The formulas for *C* and *Cbar* are

$$C_{gj} = \frac{\chi_j^2 h_{gj}}{\left(1 - h_{gj}\right)^2}, \quad j = 1,2,\cdots,J$$

$$\overline{C}_{gj} = \frac{\chi_j^2 h_{gj}}{\left(1 - h_{gj}\right)}, \quad j = 1,2,\cdots,J$$

Note that this formula matches Pregibon (1981) in the two-outcome case. In the multiple-outcome case, the two-outcome formula is applied to each outcome.

Residual Diagnostics Report

Residual Diagnostics Report For Remiss = 1

Row	Actual Remiss	Hat Diagonal		Deviance Change (DFDev)		Chi-Square Change (DFChi2)	
1	1	0.20631	||||.....	0.40098	|.....	0.24178	|.....
2	1	0.05654	|.....	0.64257	|.....	0.38576	|.....
3	0	0.26518	||||||.....	0.50425	|.....	0.31795	|.....
4	0	0.23855	|||||.....	2.81743	||||||.....	2.06910	||.....
5	1	0.12192	||.....	0.39260	|.....	0.22789	|.....
6	0	0.16277	|||.....	1.25574	||.....	0.88752	|.....
7*	1	0.04169	|.....	1.48639	|||.....	1.10084	|.....
8*	0	0.28695	||||||.....	4.06243	|||||||.....	4.33640	|||.....
9	0	0.14925	|||.....	0.74150	|.....	0.46899	|.....
10	0	0.04227	|.....	0.00642	|.....	0.00328	|.....
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.

This report gives statistics that help detect observations that have not been fitted well by the model.

Row

This is the row from the database. Rows that are starred are misclassified.

Actual Y

This is the outcome to which this row belongs (if known).

Hat Diagonal

The diagonal elements of the hat matrix can be used to detect points that are extreme in the independent variable space. They are discussed in more detail in the Residual Report.

Deviance Change (DFDev) and Chi-Square Change (DFChi2)

DFDEV and *DFCHI2* are statistics that measure the change in deviance and in Pearson's chi-square, respectively, that occurs when an observation is deleted from the dataset. Large values of these statistics indicate observations that have not been fitted well.

The formulas for these statistics are

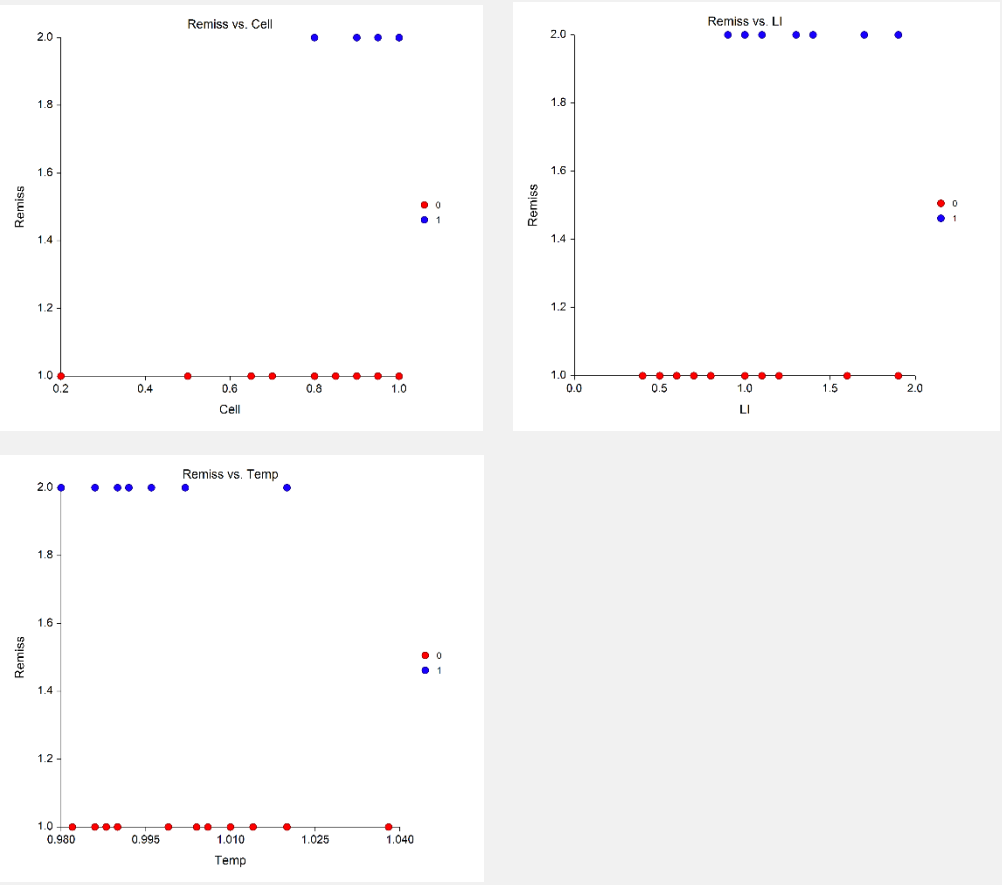
$$DFDEV_{gj} = d_j^2 + \overline{C}_{gj}, \quad j = 1,2,\cdots,J$$

