



Version 9

Design of Experiments Guide

*“The real voyage of discovery consists not in seeking new
landscapes, but in having new eyes.”*

Marcel Proust

JMP, A Business Unit of SAS
SAS Campus Drive
Cary, NC 27513

9.0.2

The correct bibliographic citation for this manual is as follows: SAS Institute Inc. 2010. *JMP® 9 Design of Experiments Guide*. Cary, NC: SAS Institute Inc.

JMP® 9 Design of Experiments Guide

Copyright © 2010, SAS Institute Inc., Cary, NC, USA

ISBN 978-1-60764-597-9

All rights reserved. Produced in the United States of America.

For a hard-copy book: No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, or otherwise, without the prior written permission of the publisher, SAS Institute Inc.

For a Web download or e-book: Your use of this publication shall be governed by the terms established by the vendor at the time you acquire this publication.

The scanning, uploading, and distribution of this book via the Internet or any other means without the permission of the publisher is illegal and punishable by law. Please purchase only authorized electronic editions and do not participate in or encourage electronic piracy of copyrighted materials. Your support of others' rights is appreciated.

U.S. Government Restricted Rights Notice: Use, duplication, or disclosure of this software and related documentation by the U.S. government is subject to the Agreement with SAS Institute and the restrictions set forth in FAR 52.227-19, Commercial Computer Software-Restricted Rights (June 1987).

SAS Institute Inc., SAS Campus Drive, Cary, North Carolina 27513.

1st printing, September 2010

2nd printing, January 2011

3rd printing, June 2011

SAS® Publishing provides a complete selection of books and electronic products to help customers use SAS software to its fullest potential. For more information about our e-books, e-learning products, CDs, and hard-copy books, visit the SAS Publishing Web site at support.sas.com/publishing or call 1-800-727-3228.

JMP®, SAS® and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are registered trademarks or trademarks of their respective companies.

Get the Most from JMP®

Whether you are a first-time or a long-time user, there is always something to learn about JMP.

Visit JMP.com to find the following:

- live and recorded Webcasts about how to get started with JMP
- video demos and Webcasts of new features and advanced techniques
- schedules for seminars being held in your area
- success stories showing how others use JMP
- a blog with tips, tricks, and stories from JMP staff
- a forum to discuss JMP with other users

<http://www.jmp.com/getstarted/>

Contents

Design of Experiments

1	Introduction to Designing Experiments	
	A Beginner's Tutorial	1
	About Designing Experiments	3
	My First Experiment	3
	The Situation	3
	Step 1: Design the Experiment	3
	Step 2: Define Factor Constraints	5
	Step 3: Add Interaction Terms	6
	Step 4: Determine the Number of Runs	7
	Step 5: Check the Design	8
	Step 6: Gather and Enter the Data	9
	Step 7: Analyze the Results	10
2	Examples Using the Custom Designer	
		15
	Creating Screening Experiments	17
	Creating a Main-Effects-Only Screening Design	17
	Creating a Screening Design to Fit All Two-Factor Interactions	19
	A Compromise Design Between Main Effects Only and All Interactions	21
	Creating 'Super' Screening Designs	23
	Screening Designs with Flexible Block Sizes	27
	Checking for Curvature Using One Extra Run	31
	Creating Response Surface Experiments	34
	Exploring the Prediction Variance Surface	35
	Introducing <i>I</i> -Optimal Designs for Response Surface Modeling	37
	A Three-Factor Response Surface Design	38
	Response Surface with a Blocking Factor	40
	Creating Mixture Experiments	44
	Mixtures Having Nonmixture Factors	44
	Experiments that are Mixtures of Mixtures	48

Special-Purpose Uses of the Custom Designer	51
Designing Experiments with Fixed Covariate Factors	51
Creating a Design with Two Hard-to-Change Factors: Split Plot	55
Technical Discussion	60

3 Building Custom Designs

The Basic Steps	63
Creating a Custom Design	65
Enter Responses and Factors into the Custom Designer	65
Describe the Model	69
Specifying Alias Terms	70
Select the Number of Runs	71
Understanding Design Evaluation	72
Specify Output Options	78
Make the JMP Design Table	79
Creating Random Block Designs	79
Creating Split Plot Designs	80
Creating Split-Split Plot Designs	81
Creating Strip Plot Designs	82
Special Custom Design Commands	83
Save Responses and Save Factors	84
Load Responses and Load Factors	85
Save Constraints and Load Constraints	85
Set Random Seed: Setting the Number Generator	85
Simulate Responses	86
Save X Matrix: Viewing the Number of Rows in the Moments Matrix and the Design Matrix (X) in the Log	87
Optimality Criterion	88
Number of Starts: Changing the Number of Random Starts	88
Sphere Radius: Constraining a Design to a Hypersphere	90
Disallowed Combinations: Accounting for Factor Level Restrictions	90
Advanced Options for the Custom Designer	91
Save Script to Script Window	93
Assigning Column Properties	93
Define Low and High Values (DOE Coding) for Columns	94
Set Columns as Factors for Mixture Experiments	95
Define Response Column Values	96

Assign Columns a Design Role	97
Identify Factor Changes Column Property	98
How Custom Designs Work: Behind the Scenes	99
4 Screening Designs	
Screening Design Examples	101
Using Two Continuous Factors and One Categorical Factor	103
Using Five Continuous Factors	105
Creating a Screening Design	109
Enter Responses	109
Enter Factors	110
Choose a Design	111
Display and Modify a Design	115
Specify Output Options	119
View the Design Table	120
Create a Plackett-Burman design	120
Analysis of Screening Data	122
Using the Screening Analysis Platform	123
Using the Fit Model Platform	124
5 Response Surface Designs	
A Box-Behnken Design: The Tennis Ball Example	127
The Prediction Profiler	132
A Response Surface Plot (Contour Profiler)	134
Geometry of a Box-Behnken Design	136
Creating a Response Surface Design	136
Enter Responses and Factors	137
Choose a Design	138
Specify Output Options	139
View the Design Table	140
6 Full Factorial Designs	
The Five-Factor Reactor Example	143
Analyze the Reactor Data	146
Creating a Factorial Design	151

Enter Responses and Factors	151
Select Output Options	152
Make the Table	152

7 Mixture Designs

.....	155
Mixture Design Types	157
The Optimal Mixture Design	157
The Simplex Centroid Design	158
Creating the Design	159
Simplex Centroid Design Examples	160
The Simplex Lattice Design	161
The Extreme Vertices Design	163
Creating the Design	164
An Extreme Vertices Example with Range Constraints	164
An Extreme Vertices Example with Linear Constraints	166
Extreme Vertices Method: How It Works	167
The ABCD Design	168
Creating Ternary Plots	168
Fitting Mixture Designs	170
Whole Model Tests and Analysis of Variance Reports	171
Understanding Response Surface Reports	171
A Chemical Mixture Example	171
Create the Design	171
Analyze the Mixture Model	174
The Prediction Profiler	175
The Mixture Profiler	176
A Ternary Plot of the Mixture Response Surface	177

8 Discrete Choice Designs

.....	179
Introduction	181
Create an Example Choice Experiment	183
Analyze the Example Choice Experiment	186
Design a Choice Experiment Using Prior Information	188
Administer the Survey and Analyze Results	190

Initial Choice Platform Analysis	191
Find Unit Cost and Trade Off Costs with the Profiler	192

9 Space-Filling Designs

.....	195
Introduction to Space-Filling Designs	197
Sphere-Packing Designs	197
Creating a Sphere-Packing Design	197
Visualizing the Sphere-Packing Design	199
Latin Hypercube Designs	200
Creating a Latin Hypercube Design	200
Visualizing the Latin Hypercube Design	202
Uniform Designs	204
Comparing Sphere-Packing, Latin Hypercube, and Uniform Methods	206
Minimum Potential Designs	207
Maximum Entropy Designs	209
Gaussian Process IMSE Optimal Designs	211
Borehole Model: A Sphere-Packing Example	211
Create the Sphere-Packing Design for the Borehole Data	212
Guidelines for the Analysis of Deterministic Data	213
Results of the Borehole Experiment	213

10 Accelerated Life Test Designs

Designing Experiments for Accelerated Life Tests	217
Overview of Accelerated Life Test Designs	219
Using the ALT Design Platform	219
Platform Options	224
Example	224

11 Nonlinear Designs

.....	229
Examples of Nonlinear Designs	231
Using Nonlinear Fit to Find Prior Parameter Estimates	231
Creating a Nonlinear Design with No Prior Data	237
Creating a Nonlinear Design	241
Identify the Response and Factor Column with Formula	241

Set Up Factors and Parameters in the Nonlinear Design Dialog	242
Enter the Number of Runs and Preview the Design	243
Make Table or Augment the Table	244
Advanced Options for the Nonlinear Designer	245

12 Taguchi Designs

.....	247
The Taguchi Design Approach	249
Taguchi Design Example	249
Analyze the Data	252
Creating a Taguchi Design	254
Detail the Response and Add Factors	254
Choose Inner and Outer Array Designs	255
Display Coded Design	256
Make the Design Table	257

13 Augmented Designs

.....	259
A <i>D</i> -Optimal Augmentation of the Reactor Example	261
Analyze the Augmented Design	264
Creating an Augmented Design	271
Replicate a Design	272
Add Center Points	275
Creating a Foldover Design	276
Adding Axial Points	277
Adding New Runs and Terms	278
Special Augment Design Commands	281
Save the Design (X) Matrix	282
Modify the Design Criterion (<i>D</i> - or <i>I</i> - Optimality)	282
Select the Number of Random Starts	283
Specify the Sphere Radius Value	283
Disallow Factor Combinations	284

14 Prospective Sample Size and Power

.....	285
Launching the Sample Size and Power Platform	287
One-Sample and Two-Sample Means	287
Single-Sample Mean	289

Sample Size and Power Animation for One Mean	292
Two-Sample Means	293
k-Sample Means	294
One Sample Standard Deviation	296
One Sample Standard Deviation Example	297
One-Sample and Two-Sample Proportions	298
One Sample Proportion	298
Two Sample Proportions	300
Counts per Unit	303
Counts per Unit Example	304
Sigma Quality Level	305
Sigma Quality Level Example	305
Number of Defects Computation Example	306
Reliability Test Plan and Demonstration	306
Reliability Test Plan	306
Reliability Demonstration	309
Index	
Design of Experiments	317

Credits and Acknowledgments

Origin

JMP was developed by SAS Institute Inc., Cary, NC. JMP is not a part of the SAS System, though portions of JMP were adapted from routines in the SAS System, particularly for linear algebra and probability calculations. Version 1 of JMP went into production in October 1989.

Credits

JMP was conceived and started by John Sall. Design and development were done by John Sall, Chung-Wei Ng, Michael Hecht, Richard Potter, Brian Corcoran, Annie Dudley Zangi, Bradley Jones, Craig Hales, Chris Gotwalt, Paul Nelson, Xan Gregg, Jianfeng Ding, Eric Hill, John Schroedl, Laura Lancaster, Scott McQuiggan, Melinda Thielbar, Clay Barker, Peng Liu, Dave Barbour, Jeff Polzin, John Ponte, and Steve Amerige.

In the SAS Institute Technical Support division, Duane Hayes, Wendy Murphrey, Rosemary Lucas, Win LeDinh, Bobby Riggs, Glen Grimme, Sue Walsh, Mike Stockstill, Kathleen Kiernan, and Liz Edwards provide technical support.

Nicole Jones, Kyoko Keener, Hui Di, Joseph Morgan, Wenjun Bao, Fang Chen, Susan Shao, Michael Crotty, Jong-Seok Lee, Tonya Mauldin, Audrey Ventura, Ani Eloyan, Bo Meng, and Sequola McNeill provide ongoing quality assurance. Additional testing and technical support are provided by Noriki Inoue, Kyoko Takenaka, Yusuke Ono, Masakazu Okada, and Naohiro Masukawa from SAS Japan.

Bob Hickey and Jim Borek are the release engineers.

The JMP books were written by Ann Lehman, Lee Creighton, John Sall, Bradley Jones, Erin Vang, Melanie Drake, Meredith Blackwelder, Diane Perhac, Jonathan Gatlin, Susan Conaghan, and Sheila Loring, with contributions from Annie Dudley Zangi and Brian Corcoran. Creative services and production was done by SAS Publications. Melanie Drake implemented the Help system.

Jon Weisz and Jeff Perkinson provided project management. Also thanks to Lou Valente, Ian Cox, Mark Bailey, and Malcolm Moore for technical advice.

Thanks also to Georges Guirguis, Warren Sarle, Gordon Johnston, Duane Hayes, Russell Wolfinger, Randall Tobias, Robert N. Rodriguez, Ying So, Warren Kuhfeld, George MacKensie, Bob Lucas, Warren Kuhfeld, Mike Leonard, and Padraic Neville for statistical R&D support. Thanks are also due to Doug Melzer, Bryan Wolfe, Vincent DelGobbo, Biff Beers, Russell Gonsalves, Mitchel Soltys, Dave Mackie, and Stephanie Smith, who helped us get started with SAS Foundation Services from JMP.

Acknowledgments

We owe special gratitude to the people that encouraged us to start JMP, to the alpha and beta testers of JMP, and to the reviewers of the documentation. In particular we thank Michael Benson, Howard Yetter (d),

Andy Mauromoustakos, Al Best, Stan Young, Robert Muenchen, Lenore Herzenberg, Ramon Leon, Tom Lange, Homer Hegedus, Skip Weed, Michael Emptage, Pat Spagan, Paul Wenz, Mike Bowen, Lori Gates, Georgia Morgan, David Tanaka, Zoe Jewell, Sky Alibhai, David Coleman, Linda Blazek, Michael Friendly, Joe Hockman, Frank Shen, J.H. Goodman, David Iklé, Barry Hembree, Dan Obermiller, Jeff Sweeney, Lynn Vanatta, and Kris Ghosh.

Also, we thank Dick DeVeaux, Gray McQuarrie, Robert Stine, George Fraction, Avigdor Cahaner, José Ramirez, Gudmunder Axelsson, Al Fulmer, Cary Tuckfield, Ron Thisted, Nancy McDermott, Veronica Czitrom, Tom Johnson, Cy Wegman, Paul Dwyer, DaRon Huffaker, Kevin Norwood, Mike Thompson, Jack Reese, Francois Mainville, and John Wass.

We also thank the following individuals for expert advice in their statistical specialties: R. Hocking and P. Spector for advice on effective hypotheses; Robert Mee for screening design generators; Roselinde Kessels for advice on choice experiments; Greg Piepel, Peter Goos, J. Stuart Hunter, Dennis Lin, Doug Montgomery, and Chris Nachtsheim for advice on design of experiments; Jason Hsu for advice on multiple comparisons methods (not all of which we were able to incorporate in JMP); Ralph O'Brien for advice on homogeneity of variance tests; Ralph O'Brien and S. Paul Wright for advice on statistical power; Keith Muller for advice in multivariate methods, Harry Martz, Wayne Nelson, Ramon Leon, Dave Trindade, Paul Tobias, and William Q. Meeker for advice on reliability plots; Lijian Yang and J.S. Marron for bivariate smoothing design; George Milliken and Yuri Bulavski for development of mixed models; Will Potts and Cathy Maahs-Fladung for data mining; Clay Thompson for advice on contour plotting algorithms; and Tom Little, Damon Stoddard, Blanton Godfrey, Tim Clapp, and Joe Ficalora for advice in the area of Six Sigma; and Josef Schmea and Alan Bowman for advice on simulation and tolerance design.

For sample data, thanks to Patrice Strahle for Pareto examples, the Texas air control board for the pollution data, and David Coleman for the pollen (eureka) data.

Translations

Trish O'Grady coordinates localization. Special thanks to Noriki Inoue, Kyoko Takenaka, Masakazu Okada, Naohiro Masukawa, and Yusuke Ono (SAS Japan); and Professor Toshiro Haga (retired, Tokyo University of Science) and Professor Hirohiko Asano (Tokyo Metropolitan University) for reviewing our Japanese translation; François Bergeret for reviewing the French translation; Bertram Schäfer and David Meintrup (consultants, StatCon) for reviewing the German translation; Patrizia Omodei, Maria Scaccabarozzi, and Letizia Bazzani (SAS Italy) for reviewing the Italian translation; RuiQi Qiao, Rula Li, and Molly Li for reviewing Simplified Chinese translation (SAS R&D Beijing); Finally, thanks to all the members of our outstanding translation and engineering teams.

Past Support

Many people were important in the evolution of JMP. Special thanks to David DeLong, Mary Cole, Kristin Nauta, Aaron Walker, Ike Walker, Eric Gjertsen, Dave Tilley, Ruth Lee, Annette Sanders, Tim Christensen, Eric Wasserman, Charles Soper, Wenjie Bao, and Junji Kishimoto. Thanks to SAS Institute quality assurance by Jeanne Martin, Fouad Younan, and Frank Lassiter. Additional testing for Versions 3 and 4 was done by Li Yang, Brenda Sun, Katrina Hauser, and Andrea Ritter.

Also thanks to Jenny Kendall, John Hansen, Eddie Routten, David Schlotzhauer, and James Mulherin. Thanks to Steve Shack, Greg Weier, and Maura Stokes for testing JMP Version 1.

Thanks for support from Charles Shipp, Harold Gugel (d), Jim Winters, Matthew Lay, Tim Rey, Rubin Gabriel, Brian Ruff, William Lisowski, David Morganstein, Tom Esposito, Susan West, Chris Fehily, Dan Chilko, Jim Shook, Ken Bodner, Rick Blahunka, Dana C. Aultman, and William Fehlner.

Technology License Notices

Scintilla is Copyright 1998-2003 by Neil Hodgson <neilh@scintilla.org>. NEIL HODGSON DISCLAIMS ALL WARRANTIES WITH REGARD TO THIS SOFTWARE, INCLUDING ALL IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS, IN NO EVENT SHALL NEIL HODGSON BE LIABLE FOR ANY SPECIAL, INDIRECT OR CONSEQUENTIAL DAMAGES OR ANY DAMAGES WHATSOEVER RESULTING FROM LOSS OF USE, DATA OR PROFITS, WHETHER IN AN ACTION OF CONTRACT, NEGLIGENCE OR OTHER TORTIOUS ACTION, ARISING OUT OF OR IN CONNECTION WITH THE USE OR PERFORMANCE OF THIS SOFTWARE.

XRender is Copyright © 2002 Keith Packard. KEITH PACKARD DISCLAIMS ALL WARRANTIES WITH REGARD TO THIS SOFTWARE, INCLUDING ALL IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS, IN NO EVENT SHALL KEITH PACKARD BE LIABLE FOR ANY SPECIAL, INDIRECT OR CONSEQUENTIAL DAMAGES OR ANY DAMAGES WHATSOEVER RESULTING FROM LOSS OF USE, DATA OR PROFITS, WHETHER IN AN ACTION OF CONTRACT, NEGLIGENCE OR OTHER TORTIOUS ACTION, ARISING OUT OF OR IN CONNECTION WITH THE USE OR PERFORMANCE OF THIS SOFTWARE.

ImageMagick software is Copyright © 1999-2010 ImageMagick Studio LLC, a non-profit organization dedicated to making software imaging solutions freely available.

bzlib software is Copyright © 1991-2009, Thomas G. Lane, Guido Vollbeding. All Rights Reserved.

FreeType software is Copyright © 1996-2002 The FreeType Project (www.freetype.org). All rights reserved.

Chapter 1

Introduction to Designing Experiments **A Beginner's Tutorial**

This tutorial chapter introduces you to the design of experiments (DOE) using JMP's custom designer. It gives a general understanding of how to design an experiment using JMP. Refer to subsequent chapters in this book for more examples and procedures on how to design an experiment for your specific project.

Contents

About Designing Experiments.	3
My First Experiment	3
The Situation	3
Step 1: Design the Experiment	3
Step 2: Define Factor Constraints	5
Step 3: Add Interaction Terms.	6
Step 4: Determine the Number of Runs	7
Step 5: Check the Design	8
Step 6: Gather and Enter the Data	9
Step 7: Analyze the Results	10

About Designing Experiments

Increasing productivity and improving quality are important goals in any business. The methods for determining *how* to increase productivity and improve quality are evolving. They have changed from costly and time-consuming trial-and-error searches to the powerful, elegant, and cost-effective statistical methods that JMP provides.

Designing experiments in JMP is centered around factors, responses, a model, and runs. JMP helps you determine if and how a factor affects a response.

My First Experiment

If you have never used JMP to design an experiment, this section shows you how to design the experiment and how to understand JMP's output.

Tip: The recommended way to create an experiment is to use the custom designer. JMP also provides classical designs for use in textbook situations.

The Situation

Your goal is to find the best way to microwave a bag of popcorn. Because you have some experience with this, it is easy to decide on reasonable ranges for the important factors:

- how long to cook the popcorn (between 3 and 5 minutes)
- what level of power to use on the microwave oven (between settings 5 and 10)
- which brand of popcorn to use (Top Secret or Wilbur)

When a bag of popcorn is popped, most of the kernels pop, but some remain unpopped. You prefer to have all (or nearly all) of the kernels popped and no (or very few) unpopped kernels. Therefore, you define “the best popped bag” based on the ratio of popped kernels to the total number of kernels.

A good way to improve any procedure is to conduct an experiment. For each experimental run, JMP's custom designer determines which brand to use, how long to cook each bag in the microwave and what power setting to use. Each *run* involves popping one bag of corn. After popping a bag, enter the total number of kernels and the number of popped kernels into the appropriate row of a JMP data table. After doing all the experimental runs, use JMP's model fitting capabilities to do the data analysis. Then, you can use JMP's profiling tools to determine the optimal settings of popping time, power level, and brand.

Step 1: Design the Experiment

The first step is to select **DOE > Custom Design**. Then, define the responses and factors.

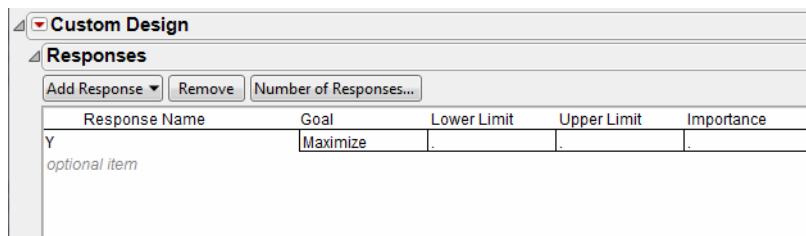
Define the Responses: Popped Kernels and Total Kernels

There are two responses in this experiment:

- the number of popped kernels
- the total number of kernels in the bag. After popping the bag add the number of unpopped kernels to the number of popped kernels to get the total number of kernels in the bag.

By default, the custom designer contains one response labeled Y (Figure 1.1).

Figure 1.1 Custom Design Responses Panel

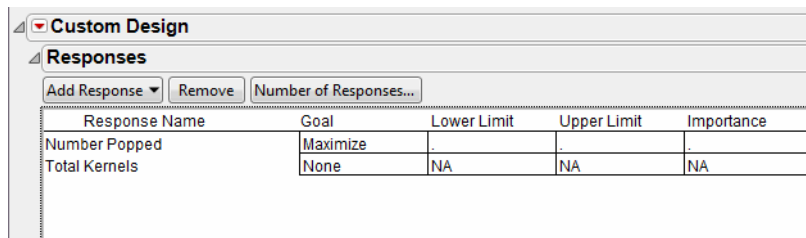


You want to add a second response to the Responses panel and change the names to be more descriptive:

1. To rename the Y response, double-click the name and type “Number Popped.” Since you want to increase the number of popped kernels, leave the goal at **Maximize**.
2. To add the second response (total number of kernels), click **Add Response** and choose **None** from the menu that appears. JMP labels this response Y2 by default.
3. Double-click Y2 and type “Total Kernels” to rename it.

The completed Responses panel looks like Figure 1.2.

Figure 1.2 Renamed Responses with Specified Goals

**Define the Factors: Time, Power, and Brand**

In this experiment, the factors are:

- brand of popcorn (Top Secret or Wilbur)
- cooking time for the popcorn (3 or 5 minutes)
- microwave oven power level (setting 5 or 10)

In the Factors panel, add **Brand** as a two-level categorical factor:

1. Click **Add Factor** and select **Categorical > 2 Level**.
2. To change the name of the factor (currently named X1), type **Brand**.
3. To rename the default levels (L1 and L2), click the level names and type **Top Secret** and **Wilbur**.

Add **Time** as a two-level continuous factor:

4. Click **Add Factor** and select **Continuous**.
5. Change the default name of the factor (X2) by typing **Time**.
6. Likewise, to rename the default levels (-1 and 1) as 3 and 5, click the current level name and type in the new value.

Add **Power** as a two-level continuous factor:

7. Click **Add Factor** and select **Continuous**.
8. Change the name of the factor (currently named X3) by typing **Power**.
9. Rename the default levels (currently named -1 and 1) as 5 and 10 by clicking the current name and typing. The completed Factors panel looks like Figure 1.3.

Figure 1.3 Renamed Factors with Specified Values

Name	Role	Changes	Values
▼ Brand	Categorical	Easy	Top Secret Wilbur
▲ Time	Continuous	Easy	3 5
▲ Power	Continuous	Easy	5 10

10. Click **Continue**.

Step 2: Define Factor Constraints

The popping time for this experiment is either 3 or 5 minutes, and the power settings on the microwave are 5 and 10. From experience, you know that

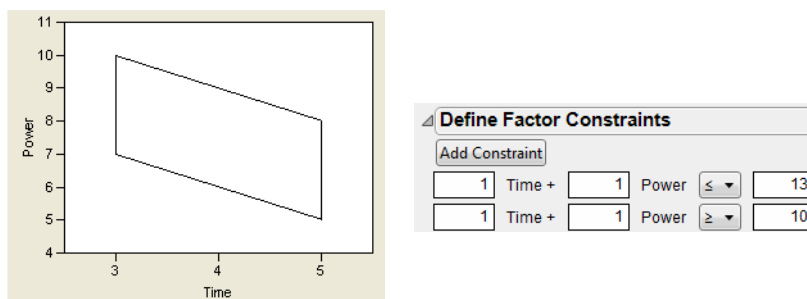
- popping corn for a long time on a high setting tends to scorch kernels.
- not many kernels pop when the popping time is brief and the power setting is low.

You want to constrain the combined popping time and power settings to be less than or equal to 13, but greater than or equal to 10. To define these limits:

1. Open the Constraints panel by clicking the disclosure button beside the **Define Factor Constraints** title bar (see Figure 1.4).
2. Click the **Add Constraint** button twice, once for each of the known constraints.
3. Complete the information, as shown to the right in Figure 1.4. These constraints tell the Custom Designer to avoid combinations of **Power** and **Time** that sum to less than 10 and more than 13. Be sure to change \leq to \geq in the second constraint.

The area inside the parallelogram, illustrated on the left in Figure 1.4, is the allowable region for the runs. You can see that popping for 5 minutes at a power of 10 is not allowed and neither is popping for 3 minutes at a power of 5.

Figure 1.4 Defining Factor Constraints



Step 3: Add Interaction Terms

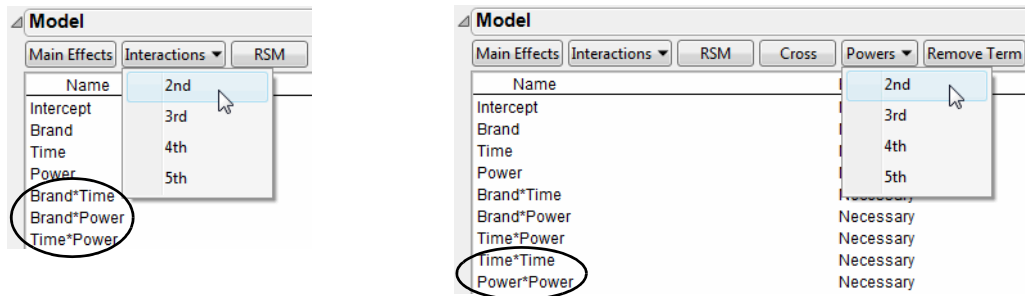
You are interested in the possibility that the effect of any factor on the proportion of popped kernels may depend on the value of some other factor. For example, the effect of a change in popping time for the Wilbur popcorn brand could be larger than the same change in time for the Top Secret brand. This kind of synergistic effect of factors acting in concert is called a *two-factor interaction*. You can examine all possible two-factor interactions in your *a priori* model of the popcorn popping process.

1. Click **Interactions** in the Model panel and select **2nd**. JMP adds two-factor interactions to the model as shown to the left in Figure 1.5.

In addition, you suspect the graph of the relationship between any factor and any response might be curved. You can see whether this kind of curvature exists with a quadratic model formed by adding the second order powers of effects to the model, as follows.

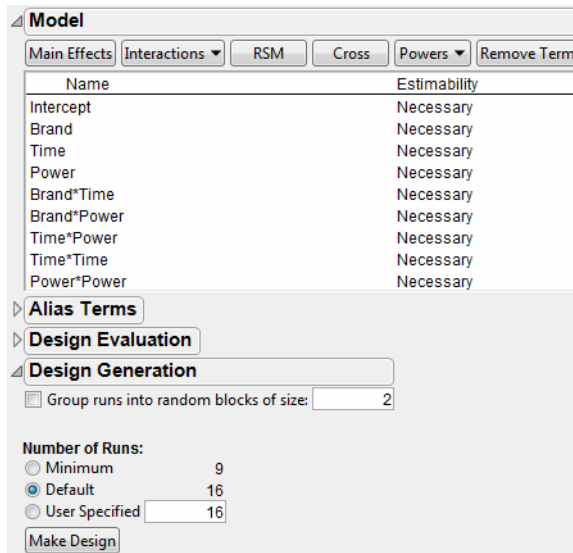
2. Click **Powers** and select **2nd** to add quadratic effects of the continuous factors, **Power** and **Time**.

The completed Model should look like the one to the right in Figure 1.5.

Figure 1.5 Add Interaction and Power Terms to the Model

Step 4: Determine the Number of Runs

The Design Generation panel in Figure 1.6 shows the minimum number of runs needed to perform the experiment with the effects you've added to the model. You can use that minimum or the default number of runs, or you can specify your own number of runs as long as that number is more than the minimum. JMP has no restrictions on the number of runs you request. For this example, use the default number of runs, 16. Click **Make Design** to continue.

Figure 1.6 Model and Design Generation Panels

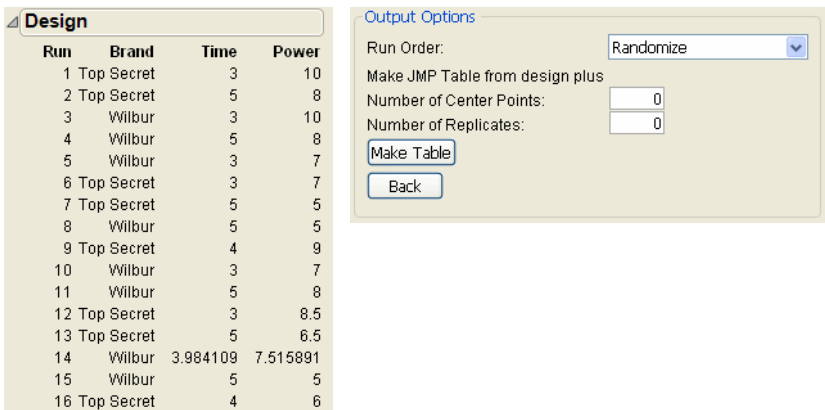
Step 5: Check the Design

When you click **Make Design**, JMP generates and displays a design, as shown on the left in Figure 1.7. Note that because JMP uses a random seed to generate custom designs and there is no unique optimal design for this problem, your table may be different than the one shown here. You can see in the table that the custom design requires 8 runs using each brand of popcorn.

Scroll to the bottom of the Custom Design window and look at the Output Options area (shown to the right in Figure 1.7. The Run Order option lets you designate the order you want the runs to appear in the data table when it is created. Keep the selection at **Randomize** so the rows (runs) in the output table appear in a random order.

Now click **Make Table** in the Output Options section.

Figure 1.7 Design and Output Options Section of Custom Designer



The resulting data table (Figure 1.8) shows the order in which you should do the experimental runs and provides columns for you to enter the number of popped and total kernels. Note that the design matrix is updated to match the order of runs, the Time and Power values, and the Number Popped and Total Kernels columns are added.

You do not have fractional control over the power and time settings on a microwave oven, so you should round the power and time settings, as shown in the data table. Although this altered design is slightly less optimal than the one the custom designer suggested, the difference is negligible.

Tip: Note that optionally, before clicking **Make Table** in the Output Options, you could select **Sort Left to Right** in the Run Order menu to have JMP present the results in the data table according to the brand. We have conducted this experiment for you and placed the results, called **Popcorn DOE Results.jmp**, in the Sample Data folder installed with JMP. These results have the columns sorted from left to right.

Figure 1.8 JMP Data Table of Design Runs Generated by Custom Designer

Custom Design		Brand	Time	Power	Number Popped	Total Kernels
Design	Custom Design	1 Wilbur	3	7	*	*
Criterion	D Optimal	2 Top Secret	3	7	*	*
Screening		3 Top Secret	5	5	*	*
Model		4 Wilbur	3	10	*	*
Constraint		5 Wilbur	3	7	*	*
DOE Dialog		6 Top Secret	3	9	*	*
		7 Top Secret	5	8	*	*
		8 Wilbur	5	5	*	*
		9 Top Secret	5	7	*	*
Columns (5/0)		10 Top Secret	3	10	*	*
Brand *		11 Top Secret	4	9	*	*
Time *		12 Top Secret	4	6	*	*
Power *		13 Wilbur	5	8	*	*
Number Popped *		14 Wilbur	4	8	*	*
Total Kernels *		15 Wilbur	5	5	*	*
		16 Wilbur	5	8	*	*

Step 6: Gather and Enter the Data

Pop the popcorn according to the design JMP provided. Then, count the number of popped and unpopped kernels left in each bag. Finally, enter the numbers shown below into the appropriate columns of the data table.

We have conducted this experiment for you and placed the results in the Sample Data folder installed with JMP. To see the results, open Popcorn DOE Results.jmp from the Design Experiment folder in the sample data. The data table is shown in Figure 1.9.

Figure 1.9 Results of the Popcorn DOE Experiment

Popcorn DOE Resu...		Brand	Time	Power	Number Popped	Total Kernels
Design	Custom Design	1 Top Secret	3	7	30	420
Criterion	D Optimal	2 Top Secret	3	7	120	420
Screening		3 Top Secret	3	10	120	400
Model		4 Top Secret	4	8	250	430
Constraint		5 Top Secret	5	5	370	420
DOE Dialog		6 Top Secret	5	5	400	420
		7 Top Secret	5	8	350	420
		8 Top Secret	5	8	370	380
		9 Wilbur	3	7	374	400
Columns (5/0)		10 Wilbur	3	8	340	440
Brand *		11 Wilbur	3	10	20	410
Time *		12 Wilbur	4	6	440	460
Power *		13 Wilbur	4	9	340	400
Number Popped *		14 Wilbur	5	5	170	370
Total Kernels *		15 Wilbur	5	7	370	380
		16 Wilbur	5	8	420	450

scripts to
analyze data

results from
experiment

Step 7: Analyze the Results

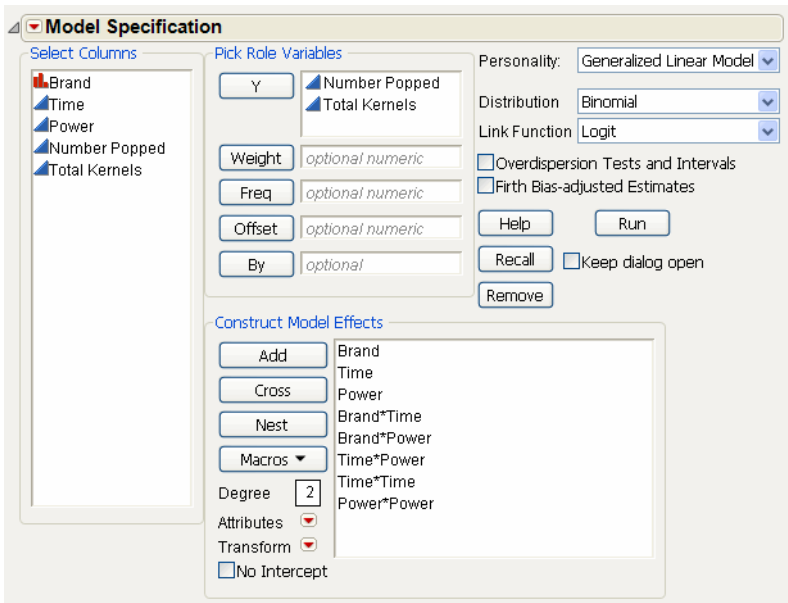
After the experiment is finished and the number of popped kernels and total kernels have been entered into the data table, it is time to analyze the data. The design data table has a script, labeled **Model**, that shows in the top left panel of the table. When you created the design, a standard least squares analysis was stored in the **Model** script with the data table.

1. Click the red triangle for **Model** and select **Run Script**.

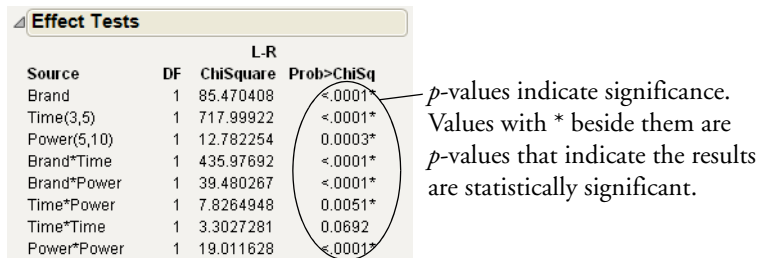
The default fitting personality in the model dialog is **Standard Least Squares**. One assumption of standard least squares is that your responses are normally distributed. But because you are modeling the proportion of popped kernels it is more appropriate to assume that your responses come from a binomial distribution. You can use this assumption by changing to a generalized linear model.

2. Change the Personality to **Generalized Linear Model**, Distribution to **Binomial**, and Link Function to **Logit**, as shown in Figure 1.10.

Figure 1.10 Fitting the Model

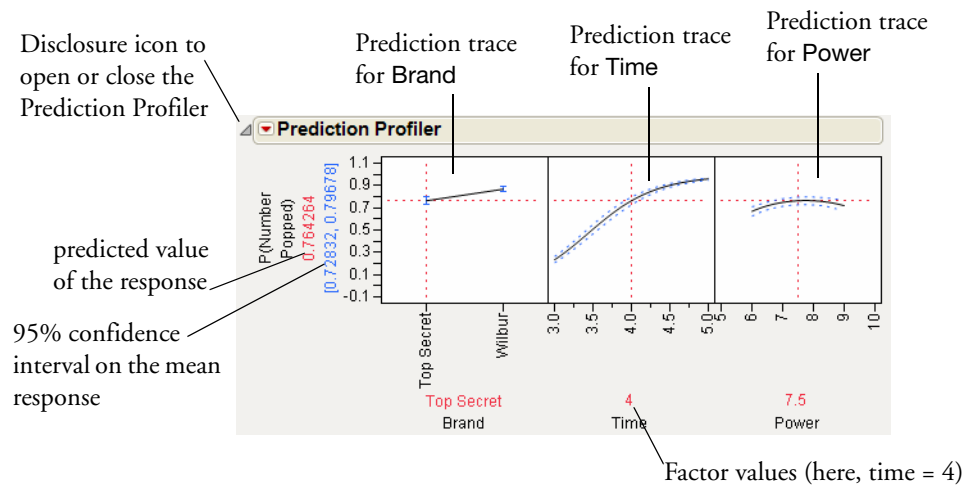


3. Click **Run**.
4. Scroll down to view the Effect Tests table (Figure 1.11) and look in the column labeled Prob>Chisq. This column lists p -values. A low p -value (a value less than 0.05) indicates that results are statistically significant. There are asterisks that identify the low p -values. You can therefore conclude that, in this experiment, all the model effects except for Time*Time are highly significant. You have confirmed that there is a strong relationship between popping time (Time), microwave setting (Power), popcorn brand (Brand), and the proportion of popped kernels.

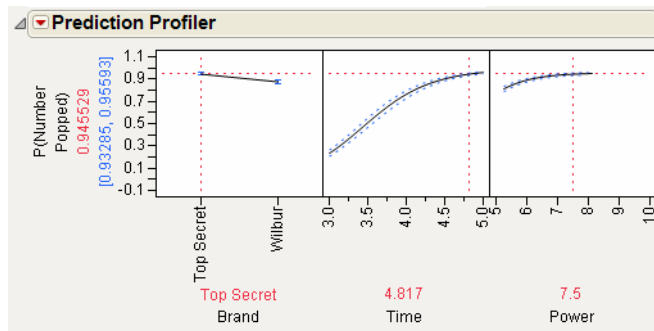
Figure 1.11 Investigating p -Values

To further investigate, use the Prediction Profiler to see how changes in the factor settings affect the numbers of popped and unpopped kernels:

1. Choose **Profilers > Profiler** from the red triangle menu on the Generalized Linear Model Fit title bar. The Prediction Profiler is shown at the bottom of the report. Figure 1.12 shows the Prediction Profiler for the popcorn experiment. Prediction traces are displayed for each factor.

Figure 1.12 The Prediction Profiler

2. Move the vertical red dotted lines to see the effect that changing a factor value has on the response. For example, drag the red line in the Time graph to the right and left (Figure 1.13).

Figure 1.13 Moving the Time Value from 4 to Near 5

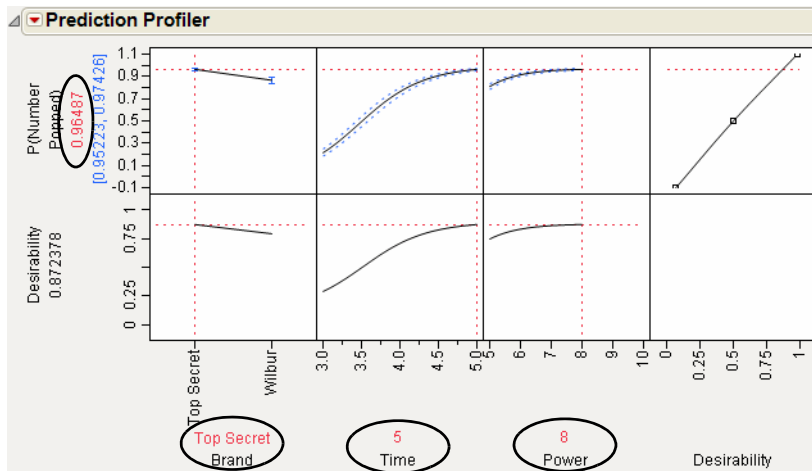
As Time increases and decreases, the curved Brand and Power prediction traces shift their slope and maximum/minimum values. The substantial slope shift tells you there is an interaction (synergistic effect) between Time and Brand and Time and Power.

Furthermore, the steepness of a prediction trace reveals a factor's importance. Because the prediction trace for Time is steeper than that for Brand or Power for the values shown in Figure 1.13, you can predict that cooking time is more important than the brand of popcorn or the microwave power setting.

Now for the final steps.

3. Click the red triangle icon in the Prediction Profiler title bar and select **Desirability Functions**.
4. Click the red triangle icon in the Prediction Profiler title bar and select **Maximize Desirability**. JMP automatically adjusts the graph to display the optimal settings at which the most kernels will be popped (Figure 1.14).

Our experiment found how to cook the bag of popcorn with the greatest proportion of popped kernels: use Top Secret, cook for five minutes, and use a power level of 8. The experiment predicts that cooking at these settings will yield greater than 96.5% popped kernels.

Figure 1.14 The Most Desirable Settings

The best settings are the Top Secret brand, cooking time at 5, and power set at 8.

Examples Using the Custom Designer

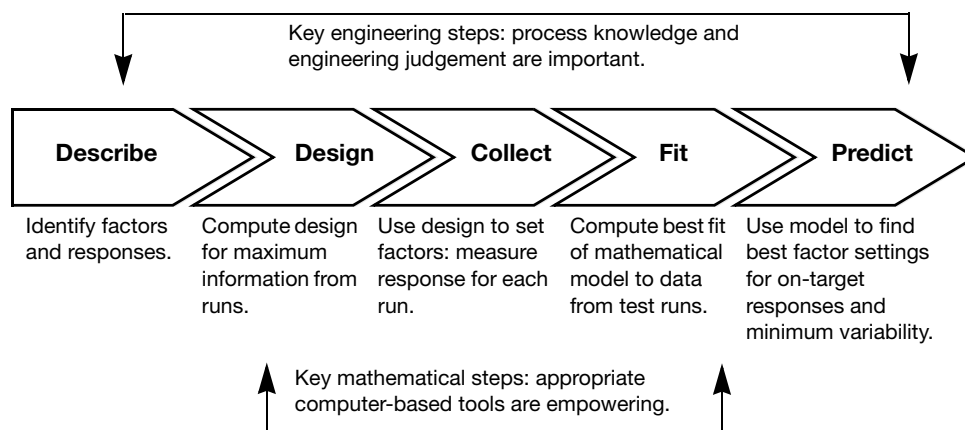
The use of statistical methods in industry is increasing. Arguably, the most cost-beneficial of these methods for quality and productivity improvement is statistical design of experiments. A trial-and-error search for the *vital few* factors that most affect quality is costly and time-consuming. The purpose of *experimental design* is to characterize, predict, and then improve the behavior of any system or process. Designed experiments are a cost-effective way to accomplish these goals.

JMP's custom designer is the recommended way to describe your process and create a design that works for your situation. To use the custom designer, you first enter the process variables and constraints, then JMP tailors a design to suit your unique case. This approach is more general and requires less experience and expertise than previous tools supporting the statistical design of experiments.

Custom designs accommodate any number of factors of any type. You can also control the number of experimental runs. This makes custom design more flexible and more cost effective than alternative approaches.

This chapter presents several examples showing the use of custom designs. It shows how to drive its interface to build a design using this easy step-by-step approach:

Figure 2.1 Approach to Experimental Design



Contents



Creating Screening Experiments	17
Creating a Main-Effects-Only Screening Design	17
Creating a Screening Design to Fit All Two-Factor Interactions	19
A Compromise Design Between Main Effects Only and All Interactions	21
Creating ‘Super’ Screening Designs	23
Screening Designs with Flexible Block Sizes	27
Checking for Curvature Using One Extra Run	31
Creating Response Surface Experiments	34
Exploring the Prediction Variance Surface	35
Introducing I-Optimal Designs for Response Surface Modeling	37
A Three-Factor Response Surface Design	38
Response Surface with a Blocking Factor	40
Creating Mixture Experiments	44
Mixtures Having Nonmixture Factors	44
Experiments that are Mixtures of Mixtures	48
Special-Purpose Uses of the Custom Designer	51
Designing Experiments with Fixed Covariate Factors	51
Creating a Design with Two Hard-to-Change Factors: Split Plot	55
Technical Discussion	60

Creating Screening Experiments

You can use the screening designer in JMP to create screening designs, but the custom designer is more flexible and general. The straightforward screening examples described below show that ‘custom’ does not mean ‘exotic.’ The custom designer is a general purpose design environment that can create screening designs.

Creating a Main-Effects-Only Screening Design

To create a main-effects-only screening design using the custom designer:

1. Select **DOE > Custom Design**.
2. Enter six continuous factors into the Factors panel (see “[Step 1: Design the Experiment](#),” p. 3, for details). Figure 2.2 shows the six factors.
3. Click **Continue**. The default model contains only the main effects.
4. Using the default of eight runs, click **Make Design**. Click the disclosure button ( on Windows and  on the Macintosh) to open the **Design Evaluation** outline node.

Note to DOE experts: The result is a resolution-three screening design. All the main effects are estimable, but they are confounded with two factor interactions.

Figure 2.2 A Main-Effects-Only Screening Design

Custom Design

Responses

Factors

Name	Role	Changes	Values
X1	Continuous	Easy	-1 1
X2	Continuous	Easy	-1 1
X3	Continuous	Easy	-1 1
X4	Continuous	Easy	-1 1
X5	Continuous	Easy	-1 1
X6	Continuous	Easy	-1 1

Define Factor Constraints

Model

Main Effects Interactions RSM Cross Powers Remove Term

Name	Estimability
Intercept	Necessary
X1	Necessary
X2	Necessary
X3	Necessary
X4	Necessary
X5	Necessary
X6	Necessary

Alias Terms

Design Generation

Group runs into random blocks of size: 2

Number of Runs:

Minimum 7
☒ Default 8
 User Specified 8

Make Design

Design Evaluation

Prediction Variance Profile
 Fraction of Design Space Plot
 Prediction Variance Surface
 Relative Variance of Coefficients
 Alias Matrix
 Color Map On Correlations
 Design Diagnostics

Output Options

Run Order: Randomize

Make JMP Table from design plus

Number of Center Points: 0
 Number of Replicates: 0

Make Table

Back

5. Click the disclosure buttons beside Design Evaluation and then beside Alias Matrix (▶ on Windows and ▶ on the Macintosh) to open the Alias Matrix. Figure 2.3 shows the Alias Matrix, which is a table of zeros, ones, and negative ones.

The Alias Matrix shows how the coefficients of the constant and main effect terms in the model are biased by any active two-factor interaction effects not already added to the model. The column labels identify interactions. For example, the columns labeled X2*X6 and X3*X4 in the table have a 1 and -1 in the row for X1. This means that the expected value of the main effect of X1 is actually the sum of the main effect of X1 and X2*X6, minus the effect of X3*X4. You are assuming that these interactions are negligible in size compared to the effect of X1.

Figure 2.3 The Alias Matrix

Alias Matrix

Effect	X1*X2	X1*X3	X1*X4	X1*X5	X1*X6	X2*X3	X2*X4	X2*X5	X2*X6	X3*X4	X3*X5	X3*X6	X4*X5	X4*X6	X5*X6
Intercept	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
X1	0	0	0	0	0	0	0	0	1	-1	0	0	0	0	0
X2	0	0	0	0	0	1	0	0	0	0	-1	0	0	0	0
X3	0	0	-1	0	0	0	0	-1	0	0	0	0	0	0	0
X4	0	-1	0	0	0	0	0	0	0	0	0	0	0	0	1
X5	0	0	0	0	0	0	-1	0	0	0	0	0	0	1	0
X6	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0

Note to DOE experts: The Alias matrix is a generalization of the confounding pattern in fractional factorial designs.

Creating a Screening Design to Fit All Two-Factor Interactions

There is risk involved in designs for main effects only. The risk is that two-factor interactions, if they are strong, can confuse the results of such experiments. To avoid this risk, you can create experiments resolving all the two-factor interactions.

Note to DOE experts: The result in this example is a resolution-five screening design. Two-factor interactions are estimable but are confounded with three-factor interactions.

1. Select **DOE > Custom Design**.
2. Enter five continuous factors into the Factors panel (see “[Step 1: Design the Experiment](#),” p. 3 in the “Introduction to Designing Experiments” chapter for details).
3. Click **Continue**.
4. In the Model panel, select **Interactions > 2nd**.
5. In the Design Generation Panel choose Minimum for **Number of Runs** and click **Make Design**.

Figure 2.4 shows the runs of the two-factor design with all interactions. The sample size, 16 (a power of two) is large enough to fit all the terms in the model. The values in your table may be different from those shown below.

Figure 2.4 All Two-Factor Interactions

Model

Main Effects Interactions RSM Cross Powers Remove Term

Name	Estimability
Intercept	Necessary
X1	Necessary
X2	Necessary
X3	Necessary
X4	Necessary
X5	Necessary
X1*X2	Necessary
X1*X3	Necessary
X1*X4	Necessary
X1*X5	Necessary
X2*X3	Necessary
X2*X4	Necessary
X2*X5	Necessary
X3*X4	Necessary
X3*X5	Necessary
X4*X5	Necessary

Alias Terms

Design Generation

☐ Group runs into random blocks of size: 2

Number of Runs:

☒ Minimum 16

☐ Default 32

☐ User Specified 16

Make Design

Design

Run	X1	X2	X3	X4	X5
1	-1	1	-1	1	-1
2	1	-1	1	1	1
3	1	1	1	-1	1
4	-1	-1	1	-1	1
5	1	-1	-1	1	-1
6	1	-1	1	-1	-1
7	-1	-1	1	1	-1
8	-1	1	1	1	1
9	-1	-1	-1	-1	-1
10	-1	1	1	-1	-1
11	-1	-1	-1	1	1
12	1	1	1	1	-1
13	1	-1	-1	-1	1
14	1	1	-1	-1	-1
15	-1	1	-1	-1	1
16	1	1	-1	1	1

Design Evaluation

- Prediction Variance Profile
- Fraction of Design Space Plot
- Prediction Variance Surface
- Relative Variance of Coefficients
- Alias Matrix
- Color Map On Correlations
- Design Diagnostics

Open outline nodes

6. Click the disclosure button (▶ on Windows and ▶▼ on the Macintosh) and to open the **Design Evaluation** outlines, then open **Alias Matrix**. Figure 2.5 shows the alias matrix table of zeros and ones. The columns labels identify an interaction. For example, the column labelled X1*X2 refers to the interaction of the first and second effect, the column labelled X2*X3 refers to the interaction between the second and third effect, and so forth.

Look at the column labelled X1*X2. There is only one value of 1 in that column. All others are 0. The 1 occurs in the row labelled X1*X2. All the other rows and columns are similar. This means that the expected value of the two-factor interaction X1*X2 is not biased by any other terms. All the rows above the row labelled X1*X2 contain only zeros, which means that the Intercept and main effect terms are not biased by any two-factor interactions.

Figure 2.5 Alias Matrix Showing all Two-Factor Interactions Clear of all Main Effects

Alias Matrix											
Effect	X1*X2	X1*X3	X1*X4	X1*X5	X2*X3	X2*X4	X2*X5	X3*X4	X3*X5	X4*X5	
Intercept	0	0	0	0	0	0	0	0	0	0	
X1	0	0	0	0	0	0	0	0	0	0	
X2	0	0	0	0	0	0	0	0	0	0	
X3	0	0	0	0	0	0	0	0	0	0	
X4	0	0	0	0	0	0	0	0	0	0	
X5	0	0	0	0	0	0	0	0	0	0	
X1*X2	1	0	0	0	0	0	0	0	0	0	
X1*X3	0	1	0	0	0	0	0	0	0	0	
X1*X4	0	0	1	0	0	0	0	0	0	0	
X1*X5	0	0	0	1	0	0	0	0	0	0	
X2*X3	0	0	0	0	1	0	0	0	0	0	
X2*X4	0	0	0	0	0	1	0	0	0	0	
X2*X5	0	0	0	0	0	0	1	0	0	0	
X3*X4	0	0	0	0	0	0	0	1	0	0	
X3*X5	0	0	0	0	0	0	0	0	1	0	
X4*X5	0	0	0	0	0	0	0	0	0	1	

A Compromise Design Between Main Effects Only and All Interactions

In a screening situation, suppose there are six continuous factors and resources for $n = 16$ runs. The first example in this section showed an eight-run design that fit all the main effects. With six factors, there are 15 possible two-factor interactions. The minimum number of runs that could fit the constant, six main effects and 15 two-factor interactions is 22. This is more than the resource budget of 16 runs. It would be good to find a compromise between the main-effects only design and a design capable of fitting all the two-factor interactions.

This example shows how to obtain such a design compromise using the custom designer.

1. Select **DOE > Custom Design**.
2. Define six continuous factors (X1 - X6).
3. Click **Continue**. The model includes the main effect terms by default. The default estimability of these terms is **Necessary**.
4. Click the **Interactions** button and choose **2nd** to add all the two-factor interactions.
5. Select all the interaction terms and click the current estimability (**Necessary**) to reveal a menu. Change **Necessary** to **If Possible**, as shown in Figure 2.6.

Figure 2.6 Model for Six-Variable Design with Two-Factor Interactions Designated If Possible

Custom Design

Responses

Factors

Define Factor Constraints

Model

Main Effects Interactions RSM Cross Powers Remove Term

Name	Estimability
Intercept	Necessary
X1	Necessary
X2	Necessary
X3	Necessary
X4	Necessary
X5	Necessary
X6	Necessary
X1*X2	If Possible
X1*X3	If Possible
X1*X4	If Possible
X1*X5	If Possible
X1*X6	If Possible
X2*X3	If Possible
X2*X4	If Possible
X2*X5	If Possible
X2*X6	If Possible
X3*X4	If Possible
X3*X5	If Possible
X3*X6	If Possible
X4*X5	If Possible
X4*X6	If Possible
X5*X6	If Possible

Alias Terms

Design Generation

☐ Group runs into random blocks of size: 2

Number of Runs:

☐ Minimum 7

☐ Default 8

☒ User Specified 16

Make Design

6. Type 16 in the **User Specified** edit box in the **Number of Runs** section, as shown. Although the desired number of runs (16) is less than the total number of model terms, the custom designer builds a design to estimate as many two-factor interactions as possible.
7. Click **Make Design**.

After the custom designer creates the design, click the disclosure button beside **Design Evaluation** to open the Alias Matrix (Figure 2.7). The values in your table may be different from those shown below, but with a similar pattern.

Figure 2.7 Alias Matrix

Alias Matrix																
Effect	X1*X2	X1*X3	X1*X4	X1*X5	X1*X6	X2*X3	X2*X4	X2*X5	X2*X6	X3*X4	X3*X5	X3*X6	X4*X5	X4*X6	X5*X6	
Intercept	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
X1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
X2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
X3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
X4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
X5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
X6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
X1*X2	0.333	0	0	0	0	0	0	0	0	0	0.333	0	0	0	0	
X1*X3	0	0.25	0	0	0	0	0	0.25	0	0	0	0	0	0.25	0	
X1*X4	0	0	0.333	0	0	0	0	0	0	0	0	0.333	0	0	0	
X1*X5	0	0	0	0.333	0	0.333	0	0	0	0	0	0	0	0	0	
X1*X6	0	0	0	0	0.333	0	0	0	0	0.333	0	0	0	0	0	
X2*X3	0	0	0	0.333	0	0.333	0	0	0	0	0	0	0	0	0	
X2*X4	0	0	0	0	0	0	0.333	0	0	0	0	0	0	0	0.333	
X2*X5	0	0.25	0	0	0	0	0	0.25	0	0	0	0	0	0.25	0	
X2*X6	0	0	0	0	0	0	0	0	0.333	0	0	0	0.333	0	0	
X3*X4	0	0	0	0	0.333	0	0	0	0	0.333	0	0	0	0	0	
X3*X5	0.333	0	0	0	0	0	0	0	0	0	0.333	0	0	0	0	
X3*X6	0	0	0.333	0	0	0	0	0	0	0	0	0.333	0	0	0	
X4*X5	0	0	0	0	0	0	0	0	0.333	0	0	0	0.333	0	0	
X4*X6	0	0.25	0	0	0	0	0	0.25	0	0	0	0	0	0.25	0	
X5*X6	0	0	0	0	0	0	0.333	0	0	0	0	0	0	0	0.333	

All the rows above the row labelled X1*X2 contain only zeros, which means that the Intercept and main effect terms are not biased by any two-factor interactions. The row labelled X1*X2 has the value 0.333 in the X1*X2 column and the same value in the X3*X5 column. That means the expected value of the estimate for X1*X2 is actually the sum of X1*X2 and any real effect due to X3*X5.

Note to DOE experts: The result in this particular example is a resolution-four screening design. Two-factor interactions are estimable but are aliased with other two-factor interactions.

Creating ‘Super’ Screening Designs

This section shows how to use the technique shown in the previous example to create ‘super’ (supersaturated) screening designs. Supersaturated designs have fewer runs than factors, which makes them attractive for factor screening when there are many factors and experimental runs are expensive.

In a saturated design, the number of runs equals the number of model terms. In a supersaturated design, as the name suggests, the number of model terms exceeds the number of runs (Lin, 1993). A supersaturated design can examine dozens of factors using fewer than half as many runs as factors.

The Need for Supersaturated Designs

The 2^{7-4} and the 2^{15-11} fractional factorial designs available using the screening designer are both saturated with respect to a main effects model. In the analysis of a saturated design, you can (barely) fit the model, but there are no degrees of freedom for error or for lack of fit. Until recently, saturated designs represented the limit of efficiency in designs for screening.

Factor screening relies on the *sparsity* principle. The experimenter expects that only a few of the factors in a screening experiment are active. The problem is not knowing which are the vital few factors and which are the trivial many. It is common for brainstorming sessions to turn up dozens of factors. Yet, in practice, screening experiments rarely involve more than ten factors. What happens to winnow the list from dozens to ten or so?

If the experimenter is limited to designs that have more runs than factors, then dozens of factors translate into dozens of runs. Often, this is not economically feasible. The result is that the factor list is reduced without the benefit of data. In a supersaturated design, the number of model terms exceeds the number of runs, and you can examine dozens of factors using less than half as many runs.

There are drawbacks:

- If the number of *active* factors approaches the number of runs in the experiment, then it is likely that these factors will be impossible to identify. A rule of thumb is that the number of runs should be at least four times larger than the number of active factors. If you expect that there might be as many as five active factors, you should have at least 20 runs.
- Analysis of supersaturated designs cannot yet be reduced to an automatic procedure. However, using forward stepwise regression is reasonable and the Screening platform (**Analyze > Modeling > Screening**) offers a more streamlined analysis.

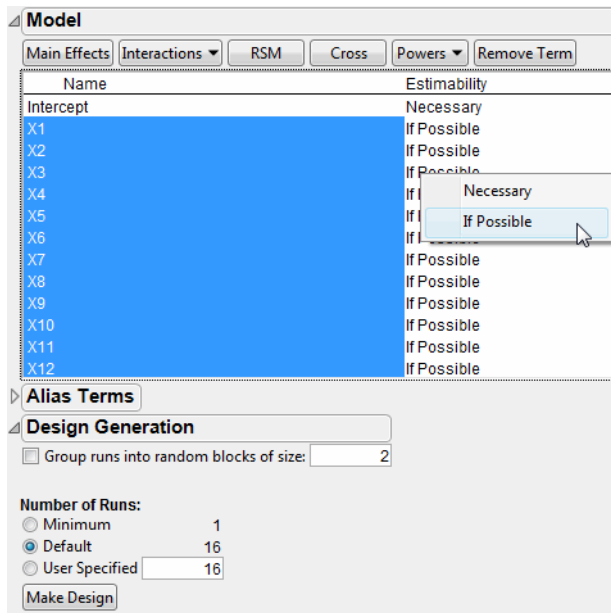
Example: Twelve Factors in Eight Runs

As an example, consider a supersaturated design with twelve factors. Using model terms designated **If Possible** provides the software machinery for creating a supersaturated design.

In the last example, two-factor interaction terms were specified **If Possible**. In a supersaturated design, all terms—including main effects—are **If Possible**. Note in Figure 2.8, the only primary term is the intercept.

To see an example of a supersaturated design with twelve factors in eight runs:

1. Select **DOE > Custom Design**.
2. Add 12 continuous factors and click **Continue**.
3. Highlight all terms except the Intercept and click the current estimability (**Necessary**) to reveal the menu. Change **Necessary** to **If Possible**, as shown in Figure 2.8.

Figure 2.8 Changing the Estimability

- The desired number of runs is eight so type 8 in the User Specified edit box in the **Number of Runs** section.
- Click the red triangle on the Custom Design title bar and select **Simulate Responses**.
- Click **Make Design**, then click **Make Table**. A window named Simulate Responses and a design table appear, similar to the one in Figure 2.9. The Y column values are controlled by the coefficients of the model in the Simulate Responses window. The values in your table may be different from those shown below.

Figure 2.9 Simulated Responses and Design Table

Custom Design

Design

Criteria

Effects

Y

	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	Y
1	1	1	-1	-1	-1	1	-1	-1	1	-1	1	1	52.21
2	1	-1	1	1	-1	-1	1	1	-1	-1	1	1	46.69
3	-1	-1	-1	1	1	-1	1	-1	1	-1	-1	1	27.91
4	1	1	-1	1	1	1	1	-1	-1	1	1	-1	21.28
5	1	1	1	-1	1	-1	-1	1	-1	1	-1	1	5.95
6	-1	1	-1	-1	-1	-1	1	1	1	1	1	-1	49.47
7	-1	1	1	1	1	1	-1	1	1	-1	1	-1	36.74
8	-1	-1	1	-1	-1	1	-1	-1	-1	1	-1	-1	22.96

Effects

Intercept 31

X1 3

X2 -1

X3 3

X4 -3

X5 -6

X6 -8

X7 0

X8 -6

X9 6

X10 -8

X11 9

X12 -7

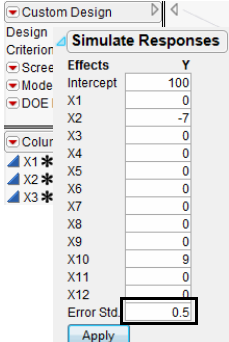
Error Std. 1

Apply

- Change the default settings of the coefficients in the Simulate Responses dialog to match those in Figure 2.10 and click **Apply**. The numbers in the Y column change. Because you have set X2 and X10 as active factors in the simulation, the analysis should be able to identify the same two factors.

Note that random noise is added to the Y column formula, so the numbers you see might not necessarily match those in the figure. The values in your table may be different from those shown below.

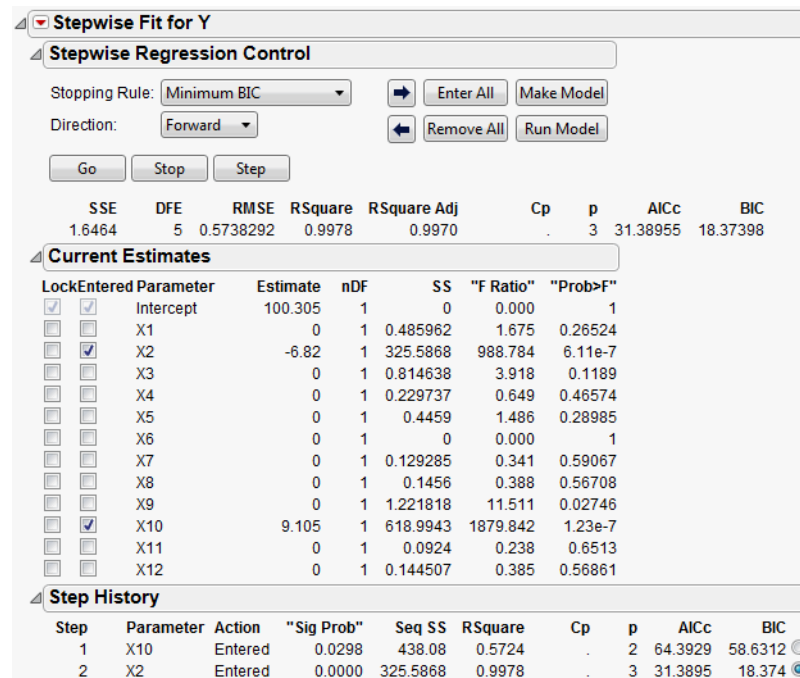
Figure 2.10 Give Values to Two Main Effects and Specify the Standard Error as 0.5



	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	Y
1	1	1	-1	-1	-1	1	-1	-1	1	-1	1	1	84.74
2	1	-1	1	1	-1	-1	1	1	-1	-1	1	1	97.64
3	-1	-1	-1	1	1	-1	1	-1	1	-1	-1	1	98.29
4	1	1	-1	1	1	1	1	-1	-1	1	1	-1	102.37
5	1	1	1	-1	1	-1	-1	1	-1	1	-1	1	101.88
6	-1	1	-1	-1	-1	-1	1	1	1	1	1	-1	103.41
7	-1	1	1	1	1	1	-1	1	1	-1	1	-1	84.13
8	-1	-1	1	-1	-1	1	-1	-1	-1	1	-1	-1	116.34

To identify active factors using stepwise regression:

- To run the Model script in the design table, click the red triangle beside **Model** and select **Run Script**.
- Change the **Personality** in the Model Specification window from **Standard Least Squares** to **Stepwise**.
- Click **Run** on the Fit Model dialog.
- In the resulting display click the **Step** button two times. JMP enters the factors with the largest effects. From the report that appears, you should identify two active factors, X2 and X10, as shown in Figure 2.11. The step history appears in the bottom part of the report. Because random noise is added, your estimates will be slightly different from those shown below.

Figure 2.11 Stepwise Regression Identifies Active Factors

Note: This example defines two large main effects and sets the rest to zero. In real-world situations, it may be less likely to have such clearly differentiated effects.

Screening Designs with Flexible Block Sizes

When you create a design using the Screening designer (**DOE > Screening**), the available block sizes for the listed designs are a power of two. However, custom designs in JMP can have blocks of any size. The blocking example shown in this section is flexible because it is using three runs per block, instead of a power of two.

After you select **DOE > Custom Design** and enter factors, the blocking factor shows only one level in the Values section of the Factors panel because the sample size is unknown at this point. After you complete the design, JMP shows the appropriate number of blocks, which is calculated as the sample size divided by the number of runs per block.

For example, Figure 2.12 shows that when you enter three continuous factors and one blocking factor with three runs per block, only one block appears in the Factors panel.

Figure 2.12 One Block Appears in the Factors Panel

Factors

Add Factor Remove Add N Factors 1

Name	Role	Changes	Values
X1	Continuous	Easy	-1 1
X2	Continuous	Easy	-1 1
X3	Continuous	Easy	-1 1
X4	Blocking	Easy	1

Specify Factors

Add a factor by clicking the Add Factor button. Double click on a factor name or level to edit it.

Continue

The default sample size of nine requires three blocks. After you click **Continue**, there are three blocks in the Factors panel (Figure 2.13). This is because the default sample size is nine, which requires three blocks with three runs each.

Figure 2.13 Three Blocks in the Factors Panel

Custom Design

Responses

Factors

Add Factor Remove Add N Factors 1

Name	Role	Changes	Values
X1	Continuous	Easy	-1 1
X2	Continuous	Easy	-1 1
X3	Continuous	Easy	-1 1
X4	Blocking	Easy	1 2 3

Define Factor Constraints

Model

Alias Terms

Design Generation

☐ Group runs into random blocks of size: 2

Number of Runs:

☐ Minimum 5

☒ Default 9

☐ User Specified 9

Make Design

If you enter 24 runs in the User Specified box of the **Number of Runs** section, the Factors panel changes and now contains 8 blocks (Figure 2.14).

Figure 2.14 Number of Runs is 24 Gives Eight Blocks

The screenshot shows the 'Custom Design' window with the following settings:

- Responses:** (Empty)
- Factors:**
 - Buttons: Add Factor, Remove, Add N Factors (1)
 - Table:
- Define Factor Constraints:** (Empty)
- Model:** (Empty)
- Alias Terms:** (Empty)
- Design Generation:**
 - ☐ Group runs into random blocks of size: 2
 - Number of Runs:**
 - ☐ Minimum: 5
 - ☐ Default: 9
 - ☒ User Specified: 24
 - Make Design

Name	Role	Changes	Values
X1	Continuous	Easy	-1 1
X2	Continuous	Easy	-1 1
X3	Continuous	Easy	-1 1
X4	Blocking	Easy	1 2 3 4 5 6 7 8

If you add all the two-factor interactions and change the number of runs to 15, three runs per block produces five blocks (as shown in Figure 2.15), so the Factors panel displays five blocks in the Values section.

Figure 2.15 Changing the Runs to 15

Custom Design

▶ **Responses**

▶ **Factors**

Add Factor ▼ Remove Add N Factors

Name	Role	Changes	Values
X1	Continuous	Easy	-1 1
X2	Continuous	Easy	-1 1
X3	Continuous	Easy	-1 1
X4	Blocking	Easy	1 2 3 4 5

▶ **Define Factor Constraints**

▶ **Model**

Main Effects Interactions ▼ RSM Cross Powers ▼ Remove Term

Name	Estimability
Intercept	Necessary
X1	Necessary
X2	Necessary
X3	Necessary
X4	Necessary
X1*X2	Necessary
X1*X3	Necessary
X2*X3	Necessary

▶ **Alias Terms**

▶ **Design Generation**

☐ Group runs into random blocks of size:

Number of Runs:

☐ Minimum 11
☐ Default 15
☒ User Specified

Make Design

Click **Make Design**, then click the disclosure button (▶ on Windows and ▶▼ on the Macintosh) to open the Design Evaluation outline node. Then, click the disclosure button to open the Relative Variance of Coefficients report. Figure 2.16 shows the variance of each coefficient in the model relative to the unknown error variance.

The values in your table may be slightly different from those shown below. Notice that the variance of each coefficient is about one-tenth the error variance and that all the variances are roughly the same size. The error variance is assumed to be 1.

Figure 2.16 Table of Relative Variance of the Model Coefficients

Relative Variance of Coefficients		
Significance Level	0.05	
Signal to Noise Ratio	1	
Effect	Variance	Power
Intercept	0.07	0.802
X1	0.078	0.763
X2	0.106	0.638
X3	0.083	0.736
X4 1	0.082	0.744
X4 2	0.081	0.75
X4 3	0.082	0.744
X4 4	0.106	0.638
X1*X2	0.083	0.736
X1*X3	0.079	0.757
X2*X3	0.078	0.763

The main question here is whether the relative size of the coefficient variance is acceptably small. If not, adding more runs (18 or more) will lower the variance of each coefficient.

For more details, see “[The Relative Variance of Coefficients and Power Table](#),” p. 75.

Note to DOE experts: There are four rows associated with X4 (the block factor). That is because X4 has 5 blocks and, therefore, 4 degrees of freedom. Each degree of freedom is associated with one unknown coefficient in the model.

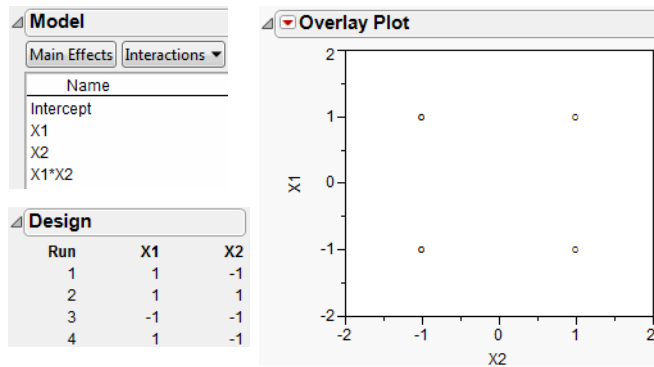
Checking for Curvature Using One Extra Run

In screening designs, experimenters often add center points and other check points to a design to help determine whether the assumed model is adequate. Although this is good practice, it is also *ad hoc*. The custom designer provides a way to improve on this *ad hoc* practice while supplying a theoretical foundation and an easy-to-use interface for choosing a design robust to the modeling assumptions.

The purpose of check points in a design is to provide a detection mechanism for higher-order effects that are contained in the assumed model. These higher-order terms are called *potential terms*. (Let q denote the potential terms, designated **If Possible** in JMP.) The assumed model consists of the *primary terms*. (Let p denote the primary terms designated **Necessary** in JMP.)

To take advantage of the benefits of the approach using **If Possible** model terms, the sample size should be larger than the number of **Necessary** (primary) terms but smaller than the sum of the **Necessary** and **If Possible** (potential) terms. That is, $p < n < p+q$. The formal name of the approach using **If Possible** model terms is *Bayesian D-Optimal design*. This type of design allows the precise estimation of all of the **Necessary** terms while providing omnibus detectability (and some estimability) for the **If Possible** terms.

For a two-factor design having a model with an intercept, two main effects, and an interaction, there are $p = 4$ primary terms. When you enter this model in the custom designer, the default minimum runs value is a four-run design with the factor settings shown in Figure 2.17.

