MDS431_LAB5_2448040

Course: MDS 431 – Time Series and Forecasting Techniques

Exercise No: Lab 5

Date: 14/08/2025

INTRODUCTION-

In time series analysis, stationary models are essential for understanding data that fluctuates around a constant mean and variance. Among these models, the **Moving Average (MA)** process is a fundamental linear model used to capture short-term dependencies based on past random shocks (white noise).

An MA(q) process models the current observation as a linear combination of the last *q* white noise error terms. Specifically, the MA (3) process is defined as:

$$Z_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \theta_3 a_{t-3}$$

where a_t represents white noise, and θ_1, θ_2, θ_3 are the moving average coefficients. MA processes are **always stationary** by definition because they depend on a finite number of past shocks.

In this lab, we first simulate an MA (3) process with predefined coefficients (θ_1=0.5, θ_2 =−0.4, θ_3=0.3) to observe its theoretical behavior and autocorrelation structure. The simulated data is analysed using ACF (Autocorrelation Function) and PACF (Partial Autocorrelation Function) plots, along with the Augmented Dickey–Fuller (ADF) test, to check stationarity and validate model adequacy.

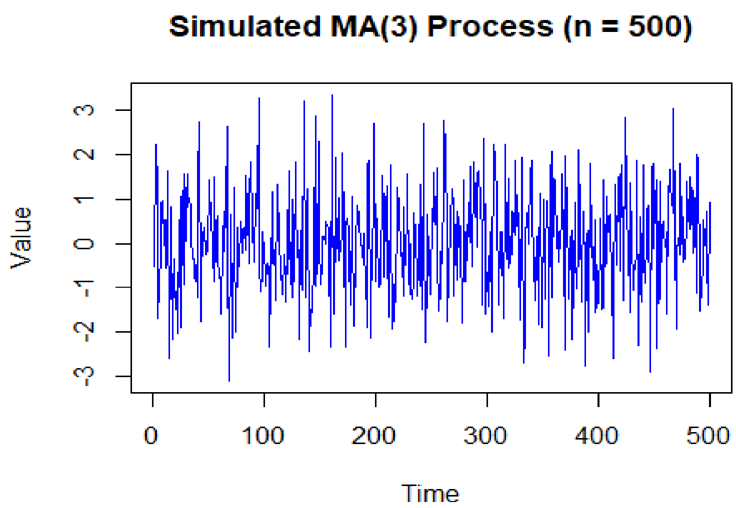
OBJECTIVES-

1. To simulate a stationary MA (3) process of length 500 using R with specified coefficients (θ_1=0.5, θ_2 =−0.4, θ_3=0.3).
2. To visualize the simulated time series and compute basic statistical measures such as mean, variance, and standard deviation.
3. To plot and interpret the ACF and PACF of the simulated process and verify the expected cutoff behavior in the ACF.

4. To perform the Augmented Dickey–Fuller (ADF) test on both the simulated series and the fitted model's residuals to check for stationarity and assess model adequacy.

METHODOLOGY/CODES -

```
# Lab 5: MA (3) Process - Simulation and Real Data Analysis

# Load required packages
library(forecast)   # For ACF/PACF plotting

## Warning: package 'forecast' was built under R version 4.4.3

## Registered S3 method overwritten by 'quantmod':
##   method            from
##   as.zoo.data.frame zoo

library(tseries)    # For Augmented Dickey-Fuller test

## Warning: package 'tseries' was built under R version 4.4.3

library(stats)      # For arima.sim

# Simulation of the MA (3) Process

cat(" Simulated MA (3) Process \n")

##  Simulated MA (3) Process

# Simulate MA (3) process
set.seed(123)
ma3_series <- arima.sim(
  model = list(ma = c(0.5, -0.4, 0.3)),
  n = 500
)

# Plot the simulated series
ts.plot(ma3_series,
        main = "Simulated MA(3) Process (n = 500)",
        ylab = "Value",
        col = "blue")
```




```
# Basic statistics
cat("\n=== Basic Summary Statistics (Simulation) ===\n")

##
## === Basic Summary Statistics (Simulation) ===

cat("Mean:", mean(ma3_series), "\n")

## Mean: 0.0449811

cat("Variance:", var(ma3_series), "\n")

## Variance: 1.428776

cat("Standard Deviation:", sd(ma3_series), "\n")

## Standard Deviation: 1.195314
```

