Macroeconometrics
Supervision 2

Samuel Lee

Question 1

(a)

The data is generated by the MA(1) model

$$y_t = \varepsilon_t - \frac{1}{2}\varepsilon_{t-1}$$

since $\varepsilon_t = y_t + \frac{1}{2}\varepsilon_{t-1}$, we can express this recursively as

$$\begin{aligned}
\varepsilon_t &= y_t + \frac{1}{2}\left(y_{t-1} + \frac{1}{2}\varepsilon_{t-2}\right) \\
&= y_t + \frac{1}{2}y_{t-1} + \frac{1}{4}\left(y_{t-2} + \frac{1}{2}\varepsilon_{t-3}\right) \\
&\vdots \\
&= \sum_{i=0}^{\infty}\left(\frac{1}{2}\right)^i y_{t-i}
\end{aligned}$$

If the process started at some time $t = 0$ with a given ε_0, then

$$\varepsilon_t = \sum_{i=0}^{t-1}\left(\frac{1}{2}\right)^i y_{t-i} + \left(\frac{1}{2}\right)^t \varepsilon_0$$

(b)

The optimal forecast of y_{T+1} is

$$\begin{aligned}
\mathrm{E}[y_{T+1}|\Omega_T] &= \mathrm{E}\left[\varepsilon_{T+1} - \frac{1}{2}\varepsilon_T \middle| \Omega_T\right] \\
&= \mathrm{E}[\varepsilon_{T+1}|\Omega_T] - \frac{1}{2}\varepsilon_T
\end{aligned}$$

Since ε_t are standard normal random variables, that is, with mean 0 and variance 1, $\mathrm{E}[\varepsilon_{T+1}|\Omega_T] = 0$ and the optimal forecast is just $-\frac{1}{2}\varepsilon_T$. This can be expressed in terms of $y_1, y_2, \ldots, y_T$ and $\varepsilon_0 = 0$ as $-\frac{1}{2}\sum_{i=0}^{t-1}\left(\frac{1}{2}\right)^i y_{t-i}$.

(c)

One criterion for determining the 'best' C is to minimize the expected mean squared error of the forecast. So we minimize the following with respect to C:

$$\mathrm{E}[(y_{T+1} - Cy_T)^2] = \mathrm{E}\left[\left(\varepsilon_{T+1} - \frac{1}{2}\varepsilon_T - C\varepsilon_T + \frac{C}{2}\varepsilon_{T-1}\right)^2\right]$$

Taking the derivative with respect to C and setting it to 0:

$$\begin{aligned}
&2 \cdot \mathrm{E}\left[\left(\varepsilon_{T+1} - \frac{1}{2}\varepsilon_T - C\varepsilon_T + \frac{C}{2}\varepsilon_{T-1}\right)\left(\frac{1}{2}\varepsilon_{T-1} - \varepsilon_T\right)\right] = 0 \\
&\mathrm{E}\left[\frac{1}{2}\varepsilon_{T+1}\varepsilon_{T-1} - \varepsilon_{T+1}\varepsilon_T - \frac{1}{4}\varepsilon_T\varepsilon_{T-1} + \frac{1}{2}\varepsilon_T^2 - \frac{C}{2}\varepsilon_T\varepsilon_{T-1} + C\varepsilon_T^2 + \frac{C}{4}\varepsilon_{T-1}^2 - \frac{C}{2}\varepsilon_{T-1}\varepsilon_T\right] = 0
\end{aligned}$$

We can ignore all the products of ε_t from different time periods, since they are independently distributed with mean 0. So the condition for optimality finally becomes

$$\begin{aligned}
\mathrm{E}\left[\left(\frac{1}{2} + C\right)\varepsilon_T^2 + \frac{C}{4}\varepsilon_{T-1}^2\right] &= 0 \\
\frac{1}{2} + C + \frac{C}{4} &= 0 \\
C &= -\frac{2}{5}
\end{aligned}$$

using the fact that $\mathrm{E}[\varepsilon_t^2] = \mathrm{Var}[\varepsilon_t] = 1$.

Question 2

(a)

The only statistics we are given are the first 8 sample autocorrelations, so our tests for serial correlation will most likely involve the Box-Pierce Q statistic or the Ljung-Box Q statistic.

