A Technical Publication for JMP® Users
Winter 2004 Issue 13

JMPer Cable

Inside This Newsletter

Breaking News

Free "JMP for Six Sigma" Webinar: Monday, March 15 at 1 pm ET.

In this webinar, you'll see how JMP can be customized to fit the varied needs and levels of practitioners in your organization and how JMP's graphical interactivity enables everyone to make a contribution to the productivity gains promised by Six Sigma.

Register at http://www.jmp.com/news/regwebinar_form.shtml or call 1-877-594-6567.

Introducing JMP 5.1

John Sall, Executive Vice President, SAS Institute

JMP 5.1, although primarily a refinement release, contains important new functionality. If you have an annual license, there is no charge for this upgrade. There are minimal compatibility issues between JMP 5 and 5.1, so do not hesitate to upgrade and take advantage of the new enhancements and platforms. Additionally, the cost to upgrade individual copies from JMP 5 to JMP 5.1 is small, so you should upgrade if JMP has a central role in your work and you want the latest new features.

The complete JMP 5.1 "New Features" document is located at www.jmp.com/product/jmp51_new_features.pdf.

JMP 5.1.1 (the most current shipping release of JMP 5.1) is available for registered users to download at www.jmp.com.

The following sections highlight some of the most compelling additions in the 5.1 release of JMP.

Linux

As part of JMP 5.1, we launched a Linux version of JMP, which provides the same set of powerful analytical capabilities as JMP for Windows and Macintosh. Additionally, JMP on Linux supports display forwarding so it can be used remotely.

Surface Plots

One of the most spectacular new features in JMP 5.1 is the surface plot (see Figure 1). The addition of surface plots is a huge step forward into three-dimensional visualization. The surface plot far surpasses the original spinning plot. Using the rendering features of OpenGL on Windows, Macintosh, and Linux, surface plots produce real-time, rotatable, three-dimensional rendered surfaces and points. The example graphs in Figure 1 illustrate smooth surfaces.

Among other things, you have complete control of individual axis scaling, color and lighting control, a variety of plot types, and the ability to plot data points, a response formula, or both. Surface plots are great for showing fitted surfaces and are used in a number of other platforms as a surface profiler. If you use the contour profiler to understand response surfaces, then you will see surface plots as a major improvement when you upgrade to JMP 5.1.

Attribute Gage R&R

Using the Variability/Gage Chart platform, you can now evaluate rater agreement through graphs, summary

reports, agreement reports, and kappa statistics for attribute (pass/fail) responses.

Continuous Gage R&R

For the Gage output on continuous variables, JMP now computes variance components for missing or unbalanced data using REML—Restricted (or Residual) Maximum Likelihood. JMP also computes bias and linearity reports for evaluating consistent measuring ability.

Random Effects Models

JMP 5.1 excels at handling random effects that are high-order crossed effects. This is important for complex multi-layer models, such as split plot designs where the whole plots are a fractional factorial.

Data Search

In 5.1, finding (searching for) data in a data table is easier in two ways:

- The new **Select Matching Cells** command selects all cells that contain the same value(s) as those you have currently selected.
- The **Select Where** command accepts multiple selection conditions.

Value Labels

If you have data whose values are codes and you would like to see longer labels associated with each code (or vice versa), then you will like the value labels feature, which is available for either numeric or character columns in JMP 5.1. For example, you might have a table with a yes/no response coded as 1 or 0. Using value labels, the data remains unchanged but reports display "yes" or "no" instead of 1 and 0. Another example might be a list of airport codes, such as LGA, that you want to see printed as "New York-Laguardia" in reports. JMP keeps the data in its original form, but uses alternate labels in the table and reports.

Item Analysis

If you develop multiple-choice tests and need to evaluate how difficult and discriminating each question is, JMP 5.1's item response theory (IRT) implementation will interest you.

Clustering

If you do large scale clustering, such as in genomics studies, then the improvements in the two clustering platforms are valuable.

- Hierarchical clustering supports the magnification of dendrogram sections.
- *K*-Means clustering supports parallel coordinate plots by cluster.

Graphics Platforms

The cell plot platform in JMP 5.1 shows color-coded rectangles for values across many rows and columns. Cell plots, also called heat maps, let you compare values of closely-related points.

The parallel plot platform in JMP 5.1 draws connected lines between column coordinates, presenting a plot of connected line segments between multiple response variables for each row in the table. Parallel plots let you visualize shapes and scales together.

Design of Experiments (DOE)

