CHAPTER 5

STATISTICS

Statistics Window

Opening the Statistics Window

Double-click on the **Statistics** folder in the **Project** window or click the **Statistics** button.

A Statistics window similar to the one shown below will appear.

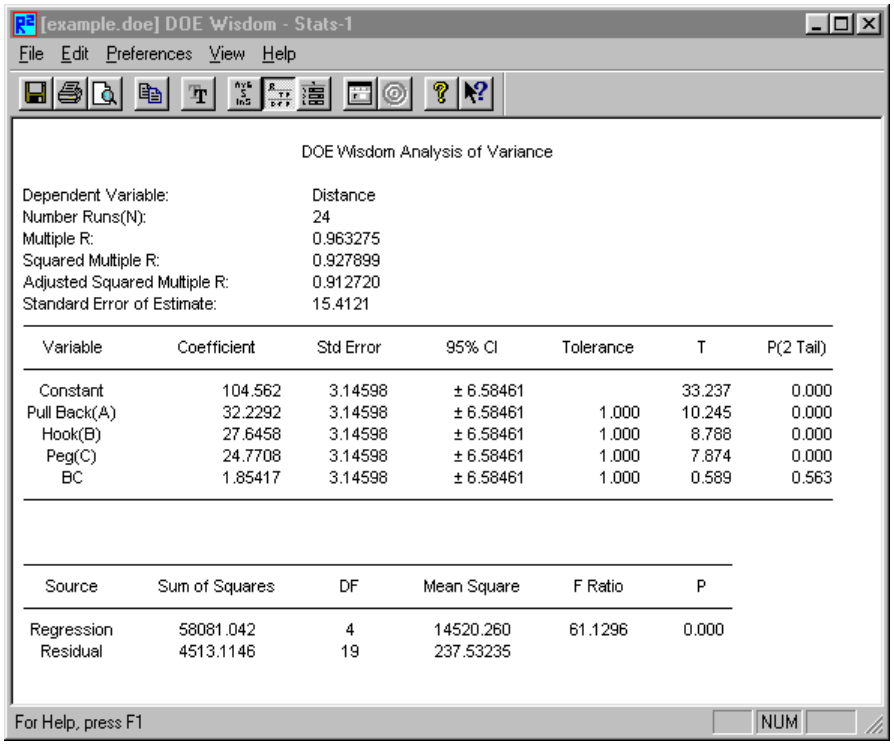

Figure 5-1

The Statistic that appears is generated according to the default statistic selected in the **Properties** section of the **File** pull-down menu in the **Project** screen. In this example, the ANOVA was selected and is the first statistic to appear.

Statistics File Menu Commands

The Statistics **File** menu offers the following commands:

Save	Saves a statistics report for future reference.
Save As	Saves the active statistics report under a new name for future reference.
Print	Prints a statistics report. Use this command to print a copy of the statistics report. DOE Wisdom displays the **Print** screen. This screen allows you to select the printer, printer properties, print range, and number of copies. When all selections have been made, choose **OK**. The statistics report will be printed.
Print Preview	Displays the document exactly as it will look when printed. With the document displayed you can view each of the pages of the document. If you wish to print the document, select the "Print" button. To close without printing the document, select the "Close" button.
Print Setup	Provides setup options for printing. DOE Wisdom displays the **Page Setup** screen. This screen allows you to select the paper size, paper source, portrait or landscape mode, and margin specifications. When all selections have been made, choose **OK**.
Properties	Displays the properties for a statistics report. When "Properties" is selected, the following window appears. The user can select either the ANOM, ANOVA, or Prediction Equation to be displayed automatically upon entering the Statistics window. The user can also have the edit locations automatically displayed. If this option is selected, edit locations will be surrounded by a "box." Double-clicking within a given box will bring up the editing options for that particular section of the statistics report.

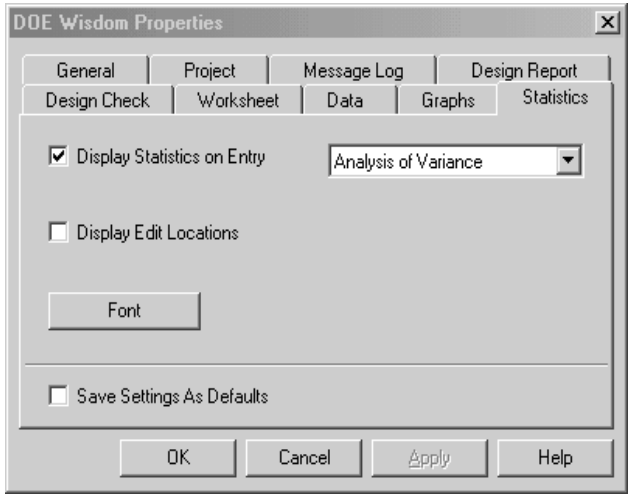

Figure 5-2

Exit	Exits the Statistics window.

Statistics Edit Menu Commands

The Statistics **Edit** menu offers the following commands:

Copy	Copies the statistics report onto the Clipboard. This allows you to easily copy the statistics report to other window programs.
Copy Regression Results	Copies raw and studentized residual data onto the Clipboard. This allows you to easily copy the residual data to other window programs for printing.
Report Variables	Edits the variables used in the statistics report.
Find Target Value	Activates the "Find Target Value" screen. This is only enabled when the prediction equation screen is selected.

Statistics Preferences Menu Commands

The Statistics **Preferences** menu offers the following commands:

Font	Changes the font attributes of the current statistics report. Font, Font Style, and Font Size may be defined.

Statistics View Menu Commands

The Statistics **View** menu offers the following commands:

Edit Locations	Highlights sections of the ANOM, ANOVA, and Prediction equation that can be modified. Edit locations will be surrounded by a "box." Double-clicking within a given box will bring up the editing options for that particular section of the statistics report.
Toolbar	Shows or hides the toolbar. The toolbar includes buttons for some of the most common statistics report commands such as Save, Print, Print Preview, Copy, Font, Analysis of Means, Analysis of Variance, Prediction Equation, Edit Variables, and Find the Target. The Statistics toolbar is shown in Figure 5-3.

