Homework Assignment

Course: Time series

Due Date: ***December2, 2025***

Professor*: Gilles Dufrénot*

Instructions:

This assignment should be completed in groups of up to four students. It is designed to assess your understanding of the concepts studied in class.

In your final submission, **you must include the full names of everyone who participated in the work**.

Please note:

The composition of the groups must be provided to the professor **by 1 October 2025 at the latest**. After this deadline, groups will be assigned at random and there will be no possibility of appeal.

Deliverables

You are expected to produce a concise written report of 18-20 pages (max : 25 pages), which should include the following:

- Written answers to the questions.

- Tables and plots summarizing your empirical findings.
- A technical appendix with the R code you used for the analysis. The latter must be provided as a markdown document, or a pdf.

Exercise 1: Extreme distributions

- Download the Excel file CAPM on AMETICE
- Consider one of the following indexes
  Oracle; Microsoft; General Electric; Ford
  NB: USTB3M is the yield of the risk-free asset (3-month US government Bond Yield) and S&P500 is the market index.

- Do a preliminary analysis of the excess returns (your asset + market excess returns) based on basics statistics, graphs of the distributions, QQ-plots, box-plot, bivariate scatterplot, graphs of time series, autocorrelation function, pp-plot
- Fit one of the following extreme distribution to the excess returns (your asset + market excess returns): Weibull, log-normal, Gamma.
- Fit a GEV distribution to the excess return series and plot the QQ-plot, density and CDF.
- Using elliptical copulas, estimate the correlation between the excess returns of your asset asset and the market excess returns.
- Fit Archimedian copulas (Clayton or Weibull).

Exercise 2 : Regime-swifting models II

Context: as a portfolio manager, you would like to test some hypotheses about some capital asset pricing model (CAPM) beta for several US stocks (data are similar). Consider the following equation:

$$E(R_{it}) - R_{ft} = \alpha_i + \beta_i(L)[E(R_{mt}) - R_{ft}] + \varepsilon_{it}$$

L is the lag operator and ε_{it} is a white noise term.

$E(R_{it})$ is the expected return on the capital asset, the index i refers to the different assets of the S&P 500.

R_{ft} is the risk-free rate of interest rate (for instance the interest rate arising from government bonds). In the database, it is denoted USTB3M : 3-month US Treasury bond;

β_i is the sensitivity of the expected excess asset returns to the expected excess market returns;

$E(R_{mt})$ is the expected return on the market (here on the S&P 500).

Questions

1. Estimate the above equation using a Markov-switching model with 2 regimes:
- Estimate the linear model and use tests on residuals to show that the dynamics is probably nonlinear.
- Fit a Markov-switching model (select the optimal lag structure using tests based on information criteria).
- Plot graphs of the posterior probabilities of each regime.
- Comment your findings.

2. We wish to apply a STAR model to the excess return on assets, the transition variable being the market excess return delayed by one or more periods (to be determined). After applying the appropriate tests, estimate an ESTAR or LSTAR model. Make a graphical representation of the transition function. Comment on your results.

Exercise 3 : Regime-swifting models II

Using data from the base_bond-equity database (to be downloaded on AMETICE), we wish to study the relationships between the bond spreads of two countries (10-year yields). Choose one of the following country pairs:

1: France/Belgium; *2* : Austria/Spain; *3*: Greece/Ireland

4: Italy/the Netherlands

Consider the following equation:

$$(Y_{it}) - Y_{gt} = \alpha_i + \beta_{ij}(L)[(Y_{jt}) - Y_{gt}] + \varepsilon_{ijt}$$

Y_{gt}: German 10-year bond

$(Y_{it}) - Y_{gt}$: yield spread of country i

$(Y_{jt}) - Y_{gt}$: yield spread of country j

L is the lag operator and ε_{ijt} is a white noise term.

