Forward Selection

- The forward selection procedure looks to add variables to the model.
- Once added, those variables stay in the model even if they become insignificant at a later step.

How does JMP do this?

- Analysis – Fit Model
- Enter the response variable in the Pick Role Variables box as Y.
- Add all the explanatory variables to the Construct Model Effects box.

JMP – Fit Model

- Make the Personality - Stepwise.
- Click on Run Model.
### Forward - Set up

#### Stepwise Fit

**Response**: MDBH  
**Prob to Enter**: 0.250  
**Prob to Leave**: 0.100  
**Direction**: Forward

#### Stepwise Regression Control

<table>
<thead>
<tr>
<th>Lock Entered Parameter</th>
<th>Estimates</th>
<th>stdF</th>
<th>SS</th>
<th>“F Ratio”</th>
<th>“Prob&gt;F”</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>6.265</td>
<td>1</td>
<td>8.265</td>
<td>0.0034</td>
<td>0.9999</td>
</tr>
<tr>
<td>X1</td>
<td>0</td>
<td>1</td>
<td>0.207045</td>
<td>20.739</td>
<td>0.0001</td>
</tr>
<tr>
<td>X2</td>
<td>0</td>
<td>1</td>
<td>0.819027</td>
<td>1.125</td>
<td>0.3006</td>
</tr>
<tr>
<td>X3</td>
<td>1</td>
<td>1</td>
<td>7.335255</td>
<td>43.287</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

#### Current Estimates

- Only the intercept is included at this point.
- The value of the intercept, 6.265 is the mean response (MDBH).

### Stepwise Regression Control

- **Direction** - Forward
- **Prob to Enter** - the P-value for a variable must be less than or equal to the Prob to Enter in order for the variable to be added to the model.

### Current Estimates

- Only the intercept is included at this point.
- The value of the intercept, 6.265 is the mean response (MDBH).
Current Estimates

- The SSE, at this point, is actually the C. Total sum of squares.
  - 10.3855
- Under the SS column are the sum of squares that will be explained if that variable is added to the model.

Note the $X_3$ will add the largest sum of squares if it is added to the model.
- 7.335
- If $X_3$ is added the SLR of MDBH on $X_3$ will have
  - $R^2 = 7.335/10.3855 = 0.7063$

Note that adding $X_3$ will be a statistically significant addition to the model.
- “F-Ratio” = 43.287
- “Prob>F” = 0.0000 (P-value)
- The P-value is small.
- Click on Step
### Stepwise Fit

**Response:** MDBH

**Stepwise Regression Control**

- **Prob to Enter:** 0.250
- **Prob to Leave:** 0.100
- **Direction:** Forward

### Current Estimates

<table>
<thead>
<tr>
<th>Step</th>
<th>Parameter</th>
<th>Action</th>
<th>Estimate</th>
<th>“Sig Prob”</th>
<th>Seq SS</th>
<th>R-Square</th>
<th>Cp</th>
<th>Adj R-Square</th>
<th>AIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>X3</td>
<td>Entered</td>
<td>3.896</td>
<td>0.0000</td>
<td>7.33255</td>
<td>0.7063</td>
<td>19.388</td>
<td>33.6102</td>
<td>3.0502449</td>
</tr>
</tbody>
</table>

### Current Estimates – Step 1

- **X3** is added to the model
- **Predicted MDBH = 3.896 + 32.937*X3**
- **R^2 = 0.7063**
- **RMSE = \sqrt{MSE} = \sqrt{0.1694581} = 0.4117**

### Current Estimates – Step 1

- Of the remaining variables not in the model **X1** will add the largest sum of squares if added to the model.
  - **SS = 1.000**
  - **“F Ratio” = 8.294**
  - **“Prob>F” = 0.0104**
JMP Forward – Step 2

- Because $X_1$ will add the largest sum of squares and that addition is statistically significant, by clicking on Step, JMP will add $X_1$ to the model with $X_3$.

