

What is the interpretation of a scale-location plot, and can you provide examples?

Authored by
stats writer

April 22, 2024

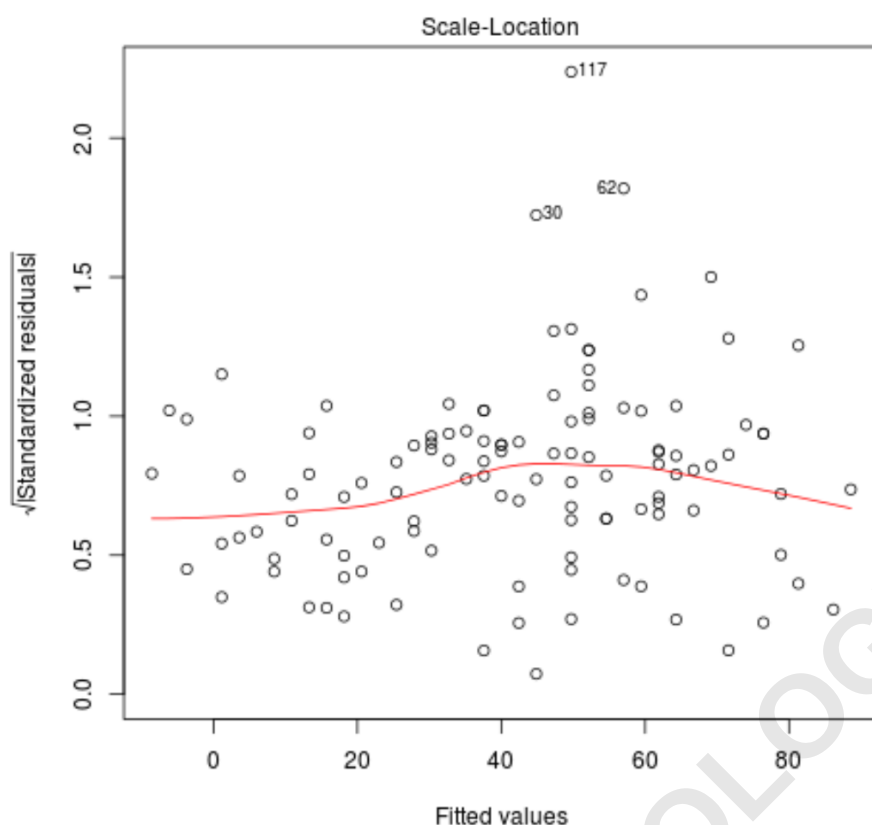
RECOMMENDED CITATION

stats writer (2024). *What is the interpretation of a scale-location plot, and can you provide examples?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=138164>

A scale-location plot is a graphical representation of data that is used to display the variability of a variable over a range of values. This type of plot is typically used in statistical analysis to assess the relationship between two variables. The x-axis of the plot represents the values of the independent variable, while the y-axis represents the variability of the dependent variable. The interpretation of a scale-location plot is that the closer the data points are to a straight line, the more likely it is that there is a linear relationship between the two variables. This type of plot is useful for identifying patterns, trends, and outliers in the data. Examples of scale-location plots include the relationship between height and weight, where the plot would show whether there is a linear relationship between the two variables, and the relationship between temperature and altitude, where the plot would show if there is a pattern in the variability of temperature as altitude increases.

Interpret a Scale-Location Plot (With Examples)

A scale-location plot is a type of plot that displays the fitted values of a regression model along the x-axis and the the square root of the standardized residuals along the y-axis.