If the time series is white noise with no serial correlation, then it can be shown that the sample autocorrelations $\hat{\rho}(\tau)$ converge in distribution to $N\left(0, \frac{1}{T}\right)$. Therefore, $\sqrt{T}\hat{\rho}(\tau) \xrightarrow{D} N(0,1)$ and $T\hat{\rho}(\tau)^2 \xrightarrow{D} \chi^2(1)$. If the series is white noise, then the sample autocorrelation at different lag times are approximately independent of one another. The sum of m independent $\chi^2(1)$ variables has a $\chi^2(m)$ distribution, and so the sum of $T\hat{\rho}(\tau)^2$ at m different τ values has a $\chi^2(m)$ distribution.

The Box-Pierce test makes use of this. We have 8 sample autocorrelations, so if we use them all the null hypothesis is that $\rho(\tau) = 0 \;\forall \tau \in \{1, \dots, 8\}$, and the alternative is hypothesis that $\rho(\tau) \neq 0$ for some $\tau \in \{1, \dots, 8\}$. The critical value for a $\chi^2(8)$ distributed test statistic at a 5% significance level is 15.51. The Box-Pierce Q statistic is $T\sum_{\tau=1}^{8}\hat{\rho}(\tau)^2$ which in this case is equal to 13.02, thus we do not reject the null hypothesis that the series has zero autocorrelation up to a lag of 8.

The Ljung-Box Q statistic is asymptotically equivalent to the Box-Pierce Q statistic, but the former has been shown to be closer to a χ^2 distribution than the latter especially for small samples. For a test of τ up to 8, the Ljung-Box Q statistic is calculated by $T(T+2)\sum_{\tau=1}^{8}\frac{1}{T-\tau}\hat{\rho}(\tau)^2$. In our example this is equal to 13.843, which is again lower than the critical value of 15.51 we found before. Therefore the null hypothesis is still not rejected.

(b)

The BIC favours the AR(0) model (which is white noise) while the AIC favours the AR(2) model. The BIC is a consistent model selection criteria while the AIC is not, that is, if the true data generating process is among the models being considered, the probability of selecting it goes to 1 as the sample size goes to infinity under a consistent model selection criteria. However, the AIC is asymptotically efficient while the BIC is not, meaning it will select a sequence of models for which the 1-step-ahead forecast error variance approaches that of the true data generating process at least as quickly as any other model selection criteria, as the sample size gets large.

With the above in consideration, Diebold suggests picking the model favoured by the BIC when there is a conflict as it is usually more parsimonious and also due to studies comparing the out-of-sample forecasting performance of the models selected by different criteria. We could listen to Diebold and blame him for whatever happens afterwards, but there is also an additional reason to favour the BIC in this case. The BIC favours the white noise model with no serial correlation, and we have already decided in (a) that there is not yet evidence strong enough to reject that y_t is serially independent. Therefore the test we did in (a) would accord with the model that the BIC favours, and it makes sense for the researcher to choose $p = 0$.

(c)

Unfortunately it seems that $p = 0$ is wrong, and the data generating process is in fact AR(2). We know that y_t is stationary, so $\mathrm{E}[y_t] = \mu \ \forall t$. Therefore, the unconditional mean is

$$\begin{aligned} \mathrm{E}[y_t] &= \mathrm{E}[1 + 0.2y_{t-1} - 0.1y_{t-2} + \varepsilon_t] \\ \mu &= 1 + 0.1\mu = \frac{10}{9} \end{aligned}$$

The model has the lag polynomial representation

$$\Phi(L)y_t = 1 + \varepsilon_t$$

where $\Phi(L) = 1 - 0.2L + 0.1L^2$. The roots of $\Phi(L)$ are the complex conjugates $1 \pm 3i$, which are obviously outside the unit circle. The inverses of the roots are $\frac{1}{1+3i} = \frac{1-3i}{10}$ and $\frac{1}{1-3i} = \frac{1+3i}{10}$, and we know the inverses of the roots are within the unit circle since the roots were outside the unit circle. Therefore the process is stable. Alternatively, we know that the process is stable because we were already told that y_t was stationary. If y_t were not stable, then $\mathrm{Cov}[y_t]$ is infinite and y_t cannot be stationary.

