Problem1:

A.

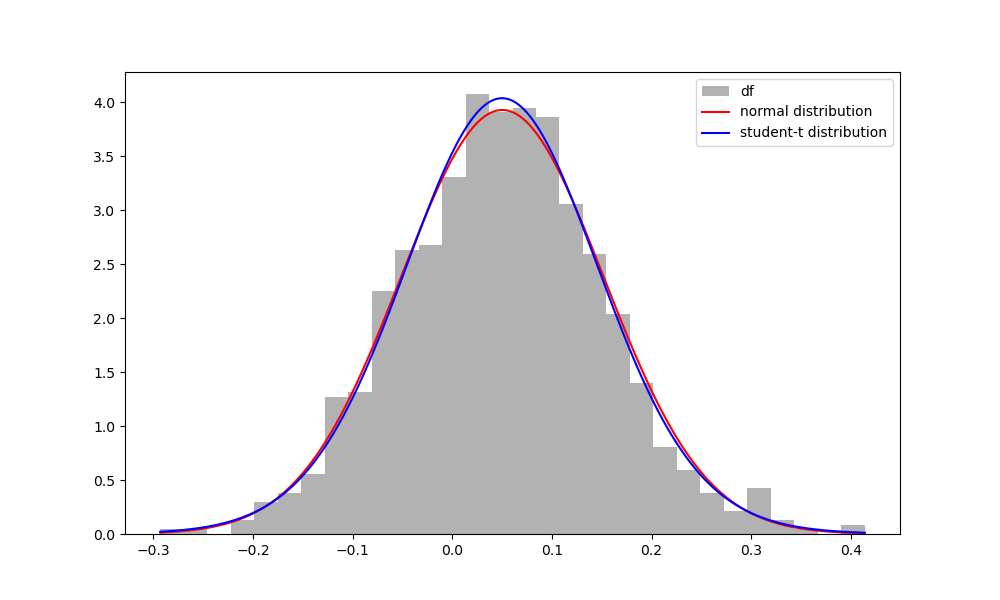
mean: 0.05019795790476916

variance: 0.010322143931072104

skewness: 0.1204447119194402

kurtosis: 0.2229270674503816

B.



Since the number of observations in the sample is 1000, it is a large sample (n > 30), and according to CLT, the sample follows a normal distribution.

C.

normal distribution:

maximum likehood: 867.793364418254

AIC: -863.793364418254

BIC: -1721.7712182785438

student-t distribution:

maximum likehood: 868.7091844597628

AIC: -862.7091844597628

BIC: -1716.6951030825792  
  
From the results of the information criterion, both AIC and BIC obtained by choosing normal distribution for fitting are lower, so it is more appropriate to choose normal distribution for fitting.

Problem2:

A.

corvarience matrix:

x1 x2 x3 x4 x5

x1 1.470484 1.454214 0.877269 1.903226 1.444361

x2 1.454214 1.252078 0.539548 1.621918 1.237877

x3 0.877269 0.539548 1.272425 1.171959 1.091912

x4 1.903226 1.621918 1.171959 1.814469 1.589729

x5 1.444361 1.237877 1.091912 1.589729 1.396186

B.

eigenvalues: [-0.31024286 -0.13323183 0.02797828 0.83443367 6.78670573]

Because some of the eigenvalues of the covariance matrix are less than 0, so the matrix is not at least semi-definite.

C.

nearest PSD matrix:

[[1.61519556 0.82885332 0.89883015 0.98871191 3.16532876]

[0.82885332 0.44470728 0.33691931 0.49599596 1.53759566]

[0.89883015 0.33691931 1.31824754 0.62230191 2.3884661 ]

[0.98871191 0.49599596 0.62230191 0.61193368 1.98545858]

[3.16532876 1.53759566 2.3884661 1.98545858 6.83687395]]

D.

covariance matrix using overlapping data:

x1 x2 x3 x4 x5

x1 0.418604 0.394054 0.424457 0.416382 0.434287

x2 0.394054 0.396786 0.409343 0.398401 0.422631

x3 0.424457 0.409343 0.441360 0.428441 0.448957

x4 0.416382 0.398401 0.428441 0.437274 0.440167

x5 0.434287 0.422631 0.448957 0.440167 0.466272

E.

Although the covariance matrix calculated by overlapping data is at least semi-definite, the Frobenius norm of difference between this two matrix is 8.61, which indicates the elements value of these two matrix have big difference.

The eigenvalues of near-PSD matrix is [5.37972191e-09, 8.98957098e-09, 5.12099767e-02, 6.92140941e-01, 1.00836071e+01], while the eigenvalues of covariance matrix using overlapping data is [0.0041588, 0.00519785, 0.01112963, 0.01897221, 2.12083697].This indicates a big difference between these two matrices.

Problem3:

μ=[0.04600157 0.09991502]

σ=[[0.0101622 0.00492354]

[0.00492354 0.02028441]]

B.

Conditional distribution method:

μ=0.3683249958609775, σ=0.133787030930085

X2|X1=0.6 ~ N(0.3683249958609775, 0.133787030930085)

Linear Regression method:

X2|X1=0.6 ~ N(0.36832499586097767, 0.017881070675442437)

