JMP

Release 8

Design of Experiments

"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

8.0.1

The correct bibliographic citation for this manual is as follows: SAS Institute Inc. 2008. *JMP*[®] 8 *Design of Experiments Guide*. Cary, NC: SAS Institute Inc.

JMP[®] 8 Design of Experiments Guide

Copyright © 2008, SAS Institute Inc., Cary, NC, USA

ISBN 978-1-59994-921-5

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.

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, November 2008

2nd printing March 2009

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.

Contents

Design of Experiments

Introduction to Designing Experiments 1 A Beginner's Tutorial About Designing Experiments 3 My First Experiment 3 Step 4: Determine the Number of Runs 7 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 18 A Compromise Design Between Main Effects Only and All Interactions 20 Creating 'Super' Screening Designs 21 Screening Designs with Flexible Block Sizes 25 Checking for Curvature Using One Extra Run 28 Creating Response Surface Experiments 31 Exploring the Prediction Variance Surface 31 Introducing *I*-Optimal Designs for Response Surface Modeling 34 A Three-Factor Response Surface Design 35 Response Surface with a Blocking Factor 36 Creating Mixture Experiments 40 Mixtures Having Nonmixture Factors 40 Experiments that are Mixtures of Mixtures 43

Creating a Custom Design Enter Responses and Factors into the Custom Designer Describe the Model Select the Number of Runs Understanding Design Evaluation Specify Output Options Make the JMP Design Table Creating Random Block Designs Creating Split Plot Designs Creating Split Plot Designs Creating Split Plot Designs Creating Split Plot Designs Creating Strip Plot Designs Special Custom Design Commands Save Responses and Save Factors Load Responses and Load Factors Save Constraints and Load Constraints Set Random Seed: Setting the Number Generator Simulate Responses Save X Matrix: Viewing the Design (X) Matrix in the Log Optimality Criterion: Changing the Number of Random Starts Sphere Radius: Constraining a Design to a Hypersphere Disallowed Combinations: Accounting for Factor Level Restrictions Advanced Options for the Custom Designer Save Script to Script Window Assigning Column Properties Define Low and High Values (DOE Coding) for Columns Set Columns as Factors for Mixture Experiments Define Response Column Values Assign Columns a Design Role		Building Custom Designs The Basic Steps
Enter Responses and Factors into the Custom Designer		
Describe the Model		
Select the Number of Runs Understanding Design Evaluation Specify Output Options Make the JMP Design Table Creating Random Block Designs Creating Split Plot Designs Creating Split Plot Designs Creating Strip Plot Designs Special Custom Design Commands Save Responses and Save Factors Load Responses and Load Factors Save Constraints and Load Constraints Set Random Seed: Setting the Number Generator Simulate Responses Save X Matrix: Viewing the Design (X) Matrix in the Log Optimality Criterion: Changing the Design Criterion (D- or I- Optimality) Number of Starts: Changing the Number of Random Starts Sphere Radius: Constraining a Design to a Hypersphere Disallowed Combinations: Accounting for Factor Level Restrictions Advanced Options for the Custom Designer Save Script to Script Window Assigning Column Properties Define Low and High Values (DOE Coding) for Columns Set Columns a Factors for Mixture Experiments Define Response Column Values Assign Columns a Design Role		Describe the Model
Understanding Design Evaluation		Select the Number of Runs
Specify Output Options		
Make the JMP Design Table		
Creating Random Block Designs		
Creating Split Plot Designs	(
Creating Strip Plot Designs		
Creating Strip Plot Designs	(Creating Split-Split Plot Designs
Special Custom Design Commands Save Responses and Save Factors Load Responses and Load Factors Save Constraints and Load Constraints Set Random Seed: Setting the Number Generator Simulate Responses Save X Matrix: Viewing the Design (X) Matrix in the Log Optimality Criterion: Changing the Design Criterion (D- or I- Optimality) Number of Starts: Changing the Number of Random Starts Sphere Radius: Constraining a Design to a Hypersphere Disallowed Combinations: Accounting for Factor Level Restrictions Advanced Options for the Custom Designer Save Script to Script Window Assigning Column Properties Define Low and High Values (DOE Coding) for Columns Set Columns as Factors for Mixture Experiments Define Response Column Values Assign Columns a Design Role		
Load Responses and Load Factors Save Constraints and Load Constraints Set Random Seed: Setting the Number Generator Simulate Responses Save X Matrix: Viewing the Design (X) Matrix in the Log Optimality Criterion: Changing the Design Criterion (<i>D</i> - or <i>I</i> - Optimality) Number of Starts: Changing the Number of Random Starts Sphere Radius: Constraining a Design to a Hypersphere Disallowed Combinations: Accounting for Factor Level Restrictions Advanced Options for the Custom Designer Save Script to Script Window Assigning Column Properties Define Low and High Values (DOE Coding) for Columns Set Columns as Factors for Mixture Experiments Define Response Column Values Assign Columns a Design Role		
