

# How to Easily Identify the 3 Types of Logistic Regression

Authored by  
**stats writer**

December 4, 2025

## RECOMMENDED CITATION

stats writer (2025). *How to Easily Identify the 3 Types of Logistic Regression*.

PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=104859>

Logistic Regression (LR) is a foundational algorithm in both statistics and machine learning, primarily categorized as a type of supervised learning method. Unlike traditional linear regression, which predicts a continuous numerical output, LR is expertly designed to predict the probability of a specific outcome belonging to a discrete class or category. When performing regression analysis, LR is utilized when the dependent variable--the target we aim to predict--is categorical outcome. It elegantly transforms input features into a probability score constrained between 0 and 1, facilitating classification decisions.

This powerful predictive capability is partitioned into three primary types, defined by the nature and structure of the categorical target variable: **binary logistic regression**, **multinomial logistic regression**, and **ordinal logistic regression**. Each model addresses unique scenarios concerning the number of categories and whether those categories possess an inherent rank or order. Understanding these distinctions is paramount for accurate model selection in data science applications.

## Understanding the Fundamentals of Logistic Regression

**Logistic regression** refers to any regression model in which the dependent variable, or the response variable, is **categorical**. It achieves classification by modeling the probability of an event occurring using the logit function, which is the natural logarithm of the odds. This transformation allows the model to map linear combinations of predictor variables to a probability scale, ensuring the output is always between zero and one.

The decision of which type of logistic regression to employ hinges entirely on the characteristics of the response variable. Specifically, we must consider both the total count of possible outcomes and whether a logical sequence or hierarchy exists among those outcomes.

## The Role of the Response Variable in Model Selection

There are three fundamental types of logistic regression models, each tailored for a specific structure of the response variable:

**Binary logistic regression:** This model is used when the response variable can only belong to one of two mutually exclusive categories. This is the most common and foundational application of LR.

**Multinomial logistic regression:** This model is deployed when the response variable can belong to one of three or more categories, but crucially, there is absolutely no natural ordering or sequence among these categories (e.g., colors, professions, types of transport).

**Ordinal logistic regression:** This specialized model is required when the response variable can belong to one of three or more categories and there *is* a natural, meaningful ordering or ranking

among the categories (e.g., satisfaction levels, educational grades).

The following table provides a quick, visual summary illustrating these critical differences based on the nature of the outcome variable:

**Types of Logistic Regression Models**

|                                            | Binomial Logistic Regression | Multinomial Logistic Regression | Ordinal Logistic Regression |
|--------------------------------------------|------------------------------|---------------------------------|-----------------------------|
| Number of Categories for Response Variable | 2                            | 3 or more                       | 3 or more                   |
| Does Order of Categories Matter?           | No                           | No                              | Yes                         |

This detailed tutorial proceeds with a comprehensive explanation of each type of logistic regression model, providing real-world examples to clarify the appropriate context for their application.

## Type #1: Binary Logistic Regression

**Binary logistic regression models** are the standard application of the algorithm, restricted to scenarios where the response variable can only have two outcomes. These outcomes are typically coded as 0 (representing failure, absence, or the reference category) and 1 (representing success, presence, or the event of interest). The core purpose is to estimate the probability that the event ( $P=1$ ) occurs, given a set of predictor variables.

This type of modeling is essential for tasks requiring a simple yes/no or pass/fail prediction. For instance, determining whether a customer will churn or stay, or predicting if a medical test result is positive or negative. The model calculates the odds ratio for the event happening versus not happening, which is then converted into a probability using the sigmoid function.

Here are two detailed examples illustrating the practical use of binary logistic regression:

### Binary Regression Example 1: NBA Draft Prediction

Suppose a **sports data scientist** is tasked with building a model to predict the career trajectory of college athletes. The scientist wants to utilize key performance indicators from their collegiate careers--specifically, (1) average points scored, (2) average rebounds, and (3) average assists--to predict the likelihood that a given college basketball player will be drafted into the National Basketball Association (NBA).

Since the response variable--the outcome for the player--is strictly limited to two possible states

(either **drafted** or **not drafted**), the data scientist must utilize a binary logistic regression model. This approach estimates the probability  $P(\text{text}\{\text{Drafted}\})$  based on the player's statistics, allowing the scientist to rank players by their predicted success likelihood.

### Binary Regression Example 2: Email Spam Detection

Consider a technology business aiming to improve its email filtering system. The goal is to accurately classify incoming emails. The business decides to use predictor variables such as (1) the total word count of the email and (2) the country of origin of the sender's IP address to predict the probability that a given email is **spam**.

In this scenario, the response variable also has only two outcomes: the email is **spam** (1) or **not spam** (0, often referred to as 'ham'). Because of this dichotomous nature, employing a binomial logistic regression model is the correct statistical choice for developing a robust and efficient spam detection system.