The space filling design platform is a new DOE platform useful for

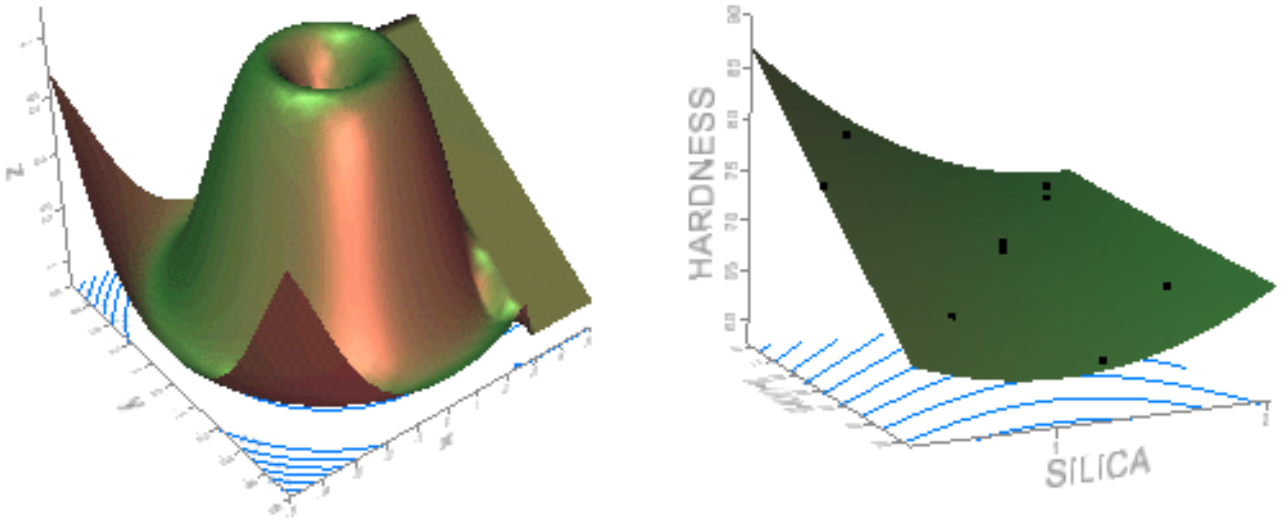


Figure 1: Surface plots in JMP 5.1 display points and surfaces in three dimensions.

modeling deterministic systems where the form of the underlying relationship between input and output variables is unknown but thought to be highly nonlinear.

JSL

There are new JSL (JMP Scripting Language) commands in JMP 5.1, and many refinements in the details. For Windows users, there is a new debugger for JSL. JSL also now supports an interface to the same 3D environment used by surface plots. You may find the demo scripts for this very interesting.

Future

The JMP development team is now working on JMP 6, and we expect its release to occur in 2005. Among other new features and enhancements, we are committed to unicode compliance, which will smoothly move JMP into the international community.

Look for JMP at these Conferences

May 9-12, 2004	SUGI (SAS Users Group International) in Montreal, Canada
May 23-26, 2004	PharmaSUG in San Diego, CA
May 24-26, 2004	American Society for Quality (ASQ)'s Annual Quality Conference (AQC) in Toronto, Canada
June 9-10, 2004	Quality Expo Detroit in Novi, MI
August 8-12, 2004	JSM (Joint Statistical Meetings (JSM) in Toronto, Canada

Tips & Techniques

Using the Add Graphics Script Command

Lee Creighton, JMP Development
Robert Allison, SAS/Graph Central Testing

JMP's graphics and analysis platforms provide a wide range of commonly-used plots and graphs. However, you can use JMP Scripting Language (JSL) to augment existing plots by adding calculated functions, additional decoration, or backgrounds. This article looks at a geographic example.

At elevated levels, mercury (Hg) has been shown to cause poisoning. Fetuses are particularly at risk and can suffer central nervous system damage, mental retardation, or a lack of physical development as a result of high-level mercury exposure. Effects on adults can also be severe and include both sensory and motor skills damage.

Eating fish is a primary vector for the transmission of mercury to humans. Government officials in the state of Maine commissioned a study to sample mercury from selected lakes across the state. Researchers captured varying numbers of fish at each of 120 randomly selected lakes and tested their mercury levels. Regulations state that mercury levels should be less than 0.5 parts per million to be safe.

The data table Maine.jmp, available for download at http://www.jmp.com/news/jmpercable, identifies each lake by name, latitude, and longitude. It also shows the mercury level (Hg level) of each lake. Other information includes the lake type, drainage area, surface area, depth, elevation, and other demographic descriptions. In addition, a column called Safe Level? computes the values Safe or Unsafe with this formula:

$$\text{If}\begin{bmatrix} Hg\ Level < 0.5 \Rightarrow \text{"Safe"} \\ \text{else} \Rightarrow \text{"Unsafe"} \end{bmatrix}$$

JMP's analysis tools make it easy to run a multitude of analyses. For example, the Distribution platform displays a histogram and two tables showing the distribution of mercury level in the sample of lakes. A single outlier, Hodgdon Pond (labeled Hodgdon.P in the data table), stands out in this graphical presentation, shown in Figure 2. It is a good idea to highlight this outlying point and select **Rows > Label** to label the point so it stands out in future analyses.

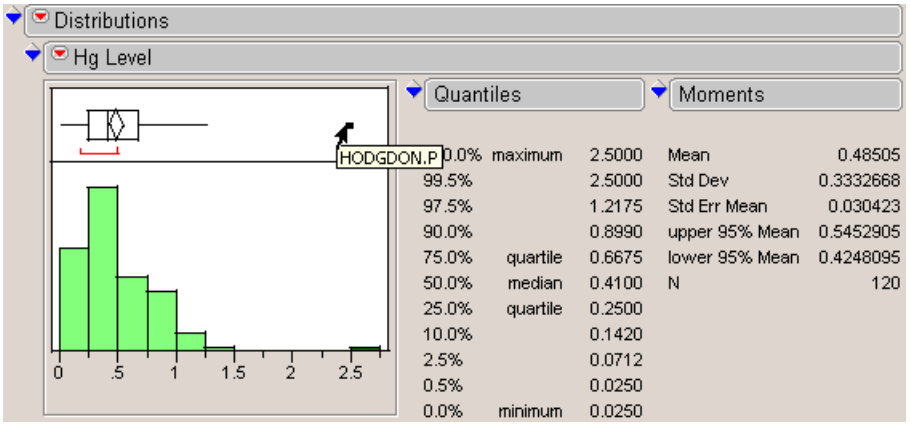


Figure 2: A distribution in JMP showing a single outlier.

