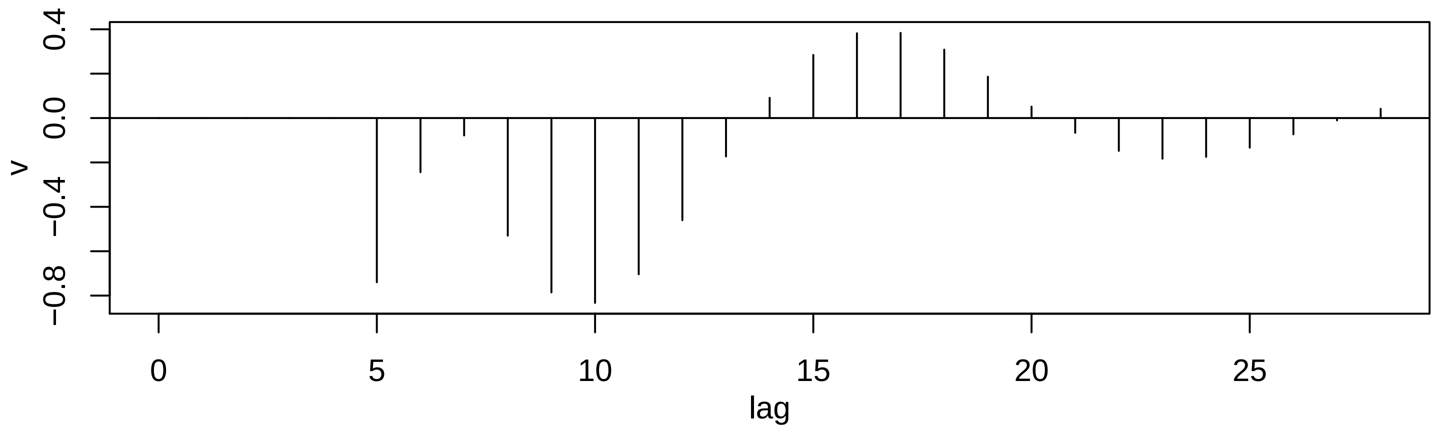
g) $(1-B^3-9B^2)z_t$

h) $(1-B^3-2B^9)z_t$

i) $(1-B-B^9)^3 z_t$

Given below is a plot of the v-weights of the transfer function

$$v(B) = \frac{B^b \omega(B)}{\delta(B)} \quad \text{where}$$
$$\omega(B) = \omega_0 - \omega_1 B - \omega_2 B^2 - \cdots - \omega_h B^h \quad \text{and}$$
$$\delta(B) = 1 - \delta_1 B - \delta_2 B^2 - \cdots - \omega_r B^r \, .$$



Based on this plot, what are plausible values of b, h, r?

Problem 8. $b =$ _________.

a) 0 **b**) 1 **c**) 2 **d**) 3 **e**)⋆ 5 **f**) 7 **g**) 10

Problem 9. $h =$ _________.

a) 0 **b**) 1 **c**)⋆ 3 **d**) 5 **e**) 10 **f**) 12 **g**) 15

Problem 10. $r =$ _________.

a) 0 **b**) 1 **c**)⋆ 2 **d**) 4 **e**) 5 **f**) 6 **g**) 10

Problem 11. If an ARMA(p, q) process is invertible, then the dual process (obtained by interchanging the roles of the AR and MA parameters) will be _________.

a) non-invertible **b**) invertible **c**) non-stationary **d**)⋆ stationary
e) non-seasonal **f**) seasonal **g**) non-normal **h**) normal

Problem 12. Suppose you have obtained a forecast for $Y = \log(X)$ and you are "back transforming" to get a forecast of $X = \exp(Y)$. If $Y \sim N(\mu, \sigma^2)$, then X has a log-normal distribution with **median** equal to ________.

a)$\star$ e^{μ} **b)** e^{σ^2} **c)** $e^{\mu+(\sigma^2/2)}$ **d)** $\log \mu$ **e)** $\log \sigma^2$ **f)** $\log \sigma$ **g)** $\log[\mu + (\sigma^2/2)]$

Problem 13. You wish to forecast a random quantity X. You have collected relevant information which does not determine X precisely, but gives you a probability distribution for the value of X. Let $\widehat{X}$ denote your forecast for X. Suppose that you will be forced to pay a penalty of $(X - \widehat{X})^2$ dollars if you guess $\widehat{X}$ and the actual value is X. Then your best forecast is the ________ of the distribution.