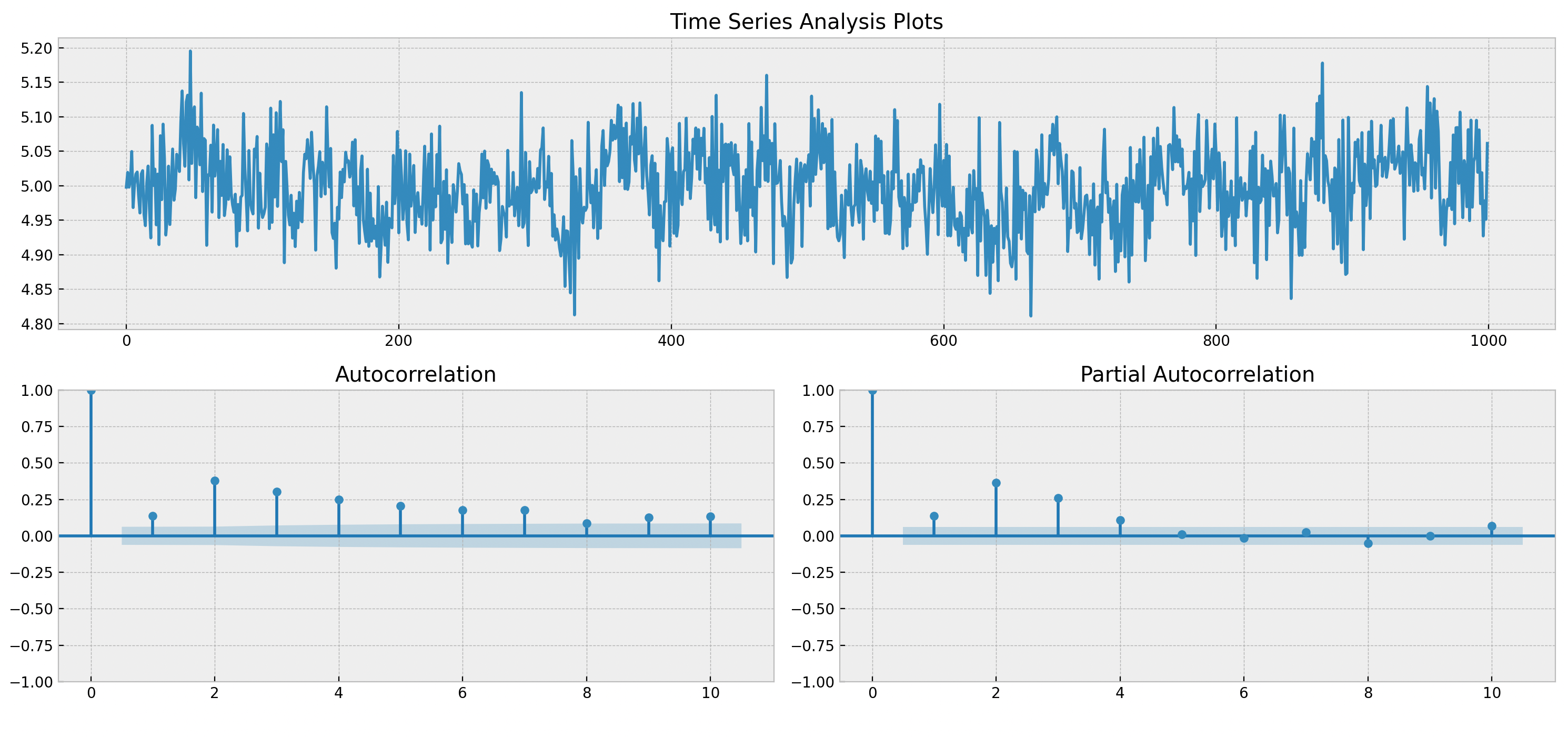
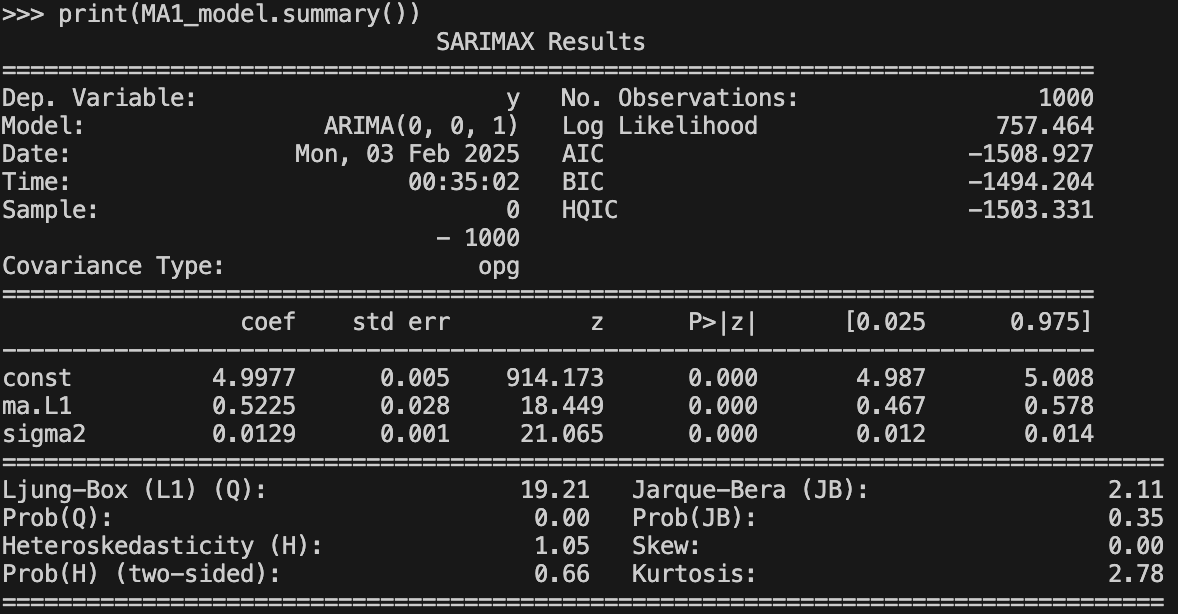
C.

after simulated, X2|X1=0.6 ~ N(0.370669366901806, 0.1302521478109876)