It can be shown that the expectation of y_{t+h} conditional on the information set Ω_t minimizes the expected mean squared error of the forecast. Therefore, the optimal forecasts are

$$\begin{aligned} \mathrm{E}[y_{101}|\Omega_{100}] &= \mathrm{E}[1 + 0.2y_{100} - 0.1y_{99} + \varepsilon_{101}|\Omega_{100}] = 1 + 0.2(0.21) - 0.1(2.15) = 0.827 \\ \mathrm{E}[y_{102}|\Omega_{100}] &= \mathrm{E}[1 + 0.2y_{101} - 0.1y_{100} + \varepsilon_{102}|\Omega_{100}] = 1 + 0.2(0.827) - 0.1(0.21) = 1.1444 \\ \mathrm{E}[y_{103}|\Omega_{100}] &= \mathrm{E}[1 + 0.2y_{102} - 0.1y_{101} + \varepsilon_{103}|\Omega_{100}] = 1 + 0.2(1.1444) - 0.1(0.827) = 1.14618 \end{aligned}$$

where we used the optimal forecasts for y_t recursively for forecasts even further ahead.

(d)

We can ignore any parameter estimation uncertainty since we have been told the (hopefully) true coefficients of the AR(2) model, and just focus on the forecast error variance. The forecast errors are

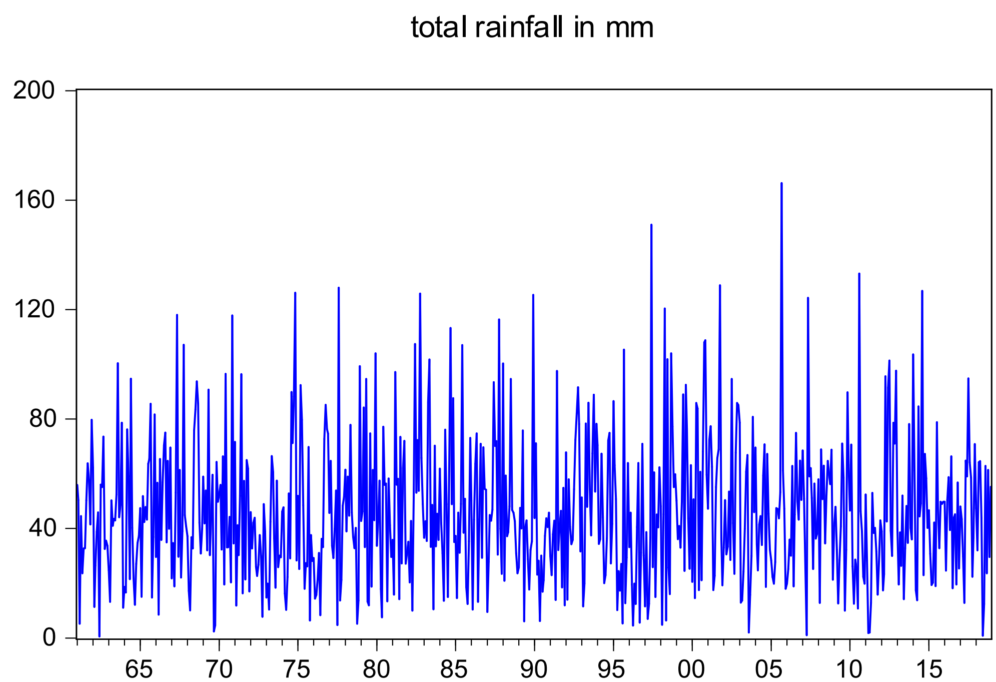
$$
\begin{aligned}
e_{101} &= y_{101} - \mathrm{E}[y_{101}|\Omega_{100}] = 1 + 0.2y_{100} - 0.1y_{99} + \varepsilon_{101} - 1 - 0.2y_{100} + 0.1y_{99} = \varepsilon_{101} \\
e_{102} &= y_{102} - \mathrm{E}[y_{102}|\Omega_{100}] = 1 + 0.2y_{101} - 0.1y_{100} + \varepsilon_{102} - 1 - 0.2\mathrm{E}[y_{101}|\Omega_{100}] + 0.1y_{100} \\
&= \varepsilon_{102} + 0.2e_{101} = \varepsilon_{102} + 0.2\varepsilon_{101} \\
e_{103} &= y_{103} - \mathrm{E}[y_{103}|\Omega_{100}] = 1 + 0.2y_{102} - 0.1y_{101} + \varepsilon_{103} - 1 - 0.2\mathrm{E}[y_{102}|\Omega_{100}] + 0.1\mathrm{E}[y_{101}|\Omega_{100}] \\
&= \varepsilon_{103} + 0.2e_{102} - 0.1e_{101} = \varepsilon_{103} + 0.2(\varepsilon_{102} + 0.2\varepsilon_{101}) - 0.1\varepsilon_{101} = \varepsilon_{103} + 0.2\varepsilon_{102} - 0.06\varepsilon_{101}
\end{aligned}
$$