Figure 5-3

Status Bar	Shows or hides the Status Bar. The Status Bar describes the action to be executed by the selected menu item or depressed toolbar button.
Analysis of Means	Selects the Analysis of Means statistic report.
Analysis of Variance	Selects the Analysis of Variance statistic report.
Prediction Equation	Selects the Prediction Equation statistic report.

Statistics Help Menu Commands

The Statistics **Help** menu offers the following commands:

Statistics Help	This allows the user to display specific statistic help topics.
Help Topics	This allows the user to view Contents, Index, and Find help features.

Analysis of Means

Analysis of Means Screen

You can select the ANOM by clicking on the **ANOM of Means** from the **View** pull-down menu.  button or by selecting **Analysis**

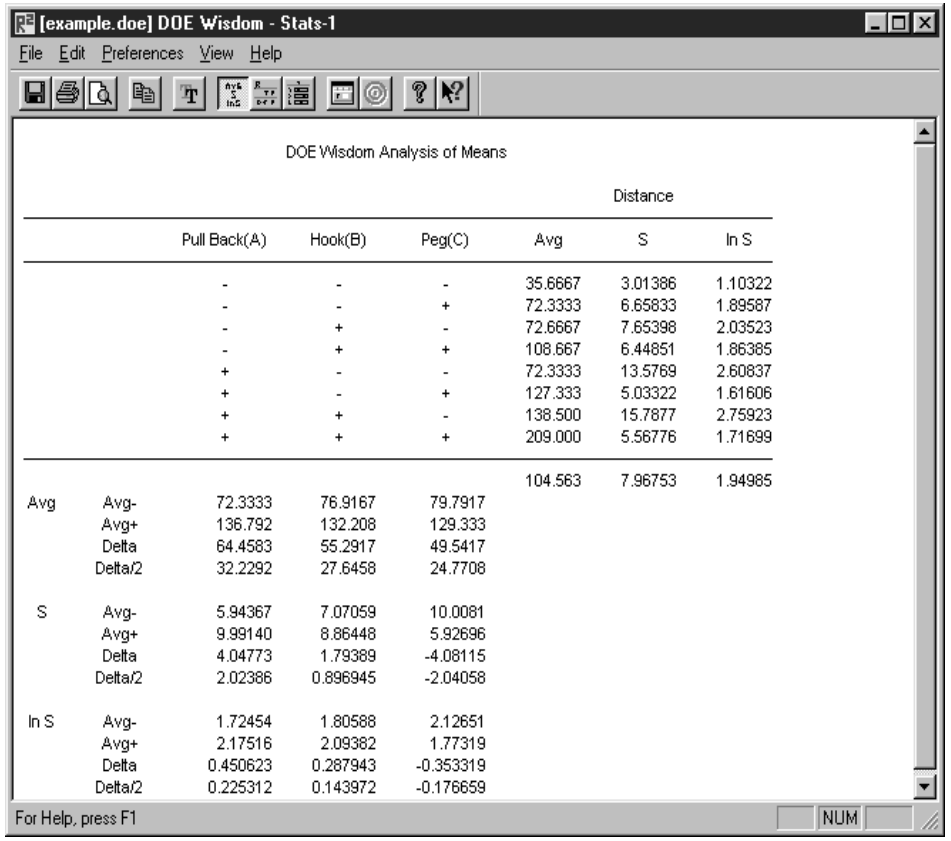

Figure 5-4

When Analysis of Means is selected, a screen similar to the one shown in Figure 5-4 will appear. If there are several factors and not all of them fit in the area, simply use the scroll bars to move to different sections of the window.

Analysis of Means Terminology

Avg The column of numbers under **Avg** gives the average value of the response for that run. The last row is the average of the averages.

S This is the standard deviation of the response values for that run. The last row is the average of the standard deviations for all the runs.

ln S This is the natural log of the standard deviation. The last row is the average of the natural logs for all the runs.

Avg - This is the average of the response values for a factor at its low (-) settings. For Pull-Back in Figure 5-4, the Avg - is 72.3333. This is the average of (35.6667 + 72.3333 + 72.6667 + 108.667).

Avg + This is the average of the response values for a factor at its high (+) settings. For Pull-Back in Figure 5-4, the Avg + is 136.792. This is the average of (72.3333 + 127.333 + 138.500 + 209.000)

Delta This is the difference between the low and high values. The greater the difference, the greater the significance of the factor.

Delta/2 Equals the delta value divided by 2. In two level designs, this is used as the coefficient for that factor in the prediction equation.

Selecting Factors/Responses/Options for the ANOM

Select the **Report Variables** option from the **Edit** pull-down menu or click the **Edit Variables** button.

A window similar to the one shown in Figure 5-5 will appear.

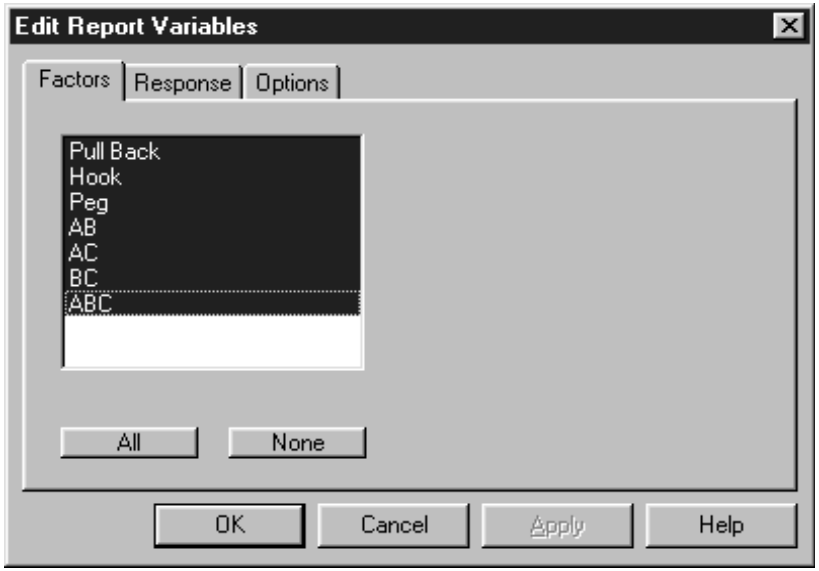

Figure 5-5

Click the **Factors** tab to display factors that can be added or removed. Highlighted factors will remain in the ANOM. Use the mouse pointer to select a factor. Click the left mouse button to remove or add the factor.