Load Responses and Load Factors Save Constraints and Load Constraints Set Random Seed: Setting the Number Generator Simulate Responses Save X Matrix: Viewing the Design (X) Matrix in the Log Optimality Criterion: Changing the Design Criterion (<i>D</i> - or <i>I</i> - Optimality) Number of Starts: Changing the Number of Random Starts Sphere Radius: Constraining a Design to a Hypersphere Disallowed Combinations: Accounting for Factor Level Restrictions Advanced Options for the Custom Designer Save Script to Script Window Assigning Column Properties Define Low and High Values (DOE Coding) for Columns Set Columns as Factors for Mixture Experiments Define Response Column Values Assign Columns a Design Role		
Set Random Seed: Setting the Number GeneratorSimulate ResponsesSave X Matrix: Viewing the Design (X) Matrix in the LogOptimality Criterion: Changing the Design Criterion (D- or I- Optimality)Number of Starts: Changing the Number of Random StartsSphere Radius: Constraining a Design to a HypersphereDisallowed Combinations: Accounting for Factor Level RestrictionsAdvanced Options for the Custom DesignerSave Script to Script WindowAssigning Column PropertiesDefine Low and High Values (DOE Coding) for ColumnsSet Columns as Factors for Mixture ExperimentsDefine Response Column ValuesAssign Columns a Design Role		Load Responses and Load Factors
Simulate Responses		Save Constraints and Load Constraints
Save X Matrix: Viewing the Design (X) Matrix in the Log		Set Random Seed: Setting the Number Generator
Optimality Criterion: Changing the Design Criterion (<i>D</i> - or <i>I</i> - Optimality)Number of Starts: Changing the Number of Random StartsSphere Radius: Constraining a Design to a HypersphereDisallowed Combinations: Accounting for Factor Level RestrictionsAdvanced Options for the Custom DesignerSave Script to Script WindowAssigning Column PropertiesDefine Low and High Values (DOE Coding) for ColumnsSet Columns as Factors for Mixture ExperimentsDefine Response Column ValuesAssign Columns a Design Role		Simulate Responses
Number of Starts: Changing the Number of Random Starts Sphere Radius: Constraining a Design to a Hypersphere Disallowed Combinations: Accounting for Factor Level Restrictions Advanced Options for the Custom Designer Save Script to Script Window Assigning Column Properties Define Low and High Values (DOE Coding) for Columns Set Columns as Factors for Mixture Experiments Define Response Column Values Assign Columns a Design Role		Save X Matrix: Viewing the Design (X) Matrix in the Log
Sphere Radius: Constraining a Design to a Hypersphere Disallowed Combinations: Accounting for Factor Level Restrictions Advanced Options for the Custom Designer Save Script to Script Window Assigning Column Properties Define Low and High Values (DOE Coding) for Columns Set Columns as Factors for Mixture Experiments Define Response Column Values Assign Columns a Design Role		Optimality Criterion: Changing the Design Criterion (<i>D</i> - or <i>I</i> - Optimality)
Disallowed Combinations: Accounting for Factor Level Restrictions Advanced Options for the Custom Designer Save Script to Script Window Assigning Column Properties Define Low and High Values (DOE Coding) for Columns Set Columns as Factors for Mixture Experiments Define Response Column Values Assign Columns a Design Role		Number of Starts: Changing the Number of Random Starts
Advanced Options for the Custom Designer Save Script to Script Window Assigning Column Properties Define Low and High Values (DOE Coding) for Columns Set Columns as Factors for Mixture Experiments Define Response Column Values Assign Columns a Design Role		Sphere Radius: Constraining a Design to a Hypersphere
Save Script to Script WindowAssigning Column PropertiesDefine Low and High Values (DOE Coding) for ColumnsSet Columns as Factors for Mixture ExperimentsDefine Response Column ValuesAssign Columns a Design Role		Disallowed Combinations: Accounting for Factor Level Restrictions
Assigning Column Properties Define Low and High Values (DOE Coding) for Columns Set Columns as Factors for Mixture Experiments Define Response Column Values Assign Columns a Design Role		Advanced Options for the Custom Designer
Define Low and High Values (DOE Coding) for ColumnsSet Columns as Factors for Mixture ExperimentsDefine Response Column ValuesAssign Columns a Design Role		Save Script to Script Window
Set Columns as Factors for Mixture ExperimentsDefine Response Column ValuesAssign Columns a Design Role	1	Assigning Column Properties
Define Response Column Values		Define Low and High Values (DOE Coding) for Columns
Assign Columns a Design Role		Set Columns as Factors for Mixture Experiments
		Define Response Column Values
Identify Factor Changes Column Property		Assign Columns a Design Role
		Identify Factor Changes Column Property