Questions

3. Estimate the above equation using a Markov-switching model with 2 regimes:
- Estimate the linear model and use tests on residuals to show that the dynamics is probably nonlinear.
- Fit a Markov-switching model (select the optimal lag structure using tests based on information criteria).
- Plot graphs of the posterior probabilities of each regime
- Comment your findings.
4. We wish to apply a STAR model to the excess return on assets, the transition variable being the market excess return delayed by one or more periods (to be determined). After applying the appropriate tests, estimate an ESTAR or LSTAR model. Make a graphical representation of the transition function. Comment on your results.
5. Use a bivariate GARCH model to study the correlation between excess return of your assets and excess return on the market. Comment on your results.

Exercise 4 : Extreme distributions

Now, we wish to model the correlation between the spreads of countries i and j using copulas.

- Do a preliminary analysis of the spreads based on basics statistics, graphs of the distributions, QQ-plots, box-plot, bivariate scatterplot, graphs of time series, autocorrelation function, pp-plot
- Fit one of the following extreme distribution to the spreads: Weibull, log-norma, Gamma.
- Fit a GEV distribution to the excess return series and plot the QQ-plot, density and CDF.
- Using elliptical copulas, estimate the correlation between the excess returns of your asset asset and the market excess returns.
- Fit Archimedian copulas (Clayton or Weibull).

Exercise 5: Long-memory models

A – Theoretical Foundations

1. Definition and Characteristics
   Begin by defining what is meant by a long-memory process. In your answer, explain the statistical features that distinguish long-memory processes from short-memory ones. You should highlight the behavior of the autocorrelation function (slow hyperbolic decay) and the spectral density (divergence near the origin). Clarify the role of the fractional

differencing parameter d, and explain why values of d between 0 and 0.5 indicate long memory.

2. Examples of Long-Memory Processes
   Present the formal definition of an ARFIMA(p,d,q) model, clearly specifying the autoregressive, moving-average, and fractional differencing operators. Then provide the formal definition of a Gegenbauer (GARMA or GARFIMA) process. Explain in words the essential difference between these two classes: while ARFIMA models capture persistence around the zero frequency, Gegenbauer processes allow for cyclical long memory with persistence around non-zero frequencies.

3. Estimation Methods
   ARFIMA models can be estimated using several approaches, including exact maximum likelihood estimation, the approximate Whittle estimator, and semiparametric methods such as the Geweke–Porter–Hudak (GPH) regression. Discuss which estimator you would prefer in practice and why. In your answer, weigh the trade-offs between efficiency, robustness, and computational complexity, as well as the sample size requirements of each method.

B – Empirical Analysis

The second part of the assignment is empirical. You will work with two datasets that can be downloaded from **AMETICE** (see the introductory section of the course):

- Monthly airline passengers.
- ENSO anomalies index (ONI).

Your task is to apply long-memory techniques to these series. Proceed as follows:

1. Exploratory Analysis
   Plot the two series and comment on their visual characteristics. Do you observe trends, seasonality, or volatility clustering? Compute and plot the autocorrelation and partial autocorrelation functions. Based on their behavior, comment on whether the series appear to exhibit long memory.

2. Model Estimation
   Estimate ARFIMA models for both series. Report the estimated fractional differencing parameter ddd, along with the autoregressive and moving-average coefficients. Provide the standard errors and confidence intervals. Interpret your results carefully: what do the estimates of ddd suggest about the degree of persistence in each series?

3. Model Diagnostics
   Perform misspecification tests on the residuals, including the Ljung–Box test for serial correlation, a test for normality, and a test for heteroskedasticity (e.g. ARCH effects). Comment on whether the residuals behave as white noise and whether the ARFIMA model is an adequate specification.

C – Comparative Analysis

In this section, focus exclusively on the airline passengers series.

1. Estimate an ARFIMA model on the series in levels.
2. Transform the series by taking log-differences (i.e. growth rates) and estimate an ARFIMA model on this transformed series.
3. Compare the results. How does differencing affect the estimate of the fractional parameter ddd? Which specification seems more appropriate for capturing long memory in the data? Discuss the trade-off between modeling the series in levels versus growth rates, especially in terms of interpretation and stationarity.

Good luck