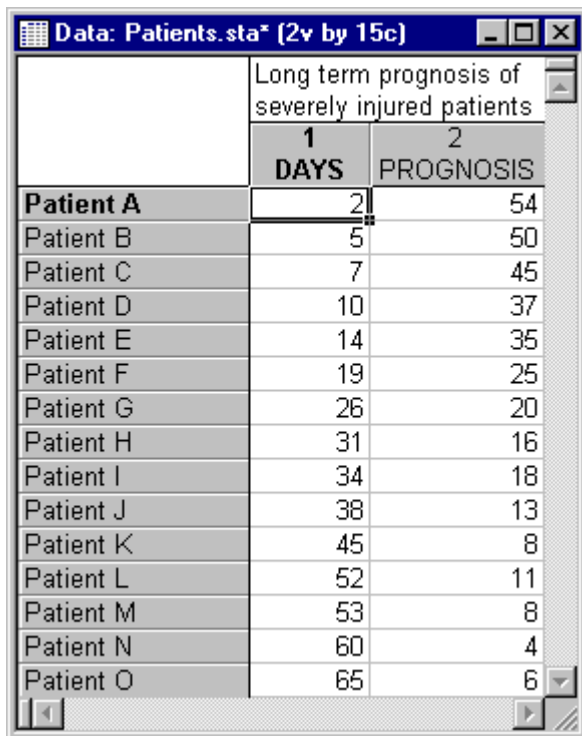


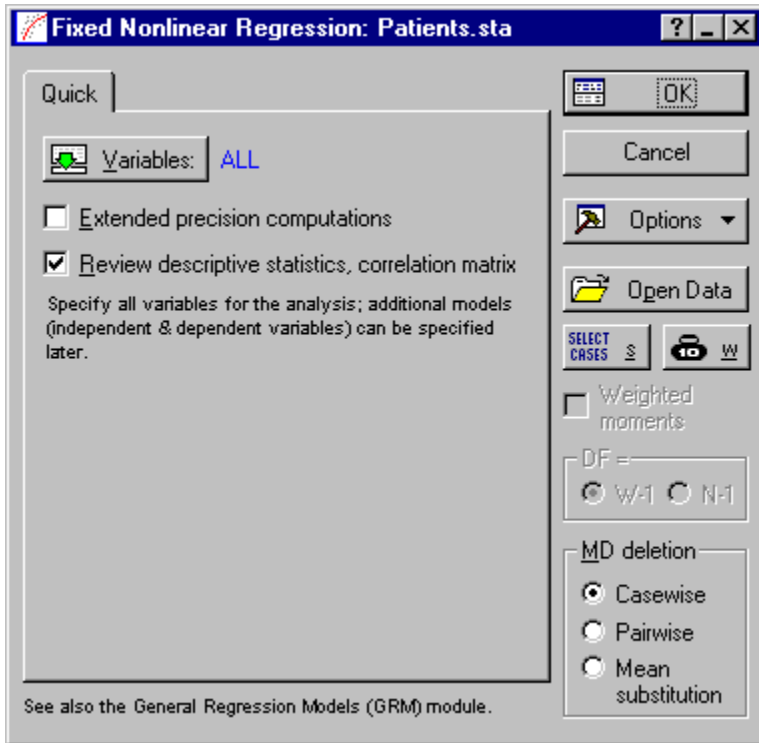
PREDICTING RECOVERY FROM INJURY (PATIENTS.STA)

Predicting Recovery from Injury. This example is based on a data set reported in Neter, Wasserman, and Kutner (1985, page 649). Suppose a hospital administrator wanted to explore the relationship between the chances for long-term recovery of severely injured patients and the number of days spent in the hospital. The data file Patients.sta contains data for 15 patients; specifically, the file contains information on the number of days that each patient was hospitalized (in the variable Days) and an index of the prognosis for long-term recovery for each patient (in the variable Prognosis; larger values reflect a better prognosis). Open this data file via the File - Open Examples menu; it is in the Datasets folder.