and since the ε_t are standard normal with 0 covariance, the forecast error variances are 1, 1.04, and 1.0436, and the 95% interval forecasts of y_{101}, y_{102}, and y_{103} are 0.827 ± 1.96, $1.1444 \pm 1.96 \times 1.04$, and $1.14618 \pm 1.96 \times 1.0436$.

Question 3

(a)

Below are the time series and correlogram plots:



Correlogram of RAIN

Date: 03/02/19 Time: 22:47
Sample: 1961M01 2018M12
Included observations: 696

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.034	0.034	0.8101	0.368
		2	0.044	0.043	2.1708	0.338
		3	-0.021	-0.024	2.4926	0.477
		4	-0.071	-0.071	6.0025	0.199
		5	0.004	0.011	6.0163	0.305
		6	-0.025	-0.020	6.4566	0.374
		7	-0.010	-0.013	6.5292	0.479
		8	-0.029	-0.032	7.1398	0.522
		9	0.015	0.019	7.3055	0.605
		10	0.064	0.062	10.164	0.426
		11	0.043	0.036	11.492	0.403
		12	0.009	-0.003	11.556	0.482
		13	-0.040	-0.039	12.675	0.473
		14	-0.044	-0.034	14.067	0.445
		15	0.058	0.070	16.455	0.352
		16	-0.017	-0.017	16.651	0.408
		17	-0.001	-0.010	16.653	0.478
		18	-0.007	-0.003	16.683	0.545
		19	-0.008	0.001	16.729	0.608
		20	-0.014	-0.024	16.876	0.661
		21	-0.009	-0.015	16.940	0.715
		22	0.047	0.047	18.556	0.673
		23	0.028	0.035	19.108	0.695
		24	-0.009	-0.014	19.170	0.743
		25	0.025	0.019	19.617	0.767
		26	0.021	0.024	19.943	0.794
		27	0.013	0.011	20.067	0.828
		28	0.004	0.005	20.080	0.862
		29	-0.097	-0.091	26.990	0.572
		30	-0.060	-0.054	29.658	0.483
		31	-0.021	-0.001	29.993	0.518
		32	-0.066	-0.069	33.162	0.410
		33	-0.022	-0.042	33.531	0.442
		34	0.039	0.039	34.625	0.438
		35	-0.027	-0.029	35.167	0.460
		36	-0.011	-0.025	35.250	0.504

The two plots seem to suggest that there is no serial correlation within the monthly rainfall values. The autocorrelation and partial autocorrelation at lags up to 36 mostly lie within the 1.96-standard-deviations bound, and none of them are statistically significant at the 10% level (even at the 20% level only the autocorrelation at lag 4 is barely significant).

(b)

Below are the results of the estimated linear trend model:

Table 1: Estimation of linear trend

Dependent Variable: RAIN
Method: Least Squares
Date: 03/02/19 Time: 23:12
Sample: 1961M01 2018M12
Included observations: 696

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	45.38381	2.000487	22.68638	0.0000
@TREND	0.002100	0.004984	0.421339	0.6736

R-squared	0.000256	Mean dependent var	46.11351
Adjusted R-squared	−0.001185	S.D. dependent var	26.40105
S.E. of regression	26.41668	Akaike info criterion	9.388738
Sum squared resid	484301.7	Schwarz criterion	9.401799
Log likelihood	−3265.281	Hannan-Quinn criter.	9.393788
F-statistic	0.177527	Durbin-Watson stat	1.932040
Prob(F-statistic)	0.673638		

Not only is the coefficent on the linear trend statistically insignificant, the magnitude is also very small (it suggests just a 1.5 mm increase in rainfall from 1961 to the end of 2018). The results suggest that a linear trend would not have to be included in the forecasting model for rainfall.

Below are the results from estimating the seasonal dummy variable model and testing that the total rainfall in Cambridge is the same for all months of the year. Since the specified equation has no intercept, this means testing that all the estimated coefficients are equal. The p-value from the test is 0.0004, so we reject the null hypothesis that total rainfall in Cambridge is the same for all the months (the value of the F-statistic is the same as the reported F-statistic when we drop one dummy variable and estimate the seasonal effects with an intercept). From the estimation of the seasonal dummy variables, we can see that the three wettest months are August, October, and November, while the three driest months are February, March, and April.

(c)

The results of the procedure are below. The results suggest that among the ARMA(p, q) models, $p, q \in \{0, 1, 2, 3, 4, 5\}$, with a seasonal dummy variable component, the one favoured by the SIC is simply the ARMA(0,0) model, which is the same as the one estimated in (c).

(d)

If we assume the plain seasonal dummy variable model, then the point forecast for the amount of rain in Cambridge in March of 2019 is simple the coefficient “@MONTH=3”, which is 37.56552. We ignore the parameter estimation uncertainty, so we take it that the remaining variance will come

Table 2: Estimation of seasonal dummy variable model

Dependent Variable: RAIN
Method: Least Squares
Date: 03/02/19 Time: 23:19
Sample: 1961M01 2018M12
Included observations: 696

Variable	Coefficient	Std. Error	t-Statistic	Prob.
@MONTH=1	46.12931	3.409600	13.52924	0.0000
@MONTH=2	33.61724	3.409600	9.859585	0.0000
@MONTH=3	37.56552	3.409600	11.01757	0.0000
@MONTH=4	40.75172	3.409600	11.95205	0.0000
@MONTH=5	46.67931	3.409600	13.69055	0.0000
@MONTH=6	48.91724	3.409600	14.34692	0.0000
@MONTH=7	46.04655	3.409600	13.50497	0.0000
@MONTH=8	53.94138	3.409600	15.82044	0.0000
@MONTH=9	47.91552	3.409600	14.05312	0.0000
@MONTH=10	51.86897	3.409600	15.21263	0.0000
@MONTH=11	51.82759	3.409600	15.20049	0.0000
@MONTH=12	48.10172	3.409600	14.10773	0.0000

R-squared	0.047941	Mean dependent var	46.11351
Adjusted R-squared	0.032630	S.D. dependent var	26.40105
S.E. of regression	25.96674	Akaike info criterion	9.368601
Sum squared resid	461201.7	Schwarz criterion	9.446969
Log likelihood	−3248.273	Hannan-Quinn criter.	9.398903
Durbin-Watson stat	1.983180		

Table 3: Results of Wald test

Wald Test:

Test Statistic	Value	df	Probability
F-statistic	3.131170	(11, 684)	0.0004
Chi-square	34.44287	11	0.0003

Null Hypothesis: C(1)=C(2)=C(3)=C(4)=C(5)=C(6)=C(7)=C(8)=C(9)=C(10)=C(11)=C(12)
Null Hypothesis Summary:

Normalized Restriction (= 0)	Value	Std. Err.
C(1) - C(12)	−1.972414	4.821902
C(2) - C(12)	−14.48448	4.821902
C(3) - C(12)	−10.53621	4.821902
C(4) - C(12)	−7.350000	4.821902
C(5) - C(12)	−1.422414	4.821902
C(6) - C(12)	0.815517	4.821902
C(7) - C(12)	−2.055172	4.821902
C(8) - C(12)	5.839655	4.821902
C(9) - C(12)	−0.186207	4.821902
C(10) - C(12)	3.767241	4.821902
C(11) - C(12)	3.725862	4.821902

Restrictions are linear in coefficients.

