

Regression

A scatter plot illustrating regression analysis. The plot features a set of data points represented by small circles in various colors (orange, blue, green, red, yellow). A solid red line with arrowheads at both ends represents the regression line, showing a positive linear correlation between the variables on the x and y axes. The axes are represented by black lines with arrowheads. The word "Regression" is written in a large, bold, black serif font across the center of the plot. A light gray horizontal band is visible in the upper portion of the plot area.

Regression Analysis

Regression analysis is a statistical technique used to model and analyze the relationship between a dependent variable (or target) and one or more independent variables (or predictors). Regression analysis is a mathematical approach that helps us to identify the variables that truly influence a situation.

It addresses essential questions:

- Which factors are the most significant?
- Which ones can be disregarded?
- How do these factors interrelate?
- Moreover, it allows us to determine the level of confidence associated with these variables

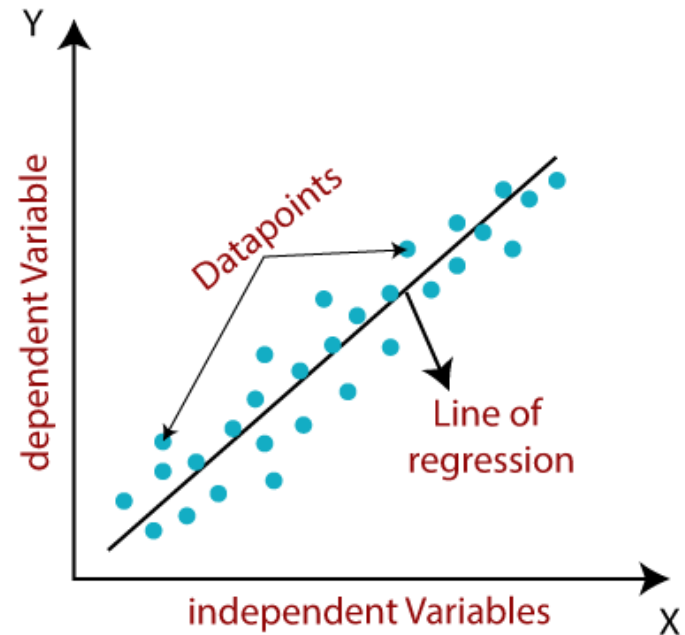
Types of Regression Analysis

Here are some common types of regression analysis based on different types of relationships and modeling scenarios

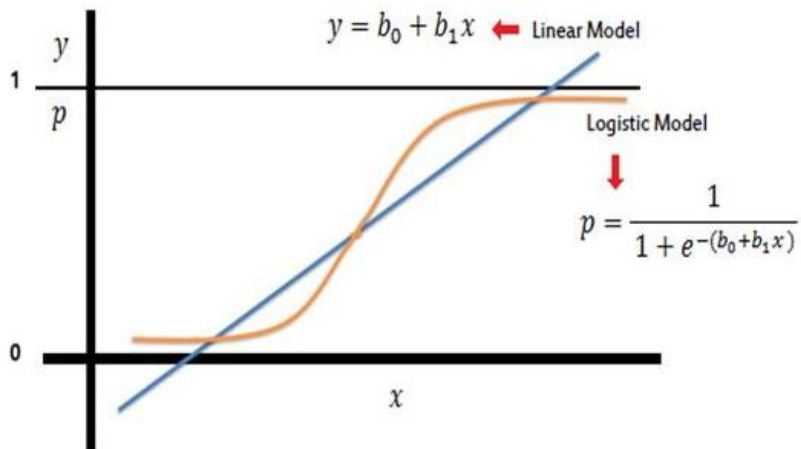
- | | |
|--------------------------|---------------------------|
| 1. Linear Regression | 7. Ordinal Regression |
| 2. Logistic Regression | 8. Poisson Regression |
| 3. Polynomial Regression | 9. Time Series Regression |
| 4. Ridge Regression | 10. Quantile Regression |
| 5. Lasso Regression | 11. Nonlinear Regression |

Linear Regression

Linear regression is used when the relationship between the dependent variable and the independent variables is linear. It's the most basic and widely used form of regression analysis.



