Print

The All-Possible Regressions Approach to Model Building

Let's start by examining how the all-possible regressions approach to model selection works. Suppose you have a fitness data set that includes the response variable, Oxygen_Consumption, and seven predictor variables: RunTime, Age, Weight, Run_Pulse, Rest_Pulse, Maximum_Pulse, and Performance. For k predictors, there are 2^k possible models, including the intercept-only model. That means that for the fitness data, there are 2^7, or 128, possible regression models, excluding polynomial and interaction effects.

You can specify one possible model that includes only the intercept and no predictor variables, seven possible models with one predictor variable, 21 possible models with two predictor variables, 35 possible models with three predictor variables, and so on, up to the full model.

As you might expect, when you use the all-possible regressions approach, SAS calculates all possible regression models. However, you can reduce the number of models in the output by specifying the BEST= option in the MODEL statement of PROC REG. SAS will still evaluate all possible models, but display only the requested subset. You'll learn how to do this in the self-study section of this lesson.

Close