Introduction to Regression Analysis

Understanding Simple Linear Regression

Kazeem Adesina DAUDA

AI4WIA BOOTCAMP GROUP

September 9, 2023

Outline

Regression Analysis

- Regression analysis is a statistical method used to examine relationships between variables.
- It helps us understand how changes in independent variables are associated with changes in a dependent variable.
- Regression models can be linear or nonlinear, but we'll focus on simple linear regression today.

Simple Linear Regression

Model

$$Y = \beta_0 + \beta_1 X + \epsilon$$

Y : Dependent Variable

X : Independent Variable (Predictor)

β_0 : Intercept

β_1 : Slope

ϵ : Error Term (Unexplained Variation)

Estimating Parameters

- The goal of regression analysis is to estimate the parameters β_0 and β_1 that best fit the data.
- Ordinary Least Squares (OLS) is a common method to estimate these parameters.
- OLS minimizes the sum of squared differences between observed values and predicted values.

Assumptions of Linear Regression

Linear regression makes several key assumptions about the data:

1. Linearity: The relationship between X and Y is linear.
2. Independence: Observations are independent.
3. Homoscedasticity: Constant variance of the error term ϵ across X.
4. Normality: Errors are normally distributed.

Inference and Hypothesis Testing

- Hypothesis testing assesses the significance of coefficients.
- For example, testing whether β_1 is significantly different from zero.
- A significant β_1 suggests a statistically significant relationship between X and Y.

Goodness of Fit

- Goodness of fit measures how well the regression model explains variation in Y.
- R^2 (Coefficient of Determination) represents the proportion of variation in Y explained by the model.
- Higher R^2 indicates a better fit.

Conclusion

- Regression analysis is a powerful statistical method for understanding relationships between variables.
- Simple linear regression models linear relationships between one independent variable and a dependent variable.
- OLS is used to estimate parameters, and several assumptions must be met.
- Hypothesis testing and R^2 help assess model significance and goodness of fit.

Questions?