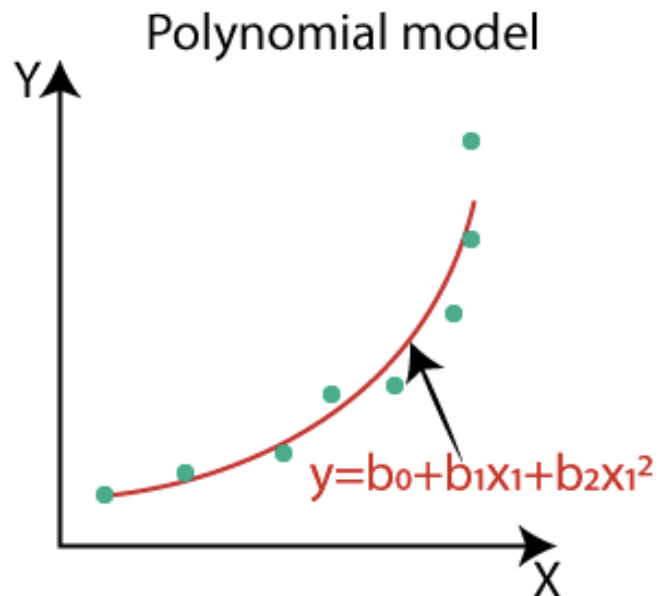
Logistic Regression



Logistic Regression: Logistic regression is used when the dependent variable is binary or categorical, representing a probability of one of two outcomes. It's commonly used for classification tasks and estimating probabilities.

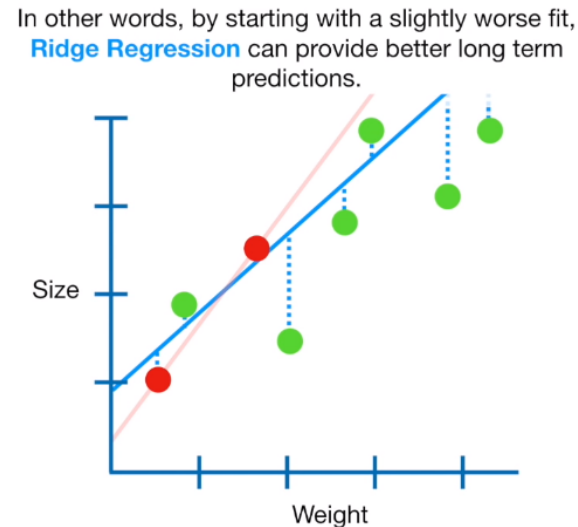
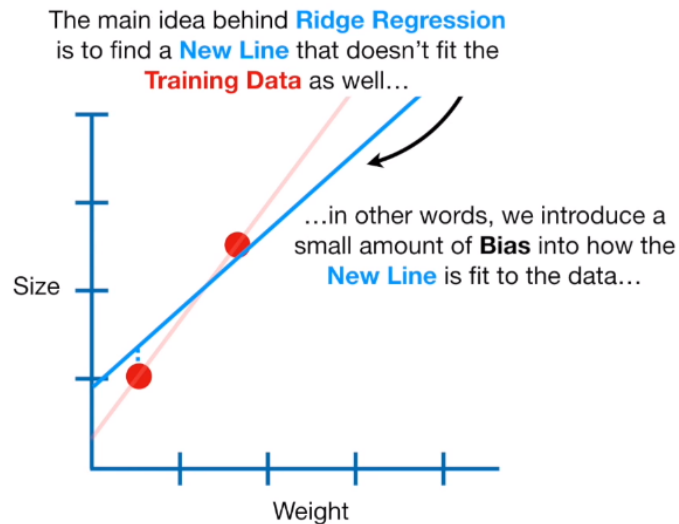
Polynomial Regression

Polynomial regression is an extension of linear regression that models the relationship between variables as a polynomial equation. It's used when a linear relationship is not sufficient to capture the underlying patterns in the data.