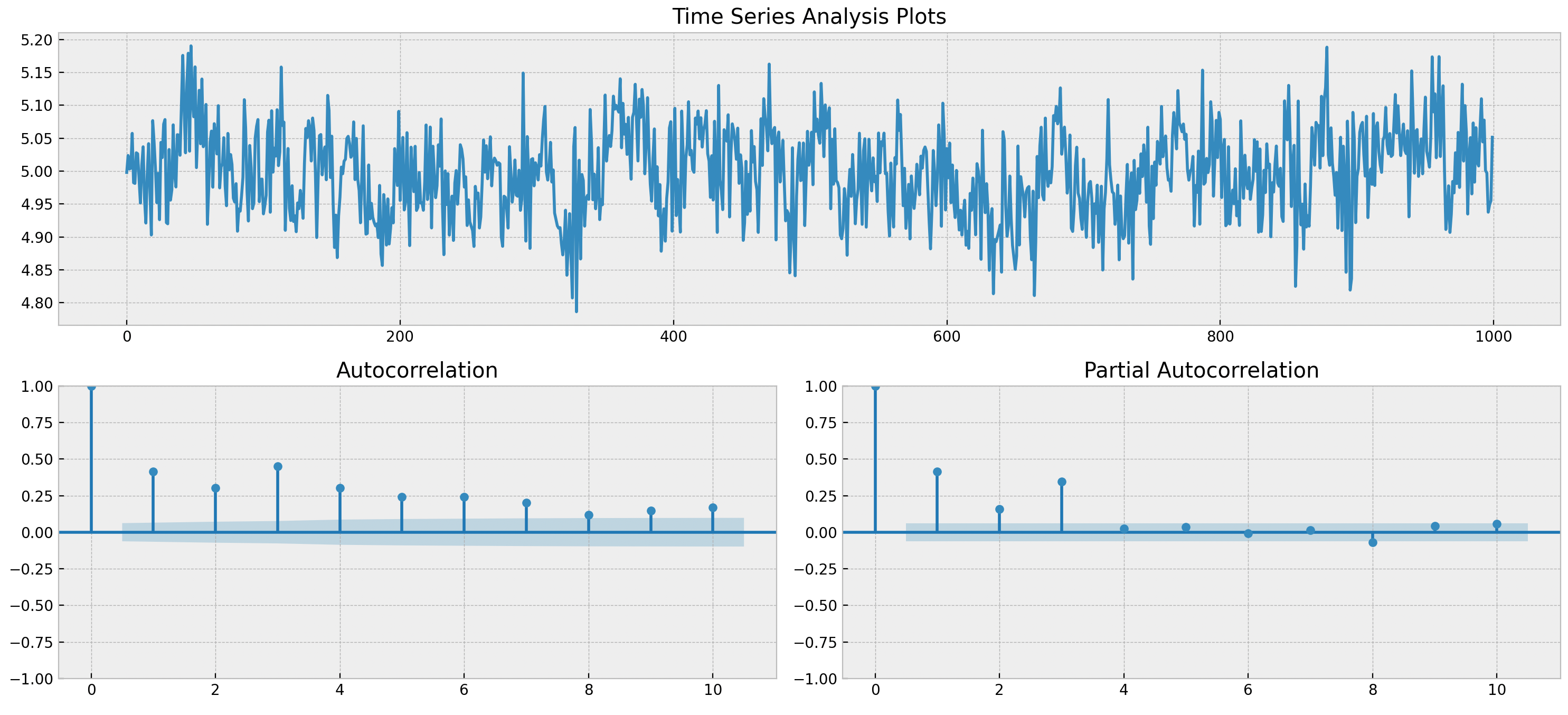
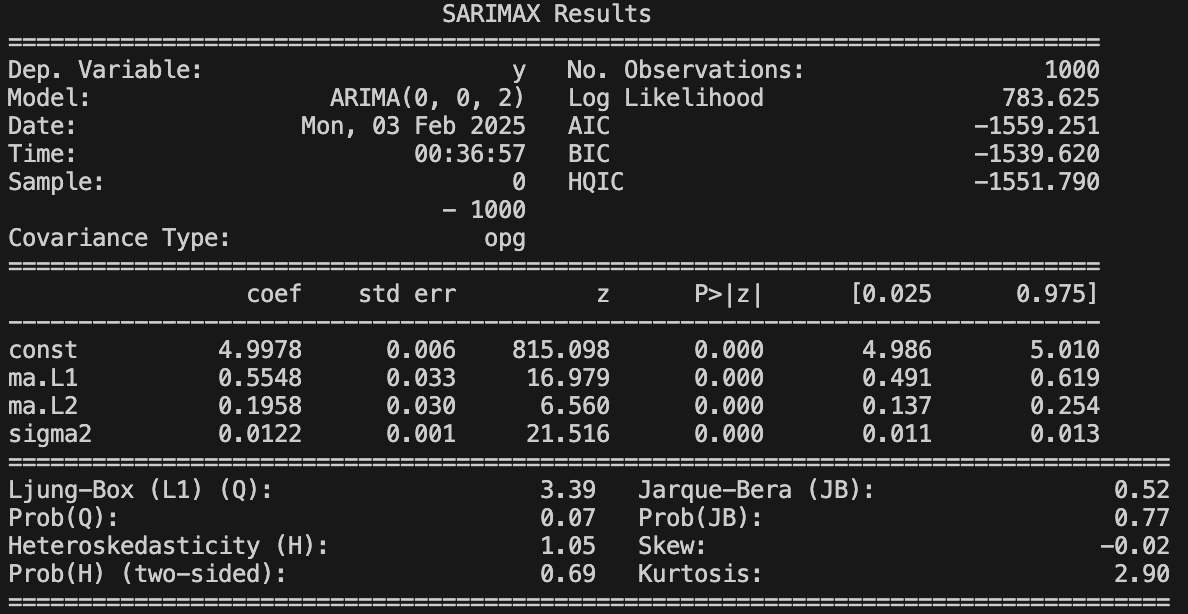
#Problem4

A.

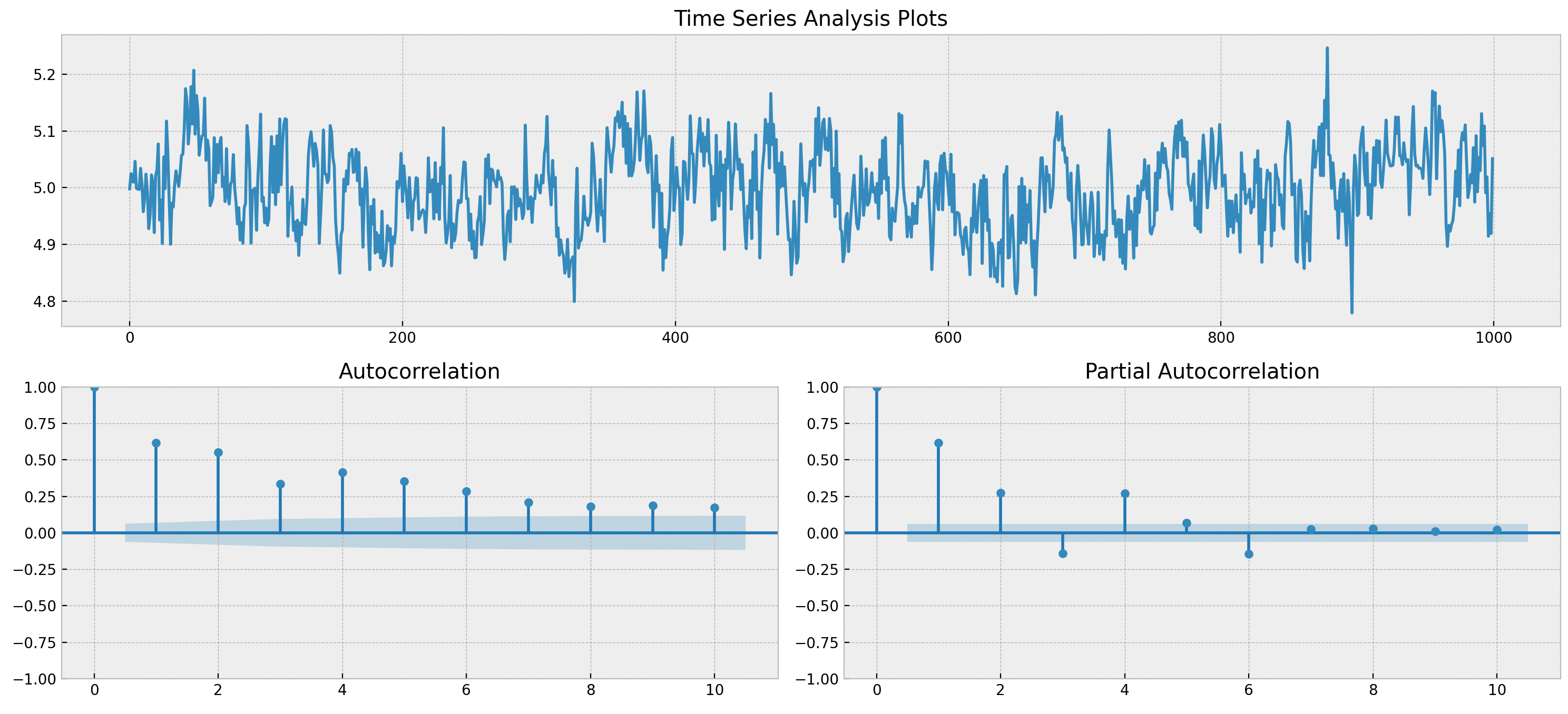
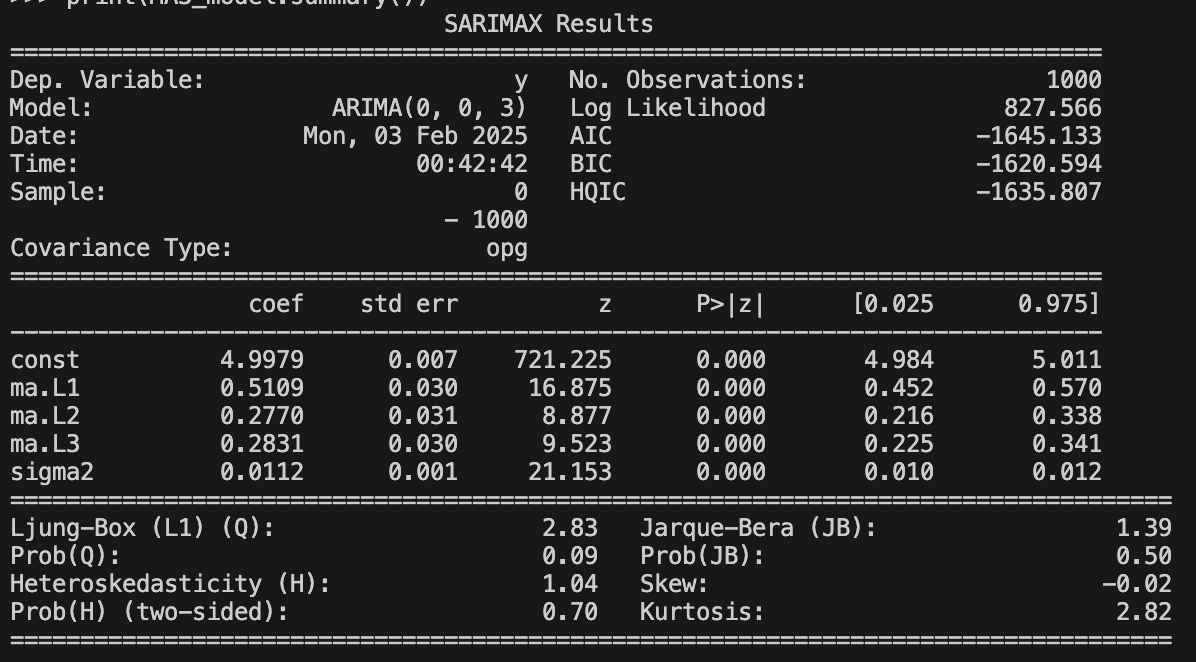
MA(1):



