Practice: Fitting a Logistic Regression Model with Interactions

Open the file **MetalCoatings.jmp**. Focus on building a logistic model with interactions in this exercise, but it's a good practice to become familiar with the data before you conduct a formal analysis. Explore the variables (**Columns Viewer**, **Distribution**, **Graph Builder**, and **Fit Y by X**) and the methods that were introduced in the Exploratory Data Analysis module.

Fit a logistic regression model in Fit Model for **Outcome**. Include all continuous and categorical main effects (**humidity** through **paint supplier**) and all possible two-way interactions.

1. How many effects are in the full model? (Hint: See the Effect Summary table.)

   There are 28 effects: seven main effects and 21 two-way interactions. The Parameter Estimates table reports the model coefficients. There are 35 model coefficients.

2. What is the misclassification rate for the full model?

   The misclassification rate for the full model is 0.2126.

3. Using the Effect Summary table, slowly reduce the model. Use a *p*-value stopping rule of 0.10. How many effects are in the reduced model?

   There are nine effects in the reduced model: three two-way interactions and six main effects.

4. What is the misclassification rate for the reduced model?

   The misclassification rate for the reduced model is 0.2350.

5. Compare the results of these two models. If the goal is to identify potential causes of defective parts, which model is better?

   The reduced model is better. Both the full and the reduced model can be used to classify parts as good or defective, but the full model is overly complicated. The full model is difficult to interpret and includes many effects that are not significant.

Hide Solution

Statistical Thinking for Industrial Problem Solving

Close