89Screening Design Examples91Using Two Continuous Factors and One Categorical Factor91Using Five Continuous Factors93

Create a Fractional Factorial Design
Enter Responses
Enter Factors
Choose a Design
Display and Modify a Design 102
Specify Output Options 105
View the Design Table 106
Create a Plackett-Burman design 106
Analysis of Screening Data
Using the Screening Analysis Platform 109
Using the Fit Model Platform 110

5 Response Surface Designs

	115
A Box-Behnken Design: The Tennis Ball Example	117
The Prediction Profiler	119
A Response Surface Plot (Contour Profiler)	121
Geometry of a Box-Behnken Design	122
Creating a Response Surface Design	123
Enter Responses and Factors	124
Choose a Design	124
Specify Output Options	126
View the Design Table	126

6 Full Factorial Designs

	129
The Five-Factor Reactor Example	131
Analyze the Reactor Data	132
Creating a Factorial Design	136
Enter Responses and Factors	136
Select Output Options	137
Make the Table	138

7 Mixture Designs

	139
Mixture Design Types	141
The Optimal Mixture Design	141
The Simplex Centroid Design	142
Creating the Design	142
Simplex Centroid Design Examples	143
The Simplex Lattice Design	145
The Extreme Vertices Design	147

8 Discrete Choice Designs

•••••••••••••••••••••••••••••••••••••••	163
Introduction	165
Create an Example Choice Experiment	166
Analyze the Example Choice Experiment	169
Design a Choice Experiment Using Prior Information	171
Administer the Survey and Analyze Results	173
Initial Choice Platform Analysis	1 74
Find Unit Cost and Trade Off Costs with the Profiler	175

9 Space-Filling Designs

I79
Introduction to Space-Filling Designs 181
Sphere-Packing Designs 181
Creating a Sphere-Packing Design 181
Visualizing the Sphere-Packing Design 183
Latin Hypercube Designs 184
Creating a Latin Hypercube Design 184
Visualizing the Latin Hypercube Design 185
Uniform Designs 186
Comparing Sphere-Packing, Latin Hypercube, and Uniform Methods 187
Minimum Potential Designs 189
Maximum Entropy Designs 190
Gaussian Process IMSE Optimal Designs 192
Borehole Model: A Sphere-Packing Example 192

Create the Sphere-Packing Design for the Borehole Data	193
Guidelines for the Analysis of Deterministic Data	194
Results of the Borehole Experiment	195

10 Nonlinear Designs

201
201
205
208
208
209
210
211
211
2 2 2 2 2 2 2 2

11 Taguchi Designs

	213
The Taguchi Design Approach	215
Taguchi Design Example	215
Analyze the Data	217
Creating a Taguchi Design	219
Detail the Response and Add Factors	219
Choose Inner and Outer Array Designs	220
Display Coded Design	221
Make the Design Table	221

12 Augmented Designs

•••••••••••••••••••••••••••••••••••••••	223
A D-Optimal Augmentation of the Reactor Example	225
Analyze the Augmented Design	227
Creating an Augmented Design	233
Replicate a Design	233
Add Center Points	235
Creating a Foldover Design	236
Adding Axial Points	237
Adding New Runs and Terms	238
Special Augment Design Commands	24I
Save the Design (X) Matrix	242
Modify the Design Criterion (<i>D</i> - or <i>I</i> - Optimality)	242
Select the Number of Random Starts	243
Specify the Sphere Radius Value	243

v

Disallow Factor Combinations		243
-------------------------------------	--	-----

13 Prospective Power and Sample Size

Prospective Power Analysis 247
One-Sample and Two-Sample Means 248
Single-Sample Mean 249
Power and Sample Size Animation for a Single Sample
Two-Sample Means 251
k-Sample Means 252
One Sample Standard Deviation 253
One-Sample and Two-Sample Proportions 254
Counts per Unit 258
Sigma Quality Level
Index
Design of Experiments

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, Craige Hales, Chris Gotwalt, Paul Nelson, Xan Gregg, Jianfeng Ding, Eric Hill, John Schroedl, Laura Lancaster, Scott McQuiggan, and Peng Liu.

In the SAS Institute Technical Support division, Wendy Murphrey and Toby Trott provide technical support and conduct test site administration. Statistical technical support is provided by Craig DeVault, Duane Hayes, Elizabeth Edwards, Kathleen Kiernan, and Tonya Mauldin.