MA(2):



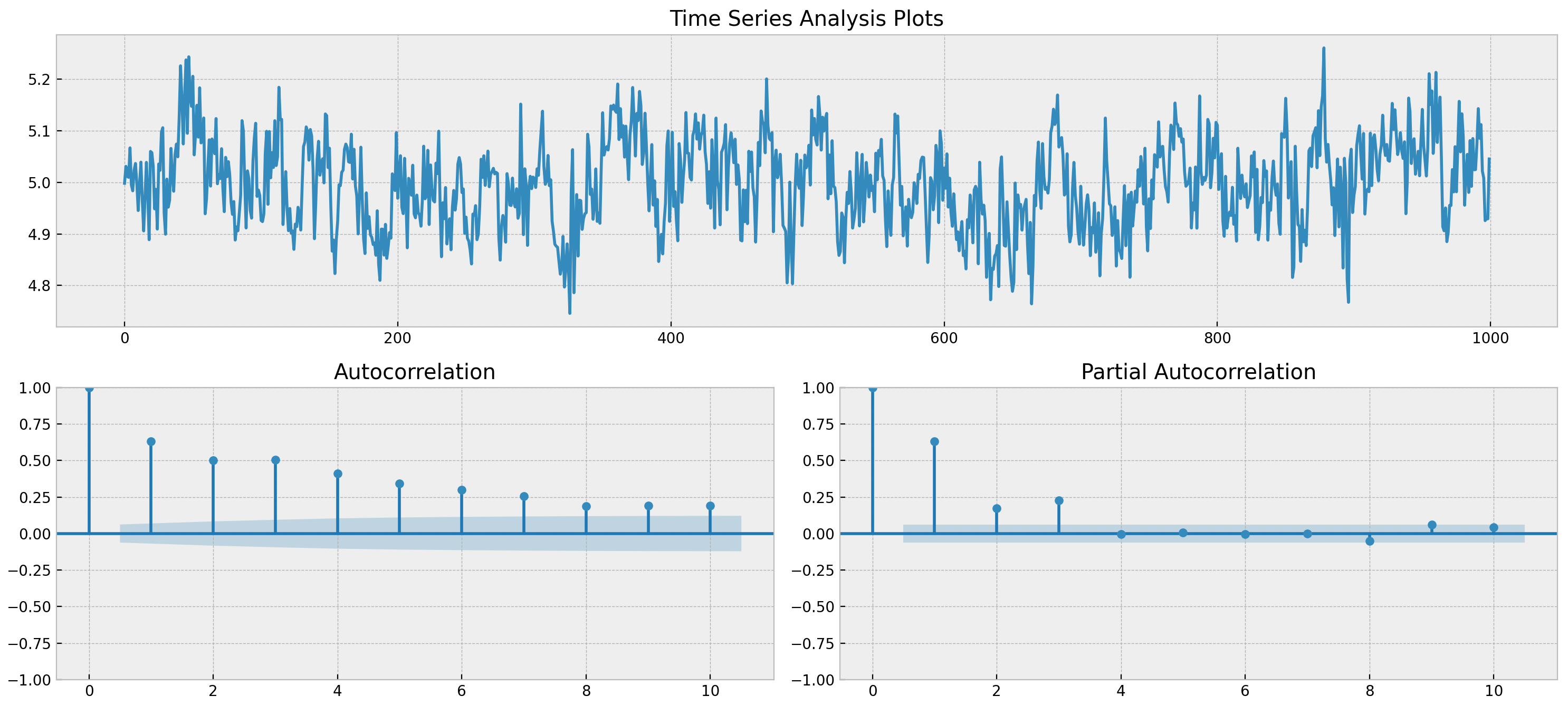
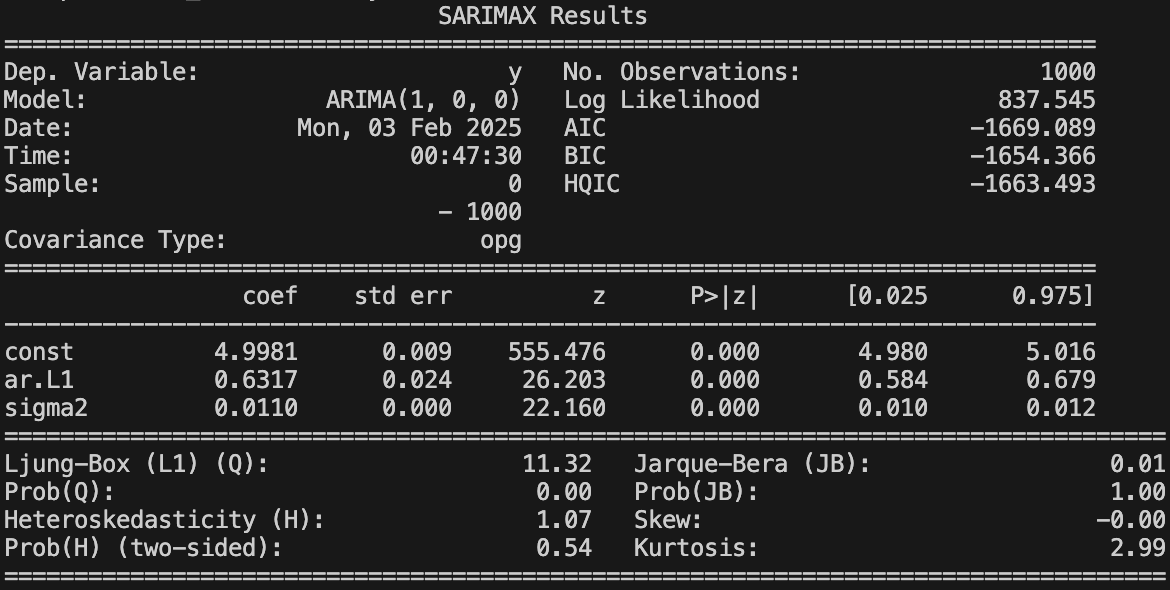
MA(3):