Figure 2.17 Two Continuous Factors with Interaction

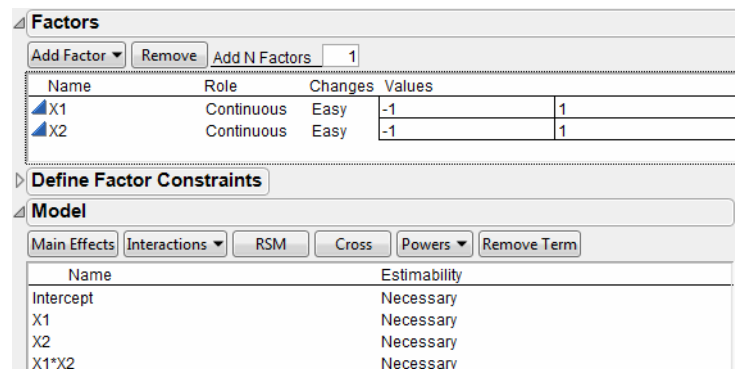
Now suppose you can afford an extra run ($n = 5$). You would like to use this point as a check point for curvature. If you leave the model the same and increase the sample size, the custom designer replicates one of the four vertices. Replicating any run is the optimal choice for improving the estimates of the terms in the model, but it provides no way to check for lack of fit.

Adding the two quadratic terms to the model makes a total of six terms. This is a way to model curvature directly. However, to do this the custom designer requires two additional runs (at a minimum), which exceeds your budget of five runs.

The Bayesian D -Optimal design provides a way to check for curvature while adding only one extra run. To create this design:

1. Select **DOE > Custom Design**.
2. Define two continuous factors (X1 and X2).
3. Click **Continue**.
4. Choose **2nd** from the **Interactions** menu in the Model panel. The results appear as shown in Figure 2.18.

Figure 2.18 Second-Level Interactions



Factors

Add Factor Remove Add N Factors 1

Name	Role	Changes	Values
X1	Continuous	Easy	-1 1
X2	Continuous	Easy	-1 1

Define Factor Constraints

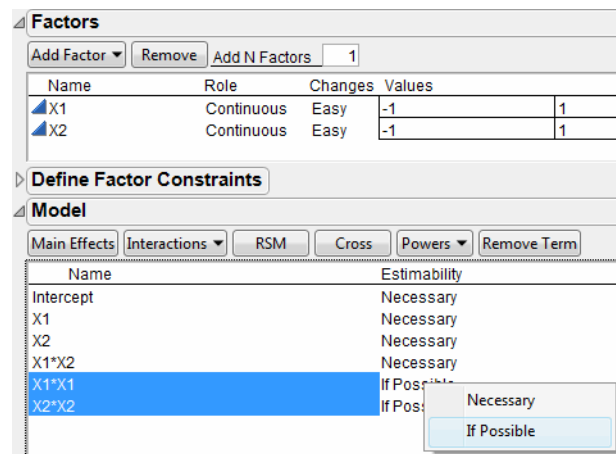
Model

Main Effects Interactions RSM Cross Powers Remove Term

Name	Estimability
Intercept	Necessary
X1	Necessary
X2	Necessary
X1*X2	Necessary

- Choose **2nd** from the **Powers** button in the Model panel. This adds two quadratic terms.
- Select the two quadratic terms ($X1^2$ and $X2^2$) and click the current estimability (**Necessary**) to see the menu and change **Necessary** to **If Possible**, as shown in Figure 2.19.

Figure 2.19 Changing the Estimability



Factors

Add Factor Remove Add N Factors 1

Name	Role	Changes	Values
X1	Continuous	Easy	-1 1
X2	Continuous	Easy	-1 1

Define Factor Constraints

Model

Main Effects Interactions RSM Cross Powers Remove Term

Name	Estimability
Intercept	Necessary
X1	Necessary
X2	Necessary
X1*X2	Necessary
X1^2	If Possible
X2^2	If Possible

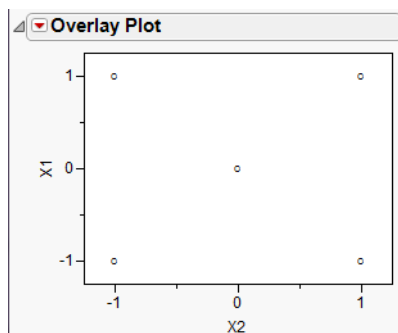
Now, the $p = 4$ primary terms (the intercept, two main effects, and the interaction) are designated as **Necessary** while the $q = 2$ potential terms (the two quadratic terms) are designated as **If Possible**. The desired number of runs, five, is between $p = 4$ and $p + q = 6$.

- Enter 5 into the User Specified edit box in the Number of Runs section of the Design Generation panel.
- Click **Make Design**. The resulting factor settings appear in Figure 2.20. The values in your design may be different from those shown below.

Figure 2.20 Five-Run Bayesian *D*-Optimal Design

Design		
Run	X1	X2
1	-1	1
2	1	-1
3	0	0
4	1	1
5	-1	-1

9. Click **Make Table** to create a JMP data table of the runs.
10. Create the overlay plot in Figure 2.21 with **Graph > Overlay Plot**, and assign X1 as Y and X2 as X. The overlay plot illustrates how the design incorporates the single extra run. In this example the design places the factor settings at the center of the design instead of at one of the corners.

Figure 2.21 Overlay Plot of Five-run Bayesian *D*-Optimal Design

Creating Response Surface Experiments

Response surface experiments traditionally involve a small number (generally 2 to 8) of continuous factors. The *a priori* model for a response surface experiment is usually quadratic.

In contrast to screening experiments, researchers use response surface experiments when they already know which factors are important. The main goal of response surface experiments is to create a predictive model of the relationship between the factors and the response. Using this predictive model allows the experimenter to find better operating settings for the process.

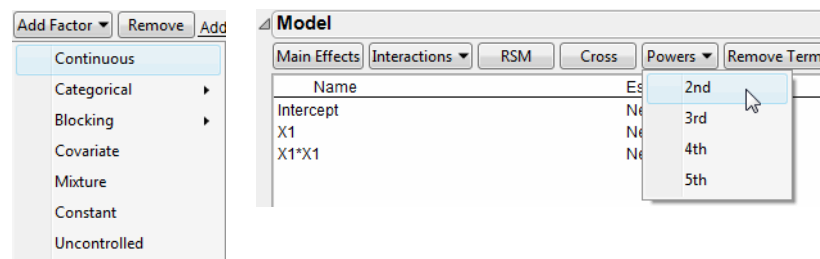
In screening experiments one measure of the quality of the design is the size of the relative variance of the coefficients. In response surface experiments, the prediction variance over the range of the factors is more important than the variance of the coefficients. One way to visualize the prediction variance is JMP's prediction variance profile plot. This plot is a powerful diagnostic tool for evaluating and comparing response surface designs.

Exploring the Prediction Variance Surface

The purpose of the example below is to generate and interpret a simple Prediction Variance Profile Plot. Follow the steps below to create a design for a quadratic model with a single continuous factor.

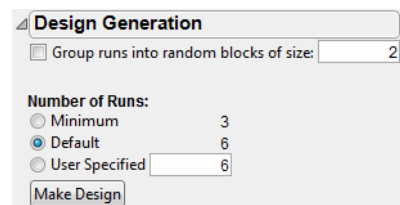
1. Select **DOE > Custom Design**.
2. Add one continuous factor by selecting **Add Factor > Continuous** (Figure 2.22), and click **Continue**.
3. In the Model panel, select **Powers > 2nd** to create a quadratic term (Figure 2.22).

Figure 2.22 Adding a Factor and a Quadratic Term



4. In the Design Generation panel, use the default number of runs (six) and click **Make Design** (Figure 2.23). The number of runs is inversely proportional to the size of variance of the predicted response. As the number of runs increases, the prediction variances decrease.

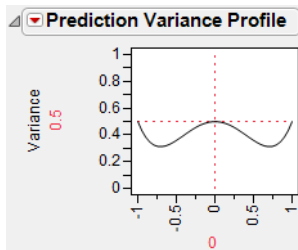
Figure 2.23 Using the Default Number of Runs



5. Click the disclosure button (▶ on Windows and ▶▼ on the Macintosh) to open the **Design Evaluation** outline node, and then the **Prediction Variance Profile**, as shown in Figure 2.24.

For continuous factors, the initial setting is at the mid-range of the factor values. For categorical factors, the initial setting is the first level. If the design model is quadratic, then the prediction variance function is quartic. The y-axis is the relative variance of prediction of the expected value of the response.

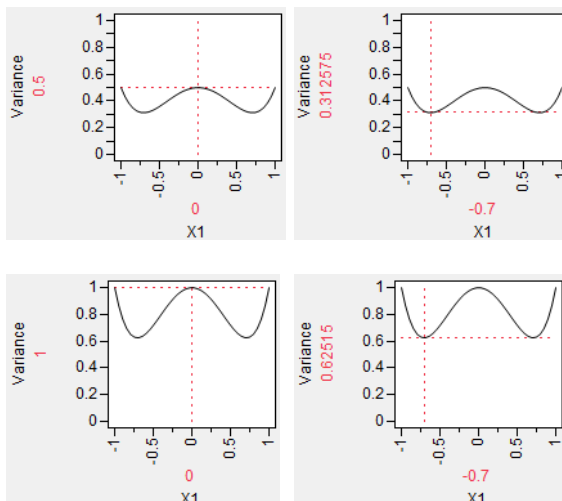
In this design, the three design points are -1 , 0 , and 1 . The prediction variance profile shows that the variance is a maximum at each of these points on the interval -1 to 1 .

Figure 2.24 Prediction Profile for Single Factor Quadratic Model

The prediction variance is relative to the error variance. When the relative prediction variance is one, the absolute variance is equal to the error variance of the regression model. More detail on the Prediction Variance Profiler is in [“Understanding Design Evaluation,” p. 72.](#)

6. To compare profile plots, click the **Back** button and choose **Minimum** in the Design Generation panel, which gives a sample size of three.
7. Click **Make Design** and then open the Prediction Variance Profile again.

Now you see a curve that has the same shape as the previous plot, but the maxima are at one instead of 0.5. Figure 2.25 compares plots for a sample size of six and sample size of three for this quadratic model. You can see the prediction variance increase as the sample size decreases. Since the prediction variance is inversely proportional to the sample size, doubling the number of runs halves the prediction variance. These profiles show settings for the maximum variance and minimum variance, for sample sizes six (top charts) and sample size three (bottom charts). The axes on the bottom plots are adjusted to match the axes on the top plot.

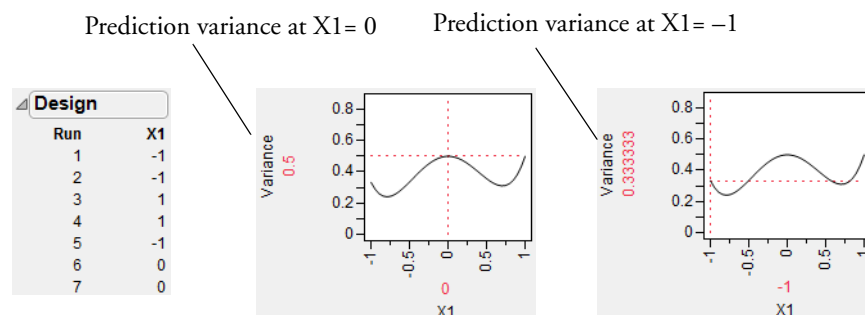
Figure 2.25 Comparison of Prediction Variance Profiles

Tip: Hold the CTRL key and click (hold the COMMAND key and click on Macintosh) on the factor to set a factor level precisely.

8. To create an unbalanced design, click the **Back** button and enter a sample size of 7 in the User Specified text edit box in the Design Generation panel, then click **Make Design**. The results are shown in Figure 2.26.

You can see that the variance of prediction at -1 is lower than the other sample points (its value is 0.33 instead of 0.5). The symmetry of the plot is related to the balance of the factor settings. When the design is balanced, the plot is symmetric, as shown in Figure 2.25. When the design is unbalanced, the prediction plot might not be symmetric, as shown in Figure 2.26.

Figure 2.26 Sample Size of Seven for the One-Factor Quadratic Model



Introducing *I*-Optimal Designs for Response Surface Modeling

The custom designer generates designs using a mathematical optimality criterion. All the designs in this chapter so far have been *D*-Optimal designs. *D*-Optimal designs are most appropriate for screening experiments because the optimality criterion focuses on precise estimates of the coefficients. If an experimenter has precise estimates of the factor effects, then it is easy to tell which factors' effects are important and which are negligible. However, *D*-Optimal designs are not as appropriate for designing experiments where the primary goal is prediction.

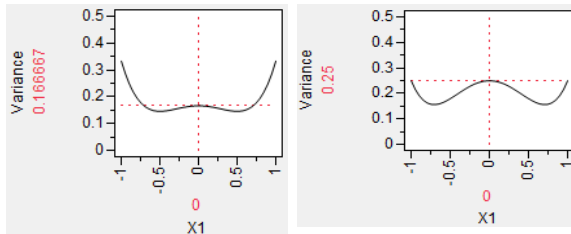
I-Optimal designs minimize the average prediction variance inside the region of the factors. This makes *I*-Optimal designs more appropriate for prediction. As a result *I*-Optimality is the recommended criterion for JMP response surface designs.

An *I*-Optimal design tends to place fewer runs at the extremes of the design space than does a *D*-Optimal design. As an example, consider a one-factor design for a quadratic model using $n = 12$ experimental runs. The *D*-Optimal design for this model puts four runs at each end of the range of interest and four runs in the middle. The *I*-Optimal design puts three runs at each end point and six runs in the middle. In this case, the *D*-Optimal design places two-thirds of its runs at the extremes versus one-half for the *I*-Optimal design.

Figure 2.27 compares prediction variance profiles of the one-factor *I*- and *D*-Optimal designs for a quadratic model with ($n = 12$) runs. The variance function for the *I*-Optimal design is less than the

corresponding function for the *D*-Optimal design in the center of the design space; the converse is true at the edges.

Figure 2.27 Prediction Variance Profiles for 12-Run *I*-Optimal (left) and *D*-Optimal (right) Designs



At the center of the design space, the average variance (relative to the error variance) for the *I*-Optimal design is 0.1667 compared to the *D*-Optimal design, which is 0.25. This means that confidence intervals for prediction will be nearly 10% shorter on average for the *I*-Optimal design.

To compare the two design criteria, create a one-factor design with a quadratic model that uses the *I*-Optimality criterion, and another one that uses *D*-Optimality:

1. Select **DOE > Custom Design**.
2. Add one continuous factor: X1.
3. Click **Continue**.
4. Click the **RSM** button in the Model panel to make the design *I*-Optimal.
5. Change the number of runs to 12.
6. Click **Make Design**.
7. Click the disclosure button (▶ on Windows and ▶▼ on the Macintosh) to open the **Design Evaluation** outline node.
8. Click the disclosure button (▶ on Windows and ▶▼ on the Macintosh) to open the **Prediction Variance Profile**. (The Prediction Variance Profile is shown on the left in Figure 2.27.)
9. Repeat the same steps to create a *D*-Optimal design, but select **Optimality Criterion > Make D-Optimal Design** from the red triangle menu on the custom design title bar. The results in the Prediction Variance Profile should look the same as those on the right in Figure 2.27.

A Three-Factor Response Surface Design

In higher dimensions, the *I*-Optimal design continues to place more emphasis on the center of the region of the factors. The *D*-Optimal and *I*-Optimal designs for fitting a full quadratic model in three factors using 16 runs are shown in Figure 2.28.

To compare the two design criteria, create a three-factor design that uses the *I*-Optimality criterion, and another one that uses *D*-Optimality:

1. Select **DOE > Custom Design**.
2. Add three continuous factors: X1, X2, and X3.
3. Click **Continue**.
4. Click the **RSM** button in the Model panel to add interaction and quadratic terms to the model and to change the default optimality criterion to *I*-Optimal.
5. Use the default of 16 runs.
6. Click **Make Design**.

The design is shown in the Design panel (the left in Figure 2.28).

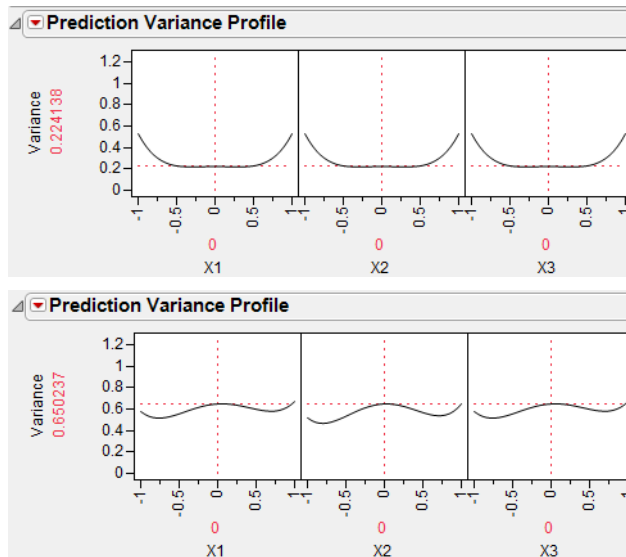
7. If you want to create a *D*-Optimal design for comparison, repeat the same steps but select **Optimality Criterion > Make D-Optimal Design** from the red triangle menu on the custom design title bar. The design should look similar to those on the right in Figure 2.28. The values in your design may be different from those shown below.

Figure 2.28 16-run *I*-Optimal and *D*-Optimal designs for RSM Model

I-Optimal Design				D-Optimal Design			
Run	X1	X2	X3	Run	X1	X2	X3
1	1	0	0	1	1	-1	-1
2	-1	1	1	2	-1	1	1
3	1	-1	1	3	1	-1	1
4	1	1	-1	4	1	1	-1
5	0	0	-1	5	1	-1	0
6	0	0	0	6	-1	1	-1
7	-1	1	-1	7	1	1	1
8	0	0	1	8	0	-1	-1
9	0	1	0	9	-1	0	0
10	-1	0	0	10	0	1	0
11	1	-1	-1	11	0	0	-1
12	-1	-1	1	12	1	0	1
13	0	0	0	13	-1	-1	-1
14	1	1	1	14	-1	-1	1
15	0	-1	0	15	0	-1	1
16	-1	-1	-1	16	-1	-1	0

Profile plots of the variance function are displayed in Figure 2.29. These plots show slices of the variance function as a function of each factor, with all other factors fixed at zero. The *I*-Optimal design has the lowest prediction variance at the center. Note that there are two center points in this design.

The *D*-Optimal design has no center points and its prediction variance at the center of the factor space is almost three times the variance of the *I*-Optimal design. The variance at the vertices of the *D*-Optimal design is not shown. However, note that the *D*-Optimal design predicts better than the *I*-Optimal design near the vertices.

Figure 2.29 Variance Profile Plots for 16 run *I*-Optimal and *D*-Optimal RSM Designs

I-Optimal RSM Design
with 16 runs

D-Optimal RSM Design
with 16 runs

Response Surface with a Blocking Factor

It is not unusual for a process to depend on both qualitative and quantitative factors. For example, in the chemical industry, the yield of a process might depend not only on the quantitative factors temperature and pressure, but also on such qualitative factors as the batch of raw material and the type of reactor. Likewise, an antibiotic might be given orally or by injection, a qualitative factor with two levels. The composition and dosage of the antibiotic could be quantitative factors (Atkinson and Donev, 1992).

The response surface designer (described in “[Response Surface Designs](#),” p. 127) only deals with quantitative factors. You could use the response surface designer to produce a Response Surface Model (RSM) design with a qualitative factor by replicating the design over each level of the factor. But, this is unnecessarily time-consuming and expensive. Using custom designer is simpler and more cost-effective because fewer runs are required. The following steps show how to accommodate a blocking factor in a response surface design using the custom designer:

1. First, define two continuous factors (X1 and X2).
2. Now, click **Add Factor** and select **Blocking > 4 runs per block** to create a blocking factor(X3). The blocking factor appears with one level, as shown in Figure 2.30, but the number of levels adjusts later to accommodate the number of runs specified for the design.

Figure 2.30 Add Two Continuous Factors and a Blocking Factor

Name	Role	Changes	Values
X1	Continuous	Easy	-1 1
X2	Continuous	Easy	-1 1
X3	Blocking	Easy	1 1

- Click **Continue**, and then click **RSM** in the Model panel to add the quadratic terms to the model (Figure 2.31). This automatically changes the recommended optimality criterion from *D*-Optimal to *I*-Optimal. Note that when you click RSM, a message reminds you that nominal factors (such as the blocking factor) cannot have quadratic effects.

Figure 2.31 Add Response Surface Terms

Name	Estimability
Intercept	Necessary
X1	Necessary
X2	Necessary
X3	Necessary
X1*X1	Necessary
X1*X2	Necessary
X2*X2	Necessary

- Enter 12 in the User Specified text edit box in the Design Generation panel, and note that the Factors panel now shows the Blocking factor, X3, with three levels (Figure 2.32). Twelve runs defines three blocks with four runs per block.

Figure 2.32 Blocking Factor Now Shows Three Levels

Factors

Add Factor Remove Add N Factors 1

Name	Role	Changes	Values
X1	Continuous	Easy	-1 1
X2	Continuous	Easy	-1 1
X3	Blocking	Easy	1 2 3

Define Factor Constraints

Model

Main Effects Interactions RSM Cross Powers Remove Term

Name	Estimability
Intercept	Necessary
X1	Necessary
X2	Necessary
X3	Necessary
X1*X1	Necessary
X1*X2	Necessary
X2*X2	Necessary

- Click **Make Design**.
- In the Output Options, select **Sort Right to Left** from the **Run Order** list.
- Click **Make Table** to see an *I*-Optimal table similar to the one on the left in Figure 2.33.

Figure 2.33 compares the results of a 12-run *I*-Optimal design and a 12-run *D*-Optimal Design.

To see the *D*-Optimal design:

- Click the **Back** button.
- Click the red triangle icon on the Custom Design title bar and select **Optimality Criterion > Make D-Optimal Design**.
- Click **Make Design**, then click **Make Table**.

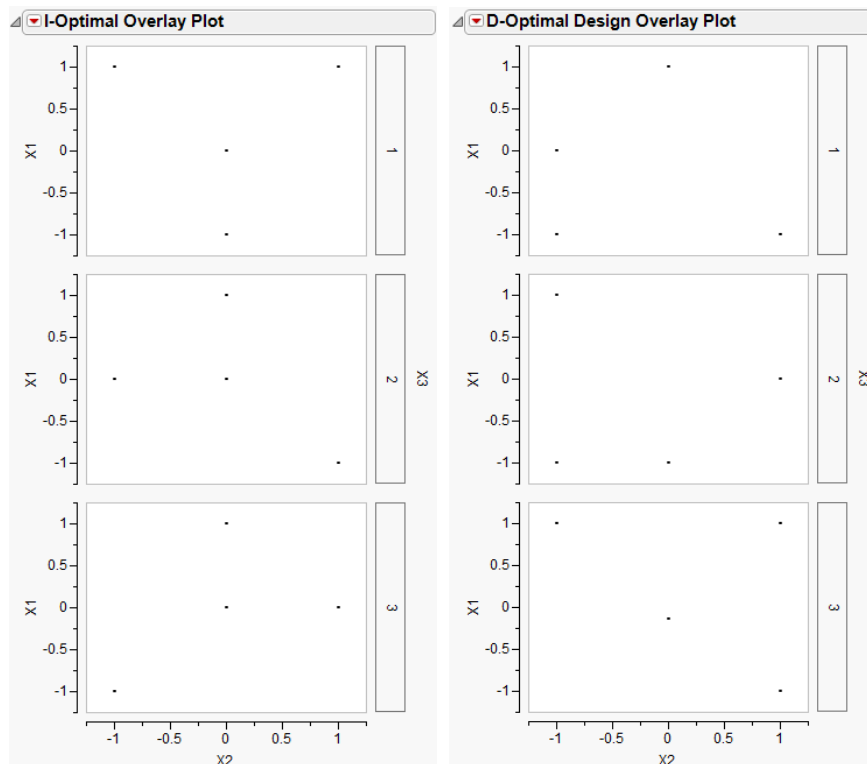
Figure 2.33 JMP Design Tables for 12-Run *I*-Optimal and *D*-Optimal Designs

I-Optimal Design				D-Optimal Design			
Run	X1	X2	X3	Run	X1	X2	X3
1	0	0	2	1	-0.14	0	3
2	1	0	3	2	1	-1	3
3	0	0	3	3	-1	0	2
4	-1	-1	3	4	-1	1	3
5	1	-1	1	5	-1	1	1
6	-1	0	1	6	-1	-1	1
7	0	1	1	7	1	-1	2
8	1	1	2	8	0	-1	1
9	-1	1	2	9	1	0	1
10	0	1	3	10	1	1	3
11	0	0	1	11	-1	-1	2
12	0	-1	2	12	0	1	2

Figure 2.34 gives a graphical view of the designs generated by this example. These plots were generated for the runs in each JMP table by choosing **Graph > Overlay Plot** from the main menu and using the blocking factor (X3) as the **Grouping** variable.

Note that there is a center point in each block of the *I*-Optimal design. The *D*-Optimal design has only one center point. The values in your graph may be different from those shown in Figure 2.34.

Figure 2.34 Plots of *I*-Optimal (left) and *D*-Optimal (right) Design Points by Block.



Either of the designs in Figure 2.34 supports fitting the specified model. The *D*-Optimal design does a slightly better job of estimating the model coefficients. The diagnostics (Figure 2.35) for the designs show beneath the design tables. In this example, the *D*-efficiency of the *I*-Optimal design is about 51%, and is 55% for the *D*-Optimal design.

The *I*-Optimal design is preferable for predicting the response inside the design region. Using the formulas given in “[Technical Discussion](#),” p. 60, you can compute the relative average variance for these designs. The average variance (relative to the error variance) for the *I*-Optimal design is 0.5 compared to 0.59 for the *D*-Optimal design (See Figure 2.35). This means confidence intervals for prediction will be almost 20% longer on average for *D*-Optimal designs.

Figure 2.35 Design Diagnostics for I-Optimal and D-Optimal Designs

Design Diagnostics		Design Diagnostics	
I Optimal Design		D Optimal Design	
D Efficiency	50.8965	D Efficiency	54.98822
G Efficiency	76.54825	G Efficiency	78.81296
A Efficiency	37.56129	A Efficiency	32.5789
Average Variance of Prediction	0.498888	Average Variance of Prediction	0.59188
Design Creation Time (seconds)	0.35	Design Creation Time (seconds)	0.4

Creating Mixture Experiments

If you have factors that are ingredients in a mixture, you can use either the custom designer or the specialized mixture designer. However, the mixture designer is limited because it requires all factors to be mixture components and you might want to vary the process settings along with the percentages of the mixture ingredients. The optimal formulation could change depending on the operating environment. The custom designer can handle mixture ingredients and process variables in the same study. You are not forced to modify your problem to conform to the restrictions of a special-purpose design approach.

Mixtures Having Nonmixture Factors

The following example from Atkinson and Donev (1992) shows how to create designs for experiments with mixtures where one or more factors are not ingredients in the mixture. In this example:

- The response is the electromagnetic damping of an acrylonitrile powder.
- The three mixture ingredients are copper sulphate, sodium thiosulphate, and glyoxal.
- The nonmixture environmental factor of interest is the wavelength of light.

Though **wavelength** is a continuous variable, the researchers were only interested in predictions at three discrete wavelengths. As a result, they treated it as a categorical factor with three levels. To create this custom design:

1. Select **DOE > Custom Design**.
2. Create **Damping** as the response. The authors do not mention how much damping is desirable, so right-click the goal and create **Damping's** response goal to be **None**.
3. In the Factors panel, add the three mixture ingredients and the categorical factor, **Wavelength**. The mixture ingredients have range constraints that arise from the mechanism of the chemical reaction. Rather than entering them by hand, load them from the Sample Data folder that was installed with JMP: click the red triangle icon on the Custom Design title bar and select **Load Factors**. Open **Donev Mixture Factors.jmp**, from the Design Experiment sample data folder. The custom design panels should now look like those shown in Figure 2.36.

Figure 2.36 Mixture Experiment Response Panel and Factors Panel

Custom Design

Responses

Add Response ▼ Remove Number of Responses...

Response Name	Goal	Lower Limit	Upper Limit	Importance
Damping	None	NA	NA	NA

Factors

Add Factor ▼ Remove Add N Factors 1

Name	Role	Changes	Values
▲ CuSO4	Mixture	Easy	0.2 0.8
▲ Na2S2O3	Mixture	Easy	0.2 0.8
▲ Glyoxal	Mixture	Easy	0 0.6
▼ Wavelength	Categorical	Easy	L1 L2 L3

The model, shown in Figure 2.37 is a response surface model in the mixture ingredients along with the additive effect of the wavelength. To create this model:

1. Click **Interactions**, and choose **2nd**. A warning dialog appears telling you that JMP removes the main effect terms for non-mixture factors that interact with all the mixture factors. Click **OK**.
2. In the Design Generation panel, type 18 in the **User Specified** text edit box (Figure 2.37), which results in six runs each for the three levels of the wavelength factor.

Figure 2.37 Mixture Experiment Design Generation Panel

Custom Design

▶ **Responses**

▲ **Factors**

Add Factor ▼ Remove Add N Factors

Name	Role	Changes	Values
▲ CuSO4	Mixture	Easy	0.2 0.8
▲ Na2S2O3	Mixture	Easy	0.2 0.8
▲ Glyoxal	Mixture	Easy	0 0.6
▼ Wavelength	Categorical	Easy	L1 L2 L3

▶ **Define Factor Constraints**

▲ **Model**

Main Effects Interactions ▼ Cross Powers ▼ Scheffe Cubic Remove Term

Name	Estimability
CuSO4	Necessary
Na2S2O3	Necessary
Glyoxal	Necessary
CuSO4*Na2S2O3	Necessary
CuSO4*Glyoxal	Necessary
CuSO4*Wavelength	Necessary
Na2S2O3*Glyoxal	Necessary
Na2S2O3*Wavelength	Necessary
Glyoxal*Wavelength	Necessary

▶ **Alias Terms**

▲ **Design Generation**

☐ Group runs into random blocks of size:

Number of Runs:

☐ Minimum 12
☐ Default 24
☒ User Specified

Make Design

3. Click **Make Design**, and then click **Make Table**.

The resulting data table is shown in Figure 2.38. The values in your table may be different from those shown below.

Figure 2.38 Mixture Experiment Design Table

Custom Design						
Design	Custom Design					
Criterion	D Optimal					
Model						
DOE Dialog						
Columns (5/0)						
CuSO4 *						
Na2S2O3 *						
Glyoxal *						
Wavelength *						
Damping *						
Rows						
All rows	18					
Selected	0					
Excluded	0					
Hidden	0					
Labelled	0					

	CuSO4	Na2S2O3	Glyoxal	Wavelength	Damping
1	0.2	0.8	0	L1	•
2	0.5	0.2	0.3	L1	•
3	0.2	0.8	0	L2	•
4	0.5	0.5	0	L1	•
5	0.8	0.2	0	L1	•
6	0.5	0.2	0.3	L2	•
7	0.2	0.5	0.3	L1	•
8	0.2	0.5	0.3	L2	•
9	0.5	0.5	0	L2	•
10	0.2	0.2	0.6	L2	•
11	0.2	0.2	0.6	L1	•
12	0.5	0.2	0.3	L3	•
13	0.5	0.5	0	L3	•
14	0.8	0.2	0	L3	•
15	0.8	0.2	0	L2	•
16	0.2	0.5	0.3	L3	•
17	0.2	0.8	0	L3	•
18	0.2	0.2	0.6	L3	•

Atkinson and Donev also discuss the design where the number of runs is limited to 10. In that case, it is not possible to run a complete mixture response surface design for every wavelength.

To view this:

1. Click the **Back** button.
2. Remove all the effects by highlighting them and clicking **Remove Term**.
3. Add the main effects by clicking the **Main Effects** button.
4. In the Design Generation panel, change the number of runs to **10** (Figure 2.39) and click **Make Design**. The Design table to the right in Figure 2.39 shows the factor settings for 10 runs.

Figure 2.39 Ten-Run Mixture Response Surface Design

Design Generation

☐ Group runs into random blocks of size: 2

Number of Runs:

☐ Minimum 5

☐ Default 6

☒ User Specified 10

Make Design

Design

Run	CuSO4	Na2S2O3	Glyoxal	Wavelength	Damping
1	0.2	0.2	0.6	L2	
2	0.2	0.8	0	L2	
3	0.2	0.2	0.6	L3	
4	0.2	0.8	0	L1	
5	0.8	0.2	0	L3	
6	0.8	0.2	0	L1	
7	0.2	0.8	0	L3	
8	0.8	0.2	0	L1	
9	0.8	0.2	0	L2	
10	0.2	0.2	0.6	L1	

Note that there are necessarily unequal numbers of runs for each wavelength. Because of this lack of balance it is a good idea to look at the prediction variance plot (top plot in Figure 2.40).

5. Open the **Design Evaluation** outline node, then open the **Prediction Variance Profile**.

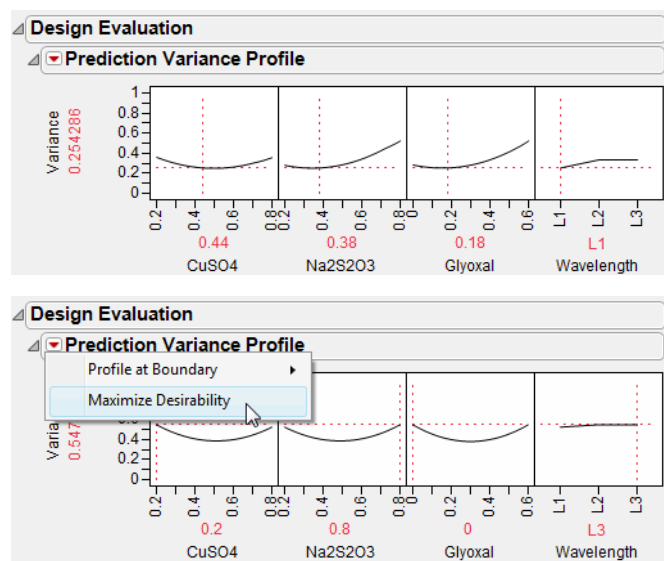
The prediction variance is almost constant across the three wavelengths, which is a good indication that the lack of balance is not a problem.

The values of the first three ingredients sum to one because they are mixture ingredients. If you vary one of the values, the others adjust to keep the sum constant.

6. Select **Maximize Desirability** from red triangle menu on the Prediction Variance Profile title bar, as shown in the bottom profiler in Figure 2.40.

The most desirable wavelength is L3, with the CuSO₄ percentage decreasing from about 0.4 to 0.2, Glyoxal percentage is zero, and Na₂S₂O₃ is 0.8, which maintains the mixture.

Figure 2.40 Prediction Variance Plots for Ten-Run Design



Experiments that are Mixtures of Mixtures

As a way to illustrate the idea of a 'mixture of mixtures' situation, imagine the ingredients that go into baking a cake and assume the following:

- dry ingredients composed of flour, sugar, and cocoa
- wet (or non-dry) ingredients consisting of milk, melted butter, and eggs.

These two components (wet and dry) of the cake are two mixtures that are first mixed separately and then blended together.

The dessert chef knows that the dry component (the mixture of flour, sugar, and cocoa) contributes 45% of the combined mixture and the wet component (butter, milk, and eggs) contributes 55%.

The objective of such an experiment might be to identify proportions within the two components that maximize some measure of taste or consistency.

This is a main effects model except that you must leave out one of the factors in order to avoid singularity. The choice of which factor to leave out of the model is arbitrary.

For now, consider these upper and lower levels of the various factors:

Within the dry mixture:

- cocoa must be greater than 10% but less than 20%
- sugar must be greater than 0% but less than 15%
- flour must be greater than 20% but less than 30%

Within the wet mixture:

- melted butter must be greater than 10% but less than 20%
- milk must be greater than 25% and less than 35%
- eggs constitute more than 5% but less than 20%

You want to bake cakes and measure taste on a scale from 1 to 10

Use the Custom Designer to set up this example, as follows:

1. In the Response Panel, enter one response and call it **Taste**.
2. Give Taste a **Lower Limit** of 1 and an **Upper Limit** of 10. (You are assuming a taste test where the respondents reply on a scale of 1 to 10.)
3. In the Factors Panel, enter the six cake factors described above.
4. Enter the given percentage values of the factors as proportions in the **Values** section of the Factors panel.

The completed Response and Factors panels should look like those shown in Figure 2.41.

Figure 2.41 Completed Responses and Factors Panel for the Cake Example

The screenshot shows the 'Custom Design' window with two main panels: 'Responses' and 'Factors'.

Responses Panel:

- Buttons: Add Response, Remove, Number of Responses...
- Table:

Response Name	Goal	Lower Limit	Upper Limit	Importance
Taste	Maximize	1	10	.

optional item

Factors Panel:

- Buttons: Add Factor, Remove, Add N Factors, 1
- Table:

Name	Role	Changes	Values
▲ Cocoa	Mixture	Easy	0.1 0.2
▲ Sugar	Mixture	Easy	0 0.15
▲ Flour	Mixture	Easy	0.2 0.3
▲ Butter	Mixture	Easy	0.1 0.2
▲ Milk	Mixture	Easy	0.25 0.35
▲ Eggs	Mixture	Easy	0.05 0.2

5. Next, click **Continue**.
6. Open the Define Factor Constraints pane and click **Add Constraint** twice.
7. Enter the constraints as shown in Figure 2.42. For the second constraint setting, click on the less than or equal to button and select the greater than or equal to direction.

By confining the dry factors to exactly 45% in this way, the mixture role of all the factors ensures that the wet factors constitute the remaining 55%.

8. Open the Model dialog and note that it lists all 6 effects. Because these are mixture factors, including all effects would render the model singular. Highlight any one of the terms in the model and click **Remove Term**, as shown.

Figure 2.42 Constraints to Define the Double Mixture Experiment

Define Factor Constraints

Add Constraint

1	Cocoa +	1	Sugar +	1	Flour +	0	Butter +	0	Milk +	0	Eggs	≤	0.45
1	Cocoa +	1	Sugar +	1	Flour +	0	Butter +	0	Milk +	0	Eggs	≥	0.45

Model

Main Effects Interactions Cross Powers Scheffe Cubic Remove Term

Name	Estimability
Cocoa	Necessary
Sugar	Necessary
Flour	Necessary
Butter	Necessary
Milk	Necessary
Eggs	Necessary

9. To see a completed example, choose **Simulate Responses** from the menu on the Custom Design title bar.
10. In the Design Generation panel, enter 10 as the number of runs for the example. That is, you would bake cakes with 10 different sets of ingredient proportions.
11. Click **Make Design** in the Design Generation panel, and then click **Make Table**.

The table in Figure 2.43 shows that the two sets of cake ingredients (dry and wet) adhere to the proportions 45% and 55% as defined by the entered constraints. In addition, the amount of each ingredient in each cake recipe (run) conforms to the upper and lower limits given in the factors dialog.

Figure 2.43 Cake Experiment Conforming to a Mixture of Mixture Design

	Cocoa	Sugar	Flour	Butter	Milk	Eggs	Taste
1	0.15	0	0.3	0.1	0.35	0.1	5.13
2	0.2	0.05	0.2	0.2	0.25	0.1	5.7
3	0.2	0	0.25	0.2	0.25	0.1	5.95
4	0.1	0.05	0.3	0.1	0.25	0.2	1.1
5	0.2	0.05	0.2	0.1	0.35	0.1	5.64
6	0.1	0.15	0.2	0.15	0.35	0.05	2.88
7	0.1	0.05	0.3	0.2	0.3	0.05	3.49
8	0.1	0.15	0.2	0.2	0.25	0.1	3.62
9	0.1	0.15	0.2	0.1	0.25	0.2	5.07
10	0.2	0.05	0.2	0.1	0.25	0.2	7.18

Each run sums to 0.45 (45%) Each run sums to 0.55 (55%)

Note: As a word of caution, keep in mind that it is easy to define constraints in such a way that it is impossible to construct a design that fits the model. In such a case, you will get a message saying “Could not find a valid starting design. Please check your constraints for consistency.”

Special-Purpose Uses of the Custom Designer

While some of the designs discussed in previous sections can be created using other designers in JMP or by looking them up in a textbook containing tables of designs, the designs presented in this section cannot be created without using the custom designer.

Designing Experiments with Fixed Covariate Factors

Pre-tabulated designs rely on the assumption that the experimenter controls all the factors. Sometimes you have quantitative measurements (a covariate) on the experimental units before the experiment begins. If this variable affects the experimental response, the covariate should be a design factor. The pre-defined design that allows only a few discrete values is too restrictive. The custom designer supplies a reasonable design option.

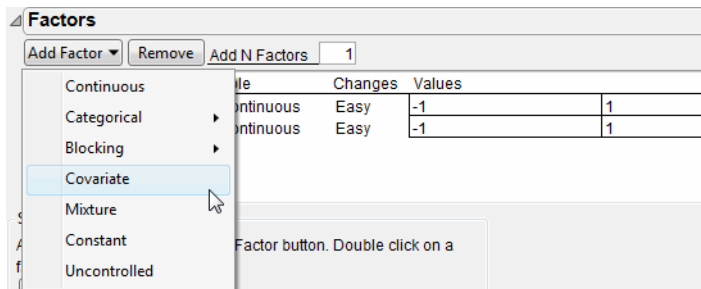
For this example, suppose there are a group of students participating in a study. A physical education researcher has proposed an experiment where you vary the number of hours of sleep and the calories for breakfast and ask each student to run 1/4 mile. The weight of the student is known and it seems important to include this information in the experimental design.

To follow along with this example that shows column properties, open **Big Class.jmp** from the Sample Data folder that was installed when you installed JMP.

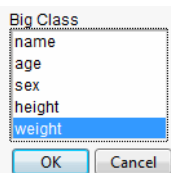
Build the custom design as follows:

1. Select **DOE > Custom Design**.

2. Add two continuous variables to the models by entering 2 beside Add N Factors, clicking **Add Factor** and selecting **Continuous**, naming them **calories** and **sleep**.
3. Click **Add Factor** and select **Covariate**, as shown in Figure 2.44. The **Covariate** selection displays a list of the variables in the current data table.

Figure 2.44 Add a Covariate Factor


4. Select **weight** from the variable list (Figure 2.45) and click **OK**.

Figure 2.45 Design with Fixed Covariate


5. Click **Continue**.
6. Add the interaction to the model by selecting **calories** in the Factors panel, selecting **sleep** in the Model panel, and then clicking the **Cross** button (Figure 2.46).

Figure 2.46 Design With Fixed Covariate Factor

Factors

Add Factor Remove Add N Factors 1

Name	Role	Changes	Values
calories	Continuous	Easy	-1 1
sleep	Continuous	Easy	-1 1
weight	Covariate	Easy	64 172

Define Factor Constraints

Model

Main Effects Interactions RSM Cross Powers Remove Term

Name	Estimability
Intercept	Necessary
calories	Necessary
sleep	Necessary
weight	Necessary

7. Click **Make Design**, then click **Make Table**. The data table in Figure 2.47 shows the design table. Your runs might not look the same because the order of the runs has been randomized.

Figure 2.47 Design Table for Covariate Example

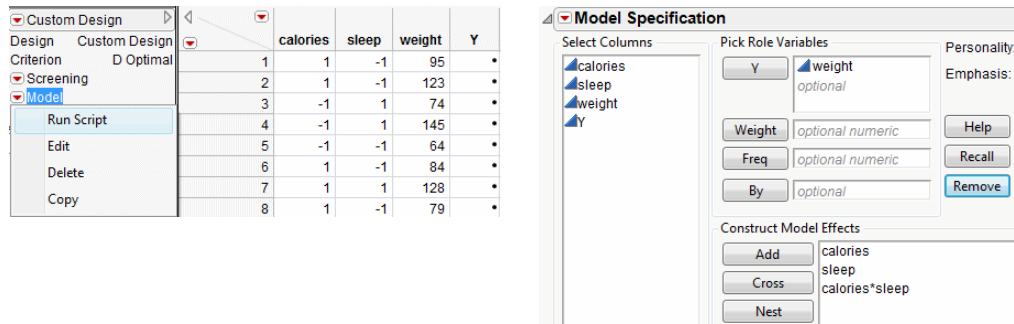
Custom Design						
Design	Custom Design					
Criterion	D Optimal					
Screening						
Model						
DOE Dialog						
Columns (4/0)						
calories *						
sleep *						
weight *						
Y *						
Rows						
All rows	40					
Selected	0					
Excluded	0					
Hidden	0					
Labelled	0					

	calories	sleep	weight	Y
1	1	1	-1	95
2	1	1	-1	123
3	-1	1	1	74
4	-1	1	1	145
5	-1	-1	64	•
6	1	-1	84	•
7	1	1	128	•
8	1	-1	79	•
9	1	-1	112	•
10	-1	1	107	•
11	-1	1	67	•
12	1	1	98	•
13	1	1	105	•
14	-1	-1	95	•
15	1	1	79	•

Note: Covariate factors cannot have missing values.

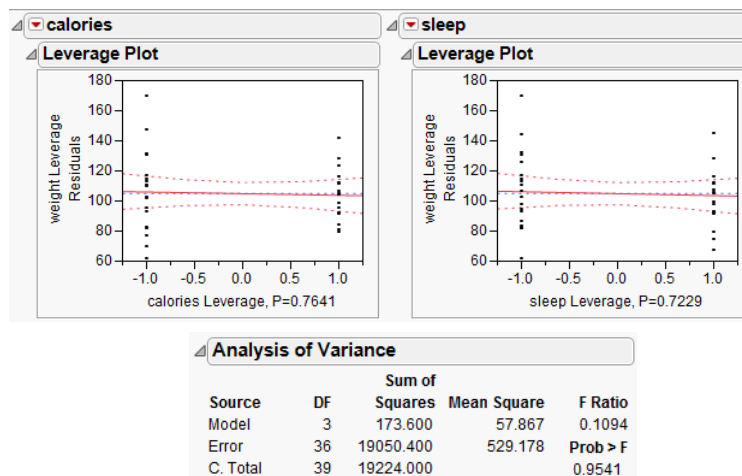
Remember that **weight** is the covariate factor, measured for each student, but it is not controlled. The custom designer has calculated settings for **calories** and **sleep** for each student. It would be desirable if the correlations between **calories**, **sleep** and **weight** were as small as possible. You can see how well the custom designer did by fitting a model of **weight** as a function of **calories** and **sleep**. If that fit has a small model sum of squares, that means the custom designer has successfully separated the effect of **weight** from the effects of **calories** and **sleep**.

8. Click the red triangle icon beside **Model** in the data table and select **Run Script**, as shown on the left in Figure 2.48.

Figure 2.48 Model Script


9. Rearrange the dialog so **weight** is Y and **calories**, **sleep**, and **calories*sleep** are the model effects, as shown to the right in Figure 2.48. Click **Run**.

The leverage plots are nearly horizontal, and the analysis of variance table shows that the model sum of squares is near zero compared to the residuals (Figure 2.49). Therefore, **weight** is independent of **calories** and **sleep**. The values in your analysis may be a little different from those shown below.

Figure 2.49 Analysis to Check That Weight is Independent of Calories and Sleep


Creating a Design with Two Hard-to-Change Factors: Split Plot

While there is substantial research literature covering the analysis of split plot designs, it has only been possible in the last few years to create optimal split plot designs (Goos 2002). The split plot design capability accessible in the JMP custom designer is the first commercially available tool for generating optimal split plot designs.

The split plot design originated in agriculture, but is commonplace in manufacturing and engineering studies. In split plot experiments, hard-to-change factors only change between one whole plot and the next. The whole plot is divided into subplots, and the levels of the easy-to-change factors are randomly assigned to each subplot.

The example in this section is adapted from Kowalski, Cornell, and Vining (2002). The experiment studies the effect of five factors on the thickness of vinyl used to make automobile seat covers. The response and factors in the experiment are described below:

- Three of the factors are ingredients in a mixture. They are plasticizers whose proportions, **m1**, **m2**, and **m3**, sum to one. Additionally, the mixture components are the subplot factors of the experiment.
- Two of the factors are process variables. They are the rate of extrusion (**extrusion rate**) and the temperature (**temperature**) of drying. These process variables are the whole plot factors of the experiment. They are hard to change.
- The response in the experiment is the thickness of the vinyl used for automobile seat covers. The response of interest (**thickness**) depends both on the proportions of the mixtures and on the effects of the process variables.

To create this design in JMP:

1. Select **DOE > Custom Design**.
2. By default, there is one response, **Y**, showing. Double-click the name and change it to **thickness**. Use the default goal, **Maximize** (Figure 2.50).
3. Enter the lower limit of 10.
4. To add three mixture factors, type 3 in the box beside **Add N Factors**, and click **Add Factor > Mixture**.
5. Name the three mixture factors **m1**, **m2**, and **m3**. Use the default levels 0 and 1 for those three factors.
6. Add two continuous factors by typing 2 in the box beside **Add N Factors**, and click **Add Factor > Continuous**. Name these factors **extrusion rate** and **temperature**.
7. Ensure that you are using the default levels, -1 and 1, in the Values area corresponding to these two factors.
8. To make **extrusion rate** a whole plot factor, click **Easy** and select **Hard**.
9. To make **temperature** a whole plot factor, click **Easy** and select **Hard**. Your dialog should look like the one in Figure 2.50.

Figure 2.50 Entering Responses and Factors

The Custom Design dialog box is shown with the 'Responses' and 'Factors' sections expanded. The 'Responses' section has buttons for 'Add Response', 'Remove', and 'Number of Responses...'. Below these is a table with columns: Response Name, Goal, Lower Limit, Upper Limit, and Importance. The 'Factors' section has buttons for 'Add Factor', 'Remove', and 'Add N Factors' (set to 1). Below these is a table with columns: Name, Role, Changes, and Values.

Response Name	Goal	Lower Limit	Upper Limit	Importance
thickness	Maximize	10	.	.
<i>optional item</i>				

Name	Role	Changes	Values
m1	Mixture	Easy	0 1
m2	Mixture	Easy	0 1
m3	Mixture	Easy	0 1
extrusion rate	Continuous	Hard	-1 1
temperature	Continuous	Hard	-1 1

10. Click **Continue**.
11. Next, add interaction terms to the model by selecting **Interactions > 2nd** (Figure 2.51). This causes a warning that JMP removes the main effect terms for non-mixture factors that interact with all the mixture factors. Click **OK**.

Figure 2.51 Adding Interaction Terms

The Model dialog box is shown with the 'Interactions' dropdown menu open, displaying options from 2nd to 5th. The main table lists various terms and their estimability.

Name	Estimability
m1	Necessary
m2	Necessary
m3	Necessary
m1*m2	Necessary
m1*m3	Necessary
m1*extrusion rate	Necessary
m1*temperature	Necessary
m2*m3	Necessary
m2*extrusion rate	Necessary
m2*temperature	Necessary
m3*extrusion rate	Necessary
m3*temperature	Necessary
extrusion rate*temperature	Necessary

12. In the Design Generation panel, type 7 in the Number of Whole Plots text edit box.
13. For **Number of Runs**, type 28 in the User Specified text edit box (Figure 2.52).

Figure 2.52 Assigning the Number of Whole Plots and Number of Runs

Design Generation

Number of Whole Plots

Number of Runs:

☐ Minimum 13

☐ Default 24

☒ User Specified

Make Design

Note: If you enter a missing value in the Number of Whole Plots edit box, then JMP considers many different numbers of whole plots and chooses the number that maximizes the information about the coefficients in the model. It maximizes the determinant of $\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}$ where \mathbf{V}^{-1} is the inverse of the variance matrix of the responses. The matrix, \mathbf{V} , is a function of how many whole plots there are, so changing the number of whole plots changes \mathbf{V} , which can make a difference in the amount of information a design contains.

14. Click **Make Design**. The result is shown in Figure 2.53.

Figure 2.53 Partial Listing of the Final Design Structure

Run	Whole Plots	m1	m2	m3	extrusion rate	temperature
1	1	0	1	0	1	1
2	1	0.5	0	0.5	1	1
3	1	0	0	1	1	1
4	1	1	0	0	1	1
5	2	0	0	1	-1	1
6	2	0	1	0	-1	1
7	2	1	0	0	-1	1
8	2	0.5	0.5	0	-1	1
9	3	1	0	0	-1	-1
10	3	0	0	1	-1	-1
11	3	0	1	0	-1	-1
12	3	0.5	0.5	0	-1	-1
13	4	0	1	0	1	-1

15. Click **Make Table**.

16. From the Sample Data folder that was installed with JMP, open *Vinyl Data.jmp* from the Design Experiment folder, which contains 28 runs as well as response values. The values in the table you generated with the custom designer may be different from those from the Sample Data folder, shown in Figure 2.54.

Figure 2.54 The Vinyl Data Design Table

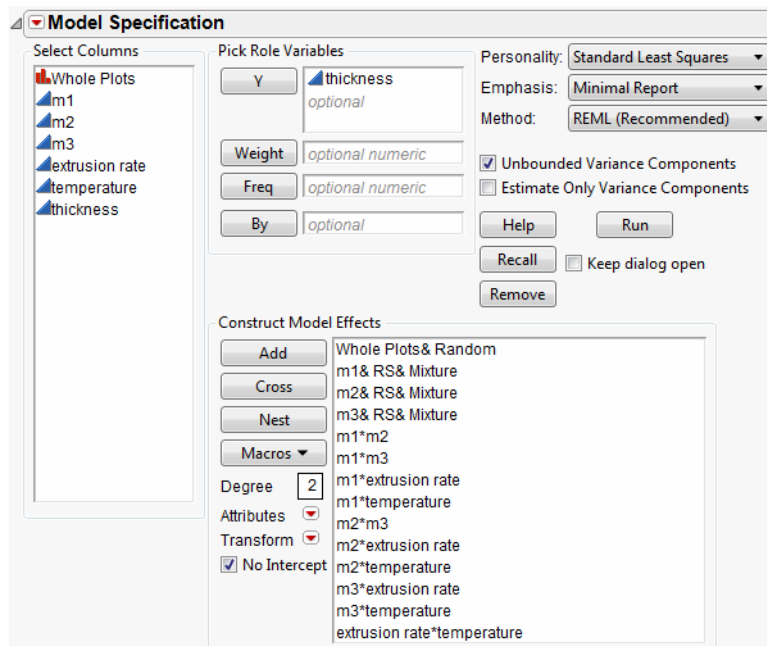
Vinyl Data								
Design	Custom Design							
Criterion	D Optimal							
Model								
Columns (7/0)								
Whole Plots *								
m1 *								
m2 *								
m3 *								
extrusion rate *								
temperature *								
thickness *								
Rows								
All rows	28							
Selected	0							
Excluded	0							
Hidden	0							
Labelled	0							

	Whole Plots	m1	m2	m3	extrusion rate	temperature	thickness
1	1	0	1	0	-1	-1	4
2	1	1	0	0	-1	-1	10
3	1	0	0	1	-1	-1	3
4	1	0.5	0	0.5	-1	-1	9
5	2	0.5	0.5	0	-1	1	8
6	2	0	1	0	-1	1	5
7	2	1	0	0	-1	1	6
8	2	0	0	1	-1	1	2
9	3	0	1	0	1	-1	7
10	3	0	0.5	0.5	1	-1	9
11	3	0	0	1	1	-1	14
12	3	1	0	0	1	-1	12
13	4	0	1	0	1	1	5
14	4	0	0	1	1	1	2
15	4	1	0	0	1	1	6
16	4	0.5	0	0.5	1	1	7

17. Click the red triangle icon next to the Model script and select **Run Script**. The dialog in Figure 2.55 appears.

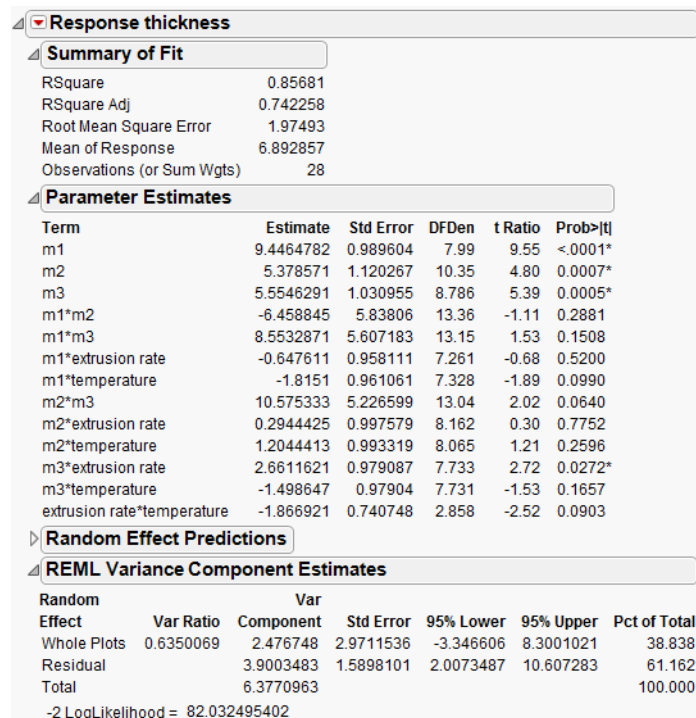
The models for split plots have a random effect associated with the whole plots' effect.

As shown in the dialog in Figure 2.55, JMP designates the error term by appending **&Random** to the name of the effect. REML will be used for the analysis, as indicated in the menu beside **Method** in Figure 2.55. For more information about REML models, see *Modeling and Multivariate Methods*.

Figure 2.55 Define the Model in the Fit Model Dialog

18. Click **Run** to run the analysis. The results are shown in Figure 2.56.

Figure 2.56 Split Plot Analysis Results



Technical Discussion

This section provides information about *I*-, *D*-, Bayesian *I*-, Bayesian *D*-, and Alias-Optimal designs.

D-Optimality:

- is the default design type produced by the custom designer except when the **RSM** button has been clicked to create a full quadratic model.
- minimizes the variance of the model coefficient estimates. This is appropriate for first-order models and in screening situations, because the experimental goal in such situations is often to identify the active factors; parameter estimation is key.
- is dependent on a pre-stated model. This is a limitation because in most real situations, the form of the pre-stated model is not known in advance.
- has runs whose purpose is to lower the variability of the coefficients of this pre-stated model. By focusing on minimizing the standard errors of coefficients, a *D*-Optimal design may not allow for checking that the model is correct. It will not include center points when investigating a first-order model. In the extreme, a *D*-Optimal design may have just p distinct runs with no degrees of freedom for lack of fit.

- maximizes D when

$$D = \det[\mathbf{X}'\mathbf{X}]$$

D -optimal split plot designs maximize D when

$$D = \det[\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}]$$

where \mathbf{V}^{-1} is the block diagonal variance matrix of the responses (Goos 2002).

Bayesian D -Optimality:

- is a modification of the D -Optimality criterion that effectively estimates the coefficients in a model, and at the same time has the ability to detect and estimate some higher-order terms. If there are interactions or curvature, the Bayesian D -Optimality criterion is advantageous.
- works best when the sample size is larger than the number of **Necessary** terms but smaller than the sum of the **Necessary** and **If Possible** terms. That is, $p + q > n > p$. The Bayesian D -Optimal design is an approach that allows the precise estimation of all of the **Necessary** terms while providing omnibus detectability (and some estimability) for the **If Possible** terms.
- uses the **If Possible** terms to force in runs that allow for detecting any inadequacy in the model containing only the **Necessary** terms. Let \mathbf{K} be the $(p + q)$ by $(p + q)$ diagonal matrix whose first p diagonal elements are equal to 0 and whose last q diagonal elements are the constant, k . If there are 2-factor interactions then $k = 4$. Otherwise $k = 1$. The Bayesian D -Optimal design maximizes the determinant of $(\mathbf{X}'\mathbf{X} + \mathbf{K})$. The difference between the criterion for D -Optimality and Bayesian D -Optimality is this constant added to the diagonal elements corresponding to the **If Possible** terms in the $\mathbf{X}'\mathbf{X}$ matrix.

I -Optimality:

- minimizes the average variance of prediction over the region of the data.
- is more appropriate than D -Optimality if your goal is to predict the response rather than the coefficients, such as in response surface design problems. Using the I -Optimality criterion is more appropriate because you can predict the response anywhere inside the region of data and therefore find the factor settings that produce the most desirable response value. It is more appropriate when your objective is to determine optimum operating conditions, and also is appropriate to determine regions in the design space where the response falls within an acceptable range. Precise estimation of the response therefore takes precedence over precise estimation of the parameters.
- minimizes this criterion: If $f'(x)$ denotes a row of the \mathbf{X} matrix corresponding to factor combinations x , then

$$I = \int_R f'(x)(\mathbf{X}'\mathbf{X})^{-1}f(x)dx = \text{Trace}[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{M}]$$

where

$$\mathbf{M} = \int_R f(x)f(x)'dx$$

is a moment matrix that is independent of the design and can be computed in advance.

Bayesian *I*-Optimality:

Bayesian *I*-Optimality has a different objective function to optimize than the Bayesian D-optimal design, so the designs that result are different. The variance matrix of the coefficients for Bayesian *I*-optimality is $\mathbf{X}'\mathbf{X} + \mathbf{K}$ where \mathbf{K} is a matrix having zeros for the **Necessary** model terms and some constant value for the **If Possible** model terms.

The variance of the predicted value at a point x_0 is:

$$\text{var}(\hat{Y}|x_0) = \left(\mathbf{x}_0'(\mathbf{X}'\mathbf{X} + \mathbf{K})^{-1} \right) \mathbf{x}_0$$

The Bayesian *I*-Optimal design minimizes the average prediction variance over the design region:

$$I_B = \text{Trace}[(\mathbf{X}'\mathbf{X} + \mathbf{K})^{-1} \mathbf{M}]$$

where \mathbf{M} is defined as before.

Alias Optimality:

- seeks to minimize the aliasing between model effects and alias effects.

Specifically, let X_1 be the design matrix corresponding to the model effects, and let X_2 be the matrix of alias effects, and let

$$\mathbf{A} = (\mathbf{X}_1' \mathbf{X}_1)^{-1} \mathbf{X}_1' \mathbf{X}_2$$

be the alias matrix.

Then, alias optimality seeks to minimize the $\text{tr}(\mathbf{A}\mathbf{A}')$, subject to the D-Efficiency being greater than some lower bound. In other words, it seeks to minimize the sum of the squared diagonal elements of \mathbf{A} .

Chapter 3

Building Custom Designs

The Basic Steps



JMP can build a *custom* design that both matches the description of your engineering problem and remains within your budget for time and material. Custom designs are general, flexible, and good for routine factor screening or response optimization. To create these tailor-made designs, use the **Custom Design** command found on the **DOE** menu or the **Custom Design** button found on the **DOE** panel of the JMP Starter.

This chapter introduces you to the steps you need to complete to build a custom design.

Contents

Creating a Custom Design	65
Enter Responses and Factors into the Custom Designer	65
Describe the Model	69
Specifying Alias Terms	70
Select the Number of Runs	71
Understanding Design Evaluation	72
Specify Output Options	78
Make the JMP Design Table	79
Creating Random Block Designs	79
Creating Split Plot Designs	80
Creating Split-Split Plot Designs	81
Creating Strip Plot Designs	82
Special Custom Design Commands	83
Save Responses and Save Factors	84
Load Responses and Load Factors	85
Save Constraints and Load Constraints	85
Set Random Seed: Setting the Number Generator	85
Simulate Responses	86
Save X Matrix: Viewing the Number of Rows in the Moments Matrix and the Design Matrix (X) in the Log	87
Optimality Criterion	88
Number of Starts: Changing the Number of Random Starts	88
Sphere Radius: Constraining a Design to a Hypersphere	90
Disallowed Combinations: Accounting for Factor Level Restrictions	90
Advanced Options for the Custom Designer	91
Assigning Column Properties	93
Define Low and High Values (DOE Coding) for Columns	94
Set Columns as Factors for Mixture Experiments	95
Define Response Column Values	96
Assign Columns a Design Role	97
Identify Factor Changes Column Property	98
How Custom Designs Work: Behind the Scenes	99

Creating a Custom Design

To begin, select **DOE > Custom Design**, or click the **Custom Design** button on the JMP Starter **DOE** page. Then, follow the steps below.

- Enter responses and factors into the custom designer.
- Describe the model.
- Select the number of runs.
- Check the design diagnostics, if desired.
- Specify output options.
- Make the JMP design table.

The following sections describe each of these steps.

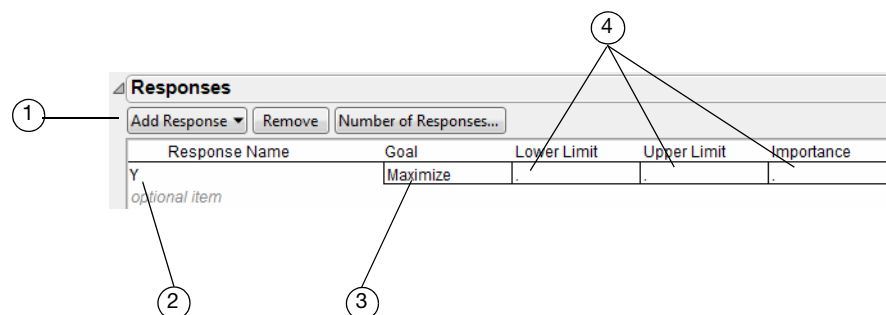
Enter Responses and Factors into the Custom Designer

How to Enter Responses

To enter responses, follow the steps in Figure 3.1.

1. To enter one response at a time, click **Add Response**, and then select a goal type. Possible goal types are **Maximize**, **Match Target**, **Minimize**, or **None**.
2. (Optional) Double-click to edit the response name.
3. (Optional) Click to change the response goal.
4. Click to enter lower and upper limits and importance weights.

Figure 3.1 Entering Responses



Tip: To quickly enter multiple responses, click **Number of Responses** and enter the number of responses you want.

Specifying Response Goal Types and Lower and Upper Limits

When entering responses, you can tell JMP that your goal is to obtain the maximum or minimum value possible, to match a specific value, or that there is no response goal.

The following description explains the relationship between the goal type (step 3 in Figure 3.1) and the lower and upper limits (step 4 in Figure 3.1):

- For responses such as strength or yield, the best value is usually the largest possible. A goal of **Maximize** supports this objective.
- The **Minimize** goal supports an objective of having the smallest value, such as when the response is impurity or defects.
- The **Match Target** goal supports the objective when the best value for a response is a specific target value, such as a dimension for a manufactured part. The default target value is assumed to be midway between the given lower and upper limits.

Note: If your target response is not equidistant from the lower and upper acceptable bounds, you can alter the default target after you make a table from the design. In the data table, open the Column Info dialog for the response column (**Cols > Column Info**) and enter the desired target value.

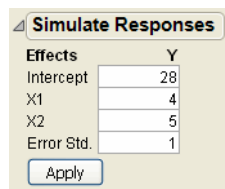
Understanding Response Importance Weights

To compute and maximize overall desirability, JMP uses the value you enter as the importance weight (step 4 in Figure 3.1) of each response. If there is only one response, then importance weight is unnecessary. With two responses you can give greater weight to one response by assigning it a higher importance value.

Adding Simulated Responses, If Desired

If you do not have values for specific responses, you might want to add simulated responses to see a prospective analysis in advance of real data collection.

1. Create the design.
2. Before you click **Make Table**, click the red triangle icon in the title bar and select **Simulate Responses**.
3. Click **Make Table** to create the design table. The Y column contains values for simulated responses.
4. For custom and augment designs, a window (Figure 3.2) appears along with the design data table. In this window, enter values you want to apply to the Y column in the data table and click **Apply**. The numbers you enter represent the coefficients in an equation. An example of such an equation, as shown in Figure 3.2, would be, $y = 28 + 4X_1 + 5X_2 + \text{random noise}$, where the *random noise* is distributed with mean zero and standard deviation one.

Figure 3.2 In Custom and Augment Designs, Specify Values for Simulated Responses


Simulate Responses

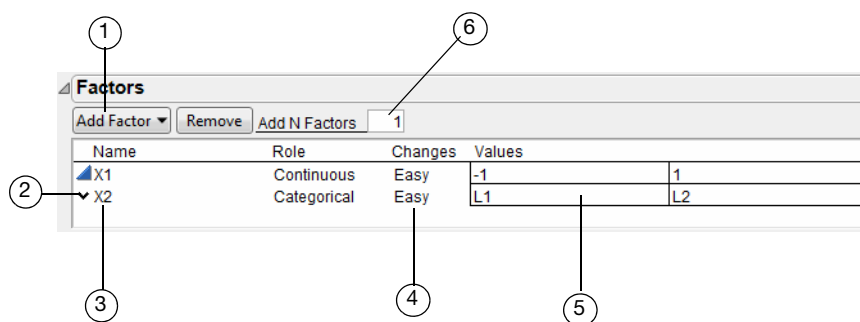
Effects	Y
Intercept	28
X1	4
X2	5
Error Std.	1

Apply

How to Enter Factors

To enter factors, follow the steps in Figure 3.3.

1. To add one factor, click **Add Factor** and select a factor type. Possible factor types are **Continuous**, **Categorical**, **Blocking**, **Covariate**, **Mixture**, **Constant**, or **Uncontrolled**. See “Types of Factors,” p. 67.
2. Click a factor and select **Add Level** to increase the number of levels.
3. Double-click a factor to edit the factor name.
4. Click to indicate that changing a factor’s setting from run to run is **Easy**, **Hard**, or **Very Hard**. Changing to Hard or Very Hard will cause the resulting design to be a split plot or split-split plot design.
5. Click to enter or change factor values. To remove a level, click it, press the delete key on the keyboard, then press the Return or Enter key on the keyboard.
6. To add multiple factors, type the number of factors in the **Add N Factors** box, click the **Add Factor** button, and select the factor type.

Figure 3.3 Entering Factors in a Custom Design


Factors

Add Factor Remove Add N Factors 1

Name	Role	Changes	Values
X1	Continuous	Easy	-1 1
X2	Categorical	Easy	L1 L2

Numbered callouts: 1 points to Add Factor, 2 points to X2, 3 points to X2, 4 points to Easy, 5 points to L1, 6 points to Add N Factors.

Types of Factors

When adding factors, click the **Add Factor** button and choose the type of factor.

Continuous Continuous factors are numeric data types only. In theory, you can set a continuous factor to any value between the lower and upper limits you supply.

Categorical Either numeric or character data types. Categorical data types have no implied order. If the values are numbers, the order is the numeric magnitude. If the values are character, the order is the sorting sequence. The settings of a categorical factor are discrete and have no intrinsic order. Examples of categorical factors are *machine*, *operator*, and *gender*.

Blocking Either numeric or character data types. Blocking factors are a special kind of categorical factor. Blocking factors differ from other categorical factors in that there is a limit to the number of runs that you can perform within one level of a blocking factor.

Covariate Either numeric or character data types. Covariate factors are not controllable, but their values are known in advance of an experiment.

Mixture Mixture factors are continuous factors that are ingredients in a mixture. Factor settings for a run are the proportion of that factor in a mixture and vary between zero and one.

Constant Either numeric or character data types. Constant factors are factors whose values are fixed during an experiment.

Uncontrolled Either numeric or character data types. Uncontrolled factors have values that cannot be controlled during an experiment, but they are factors you want to include in the model.

Factors that are Easy, Hard, or Very Hard, to Change: Creating Optimal Split-Plot and Split-Split-Plot Designs

Split plot experiments are performed in groups of runs where one or more factors are held constant within a group but vary between groups. In industrial experimentation this structure is desirable because certain factors may be difficult and expensive to change from one run to the next. It is convenient to make several runs while keeping such factors constant. Until now, commercial software has not supplied a general capability for the design and analysis of these experiments.

To indicate the difficulty level of changing a factor's setting, click in **Changes** column of the Factors panel for a given factor and select **Easy**, **Hard**, or **Very Hard** from the menu that appears. Changing to **Hard** results in a split-plot design and **Very Hard** results in a split-split-plot design.

See [“Creating Split Plot Designs,” p. 80](#), for more details.

Defining Factor Constraints, If Necessary

Sometimes it is impossible to vary factors simultaneously over their entire experimental range. For example, if you are studying the affect of cooking time and microwave power level on the number of kernels popped in a microwave popcorn bag, the study cannot simultaneously set high power and long time without burning all the kernels. Therefore, you have factors whose levels are *constrained*.

To define the constraints:



1. After you add factors and click **Continue**, click the disclosure button ( on Windows and  on the Macintosh) to open the Define Factor Constraints panel.
2. Click the **Add Constraint** button. Note that this feature is disabled if you have already controlled the design region by entering disallowed combinations or chosen a sphere radius.

Figure 3.4 Add Constraint

Custom Design

Responses

Factors

Define Factor Constraints

Add Constraint

1 X1 + 1 X2 \geq 1

3. Specify the coefficients and their limiting value in the boxes provided, as shown to the right. When you need to change the direction of the constraint, click on the default less than or equal button and select the greater than or equal to direction.
4. To add another constraint, click the **Add Constraint** button again and repeat the above steps.

Describe the Model

Initially, the Model panel lists only the main effects corresponding to the factors you entered, as shown in Figure 3.5. However, you can add factor interactions or powers of continuous factors to the model. For example, to add all the two-factor interactions and quadratic effects at once, click the **RSM** button.

Figure 3.5 Add Terms

Custom Design

Responses

Factors

Define Factor Constraints

Model

Main Effects Interactions RSM Cross Powers Remove Term

Name	Estimability
Intercept	Necessary
X1	Necessary
X2	Necessary

Table 3.1 summarizes the ways to add specific factor types to the model.

Table 3.1 How to Add Terms to a Model

Action	Instructions
Add interaction terms involving selected factors. If none are selected, JMP adds all of the interactions to the specified order.	Click the Interactions button and select 2nd , 3rd , 4th , or 5th . For example, if the factors are X1 and X2 and you click Interactions > 2nd , X1*X2 is added to the list of model terms.

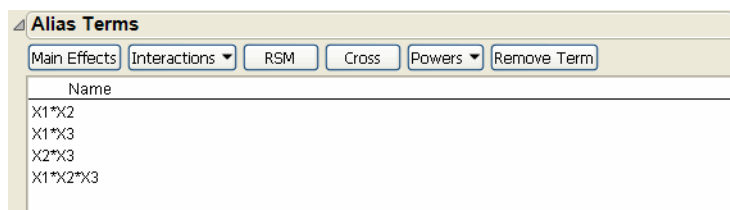
Table 3.1 How to Add Terms to a Model (*Continued*)

Action	Instructions
Add all second-order effects, including two-factor interactions and quadratic effects	Click the RSM button. The design now uses <i>I</i> -Optimality criterion rather than <i>D</i> -Optimality criterion.
Add selected cross product terms	<ol style="list-style-type: none"> 1. Highlight the factor names. 2. Highlight term(s) in the model list. 3. Click the Cross button.
Add powers of continuous factors to the model effects	Click the Powers button and select 2nd , 3rd , 4th , or 5th .

Specifying Alias Terms

You can investigate the aliasing between the model terms and terms you specify in the Alias Terms panel.

For example, suppose you specify a design with three main effects in six runs, and you want to see how those main effects are aliased by the two-way interactions and the three-way interaction. In the Alias Terms panel, specify the interactions as shown in Figure 3.6. Also, specify six runs in the Design Generation panel.

Figure 3.6 Alias Terms

After you click the **Make Design** button at the bottom of the Custom Design panel, open the Alias Matrix panel in the Design Evaluation panel to see the alias matrix. See Figure 3.7.

Figure 3.7 Aliasing

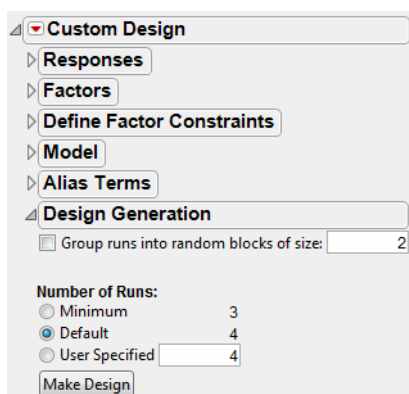
Alias Matrix					
Effect	X1*X2	X1*X3	X2*X3	X1*X2*X3	
Intercept	0	0	0.5	-0.5	
X1	0	0	-0.5	0.5	
X2	0.5	-0.5	0	0	
X3	-0.5	0.5	0	0	

In this example, all the main effects are partially aliased with two of the interactions. Also see “[The Alias Matrix \(Confounding Pattern\)](#),” p. 76.