All the factors can be selected by clicking the **All** button. If there is a long list of factors and you only want to include a few, it may be easier to select the **None** button and then highlight the desired factors.

Click the **Response** tab to display responses that can be selected. The ANOM will be displayed for the highlighted responses. Use the mouse pointer to select a response.

Click the **Options** tab to display the options for displaying qualitative factors.

Taguchi Signal-To-Noise Statistics

When a Taguchi design has been selected, the various S/N ratios can be displayed in the ANOM.

Select the **Report Variables** option from the **Edit** pull-down menu or click the **Edit Variables** button.

The following window will appear:

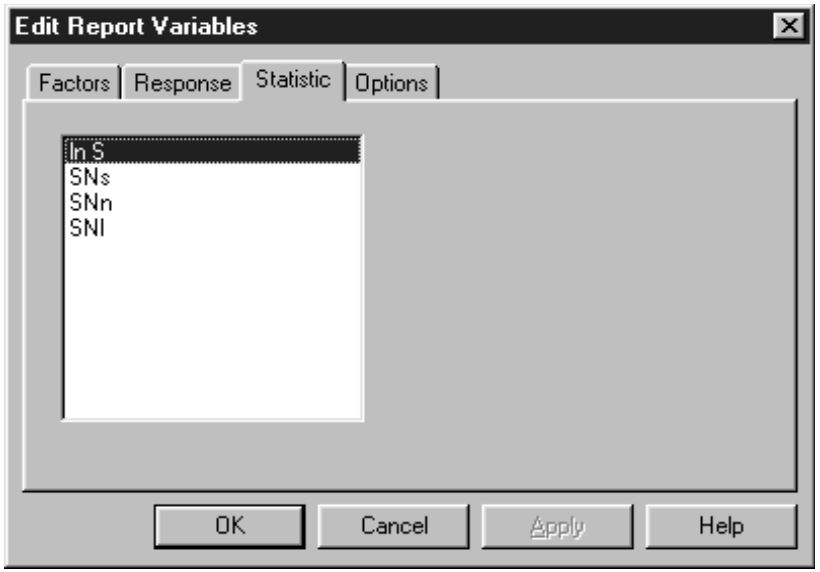

Figure 5-6

DOE Wisdom supports the following S/N ratios:

Smaller is better
Larger is Better
Nominal is Better

Use the mouse pointer to select a S/N ratio. Click the left mouse button to highlight the desired S/N ratio.

NOTE: S/N ratios are not calculated if there is only one replicate for the response.

Analysis of Variance

Analysis of Variance Screen

You can select the ANOVA by clicking on the **ANOVA**  button or by selecting **Analysis of Variance** from the **View** pull-down menu.

When Analysis of Variance is selected, a screen similar to the one shown in Figure 5-7 will appear. If there are several factors/interactions and not all of them fit in the area, simply use the scroll bars to move to different sections of the window.

There is quite a bit of valuable information in the ANOVA. Whether you are new to design of experiments or not, this information may be rather confusing. The next section gives a basic explanation of the terms listed in the ANOVA.

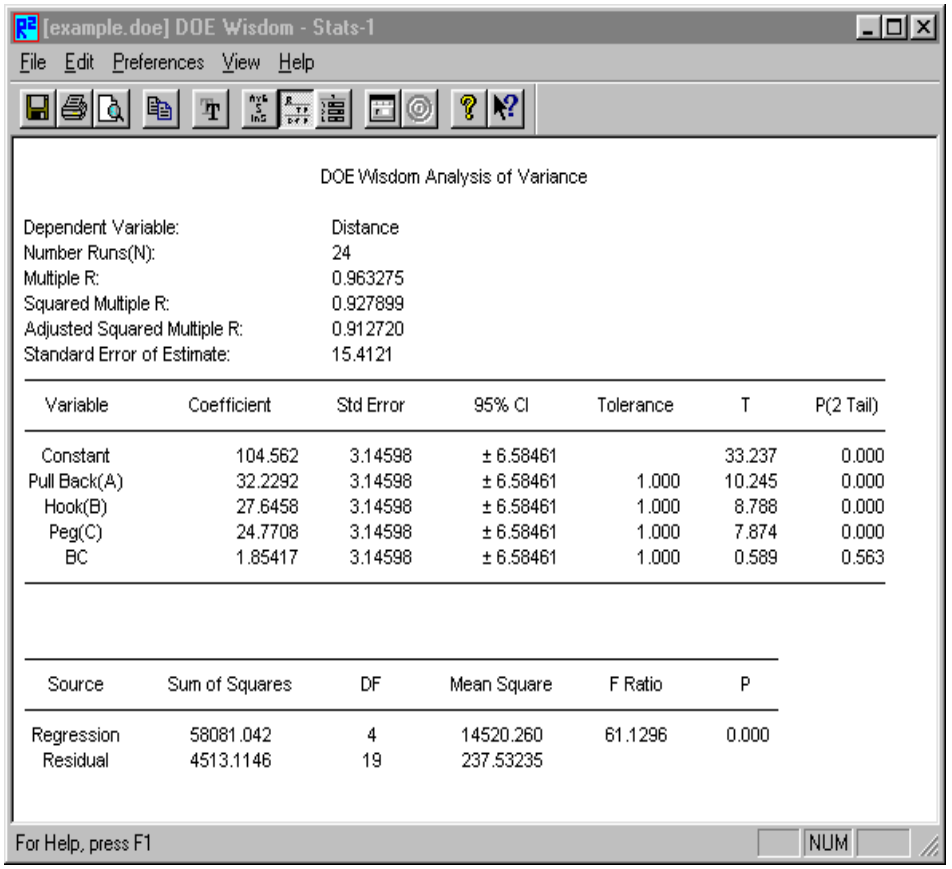

Figure 5-7

Analysis of Variance Terminology

Dependent Variable This lists the response for which the ANOVA was generated.