Nicole Jones, Jim Borek, Kyoko Keener, Hui Di, Joseph Morgan, Wenjun Bao, Fang Chen, Susan Shao, Hugh Crews, Yusuke Ono and Kelci Miclaus provide ongoing quality assurance. Additional testing and technical support is done by Noriki Inoue, Kyoko Takenaka, and Masakazu Okada from SAS Japan.

Bob Hickey is the release engineer.

The JMP manuals were written by Ann Lehman, Lee Creighton, John Sall, Bradley Jones, Erin Vang, Melanie Drake, and Meredith Blackwelder, 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 Yurii 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 Schmee 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

Erin Vang, Trish O'Grady, and Kyoko Keener coordinate 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; Professors Fengshan Bai, Xuan Lu, and Jianguo Li at Tsinghua University in Beijing, and their assistants Rui Guo, Shan Jiang, Zhicheng Wan, and Qiang Zhao; and William Zhou (SAS China) and Zhongguo Zheng, professor at Peking University, for reviewing the Simplified Chinese translation; Jacques Goupy (consultant, ReConFor) and Olivier Nuñez (professor, Universidad Carlos III de Madrid) for reviewing the French translation; Dr. Byung Chun Kim (professor, Korea Advanced Institute of Science and Technology) and Duk-Hyun Ko (SAS Korea) for reviewing the German translation; Patrizia Omodei, Maria Scaccabarozzi, and Letizia Bazzani (SAS Italy) for reviewing the Italian translation. Finally, thanks to all the members of our outstanding translation 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, Jeff Polzin, 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

The ImageMan DLL is used with permission of Data Techniques, Inc.

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 MER-CHANTABILITY AND FITNESS, IN NO EVENT SHALL NEIL HODGSON BE LIABLE FOR ANY SPECIAL, INDI-RECT 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 SOFT-WARE.

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.



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
The Situation
Step 1: Design the Experiment 3
Step 2: Define Factor Constraints 6
Step 3: Add Interaction Terms
Step 4: Determine the Number of Runs 7
Step 5: Check the Design 7
Step 6: Gather and Enter the Data
Step 7: Analyze the Results

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** (Figure 1.1). Then, define the responses and factors.

3

4 Introduction to Designing Experiments My First Experiment

Figure 1.1 Select the Custom Designer



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 1.2).

Figure 1.2 Open and Close the Responses Panel by Clicking the Disclosure Icon

Responses				
Add Response 💉 🛛 Remove	_ <u> </u>			
Response Name	Goal	Lower Limit	Upper Limit	Importance
r	Maximize			

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 1.3.

	•[Custom Design					
	1	Responses					
Maximize Match Target		Add Response 🖌 Remove	Numbe	r of Responses.			
Minimize		Response Name		Goal	Lower Limit	Upper Limit	Importance
		Number Popped		Maximize			
None		Total Kernels		None	NA	NA	NA
		optional item					

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), double-click on its name and 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 double-clicking it and 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 double-clicking it and 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 1.4.

Figure 1.4 Renamed Factors with Specified Values

Ŷ	Factors					
	Add Factor	Remove Add N Factors	1			
	Name	Role	Changes	Values		
	✓ Brand	Categorical	Easy	Top Secret	Wilbur	
	⊿ Time	Continuous	Easy	3	5	
	2 Power	Continuous	Easy	5	10	

10 Click Continue.

5

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

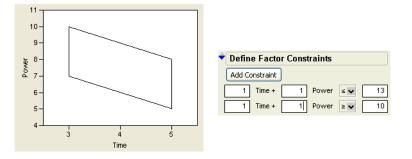
- popping corn for a long time on a high setting tends to scorch kernels.
- not many kernels pop when the popping time is brief and the power setting is low.

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 1.5).
- 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 1.5. 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 <= to >= in the second constraint.

The area inside the parallelogram, illustrated on the left in Figure 1.5, 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 1.5 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 1.6.

