PROJECT----WEEK02

Tianai Zhao

Problem 1

Compare the conditional distribution of the Multivariate Normal, to the OLS equations.
Are these values the same? Why?
Use the data in problem1.csv to prove your answer empirically.

Multivariate Normal (2 variables): $\mathbb{E}(Y|X=x) = \mu_Y + \frac{cov(X,Y)}{Var(X)}(x-\mu_X)$

OLS: $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$

$\mathbb{E}(\hat{Y}|X=x) = \hat{\beta}_0 + \hat{\beta}_1 x$

where $\begin{cases} \hat{\beta}_1 = \frac{Cov(X,Y)}{Var(X)} \\ \hat{\beta}_0 = \mu_Y - \hat{\beta}_1 \mu_X \end{cases}$

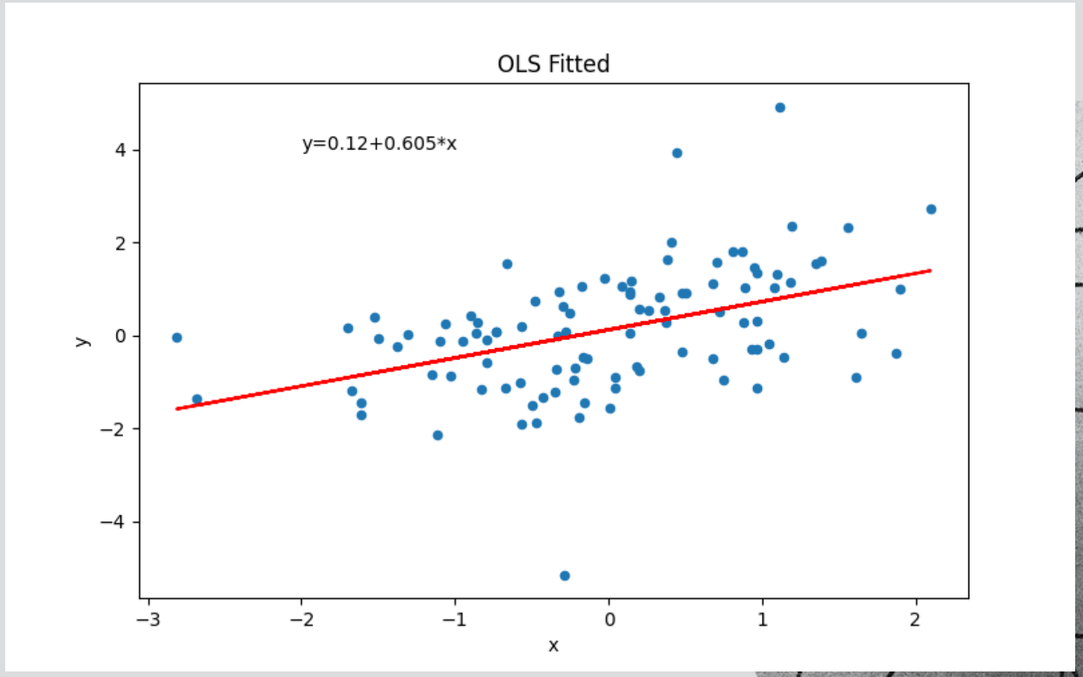
$\therefore \mathbb{E}(\hat{Y}|X=x) = \mu_Y + \frac{cov(X,Y)}{Var(X)}(x-\mu_X)$

