Birkbeck
UNIVERSITY OF LONDON

MATHEMATICS AND STATISTICS ASSIGNMENT COVER SHEET

Student to Complete

Name:	Christian Alexander Gilson
Degree Programme:	MSc. Applied Statistics
Module:	**PROBABILITY MODELS AND TIME SERIES [B.Sc./Grad. Dip]** **STOCHASTIC MODELS AND TIME SERIES [MAS Programmes]**
Lecturer:	**BROOMS**
Assignment Title/Number:	**ASSIGNMENT C [B.Sc./Grad. Dip]/ ASSIGNMENT TWO [MAS]**
Due Date:	**Thursday, 30^{th} April, 2020**

Declaration

I herby testify that, unless otherwise acknowledged, the work submitted here is entirely my own

Signature:	

LEAVE BLANK – OFFICIAL USE ONLY
Date Received:

(Unscaled) Mark/Grade	Late – up to 14 days	Late – more than 14 days

Christian Gilson - 30 April 2020 Time Series Assignment.

Deadline: Thursday, 30th April, 2020

Total marks: [25]. Marks are shown in boxes [ ]. There are 2 questions in this assignment.

1. Consider a process $\{Y_t\}$ that satisfies the model equation

AR(3) model
$$Y_t = \frac{21}{5}Y_{t-1} - \frac{19}{5}Y_{t-2} + \frac{3}{5}Y_{t-3} + \epsilon_t, \qquad t \in \mathbb{Z}.$$

   (a) Find the equation satisfied by the process, $\{W_t\}$, of first differences, where $W_t = Y_t - Y_{t-1}$. [2]

   (b) Is $\{W_t\}$ a stationary process? [3]

Question 1

1)a) Subtract Y_{t-1} from both sides

$$W_t = Y_t - Y_{t-1} = \frac{21}{5}Y_{t-1} - \frac{5}{5}Y_{t-1} - \frac{19}{5}Y_{t-2} + \frac{3}{5}Y_{t-3} + \varepsilon_t$$

$$= \frac{16}{5}Y_{t-1} - \frac{16}{5}Y_{t-2} - \frac{3}{5}Y_{t-2} + \frac{3}{5}Y_{t-3} + \varepsilon_t$$

$$= \frac{16}{5}(Y_{t-1} - Y_{t-2}) - \frac{3}{5}(Y_{t-2} - Y_{t-3}) + \varepsilon_t$$

Clearly, if $W_t = Y_t - Y_{t-1}$, then $W_{t-1} = Y_{t-1} - Y_{t-2}$, and $W_{t-2} = Y_{t-2} - Y_{t-3}$

$$\therefore \boxed{W_t = \frac{16}{5}W_{t-1} - \frac{3}{5}W_{t-2} + \varepsilon_t}$$

$\hookrightarrow W_t = Y_t - Y_{t-1} = Y_t - LY_t = (1-L)Y_t = \Delta Y_t.$

b) Necessary and sufficient condition for stationary process:

- All roots of AR characteristic eq. $\phi(z)$ are greater than unit length in the complex plane.

AR characteristic polynomial equation:

$$\phi(z) = 1 - \phi_1 z - \phi_2 z^2 - \dots - \phi_p z^p = 0$$

$$\phi_1 = \frac{16}{5}, \quad \phi_2 = \frac{-3}{5}$$

[Subbing ϕ_1 and ϕ_2 from $\{W_t\}$ equation into AR characteristic poly]

$$1 - \frac{16}{5}z - \left(\frac{-3}{5}\right)z^2 = 0$$

① $\frac{3}{5}z^2 - \frac{16}{5}z + 1 = 0$

[Multiply ① through by 5]

