sas

Print

Odds Ratios

To measure the strength of the association between a binary predictor variable and a binary response variable, you can use an odds ratio. An odds ratio indicates how much more likely it is that a certain event, or outcome, occurs in one group relative to its occurrence in another group. For example, suppose you want to see the effect that fiscal training has on spending habits. The people in group B attended a seminar about saving money, and the people in group A did not. The outcome variable indicates whether each person saved at least 10% of income in the year following the seminar.

In this example, you want to know how much more likely it is that a Yes outcome occurs in Group B relative to its occurrence in Group A, with respect to the odds. So, what are odds? Odds are not the same as probabilities. Instead, odds are calculated from probabilities. You divide the probability that the event occurs by the probability that the event does not occur.

$$\text{Odds} = \frac{P_{event}}{1 - P_{event}}$$

Group	**Outcome**		
	Yes	**No**	**Total**
A	60	20	80
B	90	10	100
Total	150	30	180

To calculate the probability of a Yes outcome in Group B, you divide the number of Yes's (90) by the total number of observations (100) to get .90, or 90%, the probability of having the outcome in Group B. The probably of a No outcome in Group B is 10 divided by 100, which is 0.10, or 10%. To calculate the probability of a Yes outcome in Group A, you divide 60 by 80, which is 0.75. The probability of a No outcome in Group A is 20 divided by 80, which is 0.25.

Now that we know the conditional probabilities of our response given the predictor, let's calculate the odds. The odds of the outcome occurring in Group B are the probability of a Yes outcome, .90, divided by the probability of a No outcome, .10. The odds are 9, or 9:1, which means that we expect nine occurrences to one non-occurrence in Group B. To calculate the odds of the outcome occurring in Group A, you divide the probability of a Yes outcome, .75, by the probability of a No outcome, .25, which is 3, or 3:1. This means that you expect three times as many occurrences as non-occurrences in Group A.

Now that you know the odds of the outcome in both groups, you can compare the odds by calculating an odds ratio. You divide the odds of an outcome in Group B, 9, by the odds in Group A, 3, with a result of 3. An odds ratio of 3 means that the odds of getting the outcome in Group B are three times those of getting the outcome in Group A. So, in this example, the odds of saving at least 10% of income are three times higher for people who received training. The odds ratio shows the strength of the association between the predictor variable and the outcome variable, and the value can range from 0 to infinity. When the odds ratio is 1, there's no association between the predictor variable and the outcome variable. If the odds ratio is greater than 1, the group in the numerator (here Group B) is more likely to have the outcome. If the odds ratio is less than 1, the group in the denominator (here Group A) is more likely to have the outcome.

The odds ratio is approximately the same regardless of sample size. To estimate the true odds ratio while considering the variability of the sample statistic, you can calculate confidence intervals just as you can for the unknown parameter youre trying to estimate. A 95% confidence interval for an odds ratio means that if you sample repeatedly and calculate a confidence interval for each sample odds ratio, 95% of the time your confidence interval will contain the true population odds ratio.

You can use an odds ratio to test for significance between two categorical variables. If the 95% confidence interval does not include 1, where a value of 1 indicates the odds are the same, then the odds ratio is significant at the 0.05 level. The true odds ratio is significantly different from 1, and therefore, you would conclude that there is an association between the two variables. If the 95% confidence interval includes 1, the odds ratio is not significant at the 0.05 level. You don't have enough evidence to conclude that the true odds are different, and the true odds ratio is significantly different from 1. So, there's not enough evidence to conclude an association between the two variables.

Close