Select the Number of Runs

The Design Generation panel (Figure 3.8) shows the minimum number of runs needed to perform the experiment based on the effects you’ve added to the model (two main effects in the example above). It also shows alternate (default) numbers of runs, or lets you choose your own number of runs. Balancing the cost of each run with the information gained by extra runs you add is a judgment call that you control.

Figure 3.8 Options for Selecting the Number of Runs



The Design Generation panel has these options for selecting the number of runs you want:

Minimum is the smallest number of terms that can create a design. When you use Minimum, the resulting design is saturated (no degrees of freedom for error). This is an extreme and risky choice, and is appropriate only when the cost of extra runs is prohibitive.

Default is a custom design suggestion for the number of runs. This value is based on heuristics for creating balanced designs with a few additional runs above the minimum.

User Specified is a value that specifies the number of runs you want. Enter that value into the Number of Runs text box.

Note: In general, the custom design suggests a number of runs that is the smallest number that can be evenly divided by the number of levels of each of the factors and is larger than the minimum possible sample size. For designs with factors at two levels only, the default sample size is the smallest power of two larger than the minimum sample size.

When the Design Generation panel shows the number of runs you want, click **Make Design**.

Understanding Design Evaluation

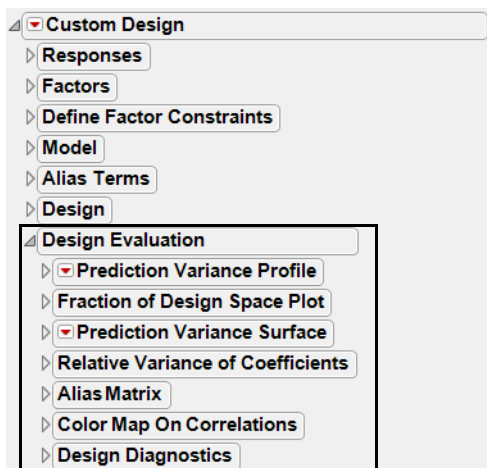
After making the design, you can preview the design and investigate details by looking at various plots and tables that serve as design diagnostic tools.

Although different tools are available depending on the model you specify, most designs display

- the Prediction Variance Profile Plot
- the Fraction of Design Space Plot
- the Prediction Variance Surface Plot
- the Relative Variance of Coefficients and Power Table
- the Alias Matrix
- Design Diagnostic Table

These diagnostic tools are outline nodes beneath the Design Evaluation panel, as shown in Figure 3.9. JMP always provides the Prediction Variance Profile, but the Prediction Surface Plot only appears if there are two or more variables.

Figure 3.9 Custom Design Evaluation and Diagnostic Tools



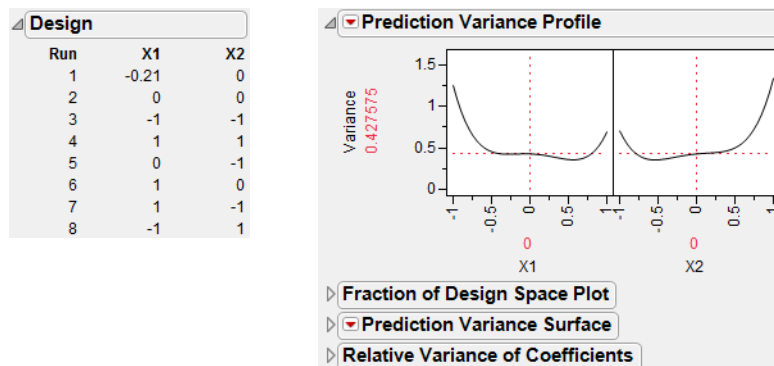
The Prediction Variance Profile

The example in Figure 3.10 shows the prediction variance profile for a response surface model (RSM) with 2 variables and 8 runs. To see a response surface design similar to this:

1. Chose **DOE > Custom Design**.
2. In the Factors panel, add 2 continuous factors.
3. Click **Continue**.
4. In the Model panel, click **RSM**.

5. Click **Make Design**.
6. Open the Prediction Variance Profile.

Figure 3.10 A Factor Design Layout For a Response Surface Design with 2 Variables



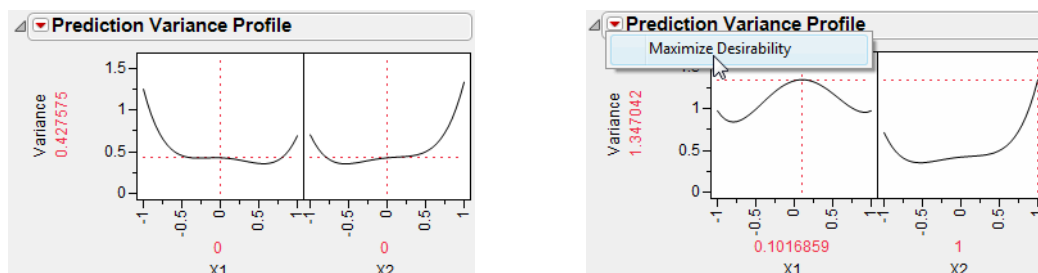
The prediction variance for any factor setting is the product of the error variance and a quantity that depends on the design and the factor setting. Before you collect the data the error variance is unknown, so the prediction variance is also unknown. However, the ratio of the prediction variance to the error variance is not a function of the error variance. This ratio, called the relative variance of prediction, depends only on the design and the factor setting and can be calculated before acquiring the data. The prediction variance profile plots the relative variance of prediction as a function of each factor at fixed values of the other factors.

After you run the experiment, collect the data, and fit the model, you can estimate the actual variance of prediction at any setting by multiplying the relative variance of prediction by the mean squared error (MSE) of the least squares fit.

It is ideal for the prediction variance to be small throughout the allowable regions of the factors. Generally, the error variance drops as the sample size increases. Comparing the prediction variance profilers for two designs side-by-side, is one way to compare two designs. A design that has lower prediction variance on the average is preferred.

In the profiler, drag the vertical lines in the plot to change the factor settings to different points. Dragging the lines reveals any points that have prediction variances that are larger than you would like.

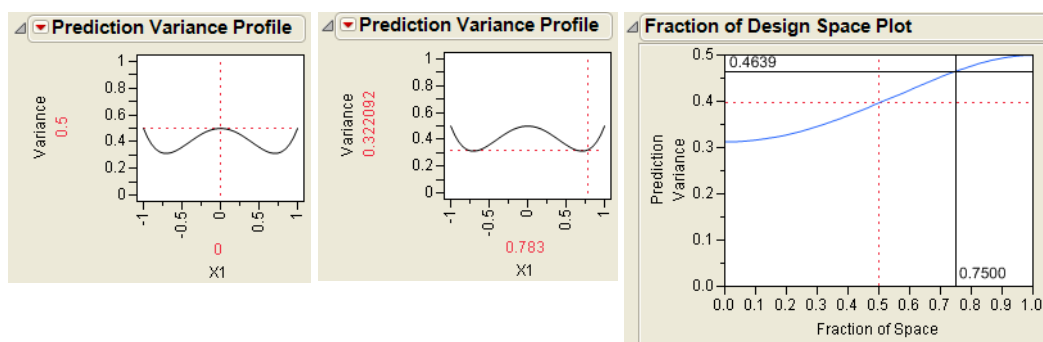
Another way to evaluate a design, or to compare designs, is to try and minimize the maximum variance. You can use the **Maximize Desirability** command on the Prediction Variance Profile title bar to identify the maximum prediction variance for a model. Consider the Prediction Variance profile for the two-factor RSM model shown in Figure 3.11. The plots on the left are the default plots. The plots on the right identify the factor values where the maximum variance (or worst-case scenario) occur, which helps you evaluate the acceptability of the model.

Figure 3.11 Find Maximum Prediction Variance

The Fraction of Design Space Plot

The Fraction of Design Space plot is a way to see how much of the model prediction variance lies above (or below) a given value. As a simple example, consider the Prediction Variance plot for a single factor quadratic model, shown on the left in Figure 3.12. The Prediction Variance plot shows that 100% of the values are smaller than 0.5. You can move the vertical trace and also see that all the values are above 0.332. The Fraction of Design Space plot displays the same information. The X axis is the proportion of prediction variance values, ranging from 0 to 100%, and the Y axis is the range of prediction variance values. In this simple example, the Fraction of Design plot verifies that 100% of the values are below 0.5 and 0% of the values are below approximately 0.3. You can use the crosshair tool and find the percentage of values for any value of the prediction variance. The example to the right in Figure 3.12 shows that 75% of the prediction variance values are below approximately 0.46.

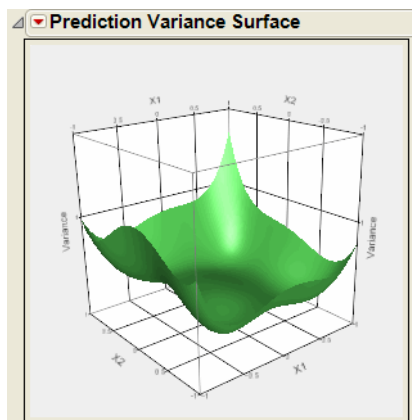
The Fraction of Design space is most useful when there are multiple factors. It summarizes the prediction variance, showing the fractional design space for all the factors taken together.

Figure 3.12 Variance Profile and Fraction of Design Space

The Prediction Variance Surface

When there are two or more factors, the Prediction Variance Surface plots the surface of the prediction variance for any two variables. This feature uses the **Graph > Surface Plot** platform in JMP, and has all its functionality. Drag on the plot to rotate and change the perspective. Figure 3.13 shows the Prediction Variance Surface plot for a two-factor RSM model. The factors are on the x and y axes, and the prediction variance is on the z axis. You can clearly see that there are high and low variance areas for both factors. Compare this plot to the Prediction Variance Profile shown in Figure 3.11.

Figure 3.13 Prediction Variance Surface Plot for Two-Factor RSM Model



You can find complete documentation for the Surface Plot platform in the *Modeling and Multivariate Methods* book.

The Relative Variance of Coefficients and Power Table

Before clicking **Make Table** in the custom designer, click the disclosure button (▶ on Windows and ►▼ on the Macintosh) to open Design Evaluation and then again to open the Relative Variance of Coefficients table.

The Relative Variance of Coefficients table (Figure 3.14) shows the relative variance of all the coefficients for the example RSM custom design (see Figure 3.10). The variances are relative to the error variance, which is unknown before the experiment, and is assumed to be one. Once you complete the experiment and have an estimate for the error variance, you can multiply it by the relative variance to get the estimated variance of the coefficient. The square root of this value should match the standard error of prediction for the coefficient when you fit a model using **Analyze > Fit Model**.

The **Power** column shows the power of the design as specified to detect effects of a certain size. In the text edit boxes, you can change the alpha level of the test and the magnitude of the effects compared to the error standard deviation. The alpha level edit box is called Significance Level. The magnitude of the effects edit box is called Signal to Noise Ratio. This is the ratio of the absolute value of the regression parameter to sigma (the square root of the error variance).

If you enter a smaller alpha (requiring a more significant test), then the power falls. If you increase the magnitude of the effect you want to detect, the power rises.

The power reported is the probability of finding a significant model parameter if the true effect is Signal to Noise Ratio times sigma. The Relative Variance of Coefficients table on the left in Figure 3.14 shows the results for the two-factor RSM model.

As another example, suppose you have a 3-factor 8-run experiment with a linear model and you want to detect any regression coefficient that is twice as large as the error standard deviation, with an alpha level of 0.05. The Relative Variance of Coefficients table on the right in Figure 3.14 shows that the resulting power is 0.984 for all the parameters.

Figure 3.14 Table of Relative Variance of Coefficients

Relative Variance of Coefficients			
Significance Level	0.05		
Signal to Noise Ratio	1		
Effect	Variance	Power	
Intercept	0.435	0.151	
X1	0.254	0.216	
X2	0.213	0.244	
X1*X1	0.704	0.114	
X1*X2	0.315	0.186	
X2*X2	0.684	0.115	

Relative Variance of Coefficients			
Significance Level	0.05		
Signal to Noise Ratio	2		
Effect	Variance	Power	
Intercept	0.125	0.984	
X1	0.125	0.984	
X2	0.125	0.984	
X3	0.125	0.984	

The Alias Matrix (Confounding Pattern)

Click the Alias Matrix disclosure button (▶ on Windows and ▶▼ on the Macintosh) to open the alias matrix (Figure 3.15).

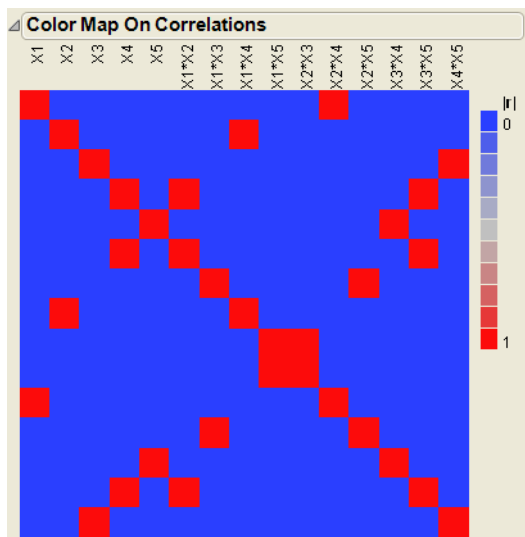
The alias matrix shows the aliasing between the model terms and the terms you specify in the Alias Terms panel (see “Specifying Alias Terms,” p. 70). It allows you to see the confounding patterns.

Figure 3.15 Alias Matrix

Alias Matrix											
Effect	X1*X2	X1*X3	X1*X4	X1*X5	X2*X3	X2*X4	X2*X5	X3*X4	X3*X5	X4*X5	
Intercept	0	0	0	0	0	0	0	0	0	0	
X1	0	0	0	0	0	-1	0	0	0	0	
X2	0	0	-1	0	0	0	0	0	0	0	
X3	0	0	0	0	0	0	0	0	0	1	
X4	-1	0	0	0	0	0	0	0	1	0	
X5	0	0	0	0	0	0	0	1	0	0	

Color Map on Correlations

The Color Map On Correlations panel (see Figure 3.16) shows the correlations between all model terms and alias terms you specify in the Alias Terms panel (see “Specifying Alias Terms,” p. 70). The colors correspond to the absolute value of the correlations.

Figure 3.16 Color Map of Correlations

The Design Diagnostics Table

Open the Design Diagnostics outline node to display a table with relative D -, G -, and A -efficiencies, average variance of prediction, and length of time to create the design. The design efficiencies are computed as follows:

$$D\text{-efficiency} = 100 \left(\frac{1}{N_D} |\mathbf{X}'\mathbf{X}|^{1/p} \right)$$

$$A\text{-efficiency} = 100 \left(\frac{p}{\text{trace}(N_D(\mathbf{X}'\mathbf{X})^{-1})} \right)$$

$$G\text{-efficiency} = 100 \left(\frac{\sqrt{\frac{p}{N_D}}}{\sigma_M} \right)$$

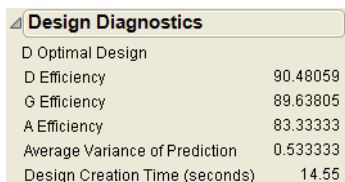
where

- N_D is the number of points in the design
- p is the number of effects in the model including the intercept
- σ_M is the maximum standard error for prediction over the design points.

These efficiency measures are single numbers attempting to quantify one mathematical design characteristic. While the maximum efficiency is 100 for any criterion, an efficiency of 100% is impossible for many design

problems. It is best to use these design measures to compare two competitive designs with the same model and number of runs rather than as some absolute measure of design quality.

Figure 3.17 Custom Design Showing Diagnostics

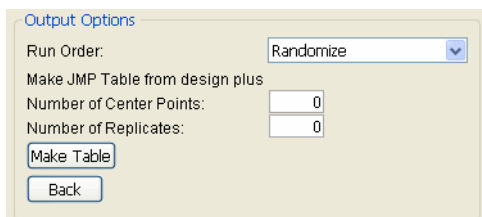


Design Diagnostics	
D Optimal Design	
D Efficiency	90.48059
G Efficiency	89.63805
A Efficiency	83.33333
Average Variance of Prediction	0.533333
Design Creation Time (seconds)	14.55

Specify Output Options

Use the Output Options panel to specify how you want the output data table to appear.

Figure 3.18 Output Options Panel



Output Options

Run Order:

Make JMP Table from design plus

Number of Center Points:

Number of Replicates:

Run Order lets you designate the order you want the runs to appear in the data table when it is created. Choices are:

- Keep the Same** the rows (runs) in the output table will appear as they do in the Design panel.
- Sort Left to Right** the rows (runs) in the output table will appear sorted from left to right.
- Randomize** the rows (runs) in the output table will appear in a random order.
- Sort Right to Left** the rows (runs) in the output table will appear sorted from right to left.
- Randomize within Blocks** the rows (runs) in the output table will appear in random order within the blocks you set up.

Add additional points using options from Make JMP Table from design plus:

- Number of Center Points:** Specifies additional runs placed at the center of each continuous factor's range.
- Number of Replicates:** Specify the number of times to replicate the entire design, including centerpoints. Type the number of times you want to replicate the design in the associated text box. One replicate doubles the number of runs.

Make the JMP Design Table

When the Design panel shows the layout you want, click **Make Table**. Parts of the table contain information you might need to continue working with the table in JMP. The upper-left of the design table can have one or more of the following scripts:

- a Screening script runs the **Analyze > Modeling > Screening** platform when appropriate for the generated design.
- a Model script runs the **Analyze > Fit Model** platform with the model appropriate for the design.
- a constraint script that shows any model constraints you entered in the Define Factor Constraints panel of the Custom Design dialog.
- a DOE Dialog script that recreates the dialog used to generate the design table, and regenerates the design table.

Figure 3.19 Example Design Table

	Design	Custom Design	Criterion	D Optimal	X1	X2	Y
1	Design	Custom Design	Criterion	D Optimal	1	-1	1
2	Screening				2	1	-1
3	Model				3	1	1
	DOE Dialog				4	1	1
					5	-1	-1
					6	1	-1

Columns (3/0)
X1 *
X2 *
Y *

1. This area identifies the design type that generated the table. Click **Custom Design** to edit.
2. Model is a script. Click the red triangle icon and select **Run Script** to open the Fit Model dialog, which is used to generate the analysis appropriate to the design.
3. DOE Dialog is a script. Click the red triangle icon and select **Run Script** to recreate the DOE Custom Dialog and generate a new design table.

Creating Random Block Designs

It is often necessary to group the runs of an experiment into blocks. Runs within a block of runs are more homogeneous than runs in different blocks. For example, the experiment described in Goos (2002), describes a pastry dough mixing experiment that took several days to run. It is likely that random day-to-day differences in environmental variables have some effect on all the runs performed on a given day. Random block designs are useful in situations like this, where there is a non-reproducible shock to the system between each block of runs. In Goos (2002), the purpose of the experiment was to understand how certain properties of the dough depend on three factors: feed flow rate, initial moisture content, and rotational

screw speed. It was only possible to conduct four runs a day. Because day-to-day variation was likely, it was important to group the runs so that this variation would not compromise the information about the three factors. Thus, blocking the runs into groups of four was necessary. Each day's experimentation was one block. The factor, Day, is an example of a random block factor.

To create a random block, use the custom design and enter responses and factors, and define your model as usual. In the Design Generation panel, check the Group runs into random blocks of size check box and enter the number of runs you want in each block. When you select or enter the sample size, the number of runs specified are assigned to the blocks.

Figure 3.20 Assigning Runs to Blocks

Design Generation

☒ Group runs into random blocks of size: 4

Number of Runs:

☐ Minimum 10

☒ Default 16

☐ User Specified 16

Make Design

In this example, the Design Generation Panel shown here designates four runs per block, and the number of runs (16) indicates there will be four days (blocks) of 4 runs. If the number of runs is not an even multiple of the random block size, some blocks will have a fewer runs than others.

Creating Split Plot Designs

Split plot experiments happen when it is convenient to run an experiment in groups of runs (called whole plots) where one or more factors stay constant within each group. Usually this is because these factors are difficult or expensive to change from run to run. JMP calls these factors **Hard** to change because this is usually how split plotting arises in industrial practice.

In a completely randomized design, any factor can change its setting from one run to the next. When certain factors are hard to change, the completely randomized design may require more changes in the settings of hard-to-change factors than desired.

If you know that a factor or two are difficult to change, then you can set the Changes setting of a factor from the default of **Easy** to **Hard**. Before making the design, you can set the number of whole plots you are willing to run.

For an example of creating a split plot design, see [“Creating a Design with Two Hard-to-Change Factors: Split Plot,” p. 55.](#)

To create a split plot design using the custom designer:

1. In the factors table there is a column called **Changes**. By default, changes are **Easy** for all factors. If, however, you click in the changes area for a factor, you can choose to make the factor **Hard** to change.

2. Once you finish defining the factors and click continue, you see an edit box for supplying the number of whole plots. You can supply any value as long as it is above the minimum necessary to fit all the model parameters. You can also leave this field empty. In this case, JMP chooses a number of whole plots to minimize the omnibus uncertainty of the fixed parameters.

Note: If you enter a missing value in the Number of Whole Plots edit box, then JMP considers many different numbers of whole plots and chooses the number that maximizes the information about the coefficients in the model. It maximizes the determinant of $X'V^{-1}X$ where V^{-1} is the inverse of the variance matrix of the responses. The matrix, V , is a function of how many whole plots there are, so changing the number of whole plots changes V , which can make a difference in the amount of information a design contains.

To create a split plot design every time you use a certain factor, save steps by setting up that factor to be “hard” in all experiments. See [“Identify Factor Changes Column Property,” p. 98](#), for details.

Creating Split-Split Plot Designs

Split-split plot designs are a three stratum extension of split plot designs. Now there are factors that are Very-Hard-to-change, Hard-to-change, and Easy-to-change. Here, in the top stratum, the Very-Hard-to-change factors stay fixed within each whole plot. In the middle stratum the Hard-to-change factors stay fixed within each subplot. Finally, the Easy-to-change factors may vary (and should be reset) between runs within a subplot. This structure is natural when an experiment covers three processing steps. The factors in the first step are Very-Hard-to-change in the sense that once the material passes through the first processing stage, these factor settings are fixed. Now the material passes to the second stage where the factors are all Hard-to-change. In the third stage, the factors are Easy-to-change.

Schoen (1999) provides an example of three-stage processing involving the production of cheese that leads to a split-split plot design. The first processing step is milk storage. Typically milk from one storage facility provides the raw material for several curds processing units—the second processing stage. Then the curds are further processed to yield individual cheeses.

In a split-split plot design the material from one processing stage passes to the next stage in such a way that nests the subplots within a whole plot. In the example above, milk from a storage facility becomes divided into two curds processing units. Each milk storage tank provided milk to a different set of curds processors. So, the curds processors were nested within the milk storage unit.

Figure 3.21 shows an example of how factors might be defined for the cheese processing example.

Figure 3.21 Example of Split-Split Response and Factors in Custom Designer Dialog

Custom Design

Responses

Add Response

Remove

Number of Responses...

Response Name	Goal	Lower Limit	Upper Limit	Importance
Consistency	None	NA	NA	NA

Factors

Add Factor

Remove

Add N Factors

1

Name	Role	Changes	Values
storage 1	Continuous	Very H	1
storage 2	Continuous	Very H	Easy
curds 1	Continuous	Hard	Hard
curds 2	Continuous	Hard	1
curds 3	Continuous	Hard	1
curds 4	Continuous	Hard	-1
curds 5	Continuous	Hard	-1
cheese 1	Continuous	Easy	-1
cheese 2	Continuous	Easy	-1

Creating Strip Plot Designs

In a strip plot design it is possible to reorder material between processing stages. Suppose units are labelled and go through the first stage in a particular order. If it is possible to collect all the units at the end of the first stage and reorder them for the second stage process, then the second stage variables are not nested within the blocks of the first stage variables. For example, in semiconductor manufacturing a boat of wafers may go through the first processing step together. However, after this step, the wafers in a given boat may be divided among many boats for the second stage.

To set up a strip plot design, enter responses and factors as usual, designating factors as Very Hard, Hard, or Easy to change. Then, in the Design Generation panel, check the box that says **Hard to change factors can vary independently of Very Hard to change factors**, as shown in Figure 3.22. Note that the Design Generation panel specified 6 whole plots, 12 subplots, and 24 runs.

When you click **Make Design**, the design table on the right in Figure 3.22 lists the run with subplots that are not nested in the whole plots.

Figure 3.22 Example of Strip Split Factors and Design Generation panel in Custom Designer Dialog

Factors
Add Factor ▼ Remove Add N Factors 1

Name	Role	Changes	Values
▲ deposition 1	Continuous	Very Hard	-1 1
▲ deposition 2	Continuous	Very Hard	-1 1
▲ etch 1	Continuous	Hard	-1 1
▲ etch 2	Continuous	Hard	-1 1
▲ etch 3	Continuous	Hard	-1 1
▲ etch 4	Continuous	Hard	-1 1
▲ clean 1	Continuous	Hard	-1 1
▲ clean 2	Continuous	Easy	-1 1
▲ clean 3	Continuous	Easy	-1 1

Define Factor Constraints
Model
Alias Terms
Design Generation
☒ Hard to change factors can vary independently of Very Hard to change factors.
Number of Whole Plots 5
Number of Subplots 10

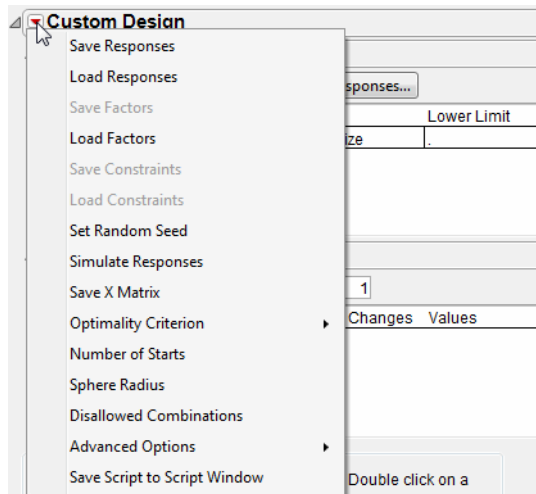
Number of Runs:
☐ Minimum 10
☒ Default 20
☐ User Specified 20
Make Design

Design

Run	Whole Plots	Subplots	deposition 1
1	1	1	-1
2	1	2	-1
3	1	3	-1
4	1	4	-1
5	2	5	-1
6	2	6	-1
7	2	7	-1
8	2	8	-1
9	3	9	1
10	3	10	1
11	3	1	1
12	3	2	1
13	4	3	1
14	4	4	1
15	4	5	1
16	4	6	1
17	5	7	1
18	5	8	1
19	5	9	1
20	5	10	1

Special Custom Design Commands

After you select **DOE > Custom Design**, click the red triangle icon on the title bar to see the list of commands available to the Custom designer (Figure 3.23). The commands found on this menu vary, depending on which DOE command you select. However, the commands to save and load responses and factors, the command to set the random seed, and the command to simulate responses are available to all designers. You should examine the red triangle menu for each designer you use to determine which commands are available. If a designer has additional commands, they are described in the appropriate chapter.

Figure 3.23 Click the Red Triangle Icon to Reveal Commands


The following sections describe these menu commands and how to use them.

Save Responses and Save Factors

If you plan to do further experiments with factors and/or responses to which you have given meaningful names and values, you can save them for later use.

To save factors or responses:

1. Select a design type from the DOE menu.
2. Enter the factors and responses into the appropriate panels (see [“Enter Responses and Factors into the Custom Designer,” p. 65](#), for details).
3. Click the red triangle icon on the title bar and select **Save Responses** or **Save Factors**.

Save Responses creates a data table containing a row for each response with a column called Response Name that identifies the responses. Four additional columns identify more information about the responses: Lower Limit, Upper Limit, Response Goal, and Importance.

Save Factors creates a data table containing a column for each factor and a row for each factor level. The columns have two column properties (noted with asterisks icons in the column panel). These properties include:

Design Role that identifies the factor as a DOE factor and lists its type (continuous, categorical, blocking, and so on).

Factor Changes that identifies how difficult it is to change the factor level. Factor Changes options are **Easy**, **Hard**, and **Very Hard**.

4. Save the data table.

Load Responses and Load Factors

If you have saved responses and factors, you can quickly apply them to your design and avoid retyping this information each time you run an experiment.

To design an experiment using responses or factors you have previously saved:

1. Open the data table that contains the factor names and levels.
2. Select a design type from the DOE menu.
3. Click the red triangle icon on the title bar and select **Load Responses** or **Load Factors**.

Tip: It is possible to create a factors table by keying data into an empty table, but remember to assign each column a factor type. Do this by right-clicking the column name, selecting **Column Info**, and then selecting **Column Properties > Design Role**. Lastly, click the button in the Design Role area and select the appropriate role.

Save Constraints and Load Constraints

In custom, augment, and mixture designs, if you set up factor constraints and plan to do further experiments with them, you can save them for later use. You can quickly apply these constraints to your design and avoid retyping this information each time you run an experiment.

To save factor constraints:

1. Select a design type from the DOE menu.
2. Enter the factor constraints into the appropriate panels (see [“Enter Responses and Factors into the Custom Designer,” p. 65](#), for details).
3. Click the red triangle icon on the title bar and select **Save Constraints**. **Save Constraints** creates a data table that contains the information you enter into a constraints panel. There is a column for each constraint. Each has a column property called Constraint State that identifies it as a ‘less than’ or a ‘greater than’ constraint. There is a row for each variable and an additional row that has the inequality condition for each variable.
4. Save the data table.

To design an experiment using factor constraints you have previously saved:

1. Open the data table that contains the constraints.
2. Select a design type from the DOE menu.
3. Click the red triangle icon on the title bar and select **Load Constraints**.

Set Random Seed: Setting the Number Generator

The design process begins with a random starting design. To set the random seed that the custom designer uses to create this starting design, click the red triangle icon in the design title bar and select **Set Random Seed**.

The window that appears shows the *generating seed* for that design (Figure 3.24). From this window, you can set a new random number and then run the design again.

If you use the same seed as a previous design, you will get the same design again.

Figure 3.24 Setting the Random Seed

Choose a positive whole number as a seed for a random starting design.

Note: The random seed is also used when you simulate responses to be used with a design, as described next.

Simulate Responses

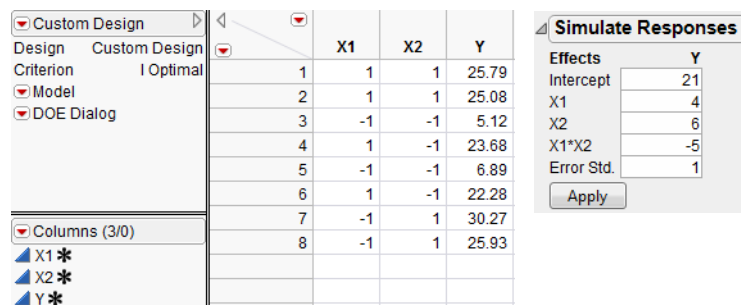
Often, when you define a custom design (or any standard design), it may be useful to look at properties of the design with response data before you have collected data. The **Simulate Responses** command adds random response values to the JMP table that the custom designer creates. To use the command, select it before you click **Make Table**. When you click **Make Table** to create the design table, the Y column contains values for simulated responses.

For custom and augment designs, an additional window appears with the design data table that lists coefficients for the design you described in the designer panels. You can enter any coefficient values you want and click **Apply** to see new Y values in the data table. An example of an equation for a model with two factors and interaction (Figure 3.25) would be,

$$y = 21 + 4X_1 + 6X_2 - 5X_1X_2 + \text{random noise},$$

where the *random noise* is distributed with mean zero and standard deviation one.

Figure 3.25 Example of a Custom Design with Simulated Responses



Save X Matrix: Viewing the Number of Rows in the Moments Matrix and the Design Matrix (X) in the Log

To create scripts for the Moments Matrix and the Design Matrix, and to save these matrices as table properties in the data table that will be generated from the design, click the red triangle icon in the Custom Design title bar (Figure 3.23) and select **Save X Matrix**. After the design and table are created, two scripts are saved as table properties and are called the Moments Matrix and the Design Matrix. Each script can be selected and run from the upper left panel of the resulting data table. Results from running each of the scripts are shown in the Log. When you run the script for the Moments Matrix, JMP shows the number of rows for the global matrix, called Moments, in the Log. Similarly, when you run the script for the Design Matrix, JMP displays the number of rows for the global matrix, called X, in the Log. The Moments Matrix and the Design Matrix are used to calculate the Average Variance of Prediction, shown in the Design Diagnostics section of the Design Evaluation. Saving these scripts to the data table provides an easy way to remember and recreate your design at a later time and to compare the matrices values with alternate designs. If you do not have the log visible, select **View > Log (Window > Log** on the Macintosh). To illustrate these features:

1. Select **DOE > Custom Design**.
2. Add 3 continuous factors and click **Continue**.
3. Click on **Interactions > 2nd** and select **Save X Matrix** from the drop-down menu of Custom Design.
4. Using the Default Number of Runs (8), click **Make Design** and then **Make Table**.
5. If it is not already open, select **View > Log (Window > Log** on the Macintosh).
6. Click on the Moments Matrix red triangle in the upper left panel of the data table under Custom Design and select **Run Script**. The result shows in the log as `N Row(:Moments):7`, which is the number of rows in the global matrix called Moments. The Moments Matrix is dependent upon the model effects but is independent of the design. (The model effects can be viewed by clicking the red triangle by Model in the upper left panel of the data table and clicking on **Run Script**.) The Moments Matrix script for this example displays the value of each moment and is shown by clicking on the red triangle of the Moments Matrix and selecting **Edit**:

```
Moments = [1 0 0 0 0 0 0,
0 0.333333333333333 0 0 0 0 0,
0 0 0.333333333333333 0 0 0 0,
0 0 0 0.333333333333333 0 0 0,
0 0 0 0 0.111111111111111 0 0,
0 0 0 0 0 0.111111111111111 0,
0 0 0 0 0 0 0.111111111111111];
```

7. Click on the Design Matrix red triangle in the upper left panel of the data table under Custom Design and select **Run Script**. The result shows in the log as `N Row(:X):8`, which is the number of rows in the global matrix called X. The X Matrix is dependent upon the design for the experiment. The script for this example shows the underlying design of the X matrix and is viewed by clicking on the red triangle of the Design Matrix and selecting **Edit**:

```
X = [1 -1 1 -1 -1 1 -1,
1 1 1 1 1 1 1,
1 -1 -1 -1 1 1 1,
1 1 -1 -1 -1 -1 1,
```

```

1 1 1 -1 1 -1 -1,
1 1 -1 1 -1 1 -1,
1 -1 1 1 -1 -1 1,
1 -1 -1 1 1 -1 -1];

```

Note that the Moments Matrix is defined as:

$$M = \int_R f(x)f(x)' dx$$

where M is a moments matrix of the parameter space that is independent of the design and can be computed in advance, and where $f(x)$ denotes a row of the design matrix corresponding to factor combinations of x . For additional details concerning moments and design matrices, see Myers, Montgomery, and Anderson-Cook (2009, pp. 365-371). Note that the moment matrix is called a matrix of region moments in this book.

Optimality Criterion

To change the design criterion, click the red triangle icon in the Custom Design title bar (Figure 3.23) and select **Optimality Criterion**, then choose one of the options:

- Make D-Optimal Design
- Make I-Optimal Design
- Make Alias Optimal Design

The default criterion for **Recommended** is D -optimal for all design types unless you have used the **RSM** button in the Model panel to add effects that make the model quadratic. For specific information about optimality criterion, see “[Technical Discussion](#),” p. 60.

Number of Starts: Changing the Number of Random Starts

To override the default number of random starts, click the red triangle icon in the Custom Design title bar (Figure 3.23) and select **Number of Starts**. When you select this command, the window shown in Figure 3.26 appears with an edit box for you to enter the number of random starts for the design you want to build. The number you enter overrides the default number of starts, which varies depending on the design.

Figure 3.26 Selecting the Number of Starts

Choose a positive whole number for the number of random starting designs.

41

OK Cancel

Note: If the design iterations are taking too long, click the **Cancel** button. The Custom Designer stops and gives the best design found at that point.

Why Change the Number of Starts?

One difficulty with the creation of optimal designs is that the methods used do not always find the globally optimal design in cases where the optimal design is known from theory. For example, orthogonal designs are D -optimal with respect to a linear additive model and a cubic design space.

As the number of factors and sample size increase, the optimization problem becomes harder. It is easy for an optimizer to converge to a local optimum instead of a global optimum.

It is useful to know that:

- If random starts are used for the optimization, the design produced at the end is not always the same. Increasing the number of random starts tends to improve the optimality of the resulting design.

- For designs with all two-level factors, there is a formula for the optimal determinant:

If D is the determinant, n is the sample size, and c is the number of columns in the design matrix, the $\text{Log}D = c \text{Log}n$.

If the determinants that result from the random starts match the formula above, the algorithm stops. The design is D -optimal and orthogonal.

Default Choice of Number of Random Starts: Technical Information

JMP does not start over with random designs until a jackpot is hit. The time it takes for one iteration of the algorithm (coordinate exchange) increases roughly as the product of the sample size and the number of terms in the model increases. By doing a large number of random starts for small sample sizes and reducing this number proportional to the square of the sample size as the designs get larger, the total time it takes to generate a design is kept roughly constant over the range of usual sample sizes.

The Custom Designer always attempts to find globally optimal designs when such designs are known from theory. For example,

- 2-level fractional factorial designs are globally D -optimal for all main effect and two-factor interaction models
- Latin-Square designs are D -optimal for main effect models assuming the right sample size and numbers of levels of the factors.
- Plackett-Burman designs are D -optimal for main effect models.

If the custom designer can identify one of these special cases, it does many more random starts. In general, however, the default number of random starts is controlled by the sample size, n , as follows:

Table 3.2 Sample Size and Random Starts

Sample Size	Number of Starts
9 or fewer	80
from 9 to 16	40
from 17 to 24	10
from 25 to 32	5

Table 3.2 Sample Size and Random Starts *(Continued)*

Sample Size	Number of Starts
more than 32	2

Note the following exceptions:

- If each factor has only two levers, the number of terms in the model is one greater than the number of factors, and the sample size is a multiple of 4, then multiply the default number of starts shown in the table above by 40.
- If the design is I-optimal divide all the Number of Starts in the table above by 2 and add 1.
- If the design could be a Latin Square, the Custom Designer does 1000 random starts.
- If the number of terms in the model is greater than 100, though, the number of random starts is 1.

After each random start, the design is checked to see if it is globally optimal, and if so, the iterations stop. Therefore, even if the default number of starts is large, it may only take a small fraction of the default number to find the globally optimal design. Again, if the process seems to be taking too long, use the **Cancel** button to see the best design found at that point.

Sphere Radius: Constraining a Design to a Hypersphere

You can constrain custom and augmented designs to a hypersphere by editing the sphere radius. Before making the design, click the red triangle icon in the Custom Design title bar (Figure 3.23) and select **Sphere Radius**. Enter the appropriate value and click **OK**.

Note that hypersphere constraints do not work with other constraints. Also, split plot designs cannot be generated with hypersphere constraints.

If you have designed any factor's changes as Hard (see [“Factors that are Easy, Hard, or Very Hard, to Change: Creating Optimal Split-Plot and Split-Split-Plot Designs,”](#) p. 68, and [“Creating Random Block Designs,”](#) p. 79), the sphere radius item becomes unavailable. Conversely, once you set the sphere radius, you cannot make a factor Hard to change.

Disallowed Combinations: Accounting for Factor Level Restrictions

JMP gives you the flexibility to disallow particular combinations of levels of factors. You can do this for custom and augmented designs except for experiments with mixture or blocking factors. This feature can also be used with continuous factors or mixed continuous and categorical factors.

For example, in a market research choice experiment, you might want to exclude a choice that allows all the best features of a product at the lowest price. In this case, the factor **Feature** has levels of worst (1), medium (2), and best (3), and the factor **Price** has levels of high (1), medium (2), and low (3). You want to exclude the third **Feature** level (best) and the third **Price** level (low).

To disallow a combination of factor levels:

1. Begin by adding the factors.

- Click the red triangle icon in the title bar (Figure 3.23) of the designer window and select **Disallowed Combinations**. Note that this menu item is not available if you have already defined linear inequality constraints.
- Enter a Boolean expression that identifies what you do not want allowed (Figure 3.27). JMP evaluates your expression, and when it sees it as true, it disallows the specified combination.

Note: When forming the expression, use the ordinal value of the level instead of the name of the level. If the level names of the factor called **Price** are high, medium, and low, their associated ordinal values are 1, 2, and 3.

For example, in Figure 3.27, **Feature==3 & Price==3** will not allow a run containing the best features at the lowest price. If there were two disallowed combinations in this example, you would use **Feature==3 & Price==3 | Quality==3 & Price==3**, which tells JMP to disallow a run with the best features at the lowest price or a run with the best quality and lowest price.

Figure 3.27 Enter a Boolean Expression

Specify constraint script or expression that returns a nonzero value for infeasible factor combinations.

Feature == 3 & Price == 3

OK Cancel

- Make the design. It excludes the combination of factors you specified, as shown in Figure 3.28.

Figure 3.28 No Row Contains L3 for Both Price and Feature

Run	Feature	Price
1	1	3
2	1	1
3	2	3
4	3	1
5	3	2
6	2	2

Advanced Options for the Custom Designer

The following options are for advanced design of experiment users.

Changing the Search Points Per Factor

For a main effects model, the coordinate exchange algorithm in the custom designer only considers the high and low values. However, you can change this so the algorithm considers more search points. For example, if you enter 5, then JMP considers five equally spaced settings for each factor. The 5 levels are considered, but all 5 levels may not appear in the output table. The Custom Designer finds a D- or I-Optimal design, which might not need to include all 5 levels.

To change the search points:

1. Select **DOE > Custom Design**.
2. Click the red triangle icon in the title bar (Figure 3.23) of the designer window and select **Advanced Options > Search Points Per Factor**.
3. Enter a positive integer and click **OK**.
4. Make the design.

Altering the Mixture Sum

If you want to keep a component of a mixture constant throughout an experiment, then the sum of the other mixture components must be between 0.001 and 1. You may have one or more fixed ingredients so that the sum of the remaining add to less than one but more than zero. To alter the mixture sum:

1. Select **DOE > Custom Design**.
2. Click the red triangle icon in the title bar (Figure 3.23) of the designer window and select **Advanced Options > Mixture Sum**.
3. Enter a positive number and click **OK**.
4. Make the design.

Split Plot Variance Ratio

The optimal split plot design depends on the ratio of the variance of the random whole plot variance to the error variance. By default, this variance is one. If you have some prior knowledge of this variance ratio, you can supply it by following these steps:

1. Select **DOE > Custom Design**.
2. Click the red triangle icon in the title bar (Figure 3.23) of the designer window and select **Advanced Options > Split Plot Variance Ratio**.
3. Enter a positive number in the resulting entry field and click **OK**.
4. Make the design.

Prior Parameter Variance

If you have specified **If Possible** as the Estimability for any factors in your model, then you can use this option to also specify the weight used for these terms. Default values are one. Larger values represent more prior information and a smaller variance. Variances are the reciprocals of the entered values.

1. Select **DOE > Custom Design**.
2. Click the red triangle icon in the title bar (Figure 3.23) of the designer window and select **Advanced Options > Prior Parameter Variance**.
3. Enter a positive number for each of the terms for which you want to specify a weight and click **OK**.
4. Make the design.

D Efficiency Weight

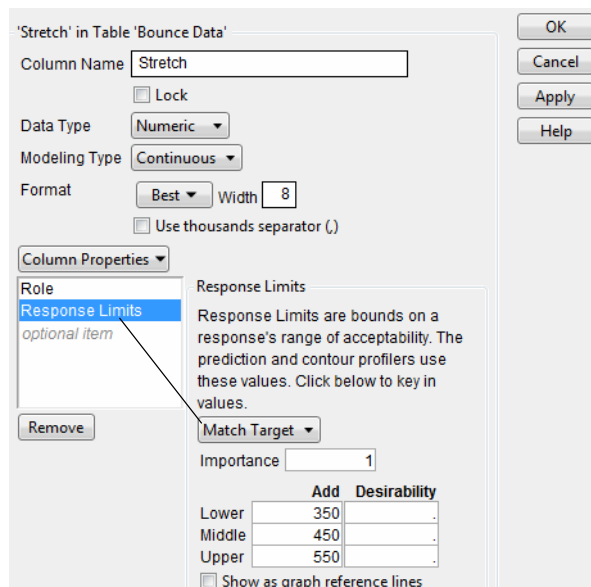
Specify the relative importance of D-Efficiency (reducing the variance of the coefficients) versus aliasing reduction. Values should be between 0 and 1, with larger values weighting more toward D-Efficiency.

Save Script to Script Window

This command creates the script for the design you described in the Custom Designer and saves it in an open script window.

Assigning Column Properties

Columns in a data table can contain special column properties. Figure 3.29 shows that a column called **Stretch** has two special properties: **Role** and **Response Limits**, that were assigned by the Custom Designer when the table was created. To see the example in Figure 3.29, open **Bounce Data.jmp** from the **Design Experiment** folder found in the sample data installed with JMP. Then, right-click the column name in the data table and select **Column Info**. When the Column Info dialog appears, click on the property you want to see.

Figure 3.29 Column Properties Menu in the Column Info Dialog


All special column properties are covered in the *Using JMP*. The following discussion gives details about properties specific to DOE and that are useful for analyzing DOE data.

Define Low and High Values (DOE Coding) for Columns

For continuous variables, the *Coding* column property transforms data in the range you specify from -1 to $+1$. When you analyze the coded variable, JMP uses those transformed data values to compute meaningful parameter estimates. You can specify the range in which the low and high values of the column are transformed.

By default, when JMP generates a design table from values entered in the Factors panel, it uses those values as the low and high values of the coding property. If a column has one or more limits missing, JMP substitutes the data's minimum and maximum for the high and low values.

You can use the Column Info dialog to manually add or delete a coding property, or change the range in which the low and high values are transformed. Figure 3.30 shows the coding values for the *Temperature* variable in the *Reactor 8 Runs* data table from the Design Experiment Sample Data.

Figure 3.30 Coding Column Property in Column Info Dialog

'Temperature' in Table 'Reactor 8 Runs'

Column Name: Temperature

☐ Lock

Data Type: Numeric

Modeling Type: Continuous

Format: Best Width: 10

☐ Use thousands separator (,)

Column Properties: Coding (selected)

Coding

Coding declares that when you fit it in a model, the range Low to High is transformed into the range -1 to 1. This property makes tests and parameters estimates more meaningful.

Low Value: 140

High Value: 180

Buttons: OK, Cancel, Apply, Help

Set Columns as Factors for Mixture Experiments

You might have a column in a data table that is one of several factors that form 100% of a mixture. You can set up the column so JMP uses it to automatically generate a no-intercept model when you analyze the data with the Fit Model platform. The following example uses the Donev Mixture Data from the Design Experiment sample data

To set up the CuS04 column as a mixture factor, first select **Cols > Column Info** to see the Column Info dialog for CuS04. Then continue as follows:

1. Select **Mixture** from the **Column Properties** drop-down menu. Upper and Lower limits, and the sum of the limits appear in a panel on the dialog, as shown in Figure 3.31. You can use these limits, or enter your own values.
2. Optionally, check the boxes beside **L PseudoComponent Coding**, **U PseudoComponent Coding**, or **both L and U PseudoComponent Coding**. Using the example in Figure 3.31, where the mixture sum value is 1, the terms are coded as:

$$X_i L = (X_i - L_i) / (1 - L) \text{ for the } L \text{ pseudocomponent}$$

$$X_i U = (U_i - X_i) / (U - 1) \text{ for the } U \text{ pseudocomponent}$$

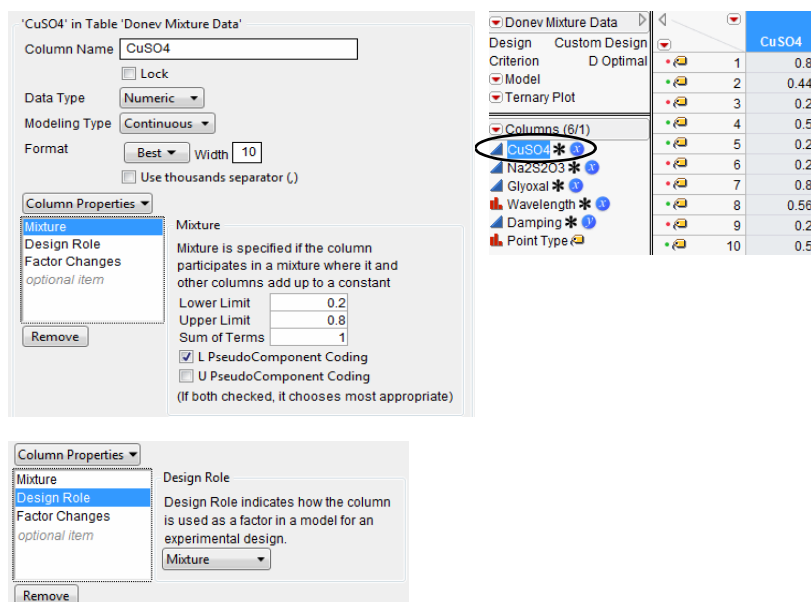
where L_i and U_i are the lower and upper bounds, L is the sum of L_i and U is the sum of U_i .

Note: If you check either **L PseudoComponent Coding** or **U PseudoComponent Coding** for the mixture coding of one mixture factor and you check the other alternative for one or more other mixture factors in the model, or if you check both boxes for one or more of the mixture factors, the Fit Model platform uses the L coding if $(1 - L) < (U - 1)$, and the U coding otherwise. If only one coding box is checked consistently for all mixture factors in the model, then only that one pseudocomponent coding is used.

In the Fit Model report, the main effects are labeled with the coding transformation. Crossed effects are not labeled, but coding values are used. All the features of fitting, such as the profilers and saved formulas, respect the pseudocomponent coding but present the un-coded values in the tables and plots.

3. Select the **Design Role** Column Property, and choose **Mixture** from its drop down menu.
4. Click **OK**. The properties icon (*) now appears next to the column name in the columns panel, indicating the column contains one or more column properties.

Figure 3.31 Column Info to Create Mixture Column For Analysis



5. Repeat the above steps for any other mixture factors that will be included in the model.

Define Response Column Values

You can save response limits in a column, which means you can run analyses without having to re-specify response limits each time. Saving these limits in a column facilitates consistency. For example, if you run an analysis that employs these limits, then come back later and change the data, you can run a new analysis using the same limits without having to reenter them. To see the example in Figure 3.32, open **Bounce Data.jmp** from the **Design Experiment** folder in the sample data installed with JMP.

Figure 3.32 shows the panel with values that specify lower, middle, and upper limits, and a desirability value. You can also select a possible goal for a DOE response variable: **Maximize**, **Match Target**, **Minimize**, or **None**. If you have more than one response, you can enter an importance value, which lets JMP know how to weigh the importance of one response against another.

To enter response limits:

1. Double-click the column name **Stretch** in the data grid. The Column Info dialog appears.
2. Select **Response Limits** from the **Column Properties** drop-down menu.
3. Select a goal for the response variable. For example, if you are in the prediction profiler and want the desired value to be close to 450, select **Match Target**.
4. When you have two responses, enter a number in the **Importance** box to indicate the amount of weight you want this response to have when JMP computes the overall desirability.
5. Enter the lower, middle and upper limits as well as the desirability values.

Figure 3.32 Completed Response Limits

'Stretch' in Table 'Bounce Data'

Column Name:

☐ Lock

Data Type:

Modeling Type:

Format: Width:

☐ Use thousands separator (,)

Column Properties:

Role: optional item

Remove

Response Limits

Response Limits are bounds on a response's range of acceptability. The prediction and contour profilers use these values. Click below to key in values.

Importance:

	Add	Desirability
Lower	<input type="text" value="350"/>	<input type="text" value="."/>
Middle	<input type="text" value="450"/>	<input type="text" value="."/>
Upper	<input type="text" value="550"/>	<input type="text" value="."/>

☐ Show as graph reference lines

OK Cancel Apply Help

6. Click **OK**. The properties icon (✱) now appears next to the column name in the column panel of the data table to indicate that the column contains a property.

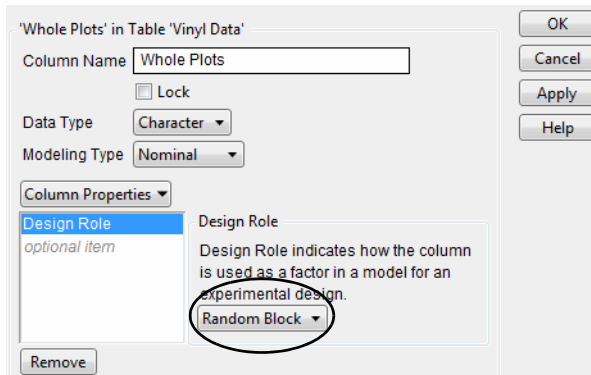
Assign Columns a Design Role

The Custom designer in JMP assigns design roles to factors when you create the design. However, you can assign a property to a column that identifies a factor column as a continuous, categorical, blocking, covariate, mixture, constant, signal, or noise factor. The example in Figure 3.33 shows the **Whole Plots** factor in the **Vinyl Data.jmp** table from the Design Experiment sample data assigned the **Random Block** design role.

To give a column a design role:

1. Double-click the column name in the data grid. The Column Info window appears.

2. Select **Design Role** from the **Column Properties** drop-down menu, as shown in Figure 3.33. Design role information appears on the right.
3. Click the Design Role drop-down menu and select how you want JMP to use the factor column: **Continuous**, **Categorical**, **Blocking**, **Covariate**, **Mixture**, **Constant**, **Signal**, **Noise**, **Uncontrolled**, or **Random Block**.
4. Click **OK** to see the property icon (✱) next to the column name in the data table's column panel.

Figure 3.33 Assign a Design Role to a Factor Column


Note: Although you can save design roles for factors, which are then automatically used each time those factors are loaded, you must always verify that the model for the design you create is correctly entered into the DOE custom designer.

Identify Factor Changes Column Property

To create split plot or split-split plot designs, you must identify a factor as having values that are hard to change, or very hard to change. This is done in the DOE design panel (see [“Creating Split Plot Designs,”](#) p. 80, for details) each time you design an experiment. However, if you know that every time you use that factor, you want it to be considered hard or very hard to change, you can save yourself steps by setting up a column property to be used in all experiments using that factor. To do this:

1. Double-click the column name in the data grid to see the Column Info dialog for that column.
2. Select **Factor Changes** from the **Column Properties** drop-down menu, as shown in Figure 3.34.
3. Click the **Factor Changes** button and select **Easy**, **Hard**, or **Very Hard** from the **Factor Changes** drop-down menu.
4. Click **OK**. The properties icon (✱) now appears next to the column name in the column panel of the data table.

Figure 3.34 Factor Changes Column Property

The screenshot shows a dialog box titled "'extrusion rate' in Table 'Vinyl Data'". It contains the following fields and controls:

- Column Name:** A text box containing "extrusion rate".
- Lock:** An unchecked checkbox.
- Data Type:** A dropdown menu set to "Numeric".
- Modeling Type:** A dropdown menu set to "Continuous".
- Format:** A dropdown menu set to "Best" and a text box for "Width" containing "10".
- Use thousands separator (,):** An unchecked checkbox.
- Column Properties:** A dropdown menu with a list of options: "Coding", "Design Role", "Factor Changes" (highlighted in blue), and "optional item".
- Factor Changes:** A sub-dropdown menu set to "Hard", which is circled in red.
- Buttons:** "OK", "Cancel", "Apply", "Help", and "Remove" are located on the right side of the dialog.

How Custom Designs Work: Behind the Scenes

The custom designer starts with a random set of points inside the range of each factor. The computational method is an iterative algorithm called *coordinate exchange* (Meyer and Nachtsheim, 1995). Each iteration of the algorithm involves testing every value of every factor in the design to determine if replacing that value increases the optimality criterion. If so, the new value replaces the old. This process continues until no replacement occurs for an entire iteration.

To avoid converging to a local optimum, the whole process is repeated several times using a different random start. The custom designer displays the best of these designs. For more details, see the section “[Optimality Criterion](#),” p. 88.

Sometimes a design problem can have several *equivalent* solutions. Equivalent solutions are designs with equal precision for estimating the model coefficients as a group. When this is true, the design algorithm may generate different (but equivalent) designs when you click the **Back** and **Make Design** buttons repeatedly.

Chapter 4

Screening Designs



Screening designs are some of the most popular designs for industrial experimentation. They examine many factors to see which have the greatest effect on the results of a process.

Screening designs generally require fewer experimental runs, which is why they cost less. Thus, they are attractive because they are a relatively inexpensive and efficient way to begin improving a process.

Often screening designs are a prelude to further experiments. It is wise to spend only about a quarter of your resource budget on an initial screening experiment. You can then use the results to guide further study.

The efficiency of screening designs depends on the critical assumption of *effect sparsity*. Effect sparsity results because real-world processes usually have only a few driving factors; other factors are relatively unimportant. To understand the importance of effect sparsity, you can contrast screening designs to full factorial designs:

- Full factorial designs consist of all combinations of the levels of the factors. The number of runs is the product of the factor levels. For example, a factorial experiment with a two-level factor, a three-level factor, and a four-level factor has $2 \times 3 \times 4 = 24$ runs.
- By contrast, screening designs reduce the number of runs by restricting the factors to two (or three) levels and by performing only a fraction of the full factorial design.

Each factor in a screening design is usually set at two levels to economize on the number of runs needed, and response measurements are taken for only a fraction of the possible combinations of levels. In the case described above, you can restrict the factors to two levels, which yield $2 \times 2 \times 2 = 8$ runs. Further, by doing half of these eight combinations you can still assess the separate effects of the three factors. So the screening approach can reduce the original 24-run experiment to four runs.

Of course, there is a price for this reduction. This chapter discusses the screening approach in detail, showing both pros and cons. It also describes how to use JMP's screening designer, which supplies a list of popular screening designs for two or more factors. These factors can be continuous or categorical, with two or three levels. The list of screening designs you can use includes designs that group the experimental runs into blocks of equal sizes where the size is a power of two.

Contents

Screening Design Examples	103
Using Two Continuous Factors and One Categorical Factor	103
Using Five Continuous Factors	105
Creating a Screening Design	109
Enter Responses	109
Enter Factors	110
Choose a Design	111
Display and Modify a Design	115
Specify Output Options	119
View the Design Table.	120
Create a Plackett-Burman design.	120
Analysis of Screening Data	122
Using the Screening Analysis Platform.	123
Using the Fit Model Platform	124

Screening Design Examples

This chapter is divided into two sections. The first section consists of two examples using screening designs. The second section outlines the procedures to follow to create a screening design to match your needs.

Using Two Continuous Factors and One Categorical Factor

Suppose an engineer wants to investigate a process that uses an electron beam welding machine to join two parts. The engineer fits the two parts into a welding fixture that holds them snugly together. A voltage applied to a beam generator creates a stream of electrons that heats the two parts, causing them to fuse. The ideal depth of the fused region is 0.17 inches. The engineer wants to study the welding process to determine the best settings for the beam generator to produce the desired depth in the fused region.

For this study, the engineer wants to explore the following three inputs, which are the *factors* for the study:

- **Operator**, who is the technician operating the welding machine
- **Rotation Speed**, which is the speed at which the part rotates under the beam
- **Beam Current**, which is a current that affects the intensity of the beam

After each processing run, the engineer cuts the part in half. This reveals an area where the two parts have fused. The **Length** of this fused area is the depth of penetration of the weld. This depth of penetration is the response for the study.

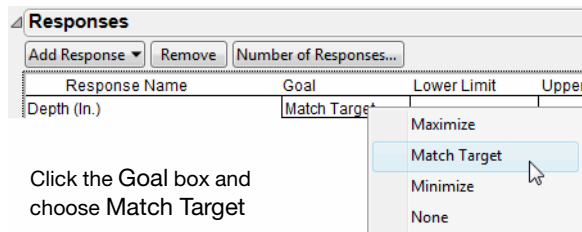
The goals of the study are to:

- find which factors affect the depth of the weld
- quantify those effects
- find specific factor settings that predict a weld depth of 0.17 inches

To begin this example, select **DOE > Screening Design** from the main menu. Note that in the Responses panel, there is a single default response called Y. Change the default response as follows:

1. Double-click the response name and change it to **Depth (In.)**.
2. The default goal for the single default response is **Maximize**, but the goal of this process is to get a target value of 0.17 inches with a lower bound of 0.12 and an upper bound of 0.22. Click the **Goal** box and choose **Match Target** from the drop-down menu, as shown in Figure 4.1.

Figure 4.1 Screening Design Response With Match Target Goal

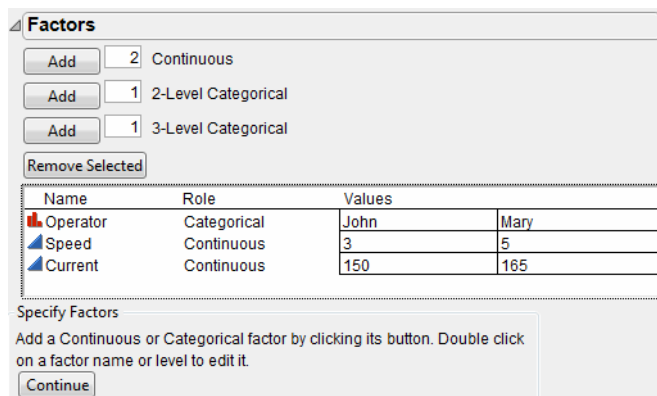


3. Click the **Lower Limit** text edit area and enter 0.12 as the lower limit (minimum acceptable value). Then click the **Upper Limit** text edit area and enter 0.22 as the upper limit (maximum acceptable value).

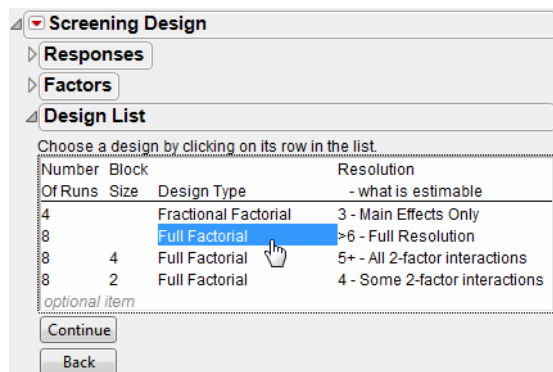
This example has one categorical factor (**Operator**) and two continuous factors (**Speed** and **Current**).

4. Add the categorical factor by clicking the **Add** button beside **2-Level Categorical**.
5. Add two continuous factors by typing **2** in the **Continuous** box and clicking the associated **Add** button.
6. Double-click the factor names and rename them **Operator**, **Speed**, and **Current**.
7. Set high and low values for **Speed** to 3 and 5 rpm. Set high and low values for **Current** to 150 and 165 amps, and assign John and Mary as values for the categorical factor called **Operator**, as shown in Figure 4.2.

Figure 4.2 Screening Design with Two Continuous and One Categorical Factor



8. Click **Continue**.
9. Select **Full Factorial** in the list of designs, as shown in Figure 4.3, and then click **Continue**.

Figure 4.3 List of Screening Designs for Two Continuous and One Categorical Factors

In the Output Options section of the Display and Modify Design panel that appears, click on the **Run Order** menu and select **Sort Left to Right**, which arranges the runs in the JMP design data table (see Figure 4.7). Then click **Make Table** to create the JMP table that contains the specified design.

The table in Figure 4.4 appears. The table uses the names for responses, factors, and levels you specified. The **Pattern** variable shows the coded design runs. You can also see the table produced in this example by selecting **Help > Sample Data > Design of Experiments > DOE Example 1**. (You can also open DOE Example 1.jmp from the sample data directory.)

Figure 4.4 The Design Data Table

Design	Pattern	Operator	Speed	Current	Depth (In.)
1	---	John	3	150	•
2	--+	John	3	165	•
3	-+-	John	5	150	•
4	+++	John	5	165	•
5	+--	Mary	3	150	•
6	++-	Mary	3	165	•
7	++-	Mary	5	150	•
8	+++	Mary	5	165	•

Using Five Continuous Factors

As illustrated in the previous section, experiments for screening the effects of many factors usually consider only two levels of each factor. This allows the examination of many factors with a minimum number of runs.

The following example, adapted from Meyer, *et al.* (1996), demonstrates how to use JMP's screening designer when you have many factors. In this study, a chemical engineer investigates the effects of five factors on the percent reaction of a chemical process. The factors are:

- feed rate, the amount of raw material added to the reaction chamber in liters per minute
- percentage of catalyst
- stir rate, the RPMs of a propeller in the chamber
- reaction temperature in degrees Celsius
- concentration of reactant

To start the example:

1. Select **DOE > Screening Design**.
2. You see one default response called Y. Change the default response name (Y) to **Percent Reacted**.
3. The **Goal** is to maximize the response, but change the minimum acceptable reaction percentage to 90 (**Lower Limit**), and upper limit to 99 (**Upper Limit**), as shown in Figure 4.5.
4. Add five continuous factors.
5. Change the default factor names (X1-X5) to **Feed Rate**, **Catalyst**, **Stir Rate**, **Temperature**, and **Concentration**.
6. Enter the high and low values, as shown in Figure 4.5.

Figure 4.5 Screening for Many Factors

Screening Design

Responses

Add Response Remove Number of Responses...

Response Name	Goal	Lower Limit	Upper Limit	Importance
Percent Reacted	Maximize	90	99	.

optional item

Factors

Add 5 Continuous

Add 1 2-Level Categorical

Add 1 3-Level Categorical

Remove Selected

Name	Role	Values
Feed Rate	Continuous	10 15
Catalyst	Continuous	1 2
Stir Rate	Continuous	100 120
Temperature	Continuous	140 180
Concentration	Continuous	3 6

Specify Factors

Add a Continuous or Categorical factor by clicking its button. Double click on a factor name or level to edit it.

Continue

7. Click **Continue**. Now, JMP lists the designs for the number of factors you specified, as shown to the left in Figure 4.6.
8. Select the first item in the list, which is an 8-run fractional factorial design with no blocks.

9. Click **Continue** to see the Output Options outline on the right in Figure 4.6.

Figure 4.6 Two-level Screening Design (left) and design output options (right)

Screening Design

Responses

Factors

Design List

Choose a design by clicking on its row in the list.

Number Of Runs	Block Size	Design Type	Resolution - what is estimable
8		Fractional Factorial	3 - Main Effects Only
8	4	Fractional Factorial	3 - Main Effects Only
12		Plackett-Burman	3 - Main Effects Only
16		Fractional Factorial	5 - All 2-factor interactions
16	8	Fractional Factorial	4 - Some 2-factor interactions
16	4	Fractional Factorial	4 - Some 2-factor interactions
16	2	Fractional Factorial	4 - Some 2-factor interactions
32		Full Factorial	>6 - Full Resolution
32	16	Full Factorial	5+ - All 2-factor interactions
32	8	Full Factorial	5+ - All 2-factor interactions
32	4	Full Factorial	4 - Some 2-factor interactions
32	2	Full Factorial	4 - Some 2-factor interactions

optional item

Continue **Back**

Output Options

Run Order: **Sort Left to Right**

Make JMP Table from design plus

Number of Center Points:

Number of Replicates:

Make Table **Back**

The design dialog has options shown in Figure 4.7 that can modify the final design table.

10. Select **Sort Left to Right** from the Run Order Menu.

Figure 4.7 Output Options for Design Table

Output Options

Run Order: **Sort Left to Right**

Make JMP Table from design plus

Number of Center Points:

Number of Replicates:

Make Table **Back**

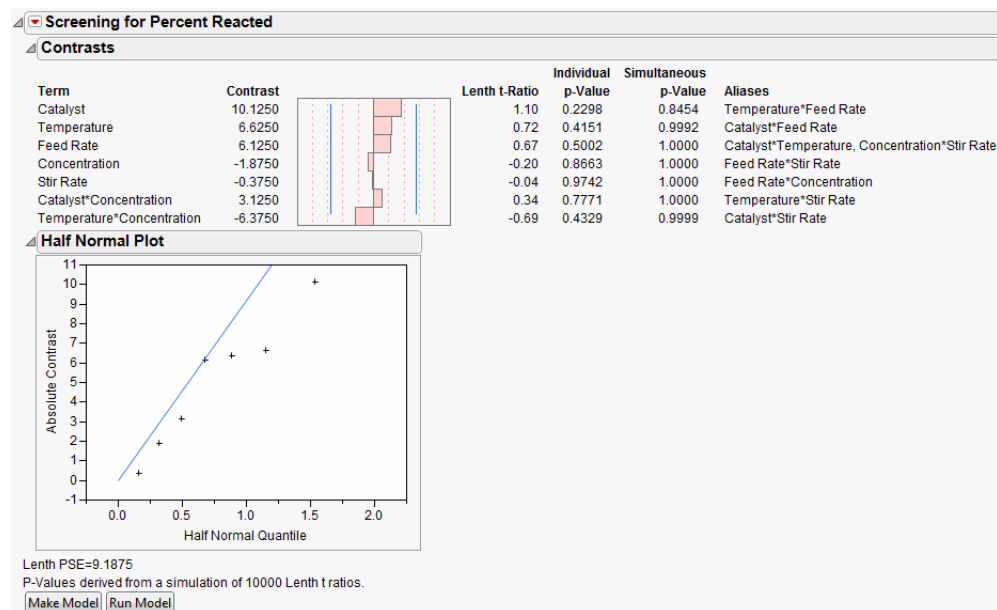
11. Click **Make Table** to create the data table shown in Figure 4.8 that lists the runs for the design you selected. Note that it also has a column called **Percent Reacted** for recording experimental results, showing as the rightmost column of the data table.

Figure 4.8 JMP Table of Runs for Screening Example

Pattern	Feed Rate	Catalyst	Stir Rate	Temperature	Concentration	Percent Reacted
1 ----+	10	1	100	140	6	•
2 ---++	10	1	120	180	3	•
3 --+++	10	2	100	180	3	•
4 -++++	10	2	120	140	6	•
5 +----	15	1	100	180	6	•
6 +++--	15	1	120	140	3	•
7 ++---	15	2	100	140	3	•
8 +++++	15	2	120	180	6	•

Of the five factors in the reaction percentage experiment, you expect a few to stand out in comparison to the others. Let's take an approach to the analysis that looks for active effects.

- To run the model generated by the data shown in Figure 4.8, open **Reactor 8 Runs.jmp** from the Design Experiment folder found in the sample data that was installed with JMP. This table has the design runs and the results of the experiment.
- In the design data table, click the **Screening** script that shows on the upper left of the data table, and select **Run Script**. Or, you can choose **Analyze > Modeling > Screening** to analyze the data. Select **Percent Reacted** as *Y* and all other continuous variables as *X*. Click **OK**. The report is shown in Figure 4.9.

Figure 4.9 Report for Screening Example

Note: Analysis of the screening data is covered in the section “[Create a Plackett-Burman design,](#)” p. 120 at the end of this chapter.

Creating a Screening Design

To begin, select **DOE > Screening Design**, or click the **Screening Design** button on the JMP Starter DOE page. Then, see the following sections for each step to create a screening design:

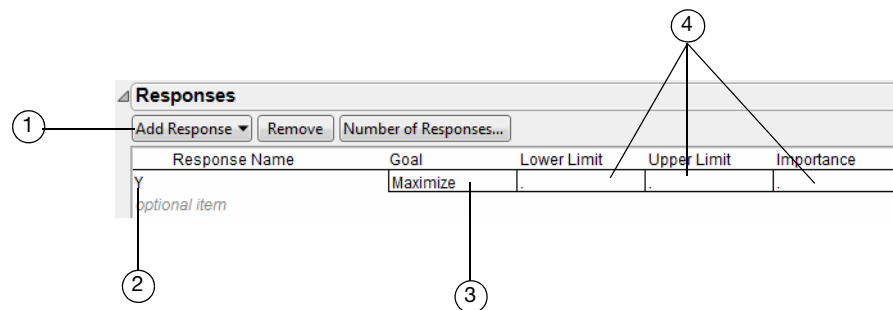
1. “[Enter Responses,](#)” p. 109
2. “[Enter Factors,](#)” p. 110
3. “[Choose a Design,](#)” p. 111
4. “[Display and Modify a Design,](#)” p. 115
5. “[Specify Output Options,](#)” p. 119
6. “[View the Design Table,](#)” p. 120

Enter Responses

To enter responses, follow the steps in Figure 4.10.

1. To enter one response at a time, click and then select a goal type. Possible goal types are **Maximize**, **Match Target**, **Minimize**, or **None**.
2. (Optional) Double-click to edit the response name.
3. (Optional) Click to change the response goal.
4. Click to enter lower and upper limits and importance weights.

Figure 4.10 Entering Responses



Tip: To quickly enter multiple responses, click the **Number of Responses** button and enter the number of responses you want.

Specifying Goal Types and Lower and Upper Limits

When entering responses, you can tell JMP that your goal is to obtain the maximum or minimum value possible, to match a specific value, or that there is no goal.

The following description explains the relationship between the goal type (step 3 in Figure 4.10) and the lower and upper limits (step 4 in Figure 4.10):

- For responses such as strength or yield, the best value is usually the largest possible. A goal of **Maximize** supports this objective.
- The **Minimize** goal supports an objective of having the smallest value be the most desirable, such as when the response is impurity or defects.
- The **Match Target** goal supports the objective when the best value for a response is a specific target value, such as dimensions of a manufactured part. The default target value is assumed to be midway between the lower and upper limits.

Note: If your target range is not symmetric around the target value, you can alter the default target after you make a table from the design. In the data table, open the response's Column Info dialog by double-clicking the column name, and enter an asymmetric target value.

Understanding Importance Weights

When computing overall desirability, JMP uses the value you enter as the importance weight (step 4 in Figure 4.10) as the weight of each response. If there is only one response, then specifying importance is unnecessary. With two responses you can give greater weight to one response by assigning it a higher importance value.

Enter Factors

Next, you enter factors. The Factors panel's appearance depends on the design you select. Entering factors is the same in **Screening Design**, **Space Filling Design**, **Mixture Design**, and **Response Surface Design**. This process is described below, in Figure 4.11.

1. To enter factors, type the number of factors and click **Add**.
2. Highlight the factor and click the **Remove Selected** button to remove a factor in the list.
3. Double-click to edit the factor name.
4. Click to enter factor values. To remove a level, click it, press the delete key on your keyboard, then press the Return or Enter key on your keyboard.

Figure 4.11 Entering Factors

Name	Role	Values
X1	Continuous	-1 1

Types of Factors

In general, when designing experiments, you can enter different types of factors in the model. Below is a description of factor types from which you can choose when creating screening designs:

Continuous Continuous factors have numeric data types only. In theory, you can set a continuous factor to any value between the lower and upper limits you supply.

Categorical Categorical factors (either numeric or categorical data types) have no implied order. If the values are numbers, the order is the numeric magnitude. If the values are character, the order is the sorting sequence. The settings of a categorical factor are discrete and have no intrinsic order. Examples of categorical factors are machine, operator, and gender.

After you enter responses and factors, click **Continue**.

Choose a Design

The list of screening designs you can use includes designs that group the experimental runs into blocks of equal sizes where the size is a power of two. Highlight the type of screening design you want to use and click **Continue**.

Figure 4.12 Choosing a Type of Screening Design

Number	Block	Size	Design Type	Resolution
4			Fractional Factorial	3 - Main Effects Only
8			Full Factorial	>6 - Full Resolution
8	4		Full Factorial	5+ - All 2-factor interactions
8	2		Full Factorial	4 - Some 2-factor interactions

optional item

The screening designer provides the following types of designs:

Two-Level Full Factorial

A full factorial design has runs for all combinations of the levels of the factors. The sample size is the product of the levels of the factors. For two-level designs, this is 2^k where k is the number of factors. This can be expensive if the number of factors is greater than 3 or 4.

These designs are orthogonal. This means that the estimates of the effects are uncorrelated. If you remove an effect in the analysis, the values of the other estimates remain the same. Their p -values change slightly, because the estimate of the error variance and the degrees of freedom are different.