Number Runs (N) The total number of runs in the experiment.

Multiple R The correlation coefficient (R) for multiple regression. It is the square root of the coefficient of determination (R^2). The value of R is limited to the interval [-1, +1]. A perfect negative correlation is -1 and +1 indicates a perfect positive correlation. Closer to ±1 is typically better.

Squared Multiple R The coefficient of determination (R^2) for multiple regression. In Figure 5-7 the R^2 value is 0.927899. This tells us that 92.8% of the variance in the data (about the mean) was due to the fact that the response varied for different levels of the factors. Closer to 1 is typically better.

Adjusted Squared Multiple R
A large number of terms for a relatively small n may lead to an overfit condition. This will mask the true variation due to error. The Adjusted R^2 inflates the error estimate using a ratio of degrees of freedom and recomputes R^2. Closer to 1 is typically better.

Standard Error of Estimate
The standard error of the estimate is the square root of the residual mean square (the mean square error). It is a measure of the unexplained variability in the dependent variable. The Standard Error of Estimate in Figure 5-7 is 15.4121. Using the empirical percentages for a normal distribution, 68.26% of any further runs should fall inside of the predicted value ±15.4121.

Variable The column under this term lists all the variables for the experiment.

Coefficient

DOE Wisdom displays coefficients in terms of the coded values. The coefficient is the number which that particular factor will be multiplied by in the prediction equation.

Std Error

$$StandardError = \sqrt{\frac{MSE}{n}}$$

Using the standard error we can precisely state the number of standard deviation units each coefficient is from zero on the t distribution. In Figure 5-7, the coefficient for Hook is 27.6458: or 8.788 standard deviation units to the right of zero on the t

$$\frac{27.6458}{3.14598}$$

distribution.

95% CI

DOE Wisdom displays the confidence interval for each coefficient. This is the range within which the value of the coefficient is expected to fall with some specified confidence level or confidence coefficient. When the confidence interval is displayed, the user can set the percent at which the confidence intervals will be calculated. The default percent is 95.

Tolerance

The Tolerance is calculated for each factor. A regression is run for a particular factor in terms of all remaining factors. Then an R^2 is calculated for that regression. Tolerance is equal to $1-R^2$ (for that regression). If there is no shift in a factor due to different levels of the other factors and interactions, $R^2=0$ and the tolerance for that factor equals 1. A tolerance of 1.00 suggests orthogonality.

T

Using the standard error we can precisely state the number of standard deviation units each coefficient is from zero on the t distribution. In Figure 5-7, the coefficient for Hook is 27.6458:

$$\frac{27.6458}{3.14598}$$

or 8.788 standard deviation units to the right of zero on the t distribution. This is the T value for Hook found in the ANOVA table in Figure 5-7.

P (2 Tail)

In the ANOVA table, the P(2 Tail) value is 0.000 for Hook. The area to the right of 8.788 is pretty close to zero. We can be at least 99.9% confident that this factor is significant. Typically P values greater than 0.10 are not included in the model. Once again, this is dependent on the significance level you want for your particular experiment.

Source

This simply identifies the source of the data for the following columns.

Sum of Squares

The Sum of Squares Error or Residual (SSE) is the summation of (**data value at run setting - predicted value at that setting)^2** for all the runs. The Sum of Squares Regression equals the sum of the number of data points collected during each run times the squared difference between the predicted value for a run and the overall mean.

DF

This refers to the degrees of freedom. Suppose you know that the mean of five numbers if 20. How many free choices do you have in selecting the numbers that will make this happen? The first four numbers are up to you; however, the fifth is predetermined by the mean. Therefore, you have four degrees of freedom.

Mean Square

The Mean Square Error (MSE) is an estimate of the variance of the population with the variance within subgroups. The MSE is found by dividing the Sum of Squares Residual (or error)(SSE) by the degrees of freedom (df).
The Mean Square Regression (MSR) is an estimate of the variance of the population with the variance between subgroups. MSR is found by dividing the Sum of Squares Regression by the degrees of freedom.

F Ratio

$$F=\frac{MSR}{MSE}$$

For Figure 5-7:

$$F=\frac{14520.260}{237.53235}=61.1296$$

If F > 6.0, there is most likely a significant shift in the response at different run settings. It indicates that the change in response at different settings didn't happen by chance. There is a high probability that the change happened because different factor levels shifted the response.
Depending on which level of confidence you want, you may wish to use an F Table. The F Table provides a means for determining, the significance of a factor to a specified level of confidence by comparing calculated F ratios to those from the F distribution. If the F ratio is greater than the table value, there is a significant effect. (Reference: *Understanding Industrial Designed Experiments*)

P

In the ANOVA table, P=0.000. The area to the right of 61.1296 is pretty close to zero. We can be 99.9% confident that the model detects a shift in data.

Selecting Factors/Responses/Options for the ANOVA

Select the **Report Variables** option from the **Edit** pull-down menu or click the **Edit Variables** button. A window similar to the one shown in Figure 5-8 will appear.

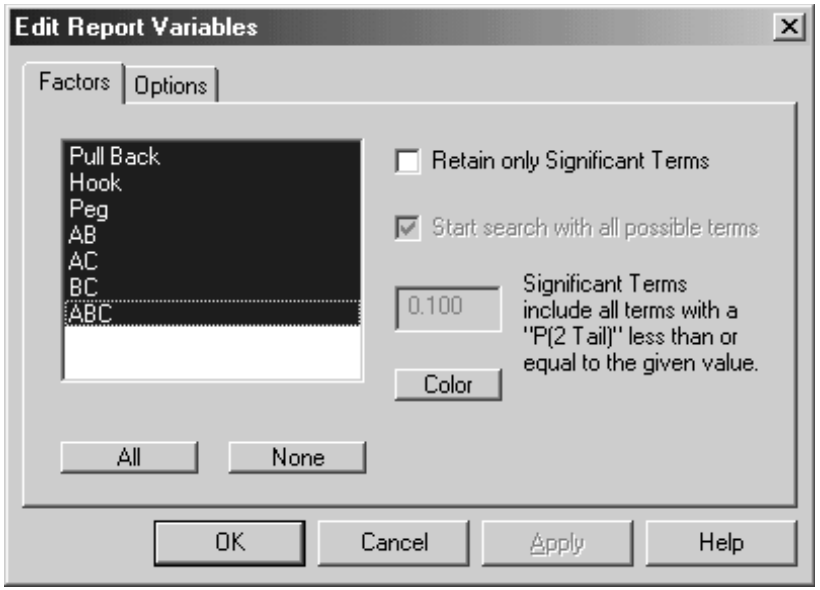

Figure 5-8

Click the **Factors** tab to display factors that can be added or removed. Highlighted factors will remain in the ANOVA. Use the mouse pointer to select a factor. Click the left mouse button to remove or add the factor.

