

Learn How to Perform a Granger Causality Test in Python for Time Series Analysis

Authored by
Mohammed loot

November 1, 2025

RECOMMENDED CITATION

Mohammed loot (2025). *Learn How to Perform a Granger Causality Test in Python for Time Series Analysis*. PSYCHOLOGICAL STATISTICS. Retrieved from <https://statistics.arabpsychology.com/?p=7813>

The **Granger Causality test** stands as a fundamental statistical tool within the domain of [time series](#) econometrics and analysis. Developed by Nobel laureate Clive Granger, its core objective is to rigorously determine whether the lagged, historical values of one specific variable (the putative predictor) contribute statistically significant information for forecasting the subsequent future values of a second, distinct variable (the response). This test moves beyond simple correlation by assessing the direction and strength of predictive influence over time, making it indispensable for modeling dynamic systems.

It is absolutely crucial for practitioners and researchers to maintain a clear distinction: the identification of a Granger-causal relationship signifies a statistically significant **predictive relationship** based on lagged data, and emphatically does not imply true physical, structural, or mechanistic causation. This predictive nature means that if past values of variable X significantly improve the prediction of variable Y, then X Granger-causes Y. Misinterpretation of this statistical criterion can lead to flawed conclusions regarding underlying economic or physical processes.

Like all formal statistical procedures, the Granger Causality test is framed around two competing hypotheses, which dictate the interpretation of the test results. These hypotheses define the scope of the predictive influence we are examining:

Null Hypothesis (H0): Time series x does not Granger-cause time series y . This assumes that the lagged values of x provide no additional explanatory power for y beyond what the history of y itself already offers.

Alternative Hypothesis (HA): Time series x Granger-causes time series y . This is the claim that we are seeking evidence for--that the historical data of x significantly enhances the forecasting accuracy of y .

The formal interpretation of the phrase "Granger-causes" fundamentally rests on the concept of **added predictive power**. Specifically, if we can build a forecasting model for time series y , and then significantly improve the accuracy of that forecast by incorporating the lagged values of time series x , we conclude that x Granger-causes y . This improvement must be statistically significant, meaning it is highly unlikely to have occurred by random chance, thereby justifying the rejection of the null hypothesis in favor of the alternative.

The rigorous execution of the Granger Causality test ultimately yields a critical **F test statistic**, which quantifies the difference in the predictive error between the restricted model (excluding x lags) and the unrestricted model (including x lags). This statistic is accompanied by a corresponding **p-value**. The decision rule is straightforward: if the calculated p-value falls below the chosen predetermined significance level (conventionally set at $\alpha = 0.05$), we possess sufficient statistical evidence to confidently reject the null hypothesis, thereby concluding that x is indeed a significant Granger-cause of y .

To seamlessly execute this powerful analysis in a computational environment, we rely on the `grangercausalitytests` function, which is readily available within the statistical toolkit provided by the [statsmodels](#) package in [Python](#). This package is meticulously designed for estimating various statistical models and performing statistical tests, making it the industry standard for econometric time series analysis. The following snippet illustrates the basic syntax required to initiate the test:

```
from statsmodels.tsa.stattools import grangercausalitytests
```

```
# Perform Granger-Causality test on selected columns  
grangercausalitytests(df, maxlag=)
```

When structuring the input DataFrame for the `grangercausalitytests` function, meticulous attention must be paid to the order of variables. By convention in [statsmodels](#), the **dependent variable** (the time series being predicted, y) must be listed first in the column selection, immediately followed by the **predictor variable** (the potential causal factor, x). Furthermore, the critical **maxlag** parameter determines the maximum number of previous periods whose values will be incorporated into the underlying Vector Autoregression (VAR) model used to compute the test statistics. Selecting an appropriate lag length is vital, as insufficient lags might fail to capture the true predictive relationship, while excessive lags can reduce statistical power.

To solidify the theoretical understanding of Granger Causality, the subsequent sections provide a detailed, step-by-step practical demonstration. We will apply the `grangercausalitytests` function to a classic real-world dataset, illustrating the entire workflow from data preparation to final result interpretation.

Step 1: Data Acquisition and Preparation for Time Series Analysis

The initial and most critical phase of any econometric modeling process is preparing the raw data. For this illustrative demonstration, we will employ a well-known, classic [time series](#) dataset that captures the annual dynamics of the U.S. agricultural sector. Specifically, this dataset tracks two highly related variables: the total number of eggs manufactured and the corresponding population count of chickens, covering an extensive period from 1930 through 1983. This data provides an excellent real-world test case for assessing predictive relationships.

To interface efficiently with the [statsmodels](#) library, the data must first be imported and structured as a pandas DataFrame. The DataFrame object serves as the standard input format for nearly all advanced time series tools in the [Python](#) ecosystem, ensuring that variables are correctly aligned and indexed prior to statistical computation. We use the `read_csv` function, specifying the non-standard separator used in the source file:

import pandas as pd

```
# Define URL where the historical dataset is located
url = "https://raw.githubusercontent.com/Statology/Miscellaneous/main/chicken_egg.txt"

# Read in dataset as pandas DataFrame, noting the ' ' separator
df = pd.read_csv(url, sep=" ")

# View the initial five rows of the DataFrame to confirm successful loading
df.head()

year chicken egg
0 1930 468491 3581
1 1931 449743 3532
2 1932 436815 3327
3 1933 444523 3255
4 1934 433937 3156
```

With the data successfully loaded, our DataFrame now contains the two critical time series variables, `chicken` (the population count) and `egg` (the number of eggs manufactured). These variables are appropriately aligned by the observation year, making them perfectly ready for the subsequent causality analysis. Before proceeding, it is crucial to ensure that time series stationarity requirements are met, though for the purpose of demonstrating the function itself, we proceed directly to the causality test.