You can also investigate if there is a relationship between amount of mercury and the lake elevation or the lake type using **Analyze > Fit Y by X**. The plots in Figure 3 don't show a significant relationship between mercury level and either of these variables possibly because of the outlier.

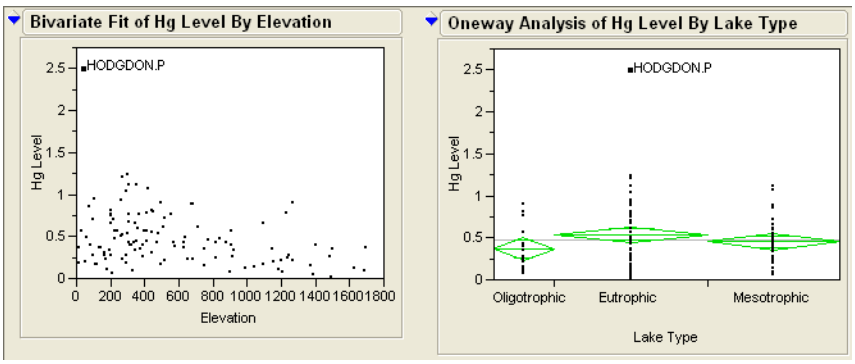


Figure 3: Relationship of mercury level with lake elevation and with lake type.

None of these analyses tell us about the geographic (physical) distribution of Hg. Therefore, it would be useful to plot the data in a geographical layout using the latitude and longitude from the data table. To do so:

1. Select **Analyze > Fit Y By X**.
2. Assign Latitude to **Y** and Longitude to **X**.
3. Click **OK**.
4. When the report appears, select **Rows > Color or Mark by Column**.
5. When the dialog appears, select the Safe Level? variable in the variable list and check the **Set Color by Value** and **Set Marker by Value** check boxes.

These commands produce the plot shown in Figure 4, with plus markers indicating unsafe ponds and square markers indicating safe ponds. By resizing the plot to the approximate proportions of Maine and using a little imagination (and, perhaps, some squinting), you can see the general shape of the state of Maine and the location of the outlier lake, Hodgdon.P, on the coast.

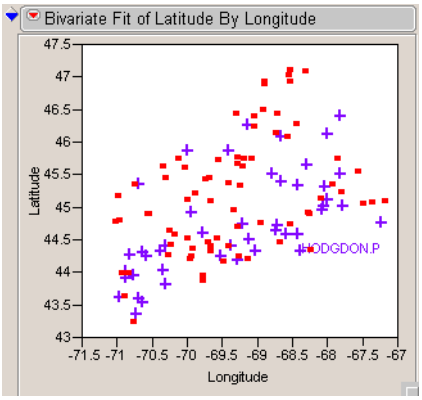


Figure 4: Plot of the data in a geographical layout.

There are several reasons why this particular graphical presentation isn't ideal for making discoveries.

- Although you can see the spatial distribution of the lakes that are safe and unsafe, you can't discern the relative level of mercury in each lake.
- The positions of the lakes are not clear in relation to the borders of the state. A map of Maine shown as the background of the plot would greatly enhance the spatial interpretation.

Adding Mercury Level Information

A useful addition to the graph might be to add markers whose size is based on the mercury level in the lake. The following example displays a circle for each lake. The circles' diameters are proportional to the lakes' mercury level.

First, create a small script by right-clicking the plot and selecting **Add Graphics Script** (Figure 5).

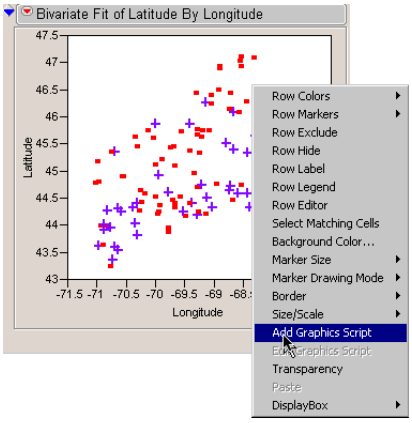


Figure 5: The Add Graphics Script command.

Enter the following script into the dialog box that appears:

```
For Each Row(Circle({:Longitude,
:Latitude}, :Hg Level/10))
```

This JSL command draws a black circle around each point (determined by its longitude and latitude), with the diameter scaled to 1/10 of the mercury level. If you want larger or smaller circles, adjust the number in the denominator of Hg Level in the script.

This script enhances the ability to pick out the lakes with large concentrations of mercury. However, it would be more informative if the circles were drawn in the same color as an assigned row state (color and marker).

To do this, first assign new markers and colors to reflect each lake's type:

1. Right-click the plot and select **Row Legend** from the menu that appears.
2. Select the Lake Type variable in the variable list and check the **Set Color by Value**, **Set Marker by Value**, and **Make Window with Legend** check boxes.
3. Click **OK** to see the new color and marker scheme, as well as a color and marker legend.
4. Right click the script name and select **Edit**. To the existing script, add the Pen Color command: `Pen Color(Color Of(Row State()));`. The script now appears as follows:

```
For Each Row(Pen Color(Color
Of(Row State())); Circle({ :Lon-
gitude, :Latitude}, :Hg Level /
10))
```

This script produces the plot in Figure 6, which displays cirlces in the same color as their respective markers.

