Sentence 1: Stochastic seasonalities exist when the roots of the ____8____ of the ARMA model are ____12____ but are not at ____2____.

Sentence 2: Polynomial stochastic trend of order k exists when there is an ____10____ of order ____14____.

Sentence 3: Time series with a deterministic seasonality is ____4____ because its ____6____ is varying over time.

Sentence 4: Optimal control of an ARMAV system with the lag between input and output equal to p is achieved by driving the ____7____ of the system output to ____1____.

Sentence 5: Existence of an autoregressive root at 1 signifies a ____13____ of order ____1____.

Table 1.1: Possible choices to complete sentences 1-5 in the part a) of Problem 1.

1. Zero	6. Expected Value	11. Stochastic seasonality
2. One	7. p-steps ahead prediction	12. On the unit circle
3. Stationary	8. Autoregressive part	13. Stochastic trend
4. Non-stationary	9. Moving average part	14. k+1
5. Two	10. Autoregressive root at 1.	15. Inside the unit circle

Part (b) (10 points)

A continuous time autoregressive A(2) system

$$\frac{d^2X(t)}{dt^2}+12\frac{dX(t)}{dt}+144X(t)=Z(t),\ E[Z(t)]=0, E[Z(t)Z(t-\tau)]=9\delta(t),$$

where $\delta(s)$ is the Dirac's delta function, is sampled equidistantly with the sampling interval $\Delta = 0.1s$

i) What is are the natural frequency ω_n and damping coefficient ξ of this system?
(3 points)

$$\omega_n^2 = 144 \Rightarrow \omega_n = 12 \ rad/s$$

$$2\xi\omega_n = 12 \Rightarrow \xi = 0.5$$

ii) Use the **impulse response equivalence** method to obtain the discrete-time model describing the sampled time-series. Please remember that the moving average part of this model will look a bit unusual (my lecture notes are relevant for it).
(7 points)

$$\mu_{1/2} = -\xi\omega_n \pm j\omega_n\sqrt{1-\xi^2} = -6 \pm j\cdot 10.3923$$

$\lambda_{1/2} = e^{\mu_{1/2}\cdot\Delta}$ and AR part of the model is

$$(1-\lambda_1 B)(1-\lambda_2 B) = 1-(\lambda_1+\lambda_2)B+\lambda_1\lambda_2 B^2$$
$$= 1-\phi_1 B-\phi_2 B^2$$

$$\phi_1 = e^{\mu_1\Delta}+e^{\mu_2\Delta} = e^{-\xi\omega_n\Delta}\cdot 2\cos(\omega_n\sqrt{1-\xi^2}\Delta) = 0.5564$$

$$\phi_2 = -e^{-2\xi\omega_n\Delta} = -0.3012$$

Form of the model is $(1-\phi_1 B-\phi_2 B^2)X_t = \theta_1 B a_t$

$$\theta_1 = \frac{\lambda_1-\lambda_2}{\mu_1-\mu_2} = 0.0455 \Rightarrow X_t - 0.5564X_{t-1} + 0.3012X_{t-2} = 0.0455a_{t-1}$$

Problem 2. (20 points)

Gate-opening in the papermaking process is sampled every $\Delta = 0.2$ hours and the resulting AR(1) model was found to be

$$X_t - 0.9X_{t-1} = a_t,\ a_t \sim NID(0, \sigma_a^2) \quad (1)$$

where $\sigma_a^2 = 4.0$.

Part (a) (2 points)

Find the corresponding Green's function.

$$G_\ell = \phi^\ell = 0.9^\ell, \quad \ell = 0, 1, \ldots$$

Part (b) (2 points)

Find the covariance function $E[X_t X_{t-l}] = \gamma_l$ for the time-series described by the AR(1) model (1).

$$\gamma_\ell = \sigma_a^2 \frac{\phi_1^2}{1-\phi_1^2} = \frac{4.0}{1-0.9^2} \cdot 0.9^\ell$$

$$= 21.0526 \cdot 0.9^\ell$$

Part (c) (4 points)

What is the stochastic differential equation describing the behavior of the gate-opening in continuous time (don't forget to identify noise characteristics).

