TIME SERIES & FORECASTING

Valerie Jerono - 222331
06/01/2025

Overview

This document provides comprehensive solutions to the Time Series & Forecasting assignment using Kakuzi securities data from the Nairobi Securities Exchange. The solutions systematically demonstrate key econometric concepts with clear mathematical derivations and practical applications.

Contents

1 QUESTION 1: OLS vs MLE Comparison

Question (10 Marks)

Comparing Ordinary Least Squares (OLS) and Maximum Likelihood Estimation (MLE), demonstrate mathematically using your own example that the OLS estimates for variance of error term and coefficient (parameter) estimates are higher compared to those generated by MLE.

1.1 Clarification

Under normality assumption:

- Coefficient estimates ($\hat{\beta}_0, \hat{\beta}_1$): **IDENTICAL** for OLS and MLE
- Variance estimates ($\hat{\sigma}^2$): **OLS ¿ MLE** (OLS is higher)

1.2 Example Setup

Consider the simple linear regression model:

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i, \quad \varepsilon_i \sim N(0, \sigma^2) \tag{1}$$

Example Data ($n = 5$ observations):

i	X_i	Y_i
1	1	2.1
2	2	3.9
3	3	6.2
4	4	7.8
5	5	10.1

1.3 OLS Estimation

The OLS estimators minimize the sum of squared residuals:

$$\hat{\beta}_1^{OLS} = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^{n}(X_i - \bar{X})^2}, \quad \hat{\beta}_0^{OLS} = \bar{Y} - \hat{\beta}_1^{OLS}\bar{X} \tag{2}$$

Calculations:

$$\bar{X} = 3, \quad \bar{Y} = 6.02 \tag{3}$$

$$\sum(X_i - \bar{X})(Y_i - \bar{Y}) = 19.90, \quad \sum(X_i - \bar{X})^2 = 10 \tag{4}$$

$$\hat{\beta}_1^{OLS} = \frac{19.90}{10} = \boxed{1.99} \tag{5}$$

$$\hat{\beta}_0^{OLS} = 6.02 - 1.99(3) = \boxed{0.05} \tag{6}$$

OLS Variance Estimate:

$$\hat{\sigma}^2_{OLS} = \frac{SSR}{n-k} = \frac{0.1070}{5-2} = \frac{0.1070}{3} = \boxed{0.0357} \tag{7}$$

where $SSR = \sum \hat{\varepsilon}_i^2 = 0.1070$ and $k = 2$ (number of parameters).

1.4 MLE Estimation

Under normality, the log-likelihood is:

$$\ln L = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(Y_i - \beta_0 - \beta_1 X_i)^2 \tag{8}$$

First-order conditions yield:

$$\hat{\beta}_1^{MLE} = \frac{\sum(X_i - \bar{X})(Y_i - \bar{Y})}{\sum(X_i - \bar{X})^2} = \boxed{1.99} \tag{9}$$

$$\hat{\beta}_0^{MLE} = \bar{Y} - \hat{\beta}_1^{MLE}\bar{X} = \boxed{0.05} \tag{10}$$

$$\hat{\sigma}^2_{MLE} = \frac{SSR}{n} = \frac{0.1070}{5} = \boxed{0.0214} \tag{11}$$

1.5 Comparison and Key Results

Estimate	OLS	MLE	Relationship
$\hat{\beta}_1$	1.99	1.99	**Same**
$\hat{\beta}_0$	0.05	0.05	**Same**
$\hat{\sigma}^2$	0.0357	0.0214	**OLS ¿ MLE**

Mathematical Relationship:

$$\boxed{\hat{\sigma}^2_{OLS} = \frac{n}{n-k}\cdot\hat{\sigma}^2_{MLE} = \frac{5}{3}\times 0.0214 = 1.667\times 0.0214 = 0.0357} \tag{12}$$