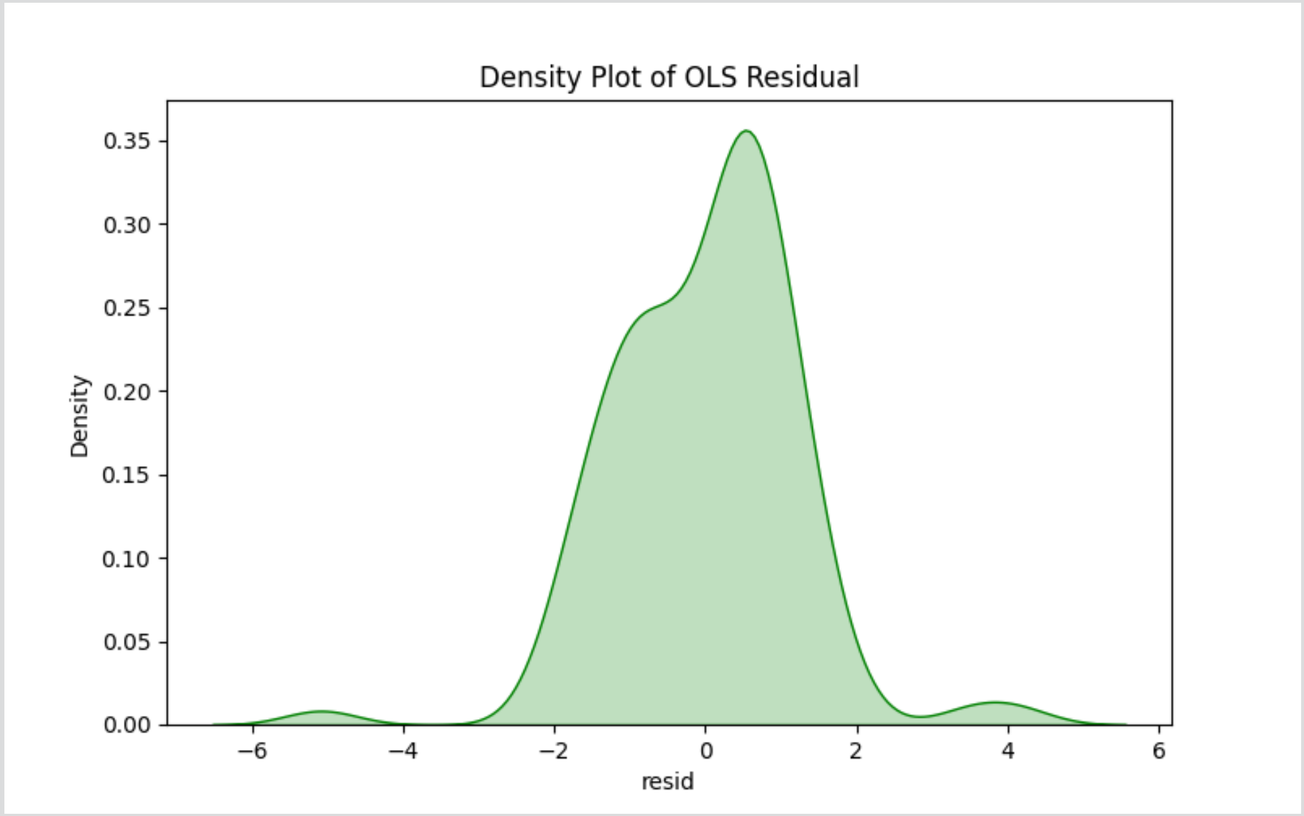
is same as the conditional mean in multivariant normal

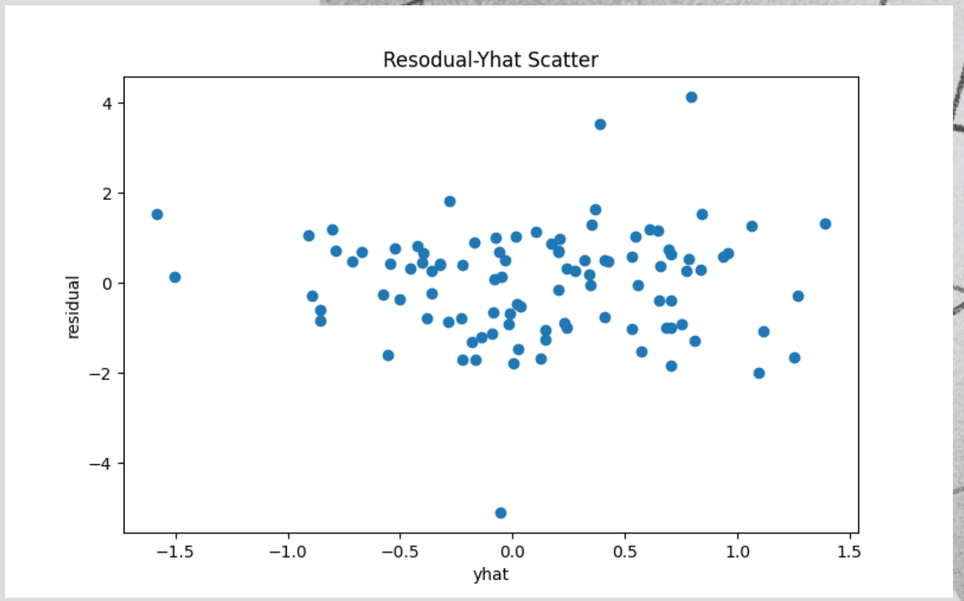
	x	y	yhat_OLS	yhat_multinorm
0	-1.166289	1.014680	-0.461299	-0.461299
1	-0.426878	0.262715	-0.144828	-0.144828
2	-1.477892	-1.044772	-0.594666	-0.594666
3	3.049119	0.804363	1.342911	1.342911
4	-2.123732	-0.689514	-0.871088	-0.871088
..	...	...	...	...
95	-0.588599	0.652704	-0.214046	-0.214046
96	-0.218138	0.067676	-0.055487	-0.055487
97	0.342822	1.214472	0.184606	0.184606
98	0.337376	0.608974	0.182275	0.182275
99	1.153817	-0.683444	0.531715	0.531715

