Likelihood Ratio Test
Testing to Determine if Coefficient Estimates for Logistic Regression Models are Equal For Two or More Models
Vartanian: SW541

Testing Differences Among Groups

If you're trying to determine if the coefficient estimates for the two groups (such as high school dropouts and high school graduates) are different from one another (you can reject the null hypothesis that they're the same), run three logistic regressions. One regression will use all the observations with all the independent variables. The other two regressions will use the separate group observations – one for high school dropouts the other for high school graduates, for example.

You will test this with a chi-square test with the number of degrees of freedom given below.

Degrees of Freedom=(U_1+U_2) - R

Where U_1 is the number of independent variables plus the intercept (k+1) for the first group, U_2 is the number of independent variables plus the intercept for the second group, and R is the number of independent variables plus the intercept for the full model. In other words, if you're examining 10 independent variables, U_1=11, U_2=11, and R=11.

The test uses the log of the likelihood function for the logistic regression estimates

$$\times^2 \rightarrow -2[L(R) - L(\hat{Q})]$$

Where R is the full model and Q hat are the separate models. So for two groups, you will subtract two L(Q) hat values. If the chi-square value is greater than the critical value, you can reject the null hypothesis.

Let's say you're examining the differential effects of the independent variables on poverty status for two groups: high school dropouts and high school graduates. The way to figure out whether the two groups are different is by expanding the above formula –

X^2 = -2 L(R) +2L(Dropouts) + 2 L(Graduates)

In your output, you're generally given the -2 Log Likelihood. You'll keep this figure for the full model (both dropouts and graduates), and subtract off the value of -2 Log Likelihood for the two separate groups. On the next few pages are logistic regression results for the likelihood of poverty status, examining two independent variables, number of children and whether the person lives in a big city. U1, U2 and R are equal to 3 (k+1 – the one for the intercept). The degrees of freedom are therefore = 3+3-3 =3.

The values for the chi-square analysis are the following:

-2 Log Likelihood for the full model=4350.35,
-2 Log Likelihood for Dropouts is 2209.60,
-2 Log Likelihood for Graduates =1889.87.

and thus, the X^2 = 4350.35 - 2209.60 - 1889.87 = 250.88 .

At 3 DFs, the critical value for a .05 test is 7.824. Since the Chi-Square value is greater than the critical value, reject the null hypothesis. This states there is support for the research hypothesis that states that the independent variables have differential effects on the dependent variable for the two groups.

Testing Differences Among Model Specifications

You're wondering if the effects of additional variables better explain the model than the model without these additional variables. Run two models: the more parsimonious model and the expanded model. Subtract -2Log Likelihood of the larger from the smaller, and test this with a chi-square test, with DF = the difference in the number of variables between the two models.