All factors can be selected by clicking the **All** button. If there is a long list of factors and you only want to include a few, it may be easier to select the **None** button and then highlight the desired factors.

DOE Wisdom also supports a feature which will automatically retain only the significant terms. Click on the box next to “Retain only Significant Terms” in the Figure 5-8 window. Click on **OK**. The software will automatically remove terms with a P(2 Tail) value greater than 0.1. It is recommended that the user first select “Start Search with all possible terms” before running the significant term feature as otherwise the software will only look at terms

currently in the model. Select **OK**. The new ANOVA will be displayed. The ANOVA will now be displayed for only the significant terms. The user can change the cut off value for P(2 Tail) by simply typing in the new value in the box shown in Figure 5-8.

Click the **Response** tab to display responses that can be selected. The ANOVA will be displayed for the highlighted responses. Use the mouse pointer to select a response.

Click the **Options** tab to display the various confidence interval options.

Prediction Equation

Prediction Equation Screen

You can select the Prediction Equation by clicking on the **Prediction Equation** button or by selecting **Prediction Equation** from the **View** pull-down menu.

When Prediction Equation is selected, a screen similar to the one shown in Figure 5-9 will appear.

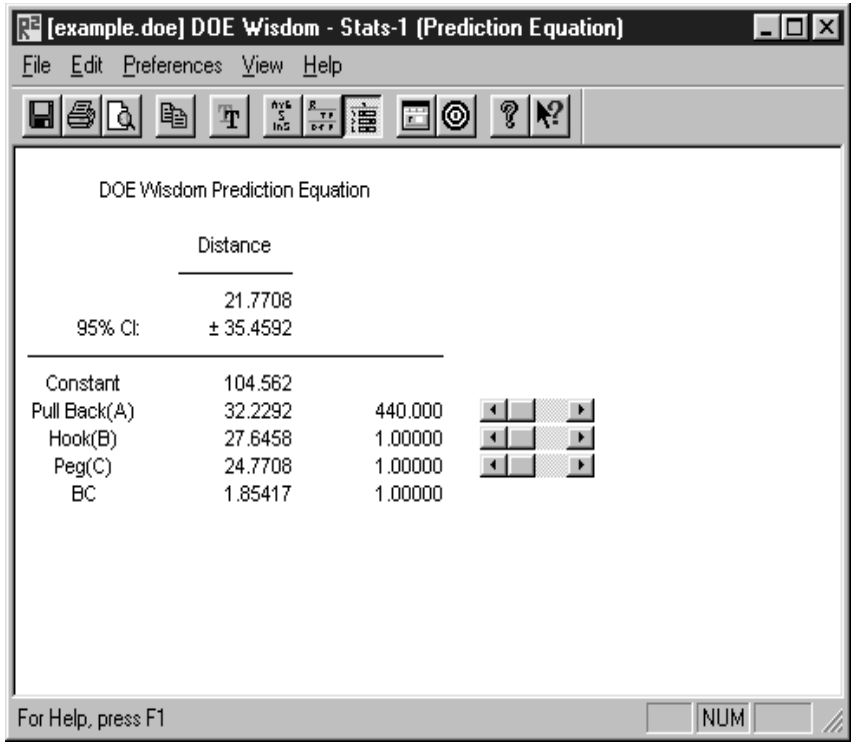

Figure 5-9

Selecting Factors/Responses/Options for the Prediction Equation

Select the **Report Variables** option from the **Edit** pull-down menu or click the **Edit Variables** button. A window similar to the one shown in Figure 5-10 will appear.

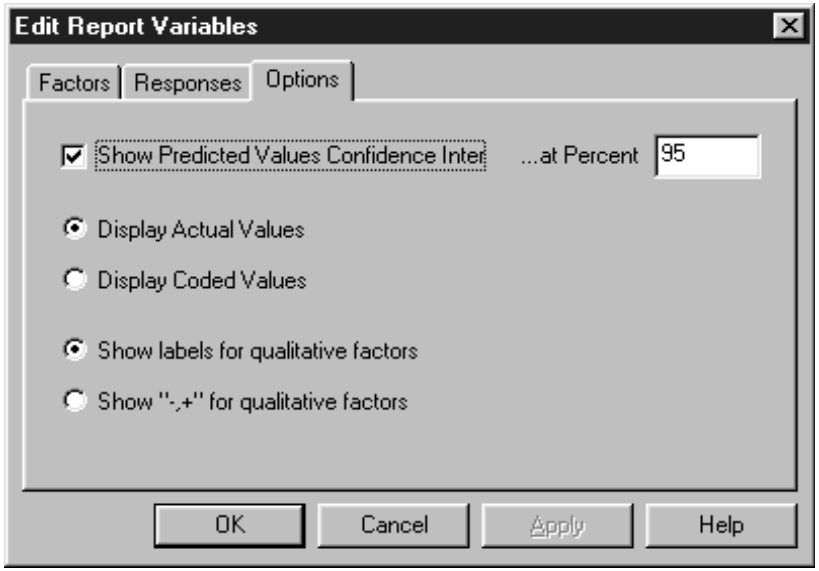

Figure 5-10

Click the **Factors** tab to display factors that can be added or removed. Highlighted factors will remain in the prediction equation. Use the mouse pointer to select a factor. Click the left mouse button to remove or add the factor.