Full factorial designs allow the estimation of interactions of all orders up to the number of factors. Most empirical modeling involves first- or second-order approximations to the true functional relationship between the factors and the responses. The figure to the left in Figure 4.13 is a geometric representation of a two-level factorial.

Two-Level Fractional Factorial

A fractional factorial design also has a sample size that is a power of two. If k is the number of factors, the number of runs is 2^{k-p} where $p < k$. The fraction of the full factorial is 2^p . Like the full factorial, fractional factorial designs are orthogonal.

The trade-off in screening designs is between the number of runs and the *resolution* of the design. If price is no object, you can run several replicates of all possible combinations of m factor levels. This provides a good estimate of everything, including interaction effects to the m th degree. But because running experiments costs time and money, you typically only run a fraction of all possible levels. This causes some of the higher-order effects in a model to become *nonestimable*. An effect is nonestimable when it is confounded with another effect. In fact, fractional factorials are designed by deciding in advance which interaction effects are confounded with the other interaction effects.

Resolution Number: The Degree of Confounding

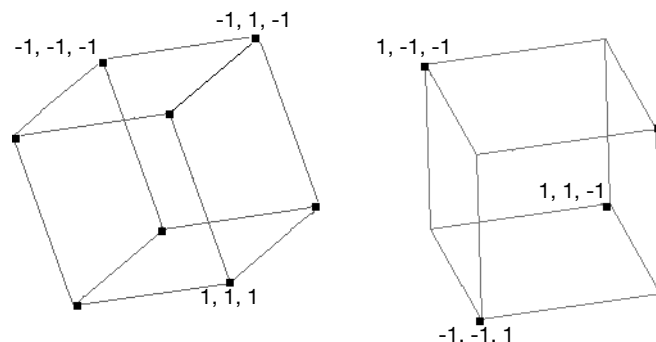
In practice, few experimenters worry about interactions higher than two-way interactions. These higher-order interactions are assumed to be zero.

Experiments can therefore be classified by *resolution number* into three groups:

- Resolution = 3 means that main effects are confounded with one or more two-way interactions, which must be assumed to be zero for the main effects to be meaningful.
- Resolution = 4 means that main effects are not confounded with other main effects or two-factor interactions. However, two-factor interactions are confounded with other two-factor interactions.
- Resolution ≥ 5 means there is no confounding between main effects, between two-factor interactions, or between main effects and two-factor interactions.

A *minimum aberration* design is one in which there are a minimum number of confoundings for a given resolution. For DOE experts, the minimum aberration design of a given resolution minimizes the number of words in the defining relation that are of minimum length.

The figure on the right in Figure 4.13 is geometric representation of a two-level fractional factorial design.

Figure 4.13 Representation of Full Factorial (Left) and Two-Level Fractional Factorial (Right) Designs

Plackett-Burman Designs

Plackett-Burman designs are an alternative to fractional factorials for screening. One useful characteristic is that the sample size is a multiple of four rather than a power of two. There are no two-level fractional factorial designs with sample sizes between 16 and 32 runs. However, there are 20-run, 24-run, and 28-run Plackett-Burman designs.

The main effects are orthogonal and two-factor interactions are only partially confounded with main effects. This is different from resolution-three fractional factorial where two-factor interactions are indistinguishable from main effects.

In cases of effect sparsity, a stepwise regression approach can allow for removing some insignificant main effects while adding highly significant and only somewhat correlated two-factor interactions. The new Screening platform in JMP, **Analyze > Modeling > Screening**, is a streamlined approach for looking at sparse data. This platform can accept multiple responses and multiple factors, then automatically fits a two-level design and shows significant effects with plots and statistics. See the chapter in the *Modeling and Multivariate Methods* book on the Screening platform for more information.

Mixed-Level Designs

If you have qualitative factors with three values, then none of the classical designs discussed previously are appropriate. For pure three-level factorials, JMP offers fractional factorials. For mixed two-level and three-level designs, JMP offers complete factorials and specialized orthogonal-array designs, listed in Table 4.1

If you have fewer than or equal to the number of factors for a design listed in the table, you can use that design by selecting an appropriate subset of columns from the original design. Some of these designs are not balanced, even though they are all orthogonal.

Table 4.1 Table of Mixed-Level Designs

Design	Two-Level Factors	Three-Level Factors
L18 John	1	7

Table 4.1 Table of Mixed-Level Designs (*Continued*)

Design	Two-Level Factors	Three-Level Factors
L18 Chakravarty	3	6
L18 Hunter	8	4
L36	11	12

Cotter Designs

Cotter designs are used when you have very few resources and many factors, and you believe there may be interactions. Suppose you believe in effect sparsity—that very few effects are truly nonzero. You believe in this so strongly that you are willing to bet that if you add up a number of effects, the sum will show an effect if it contains an active effect. The danger is that several active effects with mixed signs will cancel and still sum to near zero and give a false negative.

Cotter designs are easy to set up. For k factors, there are $2k + 2$ runs. The design is similar to the “vary one factor at a time” approach many books call inefficient and naive.

A Cotter design begins with a run having all factors at their high level. Then follow k runs each with one factor in turn at its low level, and the others high. The next run sets all factors at their low level and sequences through k more runs with one factor high and the rest low. This completes the Cotter design, subject to randomizing the runs.

When you use JMP to generate a Cotter design, the design also includes a set of extra columns to use as regressors. These are of the form *factor*Odd and *factor*Even where *factor* is a factor name. They are constructed by adding up all the odd and even interaction terms for each factor. For example, if you have three factors, A, B, and C:

Table 4.2 Cotter Design Table

AOdd = A + ABC	AEven = AB + AC
BOdd = B + ABC	BEven = AB + BC
COdd = C + ABC	CEven = BC + AC

Because these columns in a Cotter design make an orthogonal transformation, testing the parameters on these combinations is equivalent to testing the combinations on the original effects. In the example of factors listed above, AOdd estimates the sum of odd terms involving A. AEven estimates the sum of the even terms involving A, and so forth.

Because Cotter designs have a false-negative risk, many statisticians discourage their use.

How to Run a Cotter Design

By default, JMP does not include a Cotter design in the list of available screening designs (Figure 4.12). However, if you want to make a Cotter design:

1. Immediately after entering responses and factors (and before clicking **Continue**), click the red triangle icon in the Screening Design title bar.
2. Select **Suppress Cotter Designs** (to uncheck it).

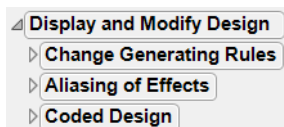
Changing the setting via the red triangle menu applies only to the *current* design. To alter the setting for *all* screening designs:

1. Select **File > Preferences**.
2. Click the **Platforms** icon.
3. Click **DOE** to highlight it.
4. Uncheck the box beside **Suppress Cotter Designs**.

Display and Modify a Design

After you select a design type, open the Display and Modify Design outline.

Figure 4.14 Display and Modification Options



Change Generating Rules Controls the choice of different fractional factorial designs for a given number of factors.

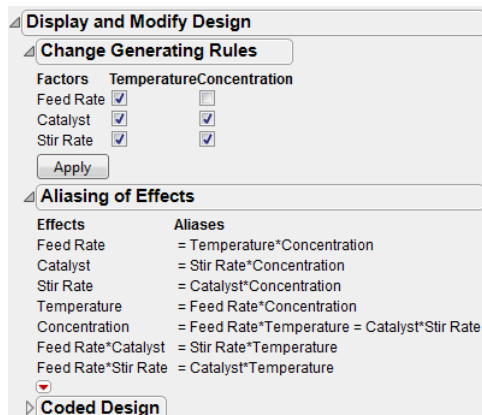
Aliasing of Effects Shows the confounding pattern for fractional factorial designs.

Coded Design Shows the pattern of high and low values for the factors in each run.

Aliasing of Effects

To see which effects are confounded with which other effects, open the Aliasing of Effects outline. It shows effects and confounding up to two-factor interactions.

Figure 4.15 Generating Rules and Aliasing of Effects Panel



For example, a full factorial with five factors requires $2^5 = 32$ runs. Eight runs can only accommodate a full factorial with three two-level factors. It is necessary to construct the two additional factors in terms of the first three factors.

The price of reducing the number of runs from 32 to eight is effect aliasing (confounding). Confounding is the direct result of the assignment of new factor values to products of the coded design columns.

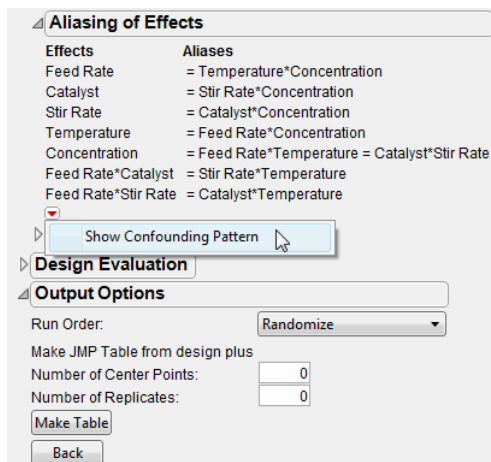
In the example above, the values for **Temperature** are the product of the values for **Feed Rate** and **Concentration**. This means that you can't tell the difference of the effect of **Temperature** and the synergistic (interactive) effect of **Feed Rate** and **Concentration**.

In the example shown in Figure 4.15, all the main effects are confounded with two-factor interactions. This is characteristic of resolution-three designs.

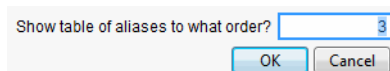
Look at the Confounding Pattern

JMP can create a data table that shows the aliasing pattern for a specified level. To create this table:

1. Click the red triangle at the bottom of the Aliasing of Effects area.
2. Select **Show Confounding Pattern** (Figure 4.16).

Figure 4.16 Show Confounding Patterns

3. Enter the order of confounding you want to see (Figure 4.17).

Figure 4.17 Enter Order of Confounding in Text Edit Box

4. Click **OK**.

Figure 4.18 shows the third order aliasing for the five-factor reactor example. The effect names begin with C (Constant) and are shown by their order number in the design. Thus, *Temperature* appears as “4”, with second order aliasing as “1 5” (Feed Rate and Concentration), and third order confounding as “1 2 3” (Feed Rate, Catalyst, and Stir Rate).

Figure 4.18 The Third Level Alias for the Five-Factor Reactor Example

Confounding Pa...		Effect Names	Alias Names
		1 C	= 2 3 5 = 1 4 5
		2 1	= 2 3 4 = 4 5
		3 2	= 1 3 4 = 3 5
		4 12	= 3 4 = 1 3 5 = 2 4 5
		5 3	= 1 2 4 = 2 5
		6 13	= 2 4 = 1 2 5 = 3 4 5
		7 23	= 1 4 = 5
		8 12 3	= 4 = 1 5
		9 4	= 1 2 3 = 1 5
		10 14	= 2 3 = 5
		11 24	= 1 3 = 1 2 5 = 3 4 5
		12 12 4	= 3 = 2 5
		13 34	= 1 2 = 1 3 5 = 2 4 5

Understanding Design Codes

In the Coded Design panel, each row represents a run. Plus signs designate high levels and minus signs represent low levels. As shown in Figure 4.19, rows for the first three columns of the coded design, which represent **Feed Rate**, **Catalyst**, and **Stir Rate** are all combinations of high and low values (a full factorial design). The fourth column (**Temperature**) of the coded design is the element-by-element product of the first three columns. Similarly, the last column (**Concentration**) is the product of the second and third columns.

Figure 4.19 Default Coded Designs

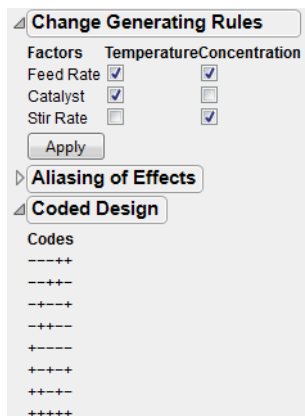
Coded Design		Feed Rate		
Codes		Catalyst	Temperature	Concentration
----		Stir Rate		
--++		----	-	+
-++-		----	+	-
++--		----	+	-
++++		----	+	+
----		----	-	+
--++		----	+	-
-++-		----	+	-
++--		----	+	+
----		----	-	+
--++		----	+	-
-++-		----	+	-
++--		----	+	+

Changing the Coded Design

In the Change Generating Rules panel, changing the check marks and clicking **Apply** changes the coded design; it changes the choice of different fractional factorial designs for a given number of factors. The Coded Design table in Figure 4.19 shows how the last two columns are constructed in terms of the first three columns. The check marks in the Change Generating Rules table shown in Figure 4.20 for **Temperature** now show it is a function of **Feed Rate**, and **Catalyst**. The check marks for **Concentration** show it is a function of **Feed Rate** and **Stir Rate**.

If you check the options as shown in Figure 4.20 and click **Apply**, the Coded Design panel changes. The first three columns of the coded design remain a full factorial for the first three factors (Feed Rate, Catalyst, and Stir Rate). Temperature is now the product of Feed Rate and Catalyst, so the fourth column of the coded design is the element by element product of the first two columns. Concentration is a function of Feed Rate and Stir Rate.

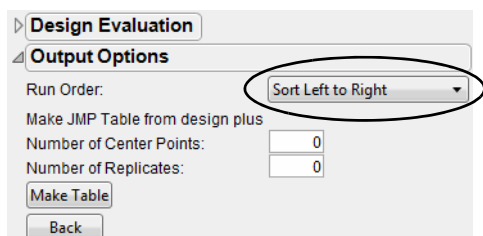
Figure 4.20 Modified Coded Designs and Generating Rules



Specify Output Options

Use the Output Options panel (Figure 4.21) to specify how you want the output data table to appear. When the options are the way you want them, click **Make Table**.

Figure 4.21 Select the Output Options



Create a Plackett-Burman design

Run Order Lets you designate the order you want the runs to appear in the data table when it is created. Choices are:

Keep the Same the rows (runs) in the output table appear as they do in the Coded Design panel.

Sort Left to Right the rows (runs) in the output table appear sorted from left to right.

Randomize the rows (runs) in the output table appear in a random order.

Sort Right to Left the rows (runs) in the output table appear sorted from right to left.

Randomize within Blocks the rows (runs) in the output table will appear in random order within the blocks you set up.

Number of Center Points Specifies additional runs placed at the center points.

Number of Replicates Specify the number of times to replicate the entire design, including centerpoints. Type the number of times you want to replicate the design in the associated text box. One replicate doubles the number of runs.

View the Design Table

Click **Make Table** to create a data table that contains the runs for your experiment. In the table, the high and low values you specified are displayed for each run.

Figure 4.22 The Design Data Table

Fractional Factorial								
Design Fractional Factorial								
Screening								
Model								
Columns (7/0)								
Pattern								
Feed Rate *								
Catalyst *								
Stir Rate *								
Temperature *								
Concentration *								
Percent Reacted *								

The name of the table is the design type that generated it. Run the Screening script to screen for active effects. The column called **Pattern** shows the pattern of low values denoted “-” and high values denoted “+”. **Pattern** is especially useful as a label variable in plots.

Create a Plackett-Burman design

The previous example shows an 8-run fractional factorial design for five continuous factors. But suppose you can afford 4 additional runs. First, repeat the steps shown in the previous sections. This time, use the Load Responses and Load Factors commands to define the design, as follows:

1. Select **DOE > Screening Design**.

2. Select **Load Responses** from the red triangle menu on the Screening Design title bar. Navigate to the Design Experiment folder in the Sample Data installed with JMP and open the file called Reactor Response.jmp.
3. Next, select **Load Factors** from red triangle menu on the Screening Design title bar. Navigate to the Design Experiment folder in the Sample Data installed with JMP and open the file called Reactor Factors.jmp.

These two commands complete the DOE screening dialog for you, with the correct response and factor names, goal and limits for the response, and the values for the factors.

4. Click **Continue** on the completed Screening design dialog to see the list of designs in Figure 4.23, and chose the Plackett-Burman, as shown.

Figure 4.23 Design List for 5-factor Plackett-Burman Screening Design

Screening Design

Responses

Factors

Name	Role	Values
Feed Rate	Continuous	10 15
Catalyst	Continuous	1 2
Stir Rate	Continuous	100 120
Temperature	Continuous	140 180
Concentration	Continuous	3 6

Design List

Choose a design by clicking on its row in the list.

Number Of Runs	Block Size	Design Type	Resolution - what is estimable
8	4	Fractional Factorial	3 - Main Effects Only
8	4	Fractional Factorial	3 - Main Effects Only
12		Plackett-Burman	3 - Main Effects Only
16		Fractional Factorial	5 - All 2-factor interactions
16	8	Fractional Factorial	4 - Some 2-factor interactions
16	4	Fractional Factorial	4 - Some 2-factor interactions
16	2	Fractional Factorial	4 - Some 2-factor interactions
32		Full Factorial	>6 - Full Resolution
32	16	Full Factorial	5+ - All 2-factor interactions
32	8	Full Factorial	5+ - All 2-factor interactions
32	4	Full Factorial	4 - Some 2-factor interactions
32	2	Full Factorial	4 - Some 2-factor interactions

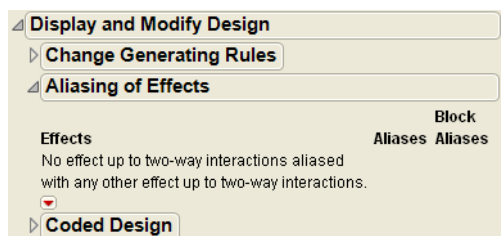
optional item

Continue

Back

5. Click **Continue**.

After you select the model from the Design list, the outlines for modifying and evaluating the model appear. In the Custom designer, you have the ability to form any model effects you want. The Screening designer creates the design effects based on the design you choose. In particular, the full factorial with all two-factor interactions has no aliasing of the included interactions, as shown in Figure 4.24.

Figure 4.24 No aliasing of the included interactions


A complete discussion of the Design Evaluation options is found in Chapter 3, “Building Custom Designs.”

To continue with this example, do the following:

6. Choose **Sort Left to Right** in the Output Options panel.
7. Click **Make Table** to see the design runs shown in Figure 4.25.

Examine the data table and note the **Pattern** variable to see the arrangement of plus and minus signs that define the runs. This table is used in the analysis sections that follow.

Figure 4.25 Listing of a 5-factor Plackett-Burman Design Table

Design	Pattern	Feed Rate	Catalyst	Stir Rate	Temperature	Concentration	Percent Reacted
1	---+	10	1	100	180	3	•
2	---+	10	1	120	140	3	•
3	--++	10	1	120	140	6	•
4	-+++	10	2	100	140	6	•
5	-+++	10	2	100	180	6	•
6	++++	10	2	120	180	3	•
7	+---+	15	1	100	140	6	•
8	+---+	15	1	100	180	3	•
9	++++	15	1	120	180	6	•
10	++++	15	2	100	140	3	•
11	++++	15	2	120	140	3	•
12	++++	15	2	120	180	6	•

Analysis of Screening Data

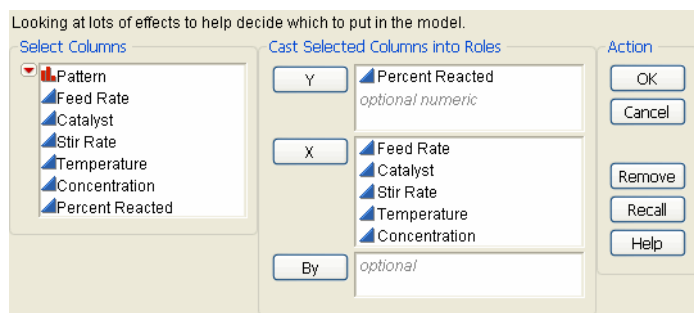
After creating and viewing the data table, you can now run analyses on the data. As an example, open the data table called **Plackett-Burman.jmp**, found in Design Experiment folder in the Sample Data installed with JMP. This table contains the design runs and the **Percent Reacted** experimental results for the 12-run Plackett-burman design created in the previous section.

Using the Screening Analysis Platform

The data table has two scripts called **Screening** and **Model**, showing in the upper-left corner of the table, that were created by the DOE Screening designer. You can use these scripts to analyze the data, however it is simple to run the analyses yourself.

1. Select **Analyze > Modeling > Screening** to see the completed launch dialog shown in Figure 4.26. When you create a DOE design table, the variable roles are saved with the data table and used by the launch platform to complete the dialog.

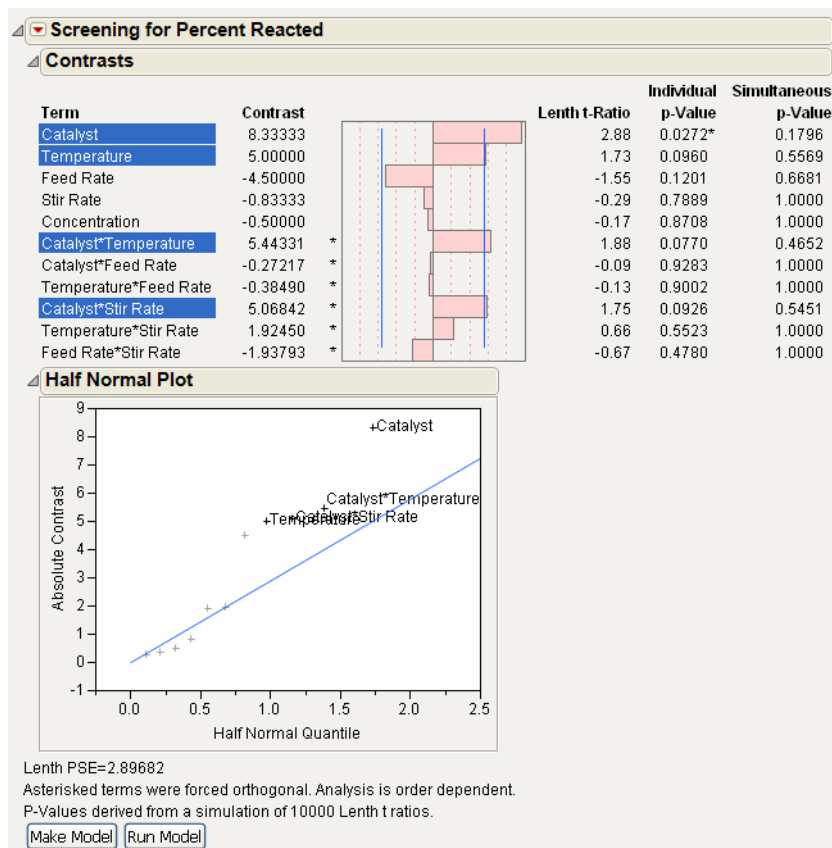
Figure 4.26 Launch Dialog for the Screening Platform



2. Click **OK** to see the Screening platform result shown in Figure 4.27.

The Contrasts section of the Screening platform results lists all possible model effects, a contrast value for each effect, Lenth t -ratios (calculated as the contrast value divided by the Lenth PSE (pseudo-standard error)), individual and simultaneous p -values, and aliases if there are any. Significant and marginally significant effects are highlighted. See the chapter on analyzing Screening designs in the *Modeling and Multivariate Methods* book for complete documentation of the Screening analysis platform.

Figure 4.27 Results of the Screening Analysis

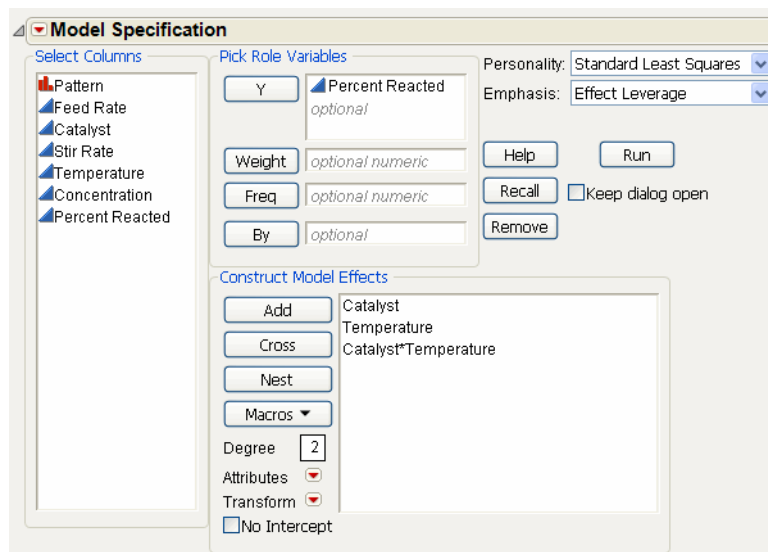


- Examine the Half Normal plot in Figure 4.27.

Using the Fit Model Platform

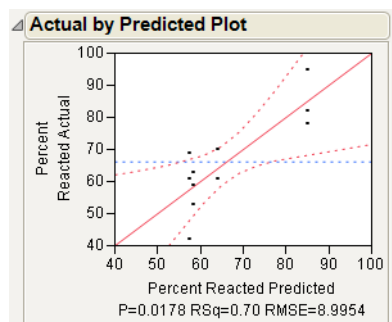
The **Make Model** button beneath the Half Normal Plot creates a Fit Model dialog that includes all the highlighted effects. However, note that the **Catalyst*Stir Rate** interaction is highlighted, but the **Stir Rate** main effect is not. Therefore, that interaction shouldn't be in the model.

- Click the **Make Model** Button beneath the Half Normal Plot.
- Highlight the **Catalyst*Stir Rate** interaction and click **Remove** on the Fit Model dialog. The dialog is shown in Figure 4.28.
- Then click **Run** to see the analysis results.

Figure 4.28 Create Fit Model Dialog and Remove Unwanted Effect

The Actual-by-Predicted Plot

The Whole Model actual-by-predicted plot, shown in Figure 4.29, appears at the top of the Fit Model report. You see at a glance that this model fits well. The blue line falls outside the bounds of the 95% confidence curves (red-dotted lines), which tells you the model is significant. The model p -value ($p = 0.0178$), R^2 , and RMSE appear below the plot. The RMSE is an estimate of the standard deviation of the process noise, assuming that the unestimated effects are negligible.

Figure 4.29 An Actual-by-Predicted Plot

The Scaled Estimates Report

To see a scaled estimates report, use **Effect Screening > Scaled Estimates** found in the red triangle menu on the **Response Percent Reacted** title bar. When there are quadratic or polynomial effects, the coefficients and the tests for them are more meaningful if effects are scaled and coded. The Scaled Estimates report includes a bar chart of the individual effects embedded in a table of parameter estimates. The last column of the table has the p -values for each effect.

Figure 4.30 Example of a Scaled Estimates Report

Scaled Estimates					
Term	Scaled Estimate		Std Error	t Ratio	Prob> t
Intercept	66.166667		2.596739	25.48	<.0001*
Catalyst(1,2)	8.3333333		2.596739	3.21	0.0124*
Temperature(140,180)	5		2.596739	1.93	0.0903
Catalyst*Temperature	5.5		2.596739	2.12	0.0670

A Power Analysis

The Fit Model report has outline nodes for the **Catalyst** and **Temperature** effects. To run a power analysis for an effect, click the red triangle icon on its title bar and select **Power Analysis**.

This example shows a power analysis for the **Catalyst** variable, using default value for α (0.05), the root mean square error and parameter estimate for **Catalyst**, for a sample size of 12. The resulting power is 0.802, which means that in similar experiments, you can expect an 89% chance of detecting a significant effect for **Catalyst**.

Figure 4.31 Example of a Power Analysis

Power Details window
 Catalyst(1,2)
 Click and Enter 1, 2 or a sequence of values for each:

	α	σ	δ	Number
From:	0.050	8.995369	8.333333	12
To:
By:

☒ Solve for Power
☐ Solve for Least Significant Number
☐ Solve for Least Significant Value
☐ Adjusted Power and Confidence Interval

Done Cancel Help

Power Details
 Test Catalyst(1,2)
Power

α	σ	δ	Number	Power
0.0500	8.995369	8.333333	12	0.8020

Refer to *Modeling and Multivariate Methods* for details.

Response Surface Designs



Response surface designs are useful for modeling a curved quadratic surface to continuous factors. A response surface model can pinpoint a minimum or maximum response, if one exists inside the factor region. Three distinct values for each factor are necessary to fit a quadratic function, so the standard two-level designs cannot fit curved surfaces.

The most popular response surface design is the central composite design, illustrated in the figure to the left below. It combines a two-level fractional factorial and two other kinds of points:

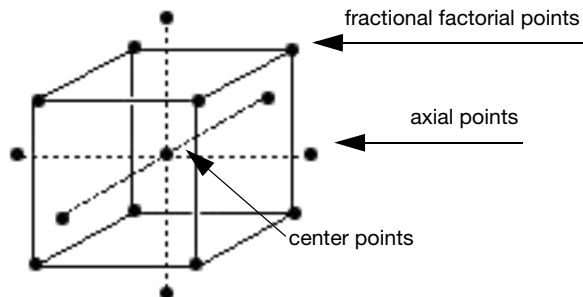
- *Center points*, for which all the factor values are at the zero (or midrange) value.
- *Axial (or star) points*, for which all but one factor are set at zero (midrange) and that one factor is set at outer (axial) values.

The Box-Behnken design, illustrated in the figure on the right below, is an alternative to central composite designs. One distinguishing feature of the Box-Behnken design is that there are only three levels per factor.

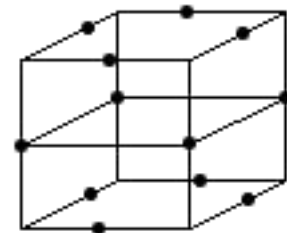
Another important difference between the two design types is that the Box-Behnken design has no points at the vertices of the cube defined by the ranges of the factors. This is sometimes useful when it is desirable to avoid these points due to engineering considerations. The price of this characteristic is the higher uncertainty of prediction near the vertices compared to the central composite design.

Figure 5.1 Response Surface Designs

Central Composite Design



Box-Behnken Design



Contents

- A Box-Behnken Design: The Tennis Ball Example129
 - The Prediction Profiler132
 - A Response Surface Plot (Contour Profiler).....134
 - Geometry of a Box-Behnken Design136
- Creating a Response Surface Design136
 - Enter Responses and Factors137
 - Choose a Design138
 - Specify Output Options139
 - View the Design Table.140

A Box-Behnken Design: The Tennis Ball Example

The `Bounce Data.jmp` sample data file has response surface data inspired by the tire tread data described in Derringer and Suich (1980). The objective of this experiment is to match a standardized target value (450) of tennis ball bounciness. The bounciness varies with amounts of Silica, Silane, and Sulfur used to manufacture the tennis balls. The experimenter wants to collect data over a wide range of values for these variables to see if a response surface can find a combination of factors that matches a specified bounce target. To follow this example:

1. Select **DOE > Response Surface Design**.
2. Load factors by clicking the red triangle icon on the Response Surface Design title bar and selecting **Load Factors**. Navigate to the Sample Data folder installed with JMP, and open `Bounce Factors.jmp` from the Design Experiment folder.
3. Load the responses by clicking the red triangle icon on the Response Surface Design title bar and selecting **Load Responses**. Navigate to the Sample Data folder, and open `Bounce Response.jmp` from the Design Experiment folder. Figure 5.2 shows the completed Response panel and Factors panel.

Figure 5.2 Response and Factors For Bounce Data

The screenshot shows the **Response Surface Design** window. The **Responses** panel has buttons for **Add Response**, **Remove**, and **Number of Responses...**. Below is a table with one row:

Response Name	Goal	Lower Limit	Upper Limit	Importance
Stretch	Match Target	350	550	1

The **Factors** panel shows a table with three factors:

Name	Role	Values
Silica	Continuous	0.7 1.7
Sulfur	Continuous	1.8 2.8
Silane	Continuous	40 60

After the response data and factors data are loaded, the Response Surface Design Choice dialog lists the designs in Figure 5.3. (Click **Continue** to see the choices on the right.)

Figure 5.3 Response Surface Design Selection

The screenshot shows the **Response Surface Design Choice** dialog. On the left, under **3 Factors** and **Choose a Design**, is a table:

Number Of Runs	Block Size	Center Points	Design Type
15		3	Box-Behnken
16		2	Central Composite Design
20		6	CCD-Uniform Precision
20	6	6	CCD-Orthogonal Blocks
23		9	CCD-Orthogonal

Below the table is an **optional item** section. On the right, under **3 Factors**, **Box-Behnken**, and **Display and Modify Design**, is the **Design Evaluation** panel. It includes **Output Options** with a **Run Order** dropdown set to **Randomize**. Below are input fields for **Make JMP Table from design plus**: **Number of Center Points** (3) and **Number of Replicates** (0). At the bottom are **Make Table** and **Back** buttons.

A Box-Behnken Design: The Tennis Ball Example

The Box-Behnken design selected for three effects generates the design table of 15 runs shown in Figure 5.4.

In real life, you would conduct the experiment and then enter the responses into the data table. Suppose you completed the experiment and the final data table is **Bounce Data.jmp**.

1. Open **Bounce Data.jmp** from the **Design Experiment** folder found in the sample data installed with JMP (Figure 5.4).

Figure 5.4 JMP Table for a Three-Factor Box-Behnken Design

Bounce Data			Pattern	Silica	Sulfur	Silane	Stretch
Design	Box-Behnken						
Model							
Columns (5/0)							
Pattern *							
Silica *							
Sulfur *							
Silane *							
Stretch *							
Rows							
All rows	15						
Selected	0						
Excluded	0						
Hidden	0						
Labelled	0						

1	-0	0.7	1.8	50	570
2	-+0	0.7	2.8	50	285
3	+0	1.7	1.8	50	260
4	++0	1.7	2.8	50	433
5	0--	1.2	1.8	40	422
6	0+-	1.2	1.8	60	351
7	0+-	1.2	2.8	40	278
8	0++	1.2	2.8	60	392
9	-0-	0.7	2.3	40	451
10	+0-	1.7	2.3	40	372
11	-0+	0.7	2.3	60	474
12	+0+	1.7	2.3	60	394
13	000	1.2	2.3	50	398
14	000	1.2	2.3	50	394
15	000	1.2	2.3	50	396

After opening the **Bounce Data.jmp** data table, run a fit model analysis on the data. The data table contains a script labeled **Model**, showing in the upper left panel of the table.

2. Click the red triangle next to **Model** and select **Run Script** to start a fit model analysis.
3. When the Fit Model dialog appears, click **Run**.

The standard Fit Model analysis results appear in tables shown in Figure 5.5, with parameter estimates for all response surface and crossed effects in the model.

The prediction model is highly significant with no evidence of lack of fit. All main effect terms are significant as well as the two interaction effects involving Sulfur.

Figure 5.5 JMP Statistical Reports for a Response Surface Analysis of Bounce Data

▼ Response Stretch

Summary of Fit

RSquare	0.999777
RSquare Adj	0.999375
Root Mean Square Error	1.987461
Mean of Response	391.3333
Observations (or Sum Wgts)	15

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	9	88453.583	9828.18	2488.146
Error	5	19.750	3.95	Prob > F
C. Total	14	88473.333		<.0001*

Lack Of Fit

Source	DF	Sum of Squares	Mean Square	F Ratio
Lack Of Fit	3	11.750000	3.91667	0.9792
Pure Error	2	8.000000	4.00000	0.5411
Total Error	5	19.750000		Max RSq
				0.9999

Parameter Estimates

Effect Tests

Source	Nparm	DF	Sum of Squares	F Ratio	Prob > F
Silica(0.7,1.7)	1	1	12880.125	3260.791	<.0001*
Silane(40,60)	1	1	968.000	245.0633	<.0001*
Sulfur(1.8,2.8)	1	1	5778.125	1462.816	<.0001*
Silane*Silica	1	1	0.250	0.0633	0.8114
Sulfur*Silica	1	1	52441.000	13276.20	<.0001*
Sulfur*Silane	1	1	8556.250	2166.139	<.0001*
Silica*Silica	1	1	2592.923	656.4362	<.0001*
Silane*Silane	1	1	0.231	0.0584	0.8186
Sulfur*Sulfur	1	1	4653.231	1178.033	<.0001*

See *Modeling and Multivariate Methods* for more information about interpretation of the tables in Figure 5.5.

Note: DOE response surface designs are available for up to eight factors only. In the DOE Response Surface Design platform, an error message is generated if more than eight factors are specified with a response surface design. Response surface designs with more than eight factors can be generated using **DOE > Custom Design**, where either a D-optimal or an I-optimal design can be specified. See “[Creating Response Surface Experiments](#),” p. 34 of *JMP Design of Experiments* for how to use the custom designer to create response surface designs. Curvature analysis is not shown (no error or warning message is given) for response surface designs of more than 20 factors when using the custom designer or the Fit Model platform; all other analyses are valid and are shown.

The Response Surface report also has the tables shown in Figure 5.6.

Figure 5.6 Statistical Reports for a Response Surface Analysis

Response Surface				
Coef	Silica(0.7,1.7)	Silane(40,60)	Sulfur(1.8,2.8)	Stretch
Silica(0.7,1.7)	26.5	-0.25	114.5	-40.125
Silane(40,60)	.	0.25	46.25	11
Sulfur(1.8,2.8)	.	.	-35.5	-26.875

Solution	
Variable	Critical Value
Silica(0.7,1.7)	1.7912411
Silane(40,60)	23.424426
Sulfur(1.8,2.8)	2.1986422
Solution is a SaddlePoint	
Critical values outside data range	
Predicted Value at Solution 360.38388	

Canonical Curvature				
Eigenvalues and Eigenvectors				
Eigenvalue	62.9095	3.2989	-74.9584	
Silica(0.7,1.7)	0.82779	-0.29879	-0.47486	
Silane(40,60)	0.19280	0.94634	-0.25937	
Sulfur(1.8,2.8)	0.52687	0.12315	0.84097	

The Response Surface report shows a summary of the parameter estimates.

The Solution report lists the critical values of the surface factors and tells the kind of solution (maximum, minimum, or saddle point). The solution for this example is a saddle point. The table also warns that the critical values given by the solution are outside the range of data values.

The Canonical Curvature report shows eigenvalues and eigenvectors of the effects. The eigenvector values show that the dominant negative curvature (yielding a maximum) is mostly in the Sulfur direction. The dominant positive curvature (yielding a minimum) is mostly in the Silica direction. This is confirmed by the prediction profiler in Figure 5.8.

See *Modeling and Multivariate Methods* for details about the response surface analysis tables in Figure 5.6.

The Prediction Profiler

Next, use the response Prediction Profiler to get a closer look at the response surface and help find the settings that produce the best response target. The Prediction Profiler is a way to interactively change variables and look at the effects on the predicted response.

1. If the Prediction Profiler is not already open, click the red triangle on the Response Stretch title bar and select **Factor Profiling > Profiler**.

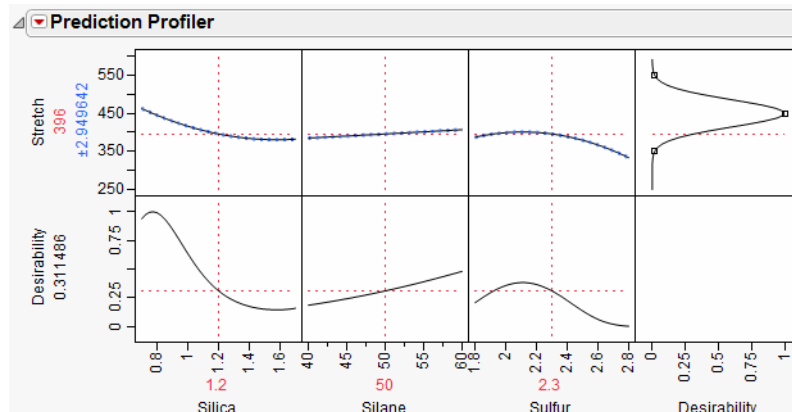
The first three plots in the top row of plots in the Prediction Profiler (Figure 5.7) display *prediction traces* for each x variable. A prediction trace is the predicted response as one variable is changed while the others are held constant at the current values (Jones 1991).

The current predicted value of Stretch, 396, is based on the default factor setting. It is represented by the horizontal dotted line that shows slightly below the desirability function target value (Figure 5.7). The

profiler shows desirability settings for the factors Silica, Silane, and Sulfur that give a value for Stretch of 396, which is quite different from the specified target of 450.

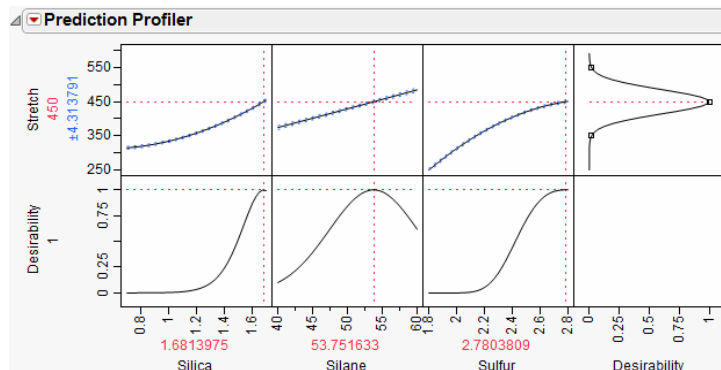
The bottom row has a plot for each factor, showing its *desirability trace*. The profiler also contains a Desirability column, which graphs desirability on a scale from 0 to 1 and has an adjustable desirability function for each y variable. The overall desirability measure is on the left of the desirability traces.

Figure 5.7 The Prediction Profiler



- To adjust the prediction traces of the factors and find a **Stretch** value that is closer to the target, click the red triangle on the Prediction Profiler title bar and select **Maximize Desirability**. This command adjusts the profile traces to produce the response value closest to the specified target (the target given by the desirability function). The range of acceptable values is determined by the positions of the upper and lower handles.

Figure 5.8 shows the result of the most desirable settings. Changing the settings of Silica from 1.2 to 1.7, Silane from 50 to 54, and Sulfur from 2.3 to 2.8 raised the predicted response from 396 to the target value of 450. Finding maximum desirability is an iterative process so your results may differ slightly from those shown below.

Figure 5.8 Prediction Profiler with Maximum Desirability Set for a Response Surface Analysis

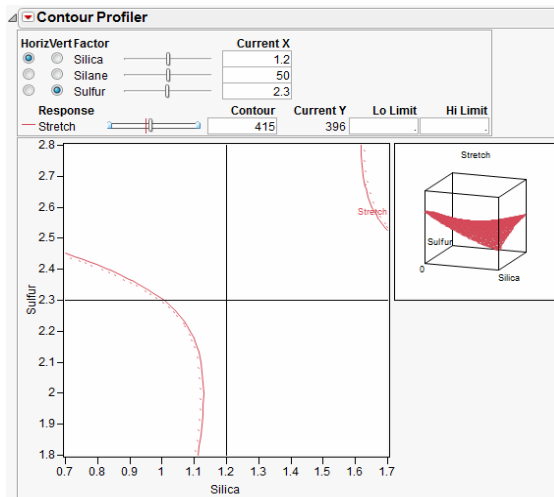
See *Modeling and Multivariate Methods* for further discussion of the Prediction Profiler.

A Response Surface Plot (Contour Profiler)

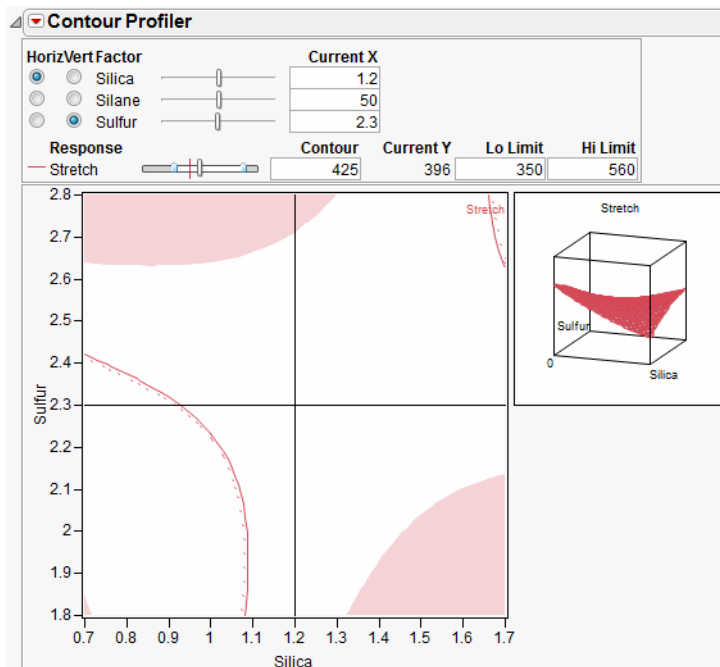
Another way to look at the response surface is to use the Contour Profiler. Click the red triangle on the Response Stretch title bar and select **Factor Profiling > Contour Profiler** to display the interactive contour profiler, as shown in Figure 5.9.

The contour profiler is useful for viewing response surfaces graphically, especially when there are multiple responses. This example shows the profile to Silica and Sulfur for a fixed value of Silane.

Options on the Contour Profiler title bar can be used to set the grid density, request a surface plot (mesh plot), and add contours at specified intervals, like those shown in the contour plot in Figure 5.9. The sliders for each factor set values for Current X and Current Y.

Figure 5.9 Contour Profiler for a Response Surface Analysis

Enter the Lo limit and Hi limit values to shade the unacceptable regions in the contour plot

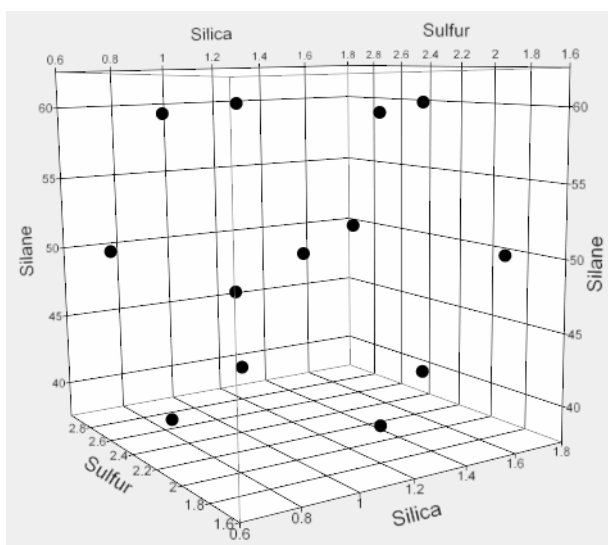
Figure 5.10 Contour Profiler with High and Low Limits

The Prediction Profiler and the Contour Profiler are discussed in more detail in *Modeling and Multivariate Methods*.

Geometry of a Box-Behnken Design

The geometric structure of a design with three effects is seen by using the Scatterplot 3D platform. The plot shown in Figure 5.11 illustrates the three Box-Behnken design columns. You can clearly see the center points and the 12 points midway between the vertices. For details on how to use the Scatterplot 3D platform, see *Basic Analysis and Graphing*.

Figure 5.11 Scatterplot 3D Rendition of a Box-Behnken Design for Three Effects



Creating a Response Surface Design

Response Surface Methodology (RSM) is an experimental technique invented to find the optimal response within specified ranges of the factors. These designs are capable of fitting a second-order prediction equation for the response. The quadratic terms in these equations model the curvature in the true response function. If a maximum or minimum exists inside the factor region, RSM can estimate it. In industrial applications, RSM designs usually involve a small number of factors. This is because the required number of runs increases dramatically with the number of factors. Using the response surface designer, you choose to use well-known RSM designs for two to eight continuous factors. Some of these designs also allow blocking.

Response surface designs are useful for modeling and analyzing curved surfaces.

To start a response surface design, select **DOE > Response Surface Design**, or click the **Response Surface Design** button on the JMP Starter DOE page. Then, follow the steps described in the following sections.

- “Enter Responses and Factors,” p. 137
- “Choose a Design,” p. 138
- “Specify Axial Value (Central Composite Designs Only),” p. 138
- “Specify Output Options,” p. 139
- “View the Design Table,” p. 140

Enter Responses and Factors

The steps for entering responses are the same in **Screening Design**, **Space Filling Design**, **Mixture Design**, **Response Surface Design**, **Custom Design**, and **Full Factorial Design**. These steps are outlined in “Enter Responses and Factors into the Custom Designer,” p. 65

Factors in a response surface design can only be continuous. The Factors panel for a response surface design appears with two default continuous factors. To enter more factors, type the number you want in the Factors panel edit box and click **Add**, as shown in Figure 5.12.

Figure 5.12 Enter Factors into a Response Surface Design

Response Surface Design

Responses

Add Response Remove Number of Responses...

Response Name	Goal	Lower Limit	Upper Limit	Importance
Y <i>optional item</i>	Maximize	.	.	.

Factors

Add 1 Continuous

Remove Selected

Name	Role	Values
X1	Continuous	-1 1
X2	Continuous	-1 1
X3	Continuous	-1 1

Specify Factors

Specify desired number of factors. Double click on a factor name or setting to edit it.

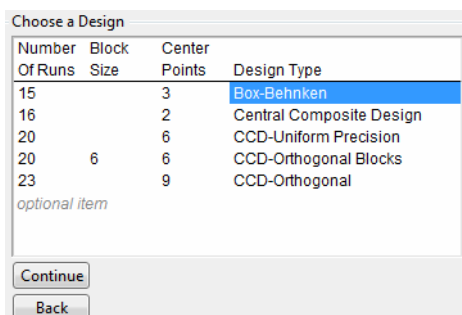
Continue

Click **Continue** to proceed to the next step.

Choose a Design

Highlight the type of response surface design you want and click **Continue**. The next sections describe the types of response surface designs shown in Figure 5.13.

Figure 5.13 Choose a Design Type



Box-Behnken Designs

The Box-Behnken design has only three levels per factor and has no points at the vertices of the cube defined by the ranges of the factors. This is sometimes useful when it is desirable to avoid extreme points due to engineering considerations. The price of this characteristic is the higher uncertainty of prediction near the vertices compared to the central composite design.

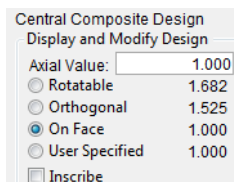
Central Composite Designs

The response surface design list contains two types of central composite designs: *uniform precision* and *orthogonal*. These properties of central composite designs relate to the number of center points in the design and to the axial values:

- Uniform precision means that the number of center points is chosen so that the prediction variance near the center of the design space is very flat.
- For orthogonal designs, the number of center points is chosen so that the second order parameter estimates are minimally correlated with the other parameter estimates.

Specify Axial Value (Central Composite Designs Only)

When you select a central composite (CCD-Uniform Precision) design and then click **Continue**, you see the panel in Figure 5.14. It supplies default axial scaling information. Entering 1.0 in the text box instructs JMP to place the axial value on the face of the cube defined by the factors, which controls how far out the axial points are. You have the flexibility to enter the values you want to use.

Figure 5.14 Display and Modify the Central Composite Design

Rotatable makes the variance of prediction depend only on the scaled distance from the center of the design. This causes the axial points to be more extreme than the range of the factor. If this factor range cannot be practically achieved, it is recommended that you choose **On Face** or specify your own value.

Orthogonal makes the effects orthogonal in the analysis. This causes the axial points to be more extreme than the -1 or 1 representing the range of the factor. If this factor range cannot be practically achieved, it is recommended that you choose **On Face** or specify your own value.

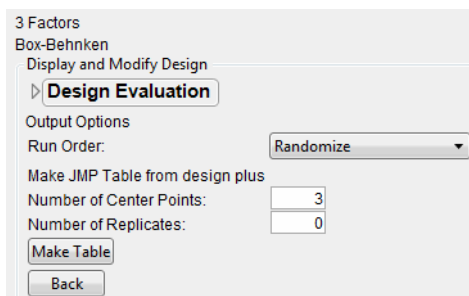
On Face leaves the axial points at the end of the -1 and 1 ranges.

User Specified uses the value you enter in the Axial Value text box.

If you want to inscribe the design, click the box beside **Inscribe**. When checked, JMP rescales the whole design so that the axial points are at the low and high ends of the range (the axials are -1 and 1 and the factorials are shrunk based on that scaling).

Specify Output Options

Use the Output Options panel to specify how you want the output data table to appear. When the options are specified the way you want them, click **Make Table**. Note that the example shown in Figure 5.15 is for a Box-Behnken design. The Box-Behnken design from the design list and the Output Options request 3 center points for a single replicate.

Figure 5.15 Select the Output Options

Run Order provides a menu with options for designating the order you want the runs to appear in the data table when it is created. Menu choices are:

Keep the Same the rows (runs) in the output table will appear in the standard order.

Sort Left to Right the rows (runs) in the output table will appear sorted from left to right.

Randomize the rows (runs) in the output table will appear in a random order.

Sort Right to Left the rows (runs) in the output table will appear sorted from right to left.

Randomize within Blocks the rows (runs) in the output table will appear in random order within the blocks you set up.

Add additional points with options given by **Make JMP Table from design** plus:

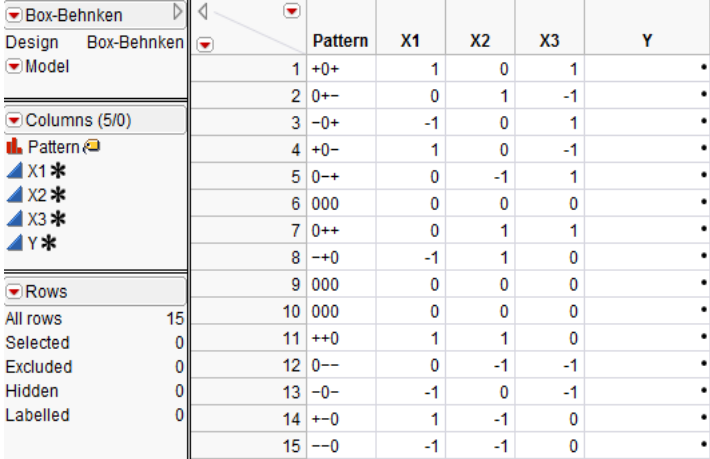
Number of Center Points Specifies additional runs placed at the center points.

Number of Replicates Specify the number of times to replicate the entire design, including centerpoints. Type the number of times you want to replicate the design in the associated text box. One replicate doubles the number of runs.

View the Design Table

Now you have a data table that outlines your experiment, as described in Figure 5.16.

Figure 5.16 The Design Data Table



	Pattern	X1	X2	X3	Y
1	+0+	1	0	1	*
2	0+-	0	1	-1	*
3	-0+	-1	0	1	*
4	+0-	1	0	-1	*
5	0-+	0	-1	1	*
6	000	0	0	0	*
7	0++	0	1	1	*
8	--0	-1	1	0	*
9	000	0	0	0	*
10	000	0	0	0	*
11	++0	1	1	0	*
12	0--	0	-1	-1	*
13	-0-	-1	0	-1	*
14	+--0	1	-1	0	*
15	--0	-1	-1	0	*

The name of the table is the design type that generated it.

Run the Model script to fit a model using the values in the design table.

The column called **Pattern** identifies the coding of the factors. It shows all the codings with “+” for high, “-” for low factor, “a” and “A” for low and high axial values, and “0” for midrange. **Pattern** is suitable to use as a label variable in plots because when you hover over a point in a plot of the factors, the pattern value

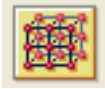
shows the factor coding of the point. The three rows whose values in the Pattern column are 000 are three center points.

The runs in the Pattern column are in the order you selected from the **Run Order** menu.

The Y column is for recording experimental results.

Chapter 6

Full Factorial Designs



A full factorial design contains all possible combinations of a set of factors. This is the most fool proof design approach, but it is also the most costly in experimental resources. The full factorial designer supports both continuous factors and categorical factors with up to nine levels.

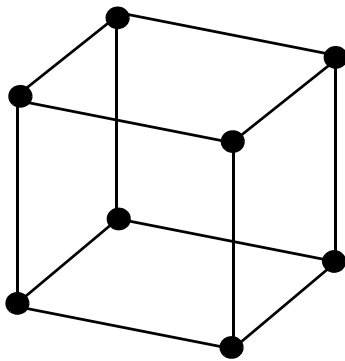
In full factorial designs, you perform an experimental run at every combination of the factor levels. The sample size is the product of the numbers of levels of the factors. For example, a factorial experiment with a two-level factor, a three-level factor, and a four-level factor has $2 \times 3 \times 4 = 24$ runs.

Factorial designs with only two-level factors have a sample size that is a power of two (specifically 2^f where f is the number of factors). When there are three factors, the factorial design points are at the vertices of a cube as shown in the diagram below. For more factors, the design points are the vertices of a hypercube.

Full factorial designs are the most conservative of all design types. There is little scope for ambiguity when you are willing to try all combinations of the factor settings.

Unfortunately, the sample size grows exponentially in the number of factors, so full factorial designs are too expensive to run for most practical purposes.

Figure 6.1 Full Factorial Design



Contents

- The Five-Factor Reactor Example 145
 - Analyze the Reactor Data146
- Creating a Factorial Design 151
 - Enter Responses and Factors 151
 - Select Output Options 152
 - Make the Table 152

The Five-Factor Reactor Example

The following example, adapted from Meyer *et al.* (1996) and Box, Hunter, and Hunter (1978), shows a five-factor reactor example.

Previously, the screening designer was used to investigate the effects of five factors on the percent reaction of a chemical process (see “[Screening Designs](#),” p. 101). The factors (Feed Rate, Catalyst, Stir Rate, Temperature, and Concentration) are all two-level continuous factors. The next example studies the same system using a full factorial design.

1. Select **DOE > Full Factorial Design**.
2. Click the red triangle icon on the Full Factorial Design title bar and select **Load Responses**.
3. In the Sample Data folder (installed with JMP), open **Reactor Response.jmp** found in the Design Experiment folder.
4. Click the red triangle icon on the Full Factorial Design title bar and select **Load Factors**.
5. In the Sample Data folder (installed with JMP), open **Reactor Factors.jmp** found in the Design Experiment folder.

The completed dialog should look like the one shown in Figure 6.2.

Figure 6.2 Full Factorial Example Response and Factors Panels

Full Factorial Design

Responses

Add Response Remove Number of Responses...

Response Name	Goal	Lower Limit	Upper Limit	Importance
Percent Reacted	Maximize	90	99	1

Factors

Continuous Categorical Remove

Name	Role	Values
Feed Rate	Continuous	10 15
Catalyst	Continuous	1 2
Stir Rate	Continuous	100 120
Temperature	Continuous	140 180
Concentration	Continuous	3 6

2x2x2x2x2 Factorial

Output Options

Run Order: Sort Left to Right

Number of Runs: 32

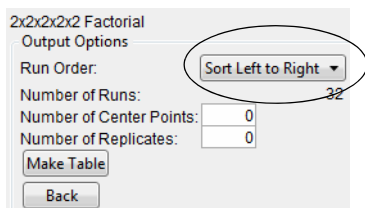
Number of Center Points: 0

Number of Replicates: 0

Make Table

Back

6. Click **Continue** to see the Output Options panel. In the Output Options panel, select **Sort Left to Right** from the Run Order menu, as shown to the right. This command defines the order of runs as they will be in the final JMP design table.



7. Click **Make Table**.

The design data table (Figure 6.3) contains a run for every combination of high and low values for the five variables, which covers all combinations of five factors with two levels each. Since there are five variables, there are $2^5=32$ runs. Initially, the table has an empty Y column named **Percent Reacted** for entering response values when the experiment is complete.

To see the completed experiment and continue this example, open **Reactor 32 Runs.jmp** found in the **Design Experiment Sample Data** folder.

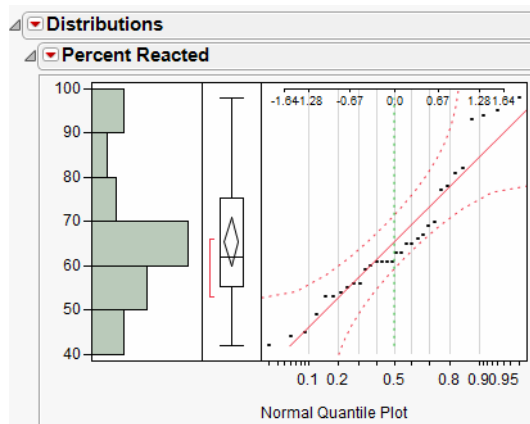
Figure 6.3 Partial Listing of Reactor 32 Runs.jmp from the Sample Data Folder

Reactor 32 Runs		Pattern	Feed Rate	Catalyst	Stir Rate	Temperature	Concentration	Percent Reacted
Design 2x2x2x2x2 Factorial								
Fit Model		1 ----	10	1	100	140	3	61
Columns (7/0)		2 ----+	10	1	100	140	6	56
Pattern		3 ---+-	10	1	100	180	3	69
Feed Rate *		4 ---++	10	1	100	180	6	44
Catalyst *		5 -+---	10	1	120	140	3	53
Stir Rate *		6 -++--	10	1	120	140	6	59
Temperature *		7 -+-++	10	1	120	180	3	66
Concentration *		8 -+++	10	1	120	180	6	49
Percent Reacted *		9 -+---	10	2	100	140	3	63
Rows		10 -++--	10	2	100	140	6	70
All rows 32		11 -+-+-	10	2	100	180	3	94
Selected 0		12 -++--	10	2	100	180	6	78
Excluded 0		13 -+-+-	10	2	120	140	3	54
Hidden 0		14 -++++	10	2	120	140	6	67
Labelled 0		15 -++--	10	2	120	180	3	95

Analyze the Reactor Data

Begin the analysis with a quick look at the response data before fitting the factorial model.

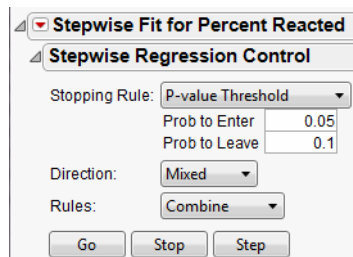
1. Select **Analyze > Distribution**.
2. Highlight **Percent Reacted** and click **Y, Columns**. Then click **OK**.
3. Click the red triangle icon on the **Percent Reacted** title bar and select **Normal Quantile Plot**. The results are shown in Figure 6.4.

Figure 6.4 Distribution of Response Variable for Reactor Data

This initial analysis shows some experimental runs with a very high percent reacted response.

Start the formal analysis with a stepwise regression. The data table has a script stored with it that automatically defines an analysis of the model with main effects and all two-factor interactions.

4. Click the red triangle icon next to the Fit Model script and select **Run Script**. The stepwise analysis begins with the Stepwise Regression Control panel shown in Figure 6.5.
5. Select **P-value Threshold** from the Stopping Rule list.
6. The probability to enter a factor (Prob to Enter) in the model should be 0.05.
7. The probability to remove a factor (Prob to Leave) should be 0.1.
8. A useful way to use the Stepwise platform is to check all the main effects in the Current Estimates table. However, make sure that the menu beside Direction in the Stepwise Regression Control panel specifies **Mixed** (see Figure 6.5).

Figure 6.5 Stepwise Control Panel

9. Check the boxes for the main effects of the factors as shown in Figure 6.6.

Figure 6.6 Starting Model For Stepwise Process

Current Estimates						
Lock	Entered	Parameter	Estimate	nDF	SS	"F Ratio" "Prob>F"
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	Intercept	65.5	1	0	0.000 1
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	Feed Rate(10,15)	-0.6875	1	15.125	0.149 0.70282
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	Catalyst(1,2)	9.75	1	3042	29.928 0.00001
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	Stir Rate(100,120)	-0.3125	1	3.125	0.031 0.86217
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	Temperature(140,180)	5.375	1	924.5	9.095 0.00566
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	Concentration(3,6)	-3.125	1	312.5	3.074 0.09132
<input type="checkbox"/>	<input type="checkbox"/>	Feed Rate*Catalyst	0	1	15.125	0.144 0.70763
<input type="checkbox"/>	<input type="checkbox"/>	Feed Rate*Stir Rate	0	1	4.5	0.043 0.83807
<input type="checkbox"/>	<input type="checkbox"/>	Feed Rate*Temperature	0	1	6.125	0.058 0.81153
<input type="checkbox"/>	<input type="checkbox"/>	Feed Rate*Concentration	0	1	0.125	0.001 0.97284
<input type="checkbox"/>	<input type="checkbox"/>	Catalyst*Stir Rate	0	1	6.125	0.058 0.81153
<input type="checkbox"/>	<input type="checkbox"/>	Catalyst*Temperature	0	1	1404.5	28.357 1.61e-5
<input type="checkbox"/>	<input type="checkbox"/>	Catalyst*Concentration	0	1	32	0.306 0.5848
<input type="checkbox"/>	<input type="checkbox"/>	Temperature*Concentration	0	1	968	14.450 0.00082

10. Click **Go**.

The mixed stepwise procedure removes insignificant main effects and adds important interactions. The end result is shown in Figure 6.7. Note that the **Feed Rate** and **Stir Rate** factors are no longer in the model.

Figure 6.7 Model After Mixed Stepwise Regression

Current Estimates						
Lock	Entered	Parameter	Estimate	nDF	SS	"F Ratio" "Prob>F"
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	Intercept	65.5	1	0	0.000 1
<input type="checkbox"/>	<input type="checkbox"/>	Feed Rate(10,15)	0	1	15.125	1.383 0.25064
<input type="checkbox"/>	<input checked="" type="checkbox"/>	Catalyst(1,2)	9.75	2	4446.5	200.362 1.6e-16
<input type="checkbox"/>	<input type="checkbox"/>	Stir Rate(100,120)	0	1	3.125	0.274 0.60543
<input type="checkbox"/>	<input checked="" type="checkbox"/>	Temperature(140,180)	5.375	3	3297	99.043 2.4e-14
<input type="checkbox"/>	<input checked="" type="checkbox"/>	Concentration(3,6)	-3.125	2	1280.5	57.700 2.7e-10
<input type="checkbox"/>	<input type="checkbox"/>	Feed Rate*Catalyst	0	2	30.25	1.406 0.26469
<input type="checkbox"/>	<input type="checkbox"/>	Feed Rate*Stir Rate	0	3	22.75	0.656 0.58722
<input type="checkbox"/>	<input type="checkbox"/>	Feed Rate*Temperature	0	2	21.25	0.954 0.39927
<input type="checkbox"/>	<input type="checkbox"/>	Feed Rate*Concentration	0	2	15.25	0.670 0.52116
<input type="checkbox"/>	<input type="checkbox"/>	Catalyst*Stir Rate	0	2	9.25	0.397 0.67635
<input type="checkbox"/>	<input checked="" type="checkbox"/>	Catalyst*Temperature	6.625	1	1404.5	126.575 1.7e-11
<input type="checkbox"/>	<input type="checkbox"/>	Catalyst*Concentration	0	1	32	3.119 0.0896
<input type="checkbox"/>	<input checked="" type="checkbox"/>	Temperature*Concentration	-5.5	1	968	87.237 8.6e-10

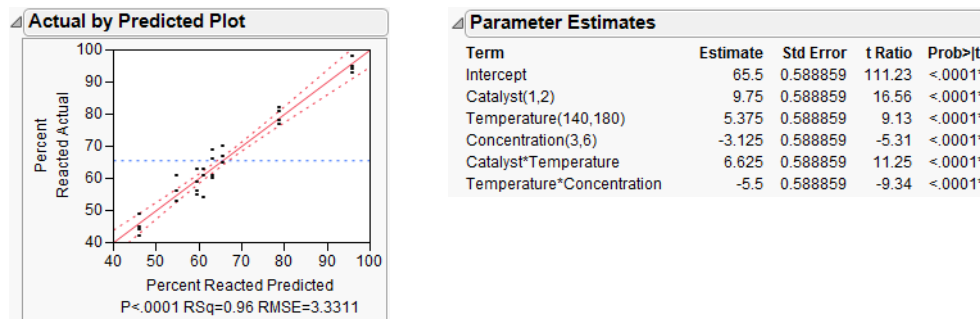
11. Click the **Make Model** button in the Stepwise Regression Control panel. The Model Specification window that appears is automatically set up with the appropriate effects (Figure 6.8).

Figure 6.8 Fitting a Prediction Model

12. Click **Run** to see the analysis for a candidate prediction model (Figure 6.9).

The figure on the left in Figure 6.9 shows the actual by predicted plot for the model. The predicted model covers a range of predictions from 40% to 95% reacted. The size of the random noise as measured by the RMSE is only 3.3311%, which is more than an order of magnitude smaller than the range of predictions. This is strong evidence that the model has good predictive capability.

The figure on the right in Figure 6.9 shows a table of model coefficients and their standard errors (labeled Parameter Estimates). All effects selected by the stepwise process are highly significant.

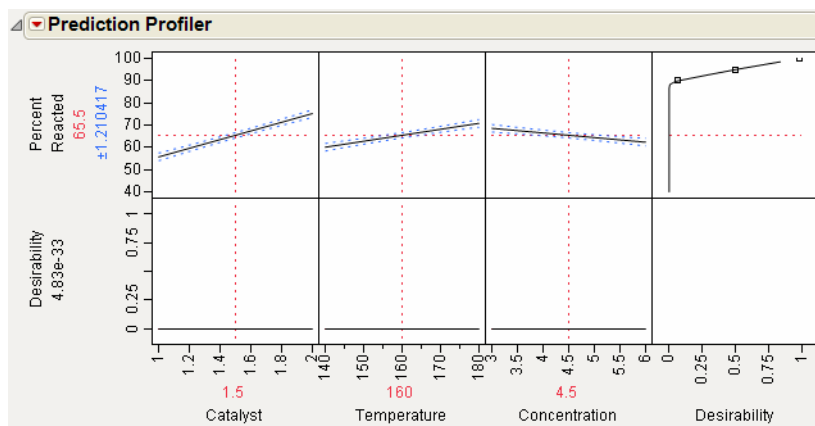
Figure 6.9 Actual by Predicted Plot and Prediction Parameter Estimates Table

The factor Prediction Profiler also gives you a way to compare the factors and find optimal settings.

1. To open the Prediction Profiler, click the red triangle on the Response Percent Reacted title bar and select **Factor Profiling > Profiler**.

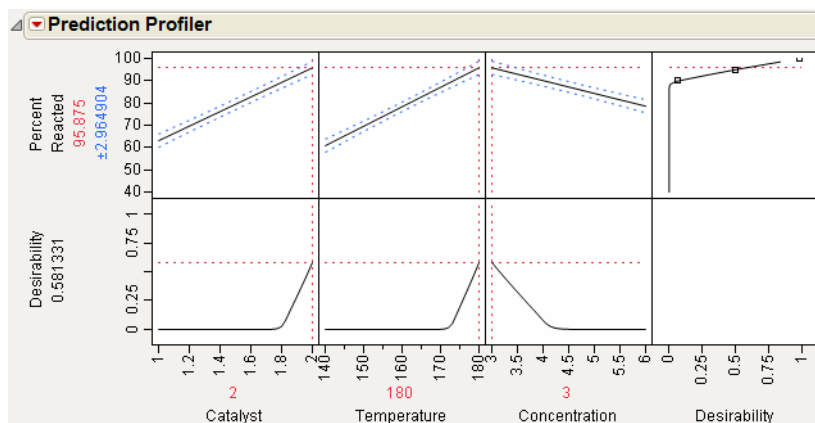
Figure 6.10 shows the profiler's initial display. The Prediction Profiler is discussed in more detail in the chapter "[Response Surface Designs](#)," p. 127, and in *Modeling and Multivariate Methods*.

Figure 6.10 Viewing the Profiler



2. Click the red triangle on the Prediction Profiler title bar and select **Maximize Desirability** to see the profiler in Figure 6.11.

Figure 6.11 Viewing the Prediction Profiles at the Optimum Settings



The goal is to maximize Percent Reacted. The reaction is unfeasible economically unless the Percent Reacted is above 90%. Percent Reacted increases from 65.5 at the center of the factor ranges to a predicted maximum of 95.875 ± 2.96 at the most desirable settings. The best settings of all three factors are

at the ends of their ranges. Future experiments could investigate what happens as you continue moving further in this direction.

Creating a Factorial Design

To start a full factorial design, select **DOE > Full Factorial Design**, or click the **Full Factorial Design** button on the JMP Starter **DOE** page. Then, follow the steps below:

- “Enter Responses and Factors,” p. 151
- “Select Output Options,” p. 152
- “Make the Table,” p. 152

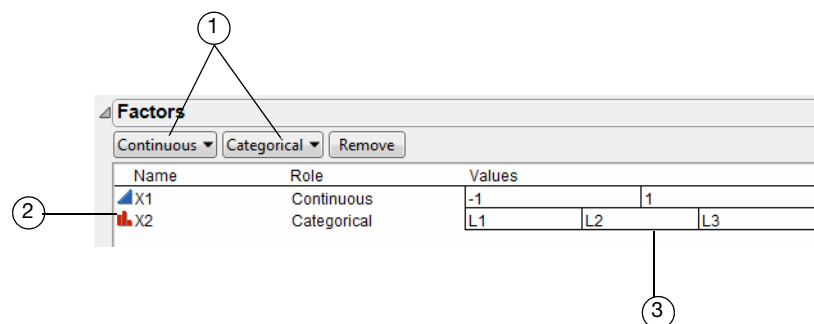
Enter Responses and Factors

The steps for entering responses are outlined in “Enter Responses and Factors into the Custom Designer,” p. 65

The steps for entering factors in a full factorial design are unique to this design. To add factors, see Figure 6.12.

1. To enter factors, click either the **Continuous** button or the **Categorical** button and select a factor type, level 2 - 9.
2. Double-click to edit the factor name.
3. Click to enter values or change the level names.

Figure 6.12 Entering Factors in a Full Factorial Design

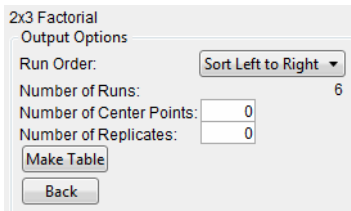


When you finish adding factors, click **Continue**.

Select Output Options

Use the Output Options panel to specify how you want the output data table to appear, as illustrated in Figure 6.13:

Figure 6.13 Output Options Panel



Run Order gives options to designate the order you want the runs to appear in the data table when it is created. Choices are:

- Keep the Same** the rows (runs) in the output table will appear in the standard order.
- Sort Left to Right** the rows (runs) in the output table will appear sorted from left to right.
- Randomize** the rows (runs) in the output table will appear in a random order.
- Sort Right to Left** the rows (runs) in the output table will appear sorted from right to left.

Add additional points to the data table with these options:

- Number of Center Points** Specifies additional runs placed at the center of each continuous factor's range.
- Number of Replicates** Specify the number of times to replicate the entire design, including centerpoints. Type the number of times you want to replicate the design in the associated text box. One replicate doubles the number of runs.

Make the Table

When you click **Make Table**, the table shown in Figure 6.14 appears.

Figure 6.14 Factorial Design Table

2x3 Factorial				
Design 2x3 Factorial				
Screening				
Model				
Columns (4/0)				
Pattern				
X1 *				
X2 *				
Y *				

Pattern	X1	X2	Y
1 -1	-1	L1	•
2 -2	-1	L2	•
3 -3	-1	L3	•
4 +1	1	L1	•
5 +2	1	L2	•
6 +3	1	L3	•

The name of the table is the design type that generated it.

Run the Model script to fit a model using values in the design table.

Values in the **Pattern** column describe the run each row represents.

- For continuous factors, a plus sign represents high levels.
- For continuous factors, a minus sign represents low levels.
- Level numbers represent values of categorical factors.

Chapter 7

Mixture Designs



The mixture designer supports experiments with factors that are ingredients in a mixture. You can choose among several classical mixture design approaches, such as simplex, extreme vertices, and lattice. For the extreme vertices approach you can supply a set of linear inequality constraints limiting the geometry of the mixture factor space.

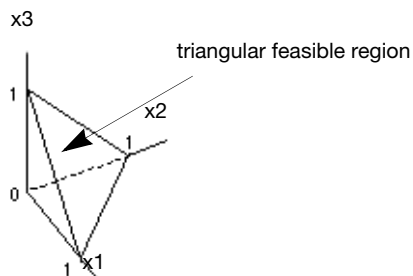
The properties of a mixture are almost always a function of the relative proportions of the ingredients rather than their absolute amounts. In experiments with mixtures, a factor's value is its proportion in the mixture, which falls between zero and one. The sum of the proportions in any mixture recipe is one (100%).