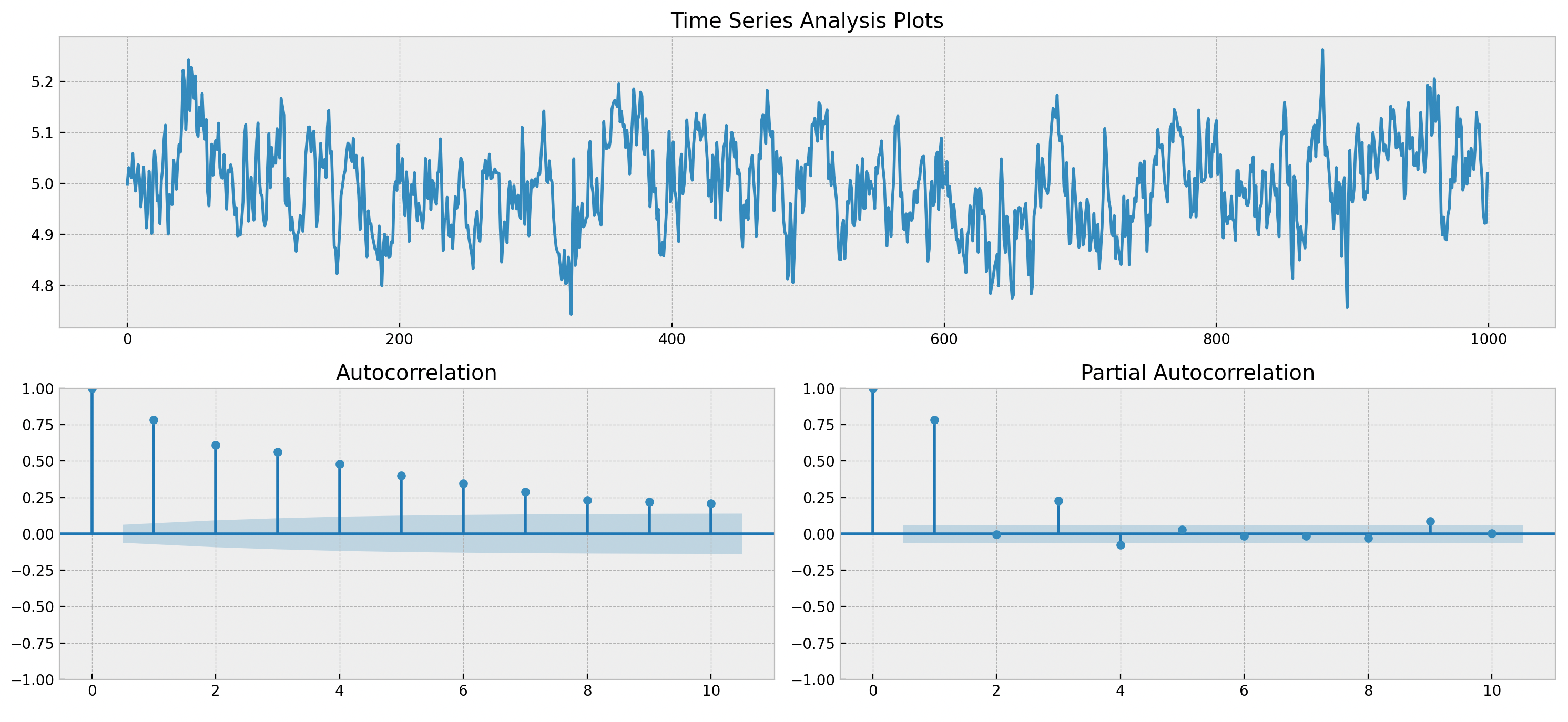
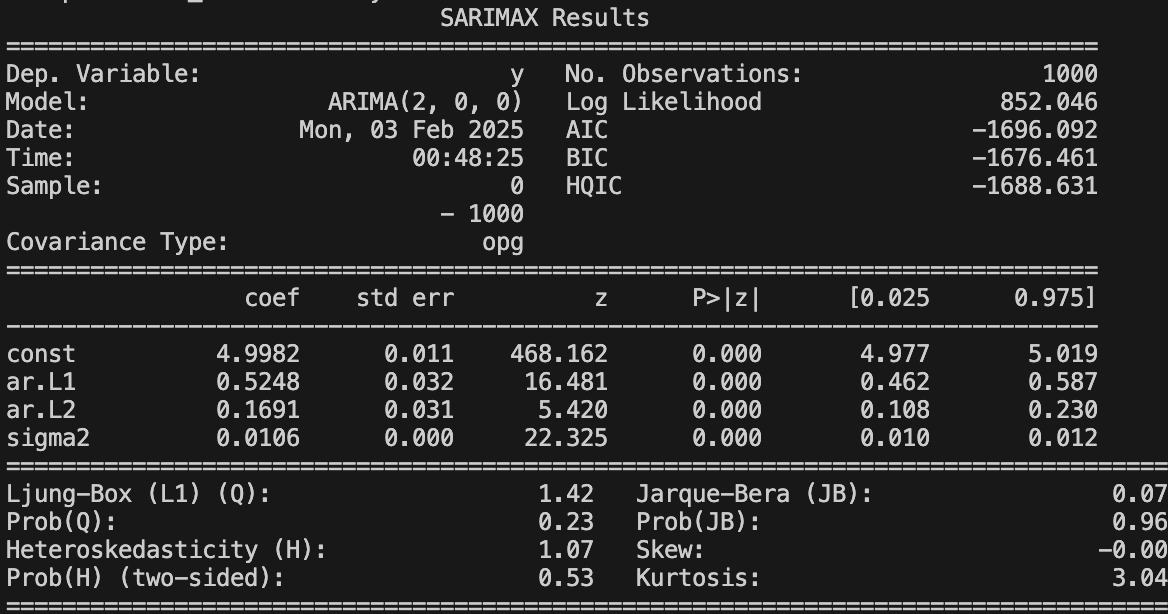
According to the information criteria, MA(3) model’s AIC, BIC is lower than others, which means that higher order MA processes do capture more time series features, and adding lag orders to the model can improve the model's goodness of fit. The PACF plot for MA(1) shows a first-order MA process, but still has the effect of higher-order lag terms.MA(2) and MA(3) gradually reduce the higher-order partial autocorrelation, indicating a more adequate model fit.