All the factors can be selected by clicking the **All** button. If there is a long list of factors and you only want to include a few, it may be easier to select the **None** button and then highlight the desired factors.

Click the **Response** tab to display responses that can be selected. The prediction equation will be displayed for the highlighted responses. Use the mouse pointer to select a response. If you have chosen the desirability function when defining your responses, the D value for each response and the D(composite) value will also be listed as possible selections. All the responses can be selected by clicking the **All** button. All responses will show up on the prediction equation screen for a given model.

Click the **Options** tab. Confidence interval percent values can be changed. Actual values or coded values can be displayed in the prediction equation. Click on the button to the left of the value you want displayed. NOTE: The coefficients are always for the coded values. Qualitative factor display

options can be selected.

Changing Factor Settings/New Predicted Response

The horizontal scroll bars to the right of each factor can be used to change the factor settings and observe the new predicted response value. For our example in Figure 5-9, the following apply:

Factor Settings	**Response Values**
Pull Back set to 440	Distance = 21.7708
Hook set to 1	
Peg set to 1	

If we slide the horizontal scroll bars to reset the factor settings, new response values will be displayed. Figure 5-11 shows the new response values for the following new factor settings.

Factor Settings	**Response Values**
Pull Back set to 440	Distance = 91.1042
Hook set to 5	
Peg set to 2	

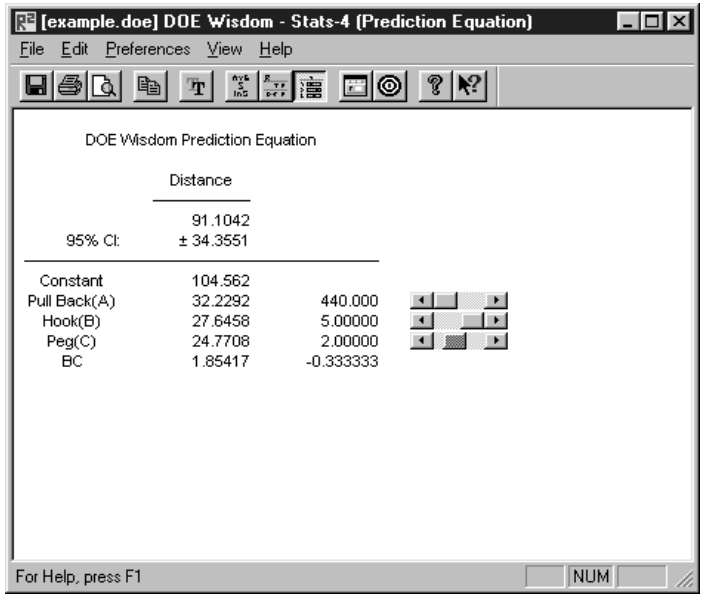

Figure 5-11

Hitting a Target

When the prediction equation screen is selected, the target button will become active. When you click on this button the following screen will appear.

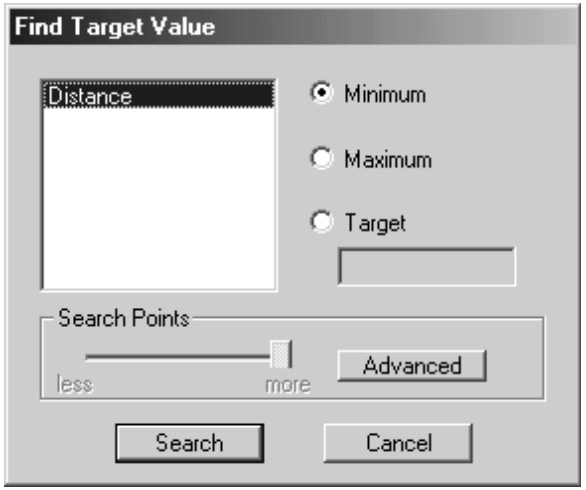

Figure 5-12

In this example, "Distance" is the response. The user can select to have the software find the minimum value, maximum value, or a target value for distance. For "Distance," we will select a target value of 55. The screen shown in Figure 5-13 will appear.

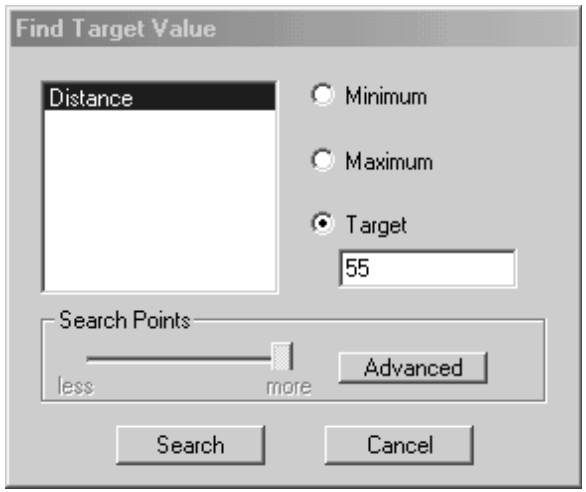

Figure 5-13

Click on the **OK** button. The software will now recommend factor settings that will "hit this target." A screen similar to the one shown in Figure 5-14 will appear. In this example, the software shows that a Distance of 54.9 can be "hit" by setting:

