When Is Stepwise Regression Appropriate?

Stepwise regression is an appropriate analysis when you have many variables and youre interested in identifying a useful subset of the predictors. In Minitab, the standard stepwise regression procedure both adds and removes predictors one at a time. Minitab stops when all variables not included in the model have p-values that are greater than a specified Alpha-to-Enter value and when all variables that are in the model have p-values that are less than or equal to a specified Alpha-to-Remove value.

In addition to the standard stepwise method, Minitab offers two other types of stepwise procedures:

Forward selection: Minitab starts with no predictors in the model and adds the most significant variable for each step. Minitab stops when all variables not in the model have p-values that are greater than the specified Alpha-to-Enter value. Backward elimination: Minitab starts with all predictors in the model and removes the least significant variable for each step. Minitab stops when all variables in the model have p-values that are less than or equal to the specified Alpha-to-Remove value.

On the Stepwise tab of the Multiple Regression dialog box, select the stepwise regression method. Stepwise removes and adds terms to the model for the purpose of identifying a useful subset of the terms. For more information, go to Basics of stepwise regression.

Specify the method that is used to fit the model.

Stepwise: This method starts with an empty model, or includes the terms you specified to include in the initial model or in every model. Then, Minitab adds or removes a term for each step. You can specify terms to include in the initial model or to force into every model. Minitab stops when all variables not in the model have p-values that are greater than the specified Alpha to enter value and when all variables in the model have p-values that are less than or equal to the specified Alpha to remove value.

Forward selection: This method starts with an empty model, or includes the terms you specified to include in the initial model or in every model. Then, Minitab adds the most significant term for each step. Minitab stops when all variables not in the model have p-values that are greater than the specified Alpha to enter value.

Backward selection: This method starts with all potential terms in the model and removes the least significant term for each step. Minitab stops when all variables in the model have p-values that are less than or equal to the specified Alpha to remove value.

Pitfalls of Stepwise Regression

While a lot can be learned with stepwise regression, there are some potential pitfalls to be aware of:

If two independent variables are highly correlated, only one may end up in the model even though both may be important. Because the procedure fits many models, it could be selecting models that fit the data well due to chance alone Stepwise regression may not always end with the model with the highest R2 value possible for a given number of predictors. Automatic procedures cannot take into account special knowledge the analyst may have about the data. Therefore, the model selected may not be the most practical one. Graphing individual predictors against the response is often misleading because graphs do not account for other predictors in the model.