a) variance **b)** sigma **c)** correlation **d)** ACF
e) PACF **f)** median **g)** mode **h)**$\star$ mean

Problem 14. Suppose you wish to explain a response series $\{y_t\}$ by a regression with ARMA errors using k input series $\{x_{1,t}\}$, $\{x_{2,t}\}$, ..., $\{x_{k,t}\}$. That is, you plan to use a multiple regression model

$$y_t = \beta_0 + \beta_1 x_{1,t} + \cdots + \beta_k x_{k,t} + \varepsilon_t$$

in which you assume the error series ε_t is an ARMA(p, q) process. Which of the following is a reasonable way to make an initial choice of p and q?

a) Study the ACF and PACF of the series $x_{1,t}, x_{2,t}, \ldots, x_{k,t}$.

b) Apply the MINIC method to the series y_t.

c) Apply the MINIC method to the series $x_{1,t}, x_{2,t}, \ldots, x_{k,t}$.

d) Fit an ordinary multiple regression model and use the p-values from this to drop the non-significant terms.

e) Fit an AR(2) model to y_t and then use the residual ACF and PACF to select a better choice of p and q.

f)$\star$ Fit an ordinary multiple regression model and then study the ACF and PACF of the residuals.

g) Study the ACF and PACF of the series y_t.

Problem 15. Suppose you are given information $\mathcal{I}$ about a series Z_t consisting of all values of the series up to time 150:

$$\mathcal{I} = \{Z_{150}, Z_{149}, Z_{148}, \ldots\}$$

You are interested in making a prediction at time 160. Suppose an ARMA model for the series Z_t tells you that the conditional distribution of Z_{160} given $\mathcal{I}$ is Normal with mean $= E(Z_{160} \,|\, \mathcal{I}) = 200$ and variance $= \mathrm{Var}(Z_{160} \,|\, \mathcal{I}) = 100$. Calculate a 95% confidence interval for Z_{160}.

a)$\star$ $(180.4, 219.6)$ **b)** $(198.04, 201.96)$ **c)** $(4, 396)$
d) $(106, 304)$ **e)** $(190.6, 209.4)$ **f)** $(199.06, 201.04)$

Problem 16. If we form a new series w_t by differencing the series x_t 3 times at lag 1 and then 2 times at lag 12, we may write $w_t =$ ________.

a) $(1-3B)(1-2B^{12})x_t$ **b)** $(1-2B)(1-3B)^{12}x_t$ **c)** $(1-B^2)(1-B^3)^{12}x_t$
d) $(1-B)^{12}(1-2B)^3x_t$ **e)**$\star$ $(1-B^{12})^2(1-B)^3x_t$ **f)** $(1-B^3)(1-B^2)^{12}x_t$

Problem 17. For a stationary ARIMA$(2,0,0)(1,0,0)_{12}$ process, the last nonzero value ________.

a) in the theoretical ACF occurs at lag 2
b) in the theoretical ACF occurs at lag 3
c) in the theoretical ACF occurs at lag 12
d) in the theoretical ACF occurs at lag 13
e) in the theoretical ACF occurs at lag 14
f) in the theoretical ACF occurs at lag 24
g) in the theoretical PACF occurs at lag 2
h) in the theoretical PACF occurs at lag 3
i) in the theoretical PACF occurs at lag 12
j) in the theoretical PACF occurs at lag 13
k)$\star$ in the theoretical PACF occurs at lag 14
l) in the theoretical PACF occurs at lag 24

Problem 18. By very carefully reading the table given below, you can determine that the ESTIMATE statement which produced this table contained ________

a) `p=(1)(2,12)` **b)** `p=(1,2,12)` **c)** `p=(1,2)(12)` **d)** `p=(2)(12)`
e) `q=(1)(2,12)` **f)** `q=(1,2,12)` **g)**$\star$ `q=(1,2)(12)` **h)** `q=(2)(12)`

Maximum Likelihood Estimation					
Parameter	**Estimate**	**Standard Error**	**t Value**	**Approx Pr > $\|t\|$**	**Lag**
MU	0.27419	0.02149	12.76	<.0001	0
MA1,1	-0.33681	0.09313	-3.62	0.0003	1
MA1,2	-0.57181	0.10020	-5.71	<.0001	2
MA2,1	0.54778	0.14127	3.88	0.0001	12

Problem 19. Another example: By very carefully reading the table given below, you can determine that the ESTIMATE statement which produced this table contained ________

a) `p=(1)(2,12)` b)⋆ `p=(1,2,12)` c) `p=(1,2)(12)` d) `p=(2)(12)`
e) `q=(1)(2,12)` f) `q=(1,2,12)` g) `q=(1,2)(12)` h) `q=(2)(12)`