Pull Back = 442
Hook = 1
Peg = 3

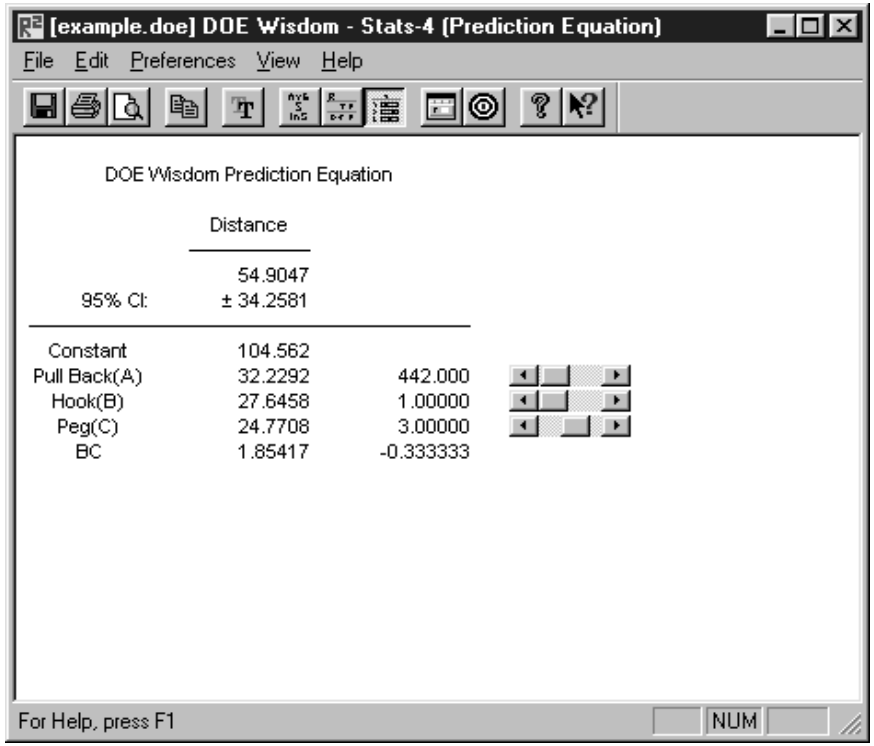

Figure 5-14

It should be noted that these factor settings may not be the ONLY settings that will generate a distance value of 55. The Target screen shown in Figure 5-13 has an **Advanced** button at the bottom of the screen. Click on this button. A screen similar to the one found in Figure 5-15 will appear.

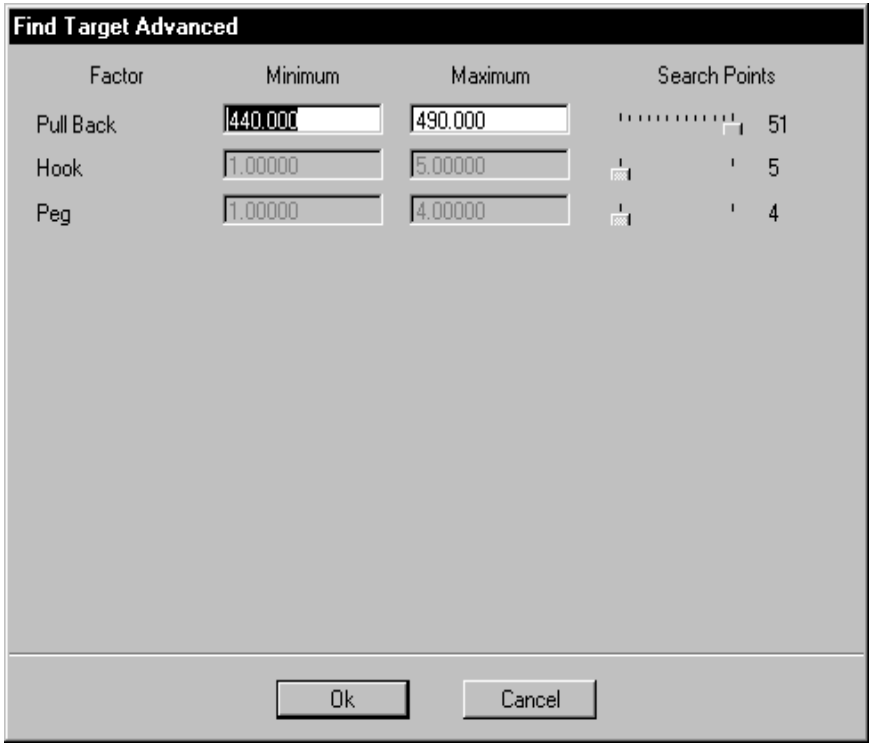

Figure 5-15

Let's say that it is not cost effective to run the Pull Back at 442. The user can ask the software to search Pull Back at settings between 450 and 490 instead of 440 and 490. Type the desired range in the minimum and maximum boxes next to Pull Back. The software now gives a default value of 41 "search points" for Pull Back. This means that the software will look at 41 different points at equal intervals between 450 and 490. This is determined by the factor precision defined in the Design Definition window for the experiment. Sometimes the "search time" is too long. You can reduce the search time by reducing the number of search points for a given factor. In our example, we will reduce the number of search points to 21. A screen similar to the one found in Figure 5-16 will appear.

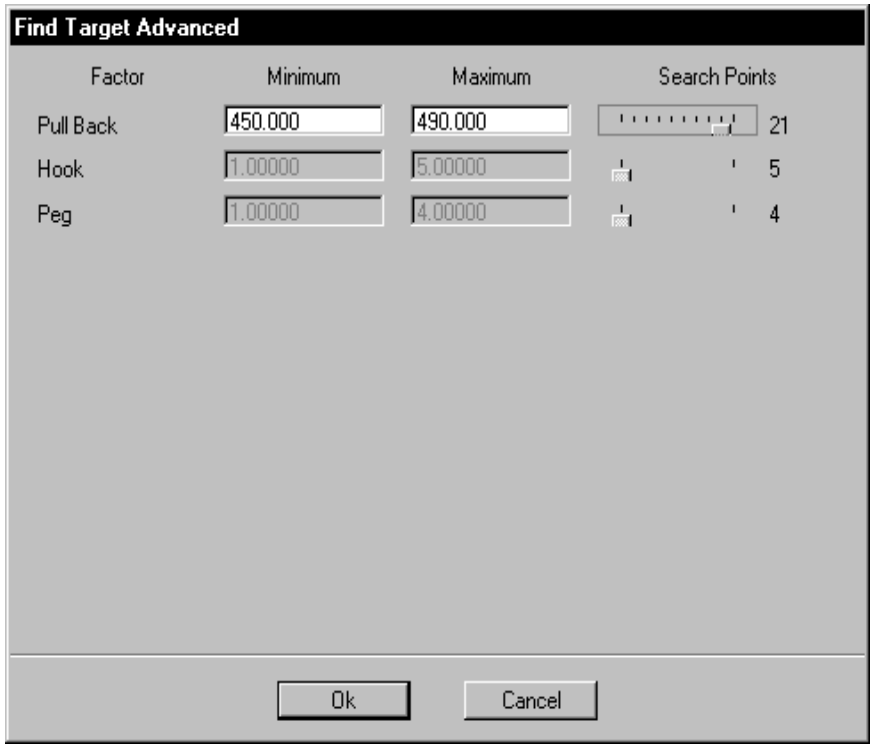

Figure 5-16

Once all factor ranges and number of search points have been defined, click on **OK**. The screen shown in Figure 5-13 will appear. Click on **OK**. The software will now recommend factor settings that will hit the target of 55 with the new factor ranges. A screen similar to the one shown in Figure 5-17 will appear. In this example, the software shows that a Distance value of 55.0969 can be "hit" by setting:

Pull Back = 454
Hook = 1
Peg = 2

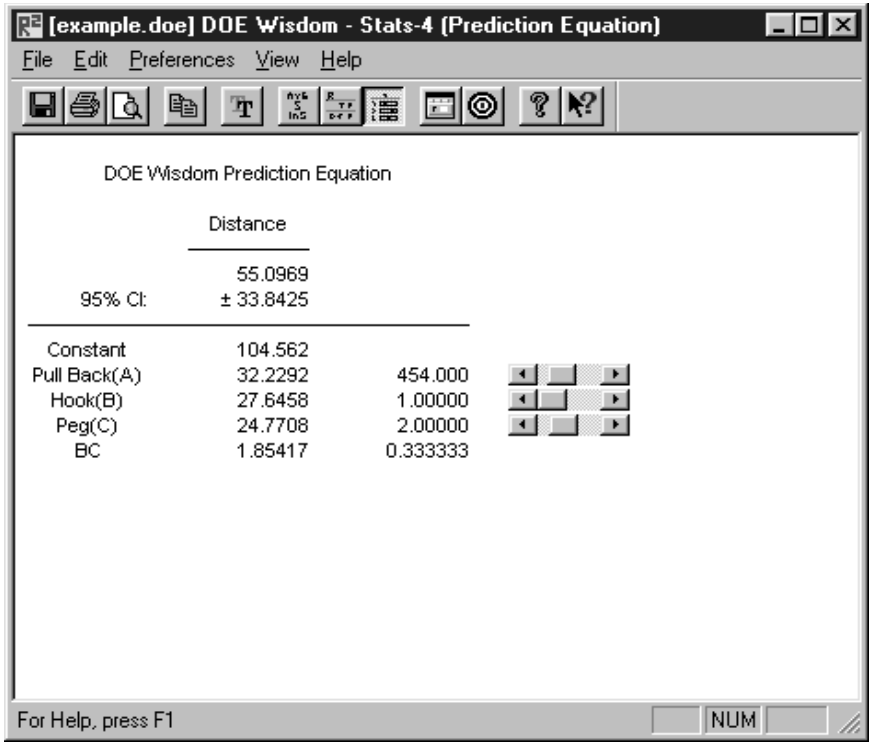

Figure 5-17

Converting Coded Equation to "Real Equation"

DOE Wisdom displays the model in a "coded" equation. This is the most statistically sound approach. Converting the equation to a real equation can be done as follows:

Fundamental Equation:

$\bar{f}_{real}$ = Average of real levels

d_r = Difference between real high and low values

$$f_{orth} = 2\left(\frac{f_{real} - \bar{f}_{real}}{d_r}\right)$$

Example:

Suppose we conducted the following DOE,

Factor	Low	High
Pressure(A)	100	200
Time(B)	10	20

Data obtained for the experiment:

Run	Pressure(A)	Time(B)	AB	Response
1	100 (-)	10 (-)	(+)	10
2	100 (-)	20 (+)	(-)	6
3	200 (+)	10 (-)	(-)	8
4	200 (+)	20 (+)	(+)	2

	Pressure	Time	AB
Avg (+)	5	4	6
Avg (-)	8	9	7
Δ	-3	-5	-1
Δ/2	-1.5	-2.5	-0.5

The equation for the orthogonal scale would be:

$$\hat{y} = 6.5 - 1.5(press) - 2.5(time) - 0.5(press)(time)$$

$$= 6.5 - 1.5(A) - 2.5(B) - 0.5(A)(B)$$

How about the real scale?

$$\text{Pressure} \Rightarrow A_{orth} = 2\left[\frac{A_{real} - \overline{A}_{real}}{d_r}\right] = 2\left[\frac{A_{real} - 150}{100}\right] = \frac{A_{real} - 150}{50}$$

$$\text{Time} \Rightarrow B_{orth} = 2\left[\frac{B_{real} - \overline{B}_{real}}{d_r}\right] = 2\left[\frac{B_{real} - 15}{10}\right] = \frac{B_{real} - 15}{5}$$

Substituting into "Orthogonal" Prediction Equation

$$\hat{y} = 6.5 - 1.5\left[\frac{A_{real} - 150}{50}\right] - 2.5\left[\frac{B_{real} - 15}{5}\right] - 0.5\left[\frac{A_{real} - 150}{50}\right]\left[\frac{B_{real} - 15}{5}\right]$$

$$= 6.5 - 0.03[A_{real}] + 4.5 - 0.5[B_{real}] + 7.5 - 0.002[(A_{real})(B_{real})] + 0.3[B_{real}] + 0.03[A_{real}] - 4.5$$

The software tells us that if we set Pressure to 100 and Time to 10 we will get a response value of 10. Let's check this.

$A_{real} = 100$, $B_{real} = 10$ Does $\hat{y} = 10$?

$= 6.5 - 0.03[100] + 4.5 - 0.5[10] + 7.5 - 0.002[(100)(10)] + 0.3[10] + 0.03[100] - 4.5$

$= 6.5 - 3.0 + 4.5 - 5 + 7.5 - 2.0 + 3.0 + 3.0 - 4.5 = 10$ (The calculation is correct!)

Exporting Statistics

To export statistics from DOE Wisdom to other applications, select the **Copy** or **Copy Regression Results** command from the **Edit** menu. Select the **Paste** command in the other application. When **Copy Regression Results** is selected, the studentized residuals will also be copied.