$$DFCHI2_{gj} = \frac{\overline{C}_{gj}}{h_{gj}}, \quad j = 1,2,\cdots,J$$

Note that this formula matches Pregibon (1981) in the two-group case. In the multiple-group case, the two-group formula is applied to each group.

Y versus X Plots

Y vs X's Plots



This section shows scatter plots with the dependent variable on the vertical axis and each of the independent variables on the horizontal axis. The plot is useful for finding typos, outliers, and other anomalies in that data.

Vertical Axis

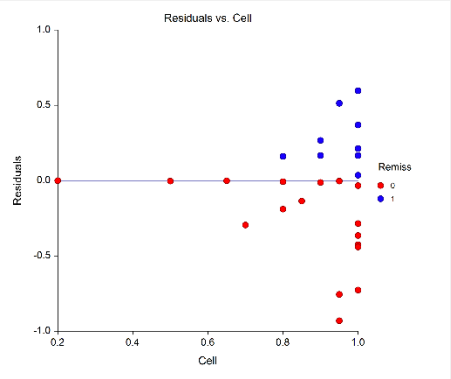
The categories of the dependent variable are shown on the vertical axis. Each category is assigned a whole number, beginning with the number one. The numbers are assigned in sorted order. Thus, if your dependent variable has values A, B, and C, it would be plotted on a numeric scale ranging from about 0.8 to 3.2. The groups would be plotted as the numbers 1, 2, and 3.

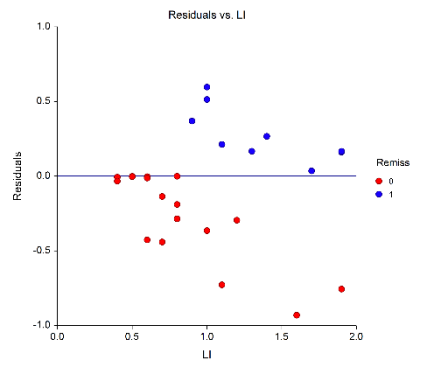
Horizontal Axis

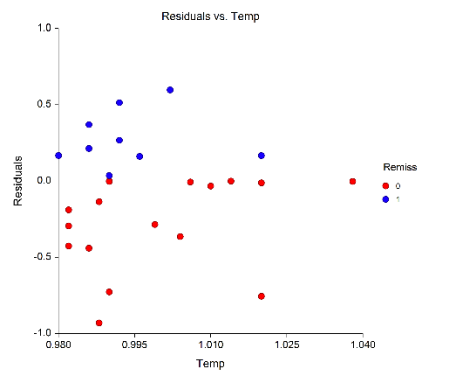
The independent variables are shown on the horizontal axis. When the independent variable is categorical, binary variables are generated for each of the categories and a separate scatter plot is generated for each binary variable.

Simple Residuals versus X Plots

Simple Residuals vs X's Plots







This section shows scatter plots with the simple residuals on the vertical axis and each of the independent variables on the horizontal axis. The plots are useful for finding outliers and other anomalies in the data.

Vertical Axis

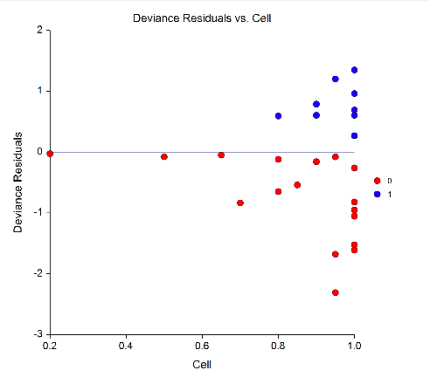
The residuals are displayed on the vertical axis. Note that the *G* residuals for each row corresponding to the simple residuals are displayed. Thus, if you have *N* rows, you will have *GN* points displayed on the plot.

