Documentation for Causal Impact Model

December 10, 2024

1. Overview

The **Causal Impact** model is used for time series analysis to measure the causal effect of an intervention (or treatment) on a target variable. It uses a Bayesian structural time-series framework to estimate the hypothetical outcome of a time series in the absence of an intervention, allowing it to measure the causal effect of the intervention by comparing observed data to the predicted behaviour. This model is especially useful for estimating the impact of events such as audits, notifications, or other changes in a system that may lead to changes in a behavior or outcome. In this context, the intervention or treatment date can represent any significant event or intervention, such as:

- Audits,
- Notifications,
- Policy changes,
- Any other treatment affecting the system.

2. Model Key Elements

1. **Pre-period and Post-period:** The model works by comparing the behavior of entities before and after the treatment. The **pre-period** is a baseline (pre-treatment), while the **post-period** is the period after the treatment (or intervention).

2. **Treatment Date:** This is the date of intervention. The function allows flexibility for the treatment date to represent various types of events, from audits to notifications or any other intervention.

3. **Data Preparation:** The data is pre-processed to handle missing values, duplicates, and date conversion. Additionally, fraud detection is integrated into the dataset for analysis.

4. **Significance Testing:** The **Causal Impact** model produces a p-value that helps in determining whether the observed change in the behavior of the entities is statistically significant.

3. Interpretation of Results

- **P-value Interpretation:**

- If the p-value is below 0.05, the treatment is considered to have a statistically significant effect on the outcome variable (e.g., sales or fraud behavior).

- A higher p-value suggests that the observed changes could be due to random variation and that the treatment did not have a significant impact.