EE627A Midterm

Name:

Q1. An MA(2) process is defined by $X_t = \epsilon_t + a_1\epsilon_{t-1} + a_2\epsilon_{t-2}$. Determine

1. a The expectation $E[X_t]$.
1. b The variance $\text{var}[X_t]$
1. c The first-order autocorrelation $\rho(1)$
1. d What is the value for the 3rd order autocorrelation $\rho(3)$?

a. $X_t = \epsilon_t + a_1\epsilon_{t-1} + a_2\epsilon_{t-2}$

$E(X_t) = E(\epsilon_t + a_1\epsilon_{t-1} + a_2\epsilon_{t-2})$

$= E(\epsilon_t) + a_1E(\epsilon_{t-1}) + a_2E(\epsilon_{t-2})$

$= 0 + 0 + 0 = 0$

b. $Var(X_t) = V_0 = E(X_t^2) - E^2(X_t)$

$= E((\epsilon_t + a_1\epsilon_{t-1} + a_2\epsilon_{t-2})^2) - 0^2$

$= E(\epsilon_t^2) + a_1^2 \cdot E(\epsilon_{t-1}^2) + a_2^2 E(\epsilon_{t-2}^2)$

$= (1 + a_1^2 + a_2^2)\sigma_t^2$

σ_t^2 is the variance of ϵ_t.

c. $V_k = E(x_t x_{t-k})$

$V_0 = E[(\epsilon_t + a_1\epsilon_{t-1} + a_2\epsilon_{t-2})^2] = (1 + a_1^2 + a_2^2)\sigma_\epsilon^2.$

$V_1 = E[(\epsilon_t + a_1\epsilon_{t-1} + a_2\epsilon_{t-2})(\epsilon_{t-1} + a_1\epsilon_{t-2} + a_2\epsilon_{t-3})] = (a_1 + a_1a_2)\sigma_t^2$

$\therefore \rho(1) = \frac{V_1}{V_0} = \frac{a_1 + a_1a_2}{1 + a_1^2 + a_2^2}$

d. $V_2 = E[(\epsilon_t + a_1\epsilon_{t-1} + a_2\epsilon_{t-2})(\epsilon_{t-2} + a_1\epsilon_{t-3} + a_2\epsilon_{t-4})] = a_2\sigma_t^2$

$V_3 = E[(\epsilon_t + a_1\epsilon_{t-1} + a_2\epsilon_{t-2})(\epsilon_{t-3} + a_1\epsilon_{t-4} + a_2\epsilon_{t-5})] = 0.$

$\therefore \rho_3 = \frac{V_3}{V_0} = 0$

Q2. (a) Find the autocorrelation function(ACF) of the time series $X_t = W_t + \frac{5}{2}W_{t-1} - \frac{3}{2}W_{t-2}$, where the Gaussian noise $W_t \sim \mathcal{N}(0,1)$.

(b) Find the autocorrelation function(ACF) of the time series $X_t = W_t - \frac{1}{6}W_{t-1} - \frac{1}{6}W_{t-2}$, where the Gaussian noise $W_t \sim \mathcal{N}(0,9)$.

a. $V_k = E(X_t X_{t+k})$

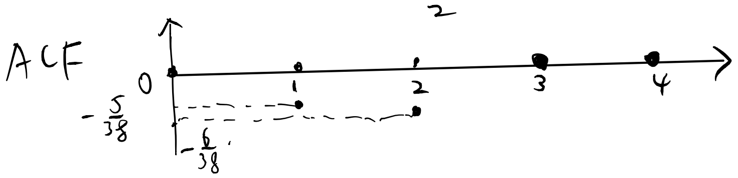
$V_0 = E[(W_t + \frac{5}{2}W_{t-1} - \frac{3}{2}W_{t-2})^2] = (1+(\frac{5}{2})^2+(-\frac{3}{2})^2)\sigma_t^2 = \frac{19}{2}\cdot\sigma_t^2 = \frac{19}{2}$

$V_1 = E[(W_t + \frac{5}{2}W_{t-1} - \frac{3}{2}W_{t-2})(W_{t-1} + \frac{5}{2}W_{t-2} - \frac{3}{2}W_{t-3})] = (\frac{5}{2} - \frac{3}{2}\times\frac{5}{2})\sigma_t^2 = -\frac{5}{4}$

