

# What are the 7 common types of regression and when should each one be used?

Authored by  
**stats writer**

May 3, 2024

## RECOMMENDED CITATION

stats writer (2024). *What are the 7 common types of regression and when should each one be used?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=142492>

Regression analysis is a statistical technique used to understand the relationship between a dependent variable and one or more independent variables. There are seven common types of regression: linear, logistic, polynomial, stepwise, ridge, lasso, and elastic net.

Linear regression is used when the relationship between the dependent and independent variables is expected to be linear. It is suitable for continuous variables and is often used for prediction and forecasting.

Logistic regression is used when the dependent variable is categorical or binary. It is commonly used in fields such as medicine and social sciences to predict the likelihood of an event occurring.

Polynomial regression is used when the relationship between the dependent and independent variables is expected to be non-linear. It involves fitting a curve to the data points and is useful for capturing complex relationships.

Stepwise regression is used when there are a large number of independent variables to choose from. It helps in selecting the most relevant variables for the model, thereby improving its accuracy.

Ridge, lasso, and elastic net regressions are used when there is multicollinearity (high correlation) among the independent variables. These techniques help in reducing the impact of multicollinearity on the model's performance.

Each type of regression has its own strengths and limitations, and the choice of regression depends on the nature of the data and the research question at hand. Understanding the different types of regression and their appropriate use can aid in making accurate and reliable predictions and inferences.

## 7 Common Types of Regression (And When to Use Each)

**Regression analysis is one of the most commonly used techniques in statistics.**

**The basic goal of regression analysis is to fit a model that best describes the relationship between one or more predictor variables and a .**

**In this article we share the 7 most commonly used regression models in real life along with when to use each type of regression.**

### **1. Linear Regression**

**Linear regression is used to fit a regression model that describes the relationship between one or more predictor variables and a numeric response variable.**

**Use when:**

**The relationship between the predictor variable(s) and the response variable is reasonably linear. The response variable is a continuous numeric variable.**

**Example: A retail company may fit a linear regression model using advertising spend to predict total sales.**

**Since the relationship between these two variables is likely linear (more money spent on advertising generally leads to an increase in sales) and the response variable (total sales) is a continuous numeric variable, it makes sense to fit a linear regression model.**

**Resource:**

## 2. Logistic Regression

**Logistic regression is used to fit a regression model that describes the relationship between one or more predictor variables and a binary response variable.**

**Use when:**

**The response variable is binary - it can only take on two values.**

**Example: Medical researchers may fit a logistic regression model using exercise and smoking habits to predict the likelihood that an individual experiences a heart attack.**

**Since the response variable (heart attack) is binary - an individual either does or does not have a heart attack - it's appropriate to fit a logistic regression model.**

**Resource:**

## 3. Polynomial Regression

**Polynomial regression is used to fit a regression model that describes the relationship between one or more predictor variables and a numeric response variable.**

## Use when:

The relationship between the predictor variable(s) and the response variable is non-linear. The response variable is a continuous numeric variable.

**Example:** Psychologists may fit a polynomial regression using 'hours worked' to predict 'overall happiness' of employees in a certain industry.

The relationship between these two variables is likely to be nonlinear. That is, as hours increases an individual may report higher happiness but beyond a certain number of hours worked, overall happiness is likely to decrease. Since this relationship between the predictor variable and response variable is nonlinear, it makes sense to fit a polynomial regression model.

## Resource:

### 4. Ridge Regression

Ridge regression is used to fit a regression model that describes the relationship between one or more predictor variables and a numeric response variable.

## Use when:

The predictor variables are highly correlated and **multicollinearity** becomes a problem. The response variable is a continuous numeric variable.

**Example:** A basketball data scientist may fit a ridge regression model using predictor variables like points, assists, and rebounds to predict player salary.

The predictor variables are likely to be highly correlated since better players tend to get more points, assists, and rebounds. Thus, multicollinearity is likely to be a problem so we can minimize this problem by using ridge regression.

## Resource:

### 5. Lasso Regression

Lasso regression is very similar to ridge regression and is used to fit a regression model that describes the relationship between one or more predictor variables and a numeric response variable.

## Use when:

The predictor variables are highly correlated and multicollinearity becomes a problem. The response variable is a continuous numeric variable.

**Example:** An economist may fit a lasso regression model using predictor variables like total years of schooling, hours worked, and cost of living to predict household income.

The predictor variables are likely to be highly correlated since individuals who receive more schooling also tend to live in cities with higher costs of living and work more hours. Thus, multicollinearity is likely to be a problem so we can minimize this problem by using lasso regression.

Note that Lasso regression and ridge regression are quite similar. When multicollinearity is a problem in a dataset, it's recommended to fit both a Lasso and Ridge regression model to see which model performs best.

**Resource:**

6. Poisson Regression

**Poisson regression is used to fit a regression model**

**that describes the relationship between one or more predictor variables and a response variable.**

**Use when:**

**The response variable consists of "count" data - e.g. number of sunny days per week, number of traffic accidents per year, number of calls made per day, etc.**

**Example: A university may use Poisson regression to examine the number of students who graduate from a specific college program based on their GPA upon entering the program and their gender.**

**In this case, since the response variable consists of count data (we can "count" the number of students who graduate - 200, 250, 300, 413, etc.) it's appropriate to use Poisson regression.**

**Resource:**

## **7. Quantile Regression**

**Quantile regression is used to fit a regression model that describes the relationship between one or more predictor variables and a response variable.**

## **Use when:**

**We would like to estimate a specific quantile or percentile of the response variable - e.g. the 90th percentile, 95th percentile, etc.**

**Example: A professor may use quantile regression to predict the expected 90th percentile of exam scores based on the number of hours studied:**

**In this case, since the professor is interested in predicting a specific percentile of the response variable (exam scores), it's appropriate to use quantile regression.**

## **Resource:**