

Linear Regression Assumptions

Contents

Assumption 1: Linear Relationship	1
Assumption 2: Homoscedasticity.....	1
Assumption 3: Normality	1
Assumption 4: Independence	1

Assumption 1: Linear Relationship

Explanation: The assumption here is that there exists a linear relationship between independent variable and dependent variable

How to check: Make a scatter plot between independent and dependent variables and check visually for linear relationship.

Future work if it is violated: We can apply nonlinear transformation log, square root, inverse etc. to variables. We can visualize this plot to see if some nonlinear modification to variables can lead to linear relationship. For example, there may not be linear relationship between x and y but it exists between $\log(x)$ and y .

Assumption 2: Homoscedasticity

Explanation: Residuals should have constant variance at every level of dependent variable.

How to check: Examine the plot of residuals versus fitted values. There should be no relationship here. If there exist some pattern here it says that assumption is violated.

Future work if it is violated: We can transform dependent variable as mentioned above. Second thing we can do is to use weighted regression.

Assumption 3: Normality

Explanation: The assumption is that residuals are normally distributed.

How to check: Plot the histogram of residuals or we can use quantile-quantile plots.

Future work if it is violated: Check for outliers and verify it. Use transformations on variables as mentioned above.

Assumption 4: Independence

Explanation: The assumption here is that residuals are independent

How to check: Examine the plot of residuals versus fitted values. There should be no relationship here. The correlation should be close to zero.