B.

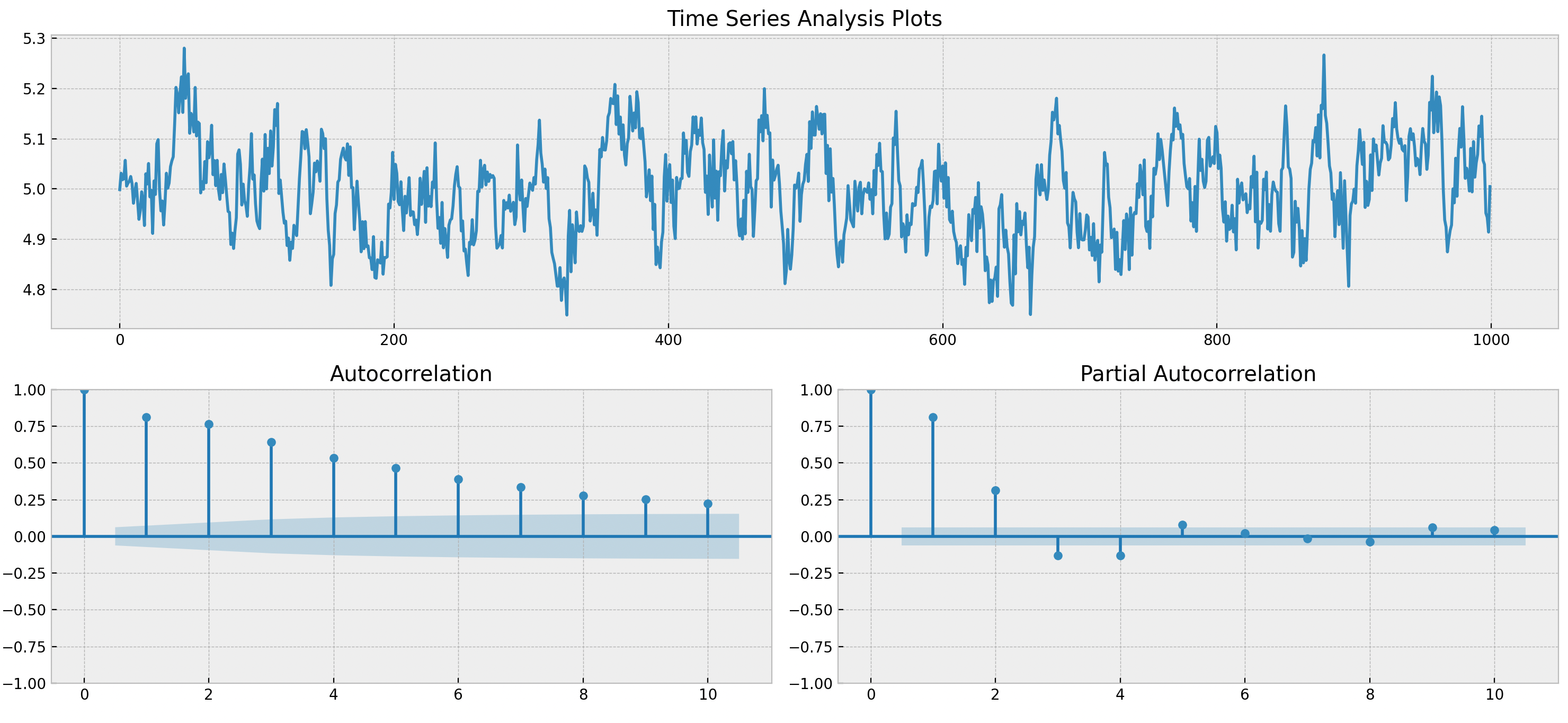
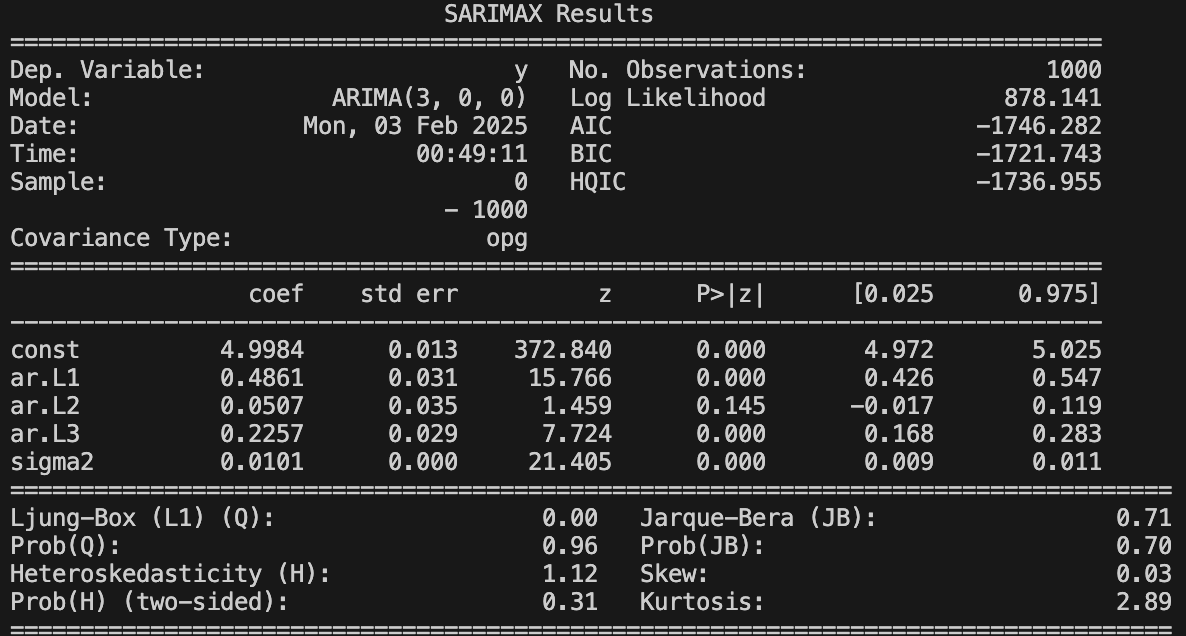
AR(1):