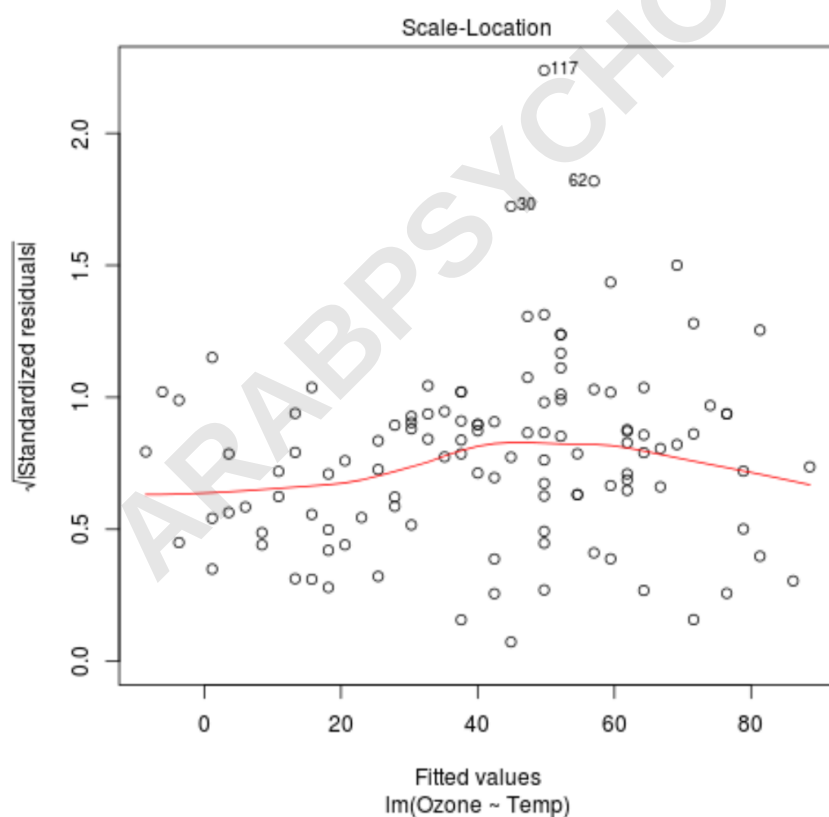
When looking at this plot, we check for two things:

1. Verify that the red line is roughly horizontal across the plot. If it is, then the assumption of homoscedasticity is likely satisfied for a given regression model. That is, the spread of the residuals is roughly equal at all fitted values.
2. Verify that there is no clear pattern among the residuals. In other words, the residuals should be randomly scattered around the red line with roughly equal variability at all fitted values.

Scale-Location Plot in R

We can use the following code to fit a simple linear regression model in R and produce a scale-location plot for the resulting model:

```
#fit simple linear regression model  
model <- lm(Ozone ~ Temp, data = airquality)  
  
#produce scale-location plot  
plot(model)
```



We can observe the following two things from the scale-

location plot for this regression model.

1. The red line is roughly horizontal across the plot. If it is, then the assumption of homoscedasticity is satisfied for a given regression model. That is, the spread of the residuals is roughly equal at all fitted values.
2. Verify that there is no clear pattern among the residuals. In other words, the residuals should be randomly scattered around the red line with roughly equal variability at all fitted values.

Technical Note

The three observations from the dataset with the highest standardized residuals are labelled in the plot.

We can see that the observations in rows 30, 62, and 117 have the highest standardized residuals.

This doesn't necessarily mean that these observations are outliers, but you may want to view the original data to take a closer look at these observations.

Although we can see that the red line is roughly

horizontal across the scale-location plot, this only serves as a visual way to see if the assumption of homoscedasticity is met.

A formal statistical test we can use to see if the assumption of homoscedasticity is met is the Breusch-Pagan Test.

Breusch-Pagan Test in R

The following code shows how to use the `bptest()` function from the `lmtest` package to perform a Breusch-Pagan Test in R:

```
#load lmtest package  
library(lmtest)
```

```
#perform Breusch-Pagan Test  
bptest(model)
```

studentized Breusch-Pagan test

data: model

BP = 1.4798, df = 1, p-value = 0.2238

A Breusch-Pagan Test uses the following null and

alternative hypotheses:

Null Hypothesis (H0): The residuals are homoscedastic (i.e. evenly spread)
Alternative Hypothesis (HA): The residuals are heteroscedastic (i.e. not evenly spread)

From the output we can see that the p-value of the test is 0.2238. Since this 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.

This result matches our visual inspection of the red line in the scale-location plot.

Understanding Heteroscedasticity in Regression Analysis

How to Create a Residual Plot in R

How to Perform a Breusch-Pagan Test in R