Designs for mixture experiments are fundamentally different from those for screening. Screening experiments are orthogonal. That is, over the course of an experiment, the setting of one factor varies independently of any other factor. Thus, the interpretation of screening experiments is relatively simple, because the effects of the factors on the response are separable.

With mixtures, it is impossible to vary one factor independently of all the others. When you change the proportion of one ingredient, the proportion of one or more other ingredients must also change to compensate. This simple fact has a profound effect on every aspect of experimentation with mixtures: the factor space, the design properties, and the interpretation of the results.

Because the proportions sum to one, mixture designs have an interesting geometry. The feasible region for the response in a mixture design takes the form of a simplex. For example, consider three factors in a 3-D graph. The plane where the sum of the three factors sum to one is a triangle-shaped slice. You can rotate the plane to see the triangle face-on and see the points in the form of a ternary plot.

Figure 7.1 Mixture Design



Contents

Mixture Design Types	157
The Optimal Mixture Design	157
The Simplex Centroid Design	158
Creating the Design.	159
Simplex Centroid Design Examples.	160
The Simplex Lattice Design.	161
The Extreme Vertices Design.	163
Creating the Design.	164
An Extreme Vertices Example with Range Constraints	164
An Extreme Vertices Example with Linear Constraints	166
Extreme Vertices Method: How It Works	167
The ABCD Design	168
Creating Ternary Plots.	168
Fitting Mixture Designs.	170
Whole Model Tests and Analysis of Variance Reports	171
Understanding Response Surface Reports	171
A Chemical Mixture Example	171
Create the Design	171
Analyze the Mixture Model	174
The Prediction Profiler	175
The Mixture Profiler	176
A Ternary Plot of the Mixture Response Surface	177

Mixture Design Types

To create a mixture design, choose **DOE > Mixture Design**, enter the responses and factors into the initial mixture designer panel, and click **Continue**. You then see the Choose Mixture Design Type panel shown in Figure 7.2. Select one of the designs from the panel:

Optimal invokes the custom designer with all the mixture variables already defined.

Simplex Centroid lets you specify the degree to which the factor combinations are made.

Simplex Lattice lets you specify how many levels you want on each edge of the grid.

Extreme Vertices lets you specify linear constraints or restrict the upper and lower bounds to be within the 0 to 1 range.

ABCD Design generates a screening design for mixtures devised by Snee (1975).

Figure 7.2 Mixture Design Selection Dialog

Choose Mixture Design Type

Optimal Create a design tailored to meet specific requirements.

Simplex Centroid Run each ingredient without mixing, then mix equal proportions of K ingredients at a time to the specified limit. **K** 2

Simplex Lattice Triangular grid. **Number of Levels** Specify number of levels per factor: 5

Extreme Vertices Find the vertices of the simplex. Then add the mid-points of the edges and averages of vertices to the specified degree. **Degree** 2

Linear Constraint Add linear constraints on the relative proportions of ingredients. Click once for each constraint.

ABCD Design A mixture design for factor screening.

Back

After you select the design type, choose the number of runs in the Design Generation panel and click **Make Design**.

The following sections describe each mixture design type and show examples.

The Optimal Mixture Design

The **Optimal** mixture design choice invokes the custom designer with the mixture variables entered into the response and factors panels. To create an optimal mixture design:

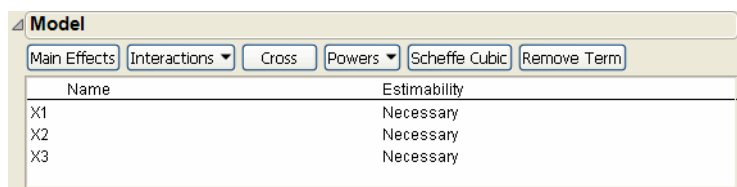
1. Select **DOE > Mixture Design**.
2. Enter factors and responses. The steps for entering responses are outlined in [“Enter Responses and Factors into the Custom Designer,” p. 65.](#)
3. After you enter responses and factors, click **Continue**.

4. Click **Optimal** on the Choose Mixture Design Type panel.
5. Add effects to the model using the instructions below.

Adding Effects to the Model

Initially, the Model panel lists only the main effects corresponding to the factors you entered, as shown in Figure 7.3.

Figure 7.3 The Model Panel



However, you can add factor interactions, specific crossed factor terms, powers, or Scheffe Cubic terms to the model.

- To add interaction terms to a model, click the **Interactions** button and select **2nd**, **3rd**, **4th**, or **5th**. For example, if you have factors X1 and X2, click **Interactions > 2nd** and X1*X2 is added to the list of model effects.
- To add crossed effects to a model, highlight the factors and effects you want to cross and click the **Cross** button.
- To add powers of continuous factors to the model, click the **Powers** button and select **2nd**, **3rd**, **4th**, or **5th**.
- When you want a mixture model with third-degree polynomial terms, the **Scheffe Cubic** button provides a polynomial specification of the surface by adding terms of the form $X1*X2*(X1-X2)$.

The Simplex Centroid Design

A simplex centroid design of degree k with n factors is composed of mixture runs with

- all one factor
- all combinations of two factors at equal levels
- all combinations of three factors at equal levels
- and so on up to k factors at a time combined at k equal levels.

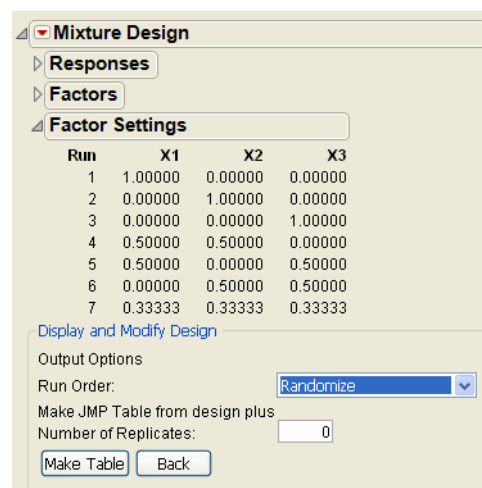
A center point run with equal amounts of all the ingredients is always included.

Creating the Design

To create a simplex centroid design:

1. Select **DOE > Mixture Design**.
2. Enter factors and responses. The steps for entering responses are outlined in [“Enter Responses and Factors into the Custom Designer,” p. 65](#).
3. After you enter responses and factors, click **Continue**.
4. Enter the number of ingredients in the box under **K**. JMP will create runs for each ingredient without mixing, then create runs that mix equal proportions of K ingredients at a time to the specified limit.
5. Click the **Simplex Centroid** button.
6. View factor settings and Output Options, as illustrated in Figure 7.4.

Figure 7.4 Example of Factor Settings and Output Options



7. Specify Run Order, which is the order you want the runs to appear in the data table when it is created. Run order choices are:
 - Keep the Same** the rows (runs) in the output table will appear as they do in the Factor Settings panel.
 - Sort Left to Right** the rows (runs) in the output table will appear sorted from left to right.
 - Randomize** the rows (runs) in the output table will appear in a random order.
 - Sort Right to Left** the rows (runs) in the output table will appear sorted from right to left.
 - Randomize within Blocks** the rows (runs) in the output table will appear in random order within the blocks you set up.
8. Specify Number of Replicates. The number of replicates is the number of times to replicate the entire design, including centerpoints. Type the number of times you want to replicate the design in the associated text box. One replicate doubles the number of runs.

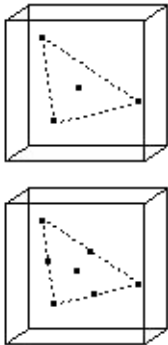
9. Click **Make Table**.

Simplex Centroid Design Examples

The table of runs for a design of degree 1 with three factors (left in Figure 7.5) shows runs for each single ingredient followed by the center point. The table of runs to the right is for three factors of degree 2. The first three runs are for each single ingredient, the second set shows each combination of two ingredients in equal parts, and the last run is the center point.

Figure 7.5 Three-Factor Simplex Centroid Designs of Degrees 1 and 2

Run	X1	X2	X3	Run	X1	X2	X3
1	1	0	0	1	1	0	0
2	0	1	0	2	0	1	0
3	0	0	1	3	0	0	1
4	0.333	0.333	0.333	4	0.5	0.5	0
				5	0.5	0	0.5
				6	0	0.5	0.5
				7	0.333	0.333	0.333



To generate the two sets of runs in Figure 7.5:

1. Choose **DOE > Mixture Design**.
2. Enter three mixture factors.
3. Click **Continue**.
4. Enter 1 in the **K** box, and click **Simplex Centroid** to see the design on the left in Figure 7.6.
5. Click the **Back** button, then click **Continue**, and enter 2 in the **K** box. Then click **Simplex Centroid** to see the design on the right in Figure 7.6.

Figure 7.6 Create Simplex Centroid Designs of Degrees 1 and 2

Factor Settings					Factor Settings				
Run	X1	X2	X3		Run	X1	X2	X3	
1	1.00000	0.00000	0.00000		1	1.00000	0.00000	0.00000	
2	0.00000	1.00000	0.00000		2	0.00000	1.00000	0.00000	
3	0.00000	0.00000	1.00000		3	0.00000	0.00000	1.00000	
4	0.33333	0.33333	0.33333		4	0.50000	0.50000	0.00000	
					5	0.50000	0.00000	0.50000	
					6	0.00000	0.50000	0.50000	
					7	0.33333	0.33333	0.33333	

As another example:

1. Choose **DOE > Mixture Design**.
2. Enter five factors and click **Continue**.
3. Use the default value, 4, in the **K** box.
4. Click **Simplex Centroid**.
5. Click **Make Table** to see the 31-run JMP data table shown in Figure 7.7.

Figure 7.7 Partial Listing of Factor Settings for Five-Factor Simplex Centroid Design

Simplex Centroid			X1	X2	X3	X4	X5	Y
Design	Simplex Centroid							
Model								
Columns (6/0)								
X1 *								
X2 *								
X3 *								
X4 *								
X5 *								
Y *								
Rows								
All rows	31							
Selected	0							
Excluded	0							
Hidden	0							
Labelled	0							

1	0	0.5	0	0	0.5	
2	0.5	0	0	0	0.5	
3	1	0	0	0	0	
4	0.2	0.2	0.2	0.2	0.2	
5	0	0	1	0	0	
6	0	0.3333333333	0.3333333333	0	0.3333333333	
7	0.25	0	0.25	0.25	0.25	
8	0	0	0.5	0	0.5	
9	0.5	0	0.5	0	0	
10	0.3333333333	0.3333333333	0.3333333333	0	0	
11	0.3333333333	0	0.3333333333	0	0.3333333333	
12	0.25	0.25	0	0.25	0.25	
13	0.5	0.5	0	0	0	
14	0.25	0.25	0.25	0	0.25	
15	0.3333333333	0.3333333333	0	0	0.3333333333	
16	0.3333333333	0	0.3333333333	0.3333333333	0	
17	0	0	0	0.5	0.5	

The Simplex Lattice Design

The simplex lattice design is a space filling design that creates a triangular grid of runs. The design is the set of all combinations where the factors' values are i / m , where i is an integer from 0 to m such that the sum of the factors is 1.

To create simplex lattice designs:

1. Select **DOE > Mixture Design**.
2. Enter factors and responses. The steps for entering responses are outlined in [“Enter Responses and Factors into the Custom Designer,” p. 65](#).
3. Click **Continue**.
4. Specify the number of levels you want in the Mixture Design Type dialog (Figure 7.2) and click **Simplex Lattice**.

Figure 7.8 shows the runs for three-factor simplex lattice designs of degrees 3, 4, and 5, with their corresponding geometric representations. In contrast to the simplex centroid design, the simplex lattice design does not necessarily include the centroid.

Figure 7.8 Three-Factor Simplex Lattice Designs for Factor Levels 3, 4, and 5

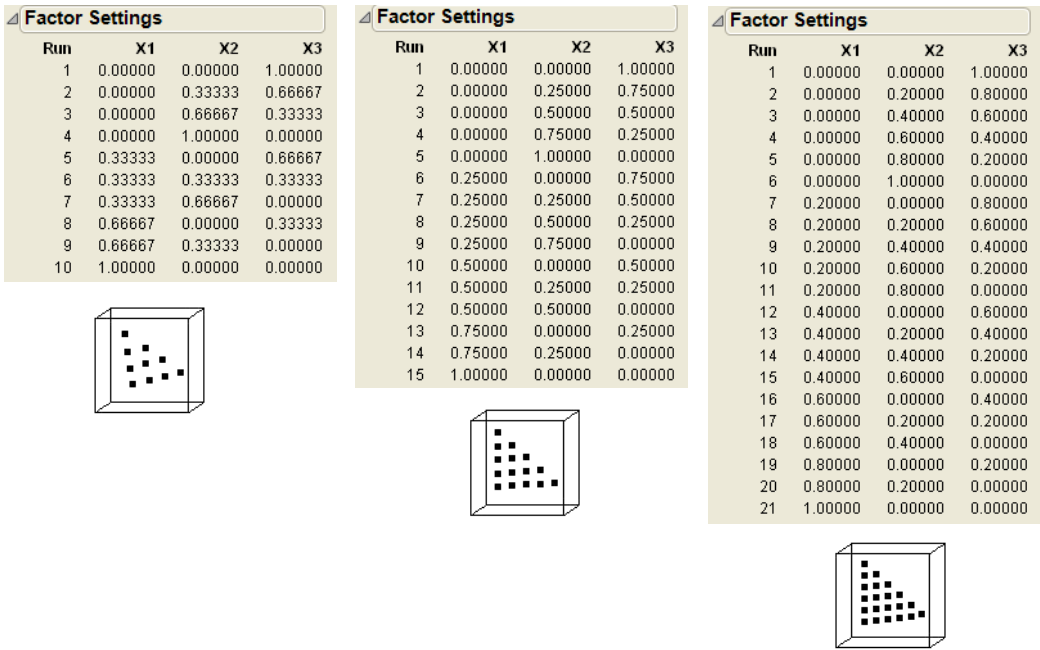


Figure 7.9 lists the runs for a simplex lattice of degree 3 for five effects. In the five-level example, the runs creep across the hyper-triangular region and fill the space in a grid-like manner.

Figure 7.9 JMP Design Table for Simplex Lattice with Five Variables, Order (Degree) 3

Factor Settings					
Run	X1	X2	X3	X4	X5
1	0.00000	0.00000	0.00000	0.00000	1.00000
2	0.00000	0.00000	0.00000	0.33333	0.66667
3	0.00000	0.00000	0.00000	0.66667	0.33333
4	0.00000	0.00000	0.00000	1.00000	0.00000
5	0.00000	0.00000	0.33333	0.00000	0.66667
6	0.00000	0.00000	0.33333	0.33333	0.33333
7	0.00000	0.00000	0.33333	0.66667	0.00000
8	0.00000	0.00000	0.66667	0.00000	0.33333
9	0.00000	0.00000	0.66667	0.33333	0.00000
10	0.00000	0.00000	1.00000	0.00000	0.00000
11	0.00000	0.33333	0.00000	0.00000	0.66667
12	0.00000	0.33333	0.00000	0.33333	0.33333
13	0.00000	0.33333	0.00000	0.66667	0.00000
14	0.00000	0.33333	0.33333	0.00000	0.33333
15	0.00000	0.33333	0.33333	0.33333	0.00000
16	0.00000	0.33333	0.66667	0.00000	0.00000
17	0.00000	0.66667	0.00000	0.00000	0.33333
18	0.00000	0.66667	0.00000	0.33333	0.00000
19	0.00000	0.66667	0.33333	0.00000	0.00000
20	0.00000	1.00000	0.00000	0.00000	0.00000
21	0.33333	0.00000	0.00000	0.00000	0.66667
22	0.33333	0.00000	0.00000	0.33333	0.33333
23	0.33333	0.00000	0.00000	0.66667	0.00000
24	0.33333	0.00000	0.33333	0.00000	0.33333
25	0.33333	0.00000	0.33333	0.33333	0.00000
26	0.33333	0.00000	0.66667	0.00000	0.00000
27	0.33333	0.33333	0.00000	0.00000	0.33333
28	0.33333	0.33333	0.00000	0.33333	0.00000
29	0.33333	0.33333	0.33333	0.00000	0.00000
30	0.33333	0.66667	0.00000	0.00000	0.00000
31	0.66667	0.00000	0.00000	0.00000	0.33333
32	0.66667	0.00000	0.00000	0.33333	0.00000
33	0.66667	0.00000	0.33333	0.00000	0.00000
34	0.66667	0.33333	0.00000	0.00000	0.00000
35	1.00000	0.00000	0.00000	0.00000	0.00000

The Extreme Vertices Design

The extreme vertices design accounts for factor limits and selects vertices and their averages (formed by factor limits) as design points. Additional limits are usually in the form of range constraints, upper bounds, and lower bounds on the factor values.

The extreme vertices design finds the corners (vertices) of a factor space constrained by limits specified for one or more of the factors. The property that the factors must be non-negative and must add up to one is the basic mixture constraint that makes a triangular-shaped region.

Sometimes other ingredients need range constraints that confine their values to be greater than a lower bound or less than an upper bound. Range constraints chop off parts of the triangular-shaped (simplex) region to make additional vertices. It is also possible to have a linear constraint, which defines a linear combination of factors to be greater or smaller than some constant.

The geometric shape of a region bound by linear constraints is called a simplex, and because the vertices represent extreme conditions of the operating environment, they are often the best places to use as design points in an experiment.

You usually want to add points between the vertices. The average of points that share a constraint boundary is called a *centroid* point, and centroid points of various degrees can be added. The centroid point for two

neighboring vertices joined by a line is a second degree centroid because a line is two dimensional. The centroid point for vertices sharing a plane is a third degree centroid because a plane is three dimensional, and so on.

If you specify an extreme vertices design but give no constraints, a simplex centroid design results.

Creating the Design

Follow these steps to create an extreme vertices design. The next sections show examples with specific constraints.

1. Select **DOE > Mixture Design**.
2. Enter factors and responses. These steps are outlined in [“Enter Responses and Factors into the Custom Designer,” p. 65](#). Remember to enter the upper and lower limits in the factors panel (see Figure 7.10).
3. Click **Continue**.
4. In the **Degree** text box, enter the degree of the centroid point you want to add. The centroid point is the average of points that share a constraint boundary.
5. If you have linear constraints, click the **Linear Constraints** button for each constraint you want to add. Use the text boxes that appear to define a linear combination of factors to be greater or smaller than some constant.
6. Click **Extreme Vertices** to see the factor settings.
7. (Optional) Change the order of the runs in the data table when it is created. Run order choices are:
 - Keep the Same**—the rows (runs) in the output table will appear as they do in the Design panel.
 - Sort Left to Right**—the rows (runs) in the output table will appear sorted from left to right.
 - Randomize**—the rows (runs) in the output table will appear in a random order.
 - Sort Right to Left**—the rows (runs) in the output table will appear sorted from right to left.
 - Randomize within Blocks**—the rows (runs) in the output table will appear in random order within the blocks you set up.
8. (Optional) Type the sample size you want in the Choose desired sample size text box.
9. (Optional) Click **Find Subset** to generate the optimal subset having the number of runs specified in sample size box described in Step 8. The **Find Subset** option uses the row exchange method (not coordinate exchange) to find the optimal subset of rows.
10. Click **Make Table**.

An Extreme Vertices Example with Range Constraints

The following example design table is for five factors with the range constraints shown in Figure 7.10, where the ranges are smaller than the default 0 to 1 range.

1. Select **DOE > Mixture Design**.
2. Add two additional factors (for a total of 5 factors) and give them the values shown in Figure 7.10.
3. Click **Continue**.

- Enter '4' in the **Degree** text box (Figure 7.10).

Figure 7.10 Example of Five-factor Extreme Vertices

Factors

Name	Role	Values
X1	Mixture	0.05 0.25
X2	Mixture	0.1 0.3
X3	Mixture	0.1 0.3
X4	Mixture	0.1 0.4
X5	Mixture	0.05 0.25

Choose Mixture Design Type

☒ **Optimal** Create a design tailored to meet specific requirements.

☐ **Simplex Centroid** Run each ingredient without mixing, then mix equal proportions of K ingredients at a time to the specified limit. **K** 4

☐ **Simplex Lattice** Triangular grid. **Number of Levels** 3

☒ **Extreme Vertices** Find the vertices of the simplex. Then add the mid-points of the edges and averages of vertices to the specified degree. **Degree** 4

☐ **Linear Constraint** Add linear constraints on the relative proportions of ingredients. Click once for each constraint.

☐ **ABCD Design** A mixture design for factor screening.

- Click **Extreme Vertices**.
- Select **Sort Left to Right** from the Run Order menu.
- Click **Make Table**. Figure 7.11 shows a partial listing of the resulting table. Note that the Rows panel in the data table shows that the table has the default 116 runs.

Figure 7.11 JMP Design Table for Extreme Vertices with Range Constraints

Extreme Vertices								
Design	Extreme Vertices		X1	X2	X3	X4	X5	Y
Model		1	0.05	0.1	0.2	0.4	0.25	
		2	0.05	0.1	0.25	0.4	0.2	
Columns (6/0)		3	0.05	0.1	0.3	0.4	0.15	
X1 *		4	0.05	0.1	0.266666667	0.366666667	0.216666667	
X2 *		5	0.05	0.15	0.15	0.4	0.25	
X3 *		6	0.05	0.15	0.3	0.4	0.1	
X4 *		7	0.05	0.2	0.1	0.4	0.25	
X5 *		8	0.05	0.2	0.2	0.4	0.15	
Y *		9	0.05	0.2	0.3	0.4	0.05	
		10	0.05	0.25	0.1	0.4	0.2	
Rows		11	0.05	0.25	0.25	0.4	0.05	
All rows	116	12	0.05	0.266666667	0.1	0.366666667	0.216666667	
Selected	0	13	0.05	0.266666667	0.266666667	0.366666667	0.05	
Excluded	0	14	0.05	0.3	0.1	0.4	0.15	
Hidden	0	15	0.05	0.3	0.15	0.4	0.1	
Labelled	0	16	0.05	0.3	0.2	0.4	0.05	

Suppose you want fewer runs. You can go back and enter a different sample size (number of runs).

8. Click **Back**, then click **Continue**.
9. Enter '4' in the **Degree** text box and click **Extreme Vertices**.
10. In the sample size text box, enter '10' as the sample size.
11. Click **Find Subset** to generate the optimal subset having the number of runs specified. The resulting design (Figure 7.12) is the optimal 10-run subset of the 116 current runs. This is useful when the extreme vertices design generates a large number of vertices.

Figure 7.12 JMP Design Table for 10-Run Subset of the 116 Current Runs

Extreme Vertices 2								
Design	Extreme Vertices							
Model								
Columns (6/0)								
X1 *								
X2 *								
X3 *								
X4 *								
X5 *								
Y *								
Rows								
All rows	10							
Selected	0							
Excluded	0							
Hidden	0							
Labelled	0							

	X1	X2	X3	X4	X5	Y
1	0.05	0.3	0.2	0.4	0.05	•
2	0.05	0.1	0.3	0.3	0.25	•
3	0.05	0.3	0.3	0.1	0.25	•
4	0.05	0.3	0.1	0.3	0.25	•
5	0.15	0.1	0.3	0.4	0.05	•
6	0.25	0.1	0.3	0.1	0.25	•
7	0.25	0.1	0.1	0.3	0.25	•
8	0.25	0.2	0.1	0.4	0.05	•
9	0.25	0.3	0.1	0.1	0.25	•
10	0.25	0.3	0.3	0.1	0.05	•

Note: The **Find Subset** option uses the row exchange method (not coordinate exchange) to find the optimal subset of rows.

An Extreme Vertices Example with Linear Constraints

Consider the classic example presented by Snee (1979) and Piepel (1988). This example has three factors, X1, X2, and X3, with five individual factor bound constraints and three additional linear constraints:

Table 7.1 Linear Constraints for the Snee and Piepel Example

$X1 \geq 0.1$	$90 \leq 85*X1 + 90*X2 + 100*X3$
$X1 \leq 0.5$	$85*X1 + 90*X2 + 100*X3 \leq 95$
	$.4 \leq 0.7*X1 + X3$
$X2 \geq 0.1$	
$X2 \leq 0.7$	
$X3 \leq 0.7$	

To enter these constraints:

1. Enter the upper and lower limits in the factors panel.
2. Click **Continue**.
3. Click the **Linear Constraint** button three times. Enter the constraints as shown in Figure 7.13.
4. Click the **Extreme Vertices** button.
5. Change the run order to **Sort Right to Left**, and keep the sample size at 13. See Figure 7.13 for the default Factor Settings and completed Output Options.
6. Click **Make Table**.

Figure 7.13 Constraints

Extreme Vertices Find the vertices of the simplex. Then add the mid-points of the edges and averages of vertices to the specified degree. **Degree** 2

Linear Constraint Add linear constraints on the relative proportions of ingredients. Click once for each constraint.

ABCD Design A mixture design for factor screening.

Back

85 X1 + 90 X2 + 100 X3 ≥ 90

85 X1 + 90 X2 + 100 X3 ≤ 95

0.7 X1 + 0 X2 + 1 X3 ≥ 0.4

Factor Settings

Run	X1	X2	X3
1	0.50000	0.10000	0.40000
2	0.50000	0.25000	0.25000
3	0.26667	0.10000	0.63333
4	0.10000	0.35000	0.55000
5	0.10000	0.57000	0.33000
6	0.33333	0.50000	0.16667
7	0.30000	0.31167	0.38833
8	0.10000	0.46000	0.44000
9	0.50000	0.17500	0.32500
10	0.38333	0.10000	0.51667
11	0.41667	0.37500	0.20833
12	0.18333	0.22500	0.59167
13	0.21667	0.53500	0.24833

Display and Modify Design

Output Options

Run Order: Sort Right to Left

D Optimal Subset

Choose desired sample size: 13

Make Table **Find Subset** **Back**

This example is best understood by viewing the design as a ternary plot, as shown at the end of this chapter, in Figure 7.15. The ternary plot shows how close to one a given component is by how close it is to the vertex of that variable in the triangle. See “[Creating Ternary Plots](#),” p. 168, for details.

Extreme Vertices Method: How It Works

If there are linear constraints, JMP uses the CONSIM algorithm developed by R.E. Wheeler, described in Snee (1979) and presented by Piepel (1988) as CONVRT. The method is also described in Cornell (1990, Appendix 10a). The method combines constraints and checks to see if vertices violate them. If so, it drops the vertices and calculates new ones. The method named CONAEV for doing centroid points is by Piepel (1988).

If there are no linear constraints (only range constraints), the extreme vertices design is constructed using the XVERT method developed by Snee and Marquardt (1974) and Snee (1975). After the vertices are found, a simplex centroid method generates combinations of vertices up to a specified order.

The XVERT method first creates a full 2^{nf-1} design using the given low and high values of the $nf-1$ factors with smallest range. Then, it computes the value of the one factor left out based on the restriction that the factors' values must sum to one. It keeps the point if it is in that factor's range. If not, it increments or decrements it to bring it within range, and decrements or increments each of the other factors in turn by the same amount, keeping the points that still satisfy the initial restrictions.

The above algorithm creates the vertices of the feasible region in the simplex defined by the factor constraints. However, Snee (1975) has shown that it can also be useful to have the centroids of the edges and faces of the feasible region. A generalized n -dimensional face of the feasible region is defined by $nf-n$ of the boundaries and the centroid of a face defined to be the average of the vertices lying on it. The algorithm generates all possible combinations of the boundary conditions and then averages over the vertices generated on the first step.

The ABCD Design

This approach by Snee (1975) generates a screening design for mixtures. To create an ABCD design:

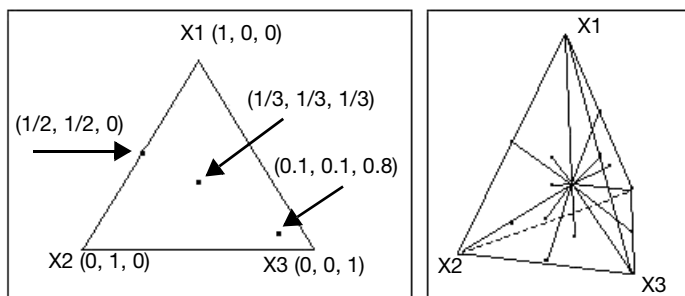
1. Select **DOE > Mixture Design**.
2. Enter factors and responses. The steps for entering responses are outlined in [“Enter Responses and Factors into the Custom Designer,” p. 65](#).
3. After you enter responses and factors, click **Continue**.
4. Click the **ABCD Design** button.
5. View factor settings and Output Options.
6. Specify Run Order, which is the order you want the runs to appear in the data table when it is created. Run order choices are:
 - Keep the Same** The rows (runs) in the output table will appear as they do in the Factor Settings panel.
 - Sort Left to Right** The rows (runs) in the output table will appear sorted from left to right.
 - Randomize** The rows (runs) in the output table will appear in a random order.
 - Sort Right to Left** The rows (runs) in the output table will appear sorted from right to left.
 - Randomize within Blocks** The rows (runs) in the output table will appear in random order within the blocks you set up.
7. Specify Number of Replicates. The number of replicates is the number of times to replicate the entire design, including centerpoints. Type the number of times you want to replicate the design in the associated text box. One replicate doubles the number of runs.
8. Click **Make Table**.

Creating Ternary Plots

A mixture problem in three components can be represented in two dimensions because the third component is a linear function of the others. The *ternary plot* in Figure 7.15 shows how close to one (1) a

given component is by how close it is to the vertex of that variable in the triangle. The plot to the left in Figure 7.14 illustrates a ternary plot.

Figure 7.14 Ternary Plot (left) and Tetrinary Plot (right) for Mixture Design



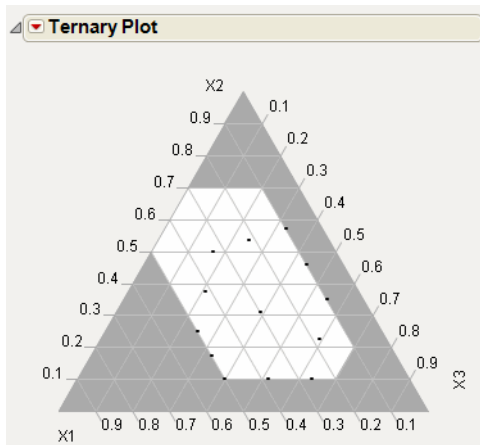
The Piepel (1979) example referenced in [“An Extreme Vertices Example with Linear Constraints,”](#) p. 166 is best understood by the ternary plot shown in Figure 7.15.

To view a mixture design as a ternary plot:

1. Create the Piepel mixture data as shown previously, or open the table called **Piepel.jmp**, found in the Design Experiments folder of the Sample Data Library.
2. Choose **Graph > Ternary Plot**.
3. In the ternary plot launch dialog, specify the three mixture components and click **OK**.

The JMP Ternary plot platform recognizes the three factors as mixture factors, and also considers the upper and lower constraints entered into the Factors panel when the design was created. The Ternary plot uses shading to exclude the unfeasible areas excluded by those constraints.

The Piepel data had additional constraints, entered as linear constraints for the extreme vertices design. There are six active constraints, six vertices, and six centroid points shown on the plot, as well as two inactive (redundant) constraints. The feasible area is the inner white polygon delimited by the design points and constraint lines.

Figure 7.15 Diagram of Ternary Plot Showing Piepel Example Constraints

Fitting Mixture Designs

When fitting a model for mixture designs, you must take into account that all the factors add up to a constant, and thus a traditional full linear model will not be fully estimable.

The recommended response surface model, called the Scheffé polynomial (Scheffé 1958), does the following:

- suppresses the intercept
- includes all the linear main-effect terms
- excludes all the square terms (such as $X_1 \cdot X_1$)
- includes all the cross terms (such as $X_1 \cdot X_2$)

To fit a model:

1. Choose **DOE > Mixture Design** and make the design data table. Remember that to fit a model, the Y column in the data table must contain values, so either assign responses or click the red triangle menu and select **Simulate Responses** before you click **Make Table**.
2. The design data table stores the model in the data table as a table property. This table property is a JSL script called **Model**, located in the left panel of the table.
3. Right-click the model and select **Run Script** to launch the Fit Model dialog, which is automatically filled with the saved model.
4. Click **Run** on the Fit Model dialog.

In this model, the parameters are easy to interpret (Cornell 1990). The coefficients on the linear terms are the fitted response at the extreme points where the mixture consists of a single factor. The coefficients on the cross terms indicate the curvature across each edge of the factor space.

The model report usually has several sections of interest, including the whole model tests, Analysis of Variance reports, and response surface reports, which are described below.

Whole Model Tests and Analysis of Variance Reports

In a whole-model Analysis of Variance table, JMP traditionally tests that all the parameters are zero except for the intercept. In a mixture model without an intercept, JMP looks for a hidden intercept, in the sense that a linear combination of effects is a constant. If it finds a hidden intercept, it does the whole model test with respect to the intercept model rather than a zero-intercept model. This test is equivalent to testing that all the parameters are zero except the linear parameters, and testing that they are equal.

The hidden-intercept property also causes the R^2 to be reported with respect to the intercept model rather than reported as missing.

Understanding Response Surface Reports

When there are effects marked as response surface effects “&RS,” JMP creates additional reports that analyze the fitted response surface. These reports were originally designed for full response surfaces, not mixture models. However, if JMP encounters a no-intercept model and finds a hidden intercept with linear response surface terms, but no square terms, then it *folds* its calculations, collapsing on the last response surface term to calculate critical values for the optimum. This can be done for any combination that yields a constant and involves the last response surface term.

A Chemical Mixture Example

Three plasticizers (p1, p2, and p3) comprise 79.5% of the vinyl used for automobile seat covers (Cornell, 1990). Within this 79.5%, the individual plasticizers are restricted by the following constraints: $0.409 \leq x_1 \leq 0.849$, $0 \leq x_2 \leq 0.252$, and $0.151 \leq x_3 \leq 0.274$.

Create the Design

To create Cornell’s mixture design in JMP:

1. Select **DOE > Mixture Design**.
2. In the Factors panel, use the three default factors but name them p1, p2, and p3, and enter the high and low constraints as shown in Figure 7.16. Or, load the factors with the **Load Factors** command in the red triangle on the Mixture Design title bar. To import the factors, open **Plastifactors.jmp**, found in the Design Experiment Sample Data folder that was installed with JMP.

Figure 7.16 Factors and Factor Constraints for the Plasticizer Experiment

Mixture Design

Responses

Add Response Remove Number of Responses...

Response Name	Goal	Lower Limit	Upper Limit	Importance
Y	Maximize	.	.	.

optional item

Factors

Add 1 Mixture Remove Selected

Name	Role	Values
p1	Mixture	0.409 0.849
p2	Mixture	0 0.252
p3	Mixture	0.151 0.274

3. Click **Continue**.
4. Enter 3 in the **Degree** text box.
5. Click **Extreme Vertices**.
6. Click **Make Table**. JMP uses the 9 factor settings to generate a JMP table (Figure 7.17).

Figure 7.17 Extreme Vertices Mixture Design

	p1	p2	p3	Y
1	0.849	0	0.151	•
2	0.6	0.126	0.274	•
3	0.7875	0	0.2125	•
4	0.6615	0.126	0.2125	•
5	0.726	0	0.274	•
6	0.474	0.252	0.274	•
7	0.5355	0.252	0.2125	•
8	0.723	0.126	0.151	•
9	0.597	0.252	0.151	•

7. Add an extra five design runs by duplicating the vertex points and center point, to give a total of 14 rows in the table.

Note: To identify the vertex points and the center (or interior) point, use the sample data script called `LabelMixturePoints.jsl` in the Sample Scripts folder installed with JMP.

- Run the `LabelMixturePoints.jsl` to see the results in Figure 7.18, and highlight the vertex points and the interior point as shown.

Figure 7.18 Identify Vertices and Center Point with Sample Script

Extreme Vertices						
Design	Extreme Vertices					
Model						
Ternary Plot						
Columns (5/0)						
p1 *						
p2 *						
p3 *						
Y *						
Point Type						
Rows						
All rows	9					
Selected	5					

		p1	p2	p3	Y	Point Type
1	0.849	0	0.151			Vertex
2	0.6	0.126	0.274			Edge
3	0.7875	0	0.2125			Edge
4	0.6615	0.126	0.2125			Interior
5	0.726	0	0.274			Vertex
6	0.474	0.252	0.274			Vertex
7	0.5355	0.252	0.2125			Edge
8	0.723	0.126	0.151			Edge
9	0.597	0.252	0.151			Vertex

- Select **Edit > Copy**, to copy the selected rows to the clipboard.
- Select **Rows > Add Rows** and type **5** as the number of rows to add.
- Click the At End radio button on the dialog, then click **OK**.
- Highlight the new rows and select **Edit > Paste** to add the duplicate rows to the table.

The Plasticizer data with the results (Y values) that Cornell obtained are available in the Sample data. Open `Plasticizer.jmp` in the Sample Data folder installed with JMP to see this table (Figure 7.19).

Figure 7.19 Plasticizer.jmp Data Table from the Sample Data Library

Plasticizer							
Design	Extreme Vertices						
Model							
Ternary Plot							
Columns (6/0)							
p1 *							
p2 *							
p3 *							
Y *							
Point Type							
Pred Formula Y							
Rows							
All rows	14						
Selected	0						
Excluded	0						
Hidden	0						
Labelled	0						

	p1	p2	p3	Y	Point Type	Pred Formula Y
1	0.474	0.252	0.274	12	Vertex	10.892307
2	0.726	0	0.274	4	Vertex	5.0923076
3	0.849	0	0.151	8	Vertex	7.4923076
4	0.597	0.252	0.151	13	Vertex	11.292307
5	0.6	0.126	0.274	13	Edge	13.030769
6	0.6615	0.126	0.2125	18	Interior	19.269230
7	0.7875	0	0.2125	12	Edge	11.830769
8	0.723	0.126	0.151	14	Edge	14.430769
9	0.5355	0.252	0.2125	16	Edge	16.630769
10	0.474	0.252	0.274	10	Vertex	10.892307
11	0.726	0	0.274	6	Vertex	5.0923076
12	0.849	0	0.151	7	Vertex	7.4923076
13	0.597	0.252	0.151	10	Vertex	11.292307
14	0.6615	0.126	0.2125	21	Interior	19.269230

Analyze the Mixture Model

Use the Cornell plasticizer data from the Sample Data library (Figure 7.19) to run the mixture model:

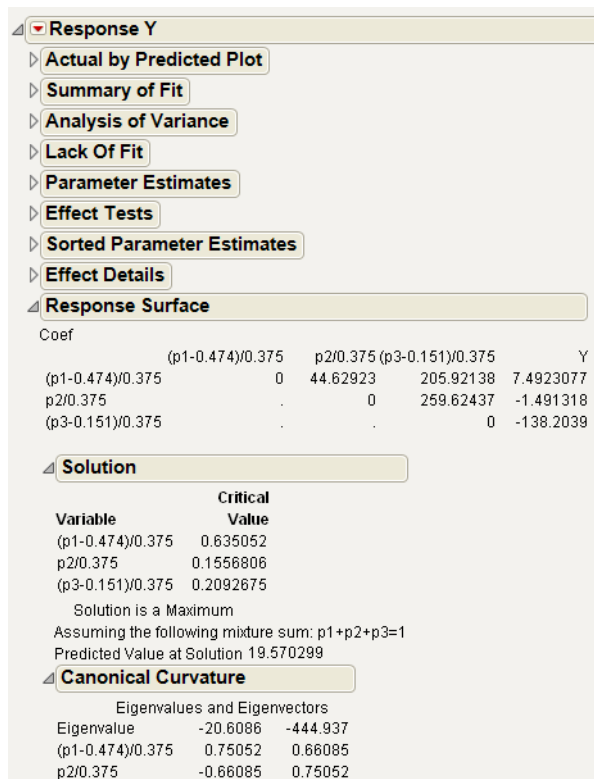
1. Click the table property named Model on the upper-left of the data table and select **Run Script**, which runs a script that displays a completed Fit Model dialog. Click **Run** to see the response surface analysis.
2. Plasticizer.jmp contains a column called Pred Formula Y. This column was added after the analysis by selecting **Save Columns > Prediction Formula** from the red triangle menu in the Response Y title bar of the analysis report. To see the prediction formula, right-click (Ctrl+click on the Mac) the column name and select **Formula**:

$$0 - 50.1465 * p1 - 282.1982 * p2 - 911.6484 * p3 + p2 * p1 * 317.363 + p3 * p1 * 1464.3298 + p3 * p2 * 1846.2177$$

Note: These results correct the coefficients reported in Cornell (1990).

The Response Surface Solution report (Figure 7.20) shows that a maximum predicted value of 19.570299 occurs at point (0.63505, 0.15568, 0.20927).

Figure 7.20 Mixture Response Surface Analysis



The Prediction Profiler

The report contains a prediction profiler.

1. If the profiler is not visible, click the red triangle in the Response Y title bar and select **Factor Profiling** > **Profiler**. You should see the initial profiler shown in Figure 7.21.

The crossed effects show as curvature in the prediction traces. Drag one of the vertical reference lines, and the other two move in the opposite direction maintaining their ratio.

Note: The axes of prediction profiler traces range from the upper and lower bounds of the factors, p1, p2, and p3, entered to create the design and the design table. When you experiment moving a variable trace, you see the other traces move such that their ratio is preserved. As a result, when the limit of a variable is reached, it cannot move further and only the third variable changes.

2. To limit the visible profile curves to bounds that use all three variables, use the **Profile at Boundary** > **Stop at Boundaries** command from the menu on the Prediction Profiler title bar.
3. If needed, select the **Desirability Functions** command to display the desirability function showing to the right of the prediction profile plots in Figure 7.22.
4. Then select **Maximize Desirability** from the Prediction Profiler menu to see the best factor settings.

The profiler in Figure 7.22, displays optimal settings (rounded) of 0.6350 for p1, 0.1557 for p2, and 0.2093 for p3, which give an estimated response of 19.5703.

Figure 7.21 Initial Prediction Profiler

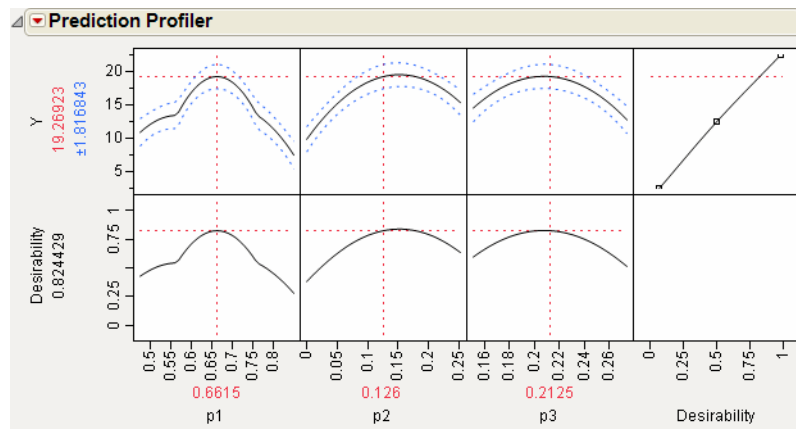
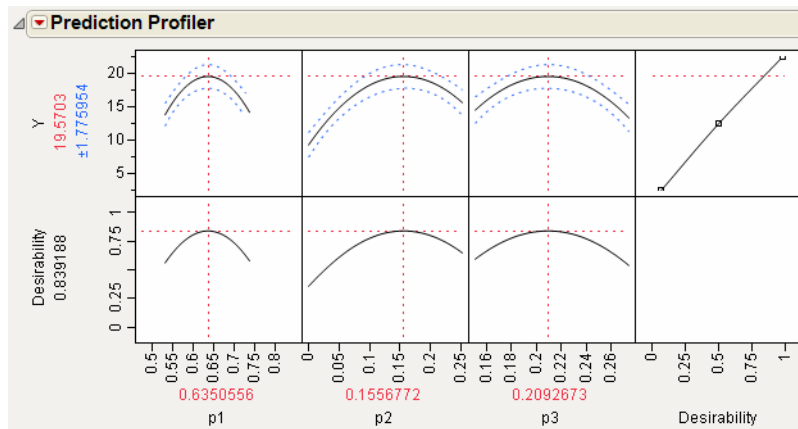


Figure 7.22 Maximum Desirability in Profiler for Mixture Analysis Example


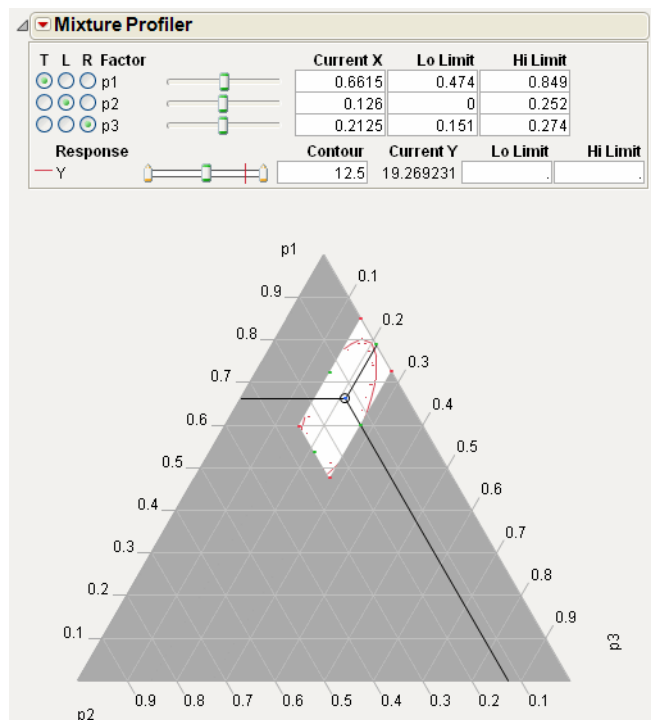
The Mixture Profiler

The Fit Model report also has a **Mixture Profiler** that is useful for visualizing and optimizing response surfaces from mixture experiments.

Many of the features are the same as those of the Contour Profiler however some are unique to the Mixture Profiler:

- A ternary plot is used instead of a Cartesian plot, which enables you to view three mixture factors at a time.
- If you have more than three factors, radio buttons let you choose which factors to plot.
- If the factors have constraints, you can enter their low and high limits in the Lo Limit and Hi Limit columns. This shades non-feasible regions in the profiler.

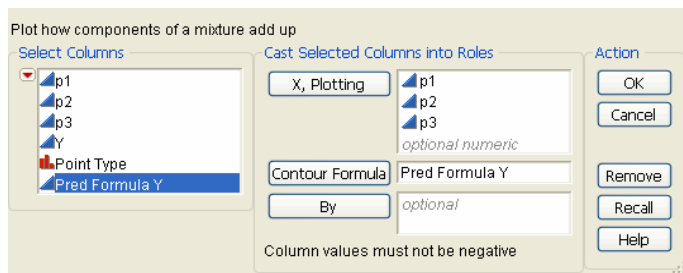
Select **Factor Profiling > Mixture Profiler** from the menu on the Response Y title bar to see the mixture profiler for the plasticizer data, shown in Figure 7.23.

Figure 7.23 Mixture Profiler for Plasticizer Example

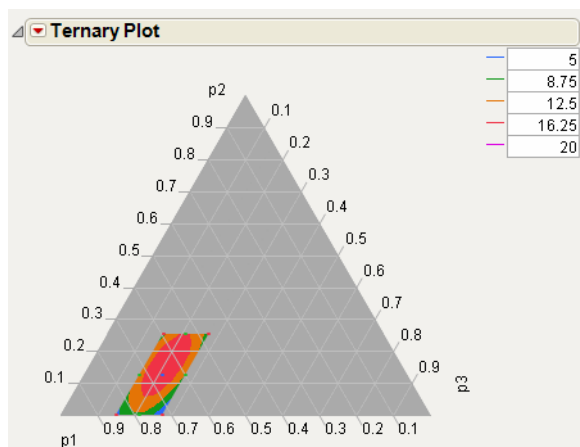
A Ternary Plot of the Mixture Response Surface

You can also plot the response surface of the plasticizer data as a ternary plot using the Ternary graph platform and contour the plot with information from an additional variable:

1. Choose **Graph > Ternary Plot**.
2. Specify plot variables (p1, p2, and p3) and click **X, Plotting**, as shown in Figure 7.24. To identify the contour variable (the prediction equation), select **Pred Formula Y** and click the **Contour Formula** button. The contour variable must have a prediction formula to form the contour lines, as shown by the ternary plots at the bottom in Figure 7.25. If there is no prediction formula, the ternary plot only shows points.

Figure 7.24 Launch Dialog for the Ternary Plot Platform

3. Click **OK** and view the results, as shown in Figure 7.25. By default, the ternary plot displays contour lines only, but you can request a **Fill Above** or **Fill Below**, as shown, with the **Contour Fill** command found in the red triangle menu on the Ternary Plot title bar.

Figure 7.25 Ternary Plot of a Mixture Response Surface

Chapter 8

Discrete Choice Designs



The Discrete Choice designer creates experiments with factors that are product attributes. A collection of attributes is called a product profile. Respondents choose one in each set of product profiles.

Industrial experimentation deals with the question of how to improve processes to deliver better products. Choice experiments help a company prioritize product features for their market. The purpose of a choice experiment is to define a product that people want to buy.

Choice experiments always involve people comparing prospective products and picking the one they prefer. For example, suppose a computer company wants to update its high-end laptop. Laptops have many features that are important to customers such as processor speed, hard disk size, screen size, battery life, and price. To build a laptop that customers want, the computer company needs to know the relative importance of each feature. Most people prefer a faster computer with more storage, longer battery life, and a low price. What the company does not know is how much more an extra hour of battery life is worth to a customer or whether doubling the hard disk size is as important as doubling the processor speed. A choice experiment can answer these questions and indicate the optimal set of trade-offs among product features.

Contents

- Introduction 181
- Create an Example Choice Experiment 183
- Analyze the Example Choice Experiment 186
- Design a Choice Experiment Using Prior Information 188
- Administer the Survey and Analyze Results 190
 - Initial Choice Platform Analysis 191
 - Find Unit Cost and Trade Off Costs with the Profiler 192

Introduction

Most choice experiments involve conducting a market research survey. The survey consists of a series of questions about attributes of interest about a potential new product or product modification. For example, a computer manufacturer might be interested in manufacturing a new laptop and wants information about customer preference before beginning an expensive development process. Computer characteristics change so rapidly that it is crucial to quickly identify the attributes that help the manufacturers to design and build a new machine most likely to capture enough market share to be profitable.

Often, the attributes are obvious. For example, the consumer wants a laptop that has a large screen, weighs almost nothing, costs almost nothing, and lasts forever on a single battery charge. The question, then, is how much is the customer willing to compromise these desires? How important is each of these attributes, and which kinds of trade off is the customer most likely to accept and still purchase a new machine?

Assume a simple situation where a computer manufacturer wants to examine preferences for four possible laptop configurations. Notice that there are no ‘right’ or ‘wrong’ selections. Instead there are just preferences. A well designed questionnaire and proper analysis of results can tell a manufacturer how to proceed. The manufacturer wants information about the following four laptop attributes.

- size of hard drive disk (40 GB or 80 GB)
- speed of processor (1.5 GHz or 2.0 GHz)
- battery life (4 Hrs or 6 Hrs)
- cost of computer (\$1000, \$1200 or \$1500)

If a survey were constructed that offered the possibility of choosing any combination of these attributes, a respondent would be forced to evaluate 24 possible combinations and make a single response. Instead each respondent usually evaluates several choice sets and for each choice set, chooses the preferred profile. In the simplest situation, each respondent chooses between sets of two profiles.

Then, you analyze the choices of multiple respondents. A well designed choice experiment, correctly analyzed is a efficient way to give the researcher the most information for the least time and expense.

Table 1 shows hypothetical results from a single *survey* designed to collect information about consumer preferences about laptop computers.

- Each column in the survey identifies a laptop *attribute*.
- Each line in the survey defines a laptop *profile*, which is a collection of attribute values.
- Each *choice set* consists of two attribute profiles.
- All of the attribute values are allowed to change across the two profiles in a choice set.

Table 8.1 Hypothetical Choice Survey Results from a single Respondent, Subject ID 2

For each pair, please check the combination of attributes you find most appealing.						
		Disk Size	Speed	Battery Life	Price	Preference
1	1	40 GB	1.5 GHz	6 hours	\$1,000	<u>X</u>
	2	80 GB	1.5 GHz	4 hours	\$1,200	—

Table 8.1 Hypothetical Choice Survey Results from a single Respondent, Subject ID 2 (Continued)

For each pair, please check the combination of attributes you find most appealing.						
		Disk Size	Speed	Battery Life	Price	Preference
2	1	40 GB	1.5 GHz	4 hours	\$1,500	—
	2	80 GB	2.0 GHz	4 hours	\$1,200	<u>X</u>
3	1	40 GB	2.0 GHz	4 hours	\$1,200	<u>X</u>
	2	80 GB	2.0 GHz	6 hours	\$1,500	—
4	1	40 GB	2.0 GHz	4 hours	\$1,000	<u>X</u>
	2	80 GB	1.5 GHz	6 hours	\$1,200	—
5	1	40 GB	1.5 GHz	6 hours	\$1,000	<u>X</u>
	2	40 GB	2.0 GHz	4 hours	\$1,500	—
6	1	40 GB	2.0 GHz	6 hours	\$1,200	<u>X</u>
	2	80 GB	1.5 GHz	4 hours	\$1,500	—
7	1	40 GB	2.0 GHz	6 hours	\$1,500	—
	2	80 GB	1.5 GHz	4 hours	\$1,000	<u>X</u>
8	1	40 GB	1.5 GHz	4 hours	\$1,200	—
	2	80 GB	2.0 GHz	4 hours	\$1,000	<u>X</u>

The DOE Choice designer can create a survey like that shown in Table 8.1. However, to create an effective design, the Choice designer needs information about the attributes. For example, most laptop attributes have values that are intrinsic preferences. That is, a bigger disk size is better, longer battery life is better, and so forth. The purpose of conducting a choice survey is to find out how the potential laptop purchasers feels about the advantages of a collection of tractates.

One way to gain prior information about attributes in a survey is to conduct a single example survey, analyze the results, and use those results as prior information to create the final survey instrument.

This chapter shows how to create a sample survey and use its results as prior information for a final survey design

Create an Example Choice Experiment

The Choice design can create a survey like the one in Table 1. As an example,

1. Choose **DOE > Choice Design**, and complete the initial dialog as shown in Figure 8.1.

Figure 8.1 Choice Design Dialog with Attributes Defined

The screenshot shows the 'Choice Design' dialog box with the 'Attributes' section expanded. It includes buttons for 'Add Factor', 'Remove', and 'Add N Factors' (set to 1). Below is a table listing attributes and their levels.

Name	Role	Attribute Levels		
▼ Disk Size	Categorical	40 GB	80 GB	
▼ Speed	Categorical	1.5 GHz	2.0 GHz	
▼ Battery Life	Categorical	4 Hrs	6 Hrs	
▼ Price	Categorical	\$1,500	\$1,200	\$1,000

Below the table, there is a 'Specify Attributes' section with instructions: 'Add an attribute by clicking the Add Factor button. Double-click an attribute name or level to edit it.' and a 'Continue' button.

2. Click **Continue**. For this example, use the default values in the Model Control panel and in the Design Generation panel, as shown in Figure 8.2.

Optionally, you can use the DOE Model Controls panel to add interactions to the choice model in situations where you expect there are interactions and want to generate profile sets that will help detect them.

Figure 8.2 Design Generation Panel for the Laptop Experiment

Choice Design

Attributes

Model

DOE Model Controls

Main Effects Interactions Remove Term

Name
Disk Size
Speed
Battery Life
Price
<i>optional item</i>

Prior Specification

Design Generation

4 Number of attributes that can change within a choice set

2 Number of profiles per choice set

8 Number of choice sets per survey

1 Number of surveys

1 Expected number of respondents per survey

Make Design

Back

The values in the Design Generation panel describe the laptop survey.

Number of attributes that can change within a choice set There are four laptop attributes.

Entering 4 as the number of attributes that can change within a choice set means that the Choice designer can change 4 or fewer attribute values within a single choice set. You can enter fewer than the total number of attributes to constrain the total number that can be changed within a choice set. This might be a reasonable thing to do if you had a large number of attributes and you want to make it easier for the respondents to make a choice. For example, a survey might be interested in 20 or more attributes of a cell phone, but show and change only 5 or fewer attributes in a choice set.

Number of profiles per choice set This example has two profiles per choice set. You can design choice experiments with more than two profiles in a choice set.

Number of choice sets per survey There are eight choice sets in the example survey but often there are many more.

Number of surveys The example only shows a single survey. Normally you expect multiple respondents and would request more than one survey.

Expected number of respondents per survey You might want to give surveys to 10 people, but use two different surveys. So you enter 2 as the Number of surveys and 5 as the Expected number of respondents per survey.

Note: Recall that this first example is used to generate prior information, then used to create a more realistic survey. This example is a single survey given to a single respondent.

3. Click **Make Design** to see the example survey results in Figure 8.3.

Figure 8.3 Survey Results based on a Simple Model and Default Prior Information

Choice Design

Attributes

Model

Design

Choice Set	Disk Size	Speed	Battery Life	Price
1	80 GB	2.0 GHz	6 Hrs	\$1,500
1	40 GB	1.5 GHz	4 Hrs	\$1,200
2	40 GB	2.0 GHz	6 Hrs	\$1,500
2	80 GB	2.0 GHz	4 Hrs	\$1,200
3	80 GB	1.5 GHz	6 Hrs	\$1,000
3	40 GB	2.0 GHz	4 Hrs	\$1,500
4	80 GB	2.0 GHz	4 Hrs	\$1,500
4	40 GB	2.0 GHz	6 Hrs	\$1,000
5	40 GB	2.0 GHz	6 Hrs	\$1,200
5	80 GB	1.5 GHz	4 Hrs	\$1,000
6	80 GB	1.5 GHz	6 Hrs	\$1,200
6	40 GB	2.0 GHz	4 Hrs	\$1,000
7	40 GB	1.5 GHz	4 Hrs	\$1,200
7	80 GB	2.0 GHz	4 Hrs	\$1,000
8	40 GB	2.0 GHz	4 Hrs	\$1,200
8	40 GB	1.5 GHz	4 Hrs	\$1,500

☐ Output separate tables for profiles and responses

☒ Combine profiles and responses in one table

Make Table

Back

The Radio buttons beneath the design settings let you choose between having the survey settings in one JMP table and gathering survey results in a second table, or generating a single table that shows the settings and has an additional column for the choice response.

You can see that there are eight choice sets, each consisting of two laptop profiles. At this point you can press **Back** and modify the design, or click **Make Table** and generate the JMP table shown in Figure 8.4.

This default design was created with no given prior information. Without prior information, the Choice designer has no way of knowing which attribute levels are better. That is, the Choice designer cannot know that a lower price might be more desirable than a higher price, a faster machine is better than a slower machine, and so forth. As a result, you can see that some choice sets might not convey useful information. The analysis results are used as prior information in a new Choice design dialog.

Figure 8.4 JMP Data Table for Preliminary Laptop Choice Survey

Choice Profiles							
Design	Discrete Choice						
Choice							
Columns (6/0)							
Choice Set							
Response Indicator							
Disk Size							
Speed							
Battery Life							
Price							
Rows							
All rows	16						
Selected	0						
Excluded	0						
Hidden	0						
Labelled	0						

	Choice Set	Response Indicator	Disk Size	Speed	Battery Life	Price
1	1		80 GB	2.0 GHz	6 Hrs	\$1,500
2	1		40 GB	1.5 GHz	4 Hrs	\$1,200
3	2		40 GB	2.0 GHz	6 Hrs	\$1,500
4	2		80 GB	2.0 GHz	4 Hrs	\$1,200
5	3		80 GB	1.5 GHz	6 Hrs	\$1,000
6	3		40 GB	2.0 GHz	4 Hrs	\$1,500
7	4		80 GB	2.0 GHz	4 Hrs	\$1,500
8	4		40 GB	2.0 GHz	6 Hrs	\$1,000
9	5		40 GB	2.0 GHz	6 Hrs	\$1,200
10	5		80 GB	1.5 GHz	4 Hrs	\$1,000
11	6		80 GB	1.5 GHz	6 Hrs	\$1,200
12	6		40 GB	2.0 GHz	4 Hrs	\$1,000
13	7		40 GB	1.5 GHz	4 Hrs	\$1,200
14	7		80 GB	2.0 GHz	4 Hrs	\$1,000
15	8		40 GB	2.0 GHz	4 Hrs	\$1,200
16	8		40 GB	1.5 GHz	4 Hrs	\$1,500

Analyze the Example Choice Experiment

Once a survey design is complete, a respondent chooses one profile from each set, entering '1' for the chosen profile and '0' for the rejected profile. Suppose a respondent completed the example survey as shown in Figure 8.5. You can now analyze these results using the Choice platform in the **Analyze** menu (**Analyze > Modeling > Choice**).

Figure 8.5 JMP Table with Survey Choice Sets and Responses

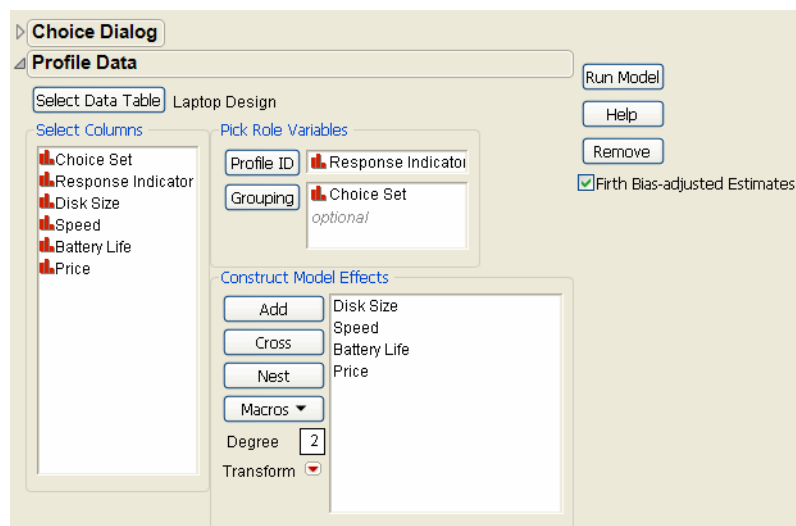
Laptop Design							
Design	Discrete Choice						
Choice							
Columns (6/0)							
Choice Set							
Response Indicator							
Disk Size							
Speed							
Battery Life							
Price							
Rows							
All rows	16						
Selected	0						
Excluded	0						
Hidden	0						
Labelled	0						

	Choice Set	Response Indicator	Disk Size	Speed	Battery Life	Price
1	1	1	40 GB	1.5 GHz	6 Hrs	\$1,000
2	1	0	80 GB	1.5 GHz	4 Hrs	\$1,200
3	2	0	40 GB	1.5 GHz	4 Hrs	\$1,500
4	2	1	80 GB	2.0 GHz	4 Hrs	\$1,200
5	3	1	40 GB	2.0 GHz	4 Hrs	\$1,200
6	3	0	80 GB	2.0 GHz	6 Hrs	\$1,500
7	4	1	40 GB	2.0 GHz	4 Hrs	\$1,000
8	4	0	80 GB	1.5 GHz	6 Hrs	\$1,200
9	5	1	40 GB	1.5 GHz	6 Hrs	\$1,000
10	5	0	40 GB	2.0 GHz	4 Hrs	\$1,500
11	6	1	40 GB	2.0 GHz	6 Hrs	\$1,200
12	6	0	80 GB	1.5 GHz	4 Hrs	\$1,500
13	7	0	40 GB	2.0 GHz	6 Hrs	\$1,500
14	7	1	80 GB	1.5 GHz	4 Hrs	\$1,000
15	8	0	40 GB	1.5 GHz	4 Hrs	\$1,200
16	8	1	80 GB	2.0 GHz	4 Hrs	\$1,000

1. Click the Choice script in the Laptop Design data table and select **Run Script** to analyze that data with the Choice modeling platform from the **Analyze** menu.

When you run the Choice model script, the Choice launch dialog shown in Figure 8.6 appears. The Choice dialog is designed to cover a variety of choice survey results, which can include data saved in multiple data tables. This example has all data contained in a single table. For details about using the Choice analysis dialog, see the Choice Modeling chapter in the *Modeling and Multivariate Methods* book.

Figure 8.6 Choice Model Fitting Dialog



2. Click **Run Model** on the Choice model fitting dialog.
3. An additional dialog then appears asking if this is a one-table analysis with all the data in the Profile Table, which is the case in this sample survey. Click **Yes** in this dialog to continue.

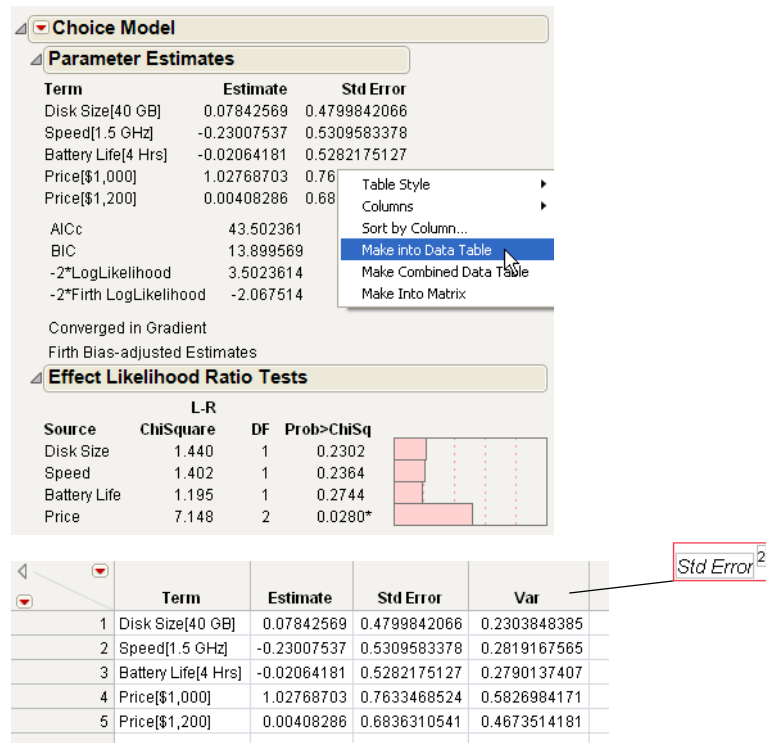
The analysis shows as in Figure 8.7.

To design the final choice survey using prior information, you will need to enter estimates of the mean and variance of the attribute parameter estimates. The analysis on the left in Figure 8.7 has estimates of the attribute means, called **Estimate**, and estimates of the standard deviation of the attributes, called **Std Error**. An easy way to see the variance of the attributes is to capture the analysis in a JMP table and compute the variance:

4. Right click on the Parameter Estimates report and choose **Make into Data Table** from the menu, as shown.
5. In the new Untitled data table, create a new column and call it Var.
6. Select **Formula** from the **Cols** menu (**Cols > Formula**), or right-click at the top of the Var column and select **Formula** for the menu that shows.

7. In the Formula Editor, click the Std Error column in the column list and click the exponent button (x^y) on the formula editor panel to compute the variance shown on the right in Figure 8.7.

Figure 8.7 Analysis of the sample Laptop Survey



This preliminary survey with its analysis gives you the information needed to design a final survey appropriate for gathering information from multiple respondents. Keep in mind that in a real situation, you might have prior information about factor attributes and not need to do a sample design.

Note: Leave the Untitled data table with the mean and variance information open to be used in the next example.

Design a Choice Experiment Using Prior Information

In some situations, you will know from previous surveys or experience how to give prior information to the Choice designer about product attributes. This example continues by designing the laptop experiment again, using the analysis information gained from the sample design.

1. Choose **DOE > Choice Design** and enter the attributes and values as before.

2. Click **Continue** to see the Choice design panels in Figure 8.8.
3. Now enter the values from the JMP table created by the previous analysis into the Prior Mean and Prior Variance Matrix panels of the Choice Design dialog, as shown in Figure 8.8. You can copy-and-paste to transfer the values from the data table to the Choice dialog panels.

Figure 8.8 Enter Prior Mean and Variance Information from Preliminary Survey

The screenshot shows the JMP Choice Design dialog box with the following panels and data:

Attributes Panel:

Name	Role	Attribute Levels			
✓ Disk Size	Categorical	40 GB	80 GB		
✓ Speed	Categorical	1.5 GHz	2.0 GHz		
✓ Battery Life	Categorical	4 Hrs	6 Hrs		
✓ Price	Categorical	\$1,000	\$1,200	\$1,500	

Model Panel:

DOE Model Controls

Prior Specification

☐ Ignore prior specifications. Generate the Utility Neutral design.

Prior Mean

Effect	Prior Mean
Disk Size	0.078
Speed	-0.23
Battery Life	-0.02
Price 1	1.028
Price 2	0.004

☐ Ignore prior variance. Generate the local design for the prior mean.

Prior Variance Matrix

Effect	Disk Size	Speed	Battery Life	Price 1	Price 2
Disk Size	0.230	0.000	0.000	0.000	0.000
Speed		0.282	0.000	0.000	0.000
Battery Life			0.279	0.000	0.000
Price 1				0.583	0.000
Price 2					0.467

JMP Table (Left):

Estimate	Std Error	Var
0.0784256	0.47998420	0.23038
-0.2300753	0.53095833	0.28191
-0.0206418	0.52821751	0.27901
1.0276870	0.76334685	0.58269
0.0040828	0.68363105	0.46735

4. Enter the values into the Design Generation panel, as shown in Figure 8.9.
 - Four or fewer attribute levels can change within a choice set.
 - There are two profiles per choice set.
 - Each survey has eight choice sets.
 - The design generates two separate surveys.
 - Five respondents are expected to complete each survey (for a total of 10 respondents).

Figure 8.9 Design Specifications for final Laptop Survey

Design Generation

4

Number of attributes that can change within a choice set

2

Number of profiles per choice set

8

Number of choice sets per survey

2

Number of surveys

5

Expected number of respondents per survey

Make Design

Back

5. Click **Make Design**. A partial view of the design is shown in Figure 8.10.

Figure 8.10 Design Runs for Two Choice Surveys

Design

Survey	Choice Set	Disk Size	Speed	Battery Life	Price
1	1	80 GB	2.0 GHz	6 Hrs	\$1,500
1	1	40 GB	1.5 GHz	4 Hrs	\$1,200
1	2	40 GB	1.5 GHz	6 Hrs	\$1,200
1	2	80 GB	2.0 GHz	4 Hrs	\$1,000
1	3	40 GB	2.0 GHz	6 Hrs	\$1,000
1	3	80 GB	1.5 GHz	4 Hrs	\$1,200
1	4	40 GB	1.5 GHz	4 Hrs	\$1,200
1	4	80 GB	2.0 GHz	6 Hrs	\$1,000
1	5	40 GB	2.0 GHz	4 Hrs	\$1,500
1	5	80 GB	1.5 GHz	6 Hrs	\$1,200
1	6	80 GB	1.5 GHz	6 Hrs	\$1,500
1	6	40 GB	2.0 GHz	4 Hrs	\$1,200
1	7	80 GB	1.5 GHz	6 Hrs	\$1,000
1	7	40 GB	2.0 GHz	4 Hrs	\$1,200
1	8	40 GB	2.0 GHz	6 Hrs	\$1,200
1	8	80 GB	1.5 GHz	4 Hrs	\$1,000

6. Click **Make Table**. The final data table will have runs for ten survey respondents, giving a total of 160 observations (2 profiles * 8 choice sets * 2 surveys * 5 respondents = 160 observations).

Administer the Survey and Analyze Results

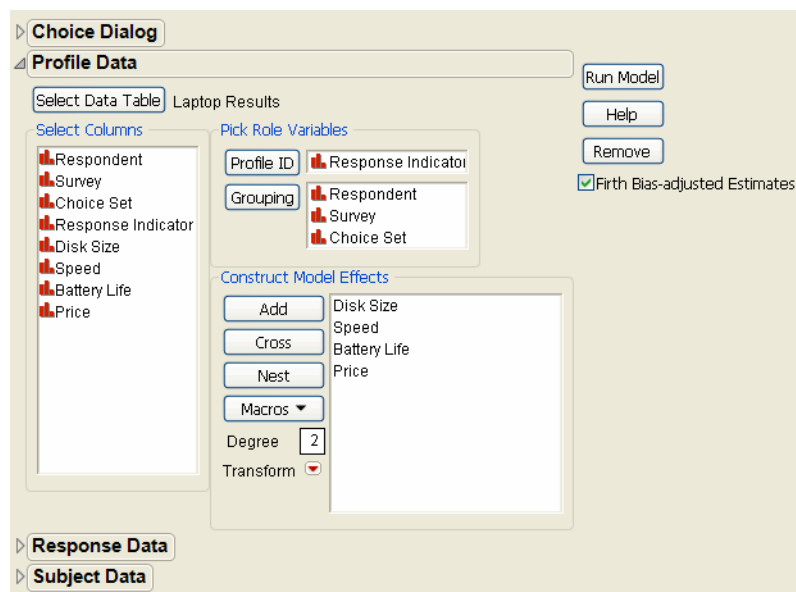
The survey data table with its results is stored in the Sample Data Design Experiment folder installed with JMP. Figure 8.11 is a partial listing of the survey data table with results. The Choice script created by the Choice designer and saved with the survey data table can be used to analyze the data. The default data table created by the Choice designer is named Choice Profiles. Note in Figure 8.11, the data table name is changed to Laptop Results.

Initial Choice Platform Analysis

1. To continue the example, open the table called **Laptop Results.jmp**, saved in the Sample Data folder.
2. To analyze the data, click the Choice script saved with the data and select **Run Script** from the menu to see the completed dialog shown in Figure 8.11.

Note that this dialog has three grouping variables (**Respondent**, **Survey**, and **Choice Set**, whereas the dialog shown Figure 8.6 had only the **Choice Set** grouping variable because there was a single survey and a single respondent. This example included multiple surveys and respondents, which must be included in the analysis.

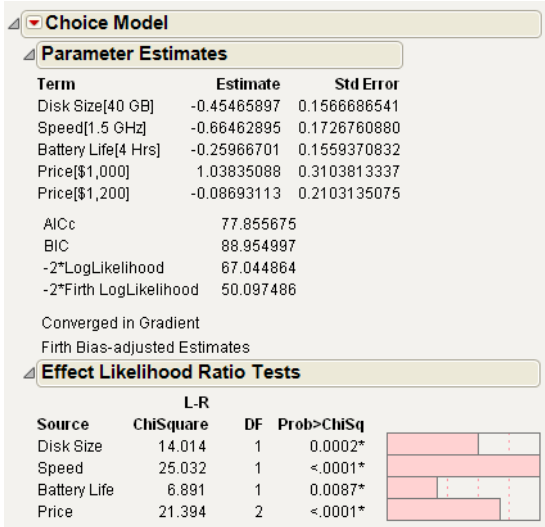
Figure 8.11 Choice Model Fitting Dialog to Analyze the Laptop Survey



3. Click **Run** on the Fit Model dialog. The query again appears asking if the analysis is a one-table analysis with all the data in the profile table. Click **Yes** to see the initial analysis result shown in Figure 8.12.

The results are clear. While all the effects are significant, the most significant attribute is **Speed**.

Figure 8.12 Initial Analysis of the Final Laptop Survey

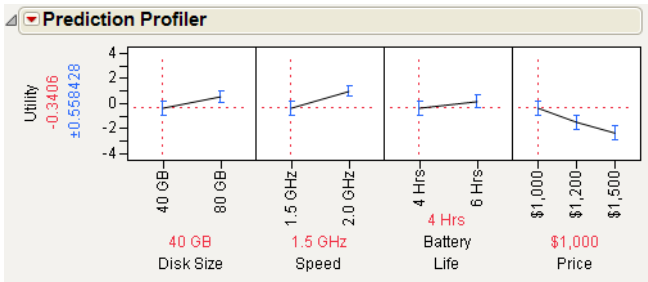


Find Unit Cost and Trade Off Costs with the Profiler

You would like to know how changing the price, or other characteristics, of a laptop affects the desirability as perceived by potential buyers. This desirability is called the *utility value* of the laptop attributes. The profiler shows the utility value and how it changes as the laptop attributes change.

