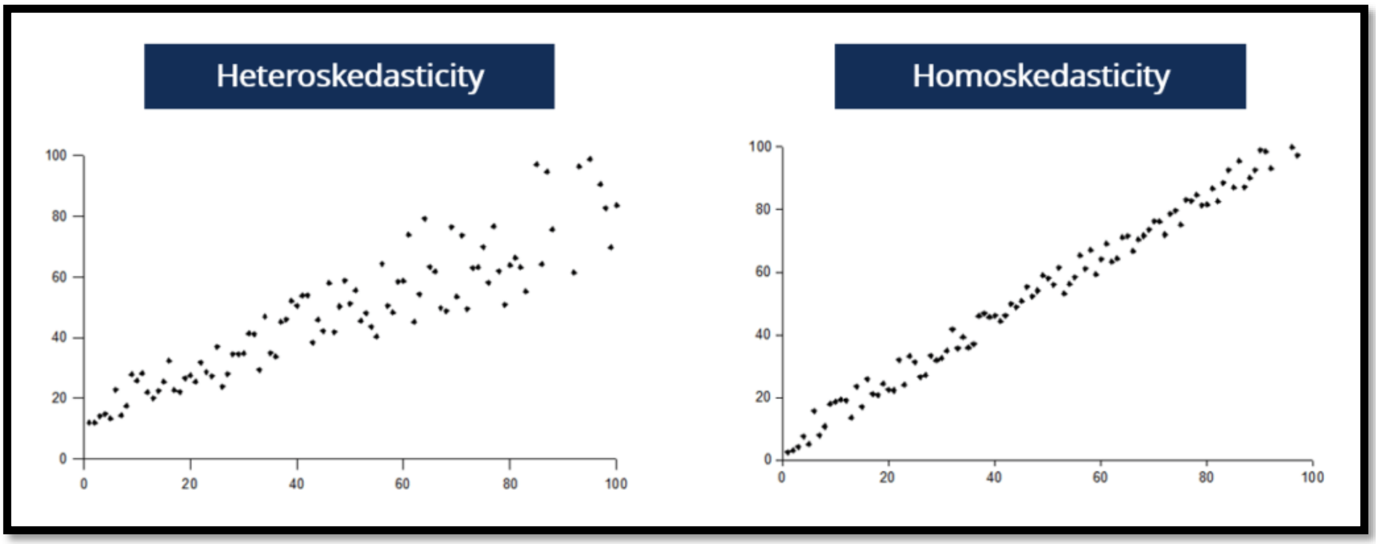
HOMOSCEDASTICITY AND HETEROSCEDASTICITY



Homoscedasticity

- **Definition**: The variance of the residuals (errors) is constant across all levels of the independent variable.
- **Implication**: It's a key assumption in linear regression and other parametric tests. When met, it ensures that model estimates are efficient and standard errors are unbiased.
- **Visual cue**: A scatterplot of residuals shows a uniform spread across predicted values.

You need to **check for homoscedasticity** when you're performing **linear regression analysis** or any statistical modelling that assumes **equal variance of errors** across observations.

Here's when and why it's important:

When to Test for Homoscedasticity

- **After fitting a regression model**: It's a key assumption of **ordinary least squares (OLS)** regression. You should test it to validate your model.
- **Before interpreting coefficients or making predictions**: Unequal error variance (heteroscedasticity) can distort standard errors, leading to unreliable confidence intervals and hypothesis tests.
- **When residual plots show patterns**: If residuals fan out or contract as fitted values increase, this suggests a violation of homoscedasticity.
- **In financial, economic, or cross-sectional data**: These often show changing variance across observations, so testing is crucial.

How to Test

- **Visual methods**: Plot residuals vs. fitted values.
- **Statistical tests**: Breusch-Pagan test, White test, Goldfeld-Quandt test.

Heteroscedasticity

- **Definition**: The variance of residuals changes across levels of the independent variable.
- **Implication**: Violates regression assumptions, leading to inefficient estimates and biased standard errors. This can affect hypothesis testing and confidence intervals.
- **Visual cue**: Residuals fan out or contract in a pattern, indicating non-constant variance.

You need to **test for heteroscedasticity** when you're performing **regression analysis** and want to ensure the reliability of your model's estimates and statistical inferences.

Here's a detailed breakdown:

Why and When to Test for Heteroscedasticity

- **After fitting a regression model**: Ordinary Least Squares (OLS) assumes that the variance of residuals is constant (homoscedasticity). If this assumption is violated, your model may produce biased standard errors and misleading results.
- **When residual plots show patterns**: If residuals fan out or contract as fitted values increase, this suggests non-constant variance—i.e., heteroscedasticity.
- **In financial, economic, or cross-sectional data**: These often exhibit changing variability across observations, making heteroscedasticity a common issue.
- **When dealing with large ranges of values**: If your independent variable spans a wide range, the variance of errors may increase with the magnitude of the variable.
- **Before conducting hypothesis, tests or building confidence intervals**: Heteroscedasticity affects the accuracy of these statistical measures, so testing is essential to ensure valid conclusions.

How to Detect It

- **Visual inspection**: Plot residuals vs. fitted values to look for patterns.

