Print

Demo: Fitting a Multiple Logistic Regression Model with All Odds Ratios Using PROC LOGISTIC

Filename: **st107d06.sas**

In this demonstration, we want to refine the multiple logistic regression model that we fit in the last demonstration. Now we want to produce the odds ratios for each value of the variables that are involved in an interaction from the final model.

PROC LOGISTIC DATA=*SAS-data-set <options>*;
    CLASS *variable <(options)> ... < / options>*;
    MODEL *variable <(variable_options)> = <effects> < / options>*;
    UNITS *<independent1=list1> ... < / options>*;
    ODDSRATIO *< 'label' > variable < / options>*;
RUN;

1. Open program st107d06.sas.

```
/*st107d06.sas*/  /*Part A*/
proc logistic data=STAT1.ameshousing3 plots(only)=(effect oddsratio);
    class Fireplaces(ref='0') Lot_Shape_2(ref='Regular') / param=ref;
    model Bonus(event='1')=Basement_Area|Fireplaces|Lot_Shape_2 @2 /
          selection=backward clodds=pl slstay=0.10;
    units Basement_Area=100;
    title 'LOGISTIC MODEL (3): Backward Elimination '
           'Bonus=Basement_Area|Fireplaces|Lot_Shape_2';
run;

/*st107d06.sas*/  /*Part B*/
proc logistic data=STAT1.ameshousing3
              plots(only)=oddsratio(range=clip);
    class Fireplaces(ref='0') Lot_Shape_2(ref='Regular') / param=ref;
    model Bonus(event='1')=Basement_Area|Lot_Shape_2 Fireplaces;
    units Basement_Area=100;
    oddsratio Basement_Area / at (Lot_Shape_2=ALL) cl=pl;
    oddsratio Lot_Shape_2 / at (Basement_Area=1000 1500) cl=pl;
    title 'LOGISTIC MODEL (3.1): Bonus=Basement_Area|Lot_Shape_2 Fireplaces';
run;
```

In the modified PROC LOGISTIC step in Part B, the PLOTS= option now specifies only an odds ratio plot, but includes the RANGE=CLIP option. This option is helpful when one or more odds ratio confidence intervals are so large that the scale makes it difficult to see the smaller ones.

The MODEL statement specifies all the significant terms that remained in the final model. At the end of the MODEL statement, notice that we've removed the SELECTION= and CLODDS= options that specify the backward elimination method and profile-likelihood confidence limits. The profile-likelihood confidence limits are now specified in the ODDSRATIO statements that we've added.

To produce the odds ratios for each value of a variable thats involved in an interaction, you specify a separate ODDSRATIO statement for each variable. In the first ODDSRATIO statement, we specify Basement_Area followed by a forward slash, and the AT option. The AT option specifies fixed levels of one or more interacting variables. For each categorical variable, you can specify a list of one or more formatted levels of the variable, the keyword REF to select the reference level, or the keyword ALL to select all levels of the variable. Here, we're requesting all levels of the variable Lot_Shape_2.

The second ODDSRATIO statement produces odds ratios for Lot_Shape_2 at 1000 and 1500 square feet of Basement_Area.

2. Submit the PROC LOGISTIC step in Part B.

3. Review the output.

   Jump to the Odds Ratios. There are four odds ratios displayed for the interaction effects. The first two show the odds ratios comparing homes with a difference of 100 square feet of basement area holding the lot shape constant. For example, the odds of being bonus eligible for a home with a regular lot shape are almost 3 times greater than a home with 100 square feet less of basement area. The last two odds ratios compare the odds ratios for irregular lots shapes compared to regular when holding the basement area constant. For example the odds of being bonus eligible are more than 20 times greater for homes with irregular lot shapes vs regular when holding the basement area fixed at 1000 square feet.

   We can assess the significance of the odds ratios from either the table or the graphic. In the table, an odds ratio is significant when the confidence interval does not include the value 1. In the PL or Profile-Likelihood plot, a vertical reference line at 1 makes it easy to assess significance. Odds ratios whose confidence intervals do not overlap the reference line are statistically significant.

   From this plot, it's clear that the lot shape effect is different at different values of basement area. The lot shape effect is highly significant when the Basement_Area is set to 1000 square feet, but not when Basement_Area is set to 1500 square feet, as you can see the confidence interval for the odds ratio covers a value of 1.

Close