1. Select **Profiler** from the menu on the Choice Model title bar to see the Prediction Profiler in Figure 8.13.

Figure 8.13 Default Prediction Profiler for Laptop Choice Analysis

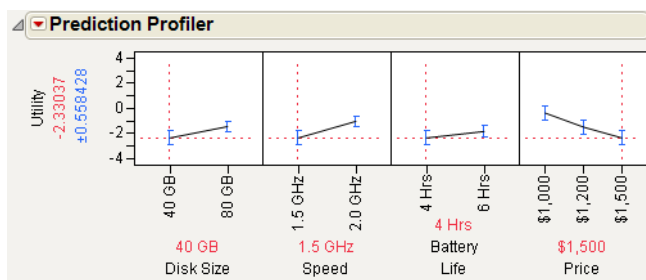


When each attribute value is set to its lowest value, the Utility value is -0.3406. The first thing you want to know is the unit utility cost.

2. To find the unit utility cost, move the trace for Price to \$1,500 and note how the Utility value changes.

Compare the Utility values in Figure 8.13 and Figure 8.14. The value of Utility changes from -0.3406 to -2.3303 when cost is raised from \$1,000 to \$1,500. That is, raising the price of a laptop \$500.00 lowers the utility (or desirability) approximately 2 units. Thus, you can say that the unit utility cost is roughly \$250.00.

Figure 8.14 Compare Change in Utility Over Price

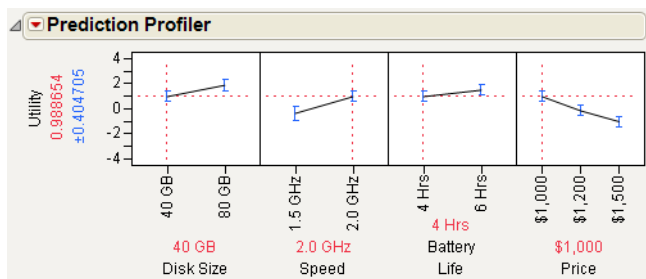


With this unit utility cost estimate you can now vary the other attributes, note the change in utility, and find an approximate dollar value associated with that attribute change. For example, the most significant attribute is speed (see Figure 8.12).

3. In the Prediction Profiler, set Price to its lowest value and change Speed to its higher value.

You can see in Figure 8.15 that the Utility value changes from the original value shown in Figure 8.13 of -0.3406 to 0.9886 , for a total change of 1.3292 units. If the unit utility cost is estimated to be \$250.00, as shown above, then the increase in price for a 2.0 GHz laptop over a 1.5 GHz laptop can be computed to be $1.3292 \times \$250.00 = \332.30 . This is the dollar value the Choice survey provides the manufacturer as a basis for pricing different laptop products. You can make similar calculations for the other attributes.

Figure 8.15 Change Speed in Profiler and Note Utility Value



This simple Choice survey and its analysis shows how this kind of information can be used to help manufacturers and retailers identify important product attributes and assign values to them.

Administer the Survey and Analyze Results

The Choice designer allows more complex designs, such as designs with interactions and other terms. The Choice analysis platform can be used to analyze complex designs, and it can be used to incorporate data from multiple data sets that include demographic information about the respondents.

Space-Filling Designs



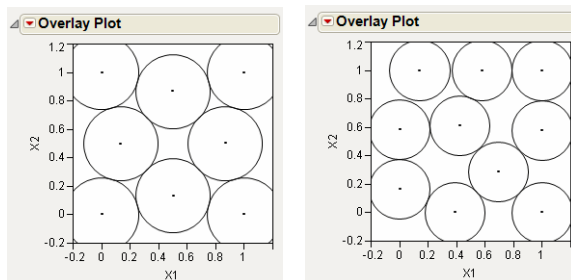
Space-filling designs are useful in situations where run-to-run variability is of far less concern than the form of the model. Sensitivity studies of computer simulations is one such situation. For this case, and any mechanistic or deterministic modeling problem, any variability is small enough to be ignored. For systems with no variability, randomization and blocking are irrelevant. Replication is undesirable because repeating the same run yields the same result. In space-filling designs, there are two objectives:

- Prevent replicate points by spreading the design points out to the maximum distance possible between any two points.
- Space the points uniformly.

The following methods are implemented for these types of designs:

- The Sphere-Packing method emphasizes spread of points.
- The Latin Hypercube method is a compromise between spread of points and uniform spacing.
- The Uniform method mimics the uniform probability distribution.
- The Minimum Potential method minimizes energy designs in a hypersphere.
- The Maximum Entropy method measures the amount of information contained in the distribution of a set of data.
- The Gaussian Process IMSE Optimal method creates a design that minimizes the integrated mean squared error of the gaussian process over the experimental region.

Figure 9.1 Space-Filling Design



Contents

Introduction to Space-Filling Designs	197
Sphere-Packing Designs.	197
Creating a Sphere-Packing Design	197
Visualizing the Sphere-Packing Design	199
Latin Hypercube Designs	200
Creating a Latin Hypercube Design.	200
Visualizing the Latin Hypercube Design	202
Uniform Designs.	204
Comparing Sphere-Packing, Latin Hypercube, and Uniform Methods	206
Minimum Potential Designs	207
Maximum Entropy Designs.	209
Gaussian Process IMSE Optimal Designs	211
Borehole Model: A Sphere-Packing Example	211
Create the Sphere-Packing Design for the Borehole Data	212
Guidelines for the Analysis of Deterministic Data	213
Results of the Borehole Experiment.	213

Introduction to Space-Filling Designs

Space-filling designs are useful for modeling systems that are deterministic or near-deterministic. One example of a deterministic system is a computer simulation. Such simulations can be very complex involving many variables with complicated interrelationships. A goal of designed experiments on these systems is to find a simpler empirical model that adequately predicts the behavior of the system over limited ranges of the factors.

In experiments on systems where there is substantial random noise, the goal is to minimize the variance of prediction. In experiments on deterministic systems, there is no variance but there is *bias*. Bias is the difference between the approximation model and the true mathematical function. The goal of space-filling designs is to bound the bias.

There are two schools of thought on how to bound the bias. One approach is to spread the design points out as far from each other as possible consistent with staying inside the experimental boundaries. The other approach is to space the points out evenly over the region of interest.

The **Space Filling** designer supports the following design methods:

Sphere Packing maximizes the minimum distance between pairs of design points.

Latin Hypercube maximizes the minimum distance between design points but requires even spacing of the levels of each factor. This method produces designs that mimic the uniform distribution. The Latin Hypercube method is a compromise between the Sphere-Packing method and the Uniform design method.

Uniform minimizes the discrepancy between the design points (which have an empirical uniform distribution) and a theoretical uniform distribution.

Minimum Potential spreads points out inside a sphere around the center.

Maximum Entropy measures the amount of information contained in the distribution of a set of data.

Gaussian Process IMSE Optimal creates a design that minimizes the integrated mean squared error of the Gaussian process over the experimental region.

Sphere-Packing Designs

The Sphere-Packing design method maximizes the minimum distance between pairs of design points. The effect of this maximization is to spread the points out as much as possible inside the design region.

Creating a Sphere-Packing Design

1. Select **DOE > Space Filling Design**.
2. Enter responses and factors. (See [“Enter Responses and Factors into the Custom Designer,” p. 65.](#))
3. Alter the factor level values, if necessary. For example, Figure 9.2 shows the two existing factors, X1 and X2, with values that range from 0 to 1 (instead of the default –1 to 1).

Figure 9.2 Space-Filling Design Dialog for Two Factors

Space Filling Design

Responses

Add Response Remove Number of Responses...

Response Name	Goal	Lower Limit	Upper Limit	Importance
Y <i>optional item</i>	Maximize	.	.	.

Factors

Add 1 Continuous Remove Selected

Name	Role	Values
X1	Continuous	0 1
X2	Continuous	0 1

[Specify Factors](#)

Specify desired number of factors. Double click on a factor name or setting to edit it.

Continue

- Click **Continue**.
- In the design specification dialog, specify a sample size (**Number of Runs**). Figure 9.3 shows a sample size of eight.

Figure 9.3 Space-Filling Design Dialog

[Space Filling Design Methods](#)

Number of Runs: 8

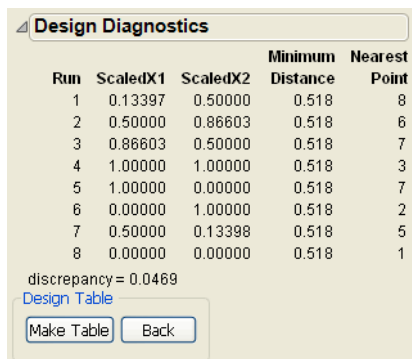
Sphere Packing	Optimal Packing of Spheres Inside of a Cube.
Latin Hypercube	Latin HyperCube with Optimal Spacing
Uniform	Uniform Design
Minimum Potential	Minimum energy designs in a spherical region.
Maximum Entropy	Maximum entropy designs for a Gaussian process
Gaussian Process IMSE Optimal	Integrated mean square error optimal designs for a Gaussian process

Back

- Click **Sphere Packing**.

JMP creates the design and displays the design runs and the design diagnostics. Figure 9.4 shows the Design Diagnostics panel open with 0.518 as the **Minimum Distance**. Your results might differ slightly from the ones below, but the minimum distance will be the same.

Figure 9.4 Sphere-Packing Design Diagnostics



Run	ScaledX1	ScaledX2	Minimum Distance	Nearest Point
1	0.13397	0.50000	0.518	8
2	0.50000	0.86603	0.518	6
3	0.86603	0.50000	0.518	7
4	1.00000	1.00000	0.518	3
5	1.00000	0.00000	0.518	7
6	0.00000	1.00000	0.518	2
7	0.50000	0.13398	0.518	5
8	0.00000	0.00000	0.518	1

discrepancy = 0.0469
Design Table
Make Table Back

- Click **Make Table**. Use this table to complete the visualization example, described next.

Visualizing the Sphere-Packing Design

To visualize the nature of the Sphere-Packing technique, create an overlay plot, adjust the plot's frame size, and add circles using the minimum distance from the diagnostic report shown in Figure 9.4 as the radius for the circles. Using the table you just created:

- Select **Graph > Overlay Plot**.
- Specify X1 as X and X2 as Y, then click **OK**.
- Adjust the frame size so that the frame is square by right-clicking the plot and selecting **Size/Scale > Size to Isometric**.
- Right-click the plot and select **Customize**. When the Customize panel appears, click the plus sign to see a text edit area and enter the following script:

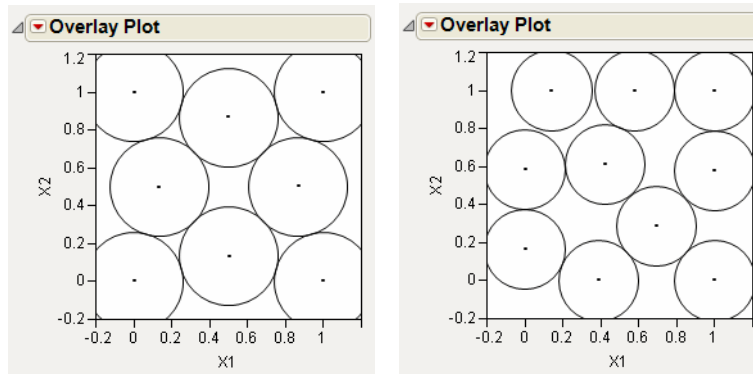
```
For Each Row(Circle({:X1, :X2}, 0.518/2))
```

where 0.518 is the minimum distance number you noted in the Design Diagnostics panel. This script draws a circle centered at each design point with radius 0.259 (half the diameter, 0.518), as shown on the left in Figure 9.5. This plot shows the efficient way JMP packs the design points.
- Now repeat the procedure exactly as described in the previous section, but with a sample size of 10 instead of eight.

Remember to change 0.518 in the graphics script to the minimum distance produced by 10 runs. When the plot appears, again set the frame size and create a graphics script using the minimum distance from the diagnostic report as the diameter for the circle. You should see a graph similar to the one on the right in Figure 9.5. Note the irregular nature of the sphere packing. In fact, you can repeat the process a

third time to get a slightly different picture because the arrangement is dependent on the random starting point.

Figure 9.5 Sphere-Packing Example with Eight Runs (left) and 10 Runs (right)



Latin Hypercube Designs

In a Latin Hypercube, each factor has as many levels as there are runs in the design. The levels are spaced evenly from the lower bound to the upper bound of the factor. Like the sphere-packing method, the Latin Hypercube method chooses points to maximize the minimum distance between design points, but with a constraint. The constraint maintains the even spacing between factor levels.

Creating a Latin Hypercube Design

To use the Latin Hypercube method:

1. Select **DOE > Space Filling Design**.
2. Enter responses, if necessary, and factors. (See [“Enter Responses and Factors into the Custom Designer,”](#) p. 65.)
3. Alter the factor level values, if necessary. For example, Figure 9.6 shows adding two factors to the two existing factors and changing their values to 1 and 8 instead of the default -1 and 1.

Figure 9.6 Space-Filling Dialog for Four Factors

Space Filling Design

Responses

Add Response Remove Number of Responses...

Response Name	Goal	Lower Limit	Upper Limit	Importance
Y <i>optional item</i>	Maximize	.	.	.

Factors

Add 2 Continuous Remove Selected

Name	Role	Values
X1	Continuous	1 8
X2	Continuous	1 8
X3	Continuous	1 8
X4	Continuous	1 8

[Specify Factors](#)

Specify desired number of factors. Double click on a factor name or setting to edit it.

Continue

4. Click **Continue**.
5. In the design specification dialog, specify a sample size (**Number of Runs**). This example uses a sample size of eight.
6. Click **Latin Hypercube** (see Figure 9.3). Factor settings and design diagnostics results appear similar to those in Figure 9.7, which shows the Latin Hypercube design with four factors and eight runs.

Note: The purpose of this example is to show that each column (factor) is assigned each level only once, and each column is a different permutation of the levels.

Figure 9.7 Latin Hypercube Design for Four Factors and Eight Runs with Eight Levels

4 Factors
Space Filling Latin Hypercube

Factor Settings

Run	X1	X2	X3	X4
1	6.00000	6.00000	8.00000	5.00000
2	7.00000	7.00000	2.00000	6.00000
3	2.00000	5.00000	1.00000	4.00000
4	3.00000	2.00000	7.00000	2.00000
5	8.00000	3.00000	4.00000	3.00000
6	1.00000	4.00000	6.00000	8.00000
7	5.00000	1.00000	3.00000	7.00000
8	4.00000	8.00000	5.00000	1.00000

Design Diagnostics

Run	ScaledX1	ScaledX2	ScaledX3	ScaledX4	Minimum Distance	Nearest Point
1	0.71429	0.71429	1.00000	0.57143	0.821	5
2	0.85714	0.85714	0.14286	0.71429	0.782	5
3	0.14286	0.57143	0.00000	0.42857	0.833	2
4	0.28571	0.14286	0.85714	0.14286	0.845	1
5	1.00000	0.28571	0.42857	0.28571	0.782	2
6	0.00000	0.42857	0.71429	1.00000	0.845	7
7	0.57143	0.00000	0.28571	0.85714	0.782	5
8	0.42857	1.00000	0.57143	0.00000	0.821	1

discrepancy = 0.0397

[Design Table](#)

Visualizing the Latin Hypercube Design

To visualize the nature of the Latin Hypercube technique, create an overlay plot, adjust the plot's frame size, and add circles using the minimum distance from the diagnostic report as the radius for the circle.

1. First, create another Latin Hypercube design using the default X1 and X2 factors.
2. Be sure to change the factor values so they are 0 and 1 instead of the default -1 and 1.
3. Click **Continue**.
4. Specify a sample size of eight (**Number of Runs**).
5. Click **Latin Hypercube**. Factor settings and design diagnostics are shown in Figure 9.8.

Figure 9.8 Latin Hypercube Design with two Factors and Eight Runs

Space Filling Latin Hypercube

Factor Settings

Run	X1	X2
1	0.85714	1.00000
2	0.14286	0.00000
3	0.57143	0.14286
4	0.42857	0.85714
5	0.00000	0.71429
6	0.28571	0.42857
7	0.71429	0.57143
8	1.00000	0.28571

Design Diagnostics

Run	ScaledX1	ScaledX2	Minimum Distance	Nearest Point
1	0.85714	1.00000	0.452	4
2	0.14286	0.00000	0.452	3
3	0.57143	0.14286	0.404	6
4	0.42857	0.85714	0.404	7
5	0.00000	0.71429	0.404	6
6	0.28571	0.42857	0.404	3
7	0.71429	0.57143	0.404	8
8	1.00000	0.28571	0.404	7

discrepancy = 0.0092

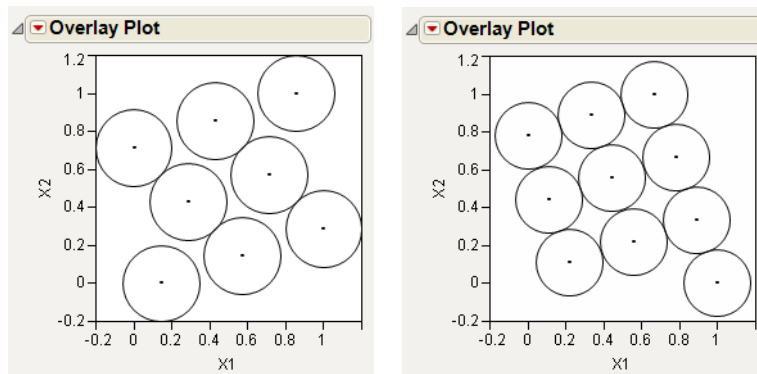
The minimum distance

- Click **Make Table**.
- Select **Graph > Overlay Plot**.
- Specify X1 as X and X2 as Y, then click **OK**.
- Right-click the plot and select **Size/Scale > Size to Isometric** to adjust the frame size so that the frame is square.
- Right-click the plot, select **Customize** from the menu. In the Customize panel, click the large plus sign to see a text edit area, and enter the following script:

```
For Each Row(Circle({:X1, :X2}, 0.404/2))
```

where 0.404 is the minimum distance number you noted in the Design Diagnostics panel (Figure 9.8). This script draws a circle centered at each design point with radius 0.202 (half the diameter, 0.404), as shown on the left in Figure 9.9. This plot shows the efficient way JMP packs the design points.
- Repeat the above procedure exactly, but with 10 runs instead of eight (step 5). Remember to change 0.404 in the graphics script to the minimum distance produced by 10 runs.

You should see a graph similar to the one on the right in Figure 9.9. Note the irregular nature of the sphere packing. In fact, you can repeat the process to get a slightly different picture because the arrangement is dependent on the random starting point.

Figure 9.9 Comparison of Latin Hypercube Designs with Eight Runs (left) and 10 Runs (right)

Note that the minimum distance between each pair of points in the Latin Hypercube design is smaller than that for the Sphere-Packing design. This is because the Latin Hypercube design constrains the levels of each factor to be evenly spaced. The Sphere-Packing design maximizes the minimum distance without any constraints.

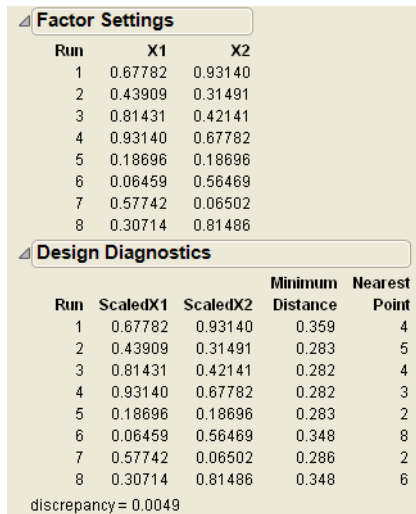
Uniform Designs

The Uniform design minimizes the discrepancy between the design points (empirical uniform distribution) and a theoretical uniform distribution.

Note: These designs are most useful for getting a simple and precise estimate of the integral of an unknown function. The estimate is the average of the observed responses from the experiment.

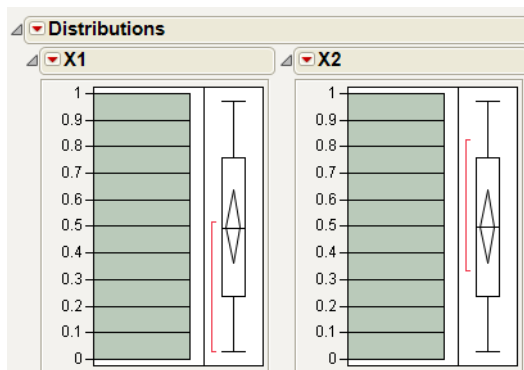
1. Select **DOE > Space Filling Design**.
2. Enter responses, if necessary, and factors. (See [“Enter Responses and Factors into the Custom Designer,”](#) p. 65.)
3. Alter the factor level values to 0 and 1.
4. Click **Continue**.
5. In the design specification dialog, specify a sample size. This example uses a sample size of eight (**Number of Runs**).
6. Click the **Uniform** button. JMP creates this design and displays the design runs and the design diagnostics as shown in Figure 9.10.

Note: The emphasis of the Uniform design method is not to spread out the points. The minimum distances in Figure 9.10 vary substantially.

Figure 9.10 Factor Settings and Diagnostics for Uniform Space-Filling Designs with Eight Runs

7. Click **Make Table**.

A Uniform design does not guarantee even spacing of the factor levels. However, increasing the number of runs and running a distribution on each factor (use **Analyze > Distribution**) shows flat histograms.

Figure 9.11 Histograms are Flat for each Factor when Number of Runs is Increased to 20

Comparing Sphere-Packing, Latin Hypercube, and Uniform Methods

To compare space-filling design methods, create the **Sphere Packing**, **Latin Hypercube**, and **Uniform** designs, as shown in the previous examples. The Design Diagnostics tables show the values for the factors scaled from zero to one. The minimum distance is based on these scaled values and is the minimum distance from each point to its closest neighbor. The discrepancy value is the integrated difference between the design points and the uniform distribution.

Figure 9.12 shows a comparison of the design diagnostics for three eight-run space-filling designs. Note that the discrepancy for the Uniform design is the smallest (best). The discrepancy for the Sphere-Packing design is the largest (worst). The discrepancy for the Latin Hypercube takes an intermediate value that is closer to the optimal value.

Also note that the minimum distance between pairs of points is largest (best) for the Sphere-Packing method. The Uniform design has pairs of points that are only about half as far apart. The Latin Hypercube design behaves more like the Sphere-Packing design in spreading the points out.

For both spread and discrepancy, the Latin Hypercube design represents a healthy compromise solution.

Figure 9.12 Comparison of Diagnostics for Three Eight-Run Space-Filling Methods

Design Diagnostics					
Run	ScaledX1	ScaledX2	Minimum Distance	Nearest Point	
1	0.86603	0.50000	0.518	6	Sphere-Packing
2	0.50000	0.13397	0.518	6	
3	0.00000	1.00000	0.518	5	
4	1.00000	1.00000	0.518	5	
5	0.50000	0.86603	0.518	8	
6	1.00000	0.00000	0.518	1	
7	0.00000	0.00000	0.518	8	
8	0.13397	0.50000	0.518	7	
discrepancy = 0.0469					

Design Diagnostics					
Run	ScaledX1	ScaledX2	Minimum Distance	Nearest Point	
1	0.57143	0.85714	0.404	4	Latin Hypercube
2	0.71429	0.42857	0.404	3	
3	1.00000	0.71429	0.404	2	
4	0.28571	0.57143	0.404	1	
5	0.42857	0.14286	0.404	2	
6	0.85714	0.00000	0.452	2	
7	0.00000	0.28571	0.404	4	
8	0.14286	1.00000	0.452	1	
discrepancy = 0.0092					

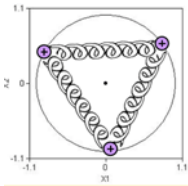
Design Diagnostics					
Run	ScaledX1	ScaledX2	Minimum Distance	Nearest Point	
1	0.43532	0.93541	0.284	8	Uniform
2	0.18514	0.69286	0.285	6	
3	0.56557	0.18590	0.277	5	
4	0.93541	0.56468	0.284	5	
5	0.81486	0.30724	0.277	3	
6	0.06460	0.43504	0.285	2	
7	0.30959	0.06908	0.281	3	
8	0.69286	0.81486	0.284	1	
discrepancy = 0.0046					

Another point of comparison is the time it takes to compute a design. The Uniform design method requires the most time to compute. Also, the time to compute the design increases rapidly with the number of runs. For comparable problems, all the space-filling design methods take longer to compute than the *D*-optimal designs in the Custom Designer.

Minimum Potential Designs

The Minimum Potential design spreads points out inside a sphere. To understand how this design is created, imagine the points as electrons with springs attached to every other point, as illustrated to the right. The coulomb force pushes the points apart, but the springs pull them together. The design is the spacing of points that minimizes the potential energy of the system.

Figure 9.13 Minimum Potential Design



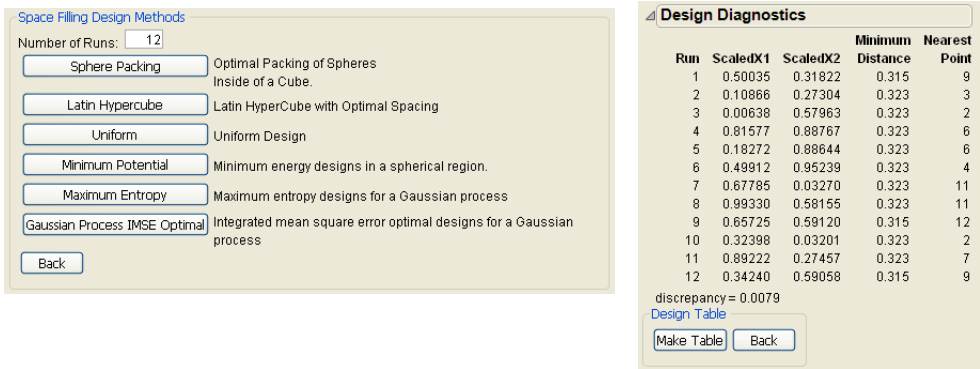
Minimum Potential designs:

- have spherical symmetry
- are nearly orthogonal
- have uniform spacing

To see a Minimum Potential example:

1. Select **DOE > Space Filling Design**.
2. Add 1 continuous factor. (See [“Enter Responses and Factors into the Custom Designer,”](#) p. 65.)
3. Alter the factor level values to 0 and 1, if necessary.
4. Click **Continue**.
5. In the design specification dialog (shown on the left in Figure 9.14), enter a sample size (**Number of Runs**). This example uses a sample size of 12.
6. Click the **Minimum Potential** button. JMP creates this design and displays the design runs (shown on the right in Figure 9.14) and the design diagnostics.

Figure 9.14 Space-Filling Methods and Design Diagnostics for Minimum Potential Design



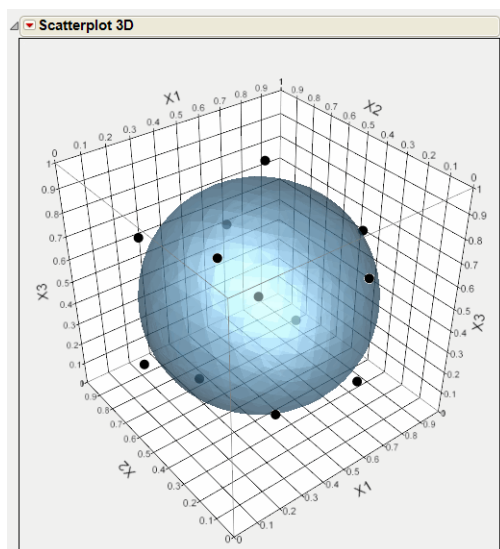
7. Click **Make Table**.

You can see the spherical symmetry of the Minimum Potential design using the Scatterplot 3D graphics platform.

1. After you make the JMP design table, choose the **Graph > Scatterplot 3D** command.
2. In the Scatterplot 3D launch dialog, select X1, X2, and X3 as **Y, Columns** and click **OK** to see the initial three dimensional scatterplot of the design points.
3. To see the results similar to those in Figure 9.15, select the **Normal Contour Ellipsoids** option from the menu in the Scatterplot 3D title bar, and make the points larger by right-clicking on the plot and selecting **Settings**, then increasing the **Marker Size** slider.

Now it is easy to see the points spread evenly on the surface of the ellipsoid.

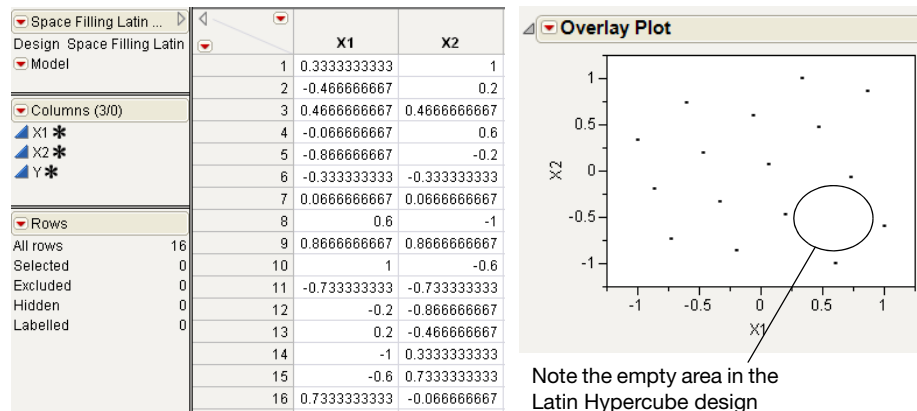
Figure 9.15 Minimum Potential Design Points on Sphere



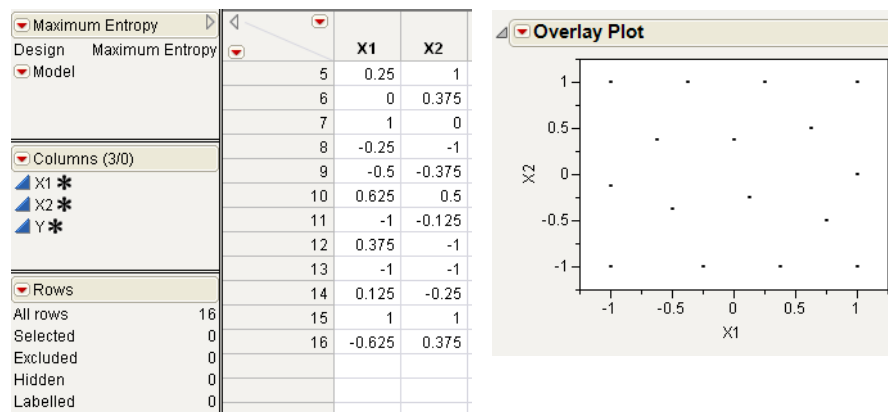
Maximum Entropy Designs

The Latin Hypercube design is currently the most popular design assuming you are going to analyze the data using a Gaussian-Process model. Computer simulation experts like to use the Latin Hypercube design because all projections onto the coordinate axes are uniform.

However, as the example at the top in Figure 9.16 shows, the Latin Hypercube design does not necessarily do a great job of space filling. This is a two-factor Latin Hypercube with 16 runs, with the factor level settings set between -1 and 1. Note that this design appears to leave a hole in the bottom right of the overlay plot.

Figure 9.16 Two-factor Latin Hypercube Design

The Maximum Entropy design is a competitor to the Latin Hypercube design for computer experiments because it optimizes a measure of the amount of information contained in an experiment. See the technical note below. With the factor levels set between -1 and 1, the two-factor Maximum Entropy design shown in Figure 9.17 covers the region better than the Latin hypercube design in Figure 9.16. The space-filling property generally improves as the number of runs increases without bound.

Figure 9.17 Two-Factor Maximum Entropy Design

Technical Note: Maximum Entropy designs maximize the Shannon information (Shewry and Wynn (1987)) of an experiment, assuming that the data come from a normal $(m, s^2 R)$ distribution, where

$$R_{ij} = \exp\left(-\sum_k \theta_k (x_{ik} - x_{jk})^2\right)$$

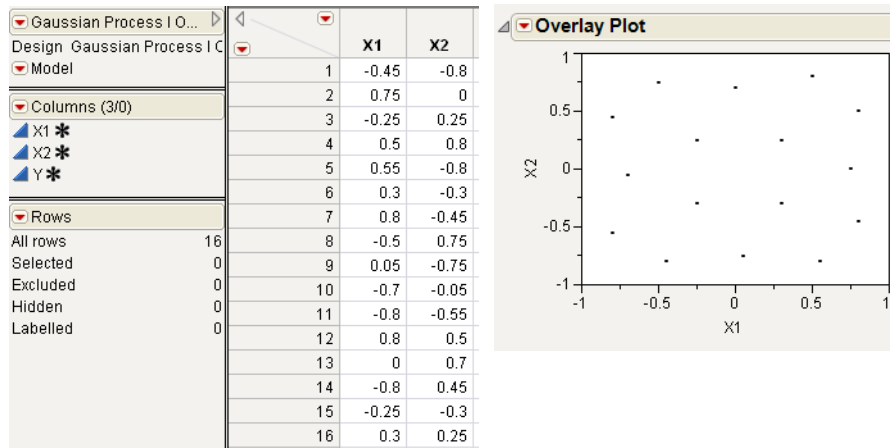
is the correlation of response values at two different design points, x_i and x_j . Computationally, these designs maximize $|R|$, the determinant of the correlation matrix of the sample. When x_i and x_j are far apart, then R_{ij} approaches zero. When x_i and x_j are close together, then R_{ij} is near one.

Gaussian Process IMSE Optimal Designs

The Gaussian process IMSE optimal design is also a competitor to the Latin Hypercube design because it minimizes the integrated mean squared error of the Gaussian process model over the experimental region.

You can compare the IMSE optimal design to the Latin Hypercube (shown previously in Figure 9.16). The table and overlay plot in Figure 9.18 show a Gaussian IMSE optimal design. You can see that the design provides uniform coverage of the factor region.

Figure 9.18 Comparison of Two-factor Latin Hypercube and Gaussian IMSE Optimal Designs



Note: Both the Maximum Entropy design and the Gaussian Process IMSE Optimal design were created using 100 random starts.

Borehole Model: A Sphere-Packing Example

Worley (1987) presented a model of the flow of water through a borehole that is drilled from the ground surface through two aquifers. The response variable y is the flow rate through the borehole in m^3/year and is determined by the following equation:

$$y = \frac{2\pi T_u (H_u - H_l)}{\ln(r/r_w) \left[1 + \frac{2LT_u}{\ln(r/r_w)r_w^2 K_w} + \frac{T_u}{T_l} \right]}$$

There are eight inputs to this model:

- r_w = radius of borehole, 0.05 to 0.15 m
- r = radius of influence, 100 to 50,000 m
- T_u = transmissivity of upper aquifer, 63,070 to 115,600 m²/year
- H_u = potentiometric head of upper aquifer, 990 to 1100 m
- T_l = transmissivity of lower aquifer, 63.1 to 116 m²/year
- H_l = potentiometric head of lower aquifer, 700 to 820 m
- L = length of borehole, 1120 to 1680 m
- K_w = hydraulic conductivity of borehole, 9855 to 12,045 m/year

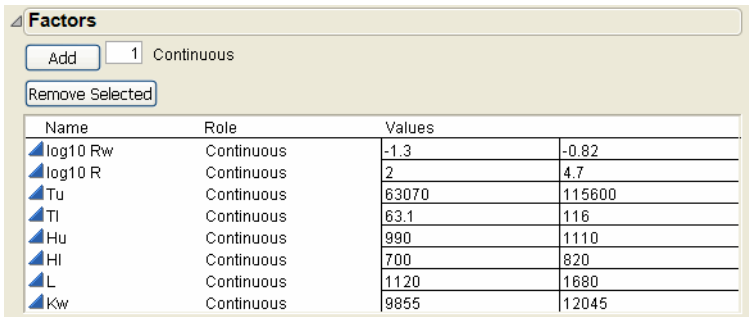
This example is atypical of most computer experiments because the response can be expressed as a simple, explicit function of the input variables. However, this simplicity is useful for explaining the design methods.

Create the Sphere-Packing Design for the Borehole Data

To create a Sphere-Packing design for the borehole problem:

1. Select **DOE > Space Filling Design**.
2. Click the red triangle icon on the **Space Filling Design** title bar and select **Load Factors**.
3. Open the Sample Data folder installed with JMP. In the Design Experiment folder, open **Borehole Factors.jmp** from the Design Experiment folder to load the factors (Figure 9.19).

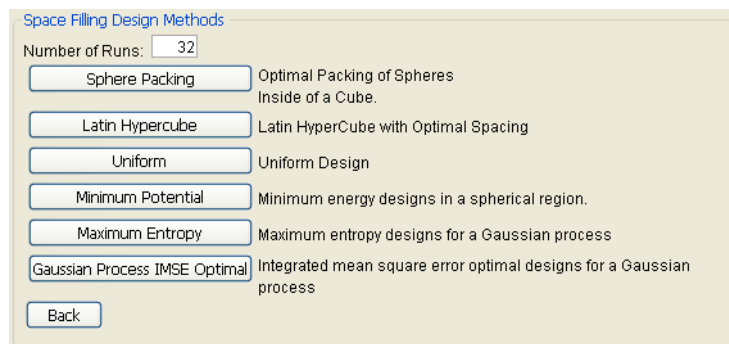
Figure 9.19 Factors Panel with Factor Values Loaded for Borehole Example



Note: The logarithm of r and r_w are used in the following discussion.

4. Click **Continue**.
5. Specify a sample size (Number of Runs) of 32 as shown in Figure 9.20.

Figure 9.20 Space-Filling Design Method Panel Showing 32 Runs



6. Click the **Sphere Packing** button to produce the design.
7. Click **Make Table** to make a table showing the design settings for the experiment. The factor settings in the example table might not have the same ones you see when generating the design because the designs are generated from a random seed.
8. To see a completed data table for this example, open *Borehole Sphere Packing.jmp* from the Design Experiment Sample Data folder installed with JMP. This table also has a table variable that contains a script to analyze the data. The results of the analysis are saved as columns in the table.

Guidelines for the Analysis of Deterministic Data

It is important to remember that deterministic data have no random component. As a result, p -values from fitted statistical models do not have their usual meanings. A large F statistic (low p -value) is an indication of an effect due to a model term. However, you cannot make valid confidence intervals about the size of the effects or about predictions made using the model.

Residuals from any model fit to deterministic data are not a measure of noise. Instead, a residual shows the model bias for the current model at the current point. Distinct patterns in the residuals indicate new terms to add to the model to reduce model bias.

Results of the Borehole Experiment

The example described in the previous sections produced the following results:

- A stepwise regression of the response, $\log y$, versus the full quadratic model in the eight factors, led to the prediction formula column.

- The prediction bias column is the difference between the true model column and the prediction formula column.
- The prediction bias is relatively small for each of the experimental points. This indicates that the model fits the data well.

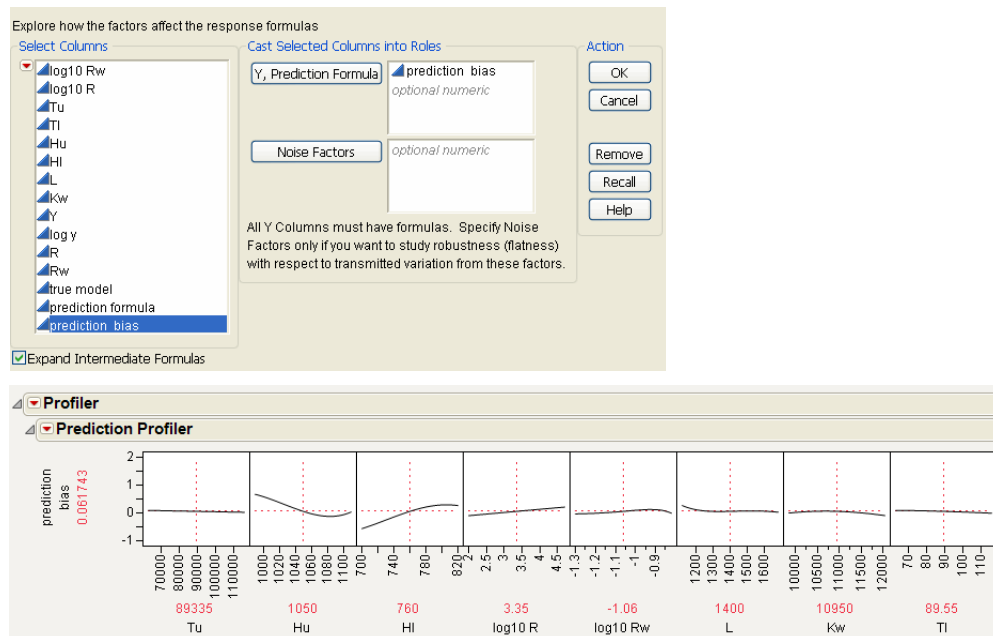
In real world examples, the true model is generally not available in a simple analytical form. As a result, it is impossible to know the prediction bias at points other than the observed data without doing additional runs.

In this case, the true model column contains a formula that allows profiling the prediction bias to find its value anywhere in the region of the data. To understand the prediction bias in this example:

1. Select **Graph > Profiler**.
2. Highlight the prediction bias column and click the **Y, Prediction Formula** button.
3. Check the **Expand Intermediate Formulas** box, as shown at the bottom on the Profiler dialog in Figure 9.21, because the prediction bias formula is a function of columns that are also created by formulas.
4. Click **OK**.

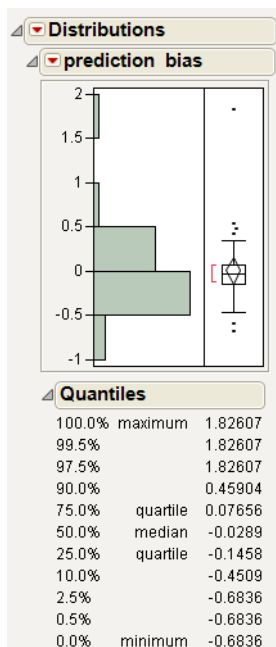
The profile plots at the bottom in Figure 9.21 show the prediction bias at the center of the design region. If there were no bias, the profile traces would be constant between the value ranges of each factor. In this example, the variables Hu and HI show nonlinear effects.

Figure 9.21 Profiler Dialog and Profile of the Prediction Bias in the Borehole Sphere-Packing Data

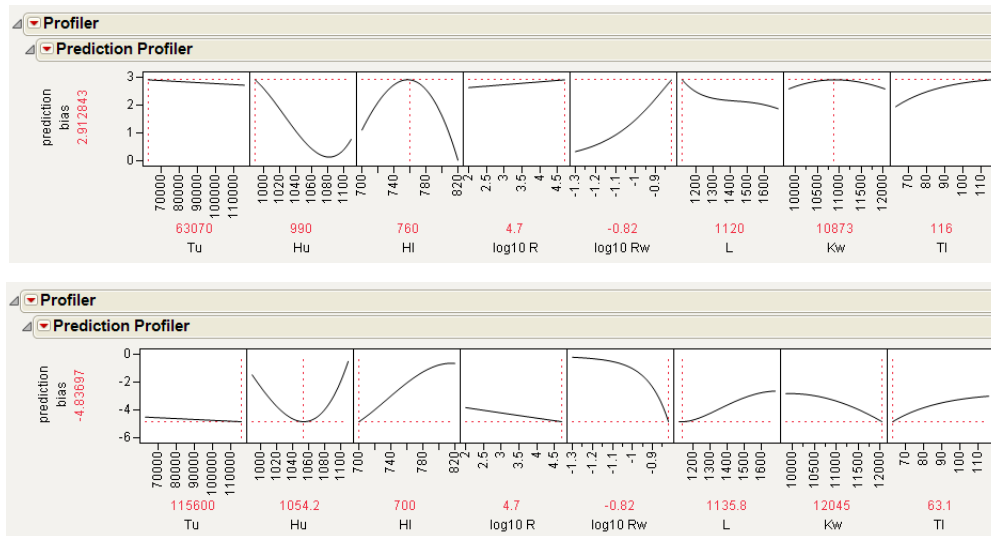


The range of the prediction bias on the data is smaller than the range of the prediction bias over the entire domain of interest. To see this, look at the distribution analysis (**Analyze > Distribution**) of the prediction bias in Figure 9.22. Note that the maximum bias is 1.826 and the minimum is -0.684 (the range is 2.51).

Figure 9.22 Distribution of the Prediction Bias



The top plot in Figure 9.23 shows the maximum bias (2.91) over the entire domain of the factors. The plot at the bottom shows the comparable minimum bias (-4.84). This gives a range of 7.75. This is more than three times the size of the range over the observed data.

Figure 9.23 Prediction Plots showing Maximum and Minimum Bias Over Factor Domains


Keep in mind that, in this example, the true model is known. In any meaningful application, the response at any factor setting is unknown. The prediction bias over the experimental data underestimates the bias throughout the design domain.

There are two ways to assess the extent of this underestimation:

- Cross-validation refits the data to the model while holding back a subset of the points and looks at the error in estimating those points.
- Verification runs (new runs performed) at different settings to assess the lack of fit of the empirical model.

Chapter 10

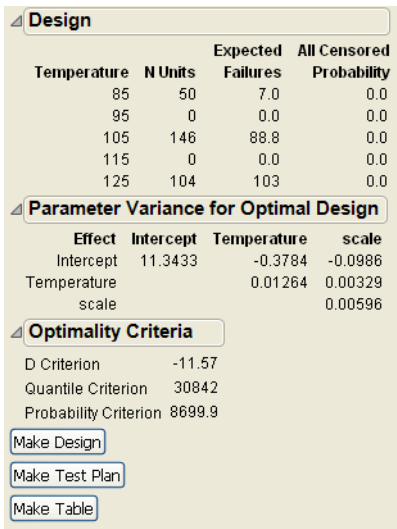
Accelerated Life Test Designs

Designing Experiments for Accelerated Life Tests



The Accelerated Life Test Design platform can be used to design experimental plans for accelerated life testing. You can design initial experiments, or augment existing experiments.

Figure 10.1 Accelerated Life Test Design



Contents

- Overview of Accelerated Life Test Designs. 219
- Using the ALT Design Platform 219
- Platform Options 224
- Example 224

Overview of Accelerated Life Test Designs

Often in reliability studies, the product reliability at use conditions is so high that the time required to test the product until it fails is prohibitive. As an alternative, you can test the product in conditions that are more extreme than normal use conditions. The extreme conditions enable the product to degrade and fail sooner, making a reliability study possible. Results are used to predict product reliability at normal use conditions.

The **Accelerated Life Testing Design** platform can be used to design experimental plans for accelerated life testing (ALT) experiments. The ALT Design platform can be used to design initial experiments, or augment existing experiments. Augmenting designs is useful if you want to obtain more data, so that you can decrease the variance associated with predicting product reliability.

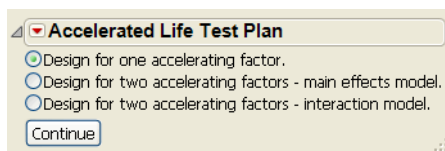
The ALT Design platform can create designs for situations involving one or two accelerating factors. For two accelerating factors, you can choose to include the interaction. You can use D-optimality or two types of I-optimality criterion.

The process requires estimates of acceleration model parameters. Since those parameters might not be known in advance, you can specify prior distributions to account for the uncertainty. Designs can be created for either Lognormal or Weibull life distributions.

Using the ALT Design Platform

1. To launch the ALT Design platform, select **DOE > Accelerated Life Test Design**.

Figure 10.2 Initial ALT Design Window



2. Select one or two accelerating factors.
For two factors, you can choose to include the interaction between the factors.
3. Click **Continue**.
A window appears for specifying details of the accelerating factor or factors.

Figure 10.3 Accelerating Factor Details Window

The screenshot shows a software window titled "Accelerated Life Test Plan". Inside, there is a table with five columns: "Factor Name", "Number of Levels", "Factor Transformation", "Low Usage Condition", and "High Usage Condition". The first row of data shows "X1" in the Factor Name column, "3" in the Number of Levels column, "Arrhenius Celsius" in the Factor Transformation column (with a dropdown arrow), "0" in the Low Usage Condition column, and "100" in the High Usage Condition column. Below the table is a "Continue" button.

Factor Name	Number of Levels	Factor Transformation	Low Usage Condition	High Usage Condition
X1	3	Arrhenius Celsius	0	100

Continue

4. Fill in these options:
- Factor Name** Enter a name for the accelerating factor.
 - Number of Levels** Enter the number of levels of the factor to include in the experiment.
 - Factor Transformation** Select a transformation for the factor. The options are Arrhenius Celsius, Reciprocal, Log, Square Root, and Linear.
 - Low Usage Condition** Enter a value for the low usage conditions.
 - High Usage Conditions** Enter a value for the high usage conditions.
5. Click **Continue**.
- A window appears for specifying additional information about the assumed distribution and desired experimental conditions.

Figure 10.4 Distribution Details

Accelerated Life Test Plan

Factors

X1 Level
Values
1
2
3

Prior Specification

Distribution Choice
☐ LogNormal
☒ Weibull

Prior Mean

Effect	Prior Mean
Intercept	4.00000
X1	-1.00000
scale	1.00000

☒ Ignore prior variance. Generate the local design for the prior mean.

Prior Variance Matrix

Effect	Intercept	X1	scale
Intercept	0.10000	0.00000	0.00000
X1		0.10000	0.00000
scale			0.10000

Diagnostic Choices

Time range of interest: 10000 20000
 Probability of interest: 0.1

Design Choices

Length of Test 10
 Number of Units Under Test 15

6. Fill in these options:

Factors Enter the values of the accelerating factor levels.

Distribution Choice Select the failure distribution, either Weibull or LogNormal.

Prior Mean Enter the acceleration model parameters. These values can be either a best guess or the estimates of the parameters from an analysis of previously obtained data.

Prior Variance Matrix Enter variances and covariances for the acceleration model parameters. These values can be either a best guess or the estimated variance matrix from an analysis of previously obtained data.

Ignore prior variance Select this option to ignore variances and covariances for model parameters. When the variances and covariances are ignored, the design is created for the specific fixed parameters entered under Prior Mean. The resulting design is said to be a locally optimal design. This design will be good if the prior mean parameters are close to the true values. However, the design will not be robust to mis-specified parameters. When the variances and covariances are used, a multivariate normal prior distribution is assumed for the acceleration model parameters. This is useful when the values entered under Prior Mean are estimates or guesses, and you want the design to reflect the associated uncertainty.

Diagnostic Choices Enter values for the following:

Time range of interest are values for which you want an estimated probability of failure. For example, if you are interested in the probability of failure by 100,000 hours, then enter 100,000 for both the lower and upper ranges.

Probability of interest is the value for which you want an estimated time of failure. For example, if you are most interested in obtaining the time until 10% of units fail, then enter 0.1.

Design Choices Enter values for the following:

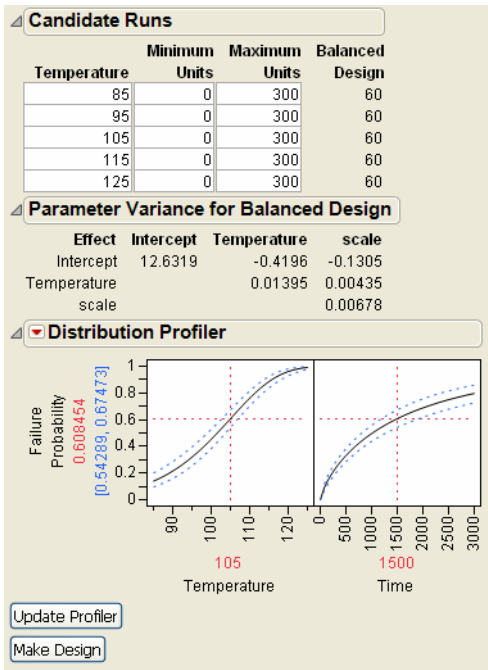
Length of Test is the length of time to run the experiment.

Number of Units Under Test is the number of units in the experiment. If augmenting a previous experiment, enter the number of units from the previous experiment plus the number of units that you want to run for the next experiment. If designing an initial experiment, enter the number of units that you want to run.

7. Click **Continue**.

New outline nodes appear as shown in Figure 10.5.

Figure 10.5 Additional Outline Nodes



Candidate Runs Enter the minimum and maximum number of runs allowed at each level of the acceleration factor. If augmenting a previous experiment, enter the number of units already run at each level for the Minimum Units.

Parameter Variance for Balanced Design Gives the expected variances and covariances for the acceleration model parameters after running and analyzing the balanced design (same number of runs at each level of the acceleration factor). These values are valid under the assumption that the values for the prior mean and variance are correct.

These values can be compared to the ones entered under Prior Variance Matrix, to see whether the balanced design can improve the variances of the parameter estimates. They can also be compared to the final parameter variances for the optimal design, after clicking **Make Design**.

Distribution Profiler Use the profiler to visualize the probability that a unit will fail at different values of the acceleration factor and time.

8. Click **Update Profiler** to update the profiler if changes are made to the distribution choice, means, variances, design choices, or candidate runs.
9. Click **Make Design** to create the optimal design and display the results.

Figure 10.6 Design Results

Design				
	X1	N Units	Expected Failures	All Censored Probability
	1	8	8.0	0.0
	2	0	0.0	0.0
	3	7	7.0	0.0

Parameter Variance for Optimal Design				
Effect	Intercept	X1	scale	
Intercept	5072.23	-120.23	-0.0171	
X1		2.84983	0.00000	
scale			0.04053	

Optimality Criteria	
D Criterion	-4.867
Quantile Criterion	132.54
Probability Criterion	203.05

Make Design

Make Test Plan

Make Table

The information below describes the results you get after clicking **Make Design**.

The **Design** report gives the expected number of failures for each level of the acceleration factor. Also given is the probability that none of the units at this setting will fail.

The **Parameter Variance for Optimal Design** report gives the variances and covariances for the acceleration model parameters for the optimal design. These values are valid under the assumption that the values for the prior mean and variance are correct. These values can be compared to those under Parameter Variance for Balanced Design to determine whether the optimal design is able to reduce the parameter variances more than the balanced design.

The **Optimality Criteria** report gives the values of the optimality criterion for the optimal design. For more information about the optimality criterion, see “Platform Options,” p. 224.

The **Make Design** button updates the optimal design if any changes are made to the distribution choice, prior means or variances, design choices, or candidate runs.

The **Make Test Plan** button creates a data table with the acceleration factor levels and the number of units to include in the experiment for each level.

The **Make Table** button creates a table that can be used for data collection during the experiment.

Platform Options

The red triangle menu for Accelerated Life Test Plan has the following options:

Simulate Responses Adds simulated responses to the table when you click **Make Table**.

ALT Optimality Criterion Gives three choices for design optimality:

Make D-Optimal Design creates a design that minimizes the variance of the model coefficients.

Make Time I-Optimal Design creates a design that minimizes the prediction variance when predicting the time to failure for the probability given in [“Diagnostic Choices,” p. 222](#).

Make Probability I-Optimal Design creates a design that minimizes the prediction variance when predicting the failure probability for the times given in [“Diagnostic Choices,” p. 222](#).

Advanced Options Gives the N Monte Carlo Spheres option, which affects the speed and accuracy of numerical integration. For more information, see [“Advanced Options for the Nonlinear Designer,” p. 245](#) in the “Nonlinear Designs” chapter.

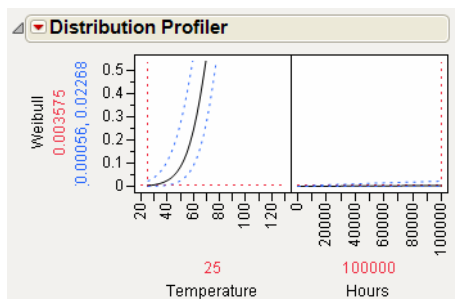
Example

This example shows how to use the Accelerated Life Test Design to augment an existing design.

An accelerated life test was performed, and the results are in the **Capacitor ALT.jmp** sample data table (in the **Design Experiment** folder). Fifty units were tested at each of three temperatures (85°, 105°, and 125° Celsius) for 1500 hours. The resulting model is used to predict the probability of failure at 100,000 hours at normal use conditions of 25°.

1. Open the **Capacitor ALT.jmp** data table in the **Design Experiment** folder.
2. Run the **Fit Life by X** table script.
3. In the **Distribution Profiler**, enter 25 for **Temperature** and 100,000 for **hours**.

The profiler is shown in Figure 10.7.

Figure 10.7 Distribution Profiler for Capacitor Model

The predicted probability of failure at 100,000 hours at 25° is 0.00358, with a confidence interval of 0.00056 to 0.0227. The analyst wants to decrease the width of the confidence interval. To do so, the experiment needs to be augmented with additional data.

To augment the design in the optimal way, use the Accelerated Life Test Design platform. Follow the steps below to use the platform:

1. Select **DOE > Accelerated Life Test Design**.
2. Select **Design for one accelerating factor** and click **Continue**.
3. Enter Temperature for **Factor Name**.
4. Enter 5 for **Number of Levels**.
5. Enter 25 for both **Low Usage Condition** and **High Usage Condition**.
6. Click **Continue**.
7. Enter 85, 95, 105, 115, and 125 for the **Temperature Level Values**.
8. Select Weibull for **Distribution Choice**.
9. Under **Prior Mean**, enter the acceleration model parameters from the Fit Life by X Estimates report. See Figure 10.8.
 - Enter –35.200 for Intercept.
 - Enter 1.389 for Temperature.
 - Enter 1.305 for scale.

Figure 10.8 Fit Life by X Estimates

Estimates				
Parameter	Estimate	Std Error	Lower 95%	Upper 95%
β_0	-35.19979	4.6912686	-44.39451	-26.00508
β_1	1.38891	0.1566297	1.08192	1.69589
σ	1.30471	0.1119128	1.08536	1.52405

In the Fit Life by X output, under the Estimates report, is the Covariance Matrix report. Note the variances for the acceleration model parameters are 22, 0.025, and 0.013. These values will be compared to our final results at the end.

10. Enter 100,000 for both boxes for **Time range of interest**.

11. Enter 1500 for **Length of Test**.

12. Enter 300 for **Number of Units Under Test**. The previous experiment used 150 units, and the next experiment uses 150 units, for a total of 300.

The completed window is shown in Figure 10.9.

Figure 10.9 Completed Window

Accelerated Life Test Plan

Factors

Temperature
Level Values

85
95
105
115
125

Prior Specification

Distribution Choice

☐ LogNormal
☒ Weibull

Prior Mean

Effect	Prior Mean
Intercept	-35.200
Temperature	1.38900
scale	1.30500

☒ Ignore prior variance. Generate the local design for the prior mean.

Prior Variance Matrix

Effect	Intercept	Temperature	scale
Intercept	0.10000	0.00000	0.00000
Temperature		0.10000	0.00000
scale			0.10000

Diagnostic Choices

Time range of interest:

Probability of interest:

Design Choices

Length of Test:

Number of Units Under Test:

13. Click **Continue**.

14. To account for the units in the previous experiment, enter the following under **Candidate Runs**.

- Enter 50 for Minimum Units for 85°.
- Enter 50 for Minimum Units for 105°.

- Enter 50 for Minimum Units for 125°.
- 15. From the red triangle menu for Accelerated Life Test Plan, select **ALT Optimality Criterion > Make Probability I-Optimal Design**.
- 16. Click **Make Design**.

The optimal experimental design is returned, along with other results. See Figure 10.10.

Figure 10.10 Optimal Design

Design			
Temperature	N Units	Expected Failures	All Censored Probability
85	184	25.3	0.0
95	0	0.0	0.0
105	50	30.0	0.0
115	0	0.0	0.0
125	66	65.4	0.0

Parameter Variance for Optimal Design			
Effect	Intercept	Temperature	scale
Intercept	11.9646	-0.3974	-0.2060
Temperature		0.01322	0.00691
scale			0.00994

Optimality Criteria	
D Criterion	-13.65
Quantile Criterion	1.0636
Probability Criterion	0.386

The optimal design is computed based on the Number of Units Under Test, Candidate Runs, and other information that you specified earlier. The optimal design consists of the following number of units at each temperature level:

- 184 units at 85°. Since the previous experiment used 50 units, 134 additional units are needed.
- 0 units at 95°. The next experiment will not utilize any units at this level.
- 50 units at 105°. Since the previous experiment already used 50 units, no additional units are needed.
- 0 units at 115°. The next experiment will not utilize any units at this level.
- 66 units at 125°. Since the previous experiment used 50 units, 16 additional units are needed.

As we entered earlier, a total of $134+16=150$ units are used for the new experiment.

An estimate of the acceleration model parameters variances is given. Note that, due to the additional data, all three variances are smaller than before from the original Fit Life by X report.

In the Profiler, enter 25 for Temperature and 100,000 for Time. The estimated probability of failure is 0.00357, with an estimated confidence interval of 0.00106 to 0.01201. This interval is narrower than the one from the previous experiment, as a result of the additional units to be tested.

To decrease the interval further, try entering more than 300 units to be tested.

Chapter 11

Nonlinear Designs



Design of experiments with models that are nonlinear in their parameters is available using either the DOE menu or the JMP Starter DOE category.

Nonlinear designs offer both advantages and disadvantages compared to designs for linear models.

On the positive side, predictions using a well chosen model are likely to be good over a wider range of factor settings. It is also possible to model response surfaces with more curvature and with asymptotic behavior.

On the negative side, the researcher needs a greater understanding of both the system and of the nonlinear design tool.

Contents

- Examples of Nonlinear Designs 231
 - Using Nonlinear Fit to Find Prior Parameter Estimates 231
 - Creating a Nonlinear Design with No Prior Data 237
- Creating a Nonlinear Design 241
 - Identify the Response and Factor Column with Formula. 241
 - Set Up Factors and Parameters in the Nonlinear Design Dialog. 242
 - Enter the Number of Runs and Preview the Design. 243
 - Make Table or Augment the Table. 244
- Advanced Options for the Nonlinear Designer 245

Examples of Nonlinear Designs

The Nonlinear Designer allows scientists to generate optimal designs and optimally augment data for fitting models that are nonlinear in their parameters. Such models, when they are descriptive of the underlying process, can yield more accurate prediction of process behavior than is possible with the standard polynomial models.

To use the Nonlinear Designer, you first need a data table that has

- one column for each factor
- one column for the response
- a column that contains a formula showing the functional relationship between the factor(s) and the response.

This is the same format for a table you would supply to the nonlinear platform for modeling.

The first example in this section describes how to approach creating a nonlinear design when there is prior data. The second example describes how to approach creating the design without data, but with reasonable starting values for the parameter estimates.

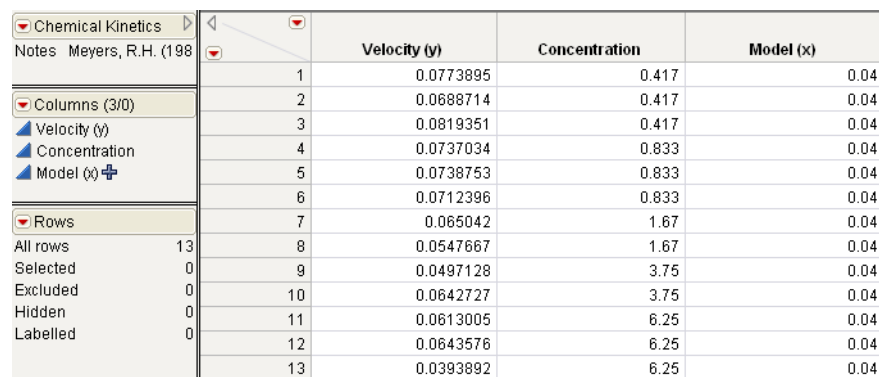
Using Nonlinear Fit to Find Prior Parameter Estimates

Suppose you have already collected experimental data and placed it in a JMP data table. That table can be used to create a nonlinear design for improving the estimates of the model's parameters.

To follow along with this example, open *Chemical Kinetics.jmp* from the Nonlinear Examples folder found in the sample data installed with JMP.

Chemical Kinetics.jmp (Figure 11.1) contains a column (Model (x)) whose values are formed by a formula with a poor guess of the parameter values.

Figure 11.1 Chemical Kinetics.jmp

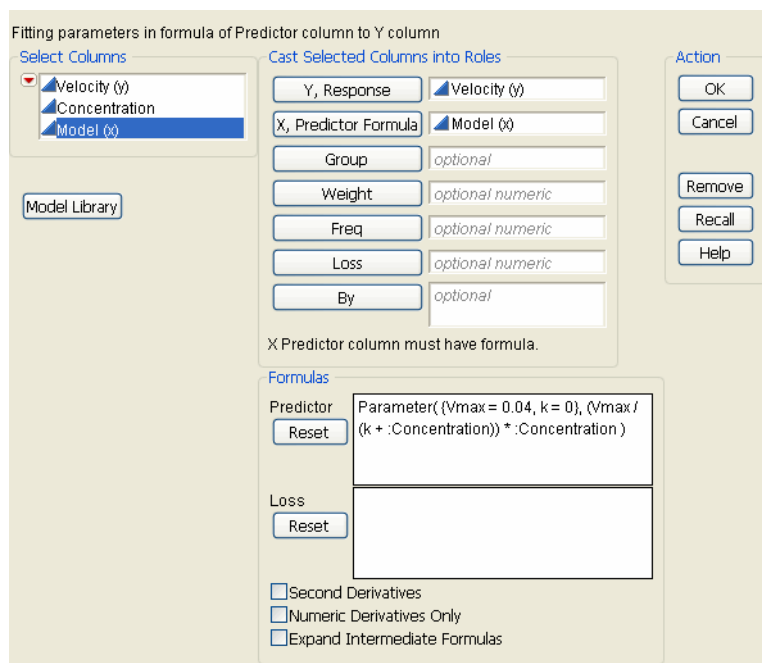


	Velocity (y)	Concentration	Model (x)
1	0.0773895	0.417	0.04
2	0.0688714	0.417	0.04
3	0.0819351	0.417	0.04
4	0.0737034	0.833	0.04
5	0.0738753	0.833	0.04
6	0.0712396	0.833	0.04
7	0.065042	1.67	0.04
8	0.0547667	1.67	0.04
9	0.0497128	3.75	0.04
10	0.0642727	3.75	0.04
11	0.0613005	6.25	0.04
12	0.0643576	6.25	0.04
13	0.0393892	6.25	0.04

First, fit the data to the model using nonlinear least squares to get better parameter values.

1. Select **Analyze > Modeling > Nonlinear**.
2. Select **Velocity (y)** and click **Y, Response** on the nonlinear launch dialog.
3. Select **Model (x)** and click **X, Predictor Formula** (see Figure 11.2). Note that the formula given by **Model (X)** shows in the launch dialog.

Figure 11.2 Initial Nonlinear Analysis Launch Dialog

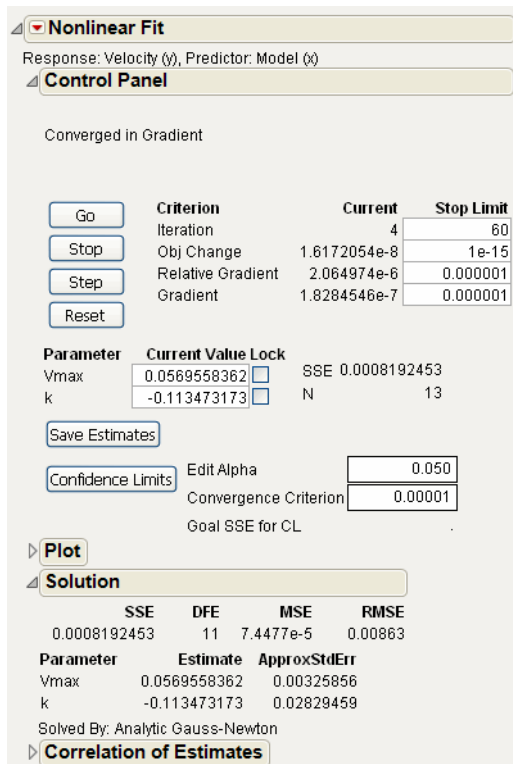


4. Click **OK** on the launch dialog to see the Nonlinear control panel.
5. Click the **Go** button on the Nonlinear control panel to see the results (Figure 11.3).

The **Save Estimates** button adds the new fitted parameter values in the **Model (x)** column in the Chemical Kinetics.jmp data table.

The **Confidence Limits** button computes confidence intervals used to create a nonlinear design.

The ranges for LowerCL and UpperCL are the intervals for Vmax and k. They are asymptotically normal. Use these limits to create a nonlinear design in JMP.

Figure 11.3 Nonlinear Fit Results

- Click the **Confidence Limits** button to produce confidence intervals.
- Click **Save Estimates** to add the new fitted parameter values in the **Model (x)** column in the Chemical Kinetics.jmp data table, which contains the formula:

Figure 11.4 Formula Produced by Save Estimates

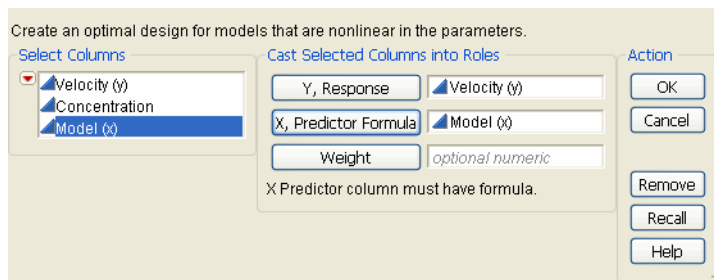
$$\frac{V_{\max}}{(k + \text{Concentration})} + \text{Concentration}$$

Note: Leave the nonlinear analysis report open because these results are needed in the DOE nonlinear design dialog described next.

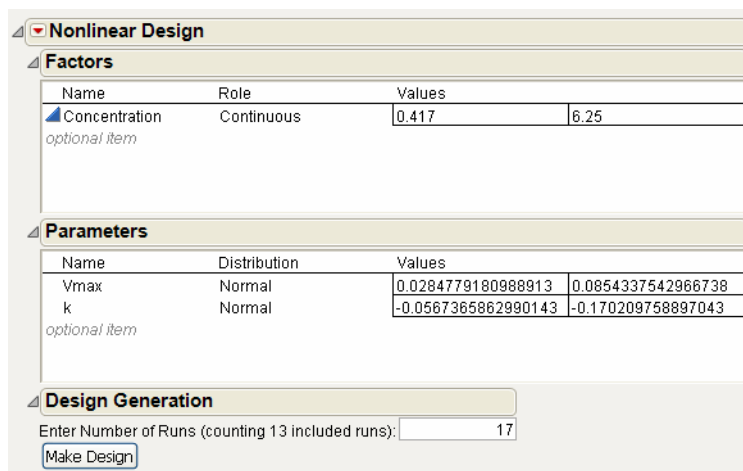
Now create a design for fitting the model's nonlinear parameters.

- With the Chemical Kinetics.jmp data table open, select **DOE > Nonlinear Design**.

- Complete the launch dialog the same way as the Nonlinear Analysis launch dialog shown previously. That is, Select **Velocity (y)** and click **Y, Response**. Select **Model (x)** and click the **X, Predictor Formula**. Figure 11.5 shows the completed dialog.

Figure 11.5 Initial Nonlinear Design Launch Dialog


- Click **OK** to see the completed Design panels for factors and parameters, as shown in Figure 11.6.

Figure 11.6 Nonlinear Design Panels for Factors and Parameters


Note that in Chemical Kinetics.jmp (Figure 11.1), the range of data for Concentration goes from 0.417 to 6.25. Therefore, these values initially appear as the high and low values in the Factors control panel as follows:

- Change the factor range for **Concentration** to a broader interval—from 0.1 to 7 (Figure 11.7).

Note that the *a priori* distribution of the parameters Vmax and k is **Normal**, which is correct for this example. Change the current level of uncertainty in the two parameters using the analysis results.

- Look back at the analysis report in Figure 11.3 and locate the upper and lower confidence limits for V_{\max} and k in the Solution table. Change the values for V_{\max} and k to correspond to those limits, as shown in Figure 11.7.

Now you have described the current level of uncertainty of the two parameters.

Figure 11.7 Change Values for Factor and Parameters

The screenshot shows the 'Nonlinear Design' window. It has three main sections: 'Factors', 'Parameters', and 'Design Generation'.

Factors Table:

Name	Role	Values
Concentration <i>optional item</i>	Continuous	0.1 7

Parameters Table:

Name	Distribution	Values
V_{\max}	Normal	0.05 0.064
k <i>optional item</i>	Normal	-0.167 -0.044

Design Generation Section:

Enter Number of Runs (counting 13 included runs):

- If necessary, type the desired number of runs (17 is the default value for this example) into the text box. Use commands from the menu on the Nonlinear Design title bar to get the best possible design:
- Select **Number of Starts** from the menu on the title bar and enter 100 in the text box.
- Select **Advanced Options > Number of Monte Carlo Samples** and enter 2 in the text box.
- Click **Make Design** to preview the design (Figure 11.8). Your results might differ from those shown for the additional runs.

Figure 11.8 Selecting the Number of Runs

Design		
Run	Concentration	Velocity (y)
1	0.417	0.07739
2	0.417	0.068871
3	0.417	0.081935
4	0.833	0.073703
5	0.833	0.073875
6	0.833	0.07124
7	1.67	0.065042
8	1.67	0.054767
9	3.75	0.049713
10	3.75	0.064273
11	6.25	0.061301
12	6.25	0.064358
13	6.25	0.039389
14	0.10551	.
15	0.205311	.
16	0.10551	.
17	0.10551	.

10. Click **Make Table**.

This creates a new JMP design table (Figure 11.9) whose rows are the runs defined by the nonlinear design.

Note: This example creates a new table to avoid altering the sample data table Chemical Kinetics.jmp. In most cases, however, you can augment the original table using the **Augment Table** option in the Nonlinear Designer instead of making a new table. This option adds the new runs shown in the Design to the existing data table.

Figure 11.9 Making a Table with the Nonlinear Designer

Output Options

Make Table

Augment Table

Back

Nonlinear Design

Design Nonlinear Design

Model

Columns (3/0)

Concentration *

Velocity (y) *

Model (x) +

Rows

All rows 17

Selected 0

Excluded 0

Hidden 0

Labelled 0

	Concentration	Velocity (y)	Model (x)
1	0.417	0.0773895	0.07824871
2	0.417	0.0688714	0.07824871
3	0.417	0.0819351	0.07824871
4	0.833	0.0737034	0.06593807
5	0.833	0.0738753	0.06593807
6	0.833	0.0712396	0.06593807
7	1.67	0.065042	0.061108
8	1.67	0.0547667	0.061108
9	3.75	0.0497128	0.05873307
10	3.75	0.0642727	0.05873307
11	6.25	0.0613005	0.05800903
12	6.25	0.0643576	0.05800903
13	6.25	0.0393892	0.05800903
14	0.1055099625	•	-0.7546464
15	0.205310643	•	0.12732972
16	0.1055099625	•	-0.7546464
17	0.1055099625	•	-0.7546464

The new runs use the wider interval of allowed concentration, which leads to more precise estimates of k and V_{\max} .

Creating a Nonlinear Design with No Prior Data

This next example describes how to create a design when you have not yet collected data, but have a guess for the unknown parameters.

To follow along with this example, open *Reaction Kinetics Start.jmp*, found in the *Design Experiment* folder in the sample data installed with JMP. Notice that the table is a template. That is, the table has columns with properties and formulas, but there are no observations in the table. The design has not yet been created and data have not been collected.

This table is used to supply the formula in the *yield model* column to the Nonlinear DOE platform. The formula is used to create a nonlinear design for fitting the model's nonlinear parameters.