Maximum Likelihood Estimation					
Parameter	**Estimate**	**Standard Error**	**t Value**	**Approx Pr > $\|t\|$**	**Lag**
MU	0.28585	0.04178	6.84	<.0001	0
AR1,1	0.34024	0.11096	3.07	0.0022	1
AR1,2	0.31666	0.11044	2.87	0.0041	2
AR1,3	-0.18476	0.09540	-1.94	0.0528	12

Problem 20. Which of the following statements is true for an ARIMA$(0,0,0)(0,0,3)_{12}$ process?

a) The first three AR coefficients will be zero.
b) The ACF will decay rapidly along the early lags (1, 2, 3, ...).
c) The PACF will decay rapidly along the early lags (1, 2, 3, ...).
d) The first three MA coefficients will be zero.
e)⋆ The ACF will be zero everywhere except at lags 12, 24, 36.
f) The IACF will be zero at lags 12, 24, 36.
g) The PACF will be zero everywhere except at lags 12, 24, 36.
h) The ACF will be zero at lags 12, 24, 36.
i) The PACF will be zero at lags 12, 24, 36.

Problem 21. The SAS code given below ________

```
PROC ARIMA DATA=STUFF;
IDENTIFY VAR=Y CROSSCOR=(X1 X2 X3) NOPRINT;
ESTIMATE INPUT=(X1 X2 X3) METHOD=ML;
QUIT;
```

a) uses pre-whitening to identify a transfer function for Y_t on $X_{t-1}, X_{t-2}, X_{t-3}$.
b) fits a transfer function model for Y_t on $X_{t-1}, X_{t-2}, X_{t-3}$ using a proxy AR(2) model for the noise.
c) fits a transfer function model for Y_t on $X_{1,t}$, $X_{2,t}$, $X_{3,t}$ using a proxy AR(2) model for the noise.
d) fits a transfer function model for Y_t on $X_{t-1}, X_{t-2}, X_{t-3}$.
e) fits a transfer function model for Y_t on $X_{1,t}$, $X_{2,t}$, $X_{3,t}$.
f)⋆ fits a multiple regression model for Y_t on the regressors $X_{1,t}$, $X_{2,t}$, $X_{3,t}$.
g) fits a multiple regression model for Y_t on the regressors $X_{t-1}, X_{t-2}, X_{t-3}$.
h) uses pre-whitening to identify a transfer function for Y_t on $X_{1,t}$, $X_{2,t}$, $X_{3,t}$.

Problem 22. An MA(1) process with $|\theta_1| < 1$ can be re-written in the form ________.

a) $\tilde{z}_t = a_t - \theta_1 a_{t-1} - \theta_1^2 a_{t-2} - \theta_1^3 a_{t-3} - \cdots$

b) $\tilde{z}_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \theta_3 a_{t-3} - \cdots$

c)$\star$ $\tilde{z}_t = a_t - \theta_1 \tilde{z}_{t-1} - \theta_1^2 \tilde{z}_{t-2} - \theta_1^3 \tilde{z}_{t-3} - \cdots$

d) $\tilde{z}_t = a_t + \theta_1 \tilde{z}_{t-1} + \theta_1^2 \tilde{z}_{t-2} + \theta_1^3 \tilde{z}_{t-3} + \cdots$

e) $\tilde{z}_t = a_t + \theta_1 \tilde{z}_{t-1} + \theta_2 \tilde{z}_{t-2} + \theta_3 \tilde{z}_{t-3} + \cdots$

f) $\tilde{z}_t = a_t + \theta_1 a_{t-1} + \theta_1^2 a_{t-2} + \theta_1^3 a_{t-3} + \cdots$

Problem 23. Suppose z_t is a **quarterly** series with a **non**stationary mean. If the first differences ∇z_t appear stationary but have substantial autocorrelations at lags 1, 4, and 8, which of the following options might you wish to pursue? (More than one may be reasonable.)

1. Try a model without any differencing.
2. Try differencing at lag 1 a second time.
3. Try replacing differencing at lag 1 by differencing at lag 4.
4. Try a seasonal model on ∇z_t which includes a seasonal term at lag 4.
5. Try an MA(1) model on ∇z_t.
6. Try an AR(1) model on ∇z_t.