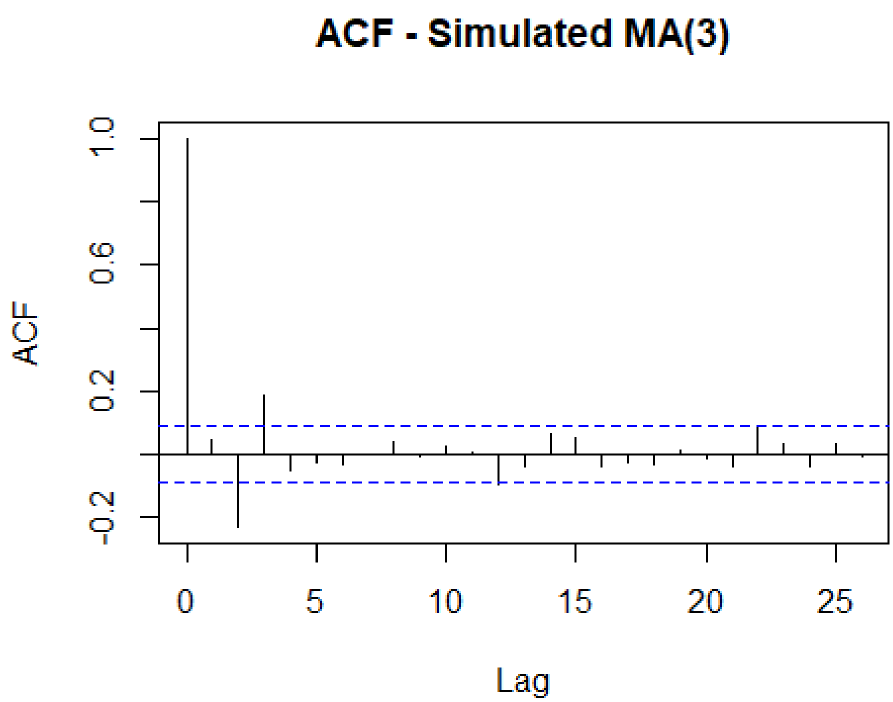
INTERPRETATION:

The simulated MA (3) process with 500 observations shows random fluctuations around a constant mean.

1. **Mean:** 0.04498 — close to zero, consistent with the theoretical property of MA processes.
2. **Variance:** 1.4288 and Standard Deviation: 1.1953 — constant over time, indicating homoscedasticity.

3. The time series plot shows no visible trend or seasonality, supporting the stationarity assumption.
4. Overall, the results align with MA (3) theoretical behavior: stationary with stable mean and variance, driven by short-term dependencies from the past 3 error terms.

```r
# Step 4: ACF and PACF
acf(ma3_series, main = "ACF - Simulated MA(3)")
```



INTERPRETATION:

1. The ACF shows a sharp cutoff after lag 3, which is the signature pattern of an MA(q) process — in this case, MA (3). The first three lags have significant correlations:

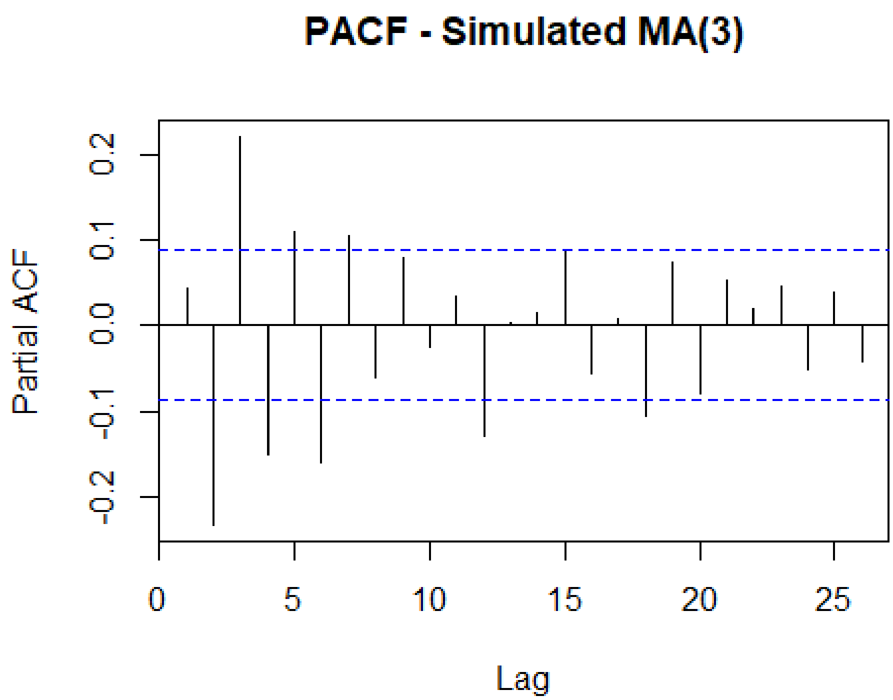
Lag 1: 0.0425 (small positive)