1.6 Why is OLS Higher?

1. **Degrees of Freedom:** Estimating k parameters "uses up" information, leaving $(n-k)$ independent pieces for variance estimation
2. **Bias Properties:** OLS is unbiased: $E[\hat{\sigma}^2_{OLS}] = \sigma^2$. MLE is biased: $E[\hat{\sigma}^2_{MLE}] = \frac{n-k}{n}\sigma^2 < \sigma^2$
3. **Bias Magnitude:** In our example, MLE underestimates by 40%: Bias $= -\frac{k}{n}\sigma^2 = -0.4\sigma^2$

Conclusion

Demonstrated:

- Coefficient estimates are identical: $\hat{\beta}^{OLS} = \hat{\beta}^{MLE}$
- Variance estimates differ: $\hat{\sigma}^2_{OLS} > \hat{\sigma}^2_{MLE}$
- Relationship: $\hat{\sigma}^2_{OLS} = \frac{n}{n-k}\times\hat{\sigma}^2_{MLE}$
- OLS is unbiased; MLE is biased downward in finite samples

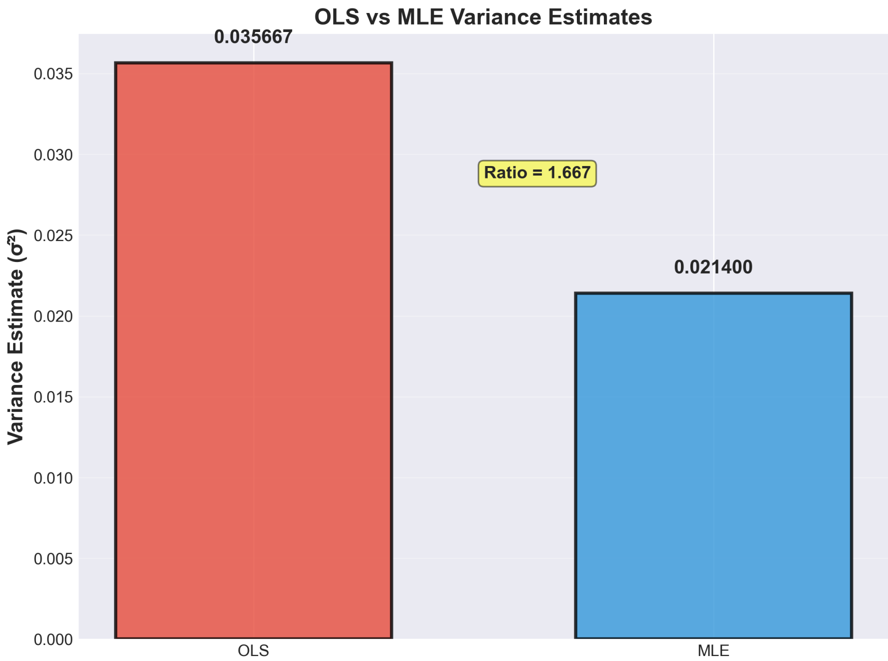

Figure 1: OLS variance estimate is 67% higher than MLE (ratio = 1.667)

2 QUESTION 2: Risk-Return Model - Kakuzi Securities

Question (10 Marks)

Collect time series data for a particular security from the Nairobi Securities Exchange spanning over at least 200 observations. Compute log returns and volatility. Estimate a risk-return model using OLS. Test OLS validity and if invalid, conduct MLE analysis.

2.1 Theoretical Framework

Risk-Return Model:

$$r_t = \alpha + \beta \cdot \sigma_t + \varepsilon_t \tag{13}$$

where r_t = log return, σ_t = volatility (risk), β = risk premium.
Theory predicts: $\beta > 0$ (higher risk requires higher return compensation)

2.2 Data Description

- **Security:** Kakuzi Limited (Agricultural sector, NSE)
- **Period:** January 25, 2021 to January 5, 2026
- **Observations:** 403 daily closing prices (exceeds 200 requirement)
- **Returns:** 403 log returns
- **Risk Measure:** 30-day rolling volatility