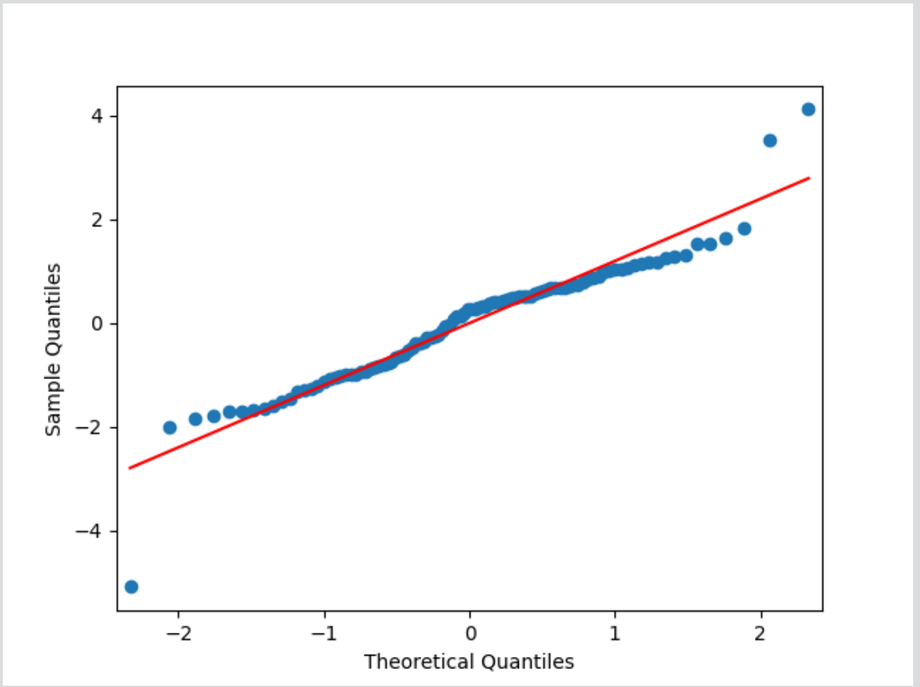
Problem 2

Fit the data in problem2.cv using OLS and calculate the error vector. Look at it's distribution.
How well does it fit the assumption of normally distributed errors?
Fit the data using MLE given the assumption of normality. Then fit the MLE using the assumption of a T distribution of the errors. Which is the best fit?
What are the fitted parameters of each and how do they compare? What does this tell us about the breaking of the normality assumption in regards to expected values in this case?




Density Plot of OLS Residual
Density
resid
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
−6
−4
−2
0
2
4
6