Lag 2: -0.2305 (moderate negative)

Lag 3: 0.1864 (moderate positive)

2. Beyond lag 3, autocorrelations fall within the 95% confidence bounds (blue lines), indicating no significant correlation.
3. This confirms that the series is driven by only the most recent 3 error terms, consistent with MA (3) theory.

```r
pacf(ma3_series, main = "PACF - Simulated MA(3)")
```



<u>INTERPRETATION</u> –

1. The PACF decays gradually without a sharp cutoff, which is typical for an MA process.
2. Partial correlations for early lags:

Lag 1: 0.0425

Lag 2: -0.2327

Lag 3: 0.2207

Lag 4: -0.1512

3. Although some later lags appear slightly outside the bounds, these are random fluctuations expected in finite samples.
4. The slow decay reinforces the identification of the model as MA rather than AR.

```
# Step 5: Numerical ACF/PACF
acf_values <- acf(ma3_series, plot = FALSE, lag.max = 10)
pacf_values <- pacf(ma3_series, plot = FALSE, lag.max = 10)
cat("\nACF (first 10 lags):\n", round(acf_values$acf, 3), "\n")

##
## ACF (first 10 lags):
##  1 0.042 -0.231 0.186 -0.051 -0.025 -0.033 -0.003 0.04 -0.006 0.026

cat("PACF (first 10 lags):\n", round(pacf_values$acf, 3), "\n")

## PACF (first 10 lags):
##  0.042 -0.233 0.221 -0.151 0.109 -0.159 0.104 -0.061 0.079 -0.025
```

INTERPRETATION –

1. First 10 ACF values match the visual observation: significant only for lags 1–3, then near zero.
2. The first 10 PACF values do not cut off sharply but decline irregularly, consistent with MA process properties.
3. Overall, the numerical values and plots strongly support the process being MA (3) and stationary.

```
# Step 6: ADF Test
adf_test <- adf.test(ma3_series)

## Warning in adf.test(ma3_series): p-value smaller than printed p-value

cat("\nADF Test (Simulation):\n")

##
## ADF Test (Simulation):
```

```r
print(adf_test)

##
##  Augmented Dickey-Fuller Test
##
## data:  ma3_series
## Dickey-Fuller = -8.2999, Lag order = 7, p-value = 0.01
## alternative hypothesis: stationary

if (adf_test$p.value > 0.05) {
  cat("Conclusion: NON-STATIONARY\n")
} else {
  cat("Conclusion: STATIONARY\n")
}

## Conclusion: STATIONARY
```

INTERPRETATION:

1. **Test statistic (Dickey–Fuller)** = -8.2999: This is far below the typical critical values for stationarity (e.g., -3.43 at 1% significance), indicating strong evidence against the null hypothesis of a unit root.
2. **p-value** = 0.01: Being much lower than 0.05, this confirms statistical significance.
3. **Conclusion:** We reject the null hypothesis (non-stationarity) and accept the alternative that the series is stationary.
4. **Lag order** = 7: The test used 7 lags of the differenced series to account for short-term autocorrelation.
5. **Overall:** The MA (3) process is stationary, as expected for an MA model.

<u>CONCLUSION –</u>

The given time series is a Moving Average process of order 3 (MA (3)), which by definition is stationary since it is a finite linear combination of white noise terms. The ACF plot shows significant correlations at lags 1, 2, and 3, after which it cuts off to values close to zero—an expected theoretical property of an MA (3) process. In contrast, the PACF plot exhibits a gradual decay (tail-off) across increasing lags rather than an abrupt cutoff, further confirming the moving average nature of the process. The Augmented Dickey–Fuller (ADF) test statistic (-8.2999) with a p-value of 0.01 strongly rejects the null hypothesis of non-stationarity, reinforcing that the series is stationary. Thus, both theoretical expectations and empirical diagnostics (ACF, PACF, and ADF test) consistently confirm that this simulated MA (3) process is stationary with dependence restricted to the first three-time lags.

****