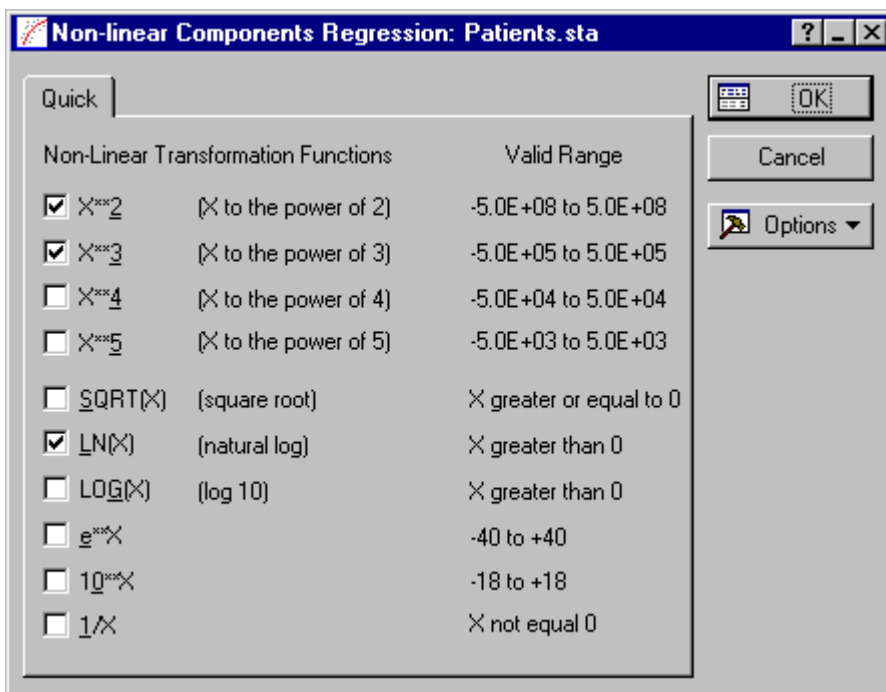


	Long term prognosis of severely injured patients	
	1	2
	DAYS	PROGNOSIS
Patient A	2	54
Patient B	5	50
Patient C	7	45
Patient D	10	37
Patient E	14	35
Patient F	19	25
Patient G	26	20
Patient H	31	16
Patient I	34	18
Patient J	38	13
Patient K	45	8
Patient L	52	11
Patient M	53	8
Patient N	60	4
Patient O	65	6

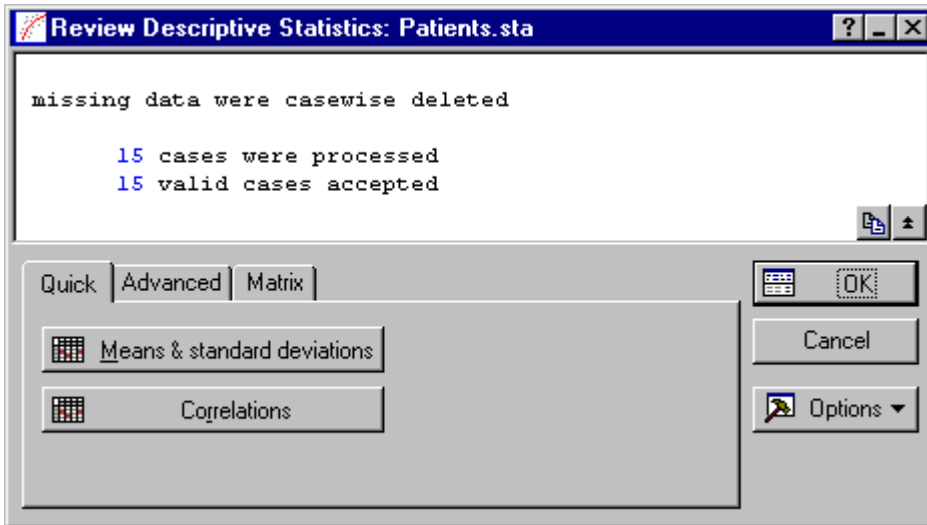
Specifying the analysis. Select Fixed Nonlinear Regression from the Statistics - Advanced Linear/Nonlinear Models menu to display the Fixed Nonlinear Regression Startup Panel. Click the Variables button to display the standard variable selection dialog. Here, select both the DAYS and PROGNOSIS variables for use in the model and then click the OK button. Note that at this point in the analysis it is not necessary to specify which variables will be the dependent or independent variables in the model. Before leaving the Startup Panel, select the Review descriptive statistics, correlation matrix check box. Selecting this option will provide opportunities to review statistics and correlations later in the analysis.



Now click the OK button to display the Non-linear Components Regression dialog (see below). On this dialog you can select up to 10 transformations to be applied to each of the designated variables. Note that for the selected transformation to be successful for each transformed case, the data must be within the range specified as valid for the transformation; nonvalid cases will be eliminated from the analysis. When the OK button is clicked on this dialog additional variables will be created in memory for each variable and transformation. For this example, select the X^{**2} , X^{**3} , and $LN(X)$ check boxes.