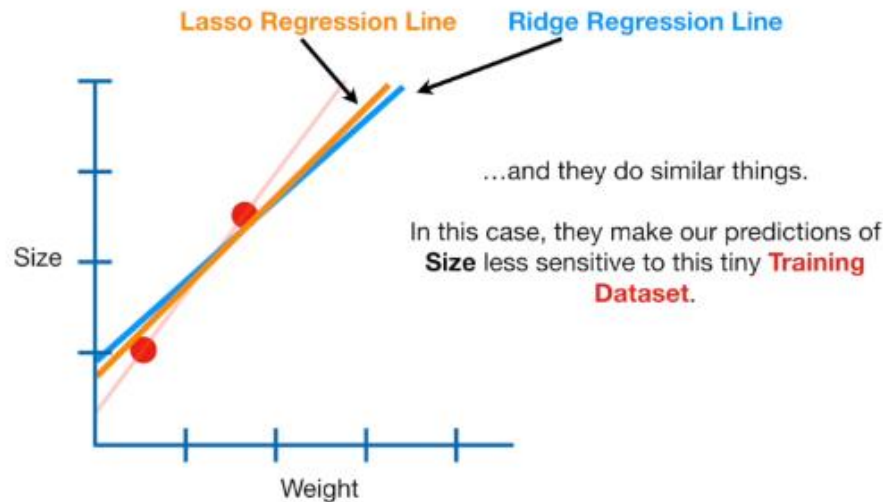


Ridge Regression

Ridge regression is a form of linear regression that includes L2 regularization to prevent overfitting. It's used when **multicollinearity** (high correlation between independent variables) is present.



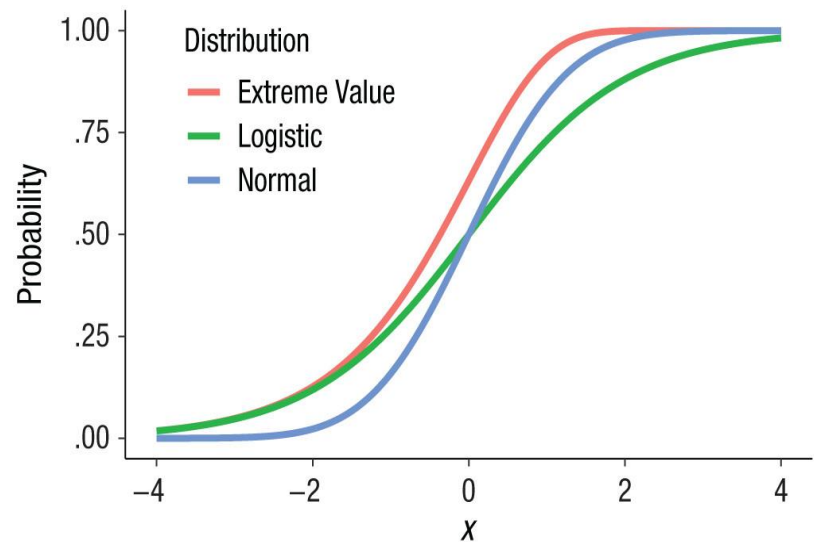
Lasso Regression



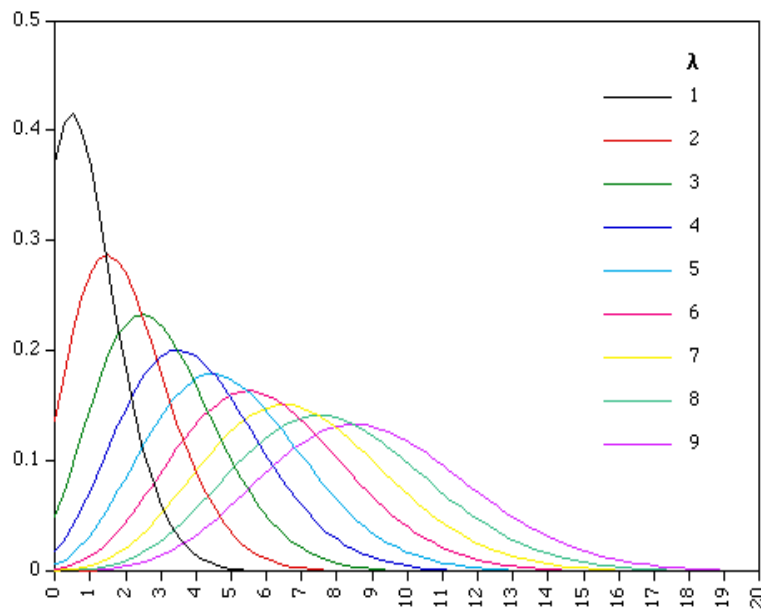
Lasso regression is similar to ridge regression but uses L1 regularization. It's often used for feature selection and can drive some regression coefficients to zero, effectively selecting a subset of important features.

Ordinal Regression

Ordinal regression is used when the dependent variable is ordinal, meaning it has ordered categories but the exact numerical difference between the categories is not defined.



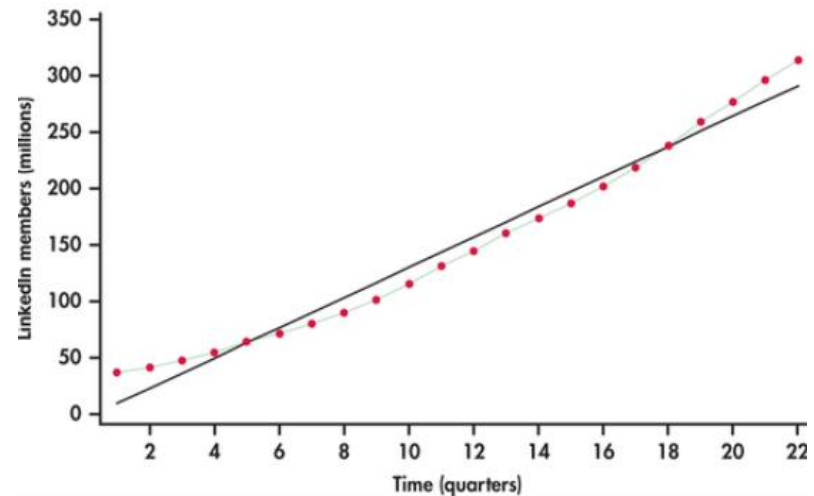
Poisson Regression



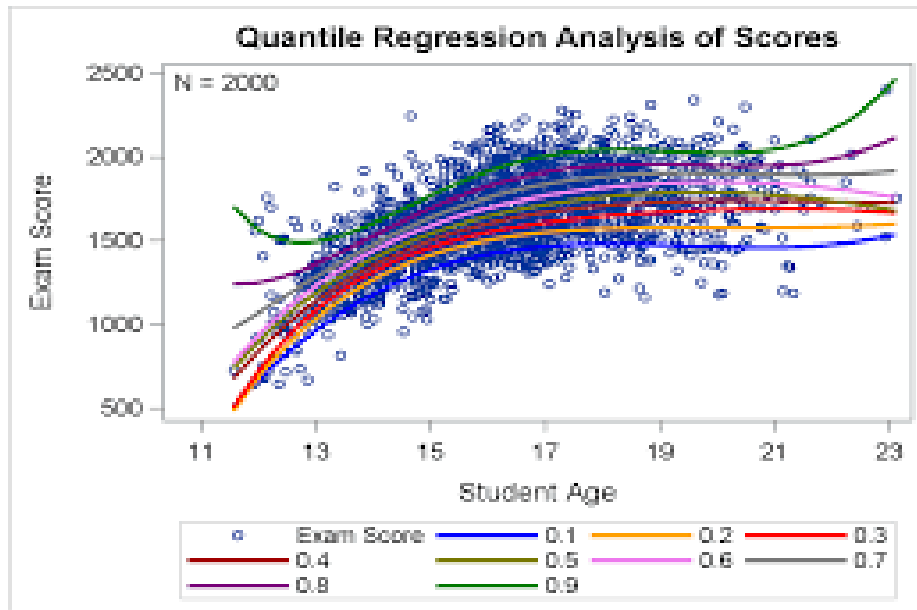
Poisson regression is used when the dependent variable represents count data and follows a Poisson distribution, such as the number of events or occurrences.

Time Series Regression

Time series regression is used when the data is collected over time, and the goal is to model and forecast time-dependent patterns and trends.



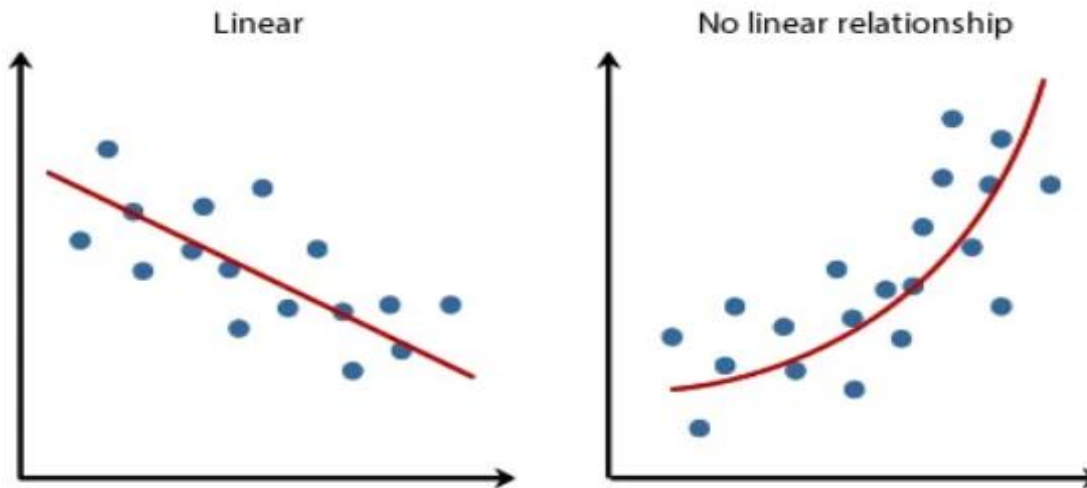
Quantile Regression



Quantile regression allows you to estimate different quantiles of the dependent variable, not just the mean, making it useful for analyzing data with [heteroscedasticity](#).

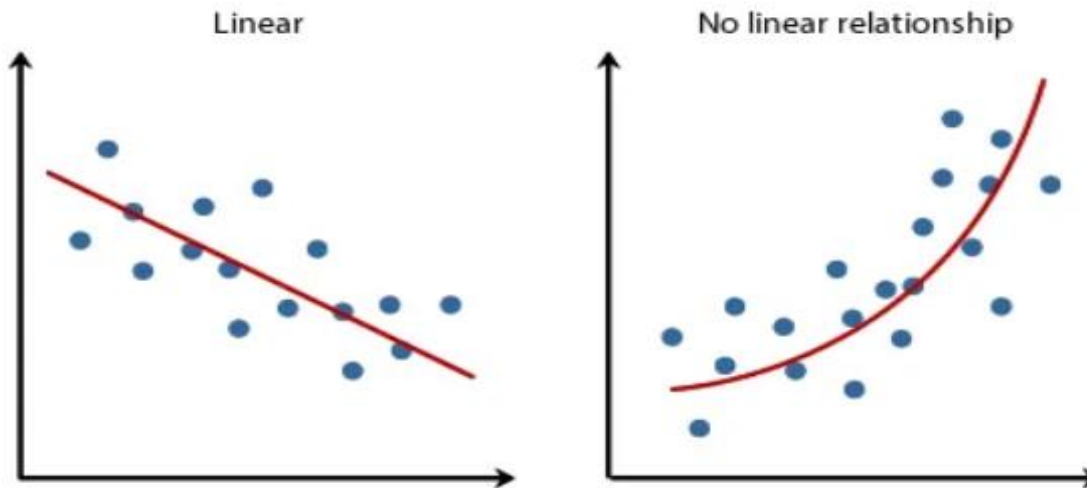
Non-linear Regression

Non-linear regression is used when the relationship between variables is non-linear. It can involve various types of nonlinear functions and modeling techniques.



Non-linear Regression

Non-linear regression is used when the relationship between variables is non-linear. It can involve various types of nonlinear functions and modeling techniques.



A scatter plot with a red regression line and the text 'Thank you'. The plot features a black L-shaped axis with arrows at the ends. A red line with an arrow at its end starts from the bottom-left and trends upwards to the right. Numerous small, semi-transparent circles in orange, blue, green, pink, and yellow are scattered across the plot area. A light gray horizontal band is positioned in the upper third of the image. The text 'Thank you' is centered in a large, bold, black serif font.

Thank you

Roksana Parvin