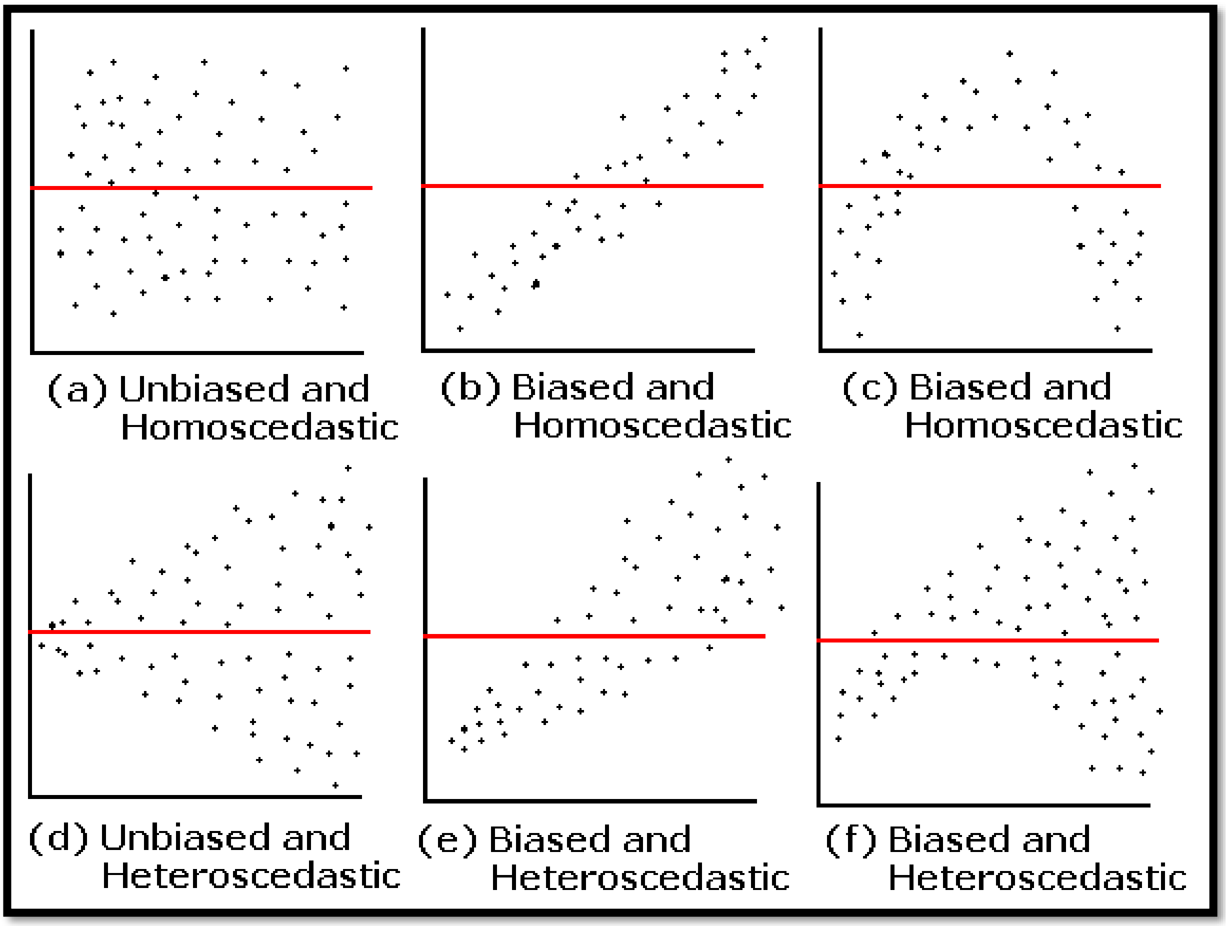
- **Statistical tests**:

    - *Breusch-Pagan test*
    - *White test*
    - *Goldfeld-Quandt test*

What to Do If Detected

- Use **robust standard errors** to correct for bias.
- Apply **weighted least squares (WLS)**.
- Transform variables (e.g., log transformation) to stabilize variance.

Here are six residual plots and their interpretations:



(a) Unbiased and homoscedastic. The residuals average to zero in each thin verical strip and the SD is the same all across the plot.

(b) Biased and homoscedastic. The residuals show a linear pattern, probably due to a lurking variable not included in the experiment.

(c) Biased and homoscedastic. The residuals show a quadratic pattern, possibly because of a nonlinear relationship. Sometimes a variable transform will eliminate the bias.

(d) Unbiased, but homoscedastic. The SD is small to the left of the plot and large to the right: the residuals are heteroscadastic.

(e) Biased and heteroscedastic. The pattern is linear.

(f) Biased and heteroscedastic. The pattern is quadratic.

We would also like the residuals to be normally distributed. We check this by looking at the normal plot of the residuals.

Here's a clear comparison table between **homoscedasticity** and **heteroscedasticity** to help you understand their differences:

Feature	Homoscedasticity	Heteroscedasticity
Definition	Constant variance of residuals across observations	Varying variance of residuals across observations
Residual Plot Pattern	Uniform spread of residuals	Funnel-shaped or uneven spread of residuals
Assumption Status	Assumes equal error variance (ideal for OLS)	Violates OLS assumption
Impact on Regression	Reliable standard errors and valid inference	Biased standard errors, invalid confidence intervals
Detection Methods	Residual plots, statistical tests (e.g., Breusch-Pagan)	Same methods, but look for non-constant variance
Remedies	No action needed if assumption holds	Use robust errors, transform variables, or WLS
Common in	Well-behaved, balanced datasets	Financial, economic, or cross-sectional data