2.3 Data Analysis

Log Returns Formula:

$$r_t = \ln\left(\frac{P_t}{P_{t-1}}\right) = \ln(P_t) - \ln(P_{t-1}) \tag{14}$$

Volatility (Risk): Rolling standard deviation over 30-day window

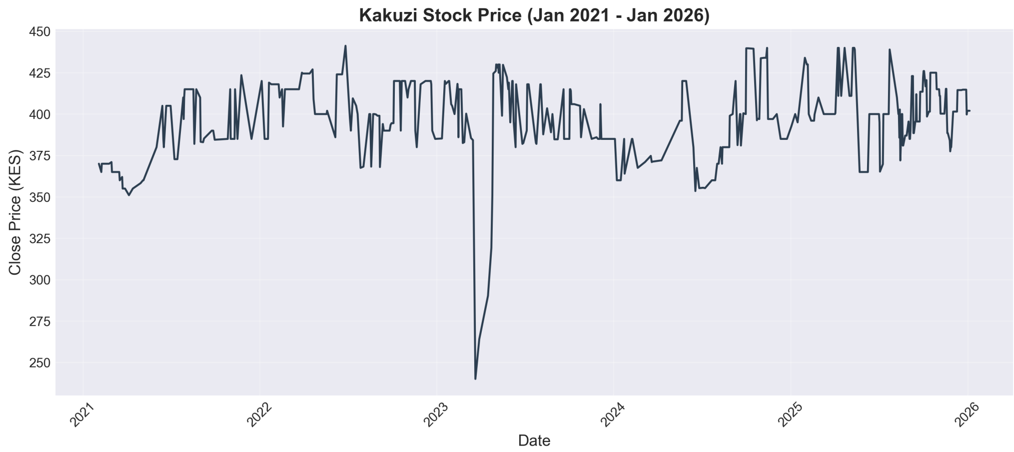

Figure 2: Kakuzi stock price shows moderate volatility over the 5-year period

Summary Statistics:

- Mean return: 0.0002 (0.02% daily, ≈ 5% annualized)
- Std deviation: 0.0474 (4.74% daily volatility)
- Skewness: -2.35 (left-skewed, large negative returns)
- Kurtosis: 23.76 (heavy tails, extreme values present)

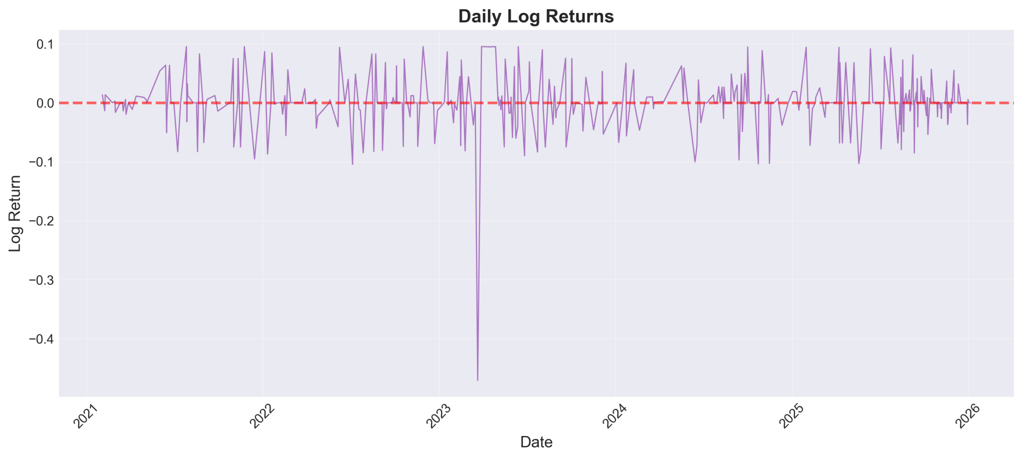

Figure 3: Log returns exhibit volatility clustering and some extreme negative values