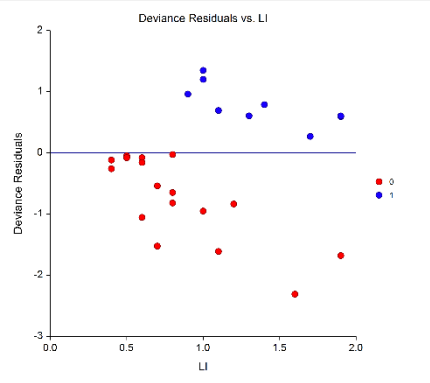
Horizontal Axis

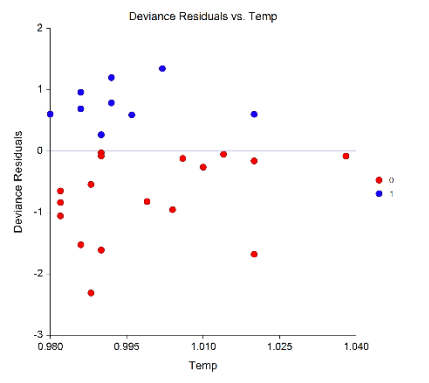
The independent variables are shown on the horizontal axis. When the independent variable is categorical, binary variables are generated for each of the categories and a separate scatter plot is generated for each binary variable.

Deviance Residuals versus X Plots

Deviance Residuals vs X's Plots







This section shows scatter plots with the deviance residuals on the vertical axis and each of the independent variables on the horizontal axis. The plots are useful for finding outliers and other anomalies in the data.

Vertical Axis

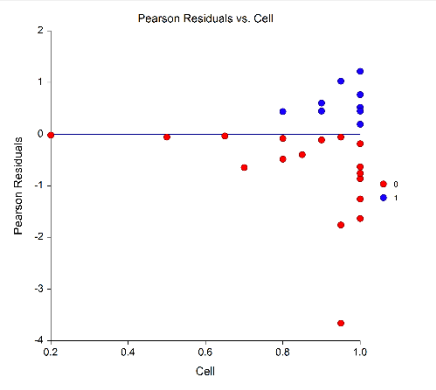
The deviance residuals are displayed on the vertical axis.

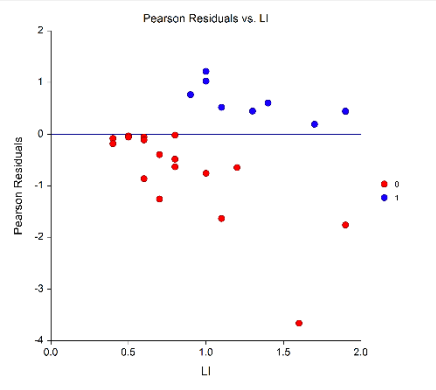
Horizontal Axis

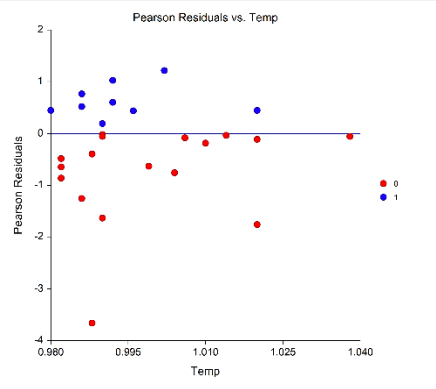
The independent variables are shown on the horizontal axis. When the independent variable is categorical, binary variables are generated for each of the categories and a separate scatter plot is generated for each binary variable.

Pearson Residuals versus X Plots

Pearson Residuals vs X's Plots







This section shows scatter plots with the Pearson residuals on the vertical axis and each of the independent variables on the horizontal axis. The plots are useful for finding outliers and other anomalies in the data.

Vertical Axis

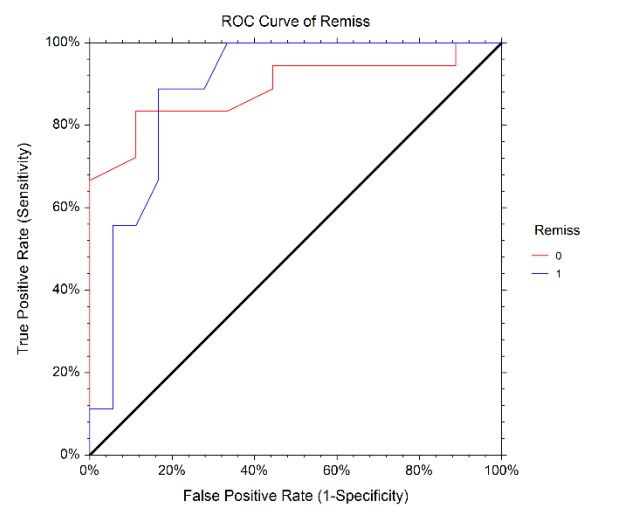
The Pearson residuals are displayed on the vertical axis.