- The model is now:

  \[ \text{Predicted } \text{MDBH} = 3.143 + 0.0314 \times X_1 + 22.954 \times X_3 \]

- \[ R^2 = 0.8026 \]

- \[ \text{RMSE} = \sqrt{\text{MSE}} = \sqrt{0.1205933} = 0.3473 \]
Current Estimates - Step 2

- Of the remaining variables not in the model $X_2$ will add the largest sum of squares if added to the model.
  - $SS = 0.671$
  - "F Ratio" = 7.784
  - "Prob>F" = 0.0131

JMP Forward - Step 3

- Because $X_2$ will add the largest sum of squares and that addition is statistically significant, by clicking on Step, JMP will add $X_2$ to the model with $X_3$ and $X_1$. 

---

<table>
<thead>
<tr>
<th>Step</th>
<th>Parameter</th>
<th>Action</th>
<th>&quot;Sig Prob&quot;</th>
<th>Seq SS</th>
<th>RSquare</th>
<th>Cp</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>X3</td>
<td>Entered</td>
<td>0.0000</td>
<td>7.335255</td>
<td>0.7083</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>X1</td>
<td>Entered</td>
<td>0.0104</td>
<td>1.000159</td>
<td>0.8026</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>X2</td>
<td>Entered</td>
<td>0.0131</td>
<td>0.870567</td>
<td>0.8672</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>
Current Estimates – Step 3
- $X_2$ is added to the model
- Predicted MDBH = $3.236 + 0.0974 \cdot X_1 - 0.000169 \cdot X_2 + 3.467 \cdot X_3$
- $R^2 = 0.8672$
- $\text{RMSE} = \sqrt{\text{MSE}} = \sqrt{0.0861949} = 0.2936$

Current Estimates – Step 3
- There are no variables remaining and so the forward selection procedure stops.
- Note that variable $X_3$ is no longer statistically significant.
- The combination of $X_1$ and $X_2$ has made $X_3$ redundant.

Backward Selection
- Start with a full model (a model that contains all of the available explanatory variables).
- Remove variables, one at a time, if they do not add significantly to the model.
Full Model - MDBH

- Predicted MDBH = 3.236 + 0.0974*X₁ - 0.000169*X₂ + 3.467*X₃
- $R^2 = 0.8672$
- $\text{RMSE} = \sqrt{\text{MSE}} = \sqrt{0.0861949} = 0.2936$

Statistical Significance

- $X₁$: $F = 14.709$, $P-value = 0.0015$
- $X₂$: $F = 7.784$, $P-value = 0.0131$
- $X₃$: $F = 0.171$, $P-value = 0.6844$

Backward Selection - Step 1

- Because the $P-value$ for $X₃$ is not small, it should be removed from the model.
- Removing $X₃$ will subtract 0.0148 from the sum of squares model.
**Backward Selection**

- Because all of the remaining variables in the model are statistically significant, the backward selection procedure stops.

**Stepwise Regression Control**

- **Direction** - Backward
- **Prob to Leave** - the P-value for a variable must be greater than the Prob to Leave in order for the variable to be removed from the model.
- **Click on Enter All**
Backward – Set up

**Response:** MDBH

**Prob to Enter:** 0.250

**Prob to Leave:** 0.100

**Direction:** Backward

<table>
<thead>
<tr>
<th>SSE</th>
<th>DFE</th>
<th>MSE</th>
<th>R Square</th>
<th>R Square Adj</th>
<th>Cp</th>
<th>AIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.3759</td>
<td>16</td>
<td>0.0610</td>
<td>0.8672</td>
<td>0.8433</td>
<td>4</td>
<td>-45.4327</td>
</tr>
</tbody>
</table>

