Homoscedasticity Heteroscedasticity Analysis

What is Homoscedasticity?

Homoscedasticity refers to the equal variance in residuals ('residual' refers to the difference between observed and predicted data)

Here Amplitude is same

What is Heteroscedasticity?

Homoscedasticity refers to the different variance in residuals ('residual' refers to the difference between observed and predicted data)

Here Amplitude is different

Thumb rule of Homoscedasticity:

If the ratio of the largest sample variance to the smallest sample variance **does not exceed 1.5**, the groups satisfy the requirement of homoscedasticity

Thumb rule of Heteroscedasticity:

If the ratio of the largest sample variance to the smallest sample variance **does exceed 1.5**, the groups satisfy the requirement of Heteroscedasticity

When to use Homoscedasticity?

> Use homoscedasticity when the assumption of constant variance is met, and valid inferences in linear regression. For example, when examining the relationship between income and expenditure in a dataset where the variance of expenditure remains constant across different income levels.

When to use Heteroscedasticity?

Use heteroscedasticity when the assumption of constant variance is violated, such as in financial data where the variability of returns increases with higher investment levels.

Which is better?

Let's take one example: Imagine you're trying to predict the price of houses based on their size.

1. As per the example, homoscedasticity would mean that the variability in house prices remains constant regardless of the size of the house. In other words, if you plotted the sizes of houses against their prices, the spread of prices around the predicted line would be roughly the same for small houses as it is for large houses. **This consistency indicates that your model is performing well across the entire range of house sizes.**
2. As per the example, Heteroscedasticity would mean that the variability in house prices changes depending on the size of the house. This could lead to misleading conclusions, as it suggests that your model is better at predicting prices for certain sizes of houses than for others. **For example, if your model consistently underestimates prices for larger houses but overestimates prices for smaller houses, it may indicate a problem with heteroscedasticity.**

   Conclusion: homoscedasticity is good because it indicates that your model is reliable and consistent across different levels of the independent or predictor variable, leading to more accurate predictions.