$X'(t) + \alpha X(t) = Z(t)$ where

$$\alpha = -\frac{\ln \phi_1}{\Delta} = -\frac{\ln 0.9}{0.2} = 0.5268$$

$E[Z(t)] = 0$ and $E[Z(t)Z(s)] = \sigma_Z^2 \delta(t-s)$

$$\frac{\sigma_Z^2}{2\alpha} = \frac{\sigma_a^2}{1-\phi_1^2} \Rightarrow \sigma_Z^2 = \frac{2\alpha}{1-\phi_1^2}\sigma_a^2 = \frac{2 \cdot 0.5268}{1-0.9^2} \cdot 4 =$$

$$= 22.1811$$

Part (d) (3 points)

What is the impulse response of the continuous-time system described by the differential equation you found in part (c).

$$G(t) = e^{-\alpha t} = e^{-0.5268 t}$$

Part (e) (3 points)

What is the connection between the Green's function found in part a and the continuous-time impulse response you found in part (d).

$$G_\ell = G(\ell \Delta)$$

(they match at appropriate time samples)

Part (f) (3 points)

Find the covariance function describing the response of the continuous-time system you found in part (d).

$$\gamma(s) = \frac{\sigma_Z^2}{2\alpha} e^{-\alpha s} =$$

$$= \frac{22.1811}{2 \cdot 0.5268} e^{-0.5268 s}, \quad s \geq 0$$

Part (g) (3 points)

What is the connection between the discrete-time covariance function you found in part (b) and the continuous-time covariance function you found in part (e).

$\gamma_\ell = \gamma(\Delta \ell)$ (they match at appropriate time-lag samples)

Problem 3. (Total 15 points)

An engineer decides to do system identification by exposing the system to a unit step input and fitting a model of the form

$$Y_t = m + \sum_{j=1}^{l}\left[A_j \sin(j\omega_0 t) + B_j \cos(j\omega_0 t)\right] + X_t$$

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \ldots + \phi_n X_{t-n} + a_t - \theta_1 a_{t-1} - \ldots - \theta_m a_{t-m}$$

to the output time-series Y_t. The output time-series has N=400 samples and the engineer successively increased the number of harmonics and the order of the ARMA(n,n-1) model describing the time-series X_t by 1 until the residual sum of square (RSS) did not reduce significantly. In other words, he first fits a model with one harmonic (one sinusoid and cosine term), one Autoregressive (AR) term and 0 Moving Average (MA) terms, then he fits a model with 2 harmonics (two sinusoidal and cosine terms), 2 AR terms and 1 MA term, followed by a model with 3 harmonics (three sinusoidal and cosine terms), 3 AR and 2 MA terms, etc. Table 1 describes the modeling procedure and order of the model (r_1, r_2, r_3) denotes a model with r_1 harmonics, r_2 AR terms and r_3 MA terms.

Table 1

Parameters \ Model Order	(1,1,0)	(2,2,1)	(3,3,2)	(2,2,0)
m	11.2 ± 0.6	10.1 ± 0.7	10.0 ± 0.7	10.2 ± 0.6
A_1	1303 ± 3.1	12.2 ± 4.2	11.9 ± 4.0	12.1 ± 4.0
B_1	0.12 ± 0.1	2.1 ± 0.7	1.9 ± 0.5	2.0 ± 0.4
A_2		2.2 ± 0.2	2.1 ± 0.1	2.0 ± 0.1
B_2		0.4 ± 0.1	0.5 ± 0.2	0.4 ± 0.2
A_3			0.2 ± 0.3	
B_3			9.0 ± 4.2	
ϕ_1	0.9 ± 0.1	-1.9 ± 0.2	-2.1 ± 0.3	-2.0 ± 0.1
ϕ_2		-1.1 ± 0.1	-1.0 ± 0.1	-1.0 ± 0.1
ϕ_3			0.2 ± 0.3	
θ_1		0.2 ± 0.3	0.2 ± 0.3	
θ_2			0.1 ± 0.2	
RSS	800	595	580	600

YOUR QUESTIONS START ON THE NEXT PAGE. PLEASE, TURN THE PAGE!

Part (a) (9 points)
What model did the engineer choose to be adequate for the time-series Y_t (conduct the necessary F-tests).