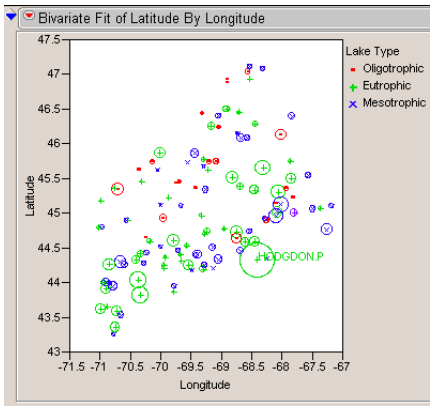


Figure 6: Markers, colors and circles show three types of lakes and mercury levels.

What are the lake types?
The lakes are typed based on feeding relationships. An *oligotropic* lake has a balance between decaying vegetation and living organisms, where the lowest layer of water never loses its oxygen and the water contains few nutrients, but sustains a fish population. A *eutrophic* lake has a high decay rate in the top layer of water and so contains little oxygen at the lowest levels. It has few fish but is rich in algae. A *mesotrophic* lake is between the oligotrophic and the richer eurtrophic state and has a moderate amount of nutrients in its water.

Adding a Map to the Plot (Fun, Advanced Topic)

Adding a map to the plot is not as simple as adding circles. First, you need coordinates for the border points of the state. If the points are not on the same scale as your data, you must scale them to be the same.

Boundary points for all 50 states (as well as a raft of other things) are available from the census department at http://www.census.gov/geo/www/cob/index.html. Unless you have specialized software, the ASCII versions of the map files are appropriate for importing into JMP. Once made into a data table, a formula may be necessary to scale the points to match the longitude and latitude values in the data. Frequently, the left/right orientation must be reversed. An alternative to scaling by eye is to use specialized software like SAS to project the data onto the appropriate scale.

Now, store the latitude and longitude coordinates in a global variable so they can be plotted. A script called MaineLines.jsl, located at http://www.jmp.com/news/jmpercable, takes the boundary values and assigns their *x*-coordinates as a matrix to a global variable called x and the *y*-coordinates as a matrix to a global variable called y. Run the MaineLines.jsl script to load the values into the global variables. In its stripped down form, MaineLines.jsl looks like this:

```
x=[…x coordinates…];
y=[…y coordinates…];
```

Then, add the following graphics script to the beginning of the Maine plot:

```
Fill Color(81); Polygon(x, y);
```

This draws polygons, based on the border coordinates, that display a map of Maine underneath the points.

If you want to draw a border around the gray map, add `Line(x, y);` to the above script. The script now appears as:

```
Line(x, y); Fill Color(81);
Polygon(x, y);
```

This draws a line connecting the boundary points and draws the polygons, as shown in Figure 7. You can change the color of the map by altering the argument of the `Fill Color` function.

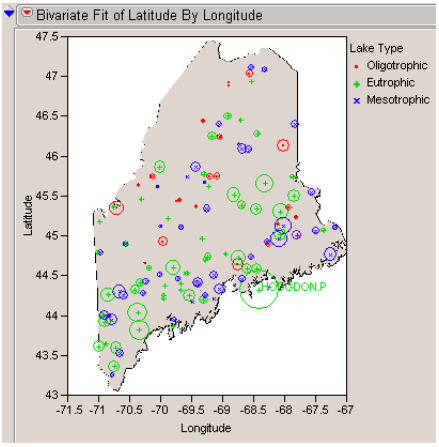


Figure 7: Map fill based on coordinates.

The point representing Hodgdon Pond appears to be in the ocean. In fact, it is a lake on a small island just off the coast of the mainland.

References

Hoeting, J. A. and Olsen, R. (1998) *Are the fish safe to eat? Assessing mercury levels in fish in Maine lakes.* In Peck, R, Haugh, L. D., and Goodman, A. *Statistical case studies: A collaboration between Academe and Industry.* Philadelphia/Alexandria: ASA/SIAM.

Wheeler, M. "Environews Focus." *Environmental health perspectives*, Volume 104, Number 8, August 1996. Available online at http://ehp.niehs.nih.gov/docs/1996/104-8/toc.html.

Teacher Talk

Creating Interactive Demonstrations of Statistical Concepts

Mark Bailey, SAS Statistical Training & Technical Services
Meredith Blackwelder, JMP Development

Some students view statistical science as profound and oftentimes counter-intuitive. Such attitudes might have stemmed from presentations by instructors who gave rigorous mathematical presentations of theories. These presentations may have been correct and complete, but not always engaging for the student, and sometimes essential ideas were lost in the transfer.

In other teaching practices, instructors deliver interactive demonstrations that help students discover topics and principles on their own; "teachers no longer cover the subject; they guide learners to discover it" (Schneiderman 2003). Using JMP to demonstrate important statistical principles is a way instructors can ensure that their teaching methods and materials help students learn by discovery. Students benefit by seeing instructors explain and use a JMP data table, create different types of analyses, and study graphical representations of their data.

In addition to capturing students' attentions and helping them learn by discovery, interactive JMP demonstrations hold advantages for instructors. They are freed from the task of setting up a demonstration by using JMP's pre-packaged meta-data, formulas, and scripts. Instructors can customize layout and format results to match forms in student textbooks.