Resodual-Yhat Scatter
residual
yhat
4
2
0
−2
−4
−1.5
−1.0
−0.5
0.0
0.5
1.0
1.5

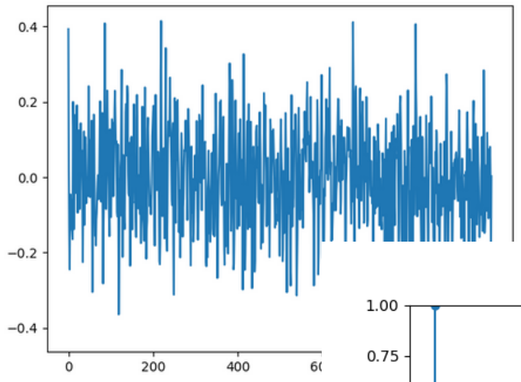

Sample Quantiles
Theoretical Quantiles
4
2
0
−2
−4
−2
−1
0
1
2

Comparing results from the three models

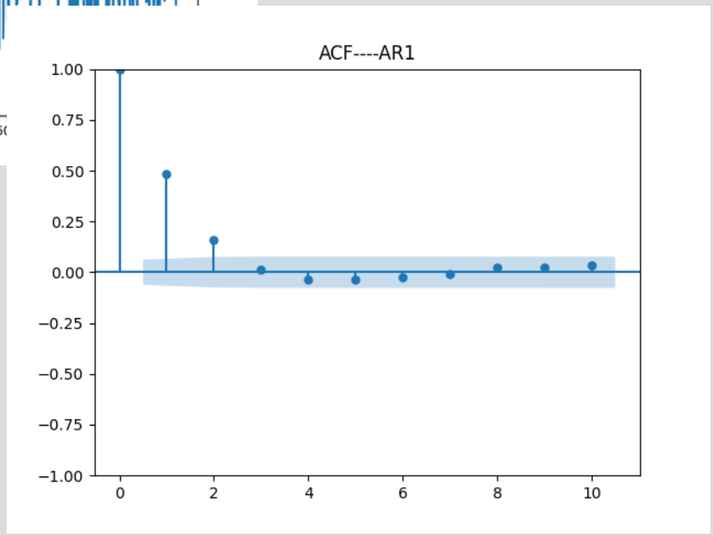
	β0	β1	Log Likelihood
OLS	0.1198362	0.6052048	-159.99
MLE (Assume normality)	0.1198362	0.6052048	-159.99209668916234
MLE (Assume T distribution)	0.12325311	0.59512445	-161.75964926470715

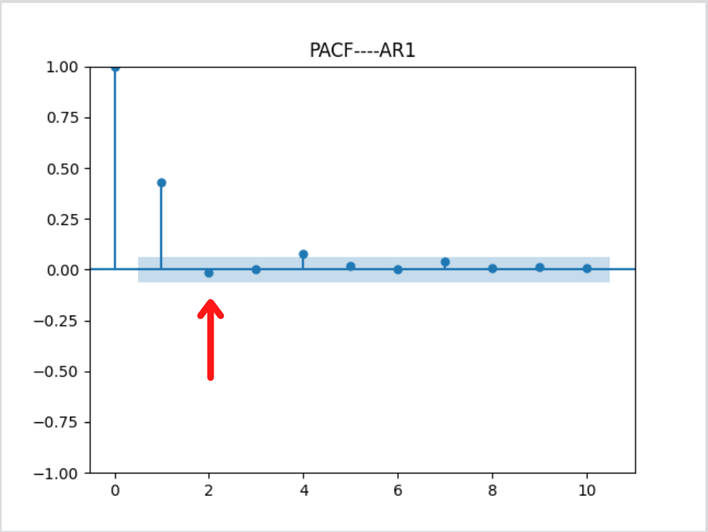
Problem 3

Simulate AR(1) through AR(3) and MA(1) through MA(3) processes. Compare their ACF and PACF graphs. How do the graphs help us to identify the type and order of each process?