AR(2):



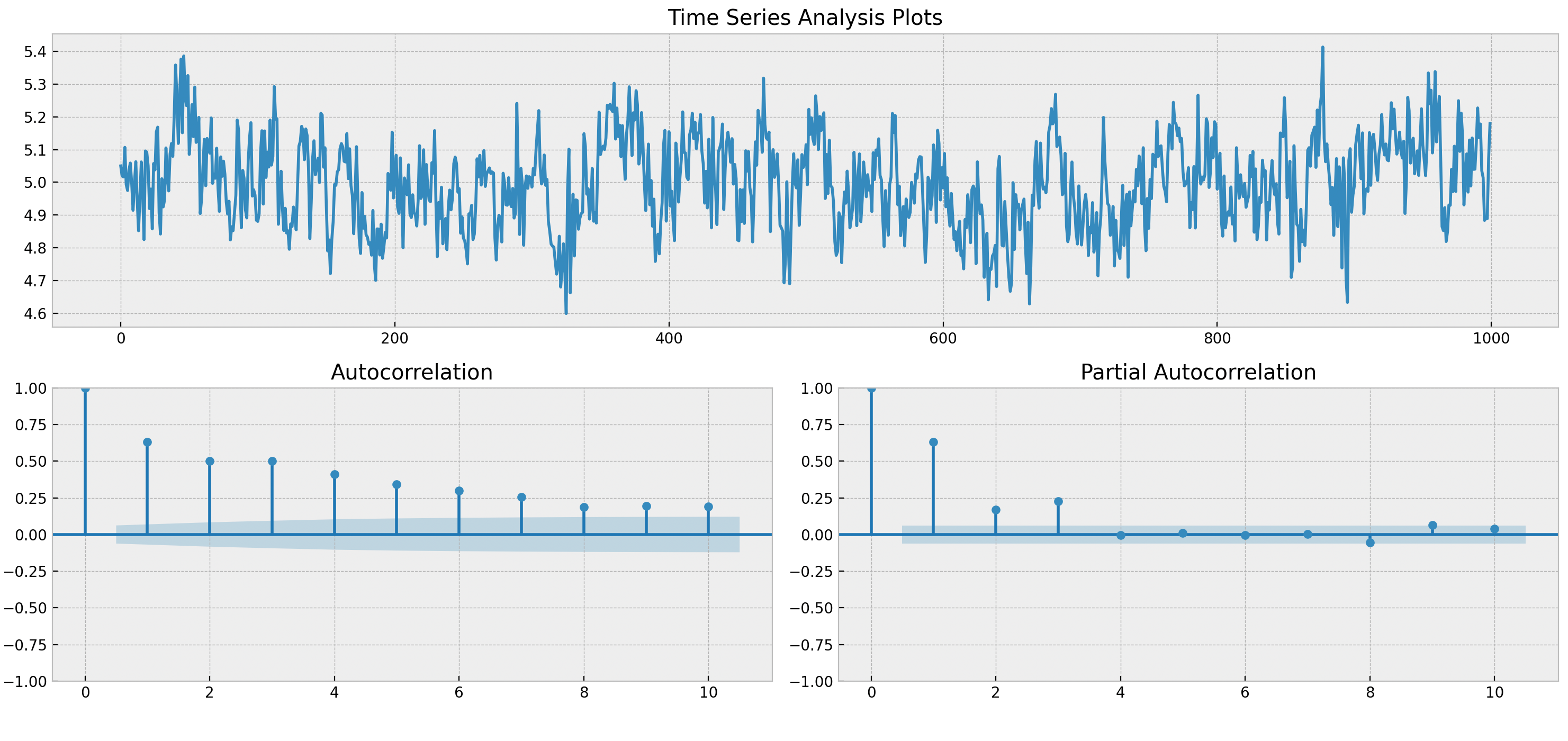
AR(3):



The gradual decrease in AIC/BIC with increasing AR order indicates that higher order AR processes improve the model fit. The PACF of AR(1) shows only one order of significant autocorrelation, suggesting that AR(1) may not be sufficient to fully explain the data structure.

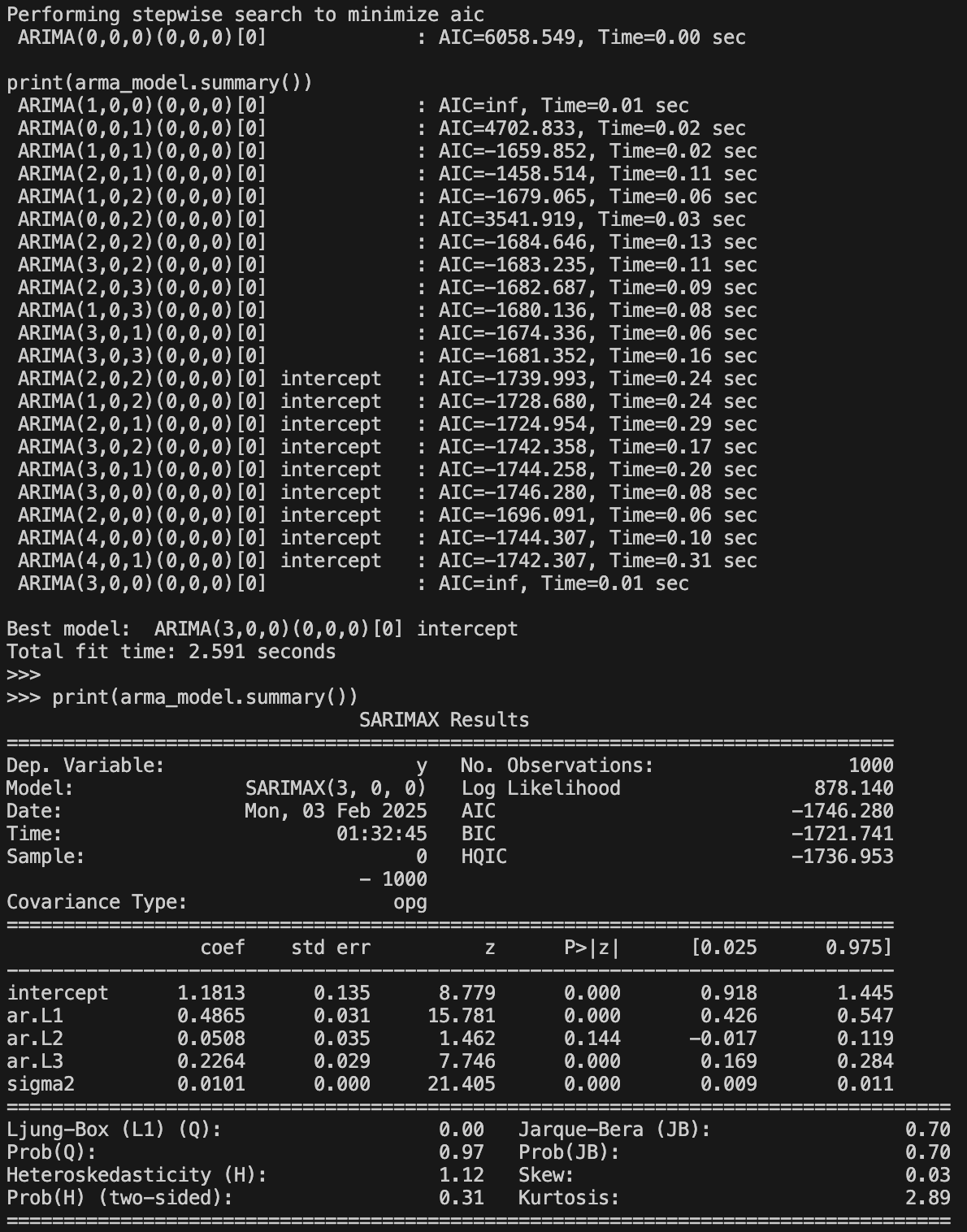
However, although the PACF of AR(3) drags its tail, suggesting that it may be capturing more lagged effects, it’s lag(2) p-value is larger than 0.05, indicates there maybe over-fitting in AR(3).