Select the pair of options which seem most reasonable and circle your choice **below**. (Do NOT circle items on the list above!)

a) 3 or 5 **b)** 2 or 5 **c)** 2 or 6 **d)**$\star$ 3 or 4

e) 1 or 2 **f)** 1 or 5 **g)** 1 or 6 **h)** 5 or 6

Problem 24. Suppose the time series z_t has the form $\ldots, 1, 3, 5, 7, 1, 3, 5, 7, 1, 3, 5, 7, \ldots$ and continues repeating the same pattern forever into both the future and the past. What is $\nabla_{\mathbf{4}} z_t$?

a)$\star$ $\ldots, 0, 0, 0, 0, 0, 0, 0, 0, \ldots$ **b)** $\ldots, -2, -4, -6, 0, -2, -4, -6, 0, \ldots$

c) $\ldots, 2, -2, 2, -2, 2, -2, 2, -2, \ldots$ **d)** $\ldots, 2, 4, 6, 0, 2, 4, 6, 0, \ldots$

e) $\ldots, 2, 2, 2, 2, 2, 2, 2, 2, \ldots$ **f)** $\ldots, 2, 2, 2, -6, 2, 2, 2, -6, \ldots$

g) $\ldots, -2, -2, -2, 6, -2, -2, -2, 6, \ldots$ **h)** $\ldots, -2, -2, -2, -2, -2, -2, -2, -2, \ldots$

Problem 25. A transfer function model has the general form

$$Y_t = C + v(B) X_t + N_t \,.$$

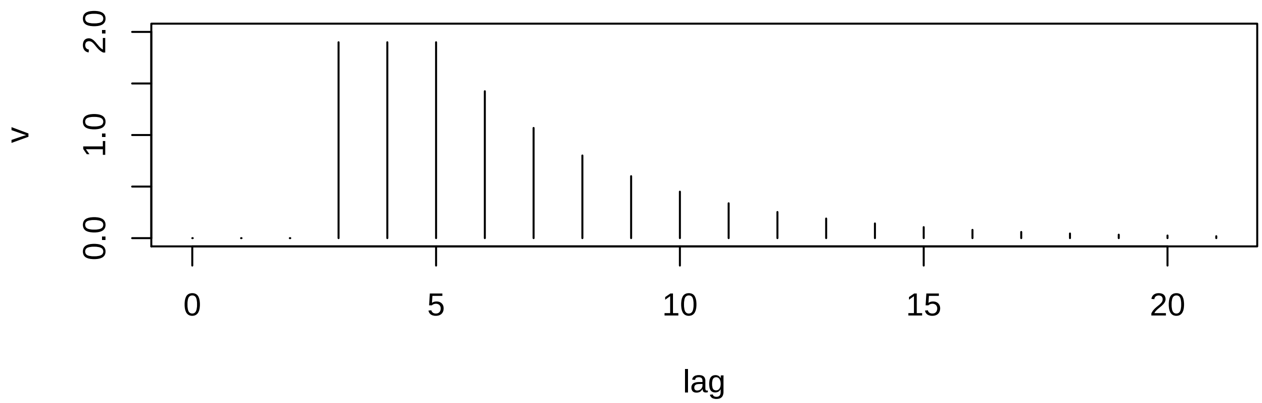
If the noise process N_t is a stationary ARMA(p, q) process with AR polynomial $\phi(B)$ and MA polynomial $\theta(B)$, then the noise process N_t can be written as ________.

a) $\dfrac{\phi(B)}{(1-B)\theta(B)} a_t$ **b)** $\dfrac{(1-B)\theta(B)}{\phi(B)} a_t$ **c)** $\dfrac{(1-B)\phi(B)}{\theta(B)} a_t$ **d)** $\dfrac{\phi(B)}{\theta(B)} a_t$

e)$\star$ $\dfrac{\theta(B)}{\phi(B)} a_t$ **f)** $\dfrac{B^b \theta(B)}{\phi(B)} a_t$ **g)** $\dfrac{B^b \phi(B)}{\theta(B)} a_t$ **h)** $\dfrac{\theta(B)}{(1-B)\phi(B)} a_t$

Problem 26. What is the value of h for the transfer function with v-weights given in the plot below?

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



Problem 27. After fitting a time series model, the plot of the residuals versus the one-step-ahead forecasts is often used to determine if the ________

a) one-step-ahead forecasts are normally distributed
b) mean of the series varies with time
c) ACF of the series varies with time
d) residuals are normally distributed
e)⋆ variability of the residuals changes with the level of the series
f) one-step-ahead forecasts change with the level of the series
g) series needs further differencing

Problem 28. For a time series with a non-stationary mean, it is common for the **P**ACF to ________

a) display strong sinusoidal oscillations.
b) decay exponentially.
c) exhibit alternating exponential decay.
d) exhibit variability which changes with the level.
e) decay very slowly.
f)⋆ have a single large spike at lag 1 with $\hat{\phi}_{11}$ close to 1.
g) have several large spikes followed by an approximate cutoff to zero.