Test 1: (2,2,1) vs (1,1,0) $A_1 = 800$; $A_0 = 595$

$s = 4$, $r = 8$, $N = 400$

$$F = \frac{(A_1 - A_0)/s}{A_0/(N-r)} = 33.76 > F_{4,\infty}^{0.95} = 2.39$$

=> keep going

Test 2: (3,3,2) vs. (2,2,1) ; $A_1 = 595$, $A_0 = 580$

$s = 4$, $r = 12$, $N = 400$

$$F = \frac{(A_1 - A_0)/s}{A_0/(N-r)} = 2.509 > F_{4,\infty}^{0.95} = 2.39$$

=> we should keep going and check (4,4,3) vs (3,3,2), but since we do not have data for it, we should keep (3,3,2) (my error with numbers)

If you went down to (2,2,1) because of the next problem, that is OK.

Part (b) (6 points)
What can you say about the deterministic and stochastic seasonalities in the time-series Y_t? Determine periodicities of each deterministic or stochastic seasonality. Identify them all based on the adequate model.

Taking (2,2,1) model, as instructed (sorry)

$$Y_t = 10.1 + 12.2 \sin \omega_0 t + 2.1 \cos \omega_0 t +$$
$$+ 2.2 \sin 2\omega_0 t + 0.4 \cos 2\omega_0 t +$$
$$- 1.9 X_{t-1} - 1.1 X_{t-2} + a_t - 0.2 a_{t-1}$$

Deterministic seasonalities of periodicities

$$T_1 = \frac{2\pi}{\omega_0}, \quad T_2 = \frac{2\pi}{2\omega_0}$$ (2 deterministic seasonalities)

AR characteristic polynomial: $s^2 + 1.9s + 1.1 = 0$

$$\lambda_{1/2} = -0.95 \pm j \cdot 0.4444 = 1.049 \cdot e^{\pm 2.704 j}$$

Assuming this is close to the unit circle (it's ok if you said it isn't), we have a stochastic seasonality of periodicity

$$T_{Stoch} = \frac{2\pi}{2.704} = 2.324 \text{ samples}$$

Problem 4. (Total 20 points)

For a single input (X_{1t}), single output (X_{2t}) system sampled at N=300 points, the following models were successively fit.

Model 1:

$$X_{2t} = 0.4X_{1t-2} + 0.7X_{2t-1} + a_{2t}$$
$$\text{RSS}_1 = 1800$$

Model 2:

$$X_{2t} = 0.32X_{1t-2} - 0.63X_{2t-1} + 0.06X_{1t-3} - 0.02X_{2t-2} + a_{2t} + 0.03a_{2t-1}$$
$$\text{RSS}_2 = 1770$$

For both models, RSS denotes the residual sum of squares for the output.

Part (a) (3 points)
Conduct the necessary statistical testing to answer if the Model 1 can be considered adequate,

$A_1 = 1800$; $A_0 = 1770$; $r = 5$; $s = 3$; $N = 300$

$$F = \frac{(A_1 - A_0)/s}{A_0/(N-r)} = 1.661 \qquad F_{3,\infty}^{95\%} = 2.60$$

$F < F_{3,\infty}^{95\%} \Rightarrow$ Model 1 can be considered as adequate

Part (b) (5 points)
Write out Model 1 in the form

$$X_{2t} = \sum_{k=0}^{\infty} G_k^X X_{1t-k} + \sum_{k=0}^{\infty} G_k a_{2t-k}$$

Interpret the two portions of the decomposition you made above.

$$(1-0.7B)X_{2_t} = 0.4B^2 X_{1_t} + a_{2_t} \Rightarrow$$

$$X_{2_t} = \frac{0.4B^2}{1-0.7B} X_{1_t} + \frac{1}{1-0.7B} a_{2_t} \Rightarrow$$

$$X_{2_t} = 0.4 \cdot (1+0.7B+0.7^2B^2+\dots) X_{1_{t-2}} +$$

$$+ (1+0.7B+0.7^2B^2+\dots) a_{2_t}$$

Part (c) (4 points)
Derive the mean-square error optimal control equation for Model 1.