Figure 11.10 Yield Model Formula

$$100 * \frac{\text{Exp}\left[\frac{t1 - 1.9279 * 10^4}{\text{Reaction Temperature}}\right] * \left[\text{Exp}\left[-\text{Exp}\left[\frac{t1 - 1.9279 * 10^4}{\text{Reaction Temperature}}\right] * \text{Reaction Time}\right] - \text{Exp}\left[-\text{Exp}\left[\frac{t3 - 1.6819 * 10^4}{\text{Reaction Temperature}}\right] * \text{Reaction Time}\right]\right]}{\left[\text{Exp}\left[\frac{t3 - 1.6819 * 10^4}{\text{Reaction Temperature}}\right] - \text{Exp}\left[\frac{t1 - 1.9279 * 10^4}{\text{Reaction Temperature}}\right]\right]}$$

This model is from Box and Draper (1987). The formula arises from the fractional yield of the intermediate product in a consecutive chemical reaction. It is written as a function of time and temperature.

1. With the *Reaction Kinetics Start.jmp* data table open, select **DOE > Nonlinear Design** to see the initial launch dialog.
2. Select **observed yield** and click **Y, Response**.
3. Select **yield model** (the column with the formula) and click **X, Predictor Formula**.

The completed dialog should look like the one in Figure 11.11.

Figure 11.11 Nonlinear Design launch Dialog

Create an optimal design for models that are nonlinear in the parameters.

Select Columns

- Reaction Time
- Reaction Temperature
- observed yield
- yield model

Cast Selected Columns into Roles

Y, Response: observed yield

X, Predictor Formula: yield model

Weight: optional numeric

X Predictor column must have formula.

Action

OK Cancel Remove Recall Help

- Click **OK** to see the nonlinear design Factors and Parameters panels in Figure 11.12.
- Change the two factors' values to be a reasonable range of values. (In your experiment, these values might have to be an educated guess.) For this example, use the values 510 and 540 for **Reaction Temperature**. Use the values 0.1 and 0.3 for **Reaction Time**.
- Change the values of the parameter **t1** to 25 and 50, and **t3** to 30 and 35.
- Click on the Distribution of each parameter and select **Uniform** from the menu to change the distribution from the default **Normal** (see Figure 11.12).
- Change the number of runs to 12 in the Design Generation panel.

Figure 11.12 Change Factor Values, Parameter Distributions, and Number of Runs

Nonlinear Design

Factors

Name	Role	Values
Reaction Temperature	Continuous	510 540
Reaction Time	Continuous	0.1 0.3

optional item

Parameters

Name	Distribution	Values
t1	Uniform	25 50
t3	Normal	30 35

optional item

Design Generation

Number of Runs: 12

Make Design

- Click **Make Design**, then **Make Table**. Your results should look similar to those in Figure 11.13.

Figure 11.13 Design Table

Nonlinear Design					
Design	Nonlinear Design				
Model					
Columns (4/0)					
Reaction Temperature *					
Reaction Time *					
observed yield *					
yield model					
Rows					
All rows	12				
Selected	0				
Excluded	0				
Hidden	0				
Labelled	0				

		Reaction Temperature	Reaction Time	observed yield	yield model
1		540	0.107270436	•	56.745725
2		540	0.3	•	58.093235
3		540	0.3	•	58.093235
4		510	0.1	•	11.2590804
5		540	0.3	•	58.093235
6		510	0.1	•	11.2590804
7		526.2712314	0.1	•	30.7102877
8		540	0.3	•	58.093235
9		540	0.107270436	•	56.745725
10		526.2712314	0.1	•	30.7102877
11		540	0.107270436	•	56.745725
12		540	0.107270436	•	56.745725

10. To analyze data that contains values for the response, observed yield, open Reaction Kinetics.jmp from the Design Experiment folder in the sample data installed with JMP (Figure 11.14).

Figure 11.14 Reaction Kinetics.jmp

Reaction Kinetics					
Design	Nonlinear Design				
Model					
Columns (4/0)					
Reaction Temperature *					
Reaction Time *					
observed yield *					
yield model					
Rows					
All rows	12				
Selected	0				
Excluded	0				
Hidden	0				
Labelled	0				

		Reaction Temperature	Reaction Time	observed yield	yield model
1		540	0.3	57	58.093235
2		540	0.11	56	57.3016426
3		540	0.3	59	58.093235
4		530	0.2	53	54.9142081
5		510	0.1	12	11.2590804
6		540	0.3	57	58.093235
7		540	0.3	59	58.093235
8		510	0.1	10	11.2590804
9		517	0.1	18	17.8546023
10		530	0.21	56	55.9838913
11		540	0.11	56	57.3016426
12		540	0.3	59	58.093235

First, examine the design region with an overlay plot.

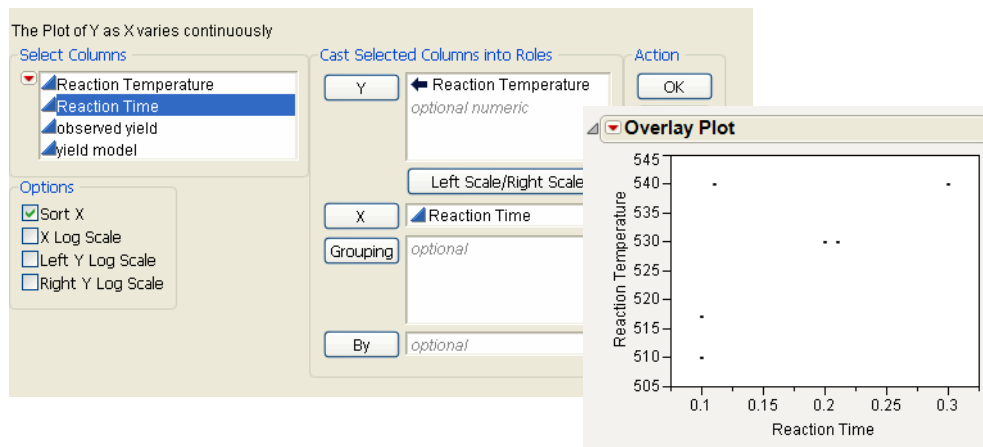
11. Select **Graph > Overlay Plot**.

12. Remove all existing column assignments.

13. Select Reaction Temperature and click **Y**

14. Select Reaction Time and click **X** as shown in the Overlay Plot launch dialog in Figure 11.15.

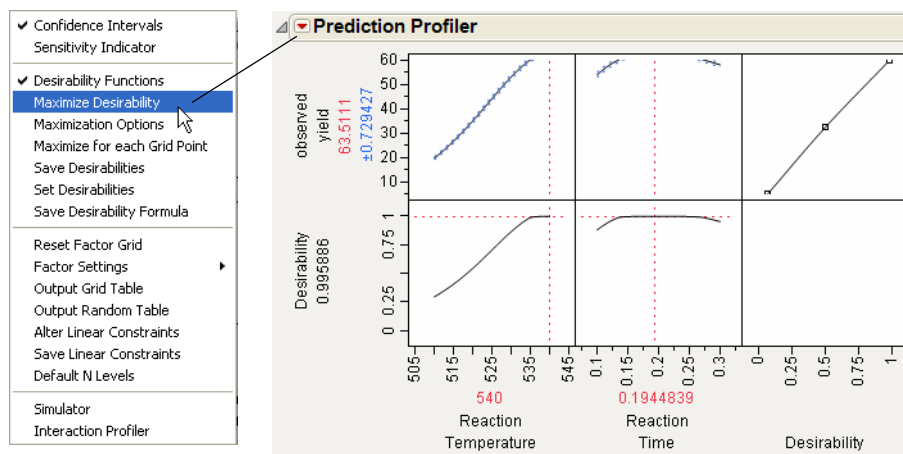
15. Click **OK** to see the overlay plot in Figure 11.15.

Figure 11.15 Create an Overlay Plot

Notice that the points are not at the corners of the design region. In particular, there are no points at low temperature and high time—the lower right corner of the graph.

16. Select **Analyze > Modeling > Nonlinear**.
17. Remove all existing column assignments.
18. Select **observed yield** and click **Y, Response**.
19. Select **yield model** and click the **X, Predictor Formula**, then click **OK**.
20. Click **Go** on the Nonlinear control panel.
21. Now, choose **Profilers > Profiler** from the red triangle menu on the Nonlinear Fit title bar.
22. To maximize the yield, choose **Maximize Desirability** from the red triangle menu on the Prediction Profiler title bar.

The maximum yield is approximately 63.5% at a reaction temperature of 540 degrees Kelvin and a reaction time of 0.1945 minutes.

Figure 11.16 Time and Temperature Settings for Maximum Yield

Creating a Nonlinear Design

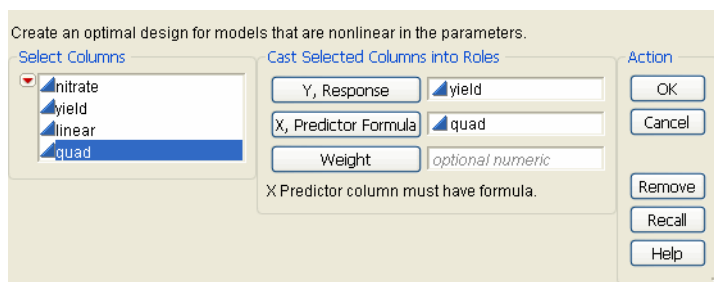
To begin, open a data table that has a column whose values are formed by a formula (for details about formulas, see the *Using JMP*). This formula must have parameters.

Select **DOE > Nonlinear Design**, or click the **Nonlinear Design** button on the JMP Starter **DOE** page. Then, follow the steps below:

- “Identify the Response and Factor Column with Formula,” p. 241
- “Set Up Factors and Parameters in the Nonlinear Design Dialog,” p. 242
- “Enter the Number of Runs and Preview the Design,” p. 243
- “Make Table or Augment the Table,” p. 244

Identify the Response and Factor Column with Formula

1. Open a data table that contains a column whose values are formed by a formula that has parameters. This example uses *Corn.jmp* from the Nonlinear Examples folder in the sample data installed with JMP.
2. Select **DOE > Nonlinear Design** to see the initial launch dialog.
3. Select **yield** and click **Y, Response**. The response column cannot have missing values.
4. Select **quad** and click **X, Predictor Formula**. The **quad** variable has a formula that includes **nitrate** and three parameters (Figure 11.17).
5. Click **OK** on the launch dialog to see the Nonlinear Design DOE panels.

Figure 11.17 Identify Response (Y) and the Column with the Nonlinear Formula (X)


Set Up Factors and Parameters in the Nonlinear Design Dialog

First, look at the formula for quad, shown in Figure 11.18, and notice there are three parameters. These parameters show in the Parameters panel of the Nonlinear design dialog, with initial parameter values.

Figure 11.18 Formula for quad has Parameters a, b, and c

$$\text{fmZ} = \begin{cases} \text{nitrate} - 29.45 < \frac{b}{2 * c} \Rightarrow a + b * (\text{nitrate} - 29.45) + c * (\text{nitrate} - 29.45)^2 \\ \text{else} \Rightarrow a - \frac{b^2}{4 * c} \end{cases}$$

Use Figure 11.19 to understand how to set up factor and parameter names and values.

- The initial values for the factor and the parameters are reasonable and do not need to be changed.
- If necessary, change the Distribution of the parameters to Uniform, as shown in Figure 11.19.

Figure 11.19 Example of Setting Up Factors and Parameters

Double-click to edit the factor or parameter name.

Click to enter or change factor values.

Nonlinear Design

Factors

Name	Role	Values
nitrate	Continuous	6.63 104.53

optional item

Parameters

Name	Distribution	Values
a	Uniform	2500 7500
b	Uniform	50 150
c	Uniform	-0.75 -2.25

optional item

Design Generation

Enter Number of Runs (counting 144 included runs): 150

Make Design

Click to edit the distribution: Uniform, Normal, Lognormal, or Exponential.

Enter the Number of Runs and Preview the Design

1. The Design Generation panel shows 150 as the default number of runs. This number includes observations in the current data. 147 runs are desired. Since there are originally 144 rows, 3 additional runs need to be added. Enter 147 in the **Number of Runs** edit box.
2. Click **Make Design** before creating the data table to preview the design. Figure 11.20 shows a partial listing of the design.

Figure 11.20 Example Preview Design

Nonlinear Design					
Factors					
Parameters					
Design					
Run	nitrate	yield			
1	9.59	4713	130	24.64	9405.75
2	7.51	1049.45	131	23.08	6278.25
3	8	1537.89	132	44.2	8368.39
4	13.93	5116.24	133	63.26	7379.55
5	29.1	8835.98	134	46	8120.8
6	32.91	9849.72	135	16.56	4771.33
7	52.68	9588.95	136	25.45	6980.81
8	64.24	9474.68	137	41.53	8616.87
9	23.41	6650.39	138	61.18	9250.9
10	20.09	7783.35	139	82.49	7967.53
11	38.36	9485.95	140	13.35	5158.1
12	47.4	9138.61	141	33.33	7627.44
13	68.23	8079.86	142	29.45	7863.4
14	24.58	2843.44	143	81.61	7780.94
15	18.56	6237.16	144	73.03	8151.96
16	26.32	7961.7	145	104.53	
17	30.98	8691.12	146	104.53	
18	26.4	8906.83	147	104.53	

Make Table or Augment the Table

- The last step is to click either **Make Table** or **Augment Table**. The **Make Table** command creates a new table (Figure 11.21) with all runs included. The **Augment Table** command adds the new runs to the existing table.

Figure 11.21 Partial Listing of an Example Nonlinear Design Table

Nonlinear Design				
Design	Nonlinear Design		nitrate	yield
Model				quad
Columns (3/0)				
nitrate *				
yield *				
quad +				
Rows				
All rows	147			
Selected	0			
Excluded	0			
Hidden	0			
Labelled	0			

1	9.59	4713	2422.3706
2	7.51	1049.45	2083.9546
3	8	1537.89	2164.84625
4	13.93	5116.24	3086.6944
5	29.1	8835.98	4964.81625
6	32.91	9849.72	5328.0426
7	52.68	9588.95	6513.55065
8	64.24	9474.68	6666.66667
9	23.41	6650.39	4341.2776
10	20.09	7783.35	3932.5856
11	38.36	9485.95	5771.91785
12	47.4	9138.61	6311.69625

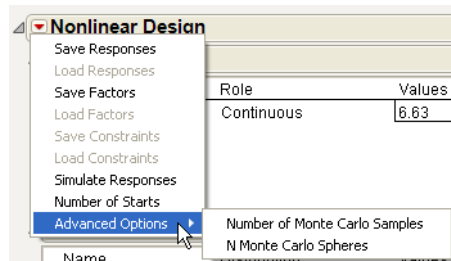
Advanced Options for the Nonlinear Designer

For advanced users, the Nonlinear Designer has the two additional options, as shown in Figure 11.22. These advanced options are included because finding nonlinear DOE solutions involves minimizing the integral of the log of the determinant of the Fisher information matrix with respect to the prior distribution of the parameters. These integrals are complicated and have to be calculated numerically.

The way the integration is done for Normal distribution priors uses a numerical integration technique where the integral is reparameterized into a radial direction, and the number of parameters minus one angular directions. The radial part of the integral is handled using Radau-Gauss-Laguerre quadrature with an evaluation at radius=0. A randomized Mysovskikh quadrature is used to calculate the integral over the spherical part, which is equivalent to integrating over the surface of a hypersphere.

Note: If some of the prior distributions are not Normal, then the integral is reparameterized so that the new parameters have normal distribution, and then the radial-spherical integration method is applied. However, if the prior distribution set for the parameters does not lend itself to a solution, sometimes the process fails and gives the message that the Fisher information is singular in a region of the parameter space, and advises changing the prior distribution or the ranges of the parameters.

Figure 11.22 Advanced Options for the Nonlinear Designer



The following is a technical description for these two advanced options:

Number of Monte Carlo Samples sets the number of octahedra per sphere. Because each octahedron is a fixed unit, this option can be thought of as setting the number of octahedra per sphere.

N Monte Carlo Spheres are the number of nonzero radius values used. The default is two spheres and one center point. Each radial value requires integration over the angular dimensions. This is done by constructing a certain number of hyperoctahedra (the generalization of an octagon in arbitrary dimensions), and randomly rotating each of them.

Technical Note: The reason for the approach given by these advanced options is to get good integral approximations much faster than using standard methods. For instance, with six parameters, using two radii and one sample per sphere, these methods give a generalized five- point rule that needs only 113 observations to get a good approximation. Using the most common approach (Simpson's rule) would need

$5^6 = 15,625$ evaluations. The straight Monte Carlo approach also requires thousands of function evaluations to get the same level of quality in the answer. For more details, see Gotwalt, Jones, and Steinberg (2009).

Keep in mind that if the number of radii is set to zero, then just the center point is used, which gives a local design that is optimal for a particular value of the parameters. For some people this is good enough for their purposes. These designs are created much faster than if the integration is performed.

Chapter 12

Taguchi Designs



Quality was the watchword of 1980s, and Genichi Taguchi was a leader in the growth of quality consciousness. One of Taguchi's technical contributions to the field of quality control was a new approach to industrial experimentation. The purpose of the *Taguchi method* was to develop products that worked well in spite of natural variation in materials, operators, suppliers, and environmental change. This is *robust* engineering.

Much of the Taguchi method is traditional. His *orthogonal arrays* are two-level, three-level, and mixed-level fractional factorial designs. The unique aspects of his approach are the use of *signal* and *noise* factors, *inner* and *outer arrays*, and *signal-to-noise ratios*.

The goal of the Taguchi method is to find control factor settings that generate acceptable responses despite natural environmental and process variability. In each experiment, Taguchi's design approach employs two designs called the *inner* and *outer* array. The Taguchi experiment is the cross product of these two arrays. The *control* factors, used to tweak the process, form the inner array. The *noise* factors, associated with process or environmental variability, form the outer array. Taguchi's *signal-to-noise ratios* are functions of the observed responses over an outer array. The Taguchi designer supports all these features of the Taguchi method. You choose from inner and outer array designs, which use the traditional Taguchi orthogonal arrays, such as L4, L8, and L16.

Dividing system variables according to their signal and noise factors is a key ingredient in robust engineering. Signal factors are system control inputs. Noise factors are variables that are typically difficult or expensive to control.

The inner array is a design in the signal factors and the outer array is a design in the noise factors. A signal-to-noise ratio is a statistic calculated over an entire outer array. Its formula depends on whether the experimental goal is to maximize, minimize or match a target value of the quality characteristic of interest.

A Taguchi experiment repeats the outer array design for each run of the inner array. The response variable in the data analysis is not the raw response or quality characteristic; it is the signal-to-noise ratio.

The **Taguchi** designer in JMP supports signal and noise factors, inner and outer arrays, and signal-to-noise ratios as Taguchi specifies.

Contents

- The Taguchi Design Approach 249
- Taguchi Design Example..... 249
 - Analyze the Data252
- Creating a Taguchi Design254
 - Detail the Response and Add Factors.....254
 - Choose Inner and Outer Array Designs.....255
 - Display Coded Design.....256
 - Make the Design Table257

The Taguchi Design Approach

The Taguchi method defines two types of factors: control factors and noise factors. An *inner* design constructed over the control factors finds optimum settings. An *outer* design over the noise factors looks at how the response behaves for a wide range of noise conditions. The experiment is performed on all combinations of the inner and outer design runs. A performance statistic is calculated across the outer runs for each inner run. This becomes the response for a fit across the inner design runs. Table 12.1 lists the recommended performance statistics.

Table 12.1 Recommended Performance Statistics

Goal	S/N Ratio Formula
nominal is best	$\frac{S}{N} = 10 \log \left(\frac{\bar{Y}^2}{s^2} \right)$
larger-is-better (maximize)	$\frac{S}{N} = -10 \log \left(\frac{1}{n} \sum_i \frac{1}{Y_i^2} \right)$
smaller-is-better (minimize)	$\frac{S}{N} = -10 \log \left(\frac{1}{n} \sum_i Y_i^2 \right)$

Taguchi Design Example

The following example is an experiment done at Baylock Manufacturing Corporation and described by Byrne and Taguchi (1986). The objective of the experiment is to find settings of predetermined control factors that simultaneously maximize the adhesiveness (pull-off force) and minimize the assembly costs of nylon tubing.

To follow along with this example, open the **Byrne Taguchi Data.jmp** table found in the Design Experiment folder of the Sample Data installed with JMP. Or, generate the original design table on your own using **DOE > Taguchi Arrays**.

Table 12.2 shows the signal and noise factors in the Byrne Taguchi Data for this example.

Table 12.2 Signal and Noise Factors

Factor Name	Type	Levels	Comment
Interfer	control	3	tubing and connector interference
Wall	control	3	the wall thickness of the connector
Depth	control	3	insertion depth of the tubing into the connector

Table 12.2 Signal and Noise Factors (Continued)

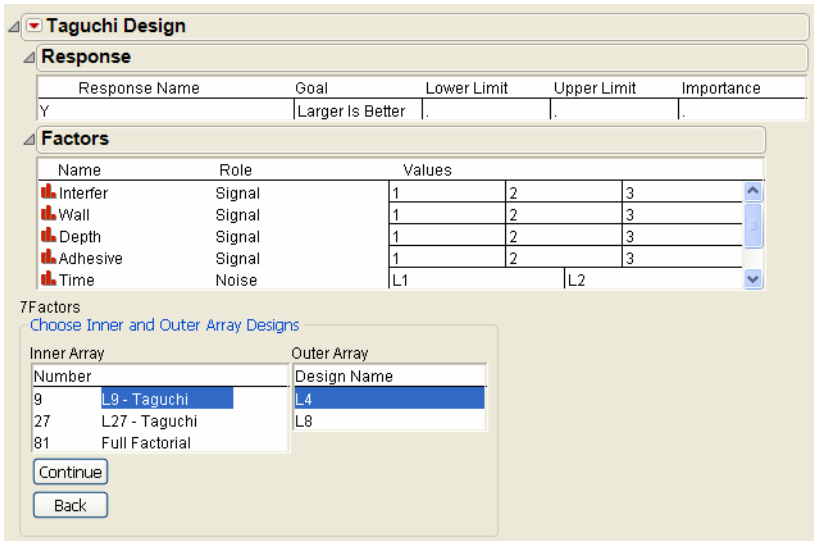
Factor Name	Type	Levels	Comment
Adhesive	control	3	percent adhesive
Time	noise	2	the conditioning time
Temperature	noise	2	temperature
Humidity	noise	2	the relative humidity

To start this example:

1. Select **DOE > Taguchi Arrays**.
2. Click the red triangle icon on the Taguchi Design title bar and select **Load Factors**.
3. When the Open File dialog appears, open the factors table, **Byrne Taguchi Factors.jmp** found in the Design Experiment Sample Data folder installed with JMP.

The factors panel then shows the four three-level control (signal) factors and three noise factors, as shown in Figure 12.1.

Figure 12.1 Response, and Signal and Noise Factors for the Byrne-Taguchi Example



4. Highlight **L9-Taguchi** to give the L9 orthogonal array for the inner design.
5. Highlight **L8** to give an eight-run outer array design.
6. Click **Continue**.

The outer design has three two-level factors. A full factorial in eight runs is generated. However, it is only used as a guide to identify a new set of eight columns in the final JMP data table—one for each combination of levels in the outer design.

- Click **Make Table** to create the design table shown in Figure 12.2.

Figure 12.2 Taguchi Design Before Data Entry

	Interfer	Wall	Depth	Adhesive	Pattern	---	--+	-+-	---+	+--	++-	+++	Mean	SN Ratio
1	1	1	1	1	----	*	*	*	*	*	*	*	*	*
2	1	2	2	2	-000	*	*	*	*	*	*	*	*	*
3	1	3	3	3	---+	*	*	*	*	*	*	*	*	*
4	2	1	2	3	0-0+	*	*	*	*	*	*	*	*	*
5	2	2	3	1	00+-	*	*	*	*	*	*	*	*	*
6	2	3	1	2	0+-0	*	*	*	*	*	*	*	*	*
7	3	1	3	2	+-+0	*	*	*	*	*	*	*	*	*
8	3	2	1	3	+0-+	*	*	*	*	*	*	*	*	*
9	3	3	2	1	++0-	*	*	*	*	*	*	*	*	*

Now, suppose the pull-off adhesive force measures are collected and entered into the columns containing missing data, as shown in Figure 12.3. The missing data column names are appended with the letter Y before the levels (+ or -) of the noise factors for that run. For example, **Y---** is the column of measurements taken with the three noise factors set at their low levels.

- To see the completed experiment, open the data table, *Byrne Taguchi Data.jmp* found in the Design Experiment Sample Data folder installed with JMP. Figure 12.3 shows the completed design.

Figure 12.3 Complete Taguchi Design Table (*Byrne Taguchi Data.jmp*)

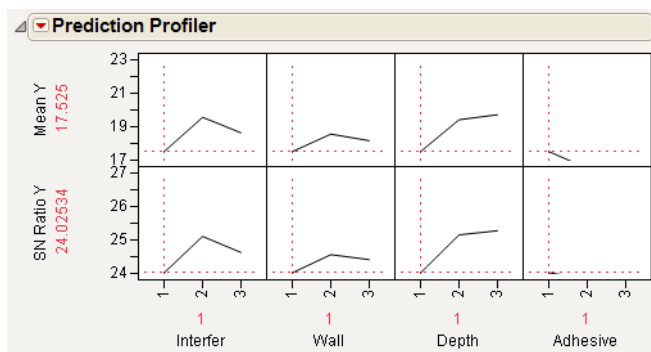
	Interfer	Wall	Depth	Adhesive	Pattern	Y---	Y-+	Y+-	Y++	Y+--	Y+-+	Y++-	Y+++	Mean Y	SN Ratio Y
1	1	1	1	1	----	15.6	9.5	16.9	19.9	19.6	19.6	20	19.1	17.525	24.02534
2	1	2	2	2	-000	15	16.2	19.4	19.6	19.7	19.8	24.2	21.9	19.475	25.52164
3	1	3	3	3	---+	16.3	16.7	19.1	15.6	22.6	18.2	23.3	20.4	19.025	25.33476
4	2	1	2	3	0-0+	18.3	17.4	18.9	18.6	21	18.9	23.2	24.7	20.125	25.90425
5	2	2	3	1	00+-	19.7	18.6	19.4	25.1	25.6	21.4	27.5	25.3	22.825	26.90753
6	2	3	1	2	0+-0	16.2	16.3	20	19.8	14.7	19.6	22.5	24.7	19.225	25.32574
7	3	1	3	2	+-+0	16.4	19.1	18.4	23.6	16.8	18.6	24.3	21.6	19.85	25.71081
8	3	2	1	3	+0-+	14.2	15.6	15.1	16.8	17.8	19.6	23.2	24.4	18.3375	24.83231
9	3	3	2	1	++0-	16.1	19.9	19.3	17.3	23.1	22.7	22.6	28.6	21.2	26.15198

The column named **SN Ratio Y** is the performance statistic computed with the formula shown below. In this case, it is the “larger-the-better” (LTB) formula, which is -10 times the common logarithm of the average squared reciprocal:

3. To open the Prediction Profiler, click the red triangle on the Response Mean Y title bar and select **Factor Profiling > Profiler**.

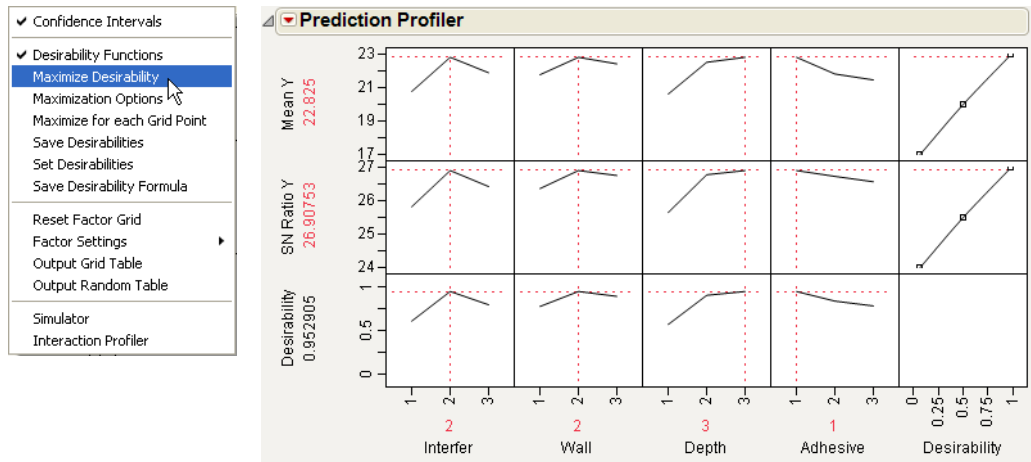
The profile traces (Figure 12.5) indicate that different settings of the first three factors would increase SN Ratio Y.

Figure 12.5 The Prediction Profiler



4. To find optimal settings, click the red triangle on the Prediction Profiler title bar and select **Desirability Functions**. This adds the row of traces and a column of function settings to the profiler, as shown in Figure 12.6. The default desirability functions are set to larger-is-better, which is what you want in this experiment. See *Modeling and Multivariate Methods* for more details about the prediction profiler.
5. Again click the red triangle on the Prediction Profiler title bar and select **Maximize Desirability** to automatically set the prediction traces that give the best results according to the desirability functions.

In this example, the settings for Interfer and Wall changed from 1 to 2. (See Figure 12.5 and Figure 12.6). The Depth setting changed from 1 to 3. The settings for Adhesive did not change. These new settings increased the signal-to-noise ratio from 24.0253 to 26.9075.

Figure 12.6 Best Factor Settings for Byrne Taguchi Data


Creating a Taguchi Design

To start a Taguchi design, select **DOE > Taguchi Arrays**, or click the **Taguchi Arrays** button on the JMP Starter **DOE** page. Then, follow the steps below:

- “Detail the Response and Add Factors,” p. 254
- “Choose Inner and Outer Array Designs,” p. 255
- “Display Coded Design,” p. 256
- “Make the Design Table,” p. 257

Detail the Response and Add Factors

The Responses panel has a single default response. The steps for setting up the details of this response are outlined in Figure 12.7. For information on importance weights and lower and upper limits, see “Understanding Response Importance Weights,” p. 66.

1. Double-click to edit the response name.
2. Click to change the response goal: Larger Is Better, Nominal is Best, Smaller is Better, or None.
3. Click to enter lower and upper limits and importance weights.

Figure 12.7 Setting Up the Response

Response Name	Goal	Lower Limit	Upper Limit	Importance
Y	Larger Is Better	.	.	.

The steps for setting up the factors are outlined in Figure 12.8.

1. Click to add a signal, then select a signal type: **2 Level**, or **3 Level**.
Or click to add a noise.
2. Double-click to edit the factor name.
3. To change the value of a signal or noise, click and then type the new value.

Figure 12.8 Entering Factors

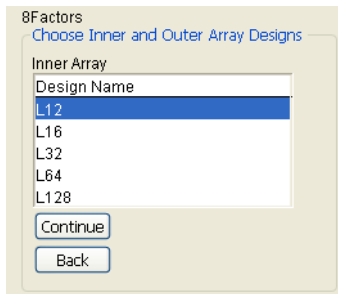
Name	Role	Values
X1	Signal	1 2
X2	Signal	1 2

When you finish adding factors, click **Continue**.

Choose Inner and Outer Array Designs

Your choice for inner and outer arrays depends on the number and type of factors you have. Figure 12.9 shows the available inner array choices when you have eight signal factors. If you also have noise factors, choices include designs for the outer array. To follow along, enter eight two-level **Signal** factors and click **Continue**. Then highlight the design you want and again click **Continue**. This example uses the L12 design.

Figure 12.9 Selecting a Design for Eight Signal Factors

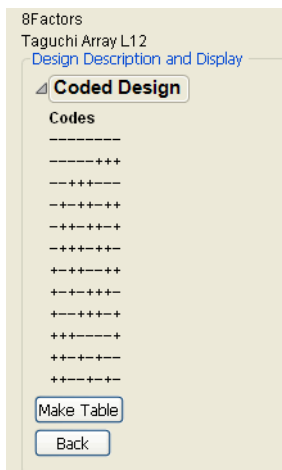


If you did not specify a noise factor, after you click **Continue**, a dialog appears that asks you to specify how many times you want to perform each inner array run. Specify two (2) for this example.

Display Coded Design

After you select a design type, the Coded Design (Figure 12.10) is shown below the Factors panel.

Figure 12.10 Coding for Eight Factor L12 Design



The Coded Design shows the pattern of high and low values for the factors in each run. For more details on the coded design, see “[Understanding Design Codes](#),” p. 118.

Chapter 13

Augmented Designs



If you treat experimentation as an iterative process, you can master the temptation to assume that one successful screening experiment has optimized your process. You can also avoid disappointment if a screening experiment leaves behind some ambiguities. The augment designer helps facilitate experimentation as an iterative process.

The augment designer modifies an existing design data table, supporting your iterative process. It gives the following five choices:

- replicate the design a specified number of times
- add center points
- create a foldover design
- add axial points together with center points to transform a screening design to a response surface design
- add runs to the design using a model that can have more terms than the original model

This chapter provides an overview of the augment designer. It also presents a case study of design augmentation.

Contents

A D-Optimal Augmentation of the Reactor Example	261
Analyze the Augmented Design	264
Creating an Augmented Design	271
Replicate a Design	272
Add Center Points	275
Creating a Foldover Design	276
Adding Axial Points	277
Adding New Runs and Terms	278
Special Augment Design Commands	281
Save the Design (X) Matrix	282
Modify the Design Criterion (D- or I- Optimality)	282
Select the Number of Random Starts	283
Specify the Sphere Radius Value	283
Disallow Factor Combinations	284

A D-Optimal Augmentation of the Reactor Example

This example, adapted from Meyer, *et al.* (1996), demonstrates how to use the augment designer in JMP to resolve ambiguities left by a screening design. In this study, a chemical engineer investigates the effects of five factors on the percent reaction of a chemical process.

To begin, open **Reactor 8 Runs.jmp** found in the Design Experiment Sample Data folder installed with JMP. Then select **Augment Design** from the **DOE** menu. When the initial launch dialog appears:

1. Select **Percent Reacted** and click **Y, Response**.
2. Select all other variables except **Pattern** and click **X, Factor**.
3. Click **OK** on the launch dialog to see the Augment Design dialog in Figure 13.1.

Note: You can check **Group new runs into separate block** to add a blocking factor to any design. However, the purpose of this example is to estimate all two-factor interactions in 16 runs, which can't be done when there is the additional blocking factor in the model.

Figure 13.1 Augment Design Dialog for the Reactor Example

Name	Role	Changes	Values
Feed Rate	Continuous	Easy	10 15
Catalyst	Continuous	Easy	1 2
Stir Rate	Continuous	Easy	100 120
Temperature	Continuous	Easy	140 180
Concentration	Continuous	Easy	3 6

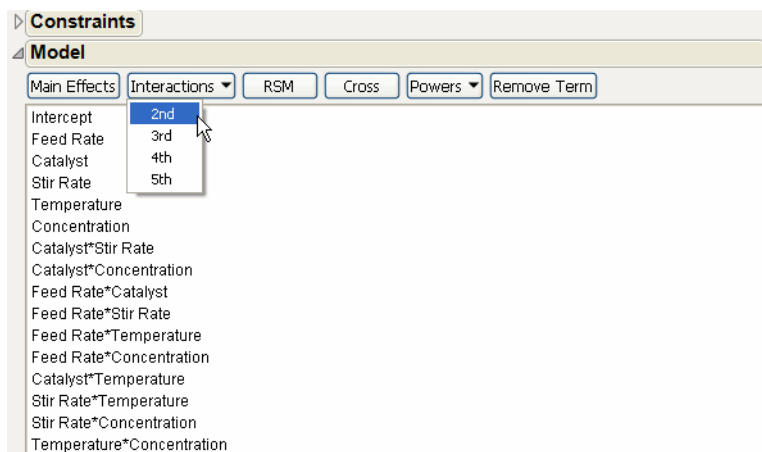
☐ Group new runs into separate block

Augmentation Choices

Replicate Add Centerpoints Fold Over Add Axial Augment

4. Now click **Augment** on the Augment Design dialog to see the display in Figure 13.2.

This model shown in Figure 13.2 is the result of the model stored with the data table when it was created by the Custom designer. However, the augmented design is to have 16 runs in order to estimate all two-factor interactions.

Figure 13.3 Augmented Model with All Two-Factor Interactions6. Click **Make Design**.

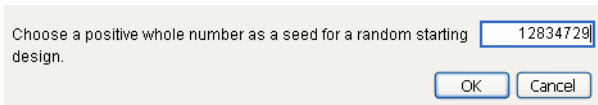
JMP now computes *D*-optimally augmented factor settings, similar to the design shown in Figure 13.4.

Figure 13.4 D-Optimally Augmented Factor Settings

Design							Percent Reacted
Run	Feed Rate	Catalyst	Stir Rate	Temperature	Concentration		
1	10	1	100	180	6		44
2	10	1	120	180	3		66
3	10	2	100	140	6		70
4	10	2	120	140	3		54
5	15	1	100	140	3		53
6	15	1	120	140	6		55
7	15	2	100	180	3		93
8	15	2	120	180	6		82
9	10	2	120	180	6		.
10	15	2	100	180	6		.
11	10	1	100	140	6		.
12	15	2	100	140	3		.
13	15	1	120	180	6		.
14	10	1	100	140	3		.
15	10	2	120	180	3		.
16	15	1	120	140	3		.

Note: The resulting design is a function of an initial random number seed. To reproduce the exact factor settings table in Figure 13.4, (or the most recent design you generated), choose **Set Random Seed** from the popup menu on the Augment Design title bar. A dialog shows the most recently used random number. Click **OK** to use that number again, or **Cancel** to generate a design based on a new random number. The dialog in Figure 13.5 shows the random number (12834729) used to generate the runs in Figure 13.4.

Figure 13.5 Specifying a Random Number



7. Click **Make Table** to generate the JMP table with *D*-Optimally augmented runs.

Analyze the Augmented Design

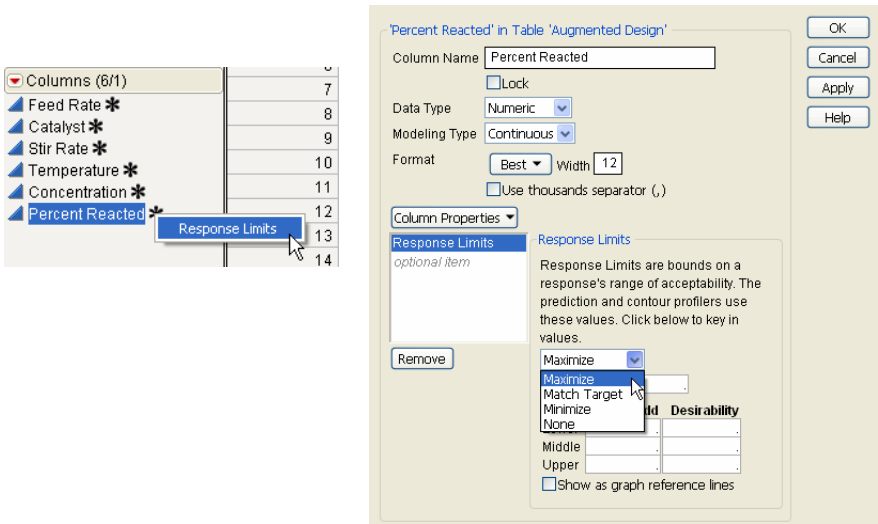
Suppose you have already run the experiment on the augmented data and recorded results in the **Percent Reacted** column of the data table.

1. To see these results, open **Reactor Augment Data.jmp** found in the **Design Experiment Sample Data** folder installed with JMP.

It is desirable to maximize **Percent Reacted**, however its column in this sample data table has a response limits column property set to **Minimize**.

2. Click the asterisk next to the **Percent Reacted** column name in the **Columns** panel of the data table and select **Response Limits**, as shown on the left in Figure 13.6.
3. In the **Column Info** dialog that appears, change the response limit to **Maximize**, as shown on the right in Figure 13.6.

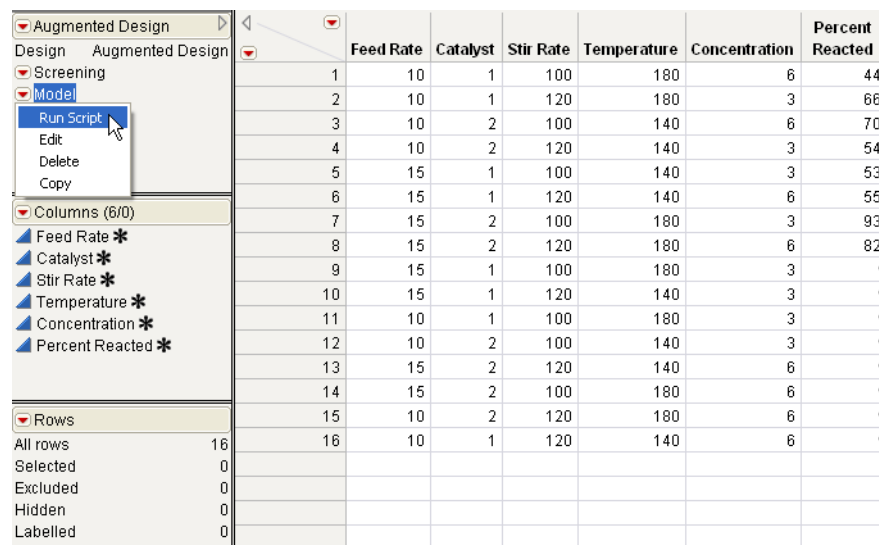
Figure 13.6 Change the Response Limits Column Property for the Percent Reacted Column



You are now ready to run the analysis.

4. To start the analysis, click the red triangle for **Model** in the upper left of the data table and select **Run Script** from the menu, as shown in Figure 13.7.

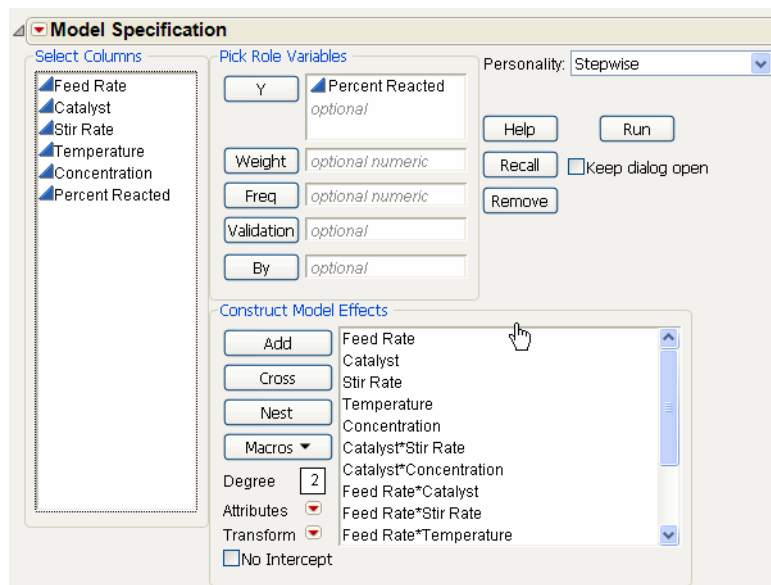
Figure 13.7 Completed Augmented Experiment (Reactor Augment Data.jmp)



	Feed Rate	Catalyst	Stir Rate	Temperature	Concentration	Percent Reacted
1	10	1	100	180	6	44
2	10	1	120	180	3	66
3	10	2	100	140	6	70
4	10	2	120	140	3	54
5	15	1	100	140	3	53
6	15	1	120	140	6	55
7	15	2	100	180	3	93
8	15	2	120	180	6	82
9	15	1	100	180	3	*
10	15	1	120	140	3	*
11	10	1	100	180	3	*
12	10	2	100	140	3	*
13	15	2	120	140	6	*
14	15	2	100	180	6	*
15	10	2	120	180	6	*
16	10	1	120	140	6	*

The **Model** script, stored as a table property with the data, contains the JSL commands that display the Fit Model dialog with all main effects and two-factor interactions as effects.

5. Change the fitting personality on the Fit Model dialog from **Standard Least Squares** to **Stepwise**, as shown in Figure 13.8.

Figure 13.8 Fit Model Dialog for Stepwise Regression on Generated Model

6. When you click **Run**, the stepwise regression control panel appears. Click the check boxes for all the main effect terms.

Important: Choose **P-value Threshold** from the **Stopping Rule** menu, **Mixed** from the **Direction** menu, and make sure **Prob to Enter** is 0.050 and **Prob to Leave** is 0.100. These are not the default values. You should see the dialog shown in Figure 13.9.

Figure 13.9 Initial Stepwise Model

Stepwise Fit for Percent Reacted

Stepwise Regression Control

Stopping Rule: P-value Threshold

Prob to Enter: 0.05
Prob to Leave: 0.1

Direction: Mixed

Rules: Combine

SSE	DFE	RMSE	RSquare	RSquare Adj	Cp	p	AICc	BIC
1257.1058	10	11.212073	0.6933	0.5399	6	143.2297	134.6378	

Current Estimates

Lock	Entered	Parameter	Estimate	nDF	SS	"F Ratio"	"Prob>F"
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	Intercept	66.8125	1	0	0.000	1
<input type="checkbox"/>	<input checked="" type="checkbox"/>	Feed Rate(10,15)	4.4375	1	315.0625	2.506	0.14448
<input type="checkbox"/>	<input checked="" type="checkbox"/>	Catalyst(1,2)	10.1153846	1	1520.198	12.093	0.00595
<input type="checkbox"/>	<input checked="" type="checkbox"/>	Stir Rate(100,120)	-0.3653846	1	1.983516	0.016	0.90253
<input type="checkbox"/>	<input checked="" type="checkbox"/>	Temperature(140,180)	6.11538462	1	555.6264	4.420	0.06183
<input type="checkbox"/>	<input checked="" type="checkbox"/>	Concentration(3,6)	-1.7884615	1	41.58173	0.331	0.57791
<input type="checkbox"/>	<input type="checkbox"/>	Catalyst*Stir Rate	0	1	346.6875	3.427	0.09715
<input type="checkbox"/>	<input type="checkbox"/>	Catalyst*Concentration	0	1	46.44643	0.345	0.57124
<input type="checkbox"/>	<input type="checkbox"/>	Feed Rate*Catalyst	0	1	2.548951	0.018	0.89541
<input type="checkbox"/>	<input type="checkbox"/>	Feed Rate*Stir Rate	0	1	0.064103	0.000	0.98338
<input type="checkbox"/>	<input type="checkbox"/>	Feed Rate*Temperature	0	1	0.730769	0.005	0.9439
<input type="checkbox"/>	<input type="checkbox"/>	Feed Rate*Concentration	0	1	0.016484	0.000	0.99157
<input type="checkbox"/>	<input type="checkbox"/>	Catalyst*Temperature	0	1	667.5208	10.190	0.01097
<input type="checkbox"/>	<input type="checkbox"/>	Stir Rate*Temperature	0	1	22.5625	0.164	0.69453
<input type="checkbox"/>	<input type="checkbox"/>	Stir Rate*Concentration	0	1	2.571429	0.018	0.89495
<input type="checkbox"/>	<input type="checkbox"/>	Temperature*Concentration	0	1	631.1429	9.074	0.01466

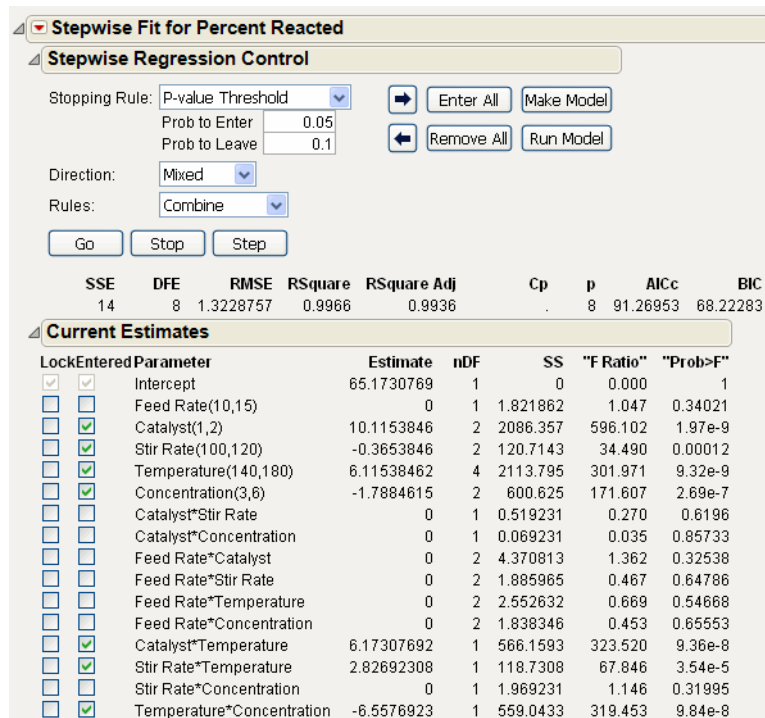
Step History

Step	Parameter	Action	"Sig Prob"	Seq SS	RSquare	Cp	p	AICc	BIC
1	Feed Rate(10,15)	Entered	0.2985	315.0625	0.0769	.	2	140.859	141.176
2	Catalyst(1,2)	Entered	0.0047	1785.063	0.5124	.	3	134.282	133.736
3	Stir Rate(100,120)	Entered	0.8049	10.5625	0.5150	.	4	138.561	136.424
4	Temperature(140,180)	Entered	0.0343	689.0625	0.6831	.	5	137.084	132.386
5	Concentration(3,6)	Entered	0.5779	41.58173	0.6933	.	6	143.23	134.638

- Click **Go** to start the stepwise regression and watch it continue until all terms are entered into the model that meet the **Prob to Enter** and **Prob to Leave** criteria in the Stepwise Regression Control panel.

Figure 13.10, shows the result of this example analysis. Note that **Feed Rate** is out of the model while the **Catalyst*Temperature**, **Stir Rate*Temperature**, and the **Temperature*Concentration** interactions have entered the model.

Figure 13.10 Completed Stepwise Model



- After Stepwise is finished, click **Make Model** on the Stepwise control panel to generate this reduced model, as shown in Figure 13.11.
- Click **Run** and fit the reduced model to do additional diagnostic work, make predictions, and find the optimal factor settings.

Figure 13.11 New Prediction Model Dialog

The Analysis of Variance and Lack of Fit Tests in Figure 13.12, indicate a highly significant regression model with no evidence of Lack of Fit.

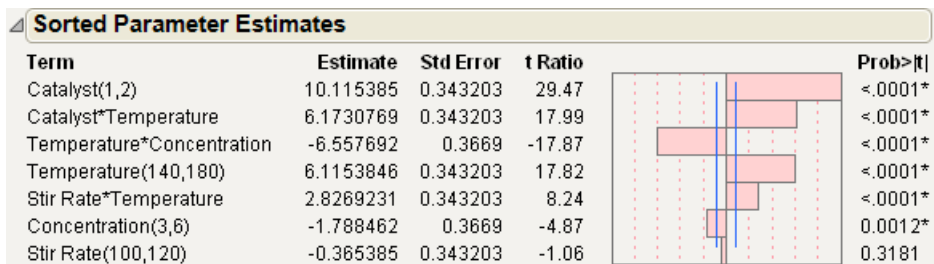
Figure 13.12 Prediction Model Analysis of Variance and Lack of Fit Tests

Response Percent Reacted				
Actual by Predicted Plot				
Summary of Fit				
Analysis of Variance				
Source	DF	Sum of Squares	Mean Square	F Ratio
Model	7	4084.4375	583.491	333.4235
Error	8	14.0000	1.750	Prob > F
C. Total	15	4098.4375		<.0001*
Lack Of Fit				
Source	DF	Sum of Squares	Mean Square	F Ratio
Lack Of Fit	6	5.500000	0.91667	Prob > F
Pure Error	2	8.500000	4.25000	0.9394
Total Error	8	14.000000		Max RSq
				0.9979

A D-Optimal Augmentation of the Reactor Example

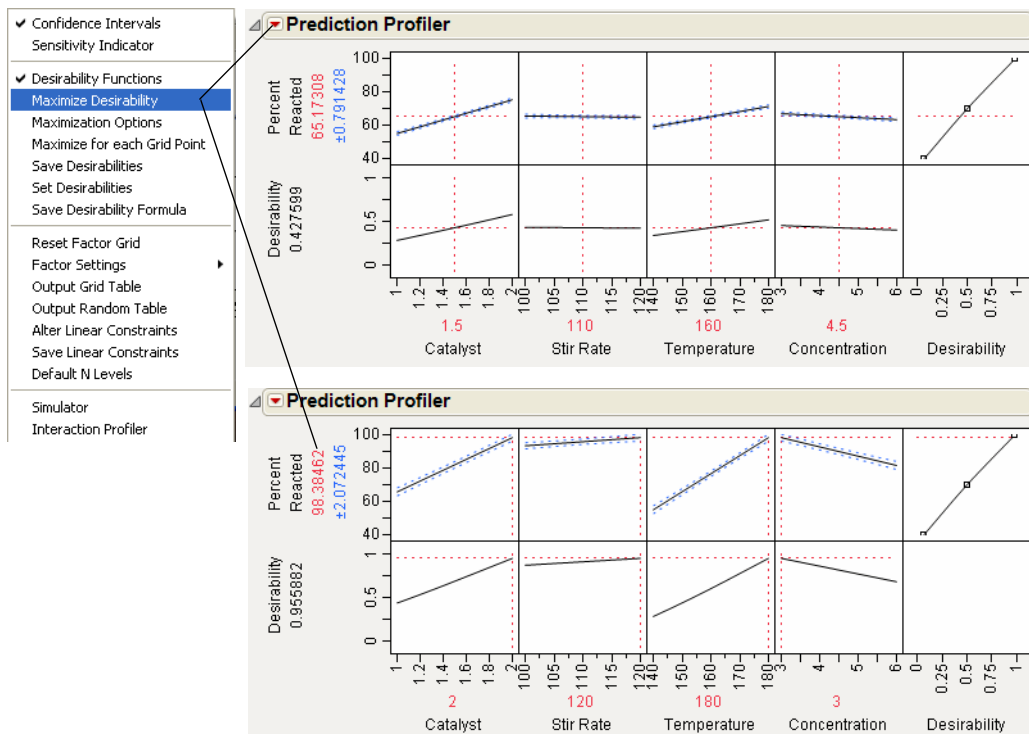
The Sorted Parameter Estimates table in Figure 13.13 shows that Catalyst has the largest main effect. However, the significance of the two-factor interactions are of the same order of magnitude as the main effects. This is the reason that the initial screening experiment, shown in the chapter “Screening Designs,” p. 101, had ambiguous results.

Figure 13.13 Prediction Model Estimates Plot



10. Choose **Maximize Desirability** from the menu on the Prediction Profiler title bar.

The prediction profile plot at the bottom in Figure 13.14 shows that maximum occurs at the high levels of Catalyst, Stir Rate, and Temperature and the low level of Concentration. At these extreme settings, the estimate of Percent Reacted increases from 65.17 to 98.38.

Figure 13.14 Maximum Percent Reacted

To summarize, compare the analysis of 16 runs with the analyses of reactor data from previous chapters:

- “Screening Designs,” p. 101, the analysis of a screening design with only 8 runs produced a model with the five main effects and two interaction effects with confounding. None of the factors effects were significant, although the Catalyst factor was large enough to encourage collecting data for further runs.
- “Full Factorial Designs,” p. 143, a full factorial of the five two-level reactor factors, 32 runs, was first subjected to a stepwise regression. This approach identified three main effects (Catalyst, Temperature, and Concentration) and two interactions (Temperature*Catalyst, Contentration*Temperature) as significant effects.
- By using a D-optimal augmentation of 8 runs to produce 8 additional runs, a stepwise analysis returned the same results as the analysis of 32 runs. The bottom line is that only half as many runs yielded the same information. Thus, using an iterative approach to DOE can save time and money.

Creating an Augmented Design

The augment designer modifies an existing design data table. It gives the following five choices:

Replicate replicates the design a specified number of times. See “Replicate a Design,” p. 272.

Add Centerpoints adds center points. See “Add Center Points,” p. 275.

Fold Over creates a foldover design. See “Creating a Foldover Design,” p. 276.

Add Axial adds axial points together with center points to transform a screening design to a response surface design. See “Adding Axial Points,” p. 277.

Augment adds runs to the design (augment) using a model, which can have more terms than the original model. See “Adding New Runs and Terms,” p. 278.

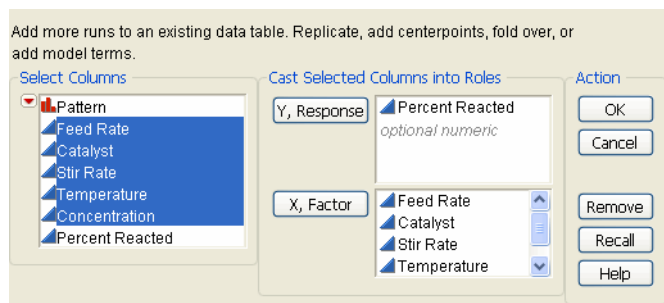
Replicate a Design

Replication provides a direct check on the assumption that the error variance is constant. It also reduces the variability of the regression coefficients in the presence of large process or measurement variability.

To replicate the design a specified number of times:

1. Open a data table that contains a design you want to augment. This example uses *Reactor 8 Runs.jmp* from the Design Experiment Sample Data folder installed with JMP.
2. Select **DOE > Augment Design** to see the initial dialog for specifying factors and responses.
3. Select **Percent Reacted** and click **Y, Response**.
4. Select all other variables (except **Pattern**) and click **X, Factor** to identify the factors you want to use for the augmented design (Figure 13.15).

Figure 13.15 Identify Response and Factors



5. Click **OK** to see the Augment Design panel shown in Figure 13.16.
6. If you want the original runs and the resulting augmented runs to be identified by a blocking factor, check the box beside **Group New Runs into Separate Block** on the Augment Design panel.

Figure 13.16 Choose an Augmentation Type

Augment Design

Factors

Name	Role	Changes	Values
Feed Rate	Continuous	Easy	10 15
Catalyst	Continuous	Easy	1 2
Stir Rate	Continuous	Easy	100 120
Temperature	Continuous	Easy	140 180
Concentration	Continuous	Easy	3 6

☐ Group new runs into separate block

Augmentation Choices

Replicate Add Centerpoints Fold Over Add Axial **Augment**

7. Click the **Replicate** button to see the dialog shown on the left in Figure 13.17. Enter the number of times you want JMP to perform each run, then click **OK**.

Note: Entering 2 specifies that you want each run to appear twice in the resulting design. This is the same as one replicate (Figure 13.17).

8. View the design, shown on the right in Figure 13.17.

Figure 13.17 Reactor Data Design Augmented With Two Replicates

Number times to perform each run.

OK Cancel

Factor Design

Run	Feed Rate	Catalyst	Stir Rate	Temperature	Concentration	Percent Reacted
1	10	1	100	180	6	44
2	10	1	120	180	3	66
3	10	2	100	140	6	70
4	10	2	120	140	3	54
5	15	1	100	140	3	53
6	15	1	120	140	6	55
7	15	2	100	180	3	93
8	15	2	120	180	6	82
9	10	1	100	180	6	.
10	10	1	120	180	3	.
11	10	2	100	140	6	.
12	10	2	120	140	3	.
13	15	1	100	140	3	.
14	15	1	120	140	6	.
15	15	2	100	180	3	.
16	15	2	120	180	6	.

Prediction Variance Profile

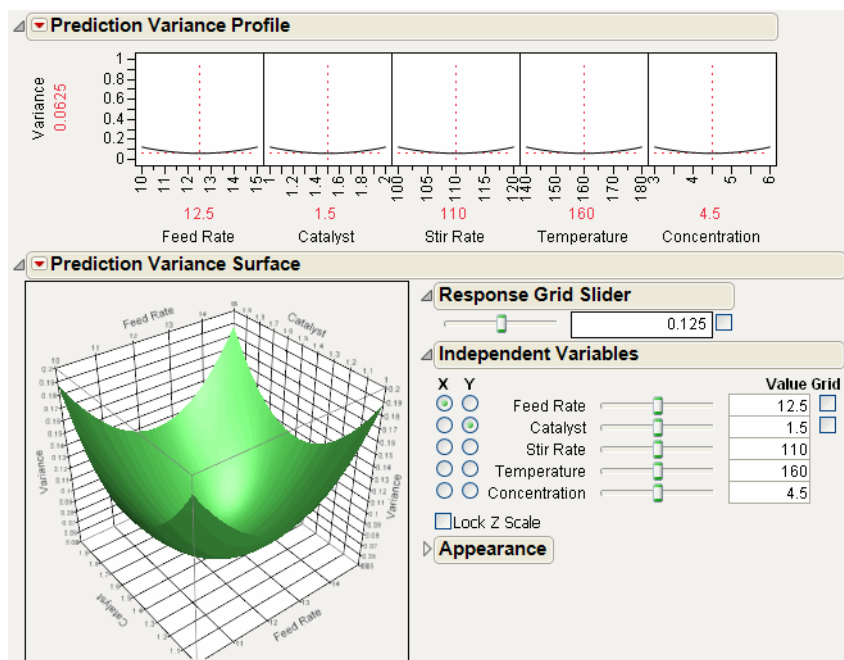
Prediction Variance Surface

Make Table

Make Table

Back

9. Click the disclosure icons next to Prediction Variance Profile and Prediction Variance Surface to see the profile and surface plots shown in Figure 13.18.

Figure 13.18 Prediction Profiler and Surface Plot


10. Click **Make Table** to produce the design table shown in Figure 13.19.

Figure 13.19 The Replicated Design

Augmented Design							
Design	Augmented Design						
Screening							
Model							
Columns (6/0)							
Feed Rate *							
Catalyst *							
Stir Rate *							
Temperature *							
Concentration *							
Percent Reacted *							
Rows							
All rows	16						
Selected	0						
Excluded	0						
Hidden	0						
Labelled	0						

	Feed Rate	Catalyst	Stir Rate	Temperature	Concentration	Percent Reacted
1	10	1	100	180	6	44
2	10	1	120	180	3	66
3	10	2	100	140	6	70
4	10	2	120	140	3	54
5	15	1	100	140	3	53
6	15	1	120	140	6	55
7	15	2	100	180	3	93
8	15	2	120	180	6	82
9	10	1	100	180	6	•
10	10	1	120	180	3	•
11	10	2	100	140	6	•
12	10	2	120	140	3	•
13	15	1	100	140	3	•
14	15	1	120	140	6	•
15	15	2	100	180	3	•
16	15	2	120	180	6	•

Add Center Points

Adding center points is useful to check for curvature and reduce the prediction error in the center of the factor region. Center points are usually replicated points that allow for an independent estimate of pure error, which can be used in a lack-of-fit test.

To add center points:

1. Open a data table that contains a design you want to augment. This example uses **Reactor 8 Runs.jmp** found in the **Design Experiment Sample Data** folder installed with JMP.
2. Select **DOE > Augment Design**.
3. In the initial Augment Design dialog, identify the response and factors you want to use for the augmented design (see Figure 13.15) and click **OK**.
4. If you want the original runs and the resulting augmented runs to be identified by a blocking factor, check the box beside **Group new runs into separate block**. (Figure 13.16 shows the check box location directly under the Factors panel.)
5. Click the **Add Centerpoints** button and type the number of center points you want to add. For this example, add two center points, and click **OK**.
6. Click **Make Table** to see the data table in Figure 13.20.

The table shows two center points appended to the end of the design.

Figure 13.20 Design with Two Center Points Added

Augmented Design		Feed Rate	Catalyst	Stir Rate	Temperature	Concentration	Percent Reacted
Design	Augmented Design						
Screening		1	10	1	100	6	44
Model		2	10	1	120	3	66
Columns (6/0)		3	10	2	100	6	70
Feed Rate *		4	10	2	140	3	54
Catalyst *		5	15	1	100	3	53
Stir Rate *		6	15	1	120	6	55
Temperature *		7	15	2	100	3	93
Concentration *		8	15	2	120	6	82
Percent Reacted *		9	12.5	1.5	160	4.5	•
Rows		10	12.5	1.5	160	4.5	•
All rows	10						
Selected	0						
Excluded	0						
Hidden	0						
Labelled	0						

Creating a Foldover Design

A foldover design removes the confounding of two-factor interactions and main effects. This is especially useful as a follow-up to saturated or near-saturated fractional factorial or Plackett-Burman designs.

To create a foldover design:

1. Open a data table that contains a design you want to augment. This example uses *Reactor 8 Runs.jmp*, found in the Design Experiment Sample Data folder installed with JMP.
2. Select **DOE > Augment Design**.
3. In the initial Augment Design dialog, identify the response and factors you want to use for the augmented design (see Figure 13.15) and click **OK**.
4. Check the box to the left of **Group new runs into separate block**. (Figure 13.16 shows the check box location directly under the Factors panel.) This identifies the original runs and the resulting augmented runs with a blocking factor.
5. Click the **Fold Over** button. A dialog appears that lists all the design factors.
6. Choose (select) which factors to fold. The default, if you choose no factors, is to fold on all design factors. If you choose a subset of factors to fold over, the remaining factors are replicates of the original runs. The example in Figure 13.21 folds on all five factors and includes a blocking factor.
7. Click **Make Table**. The design data table that results lists the original set of runs as block 1 and the new (foldover) runs are block 2.

Figure 13.21 Listing of a Foldover Design On All Factors

Augmented Design									
Design	Augmented Design		Feed Rate	Catalyst	Stir Rate	Temperature	Concentration	Block	Percent Reacted
Screening		1	10	1	100	180	6	1	44
Model		2	10	1	120	180	3	1	66
Columns (7/10)		3	10	2	100	140	6	1	70
Feed Rate *		4	10	2	120	140	3	1	54
Catalyst *		5	15	1	100	140	3	1	53
Stir Rate *		6	15	1	120	140	6	1	55
Temperature *		7	15	2	100	180	3	1	93
Concentration *		8	15	2	120	180	6	1	82
Block *		9	15	2	120	140	3	2	•
Percent Reacted *		10	15	2	100	140	6	2	•
Rows		11	15	1	120	180	3	2	•
All rows	16	12	15	1	100	180	6	2	•
Selected	0	13	10	2	120	180	6	2	•
Excluded	0	14	10	2	100	180	3	2	•
Hidden	0	15	10	1	120	140	6	2	•
Labelled	0	16	10	1	100	140	3	2	•

Adding Axial Points

You can add axial points together with center points, which transforms a screening design to a response surface design. To do this:

1. Open a data table that contains a design you want to augment. This example uses **Reactor 8 Runs.jmp**, from the **Design Experiment Sample Data** folder installed with JMP.
2. Select **DOE > Augment Design**.
3. In the initial Augment Design dialog, identify the response and factors you want to use for the augmented design (see Figure 13.15) and click **OK**.
4. If you want the original runs and the resulting augmented runs to be identified by a blocking factor, check the box beside **Group New Runs into Separate Block** (Figure 13.16).
5. Click **Add Axial**.
6. Enter the axial values in units of the factors scaled from -1 to $+1$, then enter the number of center points you want. When you click **OK**, the augmented design includes the number of center points specified and constructs two axial points for each variable in the original design.

Figure 13.22 Entering Axial Values

Axial values dialog

Please supply an axial value.

Please supply the number of center points desired.

OK Cancel

- Click **Make Table**. The design table appears. Figure 13.23 shows a table augmented with two center points and two axial points for five variables.

Figure 13.23 Design Augmented With Two Center and Ten Axial Points

	Feed Rate	Catalyst	Stir Rate	Temperature	Concentration	Percent Reacted
1	10	1	100	180	6	44
2	10	1	120	180	3	66
3	10	2	100	140	6	70
4	10	2	120	140	3	54
5	15	1	100	140	3	53
6	15	1	120	140	6	55
7	15	2	100	180	3	93
8	15	2	120	180	6	82
9	12.5	1.5	110	160	4.5	*
10	12.5	1.5	110	160	4.5	*
11	15	1.5	110	160	4.5	*
12	10	1.5	110	160	4.5	*
13	12.5	2	110	160	4.5	*
14	12.5	1	110	160	4.5	*
15	12.5	1.5	120	160	4.5	*
16	12.5	1.5	100	160	4.5	*
17	12.5	1.5	110	180	4.5	*
18	12.5	1.5	110	140	4.5	*
19	12.5	1.5	110	160	6	*
20	12.5	1.5	110	160	3	*

Adding New Runs and Terms

A powerful use of the augment designer is to add runs using a model that can have more terms than the original model. For example, you can achieve the objectives of response surface methodology by changing a linear model to a full quadratic model and adding the necessary number of runs. Suppose you start with a two-factor, two-level, four-run design. If you add quadratic terms to the model and five new points, JMP generates the 3 by 3 full factorial as the optimal augmented design.

D-optimal augmentation is a powerful tool for sequential design. Using this feature you can add terms to the original model and find optimal new test runs with respect to this expanded model. You can also group the two sets of experimental runs into separate blocks, which optimally blocks the second set with respect to the first.

To add new runs and terms to the original model:

- Open a data table that contains a design you want to augment. This example uses **Reactor Augment Data.jmp**, from the Design Experiment Sample Data folder installed with JMP.
- Select **DOE > Augment Design**.
- In the initial Augment Design dialog, identify the response and factors you want to use for the augmented design (see Figure 13.15) and click **OK**.
- If you want the original runs and the resulting augmented runs to be identified by a blocking factor, check the box beside **Group New Runs into Separate Block** (not used in this example).

- Click the **Augment** button. The original number of runs (Figure 13.24) appear in the Factor Design panel.

Figure 13.24 Viewing the Existing Design

Augment Design

Factors

Name	Role	Changes	Values
Feed Rate	Continuous	Easy	10 15
Catalyst	Continuous	Easy	1 2
Stir Rate	Continuous	Easy	100 120
Temperature	Continuous	Easy	140 180
Concentration	Continuous	Easy	3 6

☐ Group new runs into separate block

Constraints

Model

Main Effects Interactions RSM Cross Powers Remove Term

Intercept
Feed Rate
Catalyst
Stir Rate
Temperature
Concentration
Catalyst*Stir Rate
Catalyst*Concentration

Alias Terms

Factor Design

Run	Feed Rate	Catalyst	Stir Rate	Temperature	Concentration	Percent Reacted
1	10	1	100	180	6	44
2	10	1	120	180	3	66
3	10	2	100	140	6	70
4	10	2	120	140	3	54
5	15	1	100	140	3	53
6	15	1	120	140	6	55
7	15	2	100	180	3	93
8	15	2	120	180	6	82
9	15	2	120	180	3	98
10	15	2	120	140	6	65
11	10	2	100	140	3	63
12	10	1	120	180	6	49
13	15	1	100	140	6	63
14	15	1	100	180	3	61
15	10	1	120	140	6	59
16	10	2	100	180	3	94

Design Generation

Enter Number of Runs (counting 16 included runs):

Make Design

- In the Design Generation panel, enter the number of total runs you want this design to contain. The number you enter is the original number of runs plus the number of additional runs you want.
- Click the **Make Design** button. The new number of runs (Figure 13.25) appear in the Design panel.

Figure 13.25 24 Total Runs

▼ Augment Design

► Factors

► Alias Terms

▼ Design

Run	Feed Rate	Catalyst	Stir Rate	Temperature	Concentration	Percent Reacted
1	10	1	100	180	6	44
2	10	1	120	180	3	66
3	10	2	100	140	6	70
4	10	2	120	140	3	54
5	15	1	100	140	3	53
6	15	1	120	140	6	55
7	15	2	100	180	3	93
8	15	2	120	180	6	82
9	15	2	120	180	3	98
10	15	2	120	140	6	65
11	10	2	100	140	3	63
12	10	1	120	180	6	49
13	15	1	100	140	6	63
14	15	1	100	180	3	61
15	10	1	120	140	6	59
16	10	2	100	180	3	94
17	15	2	100	180	6	.
18	10	2	120	180	6	.
19	15	1	120	140	3	.
20	10	1	100	140	3	.
21	10	1	100	180	3	.
22	15	1	120	180	6	.
23	15	1	120	180	3	.
24	15	2	100	140	3	.

▼ Design Evaluation

► ▼ Prediction Variance Profile

► Fraction of Design Space Plot

► ▼ Prediction Variance Surface

► Relative Variance of Coefficients

► Alias Matrix

► Color Map On Correlations

► Design Diagnostics

Output Options

Make Table

Back

8. If desired, view the prediction variance profile and the prediction variance surface.
9. Click **Make Table** to create the augmented design JMP table (Figure 13.26) with the additional runs.

Figure 13.26 The Augmented Design Table with New Runs

Augmented Design			Feed Rate	Catalyst	Stir Rate	Temperature	Concentration	Percent Reacted
Design	Augmented Design							
Screening								
Model								
Columns (6/0)								
Feed Rate *								
Catalyst *								
Stir Rate *								
Temperature *								
Concentration *								
Percent Reacted *								
Rows								
All rows	24							
Selected	0							
Excluded	0							
Hidden	0							
Labelled	0							

1	10	1	100	180	6	44
2	10	1	120	180	3	66
3	10	2	100	140	6	70
4	10	2	120	140	3	54
5	15	1	100	140	3	53
6	15	1	120	140	6	55
7	15	2	100	180	3	93
8	15	2	120	180	6	82
9	15	2	120	180	3	98
10	15	2	120	140	6	65
11	10	2	100	140	3	63
12	10	1	120	180	6	49
13	15	1	100	140	6	63
14	15	1	100	180	3	61
15	10	1	120	140	6	59
16	10	2	100	180	3	94
17	15	2	100	180	6	•
18	10	2	120	180	6	•
19	15	1	120	140	3	•
20	10	1	100	140	3	•
21	10	1	100	180	3	•
22	15	1	120	180	6	•
23	15	1	120	180	3	•
24	15	2	100	140	3	•

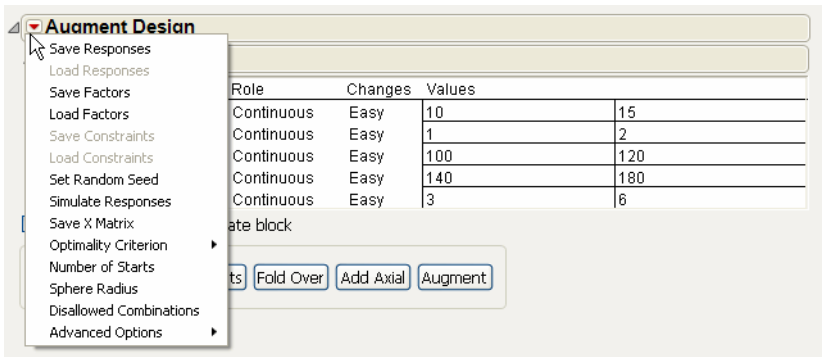
additional runs

Technical Note: The Augment designer does not support designs that have terms whose estimability has been set to If Possible instead of Necessary, as is done in some screening designs that have fewer runs than terms.

Special Augment Design Commands

After you select **DOE > Augment Design** and identify factors and responses, the window in Figure 13.27 appears. Click the red triangle icon on the Augment Design title bar to see a list of commands. Most of these commands are for saving and loading information about variables; they are available in all designs and more information is in “[Special Custom Design Commands](#),” p. 83. The following sections describe commands found in this menu that are specific to augment designs.

Figure 13.27 Click the Red Triangle Icon to Reveal Commands

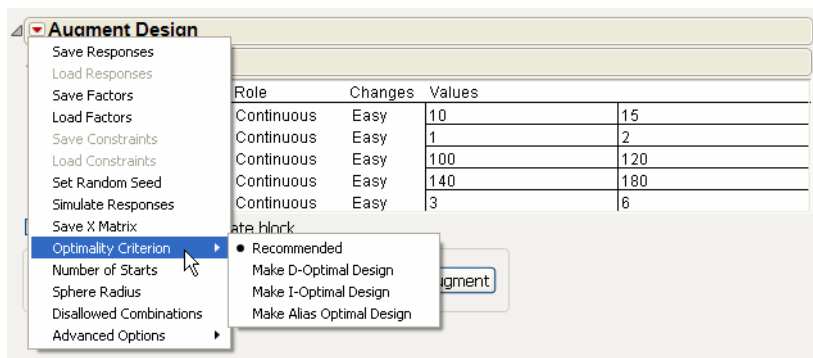


Save the Design (X) Matrix

To create a script and save it as a table property in the JMP design data table, click the red triangle icon in the Augment Design title bar (Figure 13.27) and select **Save X Matrix**. Two or three scripts are saved to the table. **Moments Matrix** and **Design Matrix** scripts are always saved. If the design is a split plot design, an additional **V Inverse** script is also saved. When you run the **Moments Matrix** script, JMP creates a matrix called Moments and displays its number of rows in the log. When you run the **Design Matrix** script, JMP creates a matrix called X and displays its number of rows in the log. When you run the **V Inverse** script, JMP creates the inverse of the variance matrix of the responses, and displays its number of rows in the log. If you do not have the log visible, select **View > Log** or **Window > Log** on the Macintosh.