$V_2 = E[(W_t + \frac{5}{2}W_{t-1} - \frac{3}{2}W_{t-2})(W_{t-2} + \frac{5}{2}W_{t-3} - \frac{3}{2}W_{t-4})] = -\frac{3}{2}\sigma_t^2 = -\frac{3}{2}$

$V_3 = 0 \quad V_4 = 0 \quad \cdots$

$\therefore r_0 = 1 \quad r_1 = \frac{-\frac{5}{4}}{\frac{19}{2}} = -\frac{5}{38} \quad r_2 = \frac{-\frac{3}{2}}{\frac{19}{2}} = -\frac{3}{19}$



b. $V_k = E(X_t X_{t+k})$

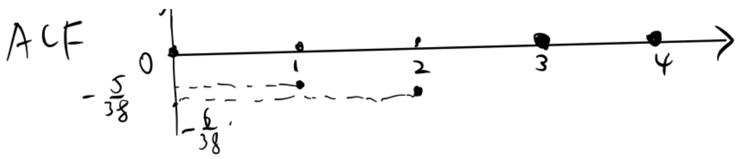
$V_0 = E[(W_t - \frac{1}{6}W_{t-1} - \frac{1}{6}W_{t-2})^2] = (1+(-\frac{1}{6})^2+(-\frac{1}{6})^2)\sigma_t^2 = \frac{19}{18}\sigma_t^2 = \frac{19}{18}\times 81 = \frac{171}{2}$

$V_1 = E[(W_t - \frac{1}{6}W_{t-1} - \frac{1}{6}W_{t-2})(W_{t-1} - \frac{1}{6}W_{t-2} - \frac{1}{6}W_{t-3})] = (-\frac{1}{6} + (-\frac{1}{6})^2)\sigma_t^2 = -\frac{5}{36}\sigma_t^2 = -\frac{45}{4}$

$V_2 = E[(W_t - \frac{1}{6}W_{t-1} - \frac{1}{6}W_{t-2})(W_{t-2} - \frac{1}{6}W_{t-3} - \frac{1}{6}W_{t-4})] = (-\frac{1}{6})\sigma_t^2 = -\frac{1}{6}\sigma_t^2 = -\frac{27}{2}$

$V_3 = 0 \quad V_4 = 0 \quad \cdots$

$\therefore r_0 = 1 \quad r_1 = \frac{-\frac{45}{4}}{\frac{171}{2}} = -\frac{5}{38} \quad r_2 = \frac{-\frac{27}{2}}{\frac{171}{2}} = -\frac{3}{19}$



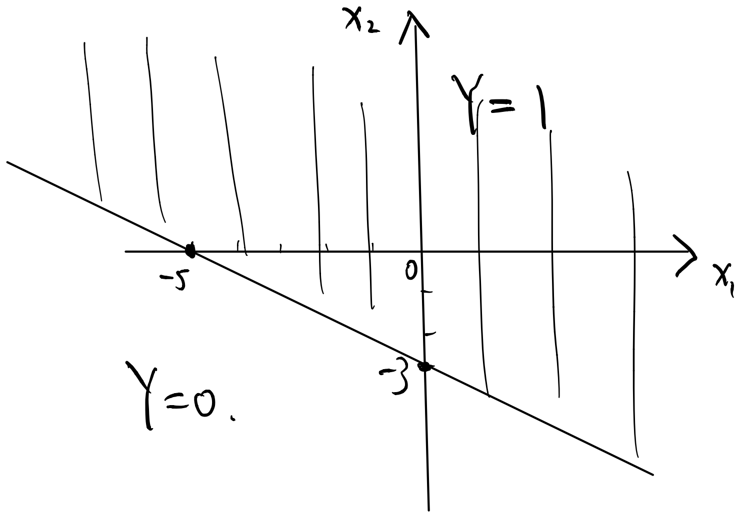
Q3. Suppose you are given the following classification task: predict the target $Y \in \{0, 1\}$, given two real valued features $X_1 \in \mathcal{R}$ and $X_2 \in \mathcal{R}$. After some training, you learn the following decision rule:

Predict $Y = 1$ if $w_1 X_1 + w_2 X_2 + w_0 \geq 0$ and $Y = 0$ otherwise, where $w_1 = 3$, $w_2 = 5$, and $w_0 = 15$.