Note on Prerequisites: For effective visualizations or any additional exploratory data analysis (EDA) involving time series, always ensure that your DataFrame is properly structured and, ideally, indexed by the time variable (e.g., converting the `year` column into a proper DateTime index).

Step 2: Executing the Unidirectional Granger-Causality Test (Egg -> Chicken)

We now proceed to execute the primary test, aiming to determine if the time series representing the number of eggs manufactured (`egg`) holds predictive power for the future number of chickens (`chicken`). Following the [statsmodels](#) convention, we place the dependent variable (`chicken`) first, followed by the potential predictor (`egg`). We utilize the `maxlag` parameter, setting it to three periods, meaning our model incorporates data from the past three years to assess the lagged predictive influence.

The specific hypothesis under examination here is formalized as: **Null Hypothesis (H0): Egg production does not Granger-cause the chicken population.** If we find evidence to reject this null, we conclude that historical egg production volumes are statistically relevant for forecasting

chicken population changes.

```
from statsmodels.tsa.stattools import grangercausalitytests
```

```
# Perform Granger-Causality test: Is Chicken dependent on Egg? (Chicken, Egg)
grangercausalitytests(df, maxlag=)
```

Granger Causality

number of lags (no zero) 3

ssr based F test: F=5.4050 , p=0.0030 , df_denom=44, df_num=3

ssr based chi2 test: chi2=18.7946 , p=0.0003 , df=3

likelihood ratio test: chi2=16.0003 , p=0.0011 , df=3

parameter F test: F=5.4050 , p=0.0030 , df_denom=44, df_num=3

Upon careful review of the statistical output generated by [statsmodels](#), particularly focusing on the results derived from the Sum of Squared Residuals (SSR) based F test, we observe the resulting statistics. The critical value for decision-making is the [p-value](#), which, in this case, is reported as **0.0030**. This p-value indicates the probability of observing the test statistic (or one more extreme) if the null hypothesis were true.

Given that 0.0030 is significantly below the standard acceptable significance threshold ($\alpha = 0.05$), we have overwhelming evidence to confidently reject the null hypothesis. This definitive rejection leads us to the conclusion that the past number of eggs manufactured is indeed a statistically useful and significant predictor for forecasting the future population of chickens. The evidence suggests a strong predictive causality running from `egg` to `chicken`.

Step 3: Assessing for Reverse Causality and Final Interpretation

A comprehensive Granger Causality analysis necessitates checking for **bidirectional causality**. Even though we established a strong predictive link from eggs to chickens, it is imperative to investigate the reverse relationship. We must formally test the alternative scenario: does the historical count of the chicken population predict the future volume of eggs manufactured? This step prevents misidentifying correlation or a common driving factor as a single-direction causal link.

Executing the reverse test is structurally simple but conceptually vital. We maintain the same maximum lag (3 periods) but strategically swap the variable order in the function call. The time series `egg` now becomes the dependent variable (response), and `chicken` is designated as the predictor variable. This adjustment allows us to test the new null hypothesis: the chicken population does not Granger-cause egg production.

```
from statsmodels.tsa.stattools import grangercausalitytests
```

```
# Perform Granger-Causality test: Is Egg dependent on Chicken? (Egg, Chicken)
grangercausalitytests(df, maxlag=)
```

Granger Causality

number of lags (no zero) 3

ssr based F test: F=0.5916 , p=0.6238 , df_denom=44, df_num=3

ssr based chi2 test: chi2=2.0572 , p=0.5606 , df=3

likelihood ratio test: chi2=2.0168 , p=0.5689 , df=3

parameter F test: F=0.5916 , p=0.6238 , df_denom=44, df_num=3

The results from this reverse test provide a clear contrast to the forward test. The SSR based **F test statistic** is reported as **0.5916**, which is considerably low. More importantly, the corresponding **p-value** stands at **0.6238**. This high p-value signifies that if the null hypothesis were true (i.e., no predictive relationship), observing this test statistic is very probable.

Since 0.6238 is vastly greater than our predetermined significance level of 0.05, we lack sufficient statistical evidence to reject the null hypothesis for this reverse relationship. Consequently, we must conclude that the historical number of chickens is not a statistically significant predictor of the future number of eggs. There is no evidence of Granger causality in this direction.

In summary, the detailed analysis demonstrates that the predictive causality flows strictly in one direction: knowing the volume of eggs produced is highly useful for forecasting the subsequent chicken population, but the reverse predictive relationship does not hold true over the investigated lag period. This highlights the asymmetric nature that Granger causality often reveals in dynamic [time series](#) systems.

Additional Resources for Advanced Time Series Analysis

Mastering the **Granger Causality test** is often just one component of a comprehensive time series modeling strategy. Researchers frequently need to perform prerequisite tests, such as checking for stationarity, or move on to forecasting models like VAR or ARIMA. The following external resources provide detailed tutorials and documentation to help you perform other common, complex tasks within [Python](#)'s analytical environment: