

ANALYZING DATA IN SPSS 13.0 USING REGRESSION ANALYSIS

Tips before you begin:

- Make sure your data set is open before attempting to run any analyses.
- During analyses, right click on terms or buttons in the dialog boxes to learn about their functions.
- The **Help** button in the dialog boxes maybe clicked at any time during analyses for more information on that particular procedure.
- Click the **Reset** button to clear the dialog box and begin a fresh analysis.
- Click the **Cancel** button to exit that dialog box without saving changes.

Choose a Procedure:

- [Linear Regression](#)
- [Curve Estimation](#)
- [Binary Logistic Regression](#)
- [Multinomial Logistic Regression](#)
- [Ordinal Regression](#)
- [Probit Analysis](#)

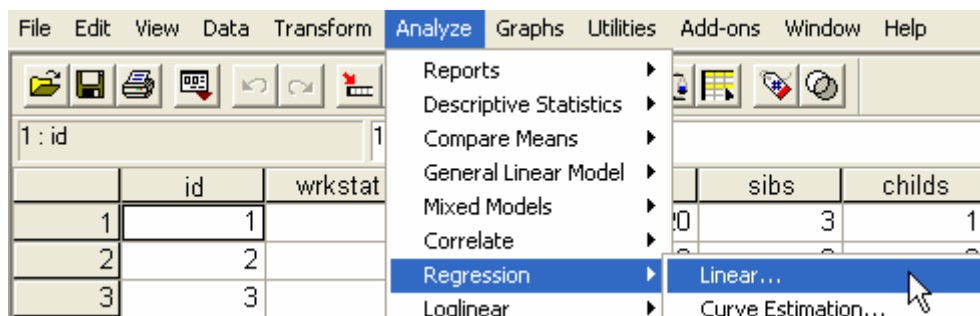
LINEAR REGRESSION

Linear Regression estimates the coefficients of the linear equation, involving one or more independent variables that best predict the value of the dependent variable.

Examples. A cost modeler wants to find the prospective cost for a new contract based on the data collected from previous contracts. Or, if the university authorities want to predict a student's grade on a freshman college calculus midterm based on his/her SAT score, then they may use linear regression.

Procedure

1. On the menu bar of the SPSS Data Editor window, click **Analyze > Regression > Linear...**



2. Select one or more variables that you want to test by clicking on the variable labels in the Linear Regression dialog box. To select multiple variables, hold down the Ctrl key and choose the variables you want. Click on the respective arrow buttons to add selected variables to the Dependent and Independent(s) windows. You may also specify the selection variable, case labels and WLS weight in this dialog box.

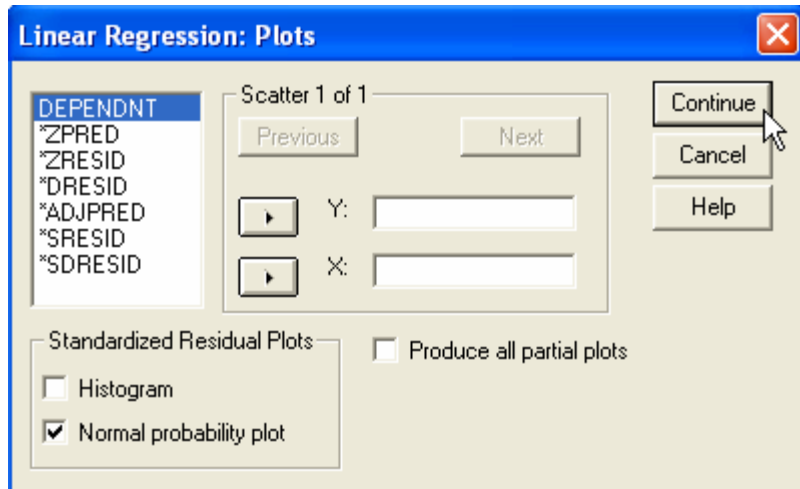
The screenshot shows the 'Linear Regression' dialog box. On the left is a 'List of Variables' containing various demographic and socioeconomic variables. In the center, there are several input fields: 'Dependent:', 'Independent(s):', 'Selection Variable:', 'Case Labels:', and 'WLS Weight:'. Each field has a right-pointing arrow button next to it. Below these fields is a 'Method:' drop-down menu set to 'Enter'. At the bottom are buttons for 'Statistics...', 'Plots...', 'Save...', and 'Options...'. On the right side, several callout ovals point to empty input fields with labels: 'Empty Dependent Window', 'Empty Independent(s) Window', 'Empty Selection Variable Window', 'Empty Case Labels Window', and 'Empty WLS Weight Window'. Another callout on the left points to the right arrow buttons with the text 'Right arrow buttons to add selected variable(s)'.

3. The Method drop-down box allows you to select the method of regression you want to run. The methods listed are Enter, Stepwise, Forward and Backward. By default, Enter method is selected.

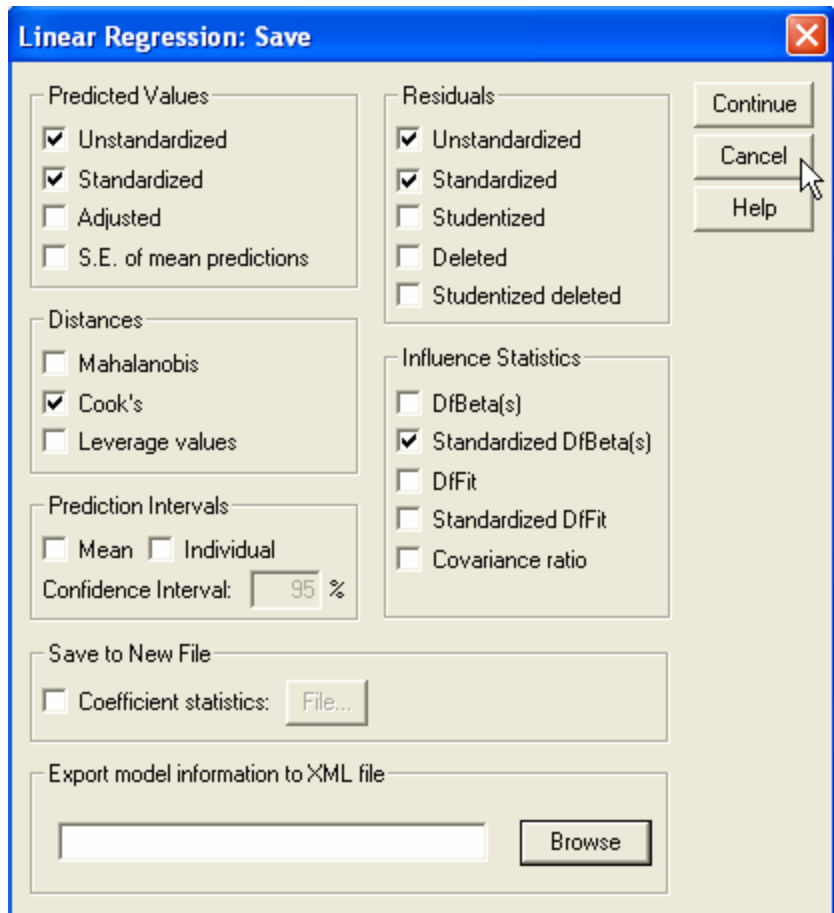
4. Click on the **Statistics...** button to select the regression coefficients, residuals and other options listed. Click on the **Continue** button.

The screenshot shows the 'Linear Regression: Statistics' dialog box. It is divided into two main sections: 'Regression Coefficients' and 'Residuals'. Under 'Regression Coefficients', there are checkboxes for 'Estimates' (checked), 'Confidence intervals', 'Covariance matrix', 'Model fit' (checked), 'R squared change', 'Descriptives', 'Part and partial correlations', and 'Collinearity diagnostics'. Under 'Residuals', there are checkboxes for 'Durbin-Watson' and 'Casewise diagnostics'. Below 'Casewise diagnostics', there are radio buttons for 'Outliers outside:' (selected) and 'All cases'. The 'Outliers outside:' option has a text box containing the number '3' followed by the text 'standard deviations'. On the right side, there are buttons for 'Continue', 'Cancel', and 'Help'.

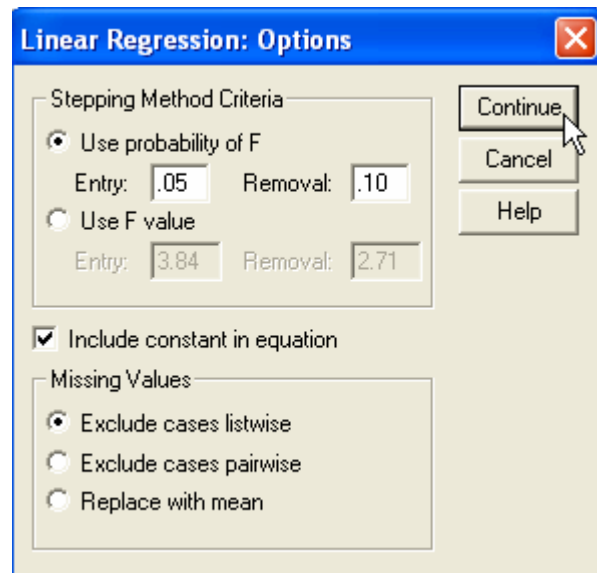
5. Click on the **Plots....** button to select the plots variables. Click on the **Continue** button.



6. Click on the **Save....** button to select Predicted Values, Residuals, Distances, Prediction Intervals and Influence Statistics. Select Save to New file if required. Click on the **Continue** button.



7. Click on the **Options....** button to select the stepping method criteria and to control the treatment of missing values. Click on the **Continue** button.



8. Click the **OK** button in the Linear Regression dialog box to run the analysis. The output will be displayed in a separate SPSS Viewer window.

[Back to Top](#)

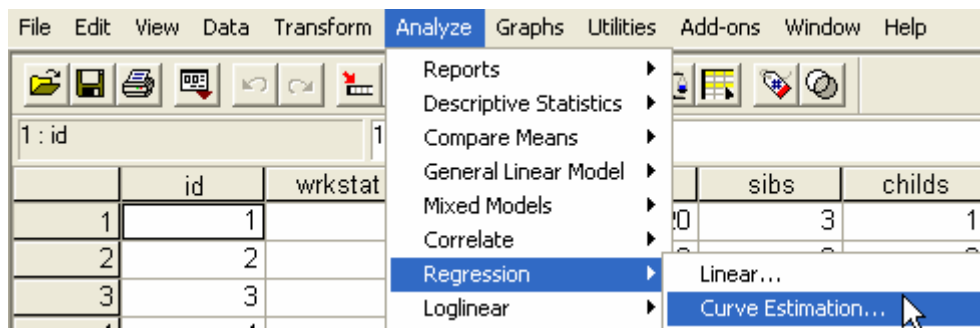
CURVE ESTIMATION

The Curve Estimation procedure produces curve estimation regression statistics and related plots for 11 different curve estimation regression models. A separate model is produced for each dependent variable. You can also save predicted values, residuals, and prediction intervals as new variables.

Example. A fire insurance company conducts a study to relate the amount of damage in serious residential fires to the distance between the closest fire station and the residence. A scatterplot reveals that the relationship between fire damage and distance to the fire station is linear.

Procedure

1. On the menu bar of the SPSS Data Editor window, click **Analyze > Regression > Curve Estimation...**



- Select one or more variables that you want to test by clicking on the variable labels in the Curve Estimation dialog box. To select multiple variables, hold down the Ctrl key and choose the variables you want. Click on the respective arrow buttons to add selected variables to the Dependent(s) and Independent Variable windows. You may also specify the case labels in this dialog box and select the model type.

The screenshot shows the 'Curve Estimation' dialog box. On the left, a list of variables is shown with right-pointing arrow buttons next to several of them. On the right, there are three empty text boxes labeled 'Dependent(s):', 'Independent Variable:', and 'Case Labels:'. Below these are buttons for 'OK', 'Paste', 'Reset', 'Cancel', and 'Help'. At the bottom, there are checkboxes for 'Include constant in equation' and 'Plot models', a 'Models' section with various model types (Linear, Quadratic, Compound, Growth, Logarithmic, Cubic, S, Exponential, Inverse, Power, Logistic), an 'Upper bound' field, and a 'Save...' button. Annotations include: 'List of Variables' pointing to the variable list; 'Right arrow buttons to add selected variable(s)' pointing to the arrow buttons; 'Empty Dependent(s) Window' pointing to the 'Dependent(s):' box; 'Empty Independent Variable Window' pointing to the 'Independent Variable:' box; and 'Empty Case Labels Window' pointing to the 'Case Labels:' box.

- Click on the **Save....** button, to save the predicted values, residuals and prediction intervals. Click on the **Continue** button.

The screenshot shows the 'Curve Estimation: Save' dialog box. It has two main sections: 'Save Variables' and 'Predict Cases'. In 'Save Variables', there are checkboxes for 'Predicted values', 'Residuals', and 'Prediction intervals', and a dropdown menu set to '95' for '% Confidence interval'. In 'Predict Cases', there are radio buttons for 'Predict from estimation period through last case' and 'Predict through:', with an 'Observation:' field below. At the bottom, there are buttons for 'Continue', 'Cancel', and 'Help'. The text 'The Estimation Period is: All cases' is displayed at the bottom left.

- Click the **OK** button in the Curve Estimation dialog box to run the analysis. The output will be displayed in a separate SPSS Viewer window.

[Back to Top](#)

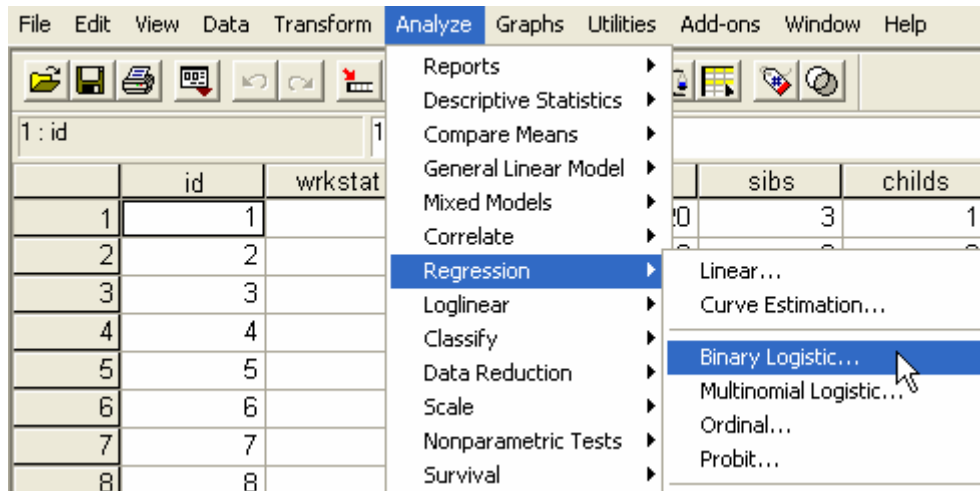
BINARY LOGISTIC REGRESSION

Logistic regression is useful for situations in which you want to be able to predict the presence or absence of a characteristic or outcome based on values of a set of predictor variables. It is similar to a linear regression model but is suited to models where the dependent variable is dichotomous. Logistic regression coefficients can be used to estimate odds ratios for each of the independent variables in the model. Logistic regression is applicable to a broader range of research situations than discriminant analysis.

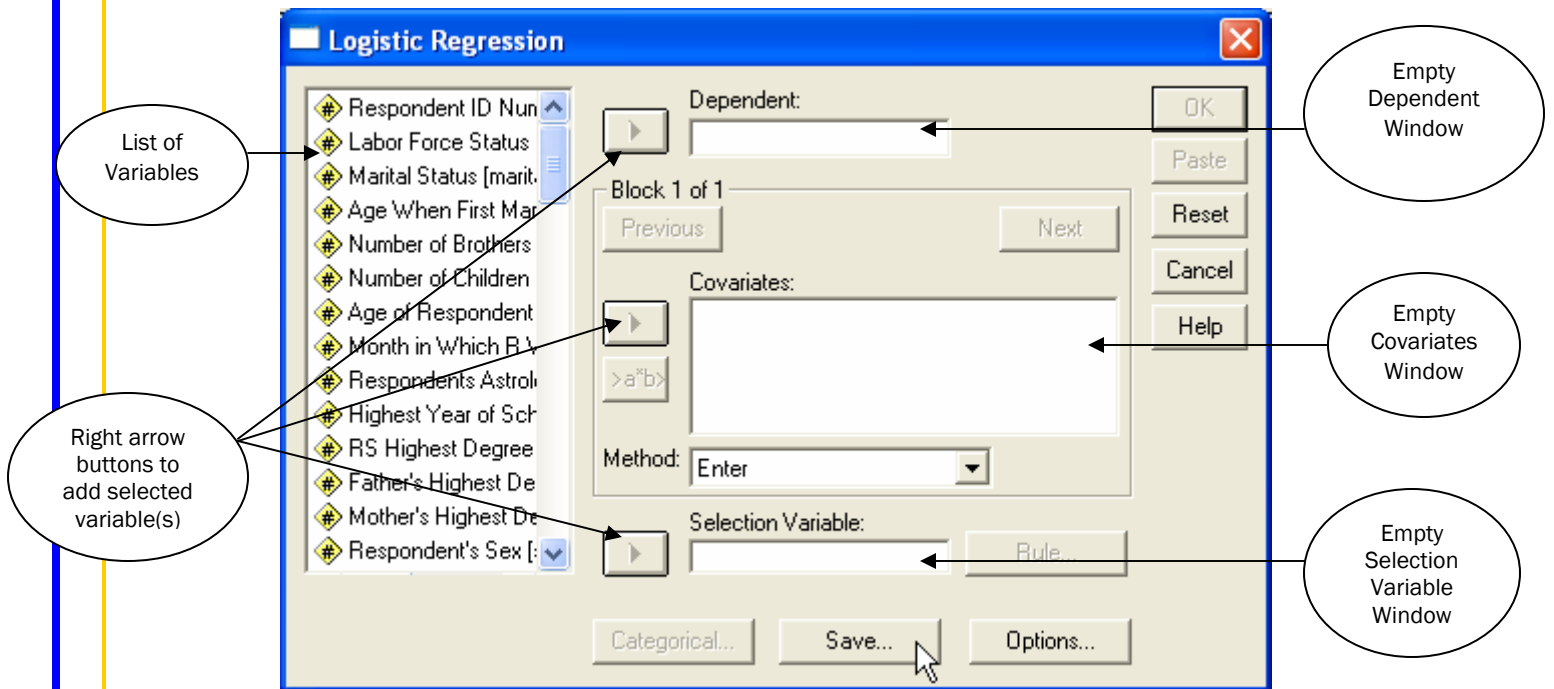
Example. What characteristics or lifestyles lead to overweight in women? For this purpose, the body perceptions and weigh management practices could be compared. The data could also compare the respondents' race and age to establish a relationship.

Procedure

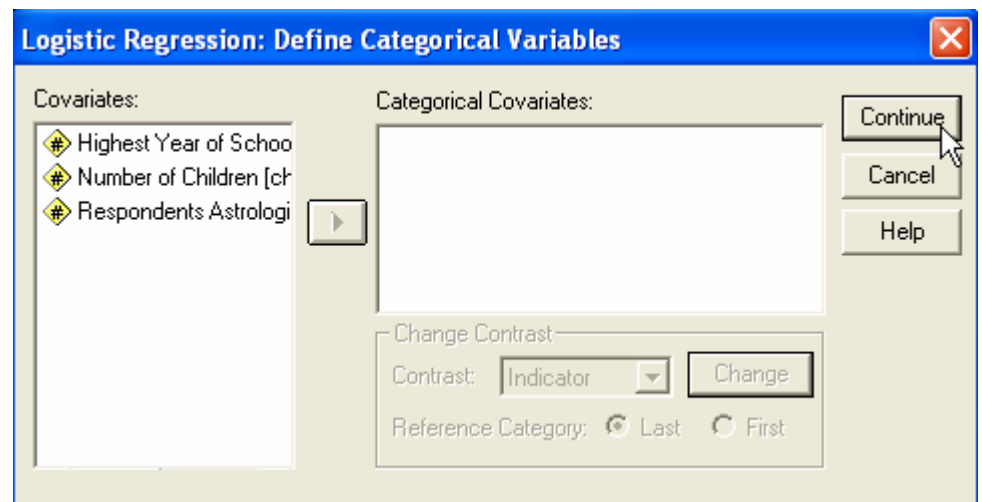
1. On the menu bar of the SPSS Data Editor window, click **Analyze > Regression > Binary Logistic...**



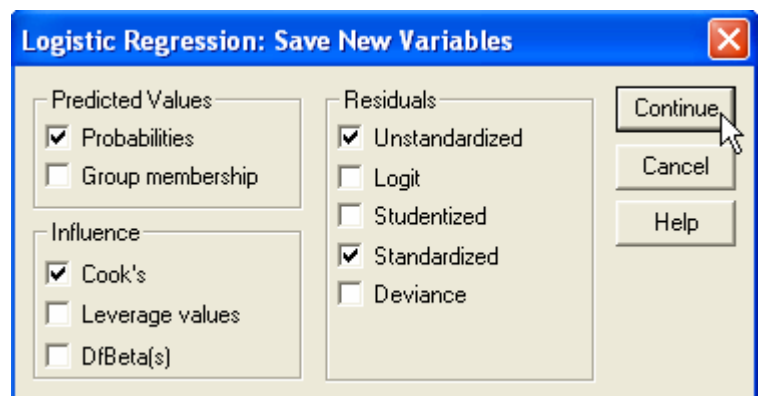
2. Select one or more variables that you want to test by clicking on the variable labels in the Logistic Regression dialog box. To select multiple variables, hold down the Ctrl key and choose the variables you want. Click on the respective arrow buttons to add selected variables to the Dependent and Covariates windows. You may also specify the selection variable in this dialog box.



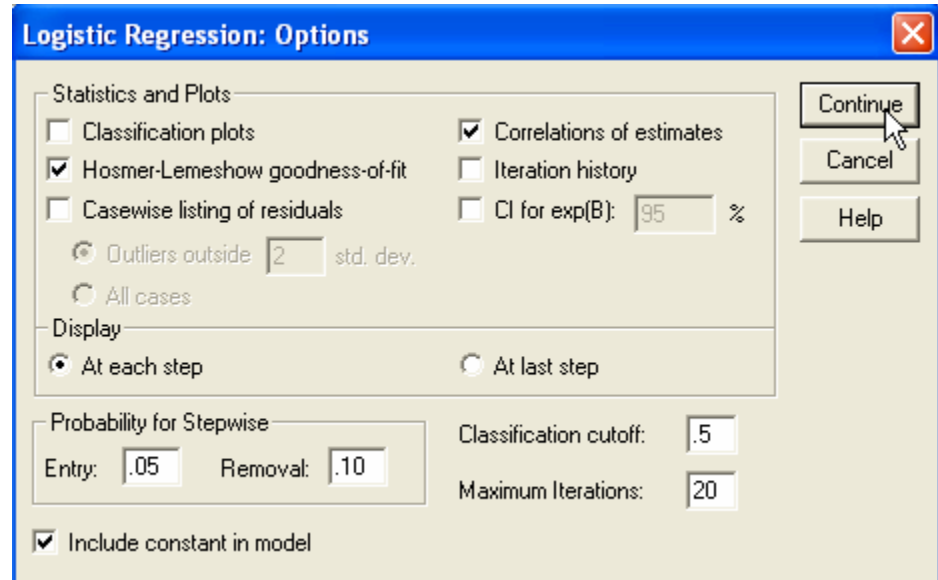
3. Click on the **Categorical ...** button to select covariate as categorical covariates. Click on the **Continue** button.



4. Click on the **Save...** button to select and save the appropriate predicted values, residuals and Influence as new variables. Click on the **Continue** button.



5. Click on the **Options....** button to select the statistics and plots, display type and probability for stepwise. Click on the **Continue** button.



6. Click the **OK** button in the main Logistic Regression window. The output will be displayed in a new SPSS viewer window.

[Back to Top](#)

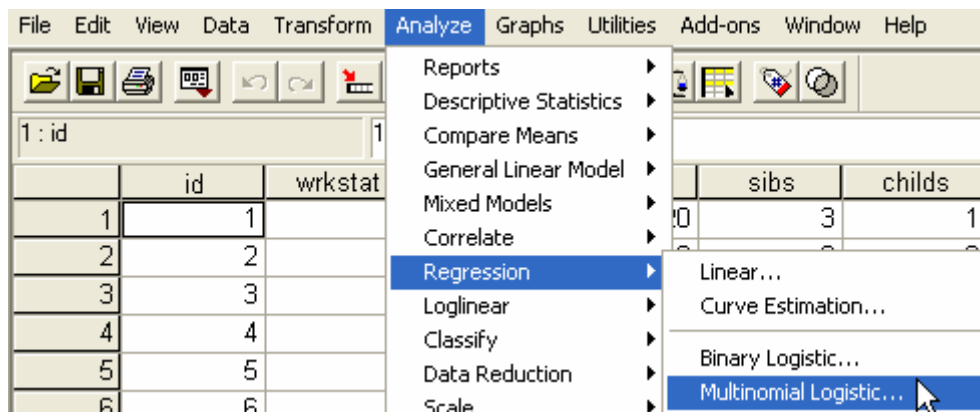
MULTINOMIAL LOGISTIC REGRESSION

Multinomial Logistic Regression is useful for situations in which you want to be able to classify subjects based on values of a set of predictor variables. This type of regression is similar to logistic regression, but it is more general because the dependent variable is not restricted to two categories.

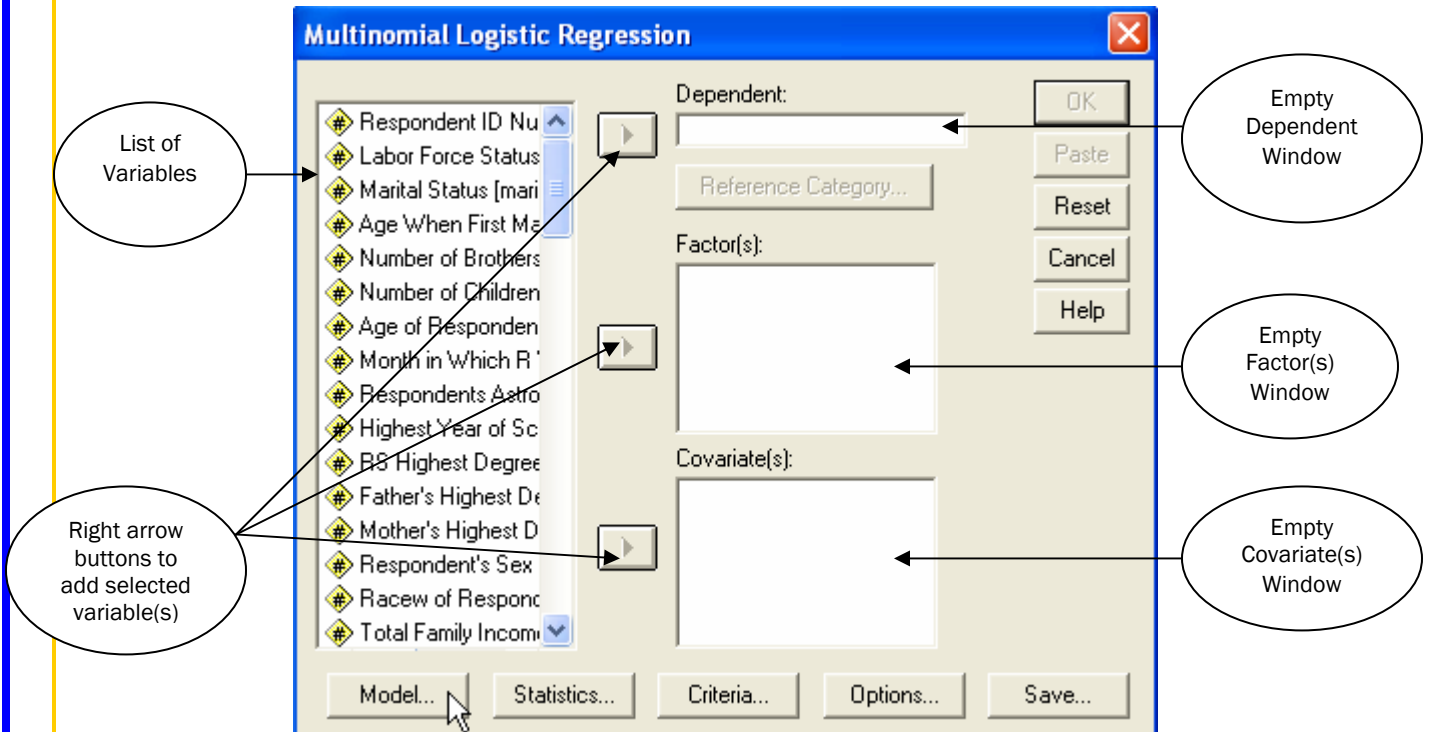
Example. In order to market films more effectively, movie studios want to predict what type of film a moviegoer is likely to see. Using a Multinomial Logistic Regression, the studio can determine the strength of influence a person's age, gender, and dating status has upon the type of film they prefer.

Procedure

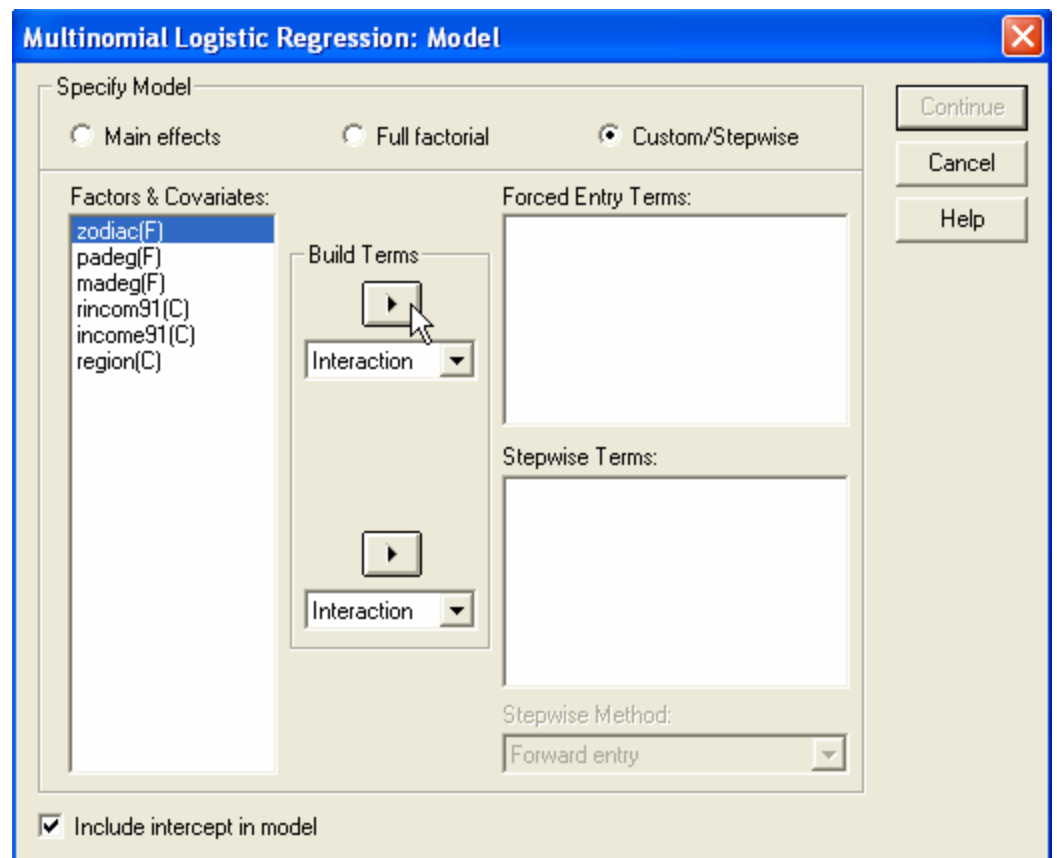
1. On the menu bar of the SPSS Data Editor window, click **Analyze > Regression > Multinomial Logistic...**



- Select one or more variables that you want to test by clicking on the variable labels in the Multinomial Logistic Regression dialog box. To select multiple variables, hold down the Ctrl key and choose the variables you want. Click on the respective arrow buttons to add selected variables to the Dependent, Factor(s) and Covariate(s) windows.



- Click on the **Model...** button to specify model type and select appropriate factors and covariates as forced entry terms and stepwise terms. Click on the **Continue** button.



4. Click on the **Statistics** button to select the appropriate options given under Model, Parameters and Define Subpopulations. The Case Processing Summary is checked by default. Click on the **Continue** button.

Multinomial Logistic Regression: Statistics

Case processing summary

Model

Pseudo R-square Cell probabilities

Step summary Classification table

Model fitting information Goodness-of-fit

Parameters

Estimates Confidence Interval (%): 95

Likelihood ratio tests

Asymptotic correlations

Asymptotic covariances

Define Subpopulations

Covariate patterns defined by factors and covariates

Covariate patterns defined by variable list below

Subpopulations:

zodiac(F)
padeg(F)
madeg(F)
rincom91(C)
income91(C)
region(C)

Continue
Cancel
Help

5. Click on the **Criteria** button to select the iterations. Click on the **Continue** button.

Multinomial Logistic Regression: Convergence Criteria

Iterations

Maximum iterations: 100

Maximum step-halving: 5

Log-likelihood convergence: 0

Parameter convergence: 0.000001

Print iteration history for every 1 step(s)

Check separation of data points from iteration 20 forward

Delta: 0 Singularity tolerance: 0.00000001

Continue
Cancel
Help

6. Click on the **Options** button to select the Dispersion Scale and Stepwise Options. Click on the **Continue** button.

7. Click on the **Save** button to save variables to the working file and export model information to an external file. Click on the **Continue** button.

8. Click the **OK** button in the Multinomial Logistic Regression dialog box to run the analysis. The output will be displayed in a separate SPSS Viewer window.

[Back to Top](#)

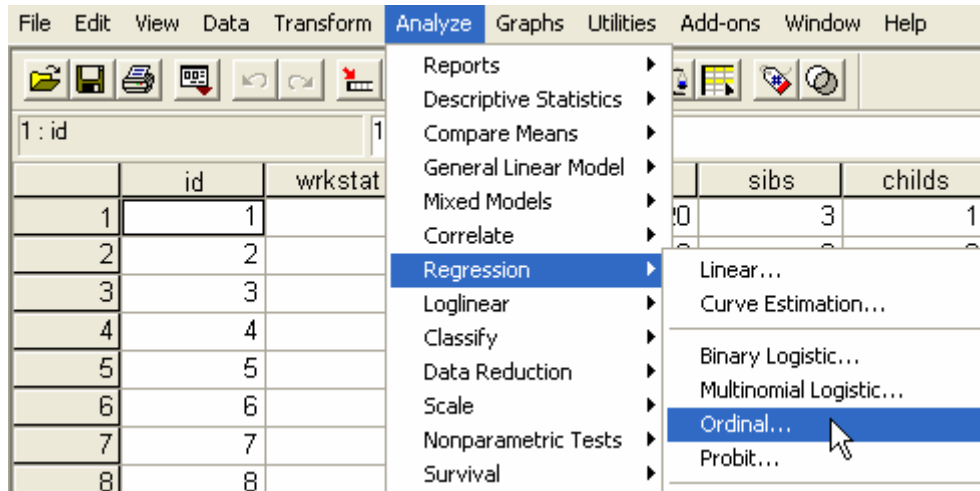
ORDINAL REGRESSION

Ordinal Regression allows you to model the dependence of a polytomous ordinal response on a set of predictors, which can be factors or covariates.

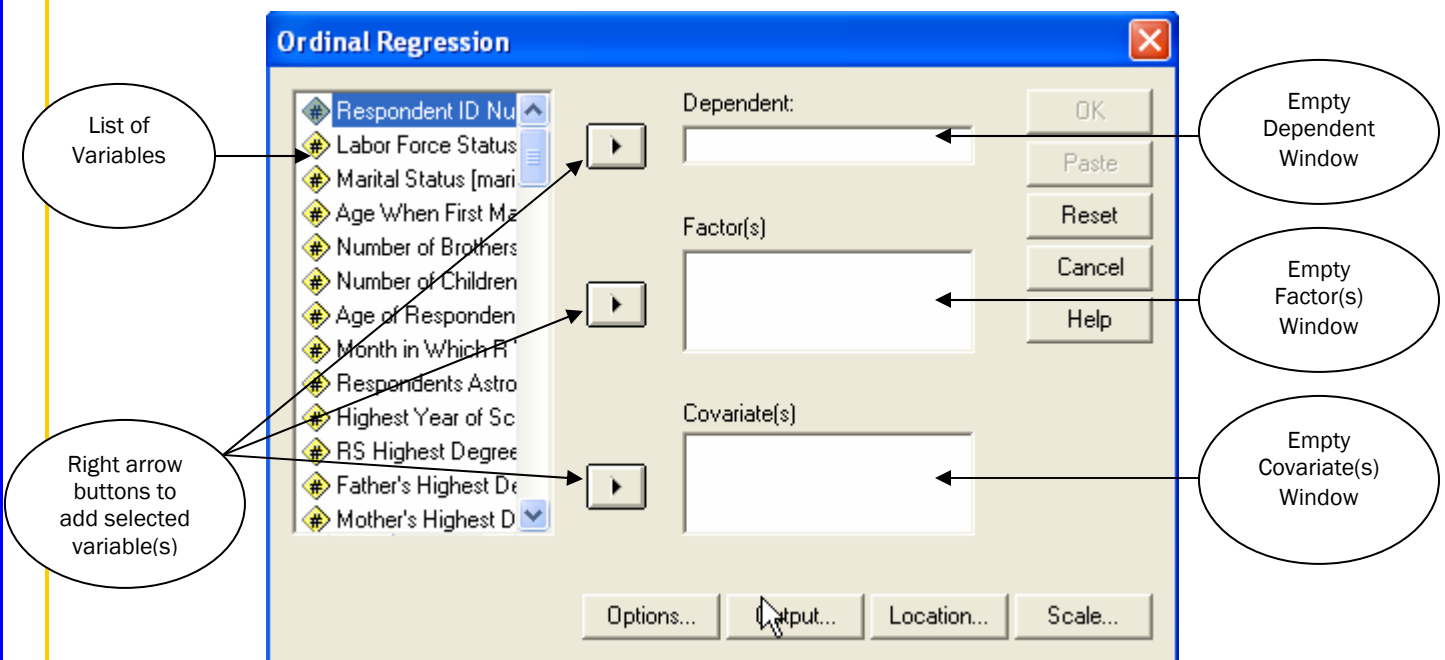
Example. Ordinal Regression could be used to study the effects on taste of various cheese additives. The possible effects on taste could be classified from strong dislike to an excellent taste.

Procedure

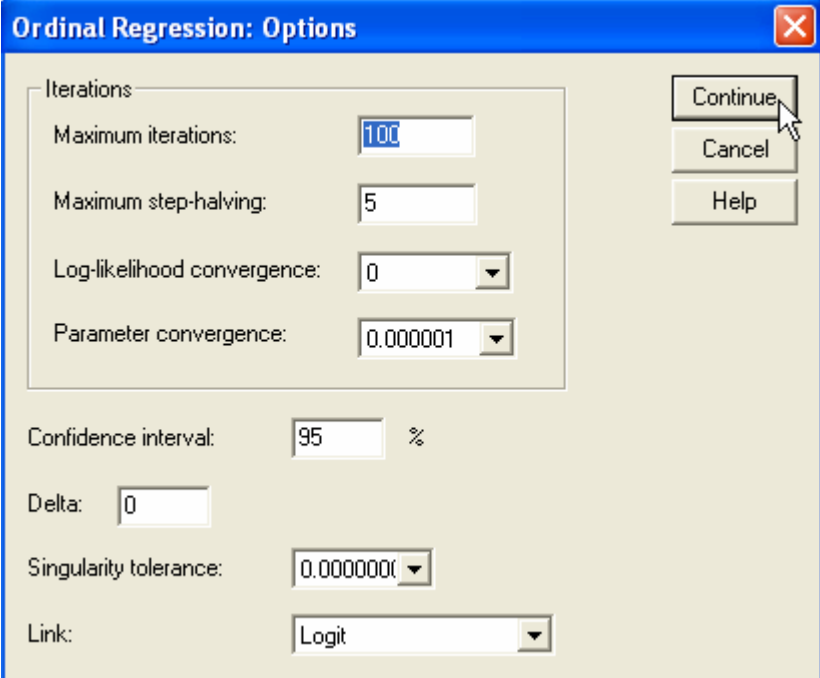
1. On the menu bar of the SPSS Data Editor window, click **Analyze > Regression > Ordinal...**



2. Select one or more variables that you want to analyze by clicking on the variable labels in the Ordinal Regression dialog box. Optionally, select one or more variables for the Factor(s) window and the Covariates window. To select multiple variables, hold down the Ctrl key and choose the variables you want. Click on the respective arrow buttons to add selected variables to the Dependent, Factor(s), and Covariate(s) windows.



3. Click on the **Options...** button to select the appropriate iterations. Click on the **Continue** button.



Ordinal Regression: Options

Iterations

Maximum iterations:

Maximum step-halving:

Log-likelihood convergence:

Parameter convergence:

Confidence interval: %

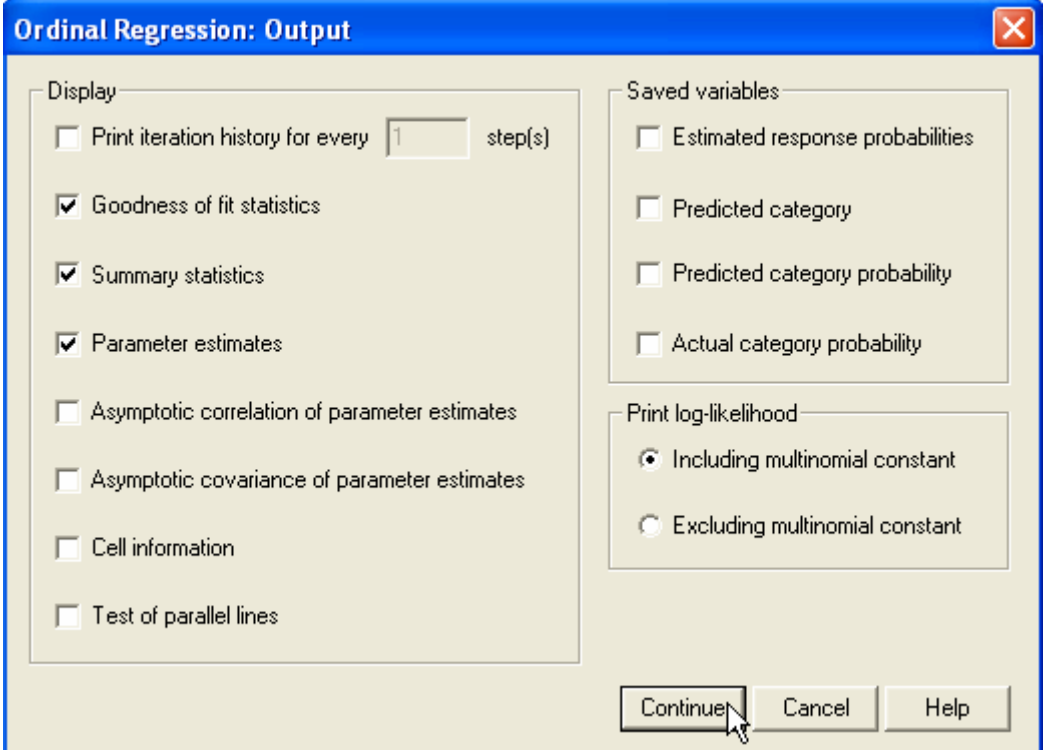
Delta:

Singularity tolerance:

Link:

Buttons: Continue, Cancel, Help

4. Click on the **Output...** button to select the appropriate options under Display, Saved Variables and Print log-likelihood. Click on the **Continue** button.



Ordinal Regression: Output

Display

Print iteration history for every step(s)

Goodness of fit statistics

Summary statistics

Parameter estimates

Asymptotic correlation of parameter estimates

Asymptotic covariance of parameter estimates

Cell information

Test of parallel lines

Saved variables

Estimated response probabilities

Predicted category

Predicted category probability

Actual category probability

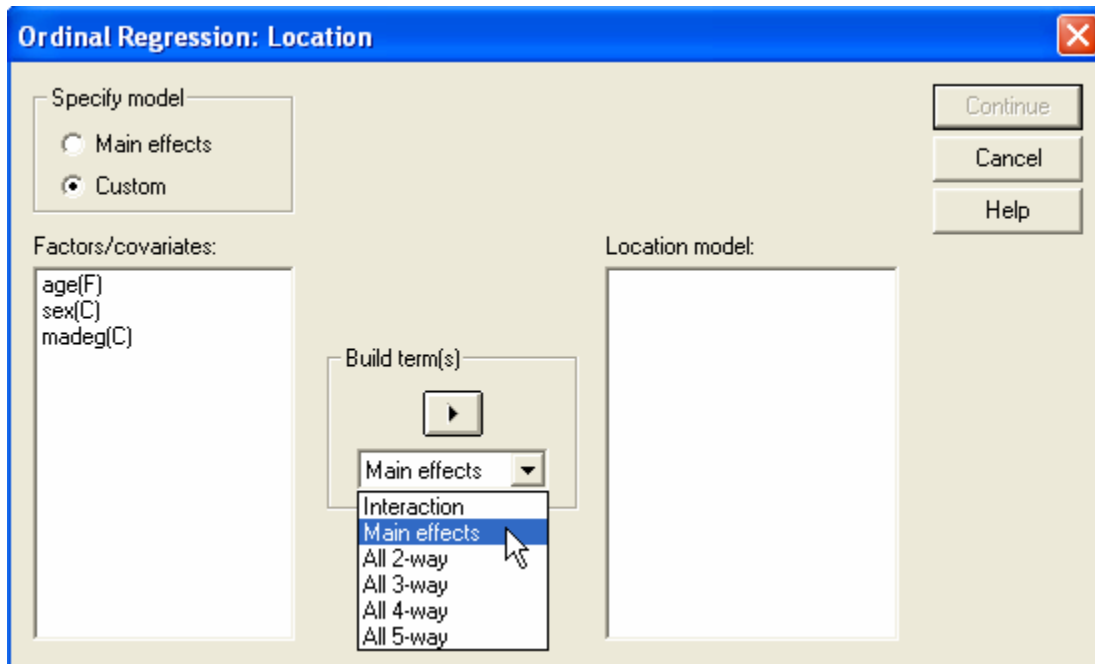
Print log-likelihood

Including multinomial constant

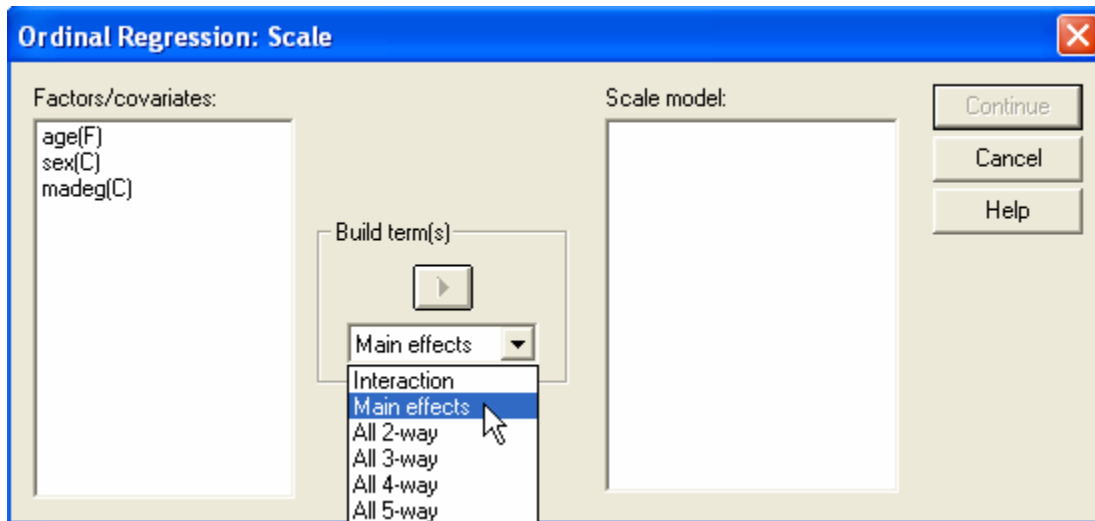
Excluding multinomial constant

Buttons: Continue, Cancel, Help

- Click on the **Location ...** button. Specify the model and select the factors/covariates onto a new Location model. Click on the **Continue** button.



- Click on the **Scale ...** button to select the factors/covariates on a Scale model. Click on the **Continue** button.



- Click on the **OK** button in the Ordinal Regression dialog box to run the analysis. The output will be displayed in a new SPSS Viewer window.

[Back to Top](#)

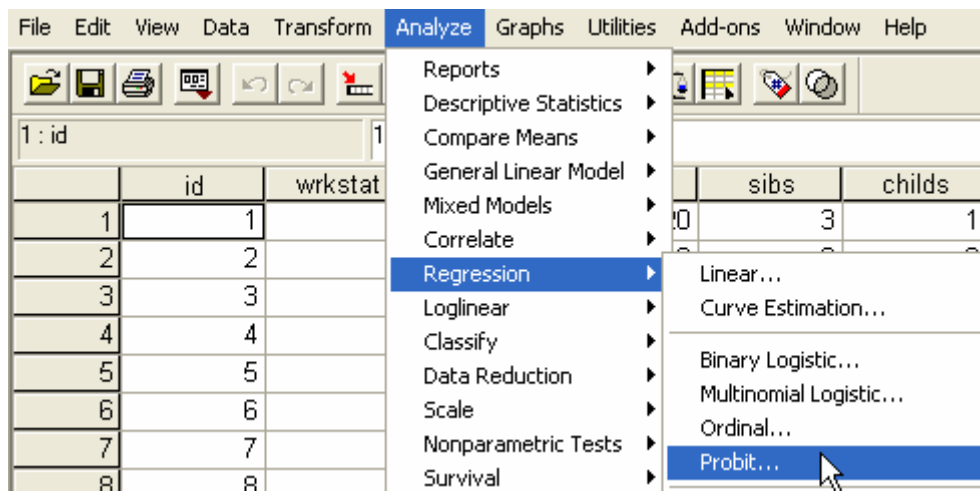
PROBIT ANALYSIS

This procedure measures the relationship between the strength of a stimulus and the proportion of cases exhibiting a certain response to the stimulus. It is useful for situations where you have a dichotomous output that is thought to be influenced or caused by levels of some independent variable(s) and is particularly well suited to experimental data. This procedure will allow you to estimate the strength of a stimulus required to induce a certain proportion of responses, such as the median effective dose.

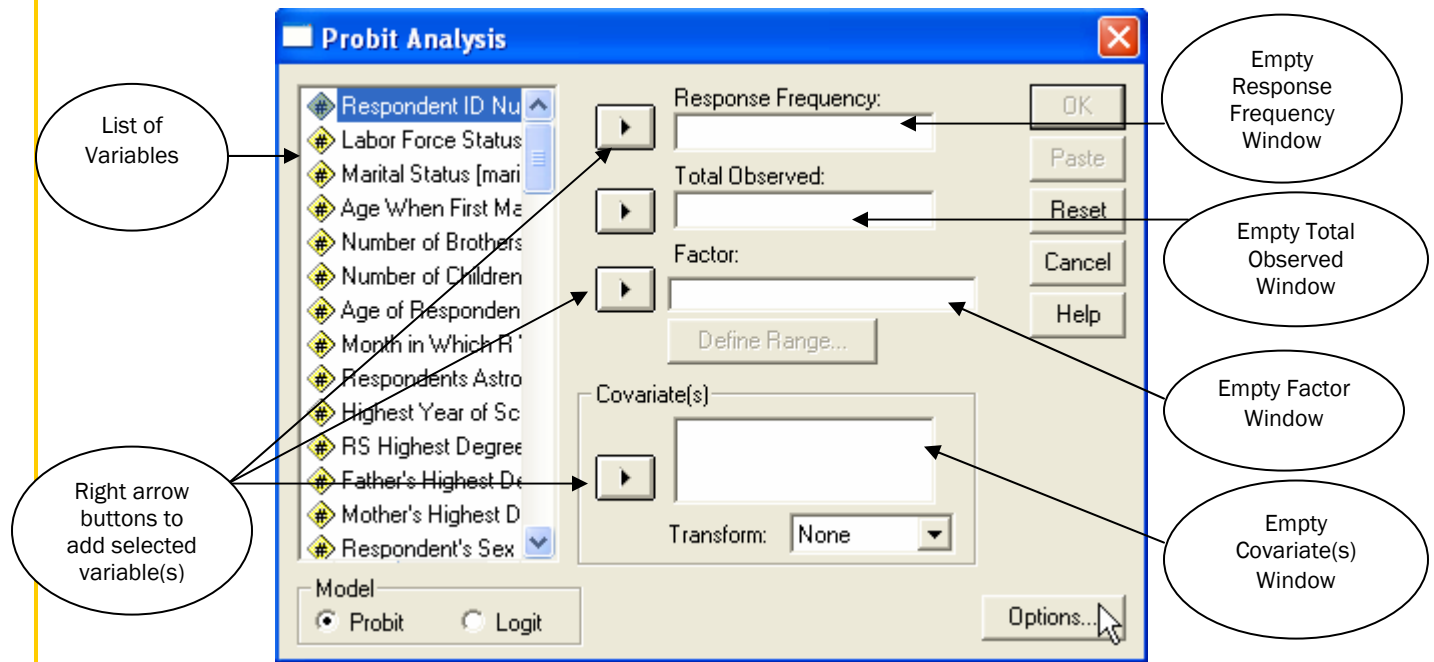
Example. How effective is a new pesticide at killing ants, and what is an appropriate concentration to use? You might perform an experiment in which you expose samples of ants to different concentrations of the pesticide and then record the number of ants killed and the number of ants exposed. Applying probit analysis to these data, you can determine the strength of the relationship between concentration and killing, and you can determine what the appropriate concentration of pesticide would be if you wanted to be sure to kill, say, 95% of exposed ants.

Procedure

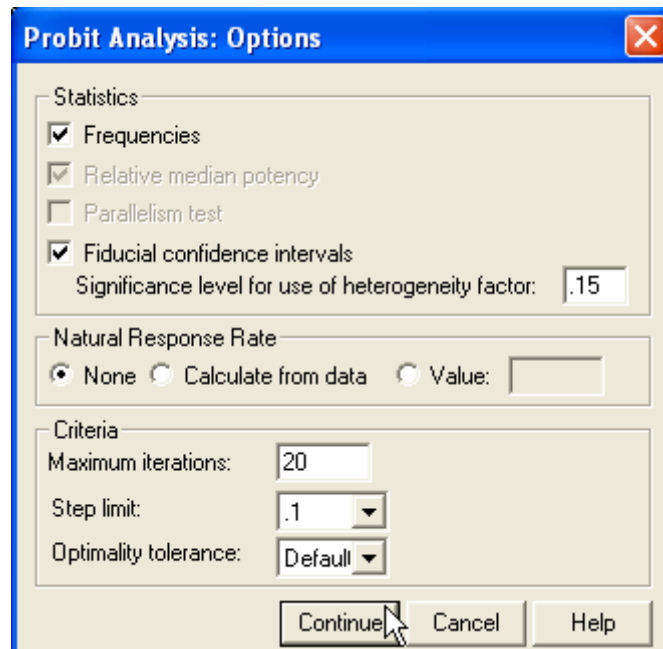
1. On the menu bar of the SPSS Data Editor window, click **Analyze > Regression > Probit...**



2. Select one or more variables that you want to analyze by clicking on the variable labels in the Probit Analysis dialog box. Select a single variable in the Response Frequency, Total Observed and Factor window. Select one or more variables for the Covariates window. To select multiple variables, hold down the Ctrl key and choose the variables you want. The model is checked Probit by default. Click on the respective arrow buttons to add selected variables to the respective windows.



3. Click on the **Options...** button to select the Statistics, Natural Response Rate and the criteria for iterations. Click on the **Continue** button.



4. Click the **OK** button in the Probit Analysis dialog box to run the analysis. The output will be displayed in a new SPSS Viewer window.

[Back to Top](#)