Plot the decision boundary on the $\{x_1, x_2\}$ feature space and label the region where we would predict $Y = 1$ and $Y = 0$.

$Y=1$ if $w_1 X_1 + w_2 X_2 + w_0 \geq 0$.

$\Rightarrow \quad 3X_1 + 5X_2 + 15 \geq 0$.



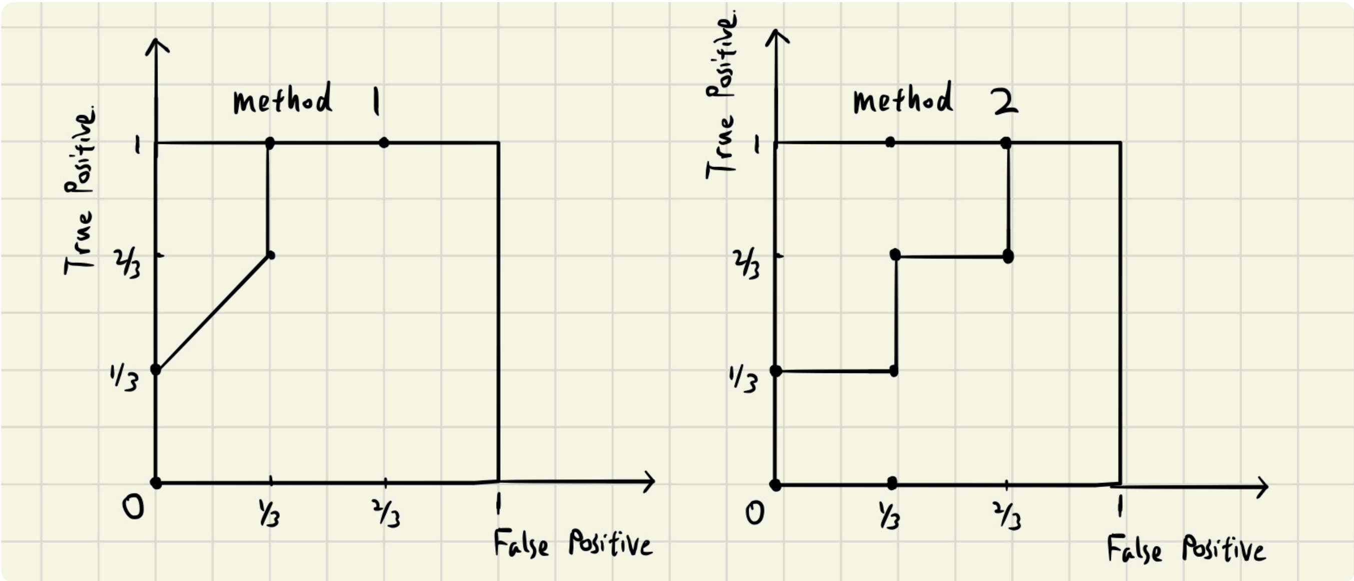
Q4. We have two different methods to predict the college enrollments in a selected school. Given the predictive scores and real outcomes for each student, please plot the ROC curves for each method and compare the AUC values. Which predictive method is relatively better in terms of AUC value?

Table 1: Predictive Scores for Collenge Enrollments

Student	Enrollment Decision	Method 1	Method 2
1	Yes	0.99	0.85
2	No	0.90	0.70
3	Yes	0.90	0.73
4	Yes	0.86	0.65
5	No	0.85	0.78
6	No	0.70	0.55

	ThreshHold	1	0.99	0.9	0.86	0.85	0.7
Method1	False Positive	0	0	1/3	1/3	2/3	1
	True Positve	0	1/3	2/3	1	1	1

	ThreshHold	1	0.85	0.78	0.73	0.7	0.65	0.55
Method2	False Positive	0	0	1/3	1/3	2/3	2/3	1
	True Positve	0	1/3	1/3	2/3	2/3	1	1