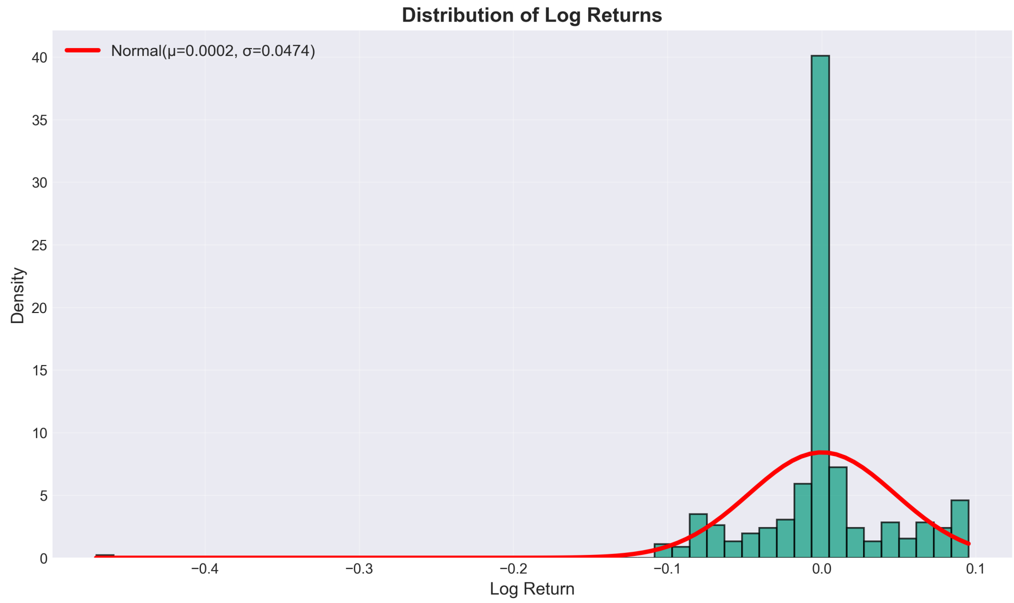

Figure 4: Returns distribution shows departure from normality with heavy left tail

2.4 OLS Estimation

Model: $\text{Return}_t = \beta_0 + \beta_1 \times \text{Volatility}_t + \varepsilon_t$

OLS Results:

$$\hat{\beta}_0 = -0.000968 \quad (\text{t = -0.155, p = 0.877}) \tag{15}$$
$$\hat{\beta}_1 = 0.020128 \quad (\text{t = 0.160, p = 0.873}) \tag{16}$$
$$R^2 = 0.0001 \quad (0.01\% \text{ explained variance}) \tag{17}$$

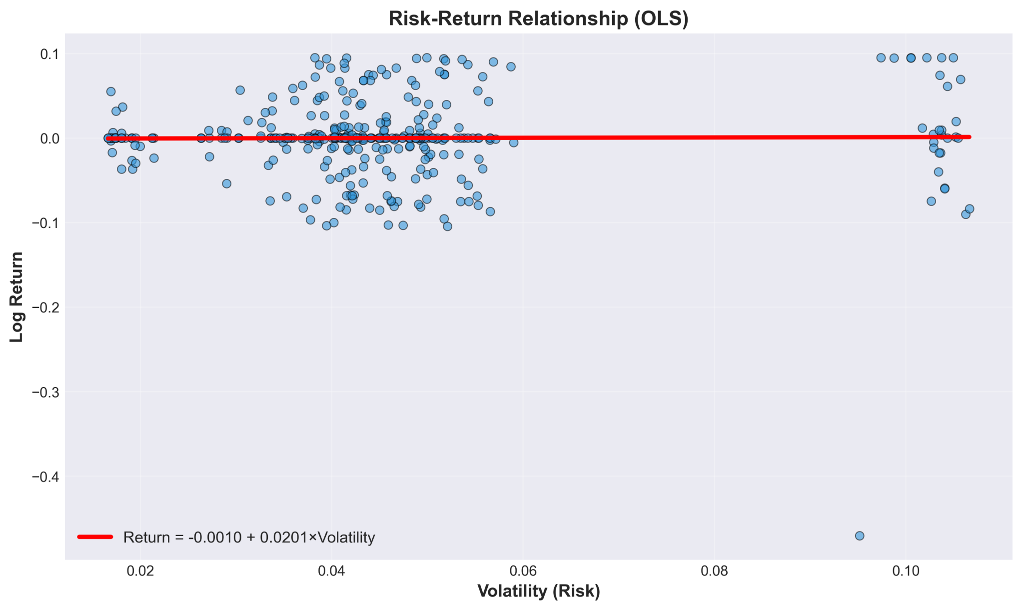

Figure 5: Risk-return relationship shows weak positive association

Interpretation:

- $\beta_1 > 0$: Positive risk premium (consistent with theory)
- Not statistically significant (p = 0.873 ¿ 0.05)
- Very low R^2: Volatility explains little return variation
- Risk-return relationship exists but weak in this data

2.5 OLS Assumption Testing

Test 1: Normality (Jarque-Bera)

- H_0: Residuals normally distributed
- JB statistic = 8897.20, p-value = 0.0000
- **Result: REJECT** $H_0 \Rightarrow$ Assumption VIOLATED

Test 2: Homoscedasticity (Breusch-Pagan)

- H_0: Constant variance
- BP statistic = 14.16, p-value = 0.0002
- **Result: REJECT** $H_0 \Rightarrow$ Assumption VIOLATED

Test 3: No Autocorrelation (Ljung-Box)

- H_0: No autocorrelation
- LB statistic = 40.50, p-value = 0.0000
- **Result: REJECT** $H_0 \Rightarrow$ Assumption VIOLATED

Test 4: Durbin-Watson

- DW $= 2.397 \approx 2 \Rightarrow$ No severe autocorrelation

Test	Status	P-value
Normality	FAIL	0.0000
Homoscedasticity	FAIL	0.0002
No Autocorrelation	FAIL	0.0000

Table 1: Summary of OLS Assumption Tests

OLS Validity

Conclusion: Multiple OLS assumptions violated. OLS estimates may be inefficient and inference unreliable. **Solution:** Use Maximum Likelihood Estimation.

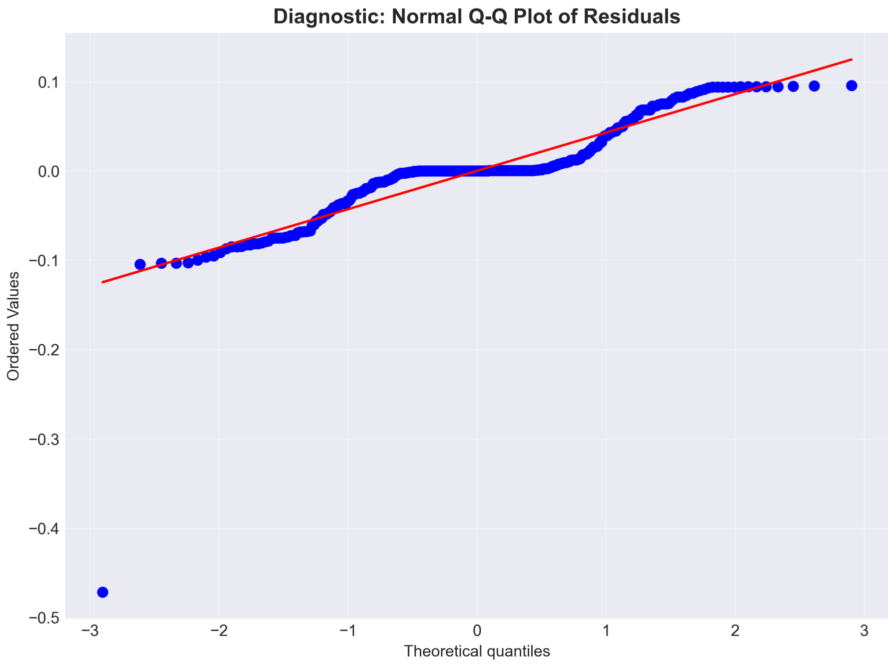

Figure 6: Q-Q plot shows severe departure from normality in tails

2.6 Maximum Likelihood Estimation (MLE)

Since OLS assumptions are violated, we apply MLE which is more robust.