Table 4: Results of automatic ARIMA forecasting

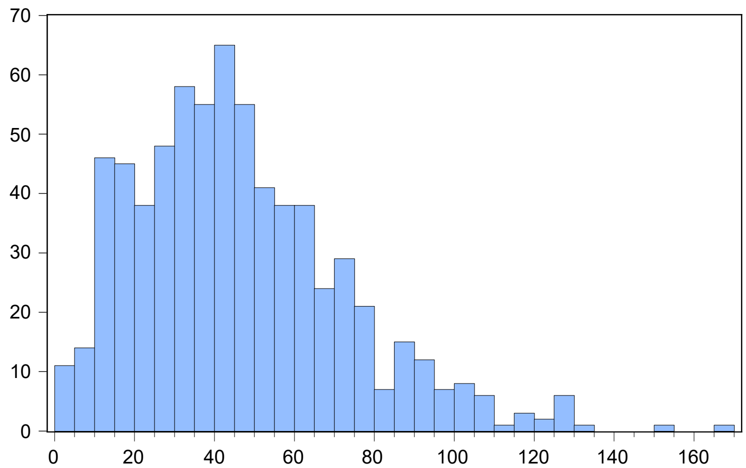
Automatic ARIMA Forecasting
Selected dependent variable: RAIN
Date: 03/03/19 Time: 00:04
Sample: 1961M01 2018M12
Included observations: 696
Forecast length: 0

Number of estimated ARMA models: 25
Number of non-converged estimations: 0
Selected ARMA model: (0,0)(0,0)
SIC value: 9.45637335689

from the random disturbance term ε_t, which we assume to be normally distributed with mean 0. The estimate of its standard deviation is the standard error of regression 25.96674, so we need to find the two-sided 50% critical values. For a standard normal distribution, $1 - \Phi(0.6745) = 0.25$, so the 50% interval forecast is $37.56552 \pm 0.675 \times 25.96674$ where we have ignored two sources of additional uncertainty: the estimate of the coefficient for the March dummy variable, and the estimate of the variance of ε_t.

(e)

Below are the histogram and stats from EViews:



Series: RAIN	
Sample 1961M01 2018M12	
Observations 696	
Mean	46.11351
Median	42.75000
Maximum	166.2000
Minimum	0.500000
Std. Dev.	26.40105
Skewness	0.896061
Kurtosis	4.030747
Jarque-Bera	123.9501
Probability	0.000000

The histogram suggests that a normally or identically distributed ε_t may not be an appropriate assumption; total rainfall cannot go below 0, so $\Pr(\varepsilon_t < @MONTH = q | month = q) = 0$. It is more likely that we have a truncated normal distribution, but it is difficult to get an analytical form for the critical values of such a distribution. We may opt to use the empirical distribution instead: first, we isolate all the observations for March, and then we take the smallest interval containing 50% of the values in March as our 50% interval forecast. This process is somewhat similar to bootstrapping.

Question 4

(a)

x_t has zero mean, so y_t also has zero mean since $\mathrm{E}[y_t] = \mathrm{E}[x_t] + \mathrm{E}[\eta_t] = 0$. Therefore, the autocovariance for y_t at lag τ is just $\mathrm{E}[y_t y_{t-\tau}]$, as such:

$$\begin{aligned} y_t y_{t-\tau} &= x_t y_{t-\tau} + \eta_t y_{t-\tau} \\ &= x_t(x_{t-\tau} + \eta_{t-\tau}) + \eta_t(x_{t-\tau} + \eta_{t-\tau}) \\ &= x_t x_{t-\tau} + x_t \eta_{t-\tau} + x_{t-\tau}\eta_t + \eta_t \eta_{t-\tau} \\ \mathrm{E}[y_t y_{t-\tau}] &= \mathrm{E}[x_t x_{t-\tau}] && \text{if } \tau \neq 0. \end{aligned}$$

since η_t is independent with zero mean. This means that the autocovariance of y_t for lag $\tau \neq= 0$ is the same as the autocovariance of x_t. We know the autocovariance function for an AR(1) process is $\gamma(\tau) = \varphi^\tau \frac{\sigma^2}{1-\varphi^2}$, and in this case we have $\varphi = 0.5, \sigma^2 = 1$.

If $\tau = 0$, we have to add the expectation of η_t^2, which is 1. Therefore, the population autocovariance function is

$$\gamma(\tau) = \begin{cases} \frac{7}{3} & \text{if } |\tau| = 0 \\ \frac{4}{3}\left(\frac{1}{2}\right)^{|\tau|} & \text{if } |\tau| \neq 0 \end{cases}$$

(b)

The formulas given appear to be wrong; for a stationary process z_t with zero mean, we have $\mathrm{Cov}[z_t, u_t] = \mathrm{E}[z_t u_t]$. This can be found by multiplying both sides of the equation for the process by u_t:

$$\begin{aligned} z_t u_t &= \varphi z_{t-1} u_t + u_t^2 + \theta u_t u_{t-1} \\ \mathrm{E}[z_t u_t] &= \varphi \mathrm{E}[z_{t-1} u_t] + \mathrm{E}[u_t^2] + \theta \mathrm{E}[u_t u_{t-1}] = \sigma^2 \end{aligned}$$

since u_t are independent and have zero mean and the expectation of their product with anything else is also 0. Knowing this, we can get the variance of z_t:

$$\begin{aligned} \mathrm{Var}[z_t] &= \mathrm{Var}[\varphi z_{t-1} + u_t + \theta u_{t-1}] \\ &= \varphi^2 \mathrm{Var}[z_t] + (1+\theta^2)\sigma^2 + 2\mathrm{Cov}[\varphi z_{t-1}, u_t + \theta u_{t-1}] \\ &= \varphi^2 \mathrm{Var}[z_t] + (1+\theta^2)\sigma^2 + 2\varphi\theta \mathrm{Cov}[z_{t-1}, u_{t-1}] \\ &= \frac{1 + 2\varphi\theta + \theta^2}{1-\varphi^2}\sigma^2 \end{aligned}$$

where we used the fact that $\mathrm{Var}[z_{t-1}] = \mathrm{Var}[z_t]$ since z_t is stationary, and our earlier derivation that $\mathrm{Cov}[z_t, u_t] = \mathrm{E}[z_t u_t] = \sigma^2$. Similarly, we can find $\mathrm{Cov}[z_t, z_{t-1}]$ which is equal to $\mathrm{E}[z_t z_{t-1}]$:

$$\begin{aligned} z_t z_{t-1} &= \varphi z_{t-1}^2 + z_{t-1} u_t + \theta z_{t-1} u_{t-1} \\ \mathrm{E}[z_t z_{t-1}] &= \varphi \mathrm{E}[z_{t-1}^2] + \mathrm{E}[z_{t-1} u_t] + \theta \mathrm{E}[z_{t-1} u_{t-1}] \\ &= \frac{\varphi(1 + 2\varphi\theta + \theta^2)}{1-\varphi^2}\sigma^2 + \theta\sigma^2 \\ &= \frac{\varphi + 2\varphi^2\theta + \varphi\theta^2 + \theta - \varphi^2\theta}{1-\varphi^2}\sigma^2 \\ &= \frac{\varphi + \varphi^2\theta + \varphi\theta^2 + \theta}{1-\varphi^2}\sigma^2 \\ &= \frac{(\varphi+\theta)(1+\varphi\theta)}{1-\varphi^2}\sigma^2 \end{aligned}$$

