

# How can I perform a Granger-Causality test in R to determine the causal relationship between two time series variables?

Authored by  
**stats writer**

April 24, 2024

## RECOMMENDED CITATION

stats writer (2024). *How can I perform a Granger-Causality test in R to determine the causal relationship between two time series variables?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=138977>

A Granger-Causality test is a statistical method used to determine the causal relationship between two time series variables. This test is commonly used in economics, finance, and other fields to analyze the relationship between a potential cause and its effect. In order to perform a Granger-Causality test in R, one must first ensure that the time series data is stationary and then use the "grangertest" function. This function will calculate the F-statistic and p-value, which can be used to determine if there is a significant causal relationship between the two variables. The results of the test can provide valuable insights into the relationship between the variables and aid in making informed decisions.

## Perform a Granger-Causality Test in R

**The Granger Causality test is used to determine whether or not one time series is useful for forecasting another.**

**This test uses the following null and alternative hypotheses:**

**Null Hypothesis (H0): Time series  $x$  does not Granger-cause time series  $y$**

**Alternative Hypothesis (HA): Time series  $x$  Granger-causes time series  $y$**

**The term "Granger-causes" means that knowing the value of time series  $x$  at a certain lag is useful for predicting the value of time series  $y$  at a later time period.**

This test produces an F test statistic with a corresponding p-value. If the p-value is less than a certain significance level (i.e.  $\alpha = .05$ ), then we can reject the null hypothesis and conclude that we have sufficient evidence to say that time series  $x$  Granger-causes time series  $y$ .

To perform a Granger-Causality test in R, we can use the `grangertest()` function from the package, which uses the following syntax:

```
grangertest(x, y, order = 1)
```

where:

**x:** The first time series  
**y:** The second time series  
**order:** The number of lags to use in the first time series. Default is 1.

The following step-by-step example shows how to use this function in practice.

**Step 1: Define the Two Time Series**

For this example, we'll use the `ChickEgg` dataset that comes pre-loaded in the `lmtest` package. This dataset

**contains values for the number of eggs manufactured along with the number of chickens in the U.S. from 1930 to 1983:**

```
#load lmtest packagelibrary(lmtest)
```

```
#load ChickEgg dataset  
data(ChickEgg)
```

```
#view first six rows of dataset  
head(ChickEgg)
```

```
chicken egg  
468491 3581  
449743 3532  
436815 3327  
444523 3255  
433937 3156  
389958 3081
```

**Step 2: Perform the Granger-Causality Test**

**Next, we'll use the `grangertest()` function to perform a Granger-Causality test to see if the number of eggs manufactured is predictive of the future number of chickens. We'll run the test using three lags:**

## #perform Granger-Causality test

```
grangertest(chicken ~ egg, order = 3, data = ChickEgg)
```

## Granger causality test

Model 1: `chicken ~ Lags(chicken, 1:3) + Lags(egg, 1:3)`

Model 2: `chicken ~ Lags(chicken, 1:3)`

Res.Df	Df	F	Pr(>F)
--------	----	---	--------

1	44		
---	----	--	--

2	47	-3	5.405 0.002966 **
---	----	----	-------------------

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

**Model 1:** This model attempts to predict the number of chickens using the number of chickens in the previous three years and the number of eggs in the previous three years as predictor variables.**Model 2:** This model attempts to predict the number of chickens using only the number of chickens in the previous three years as predictor variables.**F:** This is the F test statistic. It turns out to be 5.405.**Pr(>F):** This is the p-value that corresponds to the F test statistic. It turns out to be .002966.

**Since the p-value is less than .05, we can reject the null**

**hypothesis of the test and conclude that knowing the number of eggs is useful for predicting the future number of chickens.**

### **Step 3: Perform the Granger-Causality Test in Reverse**

**Although we rejected the null hypothesis of the test, it's actually possible that there is a case of reverse causation happening. That is, it's possible that the number of chickens is causing the number of eggs to change.**

**To rule out this possibility, we need to perform the Granger-Causality test in reverse, using chickens as the predictor variable and eggs as the :**

**#perform Granger-Causality test in reverse**

**grangertest(egg ~ chicken, order = 3, data = ChickEgg)**

**Granger causality test**

**Model 1: egg ~ Lags(egg, 1:3) + Lags(chicken, 1:3)**

**Model 2: egg ~ Lags(egg, 1:3)**

**Res.Df Df F Pr(>F)**

**1 44**

**2 47 -3 0.5916 0.6238**

**The p-value of the test is 0.6238. Since this isn't less than .05, we can't reject the null hypothesis. That is, the number of chickens isn't predictive of the future number of eggs.**

**Thus, we can conclude that knowing the number of eggs is useful for predicting the future number of chickens.**

ARABPSYCHOLOGY.COM