② $3z^2 - 16z + 5 = 0$

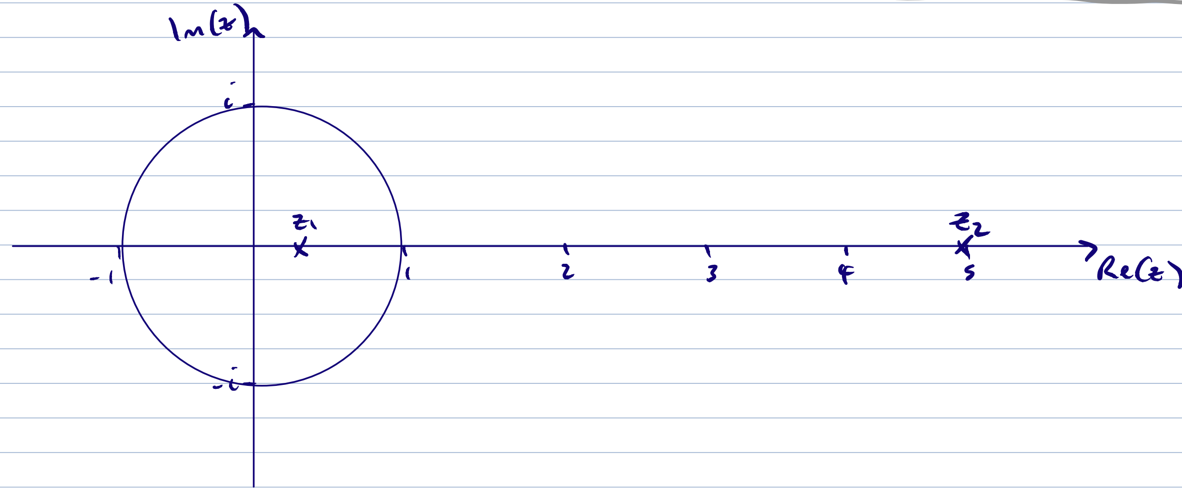
$(3z - 1)(z - 5) = 0$ Check: $3z^2 - 15z - z + 5 = 0$ ✓

$\therefore 3z_1 = 1$, $z_2 = 5$

$z_1 = \frac{1}{3}$

So, we have 2 real roots: $z_1 = \frac{1}{3}$ and $z_2 = 5$.

As not ALL roots are **greater than 1**, $\{W_t\}$ is NOT a stationary process.



2. Let $\{Z_t\}$ be an ARIMA(1,0,3) process which satisfies the relation = ARMA(1,3)

$$(1-\phi L)Z_t = \epsilon_t + \theta_1\epsilon_{t-1} + \theta_2\epsilon_{t-2} + \theta_3\epsilon_{t-3}, \qquad t \in \mathbb{Z},$$

where L is the *lag operator* such that $LZ_t = Z_{t-1}$, $\{\epsilon_t\}$ is a *white noise* process such that the ϵ_t are NID$(0,\sigma^2)$, ϕ is the *autoregressive* parameter, and θ_1, θ_2, and θ_3, are *moving average* parameters. Let $z_1, z_2, \ldots, z_T$ denote an observed realization of the process up to time T. Consider the forecasting problem with T the origin of the forecasts, h the lead time and $\hat{z}_T(h)$ the *corresponding minimum mean square error forecast*, for lead times $h \geq 1$.

(a) Deduce that an iteration scheme for the forecasts can be characterized by the following relations:

$$\begin{aligned}
\hat{z}_T(1) &= \phi z_T + \theta_1\epsilon_T + \theta_2\epsilon_{T-1} + \theta_3\epsilon_{T-2} \\
\hat{z}_T(2) &= \phi\hat{z}_T(1) + \theta_2\epsilon_T + \theta_3\epsilon_{T-1} \\
\hat{z}_T(3) &= \phi\hat{z}_T(2) + \theta_3\epsilon_T \\
\hat{z}_T(h) &= \phi\hat{z}_T(h-1), \qquad h \geq 4.
\end{aligned}$$

[8]

Question 2a

a) Recall minimum mean square forecast equation:

$\hat{z}_T(h) = E[Z_{T+h} \mid \mathcal{H}_T]$

Plan: ① First we want to express the ARMA(1,0,3) process in terms of Z_{T+h}

② Apply minimum mean square forecast equation

③ produce expressions at $h=1, 2, 3, \geq 4$.

(Could probably attack this using the recursive method which also produces the forecast error terms – but think that'll be overkill for this Q)

① $(1-\phi L)Z_t = Z_t - \phi L Z_t = Z_t - \phi Z_{t-1} = \varepsilon_t + \theta_1\varepsilon_{t-1} + \theta_2\theta\varepsilon_{t-2} + \theta_3\varepsilon_{t-3}$

$Z_t = \phi Z_{t-1} + \varepsilon_t + \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2} + \theta_3\varepsilon_{t-3}$

[in terms of Z_{T+h} : $t \to T+h$]

$Z_{T+h} = \phi Z_{T+h-1} + \varepsilon_{T+h} + \theta_1\varepsilon_{T+h-1} + \theta_2\varepsilon_{T+h-2} + \theta_3\varepsilon_{T+h-3}$

[now take conditional expectation w.r.t. the observed history, $\mathcal{H}_T$]

② $$\boxed{\begin{aligned}\hat{z}_T(h) &= E[Z_{T+h} \mid \mathcal{H}_T] \\ &= E[\phi Z_{T+h-1} + \varepsilon_{T+h} + \theta_1 \varepsilon_{T+h-1} + \theta_2 \varepsilon_{T+h-2} + \theta_3 \varepsilon_{T+h-3} \mid \mathcal{H}_T]\end{aligned}}$$

③ Let's go through this by hand for $h=1,2,3,4$, but first noting what the conditional expectations will look like for the Z terms and the ε white noise terms:

$$E[Z_{T+j} \mid \mathcal{H}_T] = \begin{cases} z_{T+j}, & j \leq 0 \quad \text{(realised observation)} \\ \hat{z}_T(j), & j > 0 \quad \text{(forecast value)} \end{cases}$$

$$E[\varepsilon_{T+j} \mid \mathcal{H}_T] = \begin{cases} \varepsilon_{T+j}, & j \leq 0 \quad \text{(realised observation, NOT a random variable)} \\ 0, & j > 0 \quad (\varepsilon_{T+j} \text{ independent of } \mathcal{H}_T, \text{ therefore: } E[\varepsilon_{T+j} \mid \mathcal{H}_T] = E[\varepsilon_{T+j}], \text{ and } \varepsilon_{T+j} \text{ is a mean zero white noise term, so } E[\varepsilon_{T+j}] = 0). \end{cases}$$

<u>When $h=1$:</u>

$$\begin{aligned}\hat{z}_T(1) &= E[\phi Z_T + \varepsilon_{T+1} + \theta_1 \varepsilon_T + \theta_2 \varepsilon_{T-1} + \theta_3 \varepsilon_{T-2} \mid \mathcal{H}_T] \\ &= \phi E[Z_T \mid \mathcal{H}_T] + \cancel{E[\varepsilon_{T+1} \mid \mathcal{H}_T]} + \theta_1 E[\varepsilon_T \mid \mathcal{H}_T] \\ &\quad + \theta_2 E[\varepsilon_{T-1} \mid \mathcal{H}_T] + \theta_3 E[\varepsilon_{T-2} \mid \mathcal{H}_T]\end{aligned}$$

$$\boxed{\hat{z}_T(1) = \phi z_T + \theta_1 \varepsilon_T + \theta_2 \varepsilon_{T-1} + \theta_3 \varepsilon_{T-2}}$$

<u>when $h=2$:</u>

$$\hat{z}_T(2) = E[\phi z_{T+1} + \varepsilon_{T+2} + \theta_1\varepsilon_{T+1} + \theta_2\varepsilon_T + \theta_3\varepsilon_{T-1} \mid \mathcal{H}_T]$$

$$= \phi E[z_{T+1} \mid \mathcal{H}_T] + \cancel{E[\varepsilon_{T+2} \mid \mathcal{H}_T]} + \theta_1 \cancel{E[\varepsilon_{T+1} \mid \mathcal{H}_T]}$$

$$+ \theta_2 E[\varepsilon_T \mid \mathcal{H}_T] + \theta_3 E[\varepsilon_{T-1} \mid \mathcal{H}_T]$$

$$\boxed{\hat{z}_T(2) = \phi \hat{z}_T(1) + \theta_2\varepsilon_T + \theta_3\varepsilon_{T-1}}$$

when $h=3$:

$$\hat{z}_T(3) = E[\phi z_{T+2} + \varepsilon_{T+3} + \theta_1\varepsilon_{T+2} + \theta_2\varepsilon_{T+1} + \theta_3\varepsilon_T \mid \mathcal{H}_T]$$

$$= \phi E[z_{T+2} \mid \mathcal{H}_T] + \cancel{E[\varepsilon_{T+3} \mid \mathcal{H}_T]} + \theta_1 \cancel{E[\varepsilon_{T+2} \mid \mathcal{H}_T]}$$

$$+ \theta_2 \cancel{E[\varepsilon_{T+1} \mid \mathcal{H}_T]} + \theta_3 E[\varepsilon_T \mid \mathcal{H}_T]$$

$$\boxed{\hat{z}_T(3) = \phi \hat{z}_T(2) + \theta_3\varepsilon_T}$$

<u>when $h=4$:</u>

$$\hat{z}_T(4) = E[\phi z_{T+3} + \varepsilon_{T+4} + \theta_1\varepsilon_{T+3} + \theta_2\varepsilon_{T+2} + \theta_3\varepsilon_{T+1} \mid \mathcal{H}_T]$$

$$= \phi E[z_{T+3} \mid \mathcal{H}_T] + \cancel{E[\varepsilon_{T+4} \mid \mathcal{H}_T]} + \theta_1 \cancel{E[\varepsilon_{T+3} \mid \mathcal{H}_T]}$$

$$+ \theta_2 \cancel{E[\varepsilon_{T+2} \mid \mathcal{H}_T]} + \theta_3 \cancel{E[\varepsilon_{T+1} \mid \mathcal{H}_T]}$$

$$\boxed{\hat{z}_T(4) = \phi \hat{z}_T(3)}$$

Generalising when $h \geq 4$:

Can see how for $h \geq 4$, that all of the moving average parameters disappear, and therefore the forecast for the next (h) time step will be the current ($h-1$) forecast multiplied by the autoregressive parameter, ϕ.

$$\therefore \boxed{\hat{z}_T(h) = \phi \, \hat{z}_T(h-1) \quad , h \geq 4}$$

Summarising above results:

$$\hat{z}_T(h) = \begin{cases} \phi z_T + \theta_1 \varepsilon_T + \theta_2 \varepsilon_{T-1} + \theta_3 \varepsilon_{T-2} & , h = 1 \\ \phi \hat{z}_T(1) + \theta_2 \varepsilon_T + \theta_3 \varepsilon_{T-1} & , h = 2 \\ \phi \hat{z}_T(2) + \theta_3 \varepsilon_T & , h = 3 \\ \phi \hat{z}_T(h-1) & , h \geq 4 \end{cases}$$

(b) By working with an appropriately explicit expression for Z_{T+2}, derive an explicit formula for each of the following:

$e_T(2)$ i) the *forecast error* at lead time $h = 2$, formulated in terms of one or more members of the white noise process; [6]

$V(2)$ ii) the *forecast error variance* at lead time $h = 2$, namely $V(2)$ (involving nothing more than σ^2, θ_1, and ϕ). [2]

Question 2b

b)

(i) Recall forecast error: $e_T(h) = Z_{T+h} - \hat{Z}_T(h)$

- want $e_T(2)$, therefore want Z_{T+2}, can use $\hat{Z}_T(2)$ from (a).
- will want to be able to do $Z_{T+2} - \hat{Z}_T(2)$, so we'll want to eliminate Y_t terms, and thus have comperable values that we can subtract from another. Can't do $Y_{T+1} - \hat{y}_T(1)$, but can do $Y_T - y_T$ (since they're both realised observations).
- So will put $\hat{Z}_T(2)$ in terms of Z_T, and Z_{T+2} in terms of Z_T.