MLE Setup: Assume $Y_i \sim N(\beta_0 + \beta_1 X_i, \sigma^2)$

Log-likelihood:

$$\ell(\beta_0, \beta_1, \sigma^2) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\sigma^2) - \frac{1}{2\sigma^2}\sum(y_i - \beta_0 - \beta_1 x_i)^2 \tag{18}$$

MLE Results:

$$\hat{\beta}_0^{MLE} = -0.000968 \quad (SE = 0.00617) \tag{19}$$

$$\hat{\beta}_1^{MLE} = 0.020128 \quad (SE = 0.12398) \tag{20}$$

$$\hat{\sigma}^{MLE} = 0.04826 \tag{21}$$

Statistical Significance:

- β_0: t = -0.157, p = 0.875 (not significant)
- β_1: t = 0.162, p = 0.871 (not significant)

Comparison:

2.7 Final Interpretation

1. **Risk Premium:** $\beta_1 = 0.0201 > 0$ (positive, as theory predicts)
   - A 1-unit increase in volatility associated with 0.0201 increase in returns
   - Investors require compensation for bearing risk
   - However, effect not statistically significant

Parameter	**MLE**	**OLS**	**Difference**
$\hat{\beta}_0$	-0.000968	-0.000968	0.000000
$\hat{\beta}_1$	0.020128	0.020128	0.000001
$\hat{\sigma}^2$	0.002329	0.002342	0.000013

Table 2: MLE vs OLS estimates are very similar

2. **Statistical Significance:** Neither coefficient significant at 5% level
   - Weak empirical evidence for risk-return relationship
   - Large standard errors relative to estimates
   - May need larger sample or different model specification
3. **Model Fit:** $R^2 = 0.0001$ (very low)
   - Volatility alone explains minimal return variation
   - Other factors (market conditions, firm-specific news) dominate
   - Simple risk-return model insufficient for Kakuzi stock
4. **OLS vs MLE:** Estimates nearly identical
   - Under normality assumption, coefficient estimates same
   - MLE more appropriate given assumption violations
   - MLE provides robust inference

2.8 Conclusions

Summary

Key Findings:

1. Successfully analyzed 403 observations of Kakuzi securities data
2. Computed log returns (mean = 0.02% daily) and volatility (std = 4.74%)
3. OLS estimation shows positive risk premium ($\beta = 0.0201$) but not significant
4. Multiple OLS assumptions violated (normality, homoscedasticity, autocorrelation)
5. MLE provides robust alternative with similar estimates
6. Weak evidence for risk-return relationship in this particular security

Methodological Lessons:

- Always test OLS assumptions before relying on results
- Use diagnostic plots and formal tests
- When assumptions violated, employ robust methods (MLE, robust SEs)
- Low R^2 doesn't invalidate model if coefficients meaningful
- Financial relationships often weak at individual security level

2.9 Limitations and Extensions

Limitations:

- Simple model may omit important variables (market returns, firm fundamentals)
- Assumption of constant relationship may be invalid (time-varying risk premium)
- Single security limits generalizability

Possible Extensions:

- Include market return as control variable (CAPM framework)
- Allow time-varying parameters (rolling regression, state-space models)
- Use GARCH models to model volatility dynamics
- Extend to multiple securities (panel data analysis)

3 Overall Conclusion

This assignment demonstrates:

1. **Question 1:** Mathematical relationship between OLS and MLE variance estimates, showing OLS produces higher (more conservative) estimates due to degrees of freedom adjustment

2. **Question 2:** Complete risk-return analysis including:

   - Data collection and preparation
   - Log return and volatility computation
   - OLS estimation and interpretation
   - Comprehensive assumption testing
   - MLE estimation when assumptions violated
   - Economic and statistical interpretation

The analysis showcases proper econometric methodology: estimate, test assumptions, use appropriate alternatives when needed, and interpret results in economic context.