

How can White's test be performed in R, and what are some examples of using it?

Authored by
stats writer

April 23, 2024

RECOMMENDED CITATION

stats writer (2024). *How can White's test be performed in R, and what are some examples of using it?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=138388>

White's test is a statistical test used to check for heteroscedasticity in linear regression models. It is commonly performed in R using the "bptest" function from the "lmtest" package. This function takes the regression model as its input and calculates the test statistic and corresponding p-value. If the p-value is less than the chosen significance level, it indicates the presence of heteroscedasticity in the model. Some examples of using White's test include analyzing the relationship between income and education level, examining the impact of advertising on sales, and studying the effectiveness of different teaching methods on students' test scores. By performing White's test, researchers can ensure the validity of their regression models and make more accurate conclusions from their data.

Perform White's Test in R (With Examples)

White's test is used to determine if is present in a regression model.

Heteroscedasticity refers to the unequal scatter of at different levels of a in a regression model, which violates one of the key that the residuals are equally scattered at each level of the response variable.

This tutorial explains how to perform White's test in R to determine whether or not heteroscedasticity is a problem in a given regression model.

Example: White's Test in R

In this example we will fit a using the built-in R dataset mtcars.

Once we've fit the model, we'll use the bptest function

from the `lmtest` library to perform White's test to determine if heteroscedasticity is present.

Step 1: Fit a regression model.

First, we will fit a regression model using `mpg` as the response variable and `disp` and `hp` as the two explanatory variables.

```
#load the dataset
```

```
data(mtcars)
```

```
#fit a regression model
```

```
model <- lm(mpg~disp+hp, data=mtcars)
```

```
#view model summary
```

```
summary(model)
```

Coefficients:

```
Estimate Std. Error t value Pr(>|t|)
```

```
(Intercept) 30.735904 1.331566 23.083 < 2e-16 ***
```

```
disp -0.030346 0.007405 -4.098 0.000306 ***
```

```
hp -0.024840 0.013385 -1.856 0.073679 .
```

```
---
```

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

Residual standard error: 3.127 on 29 degrees of freedom

Multiple R-squared: 0.7482, Adjusted R-squared: 0.7309

F-statistic: 43.09 on 2 and 29 DF, p-value: 2.062e-09

Step 2: Perform White's test.

Next, we will use the following syntax to perform White's test to determine if heteroscedasticity is present:

```
#load lmtest library
```

```
library(lmtest)
```

```
#perform White's test
```

```
bptest(model, ~ disp*hp + I(disp^2) + I(hp^2), data = mtcars)
```

```
studentized Breusch-Pagan test
```

```
data: model
```

```
BP = 7.0766, df = 5, p-value = 0.215
```

Here is how to interpret the output:

The test statistic is $X^2 = 7.0766$. The degrees of freedom is 5. The corresponding p-value is 0.215.

White's test uses the following null and alternative hypotheses:

Null (H_0): Homoscedasticity is present. Alternative (H_A): Heteroscedasticity is present.

Since the p-value is not less than 0.05, we fail to reject the null hypothesis. We do not have sufficient evidence to say that heteroscedasticity is present in the regression model.

What To Do Next

If you fail to reject the null hypothesis of White's test then heteroscedasticity is not present and you can proceed to interpret the output of the original regression.

However, if you reject the null hypothesis, this means heteroscedasticity is present in the data. In this case, the standard errors that are shown in the output table of the regression may be unreliable.

There are a couple common ways that you can fix this issue, including:

1. Transform the response variable.

You can try performing a transformation on the response variable, such as taking of the response variable. Typically this can cause heteroscedasticity to go away.

2. Use weighted regression.

assigns a weight to each data point based on the variance of its fitted value. Essentially, this gives small weights to data points that have higher variances, which shrinks their squared residuals. When the proper weights are used, this can eliminate the problem of heteroscedasticity.