Both the expressions are different from the ones given. To find the population autocovariance function, we just do the same thing above but with lag τ instead of 1, and we will probably get a recursive expression. Trying this out, we get

$$\begin{aligned} z_t z_{t-\tau} &= \varphi z_{t-1} z_{t-\tau} + u_t z_{t-\tau} + \theta u_{t-1} z_{t-\tau} \\ \mathrm{E}[z_t z_{t-\tau}] &= \varphi \mathrm{E}[z_{t-1} z_{t-\tau}] && \text{(as long as } \tau > 1\text{)} \\ &= \varphi \mathrm{E}[z_t z_{t-(\tau-1)}] \end{aligned}$$

which means that $\gamma(\tau) = \varphi\gamma(\tau - 1)$ for $\tau \in \{2, 3, \ldots\}$ (the Yule-Walker equation). Therefore, the population autocovariance function is

$$\gamma(\tau) = \begin{cases} \frac{1+2\varphi\theta+\theta^2}{1-\varphi^2}\sigma^2 & \text{for } \tau = 0 \\ \varphi^{\tau-1}\frac{(\varphi+\theta)(1+\varphi\theta)}{1-\varphi^2}\sigma^2 & \text{for } \tau \neq 0 \end{cases}$$

(c)

For the two functions to be equal, we must have

$$\frac{7}{3} = \frac{1 + 2\varphi\theta + \theta^2}{1 - \varphi^2}\sigma^2$$
$$\frac{4}{3}\left(\frac{1}{2}\right)^{\tau} = \varphi^{\tau-1}\frac{(\varphi + \theta)(1 + \varphi\theta)}{1 - \varphi^2}\sigma^2$$

We have 2 equations and 3 unknowns. We might have to set one of the parameters to some arbitrary value to get a solution, and it seems like a good idea to coerce the right-hand side of the equation into a form where the exponent $\tau - 1$ becomes τ just like the left-hand side:

$$\varphi^{\tau-1}\frac{(\varphi + \theta)(1 + \varphi\theta)}{1 - \varphi^2}\sigma^2 = \varphi^{\tau}\frac{(\varphi + \theta)(1 + \varphi\theta)}{\varphi(1 - \varphi^2)}\sigma^2$$

It then seems like a good idea to set $\varphi = \frac{1}{2}$ to make the exponentiated term the same as the left-hand side. Doing so, we get from the second equality

$$\frac{\left(\frac{1}{2} + \theta\right)\left(1 + \frac{\theta}{2}\right)}{3/8}\sigma^2 = \frac{4}{3}$$
$$\left(\frac{1}{2} + \frac{5\theta}{4} + \frac{\theta^2}{2}\right)\sigma^2 = \frac{1}{2}$$
$$\left(1 + \frac{5\theta}{2} + \theta^2\right)\sigma^2 = 1$$

and from the first equality

$$\frac{1 + \theta + \theta^2}{3/4}\sigma^2 = \frac{7}{3}$$
$$(1 + \theta + \theta^2)\sigma^2 = \frac{7}{4}$$

Dividing through, we get

$$\frac{1 + \theta + \theta^2}{1 + \frac{5\theta}{2} + \theta^2} = \frac{7}{4}$$
$$4 + 4\theta + 4\theta^2 = 7 + \frac{35\theta}{2} + 7\theta^2$$
$$3\theta^2 + \frac{27}{2}\theta + 3 = 0$$
$$2\theta^2 + 9\theta + 2 = 0$$

and we get $\theta = \frac{-9\pm\sqrt{65}}{4}$. We want the process to be invertible, so only $\theta = \frac{-9+\sqrt{65}}{4}$ is admissible. We can substitute this back to find σ^2:

$$\sigma^2 = \frac{7}{4} \frac{1}{1 + \frac{-9+\sqrt{65}}{4} + \left(\frac{-9+\sqrt{65}}{4}\right)^2}$$

Luckily, the calculator can derive the exact form for the above, which is $\frac{9+\sqrt{65}}{8}$. Therefore, one of the solutions that equate the autocovariance functions is $\varphi = \frac{1}{2}, \theta = \frac{-9+\sqrt{65}}{4}, \sigma^2 = \frac{9+\sqrt{65}}{8}$. It is probably the only solution as well, since we set $\varphi = \frac{1}{2}$ to get rid of the τ term and make the equality hold for all τ. It doesn't seem likely that any other φ could lead to a solution that is true for all τ, while still allowing σ^2, φ, and θ to be independent of τ.

Therefore, an AR(1) + white noise process can have the same autocovariance structure as an ARMA(1,1) process for some values of φ and θ. Furthermore, ε_t is normally distributed, so x_t and by extension y_t are Gaussian processes which are completely characterized by their first and second moments.