C.



From the figure, we can see that the ACF trails off, indicating that there is an AR process in the data, and that the PACF truncates quickly after lag 3, indicating that the data can be modeled using the AR(3) model.

D.



According to the AIC and auto selected arima method, AR(3) model’s AIC is the lowest

(-1746.28), so it’s the best fit of data.

#Problem5

B.

λ=0.99, PCA Explain variance:

[7.18732213e-01 7.56414069e-02 4.82033679e-02 2.87336584e-02

1.98043013e-02 1.40990410e-02 1.08218672e-02 1.04312498e-02…]

λ=0.88, PCA Explain variance:

[6.43630637e-01 1.71258019e-01 8.52365836e-02 3.11026188e-02

1.53339032e-02 1.18582545e-02 1.02036335e-02 7.58796561e-03…]

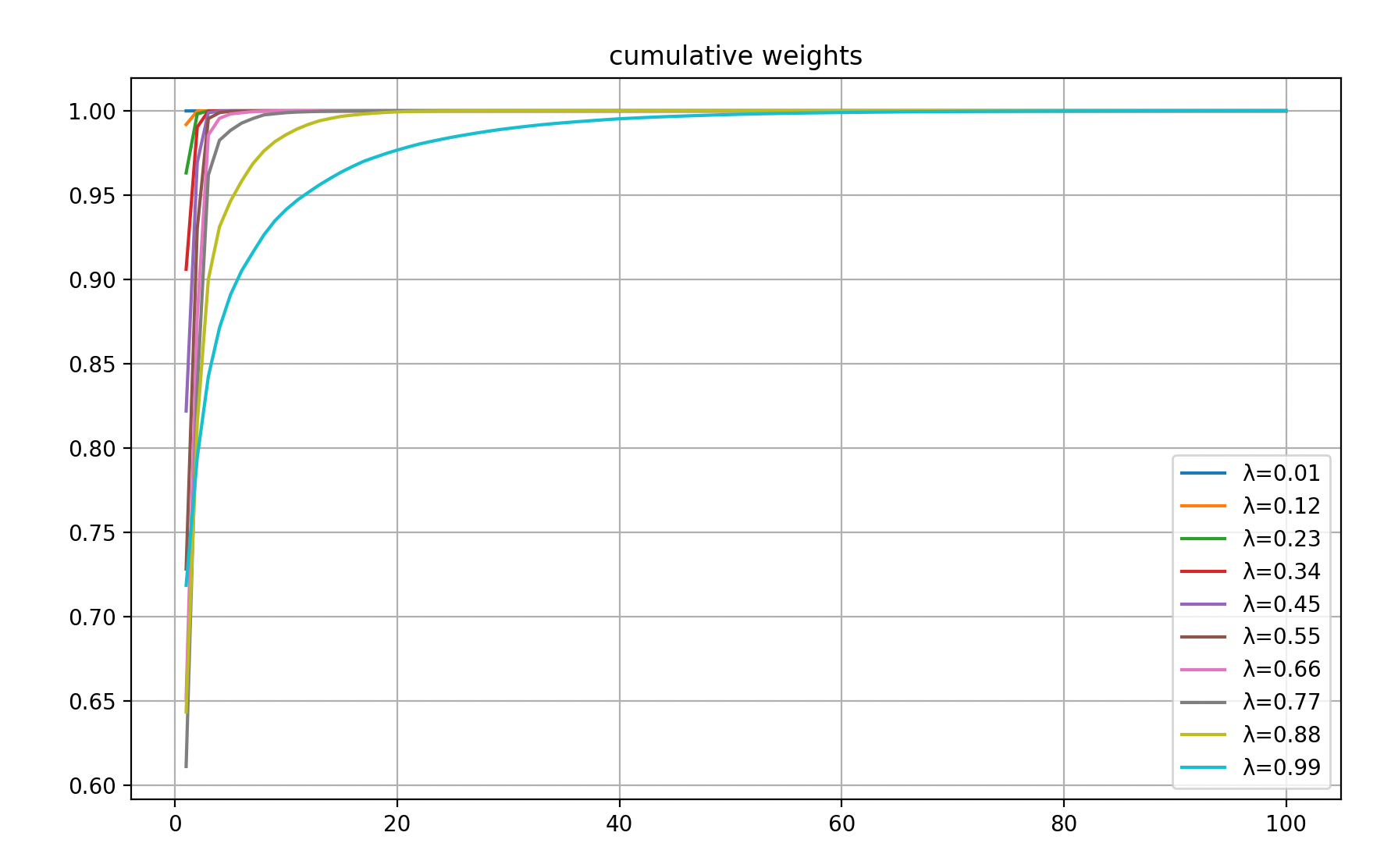
λ=0.77, PCA Explain variance:

[6.11298766e-01 2.24454381e-01 1.26163812e-01 2.06338415e-02

5.90573193e-03 4.28347589e-03 2.68240018e-03 2.17058048e-03…]

…

C.



When lambda is small, the cumulative weight function rises quickly, indicating that the exponentially weighted covariance matrix is more sensitive to recent changes. when lambda is large, the cumulative weight rises more slowly, indicating that the calculated covariance matrix is smoother, less susceptible to short-term fluctuations, and more reflective of long-term trends.

#Problem6

C.

Frobenius norm (Cholesky method): 0.2353

Frobenius norm (PCA method): 0.0212

D.

Cumulative variance explained (original): [0.02761939 0.05418435 0.07962796 0.10481583 0.12919237]

Cumulative variance explained (Cholesky): [0.00529703 0.01056292 0.01579098 0.02096686 0.02608529]

Cumulative variance explained (PCA): [0.02822073 0.05542919 0.08136647 0.1065685 0.13110493]

E.

Cholesky imulation time: 6.1680 seconds

PCA simulation time: 0.2789 seconds

F.

The speed of PCA simulation is faster, and it’s frobenius norm value is much lower.