We have a lag of 2 samples between input and output and hence, optimal regulation control law will be

$$\hat{X}_{2_t}(2) = 0; \qquad \hat{X}_{2_t}(2) = 0.4X_{1_t} + 0.7\hat{X}_{2_t}(1)$$

$$\hat{X}_{2_t}(1) = 0.4X_{1_{t-1}} + 0.7X_{2_t} \Rightarrow$$

$$\hat{X}_{2_t}(2) = 0.4X_{1_t} + 0.7(0.4X_{1_{t-1}} + 0.7X_{2_t}) = 0 \Rightarrow$$

$$\boxed{X_{1_t} = -0.7X_{1_{t-1}} - 1.225X_{2_t}}$$

Part (d) (4 points)
Write the model of the output after optimal control is applied to it (all for Model 1).

$$X_{2_{t+2}} = (1 + 0.7B)\, a_{2_{t+2}}$$

(part that cannot be annulled by driving $\hat{X}_{2_t}(2)$ to zero)

Part (e) (4 points)
What is the variance of the controlled output after control is applied?
Note: Use RSS to estimate the variance of residual noise terms.

$$Var[X_{2_t}]^{\text{with control}} = \sigma_{a_2}^2 (1 + 0.7^2)$$

$$\sigma_{a_2}^2 = \frac{RSS}{N-r} = \frac{1900}{300-2} = 6.38$$

$$Var[X_{2_t}]^{\text{with control}} = 6.38 \cdot 1.49 = \underline{\underline{9.52}}$$

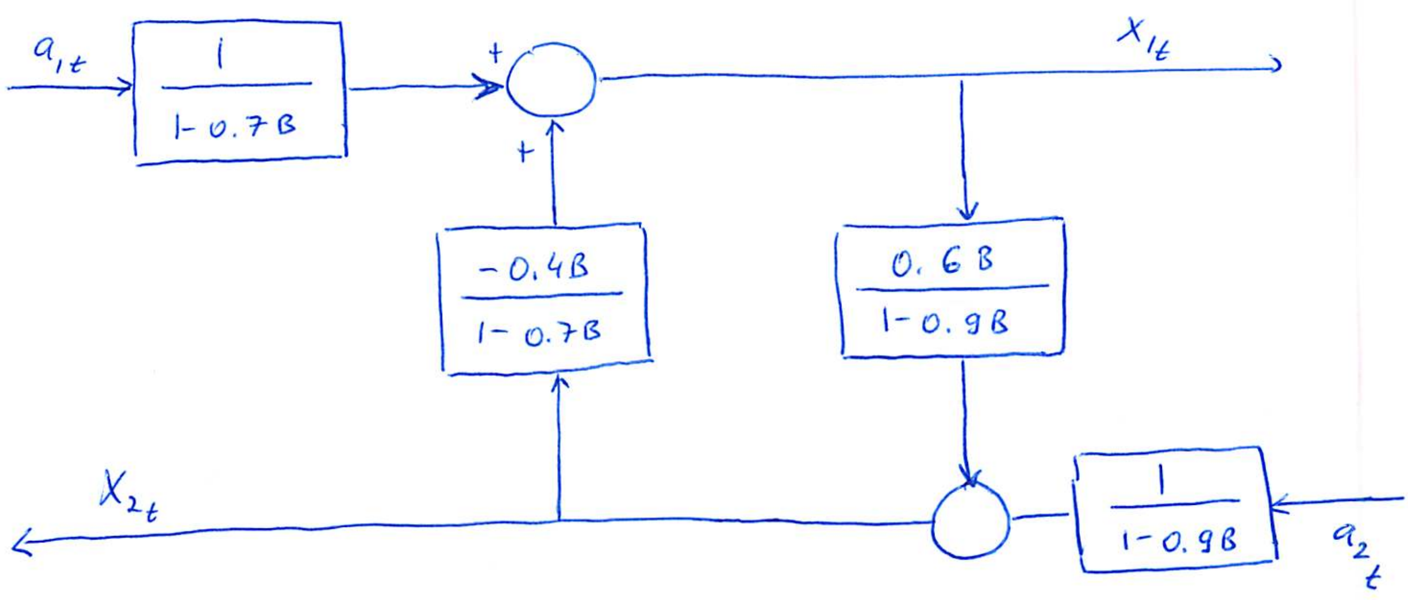
Problem 5. (Total 25 points)

It is found that N = 300 samples of two time series X_{1t} and X_{2t} can be modeled using the following vectorial ARMA model.

$$X_{1t} = 0.7X_{1t-1} - 0.4X_{2t-1} + a_{1t}$$
$$X_{2t} = 0.6X_{1t-1} + 0.9X_{2t-1} + a_{2t}$$

$\text{RSS}_1 = 100$ (corresponding to noise a_{1t}; $\text{RSS}_2 = 400$ (corresponding to noise a_{2t})

Part (a) (5 points)
Draw a block diagram of this system.



