Public Report - Module 1 (EN)

Introduction to the Project

According to the planning, the project follows the predecessor stages leading to the development of a recession prediction model. The stages that have been completed involve the process of investigation and determination of references addressed. In this sense, every methodological definition has a theoretical basis, which is divided into two sections: economic themes and statistical themes.

Introduction

Statistics, an essential tool for data analysis and interpretation, plays a crucial role in understanding economic phenomena, especially in predicting unpredictable events like crises and recessions, often associated with "black swans" (Taleb, 2007). The Monte Carlo Method, a mathematical technique developed by von Neumann and Ulam, stands out by simulating uncertain scenarios through stochastic modeling, applying from stock pricing to macroeconomic risk analysis. Studies such as those by Huub Meijers and Peter Hammond demonstrate its use in modeling economic cycles and assessing systemic risks, while more recent works explore combinations with techniques like neural networks and geometric Brownian motion. Despite its complexity and applications in sectors like space (e.g., NASA's JSTAR simulator), there are still gaps in its use for comprehensive macroeconomic predictions, especially in integrating multiple indicators and considering specific contextual factors for each country. The proposal presented here aims to fill this gap by associating the unpredictability of recessions with the simulation of multiple statistical scenarios, democratizing access to robust analyses and encouraging a critical view on the relationship between economic theory, isolated indicators, and unexpected events.

- **General Objective of the Research:**
    - Analyze the unpredictability of recessions and determine how the Monte Carlo Simulation can help interpret the randomness of economic crises.
- **Specific Objectives:**

1. Differentiate the concepts and define what constitutes an economic crisis/recession.
2. Separate and understand economic cycles.
3. Determine the factors that contribute to an economic recession.
4. Identify measurable factors.
5. Identify subjective factors and how they can be aggregated.
6. Identify why recessions are unpredictable.
7. Explain why the Monte Carlo Method can be an interesting approach for analyzing recessions.
8. Apply Monte Carlo Simulation in contexts with random variables.
9. Explain how the Monte Carlo Method works.
10. Determine a direct methodology for applying Monte Carlo Simulation to identify possible recessions in some countries.

Advances in the First Module

Going through the problem

The section on understanding the problem is essential for defining the scope of the research and its continuation. Defining a problem question is a careful process that involves identifying a problem or knowledge gap in a specific area. It begins with a broad literature review to understand the current state of research and the main issues under debate. Then, it is necessary to refine this information to focus on a particular aspect that has not been adequately explored or that presents a significant gap. The question must be clear, concise, and relevant to the field of study, as well as feasible in terms of time and available resources for the project. When formulating the question, it is essential to consider the possibility of collecting data and conducting analyses that can contribute to advancing knowledge in the chosen area.

Thus, for this project, the following problem question was chosen:

"Given the macroeconomic and internal context of a country, what is the probability of an economic recession occurring, considering the unpredictability of macroeconomic factors and the limited time window for analyzing its indicators?"

Formulation of Hypotheses and Definition of Variables:

The stage of hypothesis formulation corresponds to the process of inferring possibilities from the problem question. In other words, we list some assumptions, such as possible answers to the problem (based on literature review or expertise in the subject matter). They must be testable and falsifiable (i.e., they can be confirmed or refuted with data).

It is possible to have:

- **Main Hypotheses** (directly related to the general objective)
- **Secondary Hypotheses** (related to specific objectives)

For the project, some hypotheses were formulated based on the null and alternative hypotheses defined:

- **H_o (Null Hypothesis):** The occurrence of economic recessions is completely random and cannot be predicted based on macroeconomic and financial variables.
- **H_1 (Alternative Hypothesis):** The Monte Carlo Simulation can identify probabilistic patterns that increase the ability to interpret the randomness of economic crises.

Primary Hypotheses:

- **$H_1.1$:** Incorporating variables such as GDP, interest rates, and unemployment into Monte Carlo Simulation models allows generating probability distributions that approximate real economic cycles.
- **$H_1.2$:** The Monte Carlo Simulation, by incorporating random exogenous shocks, reproduces patterns similar to those observed in past economic crises.

- **$H_1.3$:** Stochastic modeling based on Monte Carlo presents greater predictive capacity over the onset of recessions than traditional linear models.
- **$H_1.4$:** The use of Monte Carlo allows distinguishing between normal economic fluctuations and signs of imminent systemic collapses.