$\underline{\hat{Z}_T(2)}$

$$\hat{Z}_T(2) = \phi \hat{Z}_T(1) + \theta_2 \varepsilon_T + \theta_3 \varepsilon_{T-1}$$

[sub $\hat{Z}_T(1) = \phi Z_T + \theta_1 \varepsilon_T + \theta_2 \varepsilon_{T-1} + \theta_3 \varepsilon_{T-2}$ into above]

$$\hat{Z}_T(2) = \phi\left(\phi Z_T + \theta_1 \varepsilon_T + \theta_2 \varepsilon_{T-1} + \theta_3 \varepsilon_{T-2}\right) + \theta_2 \varepsilon_T + \theta_3 \varepsilon_{T-1}$$

$$\boxed{\begin{aligned}\hat{Z}_T(2) = \phi^2 Z_T + \phi\theta_1 \varepsilon_T + \phi\theta_2 \varepsilon_{T-1} + \phi\theta_3 \varepsilon_{T-2} \\ + \theta_2 \varepsilon_T + \theta_3 \varepsilon_{T-1}\end{aligned}}$$

$\underline{Z_{T+2}}$

$$Z_t = \phi Z_{t-1} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \theta_3 \varepsilon_{t-3}$$

$[t \to T+2]$

$$Z_{T+2} = \phi Z_{T+1} + \varepsilon_{T+2} + \theta_1 \varepsilon_{T+1} + \theta_2 \varepsilon_T + \theta_3 \varepsilon_{T-1}$$

$$Z_{T+1} = \phi Z_T + \varepsilon_{T+1} + \theta_1 \varepsilon_T + \theta_2 \varepsilon_{T-1} + \theta_3 \varepsilon_{T-2}$$

[sub Z_{T+1} into expression for Z_{T+2}]

$$Z_{T+2} = \phi(\phi Z_T + \varepsilon_{T+1} + \theta_1 \varepsilon_T + \theta_2 \varepsilon_{T-1} + \theta_3 \varepsilon_{T-2})$$
$$+ \varepsilon_{T+2} + \theta_1 \varepsilon_{T+1} + \theta_2 \varepsilon_T + \theta_3 \varepsilon_{T-1}$$

$$\boxed{\begin{aligned} Z_{T+2} = {} & \phi^2 z_T + \phi \varepsilon_{T+1} + \phi\theta_1 \varepsilon_T + \phi\theta_2 \varepsilon_{T-1} + \phi\theta_3 \varepsilon_{T-2} \\ & + \varepsilon_{T+2} + \theta_1 \varepsilon_{T+1} + \theta_2 \varepsilon_T + \theta_3 \varepsilon_{T-1} \end{aligned}}$$

↳ Note: the z_T here is a realised observation from H_T.

Could factorise both expressions for $\hat{z}_T(2)$ and Z_{T+2}, but purposefully leaving both expressions as is to make subtracting terms from each other easier.

$$\underline{e_T(2) = z_{T+2} - \hat{z}_T(2)}$$

$$e_T(2) = \cancel{\phi^2 z_T} + \phi\varepsilon_{T+1} + \phi\theta_1\cancel{\varepsilon_T} + \phi\theta_2\cancel{\varepsilon_{T-1}} + \phi\theta_3\cancel{\varepsilon_{T-2}}$$
$$+ \varepsilon_{T+2} + \theta_1\varepsilon_{T+1} + \theta_2\cancel{\varepsilon_T} + \theta_3\cancel{\varepsilon_{T-1}}$$
$$- \Big(\cancel{\phi^2 z_T} + \phi\theta_1\cancel{\varepsilon_T} + \phi\theta_2\cancel{\varepsilon_{T-1}} + \phi\theta_3\cancel{\varepsilon_{T-2}}$$
$$+ \theta_2\cancel{\varepsilon_T} + \theta_3\cancel{\varepsilon_{T-1}} \Big)$$

$$e_T(2) = \phi\varepsilon_{T+1} + \theta_1\varepsilon_{T+1} + \varepsilon_{T+2}$$

$$\boxed{e_T(2) = (\phi + \theta_1)\varepsilon_{T+1} + \varepsilon_{T+2}}$$

(ii) <u>The forecast error variance at $h=2$, $V(2)$.</u>

$$V(2) = Var(e_T(2)) = Var((\phi+\theta_1)\varepsilon_{T+1} + \varepsilon_{T+2})$$