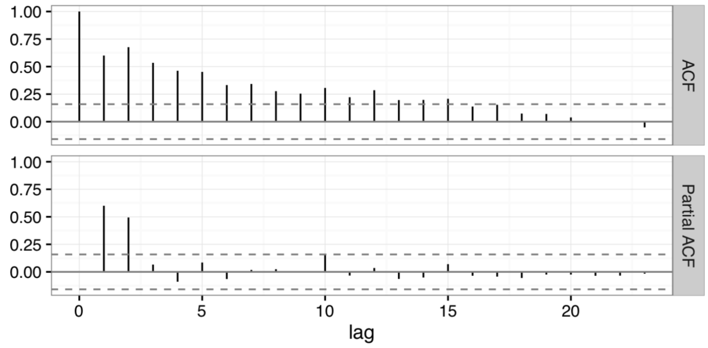


Method 1: AUC: $1 - 1 \times \frac{1}{3} \times \frac{1}{2} = \frac{5}{6}$

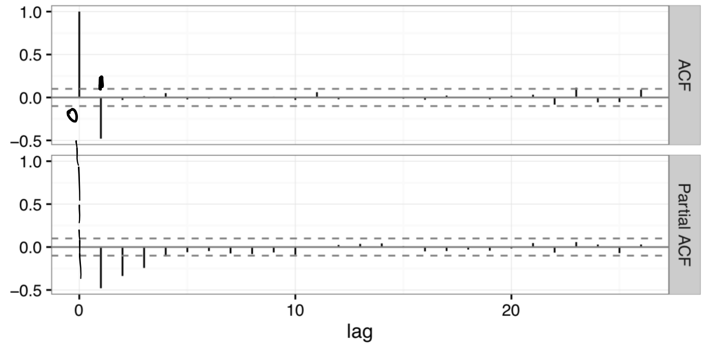
Method 2: AUC: $1 - 3 \times \frac{1}{3} \times \frac{1}{3} = \frac{2}{3}$

Obviously AUC of method 1 is greater than of method 2

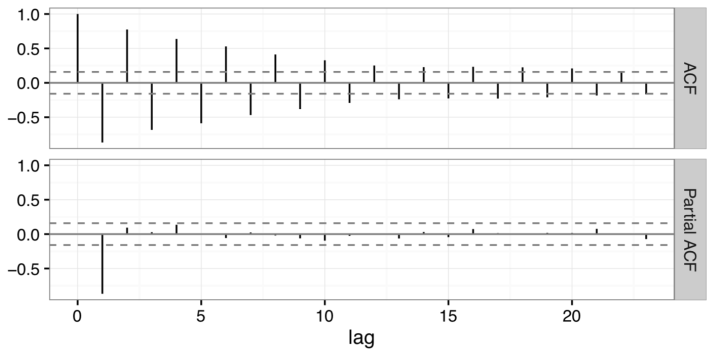
So, method 1 is better.



(a)



(b)



(c)

Q5. Above are the ACF and PACF for three time series. For each series, state whether it is autoregressive or moving average, and the order (p or q).

a. autoregressive AR(2)
$p=2$.

b. moving average. MA(1).
$q=1$

c. autoregressive. AR(1)
$p=1$

Q6. For a given predictor matrix $\mathbf{X} = [\mathbf{x}_1,\ \mathbf{x}_2,\ \cdots]$ (say, students scores list), and the binary response vector $\mathbf{y}$ (students enrolled to college or not), if we transform every column vector $\mathbf{x}_i$ in $\mathbf{X}$ to a normalized new vector

$$\mathbf{z}_i = \frac{\mathbf{x}_i - \text{mean}(\mathbf{x}_i)}{\text{standard deviation of } (\mathbf{x}_i)}$$

Does this normalization will change the AUC performance of logistic regression results? If yes, specify what the changes are? If not, explain why not?

Not change the AUC performance.

Because mean (x_i) and standard deviation of (x_i) are invariant. So this transform of x_i is a linear transform of x_i. Obviously, linear transform can not change the AUC performance of logistic regression.