This article describes four free-of-charge JMP resources that help instructors explain statistical theories in the classroom via interactive JMP demonstrations. All four resources make it possible to implement the demonstrations without learning JMP Scripting Language (JSL).

Sample Data Tables and Built-In Scripts

When you install JMP, a folder named Sample Data is placed your computer. In this folder are JMP data tables, which contain relevant and oftentimes true-to-life data that you can use in your demonstrations. Using the sample data tables, you do not have to compose a data table, create phony data, or arrange the data properly in the data table.

Most of the sample data tables contain built-in scripts that run specific analyses on the data. All you have to do is click the red triangle beside the script name in the upper left corner of the table and select **Run Script**.

Below is a glimpse of some of the sample data tables and the type of data and scripts they contain.

- Big Class.jmp This table contains data on a class of students: their name, height and weight, age, and gender. The table contains scripts that display a simple distribution, show a bivariate fit, a oneway analysis, and a contingency analysis.

- Tiretread.jmp This data table contains data on characteristics of tires, such as their abrasion, elongation, and hardness. One script in the table runs a response surface model on the data. Another creates a neural net analysis.
- Pickles.jmp Use this data table when discussing quality control. This data table contains data on which vat the pickles came from, the acidity each contained, and the time and date the sampling took place. The script stored in this table creates an individual measurement chart, where acidity is the process variable and the date is the sample identification.

The Sample Scripts Folder

In addition to giving you the Sample Data folder, the JMP installer also placed a folder named Sample Scripts on your computer. In this folder are JMP scripts you can use to illustrate statistical terms and principles. For example, the script named demoLeastSquares.jsl demonstrates regression. It shows points that are scattered in a graph and fits a line through the points. By interacting with the graph, you can add visual elements to show how far off the points are from the line. By using these scripts, you can quickly illustrate a concept and interact with the software to demonstrate qualities of different analyses.

JMP Script Library

There's a large collection of scripts in the online JMP Script Library that you can use to demonstrate different statistical concepts. To see the library, go to http://www.jmp.com and mouse over the **Downloads** link on the left, then select **JMP Script Library**.

For example, you can use demoANOVAModel.jsl to give a demonstration that shows two treatment groups; each is replicated a total of three times. The data illustrates "independence" and walks through the decomposition into each term of the model: mean, effect, and residual.

Or, if your class needs to see the parallels between the population mean and standard deviation, the demoSamplingDistribution.jsl script demonstrates these parallels as well as the parallels between the sample mean and standard error.

Collection of Educational Demonstrations

After three years of research, the Statistical Training and Technical Services department at SAS Institute developed a collection of educational demonstrations. The scope of topics in the collection continues to grow, and solutions range from simple to sublime.

Those who assimilated the collection held a workshop at the 2003 Joint Statistical Meetings (JSM) in San Francisco to showcase their use. You can access a book of course notes (and use them as lesson plans), presentation slides from the workshop, and the demonstration collection by going to http://www.jmp.com and clicking **Downloads > JSL Workshop Downloads**, or going directly to http://www.jmp.com/support/training/concepts_workshop/.

Items in the collection of educational demonstrations (the link titled **JMP Data and JSL Files**) include information on how to demonstrate:

- Simple Random Samples
- Sample Statistics and Standard Error
- Gaussian Distribution
- Central Limit Theorem
- Confidence Intervals
- Student t Distribution and Test of the Mean
- Power of Test
- One-way ANOVA Model
- Simple Linear Least Squares Regression

Conclusion

Instructors should take advantage of tools that can help them give effective interactive demonstrations. These demonstrations help students become engaged and excited about learning statistical theories. JMP can be used as a teaching tool to make statistical theories understandable, credible, compelling, experiential, dynamic, directed, and controllable. JMP is designed for research and discovery, and it is perfectly suited as a tool for statistical education. Try it in your classroom and see what positive results it produces.

References

Schneiderman, Ben. *Leonardo's Laptop: Human Needs and the New Computing Technologies*. MIT Press: 2002.

Space Filling Designs for Pre-Screen Experiments

Mark Bailey, SAS Statistical Training & Technical Services

Bradley Jones, JMP Development

JMP 5.1 introduces a new family of experiments called space filling designs. These designs are commonly used to solve the problem of sampling from systems having many factors but exhibiting little or no noise. For example, a computer simulation program always gives the same answer for any set of initial conditions. There is none of the usual experimental error associated with physical experiments. The problem lies in getting the minimum bias in your model by either spreading the design points out as far or as evenly as possible. Space filling designs can address this situation.

To further illustrate the problem described above, think of biological systems where the response can change abruptly. Imagine that you discovered a new enzyme. You expect that catalytic activity depends on the pH of the buffered solution. You know the native cellular pH, but you wonder if this enzyme exhibits activity under different conditions. In a normal screening situation, you would test at the low and high limit of the test range, and perhaps include a center point to test lack of fit. In this case, however, the naïve approach (the standard pH values 1, 7, and 14) would probably fail because the biologically important pH range is usually too narrow to be detected by using only these points. It makes more sense to set up a finer grid of test points, perhaps every half unit of pH, not because you want to fit a high-order polynomial model, but just to locate the region of interest. After you identify this region, you could include pH in a subsequent screening experiment with fewer levels over a narrower range. You might think about the initial study as a *pre-screen* experiment. This approach is a space-filling design for one factor.