The white noise terms are <u>INDEPENDENT</u> : $\varepsilon_t \sim NID(0, \sigma^2)$, and therefore the variance of the sum = sum of the variances.

$$V(2) = Var((\phi+\theta_1)\varepsilon_{T+1} + \varepsilon_{T+2})$$
$$= Var((\phi+\theta_1)\varepsilon_{T+1}) + Var(\varepsilon_{T+2})$$
$$= (\phi+\theta_1)^2 Var(\varepsilon_{T+1}) + Var(\varepsilon_{T+2})$$

$$\boxed{V(2) = (\phi+\theta_1)^2\sigma^2 + \sigma^2}$$

(won't factorise this any more as the Q requires the answer to only include ϕ, θ_1, σ^2.)

(c) Consider a stationary process $\{Y_t\}$ that satisfies the following model equation

ARMA(1,3)

$$Y_t = \phi Y_{t-1} + \epsilon_t + \omega_1\epsilon_{t-1} + \omega_2\epsilon_{t-2} + \omega_3\epsilon_{t-3}, \qquad t \in \mathbb{Z},$$

where ϕ, ω_1, ω_2, and ω_3, are parameters.

Suppose that $\hat{y}_T(2) = 12.904$, the estimated white noise variance $\hat{\sigma}^2 = 0.297$, and that the estimated values of ϕ, ω_1, ω_2, and ω_3, are 0.56, -1.383, 0.533, and -0.0625, respectively.

Stating any assumptions you make, and with reference to appropriate percentage points from statistical tables, find a 99% prediction interval for Y_{T+2}. [4]

Question 2c

(c)

- So, this is the same ARMA(1,3) model as in (a) and (b).

- we can use our expression for $V(2)$ to calculate the estimated forecast error variance.

- When producing the forecast equation in part (a) and the forecast error variance in (b), we assume that we know the real parameter values for $\phi, \theta_1, \theta_2, \theta_3$. Actually we just have estimates, $\hat{\phi}$, ω_1 $(\hat{\theta}_1)$, ω_2 $(\hat{\theta}_2)$ and ω_3 $(\hat{\theta}_3)$. These estimates should be useful for making forecasts if the history, H_T, is deep enough. (The theory often assumes that H_T goes back to the infinite past).

- We're also assuming that the process $\{Y_t\}$ is normally distributed given that the non-systematic part of the model - the white noise terms - are assumed to be normally distributed.
  - ↳ this assumption was declared in the Q, where $\epsilon_t \sim NID(0, \sigma^2)$.

- We therefore use the normal distribution to deduce how many standard errors $(\sqrt{V(h)})$ we will use to form the probability limits.

- we can then calculate the $100(1-\alpha)\% = 99\%$ prediction interval via:

$$\hat{y}_T(2) \pm z_{\alpha/2}\sqrt{V(2)}$$

α here is 0.01, and therefore $\alpha/2 = 0.005$.

The reason we divide α by 2 is because the normal distribution is symmetric and we looking at the regions of the normal distribution of equal or greater extremeness – which is in both tails.

$\boxed{z_{0.005} = 2.5758}$ (from Lindley & Scott)

Could also get the value from R via: qnorm(0.005, lower.tail=FALSE).

$$V(2) = (\phi+\theta_1)^2\sigma^2 + \sigma^2$$
$$= (0.56 - 1.383)^2 \times 0.297 + 0.297$$
$$= 0.4982$$

$$\boxed{\sqrt{V(2)} = 0.7058}$$

$$12.904 \pm 2.5758 \times 0.4982$$

$$\rightarrow \boxed{12.904 \pm 1.818}$$

And thus the 99% prediction interval for Y_{T+2}:

$$[11.086, 14.722].$$

END OF ASSIGNMENT.