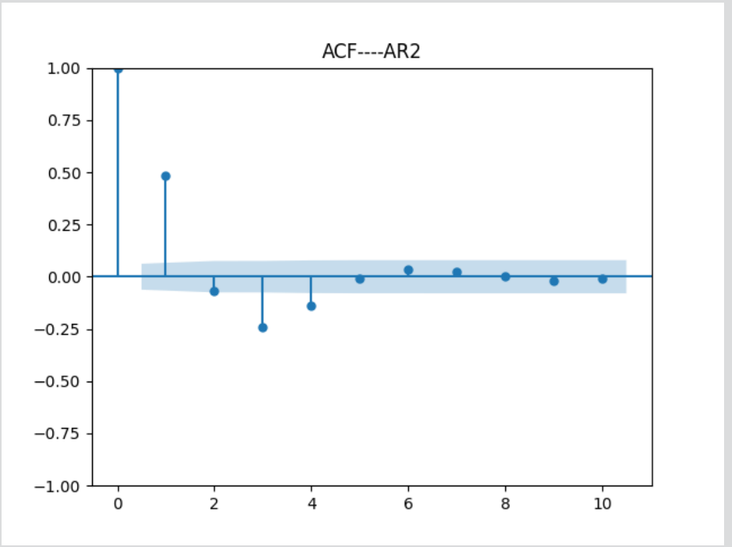


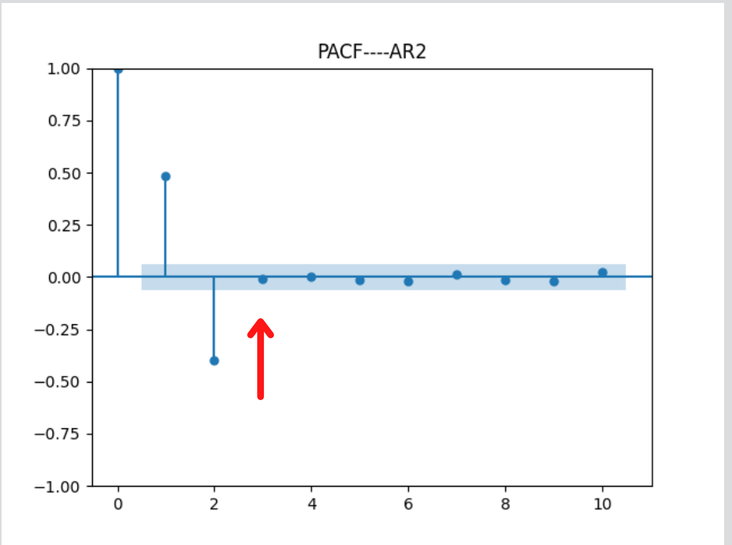
Yt = 0.5Yt-1 + et



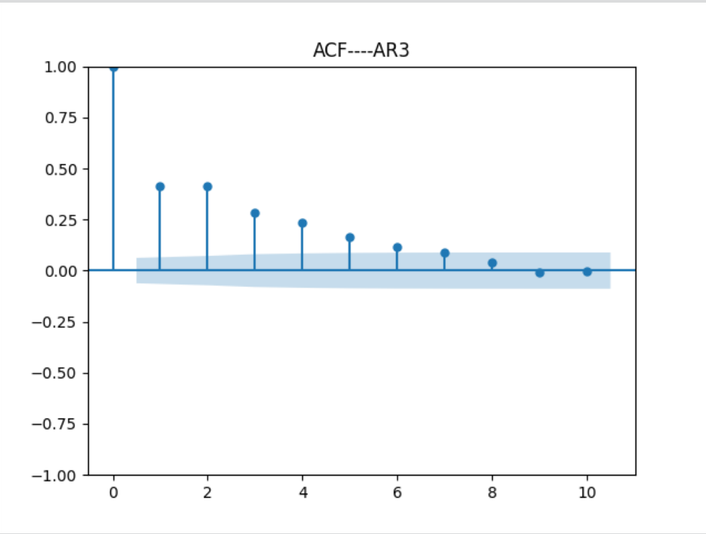


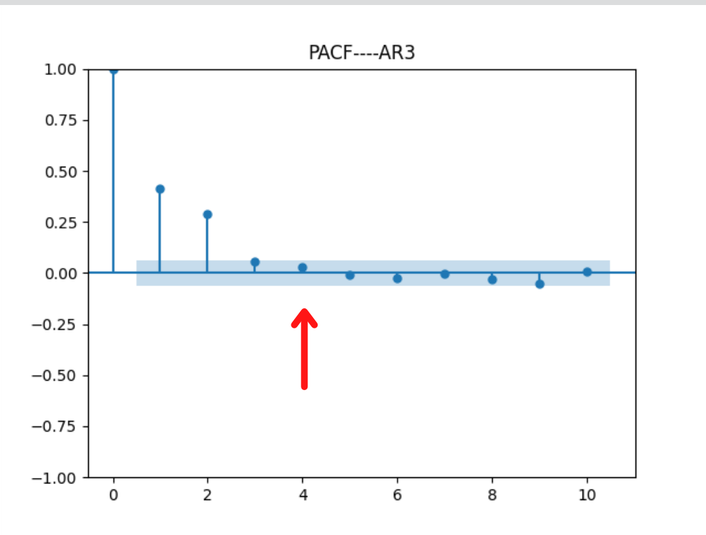
Yt = 0.7Yt-1 - 0.4Yt-2 + et