Part (b) (7 points)
If $X_{1t} = 3$ and $X_{2t} = 5$, please find 2-steps ahead predictions of both these time-series, starting at time t.

$$\hat{X}_{1_t}(1) = 0.7\,X_{1_t} - 0.4\,X_{2_t} = 0.1$$

$$\hat{X}_{2_t}(1) = 0.6\,X_{1_t} + 0.9\,X_{2_t} = 6.3$$

$$\hat{X}_{1_t}(2) = 0.7\,\hat{X}_{1_t}(1) - 0.4\,\hat{X}_{2_t}(1) = -2.45$$

$$\hat{X}_{2_t}(2) = 0.6\,\hat{X}_{1_t}(1) + 0.9\,\hat{X}_{2_t}(1) = 5.73$$

Part (c) (13 points)
Express the variance of the error of two-step ahead prediction considered in part (b).
Hint: Use residual sums of squares to obtain variances of both noise terms. Express both time-series using orthogonal decompositions regarding the noise terms a_{1t} and a_{2t}. You're on your own from then on (but solution should be easy after that).

$$(1 - 0.7B)\,X_{1_t} - 0.4B\,X_{2_t} = a_{1_t}$$

$$0.6B\,X_{1_t} + (1 - 0.9B)\,X_{2_t} = a_{2_t}$$

Solving for X_{1_t} and X_{2_t} yields

$$X_{1_t} = \frac{1 - 0.9B}{1 - 1.6B + 0.87B^2}\,a_{1_t} + \frac{-0.4B}{1 - 1.6B + 0.87B^2}\,a_{2_t}$$

$$X_{2_t} = \frac{0.6B}{1 - 1.6B + 0.87B^2}\,a_{1_t} + \frac{1 - 0.7B}{1 - 1.6B + 0.87B^2}\,a_{2_t}$$

Using Green's function like decomposition, we get

$$X_{1_t} = (G_{1_0}^{(1)} + G_{1_1}^{(1)} B + G_{1_2}^{(1)} B^2 + \dots) a_{1_t} +$$

$$+ (G_{1_0}^{(2)} + G_{1_1}^{(2)} B + G_{1_2}^{(2)} B^2 + \dots) a_{2_t}$$

$$X_{2_t} = (G_{2_0}^{(1)} + G_{2_1}^{(1)} B + G_{2_2}^{(1)} B^2 + \dots) a_{1_t} +$$

$$+ (G_{2_0}^{(2)} + G_{2_1}^{(2)} B + G_{2_2}^{(2)} B^2 + \dots) a_{2_t}$$

$$Var[\hat{e}_{1_t}(2)] = Var[a_{1_t}]([G_{1_0}^{(1)}]^2 + [G_{1_1}^{(1)}]^2) +$$

$$+ Var[a_{2_t}]([G_{1_0}^{(2)}]^2 + [G_{1_1}^{(2)}]^2) \quad \dots\dots \quad (1)$$

$$Var[\hat{e}_{2_t}(2)] = Var[a_{1_t}]([G_{2_0}^{(1)}]^2 + [G_{2_1}^{(1)}]^2) +$$

$$+ Var[a_{2_t}]([G_{2_0}^{(2)}]^2 + [G_{2_1}^{(2)}]^2) \quad \dots\dots \quad (2$$

$$[Var\, a_{1_t}] \approx \frac{RSS_1}{N - r} = \frac{100}{300 - 2} = 0.338$$

Number of estimated parameters in the first model eqn.

$$Var[a_{2_t}] \approx \frac{RSS_2}{N - r} = \frac{400}{300 - 2} = 1.351$$

Number of estimated parameters in the 2nd model eqn.

Using implicit method, we get

$$G_{1_0}^{(1)} = 1 \qquad G_{1_1}^{(1)} = -0.7$$

$$G_{1_0}^{(2)} = 0 \qquad G_{1_1}^{(2)} = -0.4$$

$$G_{2_0}^{(1)} = 0 \qquad G_{2_1}^{(1)} = 0.6$$

$$G_{2_0}^{(2)} = 1 \qquad G_{2_1}^{(2)} = -0.9$$

Hence, plugging these values into (1) and (2) gives

$$Var[\hat{e}_{1_t}(2)] = 0.7196$$

$$Var[\hat{e}_{2_t}(2)] = 2.5676$$