Secondary Hypotheses:

- **$H_2.1$:** The randomness of financial shocks, when inserted into a Monte Carlo Simulation model, leads to results more consistent with the unpredictable nature of economic crises.
- **$H_2.2$:** The degree of sensitivity of the model to small variations in input parameters can be used as a proxy to measure economic instability.
- **$H_2.3$:** Simulations that incorporate correlations between macroeconomic variables (e.g., inflation and interest rates) present greater robustness in representing economic cycles.
- **$H_2.4$:** Including historical data from previous recessions improves the performance of the Monte Carlo Simulation in generating realistic scenarios.
- **$H_2.5$:** The Monte Carlo model tends to generate asymmetric distributions that better reflect the long-tail behavior typical of recessions.

Definition of Variables:

Variables are the elements of reality that you will measure or observe to test your hypotheses. They can be divided into the following types:

- **Independent Variable:** The one you believe influences another.
- **Dependent Variable:** The one that is affected.
- **Control Variable:** External variables that can interfere with the result.

From the reading and inference in scientific literature, the following variables were selected:

Scientific Methodology

The methodology process respected the understanding of the project as a whole, according to the process of observation and development of the first analyses for the project. Below is the definition of the step-by-step process followed for defining the methodology and the process defined in carrying out this stage:

Type of Research:

- Exploratory: The objective is to investigate, understand, and gather information about a phenomenon that is still little known or studied from a certain perspective, in this case, the unpredictability of recessions and the application of Monte Carlo Simulation.

Approach:

- Quantitative: Uses numerical and statistical data to test hypotheses, identify patterns, and analyze macroeconomic and financial variables.

Method:

- Case Study (Instrumental): Focuses on the 2008 recession as a case example with emphasis on the period and its particularities.
- Purpose of the Method: The case study approach aims to extract lessons and patterns applicable to other contexts, with an emphasis on generating theoretical and practical insights.
- Analysis: Will be conducted based on the complexity, behavior, and impact of recessions.

Data Collection Instruments:

- Collection of Secondary Data: Including economic indicators (GDP, LEI, CLI, and PMI, as well as the Foreclosure Index for the 2008 recession).
- Data Corresponding to a Period of Three Years: Analysis of the period before, during, and after the recession.
- Information from Official Databases, Economic Reports, and Specialized Literature.

Example of Collected Data on LEI, CLI, and PMI Indices Before and After the Recession (Analysis of Predictability of "Shock" and Recovery):

Indicator	Release Date	Initial Value	Annual Change	Analysis
LEI (TCB)	December 2007	-4/-4	-4/-4	Drop below the two-year support level.
CLI (OECD)	November 2007	-4/-4	-4/-2	Negative trend despite limited data.
PMI (ISM)	December 2007	-6/-2	0/0	Strong signal for the start of the 2008 recession.
LEI (TCB)	June 2009	0/0	+2/+2	Increase over the last three months.
CLI (OECD)	June 2009	-2/0	0/0	Low growth with limited data.
PMI (ISM)	July 2009	+6/+6	+6/+6	Strong signal for the end of the 2008 recession.

Additional Advancements:

Initial advances for the next module, but also obtained in the process of understanding the data. One of these surveys was the provisional definition of the "population" and sample of the research:

Population and Sample (Under Review):

Initial definition of how the population and sample would be divided:

- **Population:** Economic recession events documented globally.
- **Sample:** Historical cuts of economic crises with greater documentation and impact, especially (NOT YET 100% DEFINED):
    - The 2008 global financial crisis (main case).

- Possibly other selected recessions as support (e.g., the 2020 pandemic, the 1920 crisis) → (NOT YET FULLY DEFINED).

Another important part that was mapped was the development of the model itself:

Model Development:

The intention is that in the next module, the focus will be on developing the model and capturing the prediction results from the collected data. It is a process of much adjustment and refinement to understand if the variables really aid in the process of predicting a possible recession. The initial study will be conducted with data from the 2008 recession and its particularities. During the process, the following factors will also be considered:

- **Definition of KPIs and Metrics:** For evaluating the estimate and prediction of the model.
- **Understanding the Levels of Randomness:** That the model will have, and its classification within the prediction.
- **Capturing Patterns and Analyzing Trends:** In the process of developing the model.
- **Reflection on Simulation Insights.**