**Lock Entered Parameter**

<table>
<thead>
<tr>
<th>Estimate</th>
<th>df</th>
<th>SS</th>
<th>&quot;F Ratio&quot;</th>
<th>&quot;Prob&gt;F&quot;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>2</td>
<td>0</td>
<td>0.000</td>
<td>1.0000</td>
</tr>
<tr>
<td>X1</td>
<td>1</td>
<td>0.0745</td>
<td>1.2679</td>
<td>0.0015</td>
</tr>
<tr>
<td>X2</td>
<td>1</td>
<td>0.0107</td>
<td>7.784</td>
<td>0.0191</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>nDF</th>
<th>SS</th>
<th>&quot;F Ratio&quot;</th>
<th>&quot;Prob&gt;F&quot;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>3.2357</td>
<td>0</td>
<td>14.709</td>
<td>0.0000</td>
<td></td>
</tr>
<tr>
<td>X1</td>
<td>0.0974</td>
<td>1</td>
<td>2.7846</td>
<td>0.0147</td>
<td></td>
</tr>
<tr>
<td>X2</td>
<td>-0.0002</td>
<td>1</td>
<td>0.1708</td>
<td>0.6844</td>
<td></td>
</tr>
</tbody>
</table>

**Current Estimates**

- Full model containing all of the explanatory variables.
- Click on Step

**Step History**

<table>
<thead>
<tr>
<th>Step</th>
<th>Parameter</th>
<th>Action</th>
<th>&quot;Sig Prob&quot;</th>
<th>Seq SS</th>
<th>R Square</th>
<th>Cp</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>X3</td>
<td>Removed</td>
<td>0.0044</td>
<td>0.14774</td>
<td>0.0058</td>
<td>2.1714</td>
<td>3</td>
</tr>
</tbody>
</table>

**Stepwise Fit**

**Response:** MDBH

**Prob to Enter:** 0.250

**Prob to Leave:** 0.100

**Direction:** Backward

<table>
<thead>
<tr>
<th>SSE</th>
<th>DFE</th>
<th>MSE</th>
<th>R Square</th>
<th>R Square Adj</th>
<th>Cp</th>
<th>AIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.3939</td>
<td>17</td>
<td>0.0618</td>
<td>0.8655</td>
<td>0.8550</td>
<td>2.1714</td>
<td>0.2726</td>
</tr>
</tbody>
</table>

**Lock Entered Parameter**

<table>
<thead>
<tr>
<th>Estimate</th>
<th>df</th>
<th>SS</th>
<th>&quot;F Ratio&quot;</th>
<th>&quot;Prob&gt;F&quot;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>3.2605</td>
<td>0</td>
<td>102.149</td>
<td>0.0000</td>
</tr>
<tr>
<td>X1</td>
<td>0.1069</td>
<td>1</td>
<td>33.961</td>
<td>0.0000</td>
</tr>
<tr>
<td>X2</td>
<td>-0.0002</td>
<td>1</td>
<td>0.1708</td>
<td>0.6844</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>nDF</th>
<th>SS</th>
<th>&quot;F Ratio&quot;</th>
<th>&quot;Prob&gt;F&quot;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>3.4668</td>
<td>0</td>
<td>0</td>
<td>1.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>X1</td>
<td>0.1069</td>
<td>1</td>
<td>33.961</td>
<td>0.0000</td>
<td></td>
</tr>
<tr>
<td>X2</td>
<td>-0.0002</td>
<td>1</td>
<td>0.1708</td>
<td>0.6844</td>
<td></td>
</tr>
</tbody>
</table>

**Current Estimates**

**Step History**

<table>
<thead>
<tr>
<th>Step</th>
<th>Parameter</th>
<th>Action</th>
<th>&quot;Sig Prob&quot;</th>
<th>Seq SS</th>
<th>R Square</th>
<th>Cp</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>X3</td>
<td>Removed</td>
<td>0.0044</td>
<td>0.14774</td>
<td>0.0058</td>
<td>2.1714</td>
<td>3</td>
</tr>
</tbody>
</table>
Current Estimates – Step 1

- $X_3$ is removed from the model
- Predicted MDBH = 3.261 + 0.1069$X_1$ - 0.0001898$X_2$
- $R^2 = 0.8658$
- $RMSE = \sqrt{MSE} = \sqrt{0.0819937} = 0.2863$

All of the remaining variables are statistically significant.

Clicking on Step will not change anything because no variable can be removed.

Backward Selection

Once a variable is removed it can never be entered again, even if it would add significantly to a model later on.