A Case Study

The case study in this article uses a new Polymerase Chain Reaction (PCR) to illustrate space-filling designs for pre-screen experiments. This reaction is very complex and sensitive to conditions. (This description of the PCR has been intentionally simplified to illustrate space-filling designs.) The main purpose of this reaction is to make many copies of native DNA. These copies are often subjected to a second reaction (hybridization) to incorporate labels for quantitation. The PCR factors include the raw material, a sample of native DNA, a suitable reaction buffer, and the reaction conditions.

The reaction involves repeatedly heating and cooling the reaction mixture, inducing the DNA to melt and separating the double strands. This process is known as thermal cycling. Each cycle effectively doubles the number of DNA molecules.

Scientists can be guided by experience while choosing reasonable factor ranges, but few combinations of settings work for a particular PCR. As when finding the working range for pH, scientists need to sample a large number of levels over the factor range. Specifically, they must select a number of treatments (combinations of factor levels) to satisfy the same need across multiple factors. The response might appear abruptly in a limited region, or it can form a ridge of sorts instead of the gently changing quadratic shape that is usual for other systems.

For this article, the Latin hypercube space-filling design is used to spread the design points as evenly as possible. The response is the amount of the product DNA fragment that is formed. The ultimate goal of the research is to optimize the production of the fragment. However, because there is now no yield, we are trying to detect any increase.

The DOE Space-Filling Platform

The first step in this process is to select **DOE > Space Filling Design** and specify the response and the factors, as shown in Figure 8.

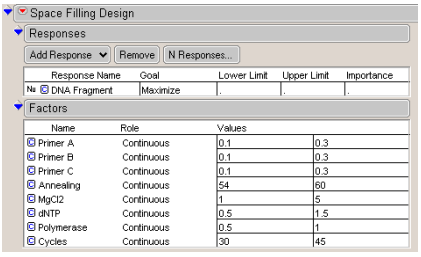


Figure 8: Definition of response and factors.

The primers and the dNTP

nucleotides are the raw material for making new DNA strands. The annealing temperature affects the kinetics of the reaction. Magnesium and polymerase are part of the catalytic reaction. To duplicate this example and quickly add factors, download Space Filling Factors.jmp, found at http://www.jmp.com/news/jmpercable. Then, click the red triangle icon on the title bar of the window, select **Load Factors**, and select the file. Now click **Continue**. JMP reveals the design methods from which you can choose.

A large number of runs ensures that the density of the test points is sufficient for this region. All of the testing could be controlled if the runs fit in a single 96-well plate. Creating 48 runs and a duplicate of those runs provides a measure of experimental error. So, first enter 48 into the **Choose Sample Size** text box, as shown in Figure 9, and click the **Latin Hypercube** button.

No matter what your computer speeed, these designs take time to compute, so please be patient!

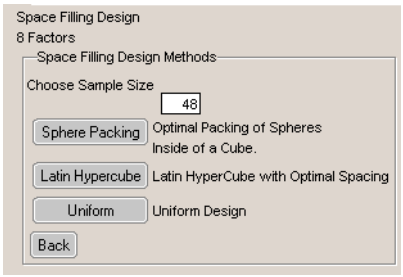


Figure 9: Specify runs and design type at the bottom of the window.

When JMP completes the design, it displays the runs in the DOE window (see Figure 10). Note that the Prediction Variance Profiler is absent, because space filling designs are usually employed when there is little or no experimental error.

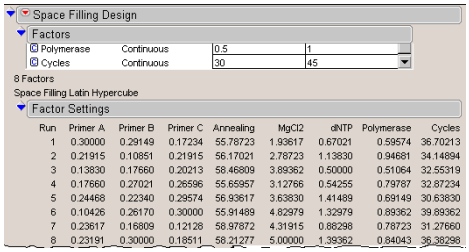


Figure 10: Latin Hypercube factor settings.

Click **Make into Data Table** to save the design in a JMP data table. To add the replicates, select the table, then:

1. Select **Edit > Copy**, which copies the whole table to the clipboard.
2. Select **Edit > Paste**, which appends the duplicate 48 rows to the table.

You can download the table (Space Filling Design.jmp) created in this example at http://www.jmp.com/news/jmpercable.

Analyze the Design

The distribution of these points can be seen in the two-dimensional scatterplot matrix in the Multivariate platform, as shown in Figure 11. To see this scatterplot matrix:

1. Choose **Analyze > Multivariate Methods > Multivariate**.
2. Select all factors (all variables except the response, DNA Fragment), as **Y** and click **OK**.

These projections cannot portray the full dispersion of the points in the hypercube, but they give a sense of even spread of points for each pair of responses. Another good tool for this purpose is the spinning plot, which shows you the point dispersion for any set of three variables (Figure 11).

Note that some of these factors, such as annealing temperature, are hard to change and may represent runs that are different in practice. Others, such as cycles, must be rounded to whole numbers.

