* [Fitting a Gamma Regression to Car Insurance Claims (Generalized Linear Models)](http://127.0.0.1:54857/help/topic/com.ibm.spss.modeler.tutorial/spss/tutorials/genlin_car-ins_intro.htm)

# Fitting a Gamma Regression to Car Insurance Claims (Generalized Linear Models)

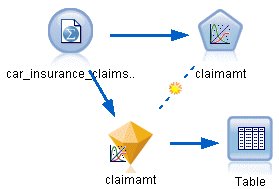
A generalized linear model can be used to fit a Gamma regression for the analysis of positive range data. For example, a dataset presented and analyzed elsewhere [1](http://127.0.0.1:54857/help/topic/com.ibm.spss.modeler.tutorial/spss/tutorials/genlin_car-ins_intro.htm#fntarg_1) concerns damage claims for cars. The average claim amount can be modeled as having a gamma distribution, using an inverse link function to relate the mean of the dependent variable to a linear combination of the predictors. In order to account for the varying number of claims used to compute the average claim amounts, you specify Number of claims as the scaling weight.

This example uses the stream named car-insurance\_genlin.str, which references the data file named car\_insurance\_claims.sav. The data file is in the Demos folder and the stream file is in the streams subfolder.

# Creating the Stream

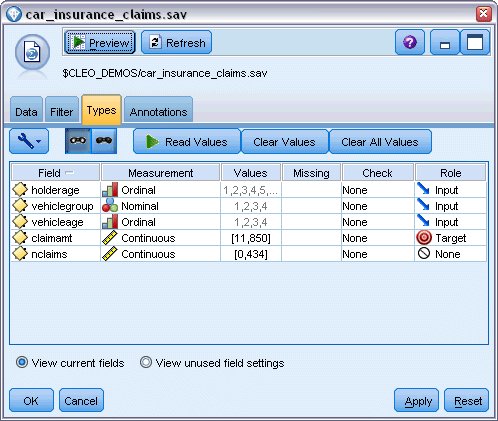
1. Add a Statistics File source node pointing to car\_insurance\_claims.sav in the Demos folder.

*Figure 1. Sample stream to predict car insurance claims*



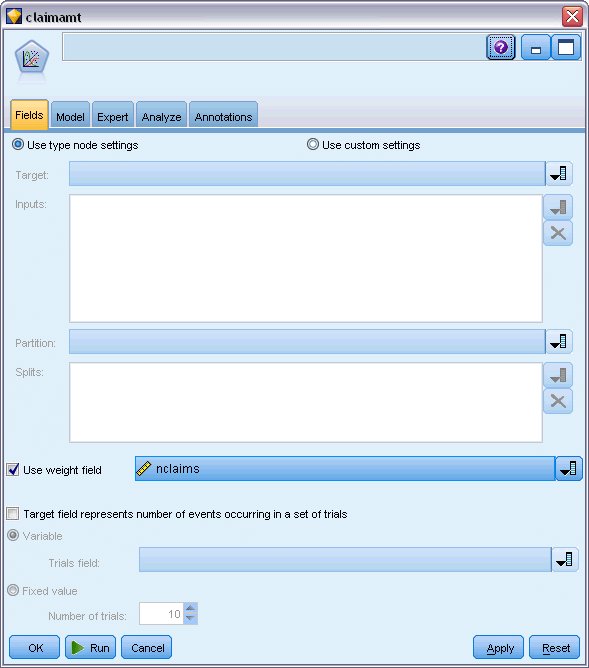
1. On the Types tab of the source node, set the role for the claimamt field to **Target**. All other fields should have their role set to **Input**.
2. Click **Read Values** to instantiate the data.

*Figure 2. Setting field role*



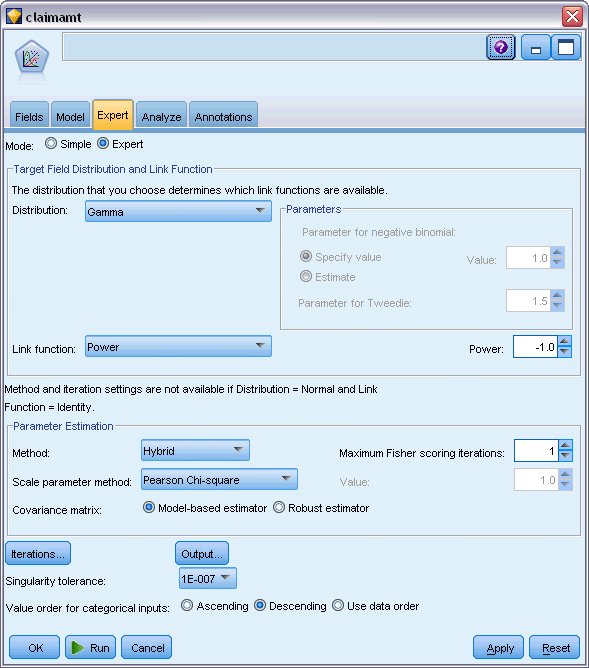
1. Attach a Genlin node to the source node; in the Genlin node, click the Fields tab.
2. Select nclaims as the scale weight field.

*Figure 3. Choosing field options*



1. Click the Expert tab and select **Expert** to activate the expert modeling options.

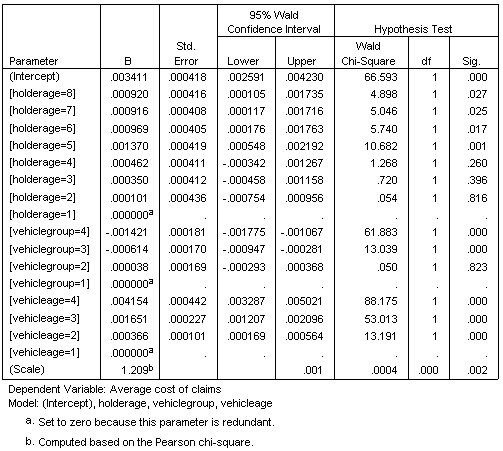
*Figure 4. Choosing expert options*



1. Select **Gamma** as the response distribution.
2. Select **Power** as the link function and type -1.0 as the exponent of the power function. This is an inverse link.
3. Select **Pearson chi-square** as the method for estimating the scale parameter. This is the method used by McCullagh and Nelder, so we follow it here in order to replicate their results.
4. Select **Descending** as the category order for factors. This indicates that the first category of each factor will be its reference category; the effect of this selection on the model is in the interpretation of parameter estimates.
5. Click **Run** to create the model nugget, which is added to the stream canvas, and also to the Models palette in the upper-right corner. To view the model details, right-click the model nugget and choose **Edit** or **Browse**, then select the Advanced tab.

# Parameter Estimates

*Figure 1. Parameter estimates*



The omnibus test and tests of model effects (not shown) indicate that the model outperforms the null model and that each of the main effects terms contribute to the model. The parameter estimates table shows the same values obtained by McCullagh and Nelder for the factor levels and the scale parameter.

# Summary

Using Generalized Linear Models, you have fit a gamma regression to the claims data. Note that while the canonical link function for the gamma distribution was used in this model, a log link will also give reasonable results. In general, it is difficult to impossible to directly compare models with different link functions; however, the log link is a special case of the power link where the exponent is 0, thus you can compare the deviances of a model with a log link and a model with a power link to determine which gives the better fit (see, for example, section 11.3 of McCullagh and Nelder).

Explanations of the mathematical foundations of the modeling methods used in IBM® SPSS® Modeler are listed in the *IBM SPSS Modeler* Algorithms Guide