### Type #2: Multinomial Logistic Regression

**Multinomial logistic regression models**, also sometimes referred to as polytomous logistic regression, are utilized when the response variable can belong to one of three or more distinct categories, and critically, there is **no natural ordering** among these categories. If an ordering were present, an ordinal model would be preferred. The multinomial model treats all outcomes equally in terms of rank, focusing only on the differentiation between them.

Modeling with multinomial LR involves simultaneously estimating multiple binary logistic regressions. Specifically, one category is selected as the reference category, and the model then calculates the relative probability of each other category occurring versus the reference category. This results in a set of coefficients defining the relationship between predictors and the log-odds for each non-reference outcome.

This approach is highly valuable in fields like market research, sociology, and political science, where choices often involve multiple, equally weighted options.

### Multinomial Regression Example 1: Political Preference Prediction

Imagine a **political scientist** conducting research to understand voter behavior. They wish to use predictor variables like (1) the individual's annual income and (2) their years of formal education to predict which one of four different presidential candidates (A, B, C, or D) an individual is most likely to vote for.

Since there are more than two possible outcomes (four potential candidates) for the response variable, and because there is no inherent or natural ordering among the candidates (i.e.,

candidate A is not inherently "better" or "higher" than candidate B), the political scientist must employ a **multinomial logistic regression model**. This model provides the probability of voting for each candidate relative to a chosen baseline candidate.

### Multinomial Regression Example 2: Preferred Sport Analysis

Suppose a **sports analyst** wants to understand demographic influences on leisure activity. They plan to use predictor variables such as (1) total TV hours viewed per week and (2) age to predict an individual's preference among three major sports: basketball, football, or baseball.

Again, the response variable involves more than two possible choices (three sports). Furthermore, there is no standardized ranking; one sport is not universally considered "higher" or "lower" than another in a statistical sense. Therefore, the sports analyst would use a **multinomial logistic regression model** to predict the likelihood of selecting each sport based on the input features.

### Type #3: Ordinal Logistic Regression

**Ordinal logistic regression models** are specialized for situations where the response variable can belong to one of three or more categories, and crucially, there *is* a clear, intrinsic, and measurable natural ordering or hierarchy among these categories. Examples of ordinal scales include rankings, satisfaction ratings (e.g., poor, fair, good, excellent), or levels of agreement (e.g., strongly disagree, disagree, neutral, agree, strongly agree).

This model takes advantage of the inherent order by estimating a single set of coefficients across all categories, rather than estimating separate coefficients for each pairwise comparison (as in multinomial regression). The most common method involves the Proportional Odds Model, which assumes that the effect of the predictor variables is consistent across the different thresholds that separate the categories.

Failing to account for this inherent order by using a multinomial model would result in a loss of statistical power and potentially misleading conclusions, as the relationship between categories would be ignored.

### Ordinal Regression Example 1: University Rating Prediction

Consider an **academic advisor** developing a predictive tool for student university placement. The advisor wants to use predictor variables such as (1) High School GPA, (2) ACT score, and (3) SAT score to predict the quality category of the university a student will gain admission to. These categories are defined as "bad," "mediocre," "good," or "great."

Since there are more than two possible outcomes (four distinct classifications of school quality) for the response variable, and since there *is* a strict, natural ordering among these outcomes (i.e.,

"bad" < "mediocre" < "good" < "great"), the academic advisor must utilize an **ordinal logistic regression model**. This model correctly incorporates the ranking inherent in the university classifications.

### Ordinal Regression Example 2: Movie Rating Scores

Suppose a **movie critic** is building a model to predict critical reception based on production characteristics. The critic intends to use predictor variables like (1) total run time and (2) movie genre to predict the probability that a given movie will receive a discrete rating score between 1 and 10.

In this scenario, the response variable involves multiple outcomes (ten possible ratings). Crucially, there *is* a precise, natural ordering among these outcomes (a rating of 1 is quantitatively worse than a rating of 10). Because the response variable is categorical yet ordered, the movie critic would accurately apply an **ordinal logistic regression model** to predict the ratings.

### Summary Comparison of Logistic Regression Models

Selecting the correct logistic regression model is vital for rigorous data analysis. The choice is always dictated by the structure of the dependent variable. If the outcome is **dichotomous** (two categories), the binary model is appropriate. If the outcome has **three or more categories with no order**, the multinomial model is required. If the outcome has **three or more categories with a definite order**, the ordinal model is the correct and most statistically efficient approach.

Mastering these distinctions allows analysts and data scientists to apply the proper statistical tool, ensuring that the model assumptions align with the data structure and yielding the most accurate and interpretable results.