

Polynomial regression terms for logistic regression in SPSS Statistics

Technote (FAQ)

Referenznummer: 1640516

Änderungsdatum: 2016-03-11

Question

I have a binary dependent variable (Y) and a continuous predictor (X) and wish to perform a logistic regression. I would like to add a quadratic term for X, i.e. X squared, to the model and test its contribution to the model. I can see how interaction terms can be added to the model in the Binary Logistic Regression dialogs, i.e., highlighting all terms involved in the interaction in the variable list on the left of the main dialog, then clicking the ">a*b>" button to enter the interaction term in the Covariates box. However, I do not see a way to entered a polynomial term for a predictor into the covariates box. How can a polynomial model be run in SPSS Statistics?

Answer

There are several procedures in SPSS Statistics which will perform a binary logistic regression. None of these procedures allow you to enter a polynomial term directly into the Model or Covariates box in the procedure dialogs, unless that polynomial term is represented by a predictor variable that is in the open data set and distinct from the variable that represents the linear term. However, some of these procedures will allow you to add a polynomial term to the model in the syntax command for the model. The options for these various procedures are described below.

#####

Binary Logistic Regression (Analyze->Regression->Binary Logistic):

In the Binary Logistic Regression procedure (LOGISTIC REGRESSION command), the only way to add a polynomial term for X to the model is to compute the polynomial term(s) as new variables and add those variables to the model. You can compute the quadratic term for X, for example, by opening the Transform->Compute variable menu. Type the name of the quadratic term variable (XSQ, for example) in the Target box, type X^{**2} in the Numeric Expression box and click OK.

When you open the Logistic Regression dialog, add both X and XSQ to the covariate box. If you wish to print a chi-square test for the improvement due to adding the quadratic term to the model, you can add the XSQ term to the model in a second block.

See Technote [1481143](#) for the steps to build a logistic regression model hierarchically.

Here are example syntax commands for computing the quadratic term variable and then including it in the Logistic Regression analysis.

X and XSQ are entered in separate blocks in this example.

```
COMPUTE XSQ = X**2,  
LOGISTIC REGRESSION VARIABLES y  
/METHOD=ENTER x  
/METHOD=ENTER xsq
```

```
/CLASSPLOT
/PRINT=GOODFIT CI(95)
/CRITERIA=PIN(0.05) POUT(0.10) ITERATE(20) CUT(0.5).
#####
```

Multinomial Logistic Regression (Analyze->Regression->Multinomial Logistic):

The Multinomial Logistic Regression procedure (NOMREG command) is designed to analyze nominal dependent variables with more than 2 categories, but you can model binary dependent variables as well. (Be careful to note that NOMREG treats the highest value of the dependent variable as the reference category by default. See Technotes [1476703](#) and [1479637](#) for a discussion of this default and instructions for choosing a different reference category.)

The Multinomial Logistic Regression procedure will not directly add a polynomial term to the model in the dialogs, if variables have not been computed which represent that polynomial term. However, you can add an X*X term. for example, to the /MODEL subcommand of the NOMREG command. You can build the model in the Multinomial Logistic Regression dialogs without the higher-order terms, then click the Paste button, rather than the OK button, to paste the NOMREG command to a syntax window. The pasted command would look something like this:

```
NOMREG Y (BASE=FIRST ORDER=ASCENDING) WITH X
/CRITERIA CIN(95) DELTA(0) MXITER(100) MXSTEP(5) CHKSEP(20) LCON-
VERGE(0) PCONVERGE(0.000001)
SINGULAR(0.00000001)
/MODEL=X
/STEPWISE=PIN(.05) POUT(0.1) MINEFFECT(0) RULE(SINGLE) ENTRY-
METHOD(LR) REMOVALMETHOD(LR)
/INTERCEPT=INCLUDE
/PRINT=PARAMETER SUMMARY LRT CPS STEP MFI.
```

You can add X*X to the /MODEL subcommand in the Syntax window as follows:

```
NOMREG Y (BASE=FIRST ORDER=ASCENDING) WITH X
/CRITERIA CIN(95) DELTA(0) MXITER(100) MXSTEP(5) CHKSEP(20) LCON-
VERGE(0) PCONVERGE(0.000001)
SINGULAR(0.00000001)
/MODEL=X X*X
/STEPWISE=PIN(.05) POUT(0.1) MINEFFECT(0) RULE(SINGLE) ENTRY-
METHOD(LR) REMOVALMETHOD(LR)
/INTERCEPT=INCLUDE
/PRINT=PARAMETER SUMMARY LRT CPS STEP MFI.
```

Note that you do not add the X*X term after the WITH keyword at the top of the NOMREG command. Only variables that exist in the active file should be listed there (or after the BY keyword if you have factors).

```
#####
```

Ordinal Regression (Analyze->Regression->Ordinal)

You can run a binary logistic regression with the Ordinal Regression procedure (PLUM command) when the dependent variable is binary.

Note that the parameterization in PLUM is different than the parameterization in Logistic Regression, so the signs of the intercepts may change across procedures. Quadratic and other higher-order terms must be represented by existing variables in the data set, so the XSQ variable must be added to the model for this example. These variables can be entered in the Ordinal Regression dialog or in the PLUM command. Here is an example of a PLUM command with the linear and quadratic terms for X entered as the variables X and SQ.

```
PLUM y WITH X XSQ
/CRITERIA=CIN(95) DELTA(0) LCONVERGE(0) MXITER(100) MXSTEP(5) PCON-
VERGE(1.0E-6) SINGULAR(1.0E-8)
/LINK=LOGIT
/PRINT=FIT PARAMETER SUMMARY.
```

#####

Generalized Linear Models (Analyze->Generalized Linear Model->Generalized Linear Model)

Like NOMREG, the Generalized Linear Model procedure (GENLIN command) does not allow you to enter a polynomial term in the dialogs but it does allow you to enter a higher-order term as an interaction of the predictor with itself in the /MODEL sub-command of the GENLIN command. For example:

```
GENLIN Y (REFERENCE=FIRST) WITH X
/MODEL X X*X INTERCEPT=YES
DISTRIBUTION=BINOMIAL LINK=LOGIT
/CRITERIA METHOD=FISHER(1) SCALE=1 COVB=MODEL MAXITERATIONS=100
MAXSTEPHALVING=5
PCONVERGE=1E-006(ABSOLUTE) SINGULAR=1E-012 ANALYSIS-
TYPE=3(WALD) CILEVEL=95 CITYPE=WALD
LIKELIHOOD=FULL
/MISSING CLASSMISSING=EXCLUDE
/PRINT CPS DESCRIPTIVES MODELINFO FIT SUMMARY SOLUTION (EXPO-
NENTIATED).
```

If you have a variable in the data set with the quadratic term for X, for example, you can add that variable to the model in either the dialogs or the GENLIN syntax command, as in the following example with the XSQ variable as defined earlier in this technical note.

```
GENLIN Y (REFERENCE=FIRST) WITH X XSQ
/MODEL X XSQ INTERCEPT=YES
DISTRIBUTION=BINOMIAL LINK=LOGIT
/CRITERIA METHOD=FISHER(1) SCALE=1 COVB=MODEL MAXITERATIONS=100
MAXSTEPHALVING=5
PCONVERGE=1E-006(ABSOLUTE) SINGULAR=1E-012 ANALYSIS-
TYPE=3(WALD) CILEVEL=95 CITYPE=WALD
LIKELIHOOD=FULL
/MISSING CLASSMISSING=EXCLUDE
/PRINT CPS DESCRIPTIVES MODELINFO FIT SUMMARY SOLUTION (EXPO-
NENTIATED).
```