

What is the interpretation of the intercept in logistic regression and can you provide an example?

Authored by
stats writer

June 23, 2024

RECOMMENDED CITATION

stats writer (2024). *What is the interpretation of the intercept in logistic regression and can you provide an example?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=148580>

Logistic regression is a statistical method used to analyze the relationship between a categorical dependent variable and one or more independent variables. In this type of regression, the intercept represents the log odds of the dependent variable when all independent variables are equal to zero. In other words, it is the expected log odds of the dependent variable when all independent variables have no effect.

For example, in a logistic regression model predicting the likelihood of a person being a smoker (dependent variable) based on their age (independent variable), the intercept would represent the log odds of a person being a smoker when their age is zero. Since age cannot be zero, the intercept in this case would not have a meaningful interpretation. However, it is still necessary to include in the model to account for the baseline log odds of the dependent variable.

Interpret Logistic Regression Intercept (With Example)

Logistic regression is a method we can use to fit a regression model when the is binary.

When we fit a logistic regression model, the intercept term in the model output represent the log odds of the response variable occurring when all predictor variables are equal to zero.

However, since log odds is hard to interpret we usually frame the intercept in terms of probability.

We can use the following formula to understand the probability that the response variable occurs, given that each predictor variable in the model is equal to zero:

$$P = e^{\beta_0} / (1 + e^{\beta_0})$$

The following example shows how to interpret a logistic regression intercept in practice.

Example: How to Interpret Logistic Regression Intercept

Suppose we would like to fit a logistic regression model using gender and number of practice exams taken to predict whether or not a student will pass a final exam in some class.

Suppose we fit the model using statistical software (such as , , , or) and receive the following output:

	Coefficient Estimate	Standard Error	Z-Value	P-value
Intercept	-1.34	0.23	5.83	<0.001
Gender (Male = 1)	-0.56	0.25	2.24	0.03
Practice Exams	1.13	0.43	2.63	0.01

We can see that the intercept term has a value of -1.34.

This means that when Gender is equal to zero (i.e. the student is female) and when Practice Exams is equal to zero (the student took no practice exams in preparation for the final exam) the log odds of the student passing the exam are -1.34.

Since log odds are hard to understand, we can instead

rewrite things in terms of probability:

Probability of Passing = $e^{\beta_0} / (1 + e^{\beta_0})$
Probability of Passing = $e^{-1.34} / (1 + e^{-1.34})$
Probability of Passing = 0.208

When both predictor variables are equal to zero (i.e. a female student who took no prep exams), the probability that the student passes the final exam is 0.208.

ARABPSYCHOLOGY.COM