Problem 29. If you difference a series and then observe that the sample ACF of the differenced series decays very slowly, what should you do?

a) Try a transformation, perhaps a log or square root.

b) Use a mixed model with both $p > 0$ and $q > 0$.

c) Conclude that the series is non-invertible.

d)⋆ Try differencing the series again.

e) Use an MA(q) model with a large value of q.

f) Conclude that the series has been over-differenced.

Problem 30. Suppose z_t is a stationary ARMA process with psi-weights $\psi_1, \psi_2, \psi_3, \ldots$ Let $e_n(3) = z_{n+3} - \hat{z}_{n+3}$ be the 3-step-ahead forecast error based on information about the series up to time n. The forecast error $e_n(3)$ is equal to _________.

a) $a_{n+3} + \psi_1 a_{n+2} + \psi_2 a_{n+1} + \psi_3 a_n$

b) $a_n + \psi_1 a_{n-1} + \psi_2 a_{n-2} + \psi_3 a_{n-3}$

c) $\sigma_a^2(1 + \psi_1^2 + \psi_2^2)$

d) $\sigma_a^2(1 + \psi_1^2 + \psi_2^2 + \psi_3^2)$

e) $\sigma_a \sqrt{1 + \psi_1^2 + \psi_2^2}$

f) $\sigma_a \sqrt{1 + \psi_1^2 + \psi_2^2 + \psi_3^2}$

g) $a_n + \psi_1 a_{n-1} + \psi_2 a_{n-2}$

h)⋆ $a_{n+3} + \psi_1 a_{n+2} + \psi_2 a_{n+1}$

Problem 31. For an AR(1) process $z_t = C + \phi_1 z_{t-1} + a_t$, if you are given $\mathcal{I}_n$ (information up to time n), the forecast $\widehat{z}_{n+2}$ is _________

a) z_n

b) $\phi_1 z_n$

c)⋆ $C + \phi_1 C + \phi_1^2 z_n$

d) $C - \phi_1 a_n - \phi_1 a_{n-1}$

e) $C - \phi_1 a_n$

f) $C + z_n$

g) C

Problem 32. An ARMA$(2, 2)$ process will be **invertible** if _________.

a) $|\phi_2| < 1$

b) $|\phi_2| < 1$, $\phi_2 + \phi_1 < 1$, and $\phi_2 - \phi_1 < 1$

c)⋆ $|\theta_2| < 1$, $\theta_2 + \theta_1 < 1$, and $\theta_2 - \theta_1 < 1$

d) it can be written as an MA(∞) process

e) the long run mean is constant

f) $|\theta_1| < 1$

g) $|\phi_1| < 1$

h) $|\theta_2| < 1$

Problem 33. For **non**-stationary ARIMA processes, as you forecast further and further into the future, the confidence interval **widths** for the forecasts _________.

a) converge to a repetitive pattern added to a straight line with nonzero slope

b) converges to a straight line with a nonzero slope

c) converge to a limiting value

d) converge to a repetitive pattern which repeats with a period of S (= the seasonality)

e)⋆ continue to gradually increase and will eventually reach arbitrarily large values

Problem 34. Attached to the end of the exam are three pages with the title "Choosing a Seasonal Model". These pages give the sample ACF, PACF, and IACF for a stationary time series with monthly data (seasonality $s = 12$). Select a reasonable seasonal ARIMA model for this series from the list below. (Circle the best choice.)

a) $(0,0,5)(0,0,4)_{12}$ **b**) $(0,0,5)(4,0,0)_{12}$ **c**) $(5,0,0)(0,0,4)_{12}$ **d**) $(5,0,0)(4,0,0)_{12}$

e)⋆ $(1,0,0)(1,0,0)_{12}$ **f**) $(1,0,0)(0,0,1)_{12}$ **g**) $(0,0,1)(1,0,0)_{12}$ **h**) $(0,0,1)(0,0,1)_{12}$

i) $(8,0,0)(4,0,0)_{12}$ **j**) $(8,0,0)(0,0,4)_{12}$ **k**) $(0,0,8)(4,0,0)_{12}$ **l**) $(0,0,8)(0,0,4)_{12}$