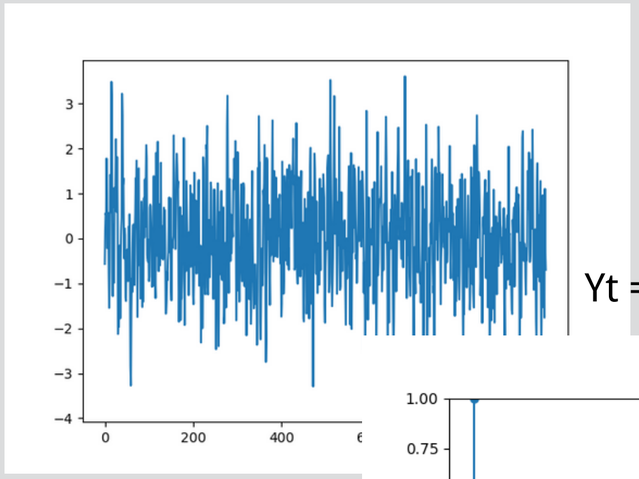




Yt = 0.3Yt-1 + 0.2Yt-2 + 0.1Yt-3 + et



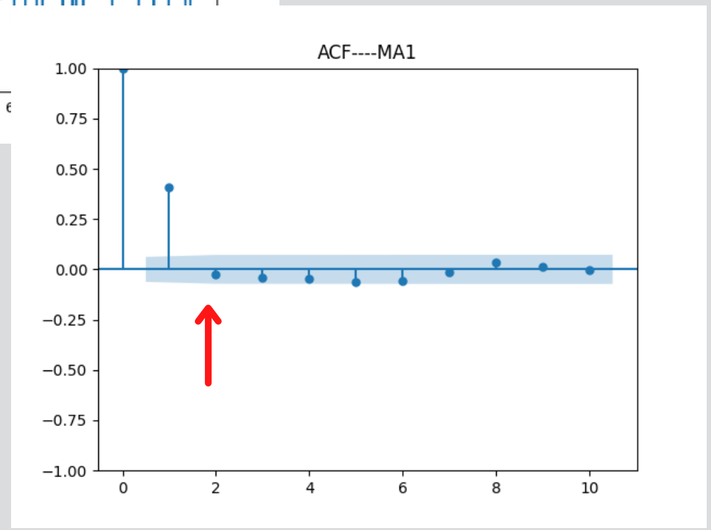


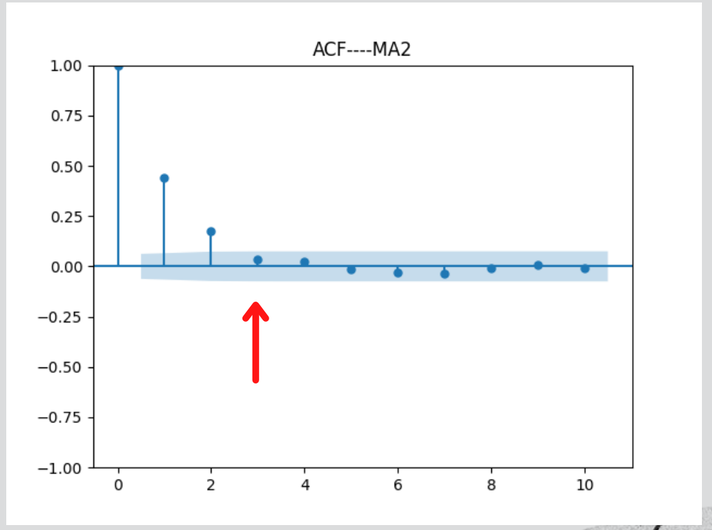


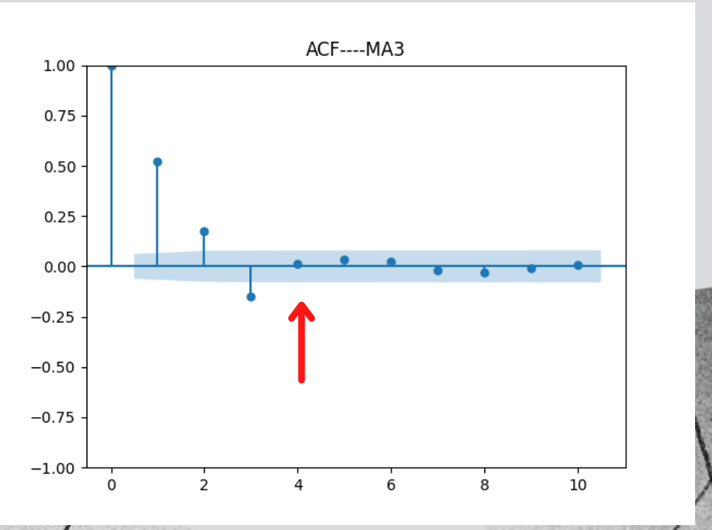
Yt = 0.5et-1 + et

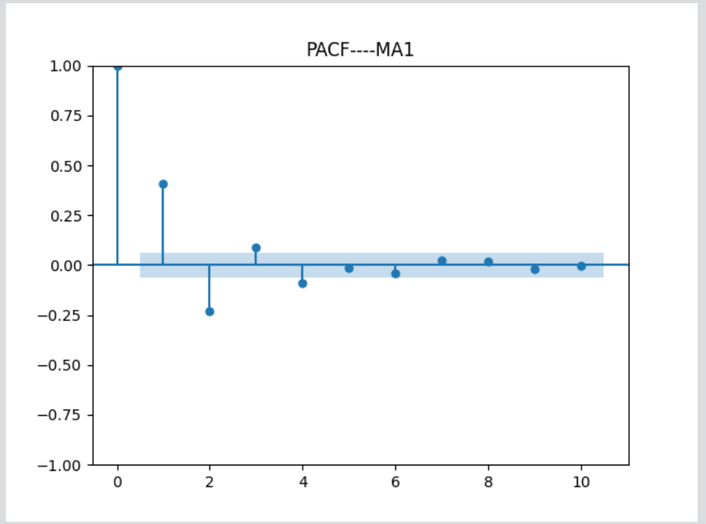
Yt = 0.5et-1 - 0.2et-2 + et

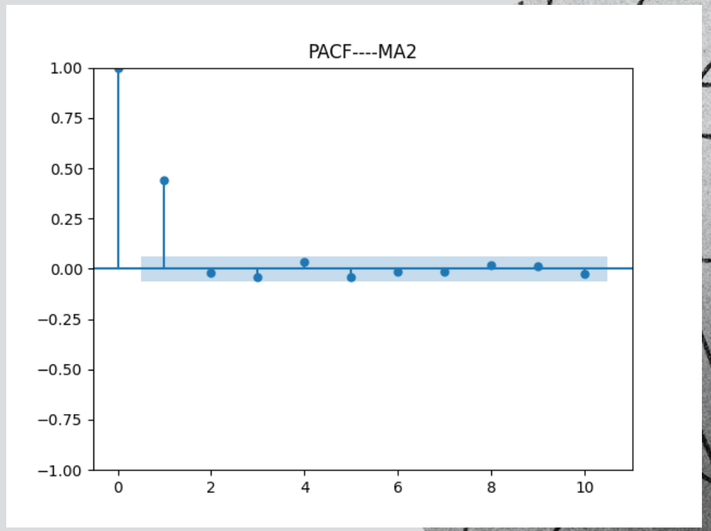
Yt = 0.7et-1 + 0.5et-2 - 0.2et-3 + et

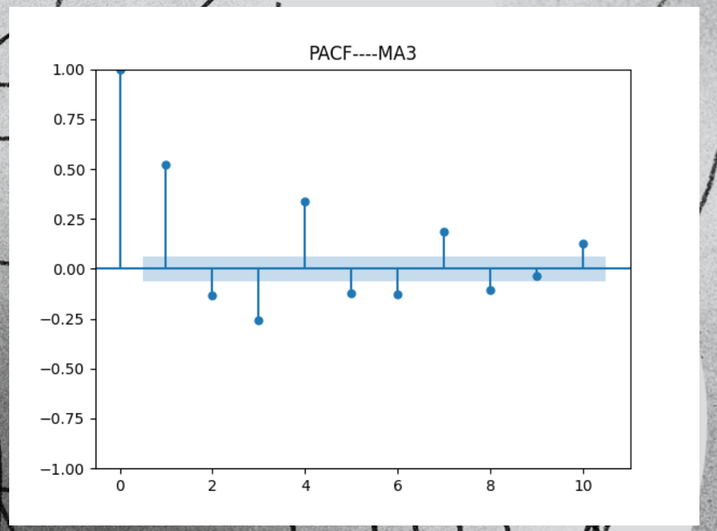












THANKS FOR LISTENING