The data table of the runs contains a Fit Model script for the additive effects. Such a hyper-plane would probably miss the important feature in this case, and the few points that yield significant amounts of the desired fragment would be outliers with high leverage and influence. This model simply cannot capture or detect

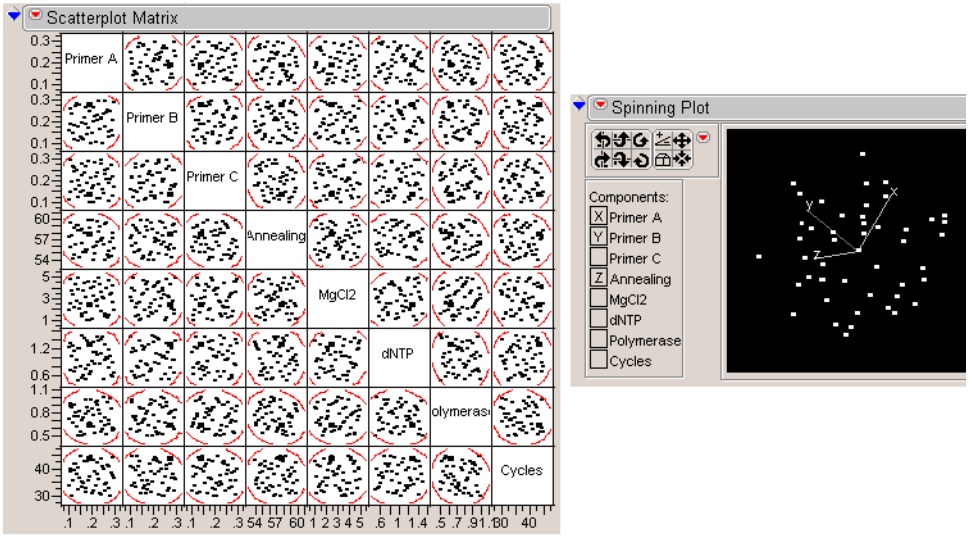


Figure 11: Scatterplot matrix and spinning plot of PCR factor levels.

features in the data that we seek.

Before fitting a model, explore the response of the fragment in a distribution analysis. Select **Analyze > Distribution** and select only the response, DNA Fragment, for the *y* role. Most of the responses, as shown on the left in Figure 12, are between zero and 50. A few more are observed between 50 and 100. There are two replicates of the same design point near 150. What are these conditions? Select these two points by Shift-clicking them in the outlier box plot, then select **Tables > Subset**. They are shown on the right in Figure 12.

The runs in the subset table indicate the conditions that define the origin in a new experiment that uses a more traditional screening design to explore the effects of these factors. The factor ranges are dramatically reduced in the next phase of testing because neighboring points did not perform nearly as well.

Another way to visualize this response is with the surface plot, a new feature in JMP 5.1. To see a surface plot:

1. Choose **Graph > Surface Plot**.
2. In the dialog, select DNA Fragment and click the **Response, Data Points Variable** button.
3. Select all factors and click the **Factors, only if no Formula** button.
4. Click **OK**.
5. Open the Appearance outline section and choose **needles**, giving the plot shown in Figure 13.

To experiment with the surface plot, open the Variables outline section and change the pair of factors involved in the plot. Use the hand tool to rotate and change the plot's orientation. Press Shift and drag the mouse to continually spin the plot.

Conclusion

JMP 5.1 includes an exciting new design platform to address experimental problems with very different characteristics than the ones usually handled by traditional or optimal design algorithms.

Refer to Chapter 7, "Space-Filling Designs," in the new *JMP 5.1 Design of Experiments Guide* to learn more about space filling designs and how to best use this platform.

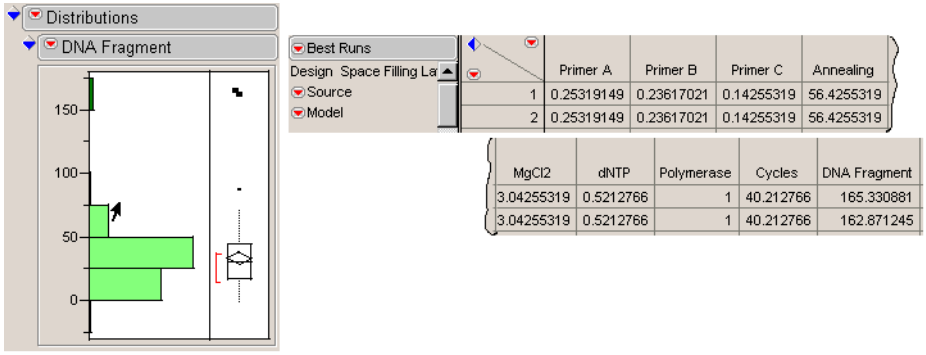


Figure 12: Distribution of DNA Fragment and subset of highest performing runs.

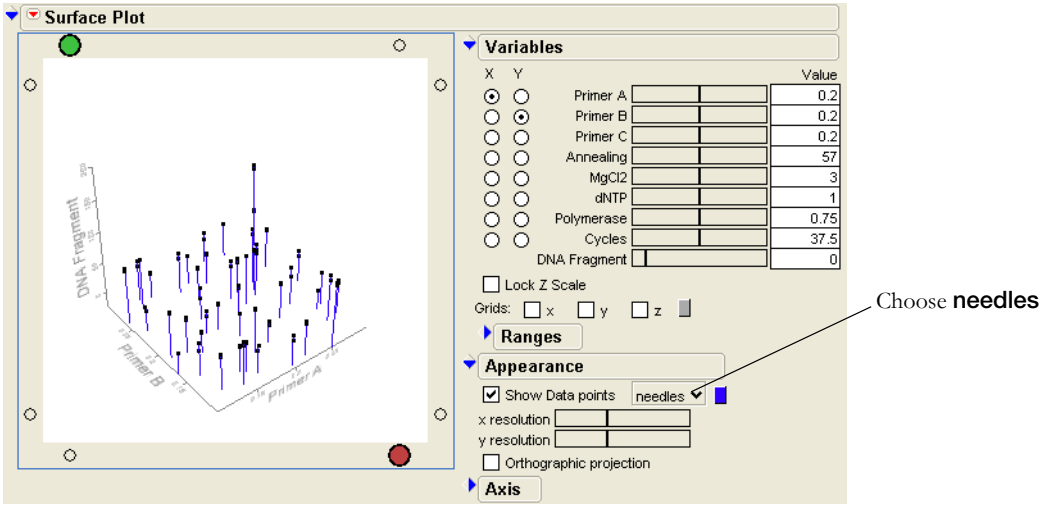
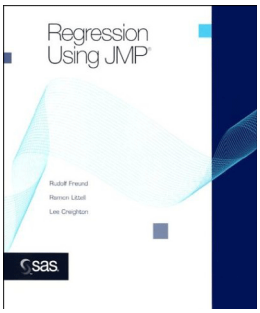


