Print

Interactions between Variables

Just as in linear regression, if you suspect there are interactions between predictors, you can fit a more complex logistic regression model by including interaction effects. Remember, an interaction is present when the effect of one variable on the outcome depends on, or changes, due to another variable. Let's look at an example.

Suppose you want to know how a customer's income affects the probability that the customer will default on a loan. In our logistic regression model, the response, which we'll call Default, is binary. A customer can either default or not default. Here's a graph of the logit of the predicted probability for defaulting on a loan by the customer's income. Predicted logit is on the Y-axis, and income is on the x-axis. As a customer's income increases, the probability of defaulting on a loan decreases. However, suppose we also look at the previous credit experience of the customer.

Now let's add the predictor BadCredit to our logistic regression model, a binary variable that specifies whether each customer has any previous defaults. As the graph shows, there's a difference in the effect of income on people with good credit compared to those with bad credit. For people who have a good credit record, as income increases, the logit of the probability for defaulting on a loan barely decreases. However, when we look at only the people who have bad credit, an increase in income is associated with a big decrease in the probability of defaulting.

There's an interaction between a person's credit record and income. This means that the effect of a person's income on the probability that the person will default on a loan depends on the person's credit record. The number of possible interaction, or combinations of variables, depends on the number of main effects. In this example, there's one possible interaction: a two-factor interaction for Income and BadCredit.

Suppose we want to look at the effect of education level, as well as income and credit history, on the probability of defaulting on a loan. How many possible interactions are there? A model with three main effects can have three two-factor interactions and one three-factor interaction. So, there are four possible interactions for this example. Keep in mind that interactions that have more than two factors might be difficult to interpret.

Close