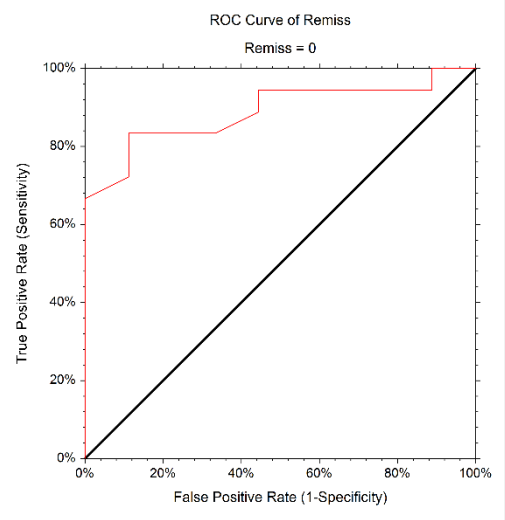
Horizontal Axis

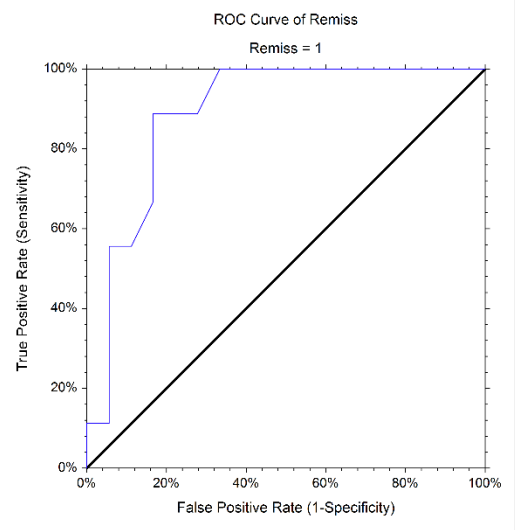
The independent variables are shown on the horizontal axis. When the independent variable is categorical, binary variables are generated for each of the categories and a separate scatter plot is generated for each binary variable.

ROC Curves - Combined and Separate

ROC Curves (Combined and Separate)







This section displays the ROC curves that can be used to help you find the best cutoff points to use for classification. The cutoff point nearest the top-left corner of the plot is the optimum cutoff. You will have to refer to the ROC Report to determine the exact value of the cutoff.

Vertical Axis

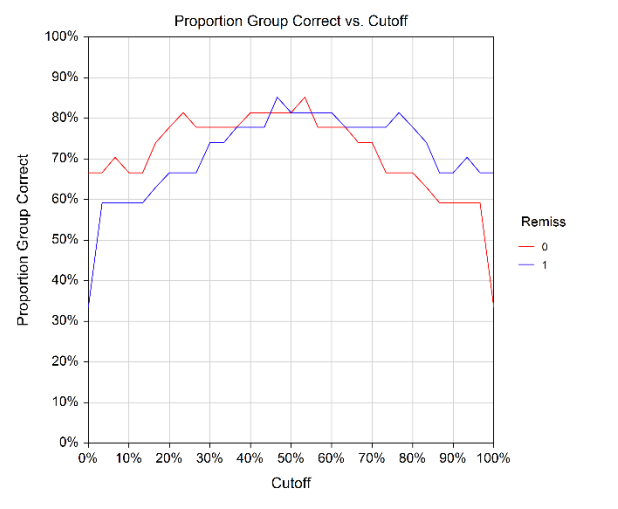
The sensitivity is displayed on the vertical axis.

Horizontal Axis

One minus the specificity is displayed on the horizontal axis.

Prob Correct versus Cutoff Plot

Prob Correct vs Cutoff Plot



This section displays a plot that shows the proportion correct versus the cutoff. It is useful to help determine the cutoff point used in classification. This plot may be difficult to use with three or more categories because of the ambiguity in the plot.

Vertical Axis

The proportion correctly classified for various cutoff values are displayed on the vertical axis.

Horizontal Axis

The cutoff values are displayed on the horizontal axis. These cutoff values are in terms of the estimated outcome-membership probabilities. Thus, a cutoff of 0.4 means that any rows with a outcome-membership probability of 0.4 or more are classified into this outcome.