

What is the method for calculating Odds Ratios in a Logistic Regression Model?

Authored by
stats writer

June 23, 2024

RECOMMENDED CITATION

stats writer (2024). *What is the method for calculating Odds Ratios in a Logistic Regression Model?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=148554>

The method for calculating Odds Ratios in a Logistic Regression Model involves using a statistical model to estimate the likelihood of an event occurring given a set of independent variables. This is achieved by fitting a logit function to the data, which transforms the probability of the event into an odds ratio. The odds ratio represents the ratio of the odds of the event occurring in one group compared to another group. This method allows for the identification of significant predictors and their impact on the outcome, providing valuable insights for making informed decisions. Additionally, the use of odds ratios in logistic regression models allows for the examination of non-linear relationships and the assessment of the strength and direction of the relationship between variables.

R: Calculate Odds Ratios in Logistic Regression Model

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

When you fit a logistic regression model in R, the coefficients in the model summary represent the average change in the log of the odds of the response variable associated with a one unit increase in each predictor variable.

However, we're often interested in calculating the odds ratio for the predictor variables in the model instead.

To quickly calculate the odds ratios for each predictor variable in the model, you can use the following syntax:

`exp(coef(model))`

You can also calculate a 95% confidence interval for each odds ratio by using the following syntax:

```
exp(cbind(Odds_Ratio = coef(model), confint(model)))
```

The following example shows how to use this syntax to calculate and interpret odds ratios for a logistic regression model in R.

Example: Calculating Odds Ratios in Logistic Regression Model in R

For this example, we'll use the Default dataset from the ISLR package in R.

We can use the following code to load and view a summary of the dataset:

```
library(ISLR)
```

```
#view first five rows of Default dataset
```

```
head(Default)
```

```
default student balance income
```

```
1 No No 729.5265 44361.625
```

```
2 No Yes 817.1804 12106.135
```

3 No No 1073.5492 31767.139

4 No No 529.2506 35704.494

5 No No 785.6559 38463.496

6 No Yes 919.5885 7491.559

This dataset contains the following information about 10,000 individuals:

default: Indicates whether or not an individual defaulted.
student: Indicates whether or not an individual is a student.
balance: Average balance carried by an individual.
income: Income of the individual.

We will use student status, bank balance, and income to build a logistic regression model that predicts the probability that a given individual defaults.

We can use the glm unction and specify family='binomial' so that R fits a logistic regression model to the dataset:

```
#fit logistic regression model  
model <- glm(default~student+balance+income,  
family='binomial', data=Default)
```

```
#disable scientific notation for model summary
options(scipen=999)
```

```
#view model summary
summary(model)
```

Call:

```
glm(formula = default ~ student + balance + income,
family = "binomial",
data = train)
```

Deviance Residuals:

```
Min 1Q Median 3Q Max
-2.5586 -0.1353 -0.0519 -0.0177 3.7973
```

Coefficients:

```
Estimate Std. Error z value Pr(>|z|)
(Intercept) -11.478101194 0.623409555 -18.412
<0.00000000000000002 ***
studentYes -0.493292438 0.285735949 -1.726 0.0843 .
balance 0.005988059 0.000293765 20.384
<0.00000000000000002 ***
income 0.000007857 0.000009965 0.788 0.4304
---
```

```
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 2021.1 on 6963 degrees of freedom

Residual deviance: 1065.4 on 6960 degrees of freedom

AIC: 1073.4

Number of Fisher Scoring iterations: 8

For example, a one unit increase in balance is associated with an average increase of 0.005988 in the log odds of defaulting.

To instead calculate the odds ratio for each predictor variable, we can use the following syntax:

```
#calculate odds ratio for each predictor variable  
exp(coef(model))
```

(Intercept) studentYes balance income

```
0.00001903854    0.52373166965    1.00575299051  
1.00000303345
```

We can also calculate each odds ratio along with a 95% confidence interval for each odds ratio:

#calculate odds ratio and 95% confidence interval for each predictor variable

exp(cbind(Odds_Ratio = coef(model), confint(model)))

Odds_Ratio 2.5 % 97.5 %

(Intercept) 0.00001903854 0.000007074481 0.0000487808
studentYes 0.52373166965 0.329882707270
0.8334223982
balance 1.00575299051 1.005308940686 1.0062238757
income 1.00000303345 0.999986952969 1.0000191246

The odds ratio for each coefficient represents the average increase in the odds of an individual defaulting, assuming all other predictor variables are held constant.

For example, the predictor variable balance has an odds ratio of 1.0057.

This means for each additional dollar in the balanced carried by an individual, the odds that the individual defaults on their loan increase by a factor of 1.0057, *assuming student status and income are held constant.*

We can interpret the odds ratios for the other predictor

variables in a similar manner.

How to Interpret $Pr(>|z|)$ in Logistic Regression Output in R

ARABPSYCHOLOGY.COM