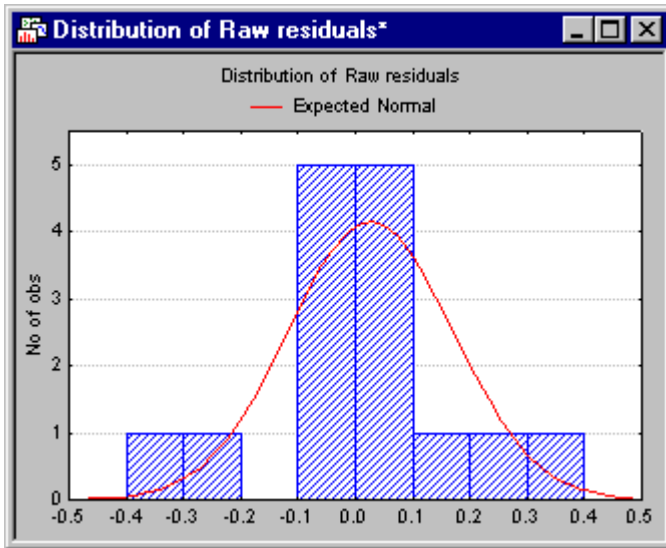
Now, click the OK button to open the Review Descriptive Statistics dialog (see below). The summary box at the top of the dialog indicates that the specified transformations were successfully applied to all cases in the data set.



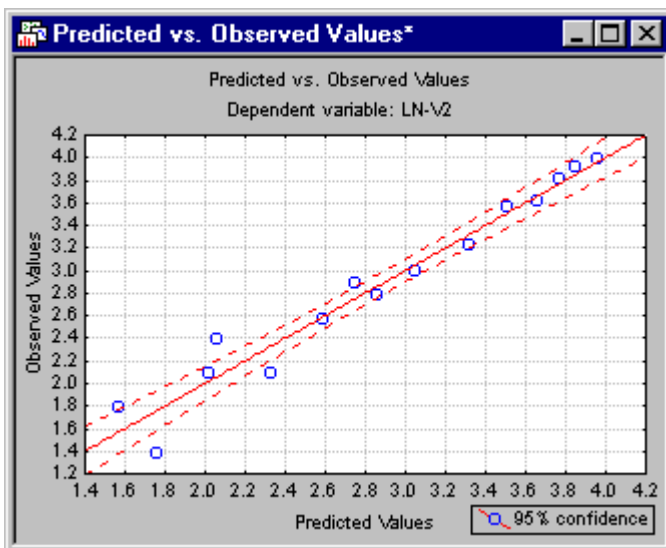
Reviewing the transformed variables. On the Quick tab, click the Correlations button to produce a spreadsheet of correlations between all combinations of the original variables and their respective transformations. On this spreadsheet (see below), note that the correlation between DAYS (V1) and PROGNOSIS (V2) is highest ($r = -0.977$) when PROGNOSIS is logarithmically transformed (LN-V2).

Data: Correlations (Patients)*								
Variable	Correlations (Patients.sta)							
	DAYS	PROGNOSIS	V1**2	V1**3	LN-V1	V2**2	V2**3	LN-V2
DAYS	1.000000	-0.941053	0.968107	0.914647	0.919686	-0.858993	-0.786973	-0.977280
PROGNOSIS	-0.941053	1.000000	-0.833385	-0.742322	-0.982535	0.976196	0.934715	0.949181
V1**2	0.968107	-0.833385	1.000000	0.985486	0.806789	-0.715873	-0.630328	-0.937609
V1**3	0.914647	-0.742322	0.985486	1.000000	0.718699	-0.613618	-0.528271	-0.880520
LN-V1	0.919686	-0.982535	0.806789	0.718699	1.000000	-0.980910	-0.958517	-0.910681
V2**2	-0.858993	0.976196	-0.715873	-0.613618	-0.980910	1.000000	0.988570	0.864646
V2**3	-0.786973	0.934715	-0.630328	-0.528271	-0.958517	0.988570	1.000000	0.793656
LN-V2	-0.977280	0.949181	-0.937609	-0.880520	-0.910681	0.864646	0.793656	1.000000

Performing the analysis. Click the OK button on the Review Descriptive Statistics dialog to proceed with the analysis. On the Model Definition - Quick tab, click the Variables button to display the standard variable selection dialog. Here, select LN-V2 as the Dependent variables, variable DAYS as the Independent variables, and then click the OK button. Now click the OK button on the Model Definition dialog to calculate the model and display the Multiple Regression Results dialog. The model fits the data very well, with roughly 95% of the variability in prediction of LN(PROGNOSIS) explained by the model (see the adjusted R^2 value in the summary box).



Finally on the Scatterplots tab, click the Predicted vs. observed button to produce scatterplot of the predicted and observed values of the dependent variable.



The resulting plot shows that the predictions are good in a general sense, especially for higher LN(PROGNOSIS) values. Predicting the prognosis for patients whose hospital stay was shorter is less accurate.