Figure 13: Surface plot showing best response.

Discussions of *Regression Using JMP*



This article presents three reviews of *Regression Using JMP* by Rudolf Freund, Ramon Littell, and Lee Creighton (Wiley Interscience 2004). Each review gives a unique perspective on the book.

Book Blends Theory and Mechanics

Eric Hamann, Principal Biostatistician, Wyeth

Regression Using JMP is a well-written book for the intermediate-level statistician who wants to get more from JMP.

The introductory chapter provides a solid review of basic regression theory. Matrix theory is often omitted from software-related books. It was a pleasant surprise to see it included. Matrix formulas are used throughout.

Many subtle JMP features (regression related or not) are incorporated into the examples. I particularly liked the combination of JMP features used in the presentation of regression diagnostics such as blending the use of overlay plots, colors, and markers with the standard Fit Y by X tools. Most JMP texts and classes limit their presentations solely to the platform being presented.

This book offers a better blend of theory and mechanics than most software-oriented texts. In the explanatory paragraphs, "English" descriptions of the parameters are regularly presented before proceeding with the more theoretical detail. They explain statistical assumptions, pitfalls, and how the newly introduced parameter relates to other terms.

The first five chapters focus on linear regression and identification/remediation of issues with the data. The latter portions of the book focus on a collection of other types of curve-fitting techniques. I was somewhat less enthusiastic about the coverage of the topics in the later section of the book. These sections adequately presented their topics, but significantly cut back on the explanatory text.

Overall, I would recommend this book to my fellow JMP users.

Meets Intermediate Users' Needs

Sharon Field, Kodak

I've been using JMP for about two years. A year ago, I started to use JMP for regression analysis. I wish I had this book when I started.

One of my favorite parts of this book was the easy to understand explanations of statistics. For example, "The objective is to find a single line that best summarizes the data—to find a linear equation that predicts values of the weight given the values of height" (page 3). These types of explanations are scattered through the book, which I found very helpful since I am somewhat new to the field of statistics. I also found the numbered explanations very helpful as well (page 25, for example). The numbering and arrows made the explanations easy to follow. I wish I had that when I was taking my last statistics class.

One of the features of JMP that I find myself using quite a bit for my regression analysis is excluding certain data and running the model. I found the explanation of that very good.

One important issue that always comes into play when I'm doing analysis is what variables do I put into the model. The explanation of the variable selection was terrific. I remember covering stepwise in my statistics course but I never knew how to do it in JMP before I read this book.

Overall I found this book very helpful. There are many great examples and illustrations used throughout this book which is great for a novice user of JMP. These examples and illustrations also make the book easier and more interesting to read. It brings to life a subject that can otherwise be somewhat dry to read about. I will continue to use this book as a resource when I'm doing future analysis.

An Example-Driven Book

Chuck Boiler, JMP Marketing

Freund, Littell, and Creighton ably introduce the key concepts of regression notation using a visual, non-math intensive approach. It is after this introduction that they use traditional mathematical descriptions. They then begin using examples,

(continued from page 11)

which are especially welcome to those who learn first through intuition and images. The authors employ a physical example of fitting using mechanical springs. They use this example as a non-mathematical description of least squares regression. Newcomers to regression could spend rewarding time reviewing the opening chapter and the interactive exercise.

The book also tells us that a web address is maintained by SAS from which example data and code can be downloaded. The text relies on JMP data table examples to introduce each regression topic using JMP software and takes advantage of JMP's interactive graphs to support learning the key concepts. The examples feed on JMP's interactive nature to support learning and exploration.

A simple regression scripting example is included in the final chapter. This example can be adapted by statisticians in the manufacturing environment to automate regression fits in production settings. This code alone might be worth the price of the book.

A welcomed surprise in this volume is a short introduction to Nonlinear fitting in JMP. This section summarizes some of the interactive features that make fitting a nonlinear model with JMP a unique joy. Try the parameter sliders that update live on the graphs!

This is an important companion to the JMP documentation set for anyone who fits linear models with JMP. I have my own personal copy and I think many others would also benefit by having it as a reference.

About JMPer Cable
Issue 13 Winter 2004

JMPer Cable is mailed to JMP users who are registered users with SAS Institute. It is also available online at www.jmp.com.

Contributors
Robert Allison
Mark Bailey
Meredith Blackwelder
Lee Creighton
Bradley Jones
John Sall

Editor
Ann Lehman

Designer
Meredith Blackwelder

Printing
SAS Institute Print Center

Questions or comments
jmp@sas.com

To order JMP software
1-877-594-6567

For more information on JMP
1-877-594-6567

www.jmp.com