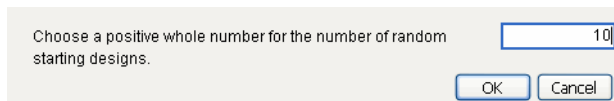
Modify the Design Criterion (D- or I- Optimality)

To modify the design optimality criterion, click the red triangle icon in the Augment Design title bar (Figure 13.28) and select **Optimality Criterion**, then choose **Make D-Optimal Design** or **Make I-Optimal Design**. The default criterion for **Recommended** is D-optimal for all design types unless you have used the **RSM** button in the Model panel to add effects that make the model quadratic.

Figure 13.28 Change the Optimality Criterion

Select the Number of Random Starts

To override the default number of random starts, click the red triangle icon in the Augment Design title bar (Figure 13.28) and select **Number of Starts**. The window in Figure 13.29 appears with an edit box for you to enter the number of random starts for the design you want to build. The number you enter overrides the default number of starts, which varies depending on the design.

Figure 13.29 Changing the Number of Starts

For additional information on the number of starts, see [“Why Change the Number of Starts?”](#), p. 89.

Specify the Sphere Radius Value

Augment designs can be constrained to a hypersphere. To edit the sphere radius for the design in units of the coded factors (–1, 1), click the red triangle icon in the Augment Design title bar (Figure 13.27) and select **Sphere Radius**. Enter the appropriate value and click **OK**.

Or, use JSL and submit the following command before you build a custom design:

```
DOE Sphere Radius = 1.0;
```

In this statement you can replace 1.0 with any positive number.

Disallow Factor Combinations

In addition to linear inequality constraints on continuous factors and constraining a design to a hypersphere, you can define general factor constraints on the factors. You can disallow any combination of levels of categorical factors if you have not already defined linear inequality constraints.

For information on how to do this, see [“Disallowed Combinations: Accounting for Factor Level Restrictions,” p. 90.](#)

Prospective Sample Size and Power



Use the **DOE > Sample Size and Power** command to answer the question “How many runs do I need?” The important quantities are sample size, power, and the magnitude of the effect. These depend on the significance level, alpha, of the hypothesis test for the effect and the standard deviation of the noise in the response. You can supply either one or two of the three values. If you supply only one of these values, the result is a plot of the other two. If you supply two values, the third value is computed.

The **Sample Size and Power** platform can answer the question, “Will I detect the group differences I am looking for, given my proposed sample size, estimate of within-group variance, and alpha level?” In this type of analysis, you must approximate the group means and sample sizes in a data table as well as approximate the within-group standard deviation (σ).

The sample size and power computations determine the sample size necessary for yielding a significant result, given that the true effect size is at least a certain size. It requires that you enter two out of three possible quantities; difference to detect, sample size, and power. The third quantity is computed for the following cases:

- difference between a one sample mean and a hypothesized value
- difference between two sample means
- differences in the means among k samples
- difference between a standard deviation and a hypothesized value
- difference between a one sample proportion and a hypothesized value
- difference between two sample proportions
- difference between counts per unit in a Poisson-distributed sample and a hypothesized value.

The calculations assume that there are equal numbers of units in each group. You can apply this platform to more general experimental designs, if they are balanced and an adjustment for the number-of-parameters is specified.

You can also compute the required sample sizes needed for reliability studies and demonstrations.

Contents

Launching the Sample Size and Power Platform	287
One-Sample and Two-Sample Means	287
Single-Sample Mean	289
Sample Size and Power Animation for One Mean	292
Two-Sample Means	293
k-Sample Means	294
One Sample Standard Deviation	296
One Sample Standard Deviation Example	297
One-Sample and Two-Sample Proportions	298
One Sample Proportion	298
Two Sample Proportions	300
Counts per Unit	303
Counts per Unit Example	304
Sigma Quality Level	305
Sigma Quality Level Example	305
Number of Defects Computation Example	306
Reliability Test Plan and Demonstration	306
Reliability Test Plan	306
Reliability Demonstration	309

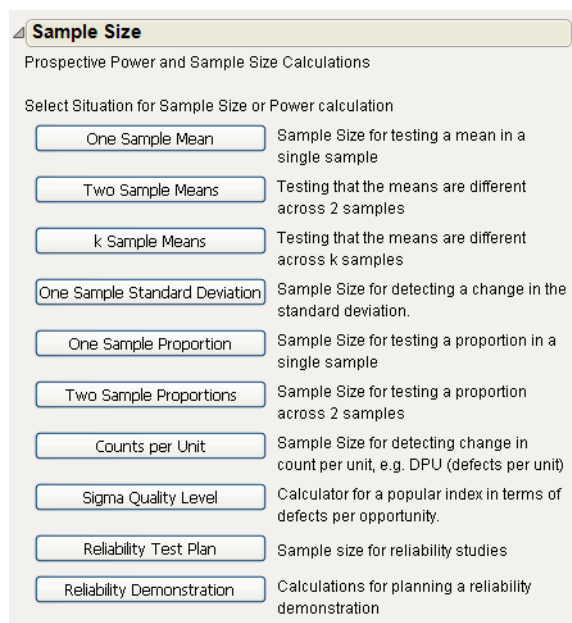
Launching the Sample Size and Power Platform

The **Sample Size and Power** platform helps you plan your study for a single mean or proportion comparison, a two sample mean or proportion comparison, a one-sample standard deviation comparison, a k sample means comparison, or a counts per unit comparison. Depending upon your experimental situation, you supply one or two quantities to obtain a third quantity. These quantities include:

- required sample size
- expected power
- expected effect size

When you select **DOE > Sample Size and Power**, the panel in Figure 14.1 appears with button selections for experimental situations. The following sections describe each of these selections and explain how to enter the quantities and obtain the desired computation.

Figure 14.1 Sample Size and Power Choices



One-Sample and Two-Sample Means

After you click either **One Sample Mean**, or **Two Sample Means** in the initial Sample Size selection list (Figure 14.1), a Sample Size and Power window appears. (See Figure 14.2.)

Figure 14.2 Initial Sample Size and Power Windows for Single Mean (left) and Two Means (right)

Sample Size
One Mean
Testing if one mean is different from the hypothesized value.
Alpha
Std Dev
Extra Parameters

Supply two values to determine the third.
Enter one value to see a plot of the other two.

Difference to detect
Sample Size
Power

Sample Size
Two Means
Testing if two means are different from each other.
Alpha
Std Dev
Extra Parameters

Supply two values to determine the third.
Enter one value to see a plot of the other two.

Difference to detect
Sample Size
Power
Sample Size is the total sample size; per group would be n/2

The windows are the same except that the One Mean window has a button at the bottom that accesses an animation script.

The initial Sample Size and Power window requires values for **Alpha**, **Std Dev** (the error standard deviation), and one or two of the other three values: **Difference to detect**, **Sample Size**, and **Power**. The Sample Size and Power platform calculates the missing item. If there are two unspecified fields, a plot is constructed, showing the relationship between these two values:

- power as a function of sample size, given specific effect size
- power as a function of effect size, given a sample size
- effect size as a function of sample size, for a given power.

The Sample Size and Power window asks for these values:

Alpha is the probability of a type I error, which is the probability of rejecting the null hypothesis when it is true. It is commonly referred to as the significance level of the test. The default alpha level is 0.05. This implies a willingness to accept (if the true difference between groups is zero) that, 5% (alpha) of the time, a significant difference is incorrectly declared.

Std Dev is the error standard deviation. It is a measure of the unexplained random variation around the mean. Even though the true error is not known, the power calculations are an exercise in probability that calculates what might happen if the true value is the one you specify. An estimate of the error standard deviation could be the root mean square error (RMSE) from a previous model fit.

Extra Parameters is only for multi-factor designs. Leave this field zero in simple cases. In a multi-factor balanced design, in addition to fitting the means described in the situation, there are other factors with extra parameters that can be specified here. For example, in a three-factor two-level design with all three two-factor interactions, the number of extra parameters is five. (This includes two parameters for the extra main effects, and three parameters for the interactions.) In

practice, the particular values entered are not that important, unless the experimental range has very few degrees of freedom for error.

Difference to Detect is the smallest detectable difference (how small a difference you want to be able to declare statistically significant) to test against. For single sample problems this is the difference between the hypothesized value and the true value.

Sample Size is the total number of observations (runs, experimental units, or samples) in your experiment. Sample size is not the number per group, but the total over all groups.

Power is the probability of rejecting the null hypothesis when it is false. A large power value is better, but the cost is a higher sample size.

Continue evaluates at the entered values.

Back returns to the previous Sample Size and Power window so that you can either redo an analysis or start a new analysis.

Animation Script runs a JSL script that displays an interactive plot showing power or sample size. See the section, [“Sample Size and Power Animation for One Mean,” p. 292](#), for an illustration of the animation script.

Single-Sample Mean

Using the Sample Size and Power window, you can test if one mean is different from the hypothesized value.

For the one sample mean, the hypothesis supported is

$$H_0: \mu = \mu_0$$

and the two-sided alternative is

$$H_a: \mu \neq \mu_0$$

where μ is the population mean and μ_0 is the null mean to test against or is the difference to detect. It is assumed that the population of interest is normally distributed and the true mean is zero. Note that the power for this setting is the same as for the power when the null hypothesis is $H_0: \mu=0$ and the true mean is μ_0 .

Suppose you are interested in testing the flammability of a new fabric being developed by your company. Previous testing indicates that the standard deviation for burn times of this fabric is 2 seconds. The goal is to detect a difference of 1.5 seconds when alpha is equal to 0.05, the sample size is 20, and the standard deviation is 2 seconds. For this example, μ_0 is equal to 1.5. To calculate the power:

1. Select **DOE > Sample Size and Power**.
2. Click the **One Sample Mean** button in the Sample Size and Power Window.
3. Leave **Alpha** as 0.05.
4. Leave **Extra Parameters** as 0.
5. Enter 2 for **Std Dev**.
6. Enter 1.5 as **Difference to detect**.
7. Enter 20 for **Sample Size**.

8. Leave **Power** blank. (See the left window in Figure 14.3.)
9. Click **Continue**.

The power is calculated as 0.8888478174 and is rounded to 0.89. (See right window in Figure 14.3.) The conclusion is that your experiment has an 89% chance of detecting a significant difference in the burn time, given that your significance level is 0.05, the difference to detect is 1.5 seconds, and the sample size is 20.

Figure 14.3 A One-Sample Example

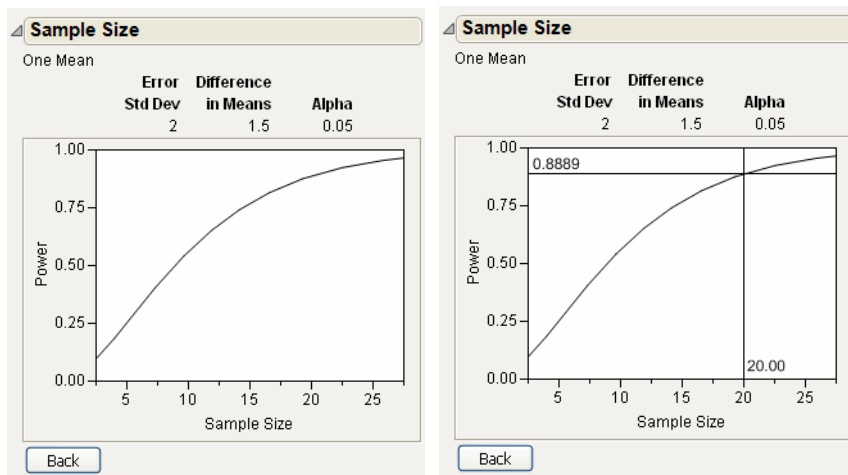
Sample Size
One Mean
Testing if one mean is different from the hypothesized value.
Alpha
Std Dev
Extra Parameters
Supply two values to determine the third.
Enter one value to see a plot of the other two.
Difference to detect
Sample Size
Power

Sample Size
One Mean
Testing if one mean is different from the hypothesized value.
Alpha
Std Dev
Extra Parameters
Supply two values to determine the third.
Enter one value to see a plot of the other two.
Difference to detect
Sample Size
Power

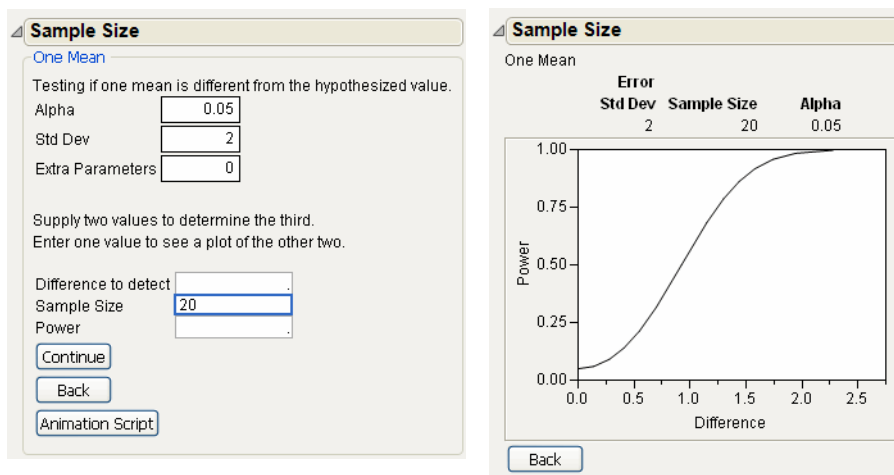
Power versus Sample Size Plot

To see a plot of the relationship of **Sample Size** and **Power**, leave both **Sample Size** and **Power** empty in the window and click **Continue**.

The plots in Figure 14.4, show a range of sample sizes for which the power varies from about 0.1 to about 0.95. The plot on the right in Figure 14.4 shows using the crosshair tool to illustrate the example in Figure 14.3.

Figure 14.4 A One-Sample Example Plot**Power versus Difference Plot**

When only **Sample Size** is specified (Figure 14.5) and **Difference to Detect** and **Power** are empty, a plot of Power by Difference appears, after clicking **Continue**.

Figure 14.5 Plot of Power by Difference to Detect for a Given Sample Size

Sample Size and Power Animation for One Mean

Clicking the **Animation Script** button on the Sample Size and Power window for one mean shows an interactive plot. This plot illustrates the effect that changing the sample size has on power. As an example of using the **Animation Script**:

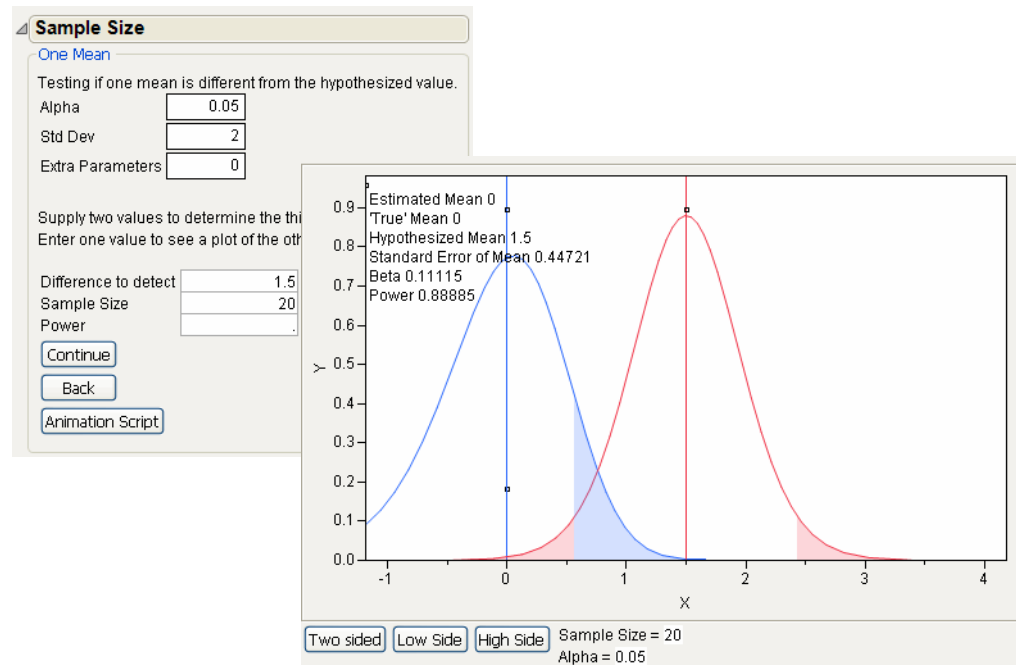
1. Select **DOE > Sample Size and Power**.
2. Click the **One Sample Mean** button in the Sample Size and Power Window.
3. Enter 2 for **Std Dev**.
4. Enter 1.5 as **Difference to detect**.
5. Enter 20 for **Sample Size**.
6. Leave **Power** blank.

The Sample Size and Power window appears as shown on the left of Figure 14.6.

7. Click **Animation Script**.

The initial animation plot shows two t-density curves. The blue curve shows the t-distribution when the true mean is zero. The red curve shows the t-distribution when the true mean is 1.5, which is the difference to be detected. The probability of committing a type II error (not detecting a difference when there is a difference) is shaded blue on this plot. (This probability is often represented as β in the literature.) Similarly, the probability of committing a type I error (deciding that the difference to detect is significant when there is no difference) is shaded as the red areas under the red curve. (The red-shaded areas under the curve are represented as α in the literature.)

Select and drag the square handles to see the changes in statistics based on the positions of the curves. To change the values of **Sample Size** and **Alpha**, click on their values beneath the plot.

Figure 14.6 Example of Animation Script to Illustrate Power

Two-Sample Means

The Sample Size and Power windows work similarly for one and two sample means; the **Difference to Detect** is the difference between two means. The comparison is between two random samples instead of one sample and a hypothesized mean.

For testing the difference between two means, the hypothesis supported is

$$H_0: \mu_1 - \mu_2 = D_0$$

and the two-sided alternative is

$$H_a: \mu_1 - \mu_2 \neq D_0$$

where μ_1 and μ_2 are the two population means and D_0 is the difference in the two means or the difference to detect. It is assumed that the populations of interest are normally distributed and the true difference is zero. Note that the power for this setting is the same as for the power when the null hypothesis is $H_0: \mu_1 - \mu_2 = 0$ and the true difference is D_0 .

Suppose the standard deviation is 2 (as before) for both groups, the desired detectable difference between the two means is 1.5, and the sample size is 30 (15 per group). To estimate the power for this example:

1. Select **DOE > Sample Size and Power**.

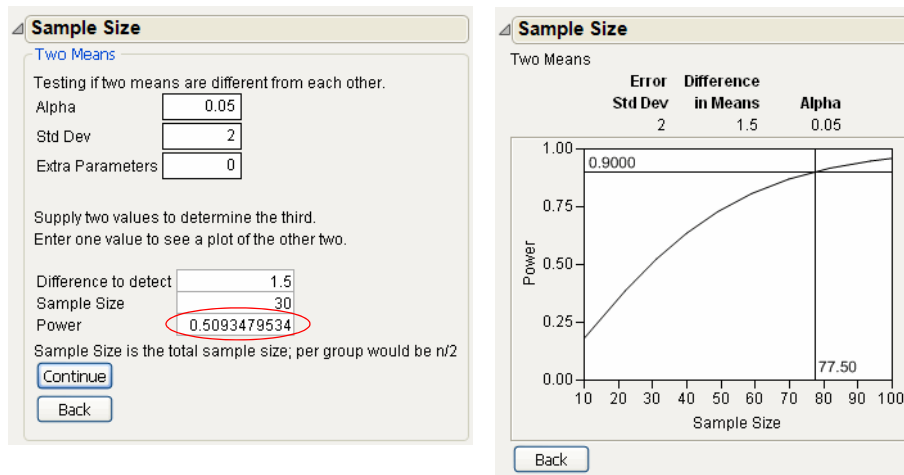
2. Click the **Two Sample Means** button in the Sample Size and Power Window.
3. Leave **Alpha** as 0.05.
4. Enter 2 for **Std Dev**.
5. Leave **Extra Parameters** as 0.
6. Enter 1.5 as **Difference to detect**.
7. Enter 30 for **Sample Size**.
8. Leave **Power** blank.
9. Click **Continue**.

The **Power** is calculated as 0.51. (See the left window in Figure 14.7.) This means that you have a 51% chance of detecting a significant difference between the two sample means when your significance level is 0.05, the difference to detect is 1.5, and each sample size is 15.

Plot of Power by Sample Size

To have a greater power requires a larger sample. To find out how large, leave both **Sample Size** and **Power** blank for this same example and click **Continue**. Figure 14.7 shows the resulting plot, with the crosshair tool estimating that a sample size of about 78 is needed to obtain a power of 0.9.

Figure 14.7 Plot of Power by Sample Size to Detect for a Given Difference



k-Sample Means

Using the **k-Sample Means** option, you can compare up to 10 means. Consider a situation where 4 levels of means are expected to be in the range of 10 to 13, the standard deviation is 0.9, and your sample size is 16.

The hypothesis to be tested is:

$H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4$ versus H_a : at least one mean is different

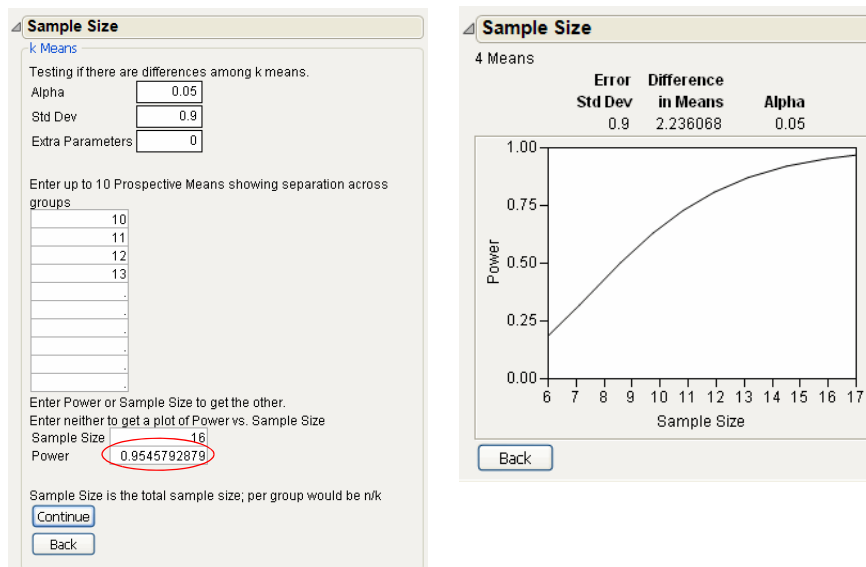
To determine the power:

1. Select **DOE > Sample Size and Power**.
2. Click the **k Sample Means** button in the Sample Size and Power Window.
3. Leave **Alpha** as 0.05.
4. Enter 0.9 for **Std Dev**.
5. Leave **Extra Parameters** as 0.
6. Enter 10, 11, 12, and 13 as the four levels of means.
7. Enter 16 for **Sample Size**.
8. Leave **Power** blank.
9. Click **Continue**.

The **Power** is calculated as 0.95. (See the left of Figure 14.8.) This means that there is a 95% chance of detecting that at least one of the means is different when the significance level is 0.05, the population means are 10, 11, 12, and 13, and the total sample size is 16.

If both **Sample Size** and **Power** are left blank for this example, the sample size and power calculations produce the **Power** versus **Sample Size** curve. (See the right of Figure 14.8.) This confirms that a sample size of 16 looks acceptable.

Notice that the difference in means is 2.236, calculated as square root of the sum of squared deviations from the grand mean. In this case it is the square root of $(-1.5)^2 + (-0.5)^2 + (0.5)^2 + (1.5)^2$, which is the square root of 5.

Figure 14.8 Prospective Power for k-Means and Plot of Power by Sample Size

One Sample Standard Deviation

Use the **One-Sample Standard Deviation** option on the Sample Size and Power window (Figure 14.1) to determine the sample size needed for detecting a change in the standard deviation of your data. The usual purpose of this option is to compute a large enough sample size to guarantee that the risk of a type II error, β , is small. (This is the probability of failing to reject the null hypothesis when it is false.)

In the Sample Size and Power window, specify:

Alpha is the significance level, usually 0.05. This implies a willingness to accept (if the true difference between standard deviation and the hypothesized standard deviation is zero) that a significant difference is incorrectly declared 5% of the time.

Hypothesized Standard Deviation is the hypothesized or baseline standard deviation to which the sample standard deviation is compared.

Alternative Standard Deviation can select **Larger** or **Smaller** from the menu to indicate the direction of the change you want to detect.

Difference to Detect is the smallest detectable difference (how small a difference you want to be able to declare statistically significant). For single sample problems this is the difference between the hypothesized value and the true value.

Sample Size is how many experimental units (runs, or samples) are involved in the experiment.

Power is the probability of declaring a significant result. It is the probability of rejecting the null hypothesis when it is false.

In the lower part of the window you enter two of the items and the Sample Size and Power calculation determines the third.

Some examples in this chapter use engineering examples from the online manual of The National Institute of Standards and Technology (NIST). You can access the NIST manual examples at <http://www.itl.nist.gov/div898/handbook>.

One Sample Standard Deviation Example

One example from the NIST manual states a problem in terms of the variance and difference to detect. The variance for resistivity measurements on a lot of silicon wafers is claimed to be 100 ohm-cm. The buyer is unwilling to accept a shipment if the variance is greater than 155 ohm-cm for a particular lot (55 ohm-cm above the baseline of 100 ohm-cm).

In the Sample Size and Power window, the One Sample Standard Deviation computations use the standard deviation instead of the variance. The hypothesis to be tested is:

$H_0: \sigma = \sigma_0$, where σ_0 is the hypothesized standard deviation. The true standard deviation is σ_0 plus the difference to detect.

In this example the hypothesized standard deviation, σ_0 , is 10 (the square root of 100) and σ is 12.4499 (the square root of $100 + 55 = 155$). The difference to detect is $12.4499 - 10 = 2.4499$.

You want to detect an increase in the standard deviation of 2.4499 for a standard deviation of 10, with an alpha of 0.05 and a power of 0.99. To determine the necessary sample size:

1. Select **DOE > Sample Size and Power**.
2. Click the **One Sample Standard Deviation** button in the Sample Size and Power Window.
3. Leave **Alpha** as 0.05.
4. Enter 10 for **Hypothesized Standard Deviation**.
5. Select Larger for **Alternate Standard Deviation**.
6. Enter 2.4499 as **Difference to Detect**.
7. Enter 0.99 for **Power**.
8. Leave **Sample Size** blank. (See the left of Figure 14.9.)
9. Click **Continue**.

The **Sample Size** is calculated as 171. (See the right of Figure 14.9.) This result is the sample size rounded up to the next whole number.

Note: Sometimes you want to detect a change to a smaller standard deviation. If you select **Smaller** from the **Alternative Standard Deviation** menu, enter a negative amount in the **Difference to Detect** field.

Figure 14.9 Window To Compare Single-Direction One-Sample Standard Deviation

Sample Size
 One Sample Standard Deviation
 Sample Size for detecting a difference in the standard deviation.
 Alpha: 0.05
 Hypothesized Standard Deviation: 10
 Alternative Standard Deviation: Larger
 Supply two values to determine the third.
 Difference to detect: 2.4499
 Sample Size: 171
 Power: 0.99
 [Continue] [Back]

Sample Size
 One Sample Standard Deviation
 Sample Size for detecting a difference in the standard deviation.
 Alpha: 0.05
 Hypothesized Standard Deviation: 10
 Alternative Standard Deviation: Larger
 Supply two values to determine the third.
 Difference to detect: 2.4499
 Sample Size: 171
 Power: 0.99
 [Continue] [Back]

One-Sample and Two-Sample Proportions

The Sample Size windows and computations to test sample sizes and power for proportions are similar to those for testing means. You enter a true **Proportion** and choose an **Alpha** level. Then, for the one-sample proportion case, enter the **Sample Size** and **Null Proportion** to obtain the **Power**. Or, enter the **Power** and **Null Proportion** to obtain the **Sample Size**. Similarly, to obtain a value for **Null Proportion**, enter values for **Sample Size** and **Power**. For the two-sample proportion case, either the two sample sizes or the desired **Power** must be entered. (See Figure 14.10 and Figure 14.11.)

The sampling distribution for proportions is normal, but the computations to determine sample size and test proportions use exact methods based on the binomial distribution. Exact methods are more reliable since using the normal approximation to the binomial can provide erroneous results when small samples or proportions are used. Exact power calculations are used in conjunction with a modified Wald test statistic described in Agresti and Coull (1998).

One Sample Proportion

Clicking the **One Sample Proportion** option on the Sample Size and Power window yields a One Proportion window. In this window, you can specify the alpha level and the true proportion. The sample size, power, or the hypothesized proportion is calculated. If you supply two of these quantities, the third is computed, or if you enter any one of the quantities, you see a plot of the other two.

For example, if you have a hypothesized proportion of defects, you can use the **One Sample Proportion** window to estimate a large enough sample size to guarantee that the risk of accepting a false hypothesis (β) is small. That is, you want to detect, with reasonable certainty, a difference in the proportion of defects.

For the one sample proportion, the hypothesis supported is

$$H_0: p = p_0$$

and the two-sided alternative is

$$H_a: p \neq p_0$$

where p is the population proportion and p_0 is the null proportion to test against. Note that if you are interested in testing whether the population proportion is greater than or less than the null proportion, you use a one-sided test. The one-sided alternative is either

$$H_a: p < p_0$$

or

$$H_a: p > p_0$$

One-Sample Proportion Window Specifications

In the top portion of the Sample Size window, you can specify or enter values for:

Alpha is the significance level of your test. The default value is 0.05.

Proportion is the true proportion, which could be known or hypothesized. The default value is 0.1.

One-Sided or Two-Sided Specify either a one-sided or a two-sided test. The default setting is the two-sided test.

In the bottom portion of the window, enter two of the following quantities to see the third, or a single quantity to see a plot of the other two.

Null Proportion is the proportion to test against (p_0) or is left blank for computation. The default value is 0.2.

Sample Size is the sample size, or is left blank for computation. If **Sample Size** is left blank, then values for **Proportion** and **Null Proportion** must be different.

Power is the desired power, or is left blank for computation.

One-Sample Proportion Example

As an example, suppose that an assembly line has a historical proportion of defects equal to 0.1, and you want to know the power to detect that the proportion is different from 0.2, given an alpha level of 0.05 and a sample size of 100.

1. Select **DOE > Sample Size and Power**.
2. Click **One Sample Proportion**.
3. Leave **Alpha** as 0.05.
4. Leave 0.1 as the value for **Proportion**.
5. Accept the default option of **Two-Sided**. (A one-sided test is selected if you are interested in testing if the proportion is either greater than or less than the **Null Proportion**.)
6. Leave 0.2 as the value for **Null Proportion**.
7. Enter 100 as the **Sample Size**.
8. Click **Continue**.

The **Power** is calculated and is shown as approximately 0.7. See Figure 14.10.

Figure 14.10 Power and Sample Window for One-Sample Proportions

Sample Size

One Proportion

Testing if one proportion is different from the hypothesized value.

Alpha Ho: P = P₀

Proportion

☒ Two-Sided
☐ One-Sided

Enter one value to see a plot of the other two.

Null Proportion
 Sample Size
 Power

Two Sample Proportions

The **Two Sample Proportions** option computes the power or sample sizes needed to detect the difference between two proportions, p_1 and p_2 .

For the two sample proportion, the hypothesis supported is

$$H_0: p_1 - p_2 = D_0$$

and the two-sided alternative is

$$H_a: p_1 - p_2 \neq D_0$$

where p_1 and p_2 are the population proportions from two populations, and D_0 is the hypothesized difference in proportions.

The one-sided alternative is either

$$H_a: (p_1 - p_2) < D_0$$

or

$$H_a: (p_1 - p_2) > D_0$$

Two Sample Proportion Window Specifications

Specifications for the Two Sample Proportions window include:

Alpha is the significance level of your test. The default value is 0.05.

Proportion 1 is the proportion for population 1, which could be known or hypothesized. The default value is 0.5.

Proportion 2 is the proportion for population 2, which could be known or hypothesized. The default value is 0.1.

One-Sided or Two-Sided Specify either a one-sided or a two-sided test. The default setting is the two-sided test.

Null Difference in Proportion is the proportion difference (D_0) to test against, or is left blank for computation. The default value is 0.2.

Sample Size 1 is the sample size for population 1, or is left blank for computation.

Sample Size 2 is the sample size for population 2, or is left blank for computation.

Power is the desired power, or is left blank for computation.

If you enter any two of the following three quantities, the third quantity is computed:

- **Null Difference in Proportion**
- **Sample Size 1** and **Sample Size 2**
- **Power**

Example of Determining Sample Sizes with a Two-Sided Test

As an example, suppose you are responsible for two silicon wafer assembly lines. Based on the knowledge from many runs, one of the assembly lines has a defect rate of 8%; the other line has a defect rate of 6%. You want to know the sample size necessary to have 80% power to reject the null hypothesis of equal proportions of defects for each line.

To estimate the necessary sample sizes for this example:

1. Select **DOE > Sample Size and Power**.
2. Click **Two Sample Proportions**.
3. Accept the default value of **Alpha** as 0.05.
4. Enter 0.08 for **Proportion 1**.
5. Enter 0.06 for **Proportion 2**.
6. Accept the default option of **Two-Sided**.
7. Enter 0.0 for **Null Difference in Proportion**.
8. Enter 0.8 for **Power**.
9. Leave **Sample Size 1** and **Sample Size 2** blank.

The completed window appears as the left window in Figure 14.11.

10. Click **Continue**.

The Sample Size window shows sample sizes of 2554. (See the figure on the right in Figure 14.11.) Testing for a one-sided test is conducted similarly. Simply select the **One-Sided** option.

Note: The computations for finding two sample sizes can take a little time so be patient.

Figure 14.11 Difference Between Two Proportions for a Two-Sided Test

Sample Size
Two Proportions

Testing if two proportions are different from each other.

Alpha $H_0: P_1 - P_2 = \Delta_0$

Proportion 1

Proportion 2

☒ Two-Sided
☐ One-Sided

Supply two of (difference, sample sizes, power) to determine the third.
When entering sample sizes, enter a value for both groups.

Null Difference in Proportion

Sample Size 1

Sample Size 2

Power

Sample Size
Two Proportions

Testing if two proportions are different from each other.

Alpha $H_0: P_1 - P_2 = \Delta_0$

Proportion 1

Proportion 2

☒ Two-Sided
☐ One-Sided

Supply two of (difference, sample sizes, power) to determine the third.
When entering sample sizes, enter a value for both groups.

Null Difference in Proportion

Sample Size 1

Sample Size 2

Power

Example of Determining Power with Two Sample Proportions Using a One-Sided Test

Suppose you want to compare the effectiveness of two chemical additives. The standard additive is known to be 50% effective in preventing cracking in the final product. The new additive is assumed to be 60% effective. You plan on conducting a study, randomly assigning parts to the two groups. You have 800 parts available to use for the study (400 parts for each additive). Your objective is to determine the power of your test, given a null difference in proportions of 0.01 and an alpha level of 0.05. Because you are interested in testing that the difference in proportions is greater than 0.01, you use a one-sided test.

1. Select **DOE > Sample Size and Power**.
2. Click **Two Sample Proportions**.
3. Accept the default value of **Alpha** as 0.05.
4. Enter 0.6 for **Proportion 1**.
5. Enter 0.5 for **Proportion 2**.
6. Select **One-Sided**.
7. Enter 0.01 as the **Null Difference in Proportion**.
8. Enter 400 for **Sample Size 1**.
9. Enter 400 for **Sample Size 2**.
10. Leave **Power** blank.

Figure 14.12 (left side) shows the completed Two Proportions window.

11. Click **Continue**.

Figure 14.12 (right side) shows the Two Proportions windows with the estimated **Power** calculation of 0.82.

Figure 14.12 Difference Between Two Proportions for a One-Sided Test

Sample Size
Two Proportions
Testing if two proportions are different from each other.
Alpha $H_0: P_1 - P_2 = \Delta_0$
Proportion 1
Proportion 2
☐ Two-Sided
☒ One-Sided
Supply two of (difference, sample sizes, power) to determine the third.
When entering sample sizes, enter a value for both groups.
Null Difference in Proportion
Sample Size 1
Sample Size 2
Power

Sample Size
Two Proportions
Testing if two proportions are different from each other.
Alpha $H_0: P_1 - P_2 = \Delta_0$
Proportion 1
Proportion 2
☐ Two-Sided
☒ One-Sided
Supply two of (difference, sample sizes, power) to determine the third.
When entering sample sizes, enter a value for both groups.
Null Difference in Proportion
Sample Size 1
Sample Size 2
Power

You conclude that there is about an 82% chance of rejecting the null hypothesis at the 0.05 level of significance, given that the sample sizes for the two groups are each 400.

Counts per Unit

You can use the **Counts per Unit** option from the Sample Size and Power window (Figure 14.1) to calculate the sample size needed when you measure more than one defect per unit. A unit can be an area and the counts can be fractions or large numbers.

Although the number of defects observed in an area of a given size is often assumed to have a Poisson distribution, it is understood that the area and count are large enough to support a normal approximation.

Questions of interest are:

- Is the defect density within prescribed limits?
- Is the defect density greater than or less than a prescribed limit?

In the Counts per Unit window, options include:

Alpha is the significance level of your test. The default value is 0.05.

Baseline Count per Unit is the number of targeted defects per unit. The default value is 0.1.

Difference to detect is the smallest detectable difference to test against and is specified in defects per unit, or is left blank for computation.

Sample Size is the sample size, or is left blank for computation.

Power is the desired power, or is left blank for computation.

In the Counts per Unit window, enter **Alpha** and the **Baseline Count per Unit**. Then enter two of the remaining fields to see the calculation of the third. The test is for a one-sided (one-tailed) change. Enter the **Difference to Detect** in terms of the **Baseline Count per Unit** (defects per unit). The computed sample size is expressed as the number of units, rounded to the next whole number.

Counts per Unit Example

As an example, consider a wafer manufacturing process with a target of 4 defects per wafer. You want to verify that a new process meets that target within a difference of 1 defect per wafer with a significance level of 0.05. In the Counts per Unit window:

1. Leave **Alpha** as 0.05 (the chance of failing the test if the new process is as good as the target).
2. Enter 4 as the **Baseline Counts per Unit**, indicating the target of 4 defects per wafer.
3. Enter 1 as the **Difference to detect**.
4. Enter a power of 0.9, which is the chance of detecting a change larger than 1 (5 defects per wafer). In this type of situation, alpha is sometimes called the *producer's risk* and beta is called the *consumer's risk*.
5. Click **Continue** to see the results in Figure 14.13, showing a computed sample size of 38 (rounded to the next whole number).

The process meets the target if there are less than 190 defects (5 defects per wafer in a sample of 38 wafers).

Figure 14.13 Window For Counts Per Unit Example

Sample Size
Counts per Unit
Detecting change in count per unit, e.g. DPU (defects per unit)
Alpha
Baseline Count per Unit
Supply two values to determine the third.
Difference to detect
Sample Size
Power
Using normal approximations

Sample Size
Counts per Unit
Detecting change in count per unit, e.g. DPU (defects per unit)
Alpha
Baseline Count per Unit
Supply two values to determine the third.
Difference to detect
Sample Size
Power
Using normal approximations

Sigma Quality Level

The Sigma Quality Level feature is a simple statistic that puts a given defect rate on a “six-sigma” scale. For example, on a scale of one million opportunities, 3.397 defects result in a six-sigma process. The computation that gives the Sigma Quality Level statistic is

$$\text{Sigma Quality Level} = \text{NormalQuantile}(1 - \text{defects/opportunities}) + 1.5$$

Two of three quantities can be entered to determine the Sigma Quality Level statistic in the Sample Size and Power window:

- **Number of Defects**
- **Number of Opportunities**
- **Sigma Quality Level**

When you click **Continue**, the sigma quality calculator computes the missing quantity.

Sigma Quality Level Example

As an example, use the Sample Size and Power feature to compute the Sigma Quality Level for 50 defects in 1,000,000 opportunities:

1. Select **DOE > Sample Size and Power**.
2. Click the **Sigma Quality Level** button.
3. Enter 50 for the **Number of Defects**.
4. Enter 1000000 as the **Number of Opportunities**. (See window to the left in Figure 14.14.)
5. Click **Continue**.

The results are a Sigma Quality Level of 5.39 defects in 1,000,000 opportunities. (See right window in Figure 14.14.)

Figure 14.14 Sigma Quality Level Example 1

The figure consists of two side-by-side screenshots of the 'Sample Size' dialog box. Both windows have a title bar 'Sample Size' and a subtitle 'Sigma Quality Level'. The formula 'NormalQuantile(1-defects/opportunities)+1.5' is displayed at the top. Below the formula, it says 'Supply two values to determine the third.' The left window shows 'Number of Defects' as 50 and 'Number of Opportunities' as 1000000, with 'Sigma Quality Level' left blank. The right window shows the same inputs, but 'Sigma Quality Level' is calculated and displayed as 5.3905918864, which is circled in red. Both windows have 'Continue' and 'Back' buttons at the bottom.

Field	Left Window Value	Right Window Value
Number of Defects	50	50
Number of Opportunities	1000000	1000000
Sigma Quality Level		5.3905918864

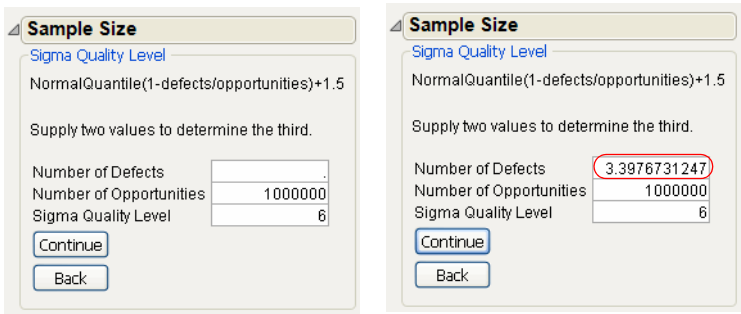
Number of Defects Computation Example

If you want to know how many defects reduce the Sigma Quality Level to “six-sigma” for 1,000,000 opportunities:

- 1. Select **DOE > Sample Size and Power**.
- 2. Click the **Sigma Quality Level** button.
- 3. Enter 6 as **Sigma Quality Level**.
- 4. Enter 1000000 as the **Number of Opportunities**. (See left window in Figure 14.14.)
- 5. Leave **Number of Defects** blank.
- 6. Click **Continue**.

The computation shows that the **Number of Defects** cannot be more than approximately 3.4. (See right window in Figure 14.15.)

Figure 14.15 Sigma Quality Level Example 2



Reliability Test Plan and Demonstration

You can compute required sample sizes for reliability tests and reliability demonstrations using the **Reliability Test Plan** and **Reliability Demonstration** features.

Reliability Test Plan

The **Reliability Test Plan** feature computes required sample sizes, censor times, or precision, for estimating failure times and failure probabilities.

To launch the Reliability Test Plan calculator, select **DOE > Sample Size and Power**, and then select **Reliability Test Plan**. Figure 14.16 shows the Reliability Test Plan window.

Figure 14.16 Reliability Test Plan Window

Sample Size

Reliability Test Plan

Calculations for designing a reliability study

Alpha

Distribution

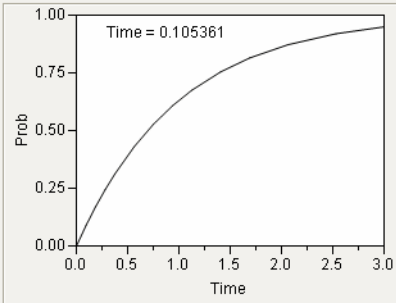
Weibull α

Weibull β

Precision Measure

Objective ☒ Estimate time associated with specified failure probability $p =$

☐ Estimate failure probability at time $t =$



Supply two values to determine the third.

Sample Size

Censor Time

Precision

Large-sample approximate covariance matrix

The Reliability Test Plan window has the following options:

Alpha is the significance level. It is also 1 minus the confidence level.

Distribution is the assumed failure distribution, with the associated parameters.

Precision Measure is the precision measure. In the following definitions, U and L correspond to the upper and lower confidence limits of the quantity being estimated (either a time or failure probability), and T corresponds to the true time or probability for the specified distribution.

Interval Ratio is $\sqrt{U/L}$, the square root of the ratio of the upper and lower limits.

Two-sided Interval Absolute Width is $U-L$, the difference of the upper and lower limits.

Lower One-sided Interval Absolute Width is $T-L$, the true value minus the lower limit.

Two-sided Interval Relative Width is $(U-L)/T$, the difference between the upper and lower limits, divided by the true value.

Lower One-sided Interval Relative Width is $(T-L)/T$, the difference between the true value and the lower limit, divided by the true value.

Objective is the objective of the study. The objective can be one of the following two:

- estimate the time associated with a specific probability of failure.
- estimate the probability of failure at a specific time.

CDF Plot is a plot of the CDF of the specified distribution. When estimating a time, the true time associated with the specified probability is written on the plot. When estimating a failure probability, the true probability associated with the specified time is written on the plot.

Sample Size is the required number of units to include in the reliability test.

Censor Time is the amount of time to run the reliability test.

Precision is the level of precision. This value corresponds to the Precision Measure chosen above.

Large-sample approximate covariance matrix gives the approximate variances and covariance for the location and scale parameters of the distribution.

Continue click here to make the calculations.

Back click here to go back to the Power and Sample Size window.

After the **Continue** button is clicked, two additional statistics are shown:

Expected number of failures is the expected number of failures for the specified reliability test.

Probability of fewer than 3 failures is the probability that the specified reliability test will result in fewer than three failures. This is important because a minimum of three failures is required to reliably estimate the parameters of the failure distribution. With only one or two failures, the estimates are unstable. If this probability is large, you risk not being able to achieve enough failures to reliably estimate the distribution parameters, and you should consider changing the test plan. Increasing the sample size or censor time are two ways of lowering the probability of fewer than three failures.

Example

A company has developed a new product and wants to know the required sample size to estimate the time till 20% of units fail, with a two-sided absolute precision of 200 hours. In other words, when a confidence interval is created for the estimated time, the difference between the upper and lower limits needs to be approximately 200 hours. The company can run the experiment for 2500 hours. Additionally, from studies done on similar products, they believe the failure distribution to be approximately Weibull (2000, 3).

To compute the required sample size, do the following steps:

1. Select **DOE > Sample Size and Power**.
2. Select **Reliability Test Plan**.
3. Select **Weibull** from the Distribution list.
4. Enter 2000 for the Weibull α parameter.
5. Enter 3 for the Weibull β parameter.
6. Select **Two-sided Interval Absolute Width** from the Precision Measure list.
7. Select **Estimate time associated with specified failure probability**.
8. Enter 0.2 for p .
9. Enter 2500 for Censor Time.
10. Enter 200 for Precision.
11. Click **Continue**. Figure 14.17 shows the results.

Figure 14.17 Reliability Test Plan Results

Sample Size	217
Censor Time	2,500
Precision	200
Expected number of failures	186.2229
Probability of fewer than 3 failures	7.4e-179

The required sample size is 217 units if the company wants to estimate the time till 20% failures with a precision of 200 hours. The probability of fewer than 3 failures is small, so the experiment will likely result in enough failures to reliably estimate the distribution parameters.

Reliability Demonstration

A reliability demonstration consists of testing a specified number of units for a specified period of time. If fewer than k units fail, you pass the demonstration, and conclude that the product reliability meets or exceeds a reliability standard.

The **Reliability Demonstration** feature computes required sample sizes and experimental run-times for demonstrating that a product meets or exceeds a specified reliability standard.

To launch the Reliability Demonstration calculator, select **DOE > Sample Size and Power**, and then select **Reliability Demonstration**. Figure 14.18 shows the Reliability Demonstration window.

Figure 14.18 Reliability Demonstration Window

Sample Size

Reliability Demonstration

Comparing product reliability to a standard

Alpha

Distribution

Weibull β

Max Failures Tolerated

Reliability Standard

Time

Probability of Surviving

Enter one value to determine the other

Time of Demonstration

Number of Units Tested

The Reliability Demonstration window has the following options:

Alpha is the alpha level.

Distribution is the assumed failure distribution. After selecting a distribution, specify the associated scale parameter in the text field under the **Distribution** menu.

Max Failures Tolerated is the maximum number of failures you want to allow during the demonstration. If we observe this many failures or fewer, then we say we passed the demonstration.

Time is the time component of the reliability standard you want to meet.

Probability of Surviving is the probability component of the reliability standard you want to meet.

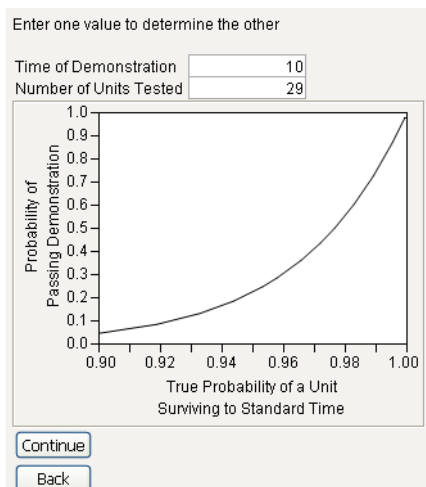
Time of Demonstration is the required time for the demonstration.

Number of Units Tested is the required number of units for the demonstration.

Continue click here to make the calculations.

Back click here to go back to the Power and Sample Size window.

After the Continue button is clicked, a plot appears (see Figure 14.19).

Figure 14.19 Reliability Demonstration Plot

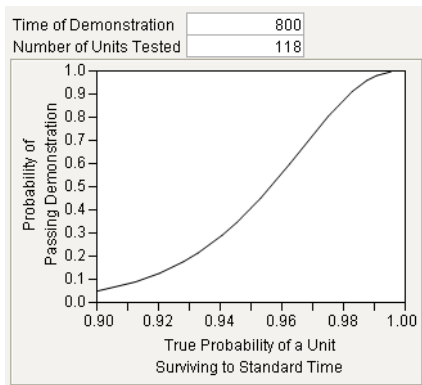
The true probability of a unit surviving to the specified time is unknown. The Y axis of the plot gives the probability of passing the demonstration (concluding the true reliability meets or exceeds the standard) as a function of the true probability of a unit surviving to the standard time. Notice the line is increasing, meaning that the further the truth is above the standard, the more likely you are to detect the difference.

Example

A company wants to get the required sample size for assessing the reliability of a new product against an historical reliability standard of 90% survival after 1000 hours. From prior studies on similar products, it is believed that the failure distribution is Weibull, with a β parameter of 3. The company can afford to run the demonstration for 800 hours, and wants the experiment to result in no more than 2 failures.

To compute the required sample size, do the following steps:

1. Select **DOE > Sample Size and Power**.
2. Select **Reliability Demonstration**.
3. Select **Weibull** from the Distribution list.
4. Enter 3 for the Weibull β .
5. Enter 2 for Max Failures Tolerated.
6. Enter 1000 for Time.
7. Enter 0.9 for Probability of Surviving.
8. Enter 800 for Time of Demonstration.
9. Click **Continue**. Figure 14.20 shows the results.

Figure 14.20 Reliability Demonstration Results

The company needs to run 118 units in the demonstration. Furthermore, if they observe 2 or fewer failures by 800 hours, we can conclude that the new product reliability is at least as reliable as the standard.

References

- Agresti, A., and Caffo, B. N. (2000), "Simple and Effective Confidence Intervals for Proportions and Differences of Proportions Result from Adding Two Successes and Two Failures," *The American Statistician*, 54, 280-288.
- Atkinson, A. C. and Donev, A. N. (1992), *Optimum Experimental Designs*, Clarendon Press, Oxford p.148.
- Bose, R.C., (1947), "Mathematical Theory of the Symmetrical Factorial Design" *Sankhya: The Indian Journal of Statistics*, Vol. 8, Part 2, pp. 107-166.
- Box, G.E.P. and Wilson, K.B. (1951), "On the Experimental Attainment of Optimum Conditions," *Journal of the Royal Statistical Society, Series B*, 13, 1-45.
- Box, G.E.P. and Behnken, D.W. (1960), "Some New Three-Level Designs for the Study of Quantitative Variables," *Technometrics* 2, 455-475.
- Box, G.E.P. and Meyer, R.D. (1986), "An Analysis of Unreplicated Fractional Factorials," *Technometrics* 28, 11-18.
- Box, G.E.P. and Draper, N.R. (1987), *Empirical Model-Building and Response Surfaces*, New York: John Wiley and Sons.
- Box, G.E.P. (1988), "Signal-to-Noise Ratio, Performance Criteria, and Transformations," *Technometrics* 30, 1-40.
- Box, G.E.P., Hunter, W.G., and Hunter, J.S. (1978), *Statistics for Experimenters*, New York: John Wiley and Sons, Inc.
- Byrne, D.M. and Taguchi, G. (1986), *ASQC 40th Anniversary Quality Control Congress Transactions*, Milwaukee, WI: American Society of Quality Control, 168-177.
- Chen, J., Sun, D.X., and Wu, C.F.J. (1993), "A Catalogue of Two-level and Three-Level Fractional Factorial Designs with Small Runs," *International Statistical Review*, 61, 1, p131-145, International Statistical Institute.
- Cochran, W.G. and Cox, G.M. (1957), *Experimental Designs*, Second Edition, New York: John Wiley and Sons.
- Cornell, J.A. (1990), *Experiments with Mixtures*, Second Edition New York: John Wiley & Sons.
- Cook, R.D. and Nachtsheim, C.J. (1990), "Letter to the Editor: Response to James M. Lucas," *Technometrics* 32, 363-364.
- Daniel, C. (1959), "Use of Half-normal Plots in Interpreting Factorial Two-level Experiments," *Technometrics*, 1, 311-314.
- Daniel C. and Wood, F. (1980), *Fitting Equations to Data*, Revised Edition, New York: John Wiley and Sons, Inc.

- Derringer, D. and Suich, R. (1980), "Simultaneous Optimization of Several Response Variables," *Journal of Quality Technology*, Oct. 1980, 12:4, 214–219.
- DuMouchel, W. and Jones, B. (1994), "A Simple Bayesian Modification of *D*-Optimal Designs to Reduce Dependence on an Assumed Model," *Technometrics*, 36, 37–47.
- Fang, K.T. and Sudhianto, A. (2006), *Design and Modeling for Computer Experiments*, Chapman and Hall CRC, Boca Ratan. Florida
- Goos, P. (2002) *The Optimal Design of Blocked and Split Plot Experiments*, New York: Springer
- Gotwalt, C. M., Jones, B.A., and Steinberg, D. M., (2009) "Fast Computation of Designs Robust to Parameter Uncertainty for Nonlinear Settings," *Technometrics*, 51:1, 88-95.
- Haaland, P.D. (1989), *Experimental Design in Biotechnology*, New York: Marcel Dekker, Inc.
- Hahn, G. J., Meeker, W.Q., and Feder, P. I., (1976), "The Evaluation and Comparison of Experimental Designs for Fitting Regression Relationships," *Journal of Quality Technology*, Vol. 8, #3, pp. 140-157.
- John, P.W.M. (1972), *Statistical Design and Analysis of Experiments*, New York: Macmillan Publishing Company, Inc.
- Johnson, M.E. and Nachtsheim, C.J. (1983), "Some Guidelines for Constructing Exact *D*-Optimal Designs on Convex Design Spaces," *Technometrics* 25, 271–277.
- Jones, Bradley (1991), "An Interactive Graph For Exploring Multidimensional Response Surfaces," 1991 Joint Statistical Meetings, Atlanta, Georgia.
- Jones, Bradley and Goos, Peter (2007), "A Candidate-set-free Algorithm for Generating *D*-optimal Designs," *Applied Statistics*, 56, Part 3, pp 1-18.
- Khuri, A.I. and Cornell, J.A. (1987) *Response Surfaces: Design and Analysis*, New York: Marcel Dekker.
- Kowalski, S.M., Cornell, J.A., and Vining, G.G. (2002) "Split Plot Designs and Estimation Methods for Mixture Experiments with Process Variables," *Technometrics* 44: 72-79.
- Lenth, R.V. (1989), "Quick and Easy Analysis of Unreplicated Fractional Factorials," *Technometrics*, 31, 469–473.
- Lin, D. K. J. (1993), "A New Class of Supersaturated Design," *Technometrics*, 35, 28-31.
- Lucas, J.M., (1990), "Letter to the Editor: Comments on Cook and Nachtsheim (1989)," *Technometrics*, 32, 363–364.
- Mahalanobis, P.C. (1947), *Sankhya, The Indian Journal of Statistics*, Vol 8, Part 2, April.
- Myers, R.H. (1976) *Response Surface Methodology*, Boston: Allyn and Bacon.
- Myers, R.H. (1988), *Response Surface Methodology*, Virginia Polytechnic and State University.
- Meyer, R.K. and Nachtsheim, C.J. (1995), "The Coordinate Exchange Algorithm for Constructing Exact Optimal Designs," *Technometrics*, Vol 37, pp. 60-69.
- Meyer, R.D., Steinberg, D.M., and Box, G. (1996), "Follow-up Designs to Resolve Confounding in Multifactor Experiments," *Technometrics*, 38:4, p307.
- Mitchell, T.J. (1974), "An algorithm for the Construction of *D*-Optimal Experimental Designs," *Technometrics*, 16:2, pp.203-210.
- Montgomery, D.C. (2009), *Design and Analysis of Experiments*, Seventh Edition, Hoboken: John Wiley & Sons, Inc.

- Morris, M.D., Mitchell, T.J., and Ylvisaker, D. (1993), "Bayesian Design and Analysis of Computer Experiments: Use of Derivatives in Surface Prediction," *Technometrics* 35:2, 243-255.
- Myers, R.H., Montgomery, D.C., and Anderson-Cook, C.M. (2009), *Response Surface Methodology, Process and Product Optimization Using Designed Experiments*, Third Edition, Hoboken: John Wiley & Sons, Inc.
- Piepel, G.F. (1988), "Programs for Generating Extreme Vertices and Centroids of Linearly Constrained Experimental Regions," *Journal of Quality Technology* 20:2, 125-139.
- Plackett, R.L. and Burman, J.P. (1947), "The Design of Optimum Multifactorial Experiments," *Biometrika*, 33, 305-325.
- Santner, T.B., Williams, B., and Notz, W. (2003), *The Design and Analysis of Computer Experiments*, Springer, New York.
- Schoen, E. D. (1999), "Designing Fractional Two-level Experiments with Nested Error Structures," *Journal of Applied Statistics* 26: 495-508.
- Shannon, C.E. and Weaver, W. (1949), "The Mathematical Theory Of Communication," University Of Illinois Press, Urbana, Illinois.
- Scheffé, H. (1958) "Experiments with Mixtures", *Journal of the Royal Statistical Society B* v20, 344-360.
- Shewry, M.C. and Wynn, H.P. (1987), "Maximum Entropy Sampling," *Journal of Applied Statistics* 14, 165-170.
- Snee, R.D. and Marquardt, D.W. (1974), "Extreme Vertices Designs for Linear Mixture Models," *Technometrics*, 16, 391-408.
- Snee, R.D. (1975), "Experimental Designs for Quadratic Models in Constrained Mixture Spaces," *Technometrics*, 17:2, 149-159.
- Snee, R.D. (1979), "Experimental Designs for Mixture Systems with Multicomponent Constraints," *Commun. Statistics*, A8(4), 303-326.
- Snee, Ronald D. (1985) *Computer Aided Design of Experiments - Some Practical Experiences*, Journal of Quality Technology, Vol 17. No. 4 October 1985 p.231.
- St. John, R.C. and Draper, N.R. (1975), "D-Optimality for Regression Designs: A Review," *Technometrics*, 17 pp 15-23.
- Taguchi, G. (1976), "An Introduction to Quality Control," Nagoya, Japan: Central Japan Quality Control Association.
- Welch, W.J. (1984), "Computer-Aided Design of Experiments for Response Estimation," *Technometrics*, 26, 217-224.
- Zahran, A., Anderson-Cook, C.M., and Myers, R.H. (2003), "Fraction of Design Space to Assess the Prediction Capability of Response Surface Designs," *Journal of Quality Technology*, 35, 377-386.

Index

Design of Experiments

A

ABCD designs 168
acceptable values *See* lower limits and upper limits
Actual-by-Predicted plots 125
adding
 center points in augment designs 275, 277
 factors 40, 67
 linear constraints 164
 runs and terms in augment designs 279
additional runs 78, 120, 140, 152
advanced options (nonlinear designer) 245
algorithms, coordinate exchange 99
alias matrix 76
 explanation 18
 showing 18
aliasing effects 115
Alias-optimal designs 60, 88
Alpha 288
animation scripts 289
Anova reports 171
assigning importance (of responses) 65, 109, 254
assigning importance of responses 84
attribute, discrete choice designs 181
augment nonlinear design table 244
Augmented Designs 259–284
 add centerpoints 272, 275
 add new runs and terms 272, 278
 adding axial points 272, 277
 disallow factor combinations 284
 D-optimal 278
 extending experiments 271
 foldover 272, 276
 modify D- or I- optimality criterion 282
 replicate 271–272
 saved design (X) matrix 282
 select number of random starts 283
 special commands 281
 specify sphere radius value 283
axial
 points 127
 scaling, central composite designs 138

B

balanced designs 71
Bayesian D-optimal designs 31, 60–61
Bayesian I-optimal designs 60, 62
Big Class.jmp 51
block sizes 27
Borehole Factors.jmp 212
Borehole Sphere Packing.jmp 213
Bounce Data.jmp 129
Bounce Factors.jmp 129–130
Box-Behnken designs 127, 129, 138
 See also Response Surface designs
Byrne Taguchi Factors.jmp 250–251

C

categorical factors 68, 111
CCD *See* central composite designs
center points
 augment designs 275, 277
 central composite designs 127
 response surface designs 127
 simplex centroid designs 158
central composite designs 127, 138
 See also response surface designs
centroid points 163
Chakravarty 114
Chemical Kinetics.jmp 231
chemical mixture, examples 171
choice designs
 example 179
 purpose 179
choice set, discrete choice designs 181
coding, column property 94
coefficients, relative variance 75
column properties
 coding 94
 constrained state 85
 design roles 84, 97
 factors for mixture experiments 95
 high and low values 94

- response limits 96
- combinations, disallowing 90, 283
- CONAEV method 167
- confounding 116
 - resolution numbers 112
- confounding pattern 76, 116
- constraints
 - adding 164
 - disallowing combinations 90, 283
 - entering 85
 - linear 166
 - loading 85
 - saving 85
- continuous factors 67, 111
- control factors 247, 249
- CONVRT method 167
- coordinate exchange algorithms 99
- Corn.jmp 241
- Cotter designs 114
- counts per unit (power and sample size) 303
- covariate factors 51
- creating factors tables 85
- criterion, optimality 88, 282
- custom design commands 83
- custom designs
 - data tables 79
 - Design Generation panel 7, 71
 - factors, defining 65
 - flexible block size 27
 - how they work 99
 - models, describing 69
 - random block design 79
 - screening 17
 - special commands 83
 - split plot 80
 - split-split plot 81
 - strip plot 82

D

- defaults
 - number of random starts 88
- defects 303
- D-efficiencies 43
- describing models 69, 158
- design
 - matrix table properties 87
 - Moments Matrix 87
 - resolutions 112
 - roles 84

- Save X Matrix 87
- design diagnostics table 206
- Design of Experiments. *See* DOE
- design roles 97
- designs
 - ABCD 168
 - aberration 112
 - augment 271
 - balanced 71
 - Bayesian D-optimal 31, 61
 - Bayesian I-optimal 62
 - Box-Behnken 127, 129, 138
 - central composite 127
 - discrete choice 179
 - foldover 276
 - fractional factorial 112
 - full factorial 101, 143, 151
 - full factorials 112
 - Gaussian process IMSE optimal space filling 211
 - I-optimal 61
 - Latin hypercube space filling design 200–204
 - maximum entropy space filling 209
 - minimum aberration 112
 - minimum potential space filling 207
 - mixed-level 113
 - mixture 155, 171
 - nonlinear 231
 - orthogonal
 - screening designs 112
 - screening experiments 155
 - surface designs 138
 - orthogonal arrays 113
 - Plackett-Burman 113
 - random block 79
 - replicating 272–273
 - response surface 127
 - saturated 71
 - screening 101
 - simplex centroid 158
 - simplex centroids 158
 - simplex lattice 161
 - space filling 197–216
 - sphere packing space filling 197
 - split plot 80
 - split-split plot 81
 - strip plot 82
 - uniform precision 138
 - uniform space filling 204
- desirability
 - functions 133, 253

- maximizing 133
- values 66, 110
- desirability traces 133
- desirability values, as column properties 96
- determinant formula 89
- diagnostics for custom designs 86, 282
- Difference to Detect** option 288–289, 293, 296
- disallowing combinations 90, 283
- discrete choice designs
 - analysis 186, 190
 - attribute 181
 - attributes panel 183
 - choice set 181
 - definition 179
 - design generation panel 184
 - dialog 183
 - example 181, 183–184
 - prior information 188
 - profile 181
 - survey 181
- DOE (Design of Experiments)
 - coding 94
 - simple examples 103
- DOE Sphere Radius 283
- Donev Mixture Factors.jmp 44
- D-optimal
 - augmentation 278
 - designs 31, 60–61
 - optimality criteria 88, 282

E

- effect
 - aliasing 116
 - eigenvalue 132
 - eigenvector 132
 - nonestimable 112
 - orthogonal 139, 247
 - sparsity 101, 113–114
- efficiencies
 - D, G, and A 77
- eigenvalue of effect 132
- eigenvector of effect 132
- equivalent solutions 99
- error standard deviation 288
- error variance 36
- excluding factor combinations 90
- extra parameters 288
- extreme vertices 163
 - algorithms 167

- finding optimal subsets 164
- range constraints 164

F

- Factor Changes 84
- factor combinations, disallowing 90
- Factor Profiling** option 132, 134, 150, 175, 253
- factorial designs
 - fractional 112
 - full 101, 112, 143, 151
 - three level 113
- factors
 - adding 40, 67
 - categorical 68, 111
 - column 98
 - continuous 67, 111
 - control factors 247, 249
 - covariate 51
 - for mixture experiments 95
 - key factors 101
 - loading 85
 - nonmixture 44
 - saving 84
 - tables, creating 85
- false negatives 114
- finding optimal subsets (extreme vertices) 164
- fitting mixture designs 170
- fixed covariate factors 51
- fixed covariates example 51
- folding the calculation 171
- foldover designs 276
- fraction of design space plot 74
- fractional factorial designs 112
- Full Factorial Designs 143–153
 - design generation panel 71
 - examples 145
- full factorial designs 112, 143
- functions, desirability 133, 253

G

- Gaussian process IMSE optimal design 195, 211
- global optimum 89
- goal types 65, 109
 - matching targets 66, 110
 - minimizing and maximizing 66, 110
- Group New Runs into Separates Block** option 275

H

hidden intercepts 171
hyperspheres 90, 283

I

identifying key factors 101
importance of responses 65, 84, 109, 254
 entering 96
industrial experimentation 179
inner arrays, inner designs 247, 249
Inscribe option 139
interactions 114
 high-order 112
intercepts, hidden 171
I-optimal
 designs 60–62
 optimality criteria 88, 282
I-optimal designs 61

J

JSL (JMP Scripting Language)
 animation scripts 289
 augmented designs 265
 sphere radius 283

K

Keep the Same 78
k-Sample Means (power and sample size) 294

L

L18 Chakravarty 114
L18 Hunter 114
L18 John 113
L36 114
L9-Taguchi 250
Label column 120, 140
larger-the-better formulas (LTB) 251
Latin hypercube space filling design 200–204
limits, responses 96
linear constraints 164, 166
loading
 constraints 85
 factors 85
 responses 85
local optimum 89

M

matching target goals 65–66, 96, 109–110
matrix, alias 76
matrix, design 87
maximize responses 65, 109
maximizing
 desirability 133
 goals 66, 110
maximum entropy space filling design 195, 209
means, one- and two-sample (power and sample size) 287
methods
 CONAEV 167
minimize responses 65, 109
minimizing goals 66, 110
minimum aberration designs 112
minimum potential design 207
minimum potential space filling design 195
mixed-level designs 113
Mixture Designs 155–179
 ABCD design 168
 compared to screening designs 155
 definition 155
 examples 171–174
 extreme vertices 163
 fitting 170
 linear constraints 166
 optimal 157
 optimal subsets 164
 response surfaces 175
 simplex centroids 158
 simplex lattice 161
 ternary plots 168
 with nonmixture factors 44
mixture designs
 examples 171
mixture experiments (column property) 95
mixture, column properties 95
Model script
 Model Specification dialog 123, 130
models
 custom design 69
 describing 69, 158

N

N factors, adding 67
N Monte Carlo Spheres (nonlinear design option) 245
N Responses 65

N Responses button 109

no-intercept model 95

noise factors 247, 249

nonestimable effect 112

nonlinear design

advanced options 245

augment table 236, 244

create using prior parameter estimates 233

creating 241

creating with no prior data 237

distribution of parameters 238

fitting to find prior parameter estimates 231

introduction 229

launch dialog 234, 241–242

number of Monte Carlo samples 235

number of starts 235

overlay plot of design points 239

using prior parameter estimates 233

Nonlinear Designs 229–246

advanced options 245

creating 241

examples 231

nonmixture factors in mixture designs 44

Number of Monte Carlo Samples (nonlinear design option) 245

number of runs 15, 71

screening designs 112

number of starts 88, 283

O**On Face** option 139

one-sample and two-sample means (power and sample size) 287

one-sample and two-sample proportion (power and sample size) 298

one-sample variance (power and sample size) 296

optimal determinant formula 89

optimal subsets (mixture designs) 164

Optimality Criterion 88, 282

optimality criterion 88

order of runs 78, 120, 140, 152, 159, 164, 168

orthogonal array designs 113, 247

orthogonal designs

screening designs 112

screening experiments 155

surface designs 138

Orthogonal option 139

outer arrays, outer designs 247, 249

P

parameters, extra 288

Pattern column 105, 120, 140, 153

patterns, confounding 116

performance statistics 249

Plackett-Burman design

creating 120

Plackett-Burman designs 113

Plastifactors.jmp 171

plots

Actual-by-Predicted 125

Scatterplot 3D 136

ternary 155, 168

points

axial 127

center *See* center points

centroid 163

potential terms (DOE) 31

power and sample size calculations 285–306

analyses 287

animation 292

counts per unit 303

in statistical tests on means 296

k-sample means 294

one-sample and two sample proportions 298

one-sample mean 288–289

one-sample variance 296, 298, 300

Sigma quality level 305

two-sample means 293

prediction

profilers 35, 38

traces 132

variances 36, 138

prediction formulas, saving 174

prediction profiler 132

prediction variance profile plot 35

prediction variance surface plot 75

primary terms (Bayesian D-optimal design) 31

profile, discrete choice designs 181

profilers

mixture response surface 175

prediction profilers 38, 132

prediction variance plot 35

prediction Variance profiler 35

properties, columns 93

proportions (power and sample size) 298

prospective power and sample size 285

prospective power and sample size calculations (see power and sample size) 306

pseudocomponent (mixture column property) 95

pure error 275

Q

quadratic model 34–76

R

radius, sphere 90, 283

random block designs 79

random seed, setting 85

random starts 88

Randomize within Blocks 78

randomizing runs 78

range constraints 163

Reaction Kinetics Start.jmp 237

Reaction Kinetics.jmp 239

Reactor 8 Runs.jmp 261

Reactor Factors.jmp 145–146

Reactor Response.jmp 145

regressor columns 114

relative proportions *See* mixture designs

relative variance of coefficients 75

relative variance of coefficients and power table 30
replicating

designs 273

replicating designs 272

reponses, loading 85

requesting additional runs 78, 120, 140, 152

rescaling designs 139

resolution numbers 112

resolutions of designs 112

response limits, column property 96

Response Surface Designs 127–141

examples 129–136

introduction 136

purpose 127

reports 131

with blocking factors 40

with categorical factors 40

response surface effects 171

Response Surface Methodology (RSM) 136

response surfaces

mixture designs 175

responses

custom designs 65, 109, 137

desirability values 66, 110

goals 65, 109

goals, desirability functions 133

lower limits 66, 110

saving 84

upper limits 66, 110

responses, multiple 96

RMSE 125, 149

robust engineering 247

roles, design 97

Rotatable option 139

RSM (Response Surface Methodology) 136

runs

additional 78, 120, 140, 152

order they appear in table 78, 120, 140, 152, 159,
164, 168

requesting additional 78, 120, 140, 152

screening designs 112

S

sample means (power and sample size) 287

sample sizes

example comparing single-direction one-sample
variances 296, 298

example with counts per unit 303

one and two sample means 289

prospective power analysis 296

screening designs 143

saturated designs 71

saving

constraints 85

factors 84

prediction formulas 174

responses 84

X Matrix 87, 282

scaling

axial 138

designs 139

Scatterplot 3D

Box-Behnken designs 136

Scheffé polynomial 170

screening designs 101

custom designs 17

design types 111

dialogs 106, 120

examples 105–140

scripts

animation 289

generating the analysis model 123, 130

Model script *See* Model table property

scripting *See* JSL

Set Random Seed command 85

sigma quality level (power and sample size) 305

signal factors 247

signal-to-noise ratios 247
 simplex 155
 centroid designs 158
 lattice designs 161
 single-sample means (power and sample size) 289
 solutions, equivalent 99
Sort Left to Right options 78
Sort Right to Left option 78
 Space Filling Designs 197–216
 borehole problem, sphere packing example 212
 Gaussian process IMSE optimal 195, 211
 Latin hypercube 195, 200–204
 maximum entropy 195, 209
 minimum potential 195, 207
 sphere packing 195, 197
 uniform 195, 204
 sparsity, effect 101, 113–114
 sphere packing design 195, 197
 sphere radius 90, 283
 split plot designs 68, 80
 split-split plot designs 81
 star points 127
 starts, number of 88, 283
 starts, random 88
 statistics, performance 249
 strip plot designs 82
 subsets, finding optimal 164
 supersaturated designs 23
 surface designs *See* response surface designs
 survey, discrete choice designs 181

T

tables
 factors table, creating 85
 making in custom designs 79
 Taguchi designs 247–254
 description 247
 methods 247
 target values 66, 96, 110
 ternary plots 155, 168
 three-dimensional scatterplot
 Box-Behnken designs 136
 traces, desirability 133
 trade-off in screening designs 112
 tutorial examples
 augment designs 261–271
 custom design 1–13
 DOE 103
 flexible block size 27

 full factorial designs 145
 mixture designs 171–174
 response surface designs 129, 134, 136
 screening designs 105
 two-level categorical 104
 two-level fractional factorials 112
 two-level full factorials 112
 two-sample and one-sample means 293
 two-sample means (power and sample size) 287
 two-sample proportion (power and sample size) 298

U

Uniform (space filling design) 195
 uniform (space filling design) 204
 uniform precision designs 138
User Defined option 139

V

values
 high and low for columns (in DOE) 94
 target 66, 110
 variance of prediction 139
 variance, error and prediction 36
 vertices
 extreme 163
 extreme, finding optimal subsets 164

W-Z

weight, importance of response 65, 84, 97, 109, 254
 whole model test 171
 X Matrix, saving 87, 282
 XVERT method 167

