

Demo: Analyzing Residuals and Outliers

In this example, we continue where we left off in the previous JMP demo. Recall that we fit a model for Impurity, with three predictors, Temp, Catalyst Conc, and Reaction Time, using Analyze, Fit Model.

Impurity is the Y variable, and Temp, Catalyst Conc, and Reaction Time are the model effects.

The Actual by Predicted plot, which was described earlier, is a plot of the actual response values against the values predicted by our current model.

The Actual by Predicted plot is useful for identifying outliers, unusual observations, and patterns in the data not explained by our model, such as curvature. For the most part, our points appear to be randomly scattered around the fitted regression line.

JMP provides the Residual by Predicted plot by default. This plot shows no obvious patterns, and the residuals appear randomly scattered around the center line of zero.

Additional residual plots described in this lesson are available from the top red triangle under Row Diagnostics. This includes the Residual by Row plot, the Studentized Residuals plot, and the Normal Quantile plot of residuals.

Note that residuals, studentized residuals, predicted values, and other values can be saved as columns to the data table using the options under Save Columns.

This is particularly useful, for example, if you want to use JMP platforms to create your own residual plots.

Cook's D Influence is another option under Save Columns.

When we select Cook's D Influence, a new column is saved to the data table with Cook's D values for each observation. We can now use the Distribution platform to determine whether we have any influential observations.

In this example, the maximum Cook's D value is 0.06. So, we don't have a problem with influential observations.

One additional type of residual plot in JMP is the effect leverage plot. These plots are displayed by default for each term in the model.

Essentially, each plot shows the impact of adding the predictor to the model, given the other effects already in the model.

Leverage plots make it easier to see unusual patterns, to see which points are exerting influence on the model, and to identify issues with multicollinearity.

For more information about interpreting leverage plots, we encourage you to visit the JMP Help.

Statistical Thinking for Industrial Problem Solving

Copyright © 2020 SAS Institute Inc., Cary, NC, USA. All rights reserved.

Close