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 1.6. Figure 1.6 Add Interaction and Power Terms to the Model

Ŷ	Model					
	Main Effects	Interactions	▼ RSM Cross Powers ▼ Remove			
	Name	2nd 🔪	Estimability			
	Intercept	3rd h	Necessary			
	Brand	4th	Necessary			
	Time	5th	Necessary			
	Power		Necessary			
1	Brand*Time		Necessary			
(Brand*Power)	Necessary			
1	Time*Power 🖌	/	Necessary			

+	Model	
	Main Effects Interactions	▼ R5M Cross Powers ▼ Remo
	Name	Estimab 2nd
	Intercept	Necess 3rd 🗸
	Brand	Necess 4th
	Time	Necess 5th
	Power	Necessary
	Brand*Time	Necessary
	Brand*Power	Necessary
	Time*Power	Necessary
1	Time*Time	Necessary
V	Power*Power	Necessary

Step 4: Determine the Number of Runs

The Design Generation panel in Figure 1.7 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 1.7 Model and Design Generation Panels

÷	Model	
	Main Effects Interactions 🗸 RSM	1 Cross Powers 🗸 Remove Term
	Name	Estimability
	Intercept	Necessary
	Brand	Necessary
	time	Necessary
	Power	Necessary
	Brand*time	Necessary
	Brand*Power	Necessary
	time*Power	Necessary
	time*time	Necessary
	Power*Power	Necessary
•	Design Generation	
	Group runs into random blocks of s	size: 2
	— .	
	Number of Runs:	
	O Minimum 9	
	Default 16	
	O User Specified 16	
	Make Design	

Step 5: Check the Design

When you click **Make Design**, JMP generates and displays a design, as shown on the left in Figure 1.8. 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.

7

8 Introduction to Designing Experiments My First Experiment

Scroll to the bottom of the Custom Design window and look at the Output Options area (shown to the right in Figure 1.8. 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 1.8 Design and Output Options Section of Custom Designer

ston	n Desigı	ı)	Design Evaluation
sign					Prediction Variance Profile
n	Brand	Time	Power		
3	Wilbur	3	8.49209		Fraction of Design Space Plot
4	Wilbur	3	10	4	Prediction Variance Surface
5 To	p Secret	3	7	4	Relative Variance of Coefficients
6 To	p Secret	5	5		
7 To	p Secret	3.983825	7.540438	1	Design Diagnostics
8	Wilbur	5	5	-9	Output Options
9 To	p Secret	5	8		
10 To	p Secret	3	7	R	un Order: Randomize
1 To	p Secret	5	5	M	lake JMP Table from design plus
12	Wilbur	5	6.50791	Nu	lumber of Center Points: 0
13	Wilbur	5	8	Nu	lumber of Replicates: 0
14	Wilbur	4.005273	8.994727	6	
15 To	p Secret	3	10		Make Table
16 To	p Secret	5	8		Back
sign	Evalua	tion			

The resulting data table (Figure 1.9) 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. According to the table, cook the first bag of Top Secret for five minutes using a power setting of eight (rounded value). Next, cook a bag of Wilbur for five minutes at a power of eight, then cook Top Secret for three minutes using a power setting of seven, and so forth.

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

💌 Custom Design		◆ •					
Design Custom De	sign	•	Brand	Time	Power	Number Popped	Total Kernels
Criterion D Opt	timal	1	Top Secret	5	8	•	
 Screening 		2	Wilbur	5	8	•	
Model		3	Top Secret	3	7	•	
Constraint		4	Top Secret	3	10	•	
■DOE Dialog			i Wilbur	3	10	-	
💌 Columns (5/0)		6	i Wilbur	5	5	•	
և Brand 🛠		7	/ Wilbur	4	6	•	
🔺 Time 🛠		8	Top Secret	5	8	•	
A Power *		9	Top Secret	3	7	•	
🚄 Number Popped 苯 🔺 Total Kernels 苯		10) Top Secret	5	5	•	
📶 Tutai Kemeis 🛧		11	Wilbur	3	7	•	
Rows		12	Wilbur	3	8	•	
All rows	16	13	8 Wilbur	4	9	•	
Belected	0		Wilbur	5	7	•	
Excluded	0	15	5 Top Secret	4	7	•	
Hidden	0	16	Top Secret	3	10	•	
Labelled	0						

Figure 1.9 JMP Data Table of Design Runs Generated by Custom Designer

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 1.10.

Figure 1.10 Results of the Popcorn DOE Experiment

Popcorn DOE Results	6 . 🛡				Margaret and	Tedal	ľ
Design Custom Design		Brand	Time	Power	Number Popped	Total Kernels	
Criterion D Optimal	-	Top Secret	3	7	30	420	
 Screening 	· · ·			· ·			scripts to
 Ocreening Model 	2		3	7	120	420	1
	3	Top Secret	3	10	120	400	analyze data
Constraint ▼DOE Dialog	4	Top Secret	4	8	250	430	
	5	Top Secret	5	5	370	420	
💌 Columns (5/0)	6	Top Secret	5	5	400	420	results from
🔥 Brand 🛠	7	Top Secret	5	8	350	420	
🚄 Time 🗚	8	Top Secret	5	8	370	380	experiment
A Power \star	9	Wilbur	3	7	374	400	
A Number Popped \star	10	Wilbur	3	8	340	440	
🔺 Total Kernels 🛠	11	Wilbur	3	10	20	410	
Rows	12	Wilbur	4	6	440	460	
All rows 16	13	Wilbur	4	9	340	400	
Selected 0	14	Wilbur	5	5	170	370	
Excluded 0	IJ	Wilbur	5	7	370	380	
Hidden 0	10	Wilbur	5	8	420	450	
Labelled 0							1

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 1.11.

🏓 Report: Fit Model _ 🗆 🗙 Model Specification Select Columns -Pick Role Variables Personality: Generalized Linear Model 🗸 🔒 Brand 🚄 Number Popped V. Distribution Binomial ⊿ Time 🚄 Total Kernels Power Link Function Logit Number Popped Weight optional numeric Total Kernels Overdispersion Tests and Intervals Freq optional numeric Firth Bias-adjusted Estimates Offset optional numeric Run Model Help By optional Recall Remove Construct Model Effects Brand Add Time Cross Power Brand*Time Nest Brand*Power Time*Power Macros ~ Time*Time 2 Degree Power*Power Attributes Transform 💌 No Intercept

Figure 1.11 Fitting the Model

3 Click Run Model.

4 Scroll down to view the Effect Tests table (Figure 1.12) 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.

Introduction

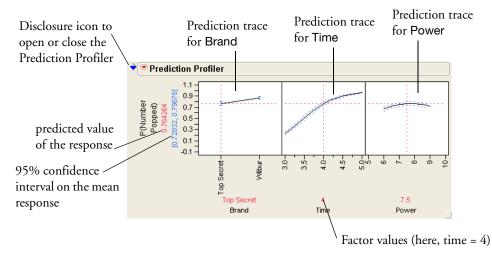
Figure 1.12 Investigating *p*-Values

Effect Tests				
		L-R		
Source	DF	ChiSquare	Prob> ChiS q	— <i>p</i> -values indicate significance.
Brand	1	85.470408	<.0001*	1 0
Time(3,5)	1	717.99922	/ <.0001* \	Values with * beside them are
Power(5,10)	1	12.782254	0.0003*	t walves that indicate the near
Brand*Time	1	435.97692	<.0001*	<i>p</i> -values that indicate the resu
Brand*Power	1	39.480267	<.0001*	are statistically significant.
Time*Power	1	7.8264948	0.0051*	7 8
Time*Time	1	3.3027281	0.0692	
Power*Power	1	19.011628	<.0001*	

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 1.13 shows the Prediction Profiler for the popcorn experiment. Prediction traces are displayed for each factor.

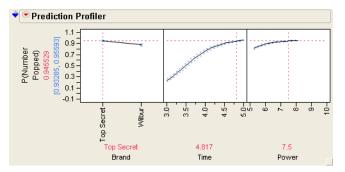
Figure 1.13 The Prediction Profiler



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 right and left (Figure 1.14).

12 Introduction to Designing Experiments My First Experiment

Figure 1.14 Moving the Time Value from 4 to Near 5



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

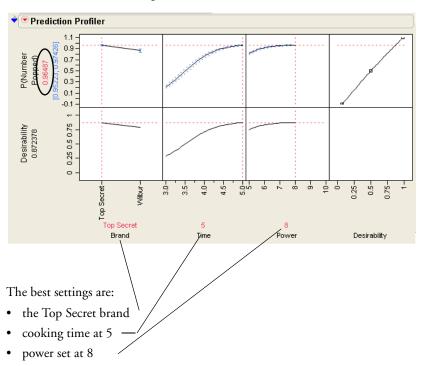
Furthermore, the steepness of a prediction trace reveals a factor's importance. Because the prediction trace for Time is steeper than that for Brand or Power (see Figure 1.14), you can see that cooking time is more important than the brand of popcorn or the microwave power setting.

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 1.15).

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.







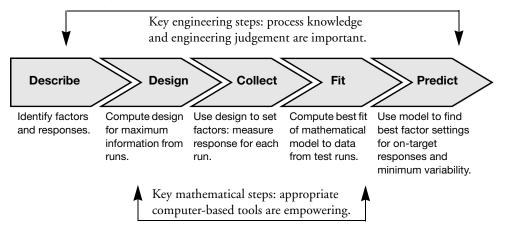
Examples Using the Custom Designer

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:



Contents

Creating Screening Experiments
Creating a Main-Effects-Only Screening Design
Creating a Screening Design to Fit All Two-Factor Interactions
A Compromise Design Between Main Effects Only and All Interactions
Creating 'Super' Screening Designs21
Screening Designs with Flexible Block Sizes25
Checking for Curvature Using One Extra Run28
Creating Response Surface Experiments 31
Exploring the Prediction Variance Surface
Introducing I-Optimal Designs for Response Surface Modeling
A Three-Factor Response Surface Design
Response Surface with a Blocking Factor
Creating Mixture Experiments 40
Mixtures Having Nonmixture Factors
Experiments that are Mixtures of Mixtures43
Special-Purpose Uses of the Custom Designer 46
Designing Experiments with Fixed Covariate Factors
Creating a Design with Two Hard-to-Change Factors: Split Plot
Technical Discussion

Creating Screening Experiments

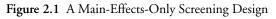
You can use the screening designer in JMP to create screening designs, but the custom designer is more flexible and general. The straightforward screening examples described below show that 'custom' does not mean 'exotic.' The custom designer is a general purpose design environment that can create screening 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 2.1 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.

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.



Custom Design Responses	Respon	Design					
•							
Factors	Factors						
Add Factor 🖌 Remove Add N Factors 1	Define I	Factor C	onstraint	ts			
Name Role Changes Values	Model 🔹						
AX1 Continuous Easy -1 1	Design						
AX2 Continuous Easy -1 1			140				
X3 Continuous Easy -1 1	Run 1	X1	X2	X3 1	X4	X5	×
AX4 Continuous Easy -1 1	2	-1			•	1	
X5 Continuous Easy -1 1	_	-1	-1	-1	1		
⊿ X6 Continuous Easy -1 1	3	-1 -1	1	-1	1	-1	-
Define Factor Constraints	4	-1	1 -1	1	-1 1	-1 -1	
				1 -1			
Model	6	1	-1	-1	-1	-1	-
Main Effects Interactions V R5M Cross Powers V Remove Term		-1	-1		-1	1	
	8		1	-1	1	1	
	🕈 Design Ev	valuatio	า				
Intercept Necessary	🔷 🕈 💌 Predi	iction Va	riance P	Profile			
X1 Necessary	Fraction	n of Doo	an Enga	o Blot			
X2 Necessary X3 Necessary			• •				
···· ·····,	🔷 🕈 💌 Predi	iction Va	riance S	Surface			
X4 Necessary	Relative	o Varian	ce of Co	efficients	2		
X5 Necessary				cincicita	·		
X6 Necessary	Alias M	atrix					
Design Generation	Design	Diagnos	tics				
Group runs into random blocks of size: 2	Output Opt						
		lions					
umber of Runs:	Run Order:			Randomize	,	~	
O Minimum 7	Make JMP T	able from c	lesian plus				
⊙ Default8	Number of C			0			
O User Specified 8	Number of F						
Make Design		_					
i lano boolgi	Make Table	•					

5 Click the disclosure button ($\blacklozenge \diamondsuit$ on Windows/Linux and $\blacktriangleright \lor$ on the Macintosh) to open the Alias Matrix. Figure 2.2 shows the Alias Matrix, which is a table of zeros, ones, and negative ones.

18 **Examples Using the Custom Designer** Creating Screening Experiments

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 2.6 and 3.4 in the table have a 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 the two-factor interactions X2*X6 and X3*X4. You are assuming that these interactions are negligible in size compared to the effect of X1.

Figure 2.2 The Alias Matrix

💙 Alias Ma	atrix														
Effect	12	13	14	15	16	23	24	25	26	34	35	36	45	46	56
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	1	0	0	0	0	0	0	0	-1	0	0
Х3	0	0	1	0	0	0	0	0	0	0	0	0	0	0	-1
X4	0	1	0	0	0	0	0	-1	0	0	0	0	0	0	0
X5	0	0	0	0	0	0	-1	0	0	0	0	-1	0	0	0
X6	1	0	0	0	0	0	0	0	0	0	-1	0	0	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" in Chapter 1, 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 2.3 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.

Main Effects Interactions 🗸	RSM Cross Powers 🗸 Rem	nove Term					
Name	Estimability	Model)		1		
Intercept	Necessary						
X1	Necessary	Design					
X2	Necessary	Run	X1	X2	Х3	X4	X
хз	Necessary	1	-1	1	1	-1	ĺ
X4	Necessary	2	1 -1	-1 -1	1 -1	-1 -1	1
x5	Necessary	4	-1	-1	-1	-1	2
x1*X2	Necessarv	5	-1	1	-1	1	-
X1*X3	Necessary	6	1	1	-1	-1	
X1*X4	Necessary	7	1	-1	-1	1	
x1 *X5		8	1	1	1	1	
	Necessary	9	-1	-1	1	1	1
X2*X3	Necessary	10	-1	-1	1	-1	-'
X2*X4	Necessary	11	1	1	-1	1	-1
X2*X5	Necessary	12	-1	-1	-1	1	-
X3*X4	Necessary	13	1	-1	-1	-1	-1
X3*X5	Necessary	14 15	-1 1	1	1	1 -1	-1 -1
X4*X5	Necessary	15	1	1 -1	1	-1	-1 -1
Design Generation						_ '	-
		/ Design E	valuatio	n			
Group runs into random block:	s of size: 2	🖉 🕨 🔍	diction V	'ariance l	Profile		
umber of Runs:		Fracti	on of De	sign Spa	ce Plot		
 Minimum 16 	Open outline /	🕨 💌 🕨	diction V	ariance 🗄	Surface		
ODefault32_	nodes <	🕨 Relati	ve Varia	nce of Co	pefficients		
O User Specified 16		Alias I	Matriu				

Figure 2.3 All Two-Factor Interactions

6 Click the disclosure button to open the Design Evaluation outlines, then open Alias Matrix. Figure 2.4 shows the alias matrix table of zeros and ones. The columns labels identify an interaction. For example, the column labelled 1 2 refers to the interaction of the first and second effect, the column labelled 2 3 refers to the interaction between the second and third effect, and so forth. Look at the column labelled 1 2. 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 2.4 Alias Matrix Showing all Two-Factor Interactions Clear of all Main Effects

Alias M	atrix									
Effect	12	13	14	15	23	24	25	34	35	4 :
Intercept	0	0	0	0	0	0	0	0	0	1
X1	0	0	0	0	0	0	0	0	0	
X2	0	0	0	0	0	0	0	0	0	
XЗ	0	0	0	0	0	0	0	0	0	
X4	0	0	0	0	0	0	0	0	0	
X5	Å	0	0	0	0	0	0	0	0	
X1*X2	(1)	0	0	0	0	0	0	0	0	
X1*X3	\checkmark	1	0	0	0	0	0	0	0	
X1*X4	0	0	1	0	0	0	0	0	0	
X1*X5	0	0	0	1	0	0	0	0	0	
X2*X3	0	0	0	0	1	0	0	0	0	
X2*X4	0	0	0	0	0	1	0	0	0	
X2*X5	0	0	0	0	0	0	1	0	0	
X3*X4	0	0	0	0	0	0	0	1	0	
X3*X5	0	0	0	0	0	0	0	0	1	
X4*X5	0	0	0	0	0	0	0	0	0	

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 2.5.

Figure 2.5 Model for Six-Variable Design with Two-Factor Interactions Designated If Possible

🕈 Custom Design	
Factors	
Define Factor Constraints	
Model	
Main Effects Interactions 🖌 RSM Cross Powe	ers 🗙 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 x1*x6	If Possible If Possib Necessary
X2*X3	
X2*X4	If Possib 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
Design Generation	
Group runs into random blocks of size: 2	1
	1
lumber of Runs:	
O Minimum 7 O 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 2.6). The values in your table may be different from those shown below, but with a similar pattern.

Figure 2.6 Alias Matrix

Alias N	latrix														
Effect	12	13	14	15	16	23	24	25	26	34	35	36	45	46	56
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	0	0	0.333	0
X1*X3	0	0.333	0	0	0	0	0	0	0	0	0	0	0	0	0.333
X1*X4	0	0	0.333	0	0	0	0	0	0.333	0	0	0	0	0	0
X1*X5	0	0	0	0.333	0	0	0	0	0	0	0	0.333	0	0	0
X1*X6	0	0	0	0	0.25	0	0.25	0	0	0	0.25	0	0	0	0
X2*X3	0	0	0	0	0	0.333	0	0	0	0	0	0	0.333	0	0
X2*X4	0	0	0	0	0.25	0	0.25	0	0	0	0.25	0	0	0	0
X2*X5	0	0	0	0	0	0	0	0.333	0	0.333	0	0	0	0	0
X2*X6	0	0	0.333	0	0	0	0	0	0.333	0	0	0	0	0	0
X3*X4	0	0	0	0	0	0	0	0.333	0	0.333	0	0	0	0	0
X3*X5	0	0	0	0	0.25	0	0.25	0	0	0	0.25	0	0	0	0
X3*X6	0	0	0	0.333	0	0	0	0	0	0	0	0.333	0	0	0
X4*X5	0	0	0	0	0	0.333	0	0	0	0	0	0	0.333	0	0
X4*X6	0.333	0	0	0	0	0	0	0	0	0	0	0	0	0.333	0
X5*X6	0	0.333	0	0	0	0	0	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 1 2 column and the same value in the 4 6 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 $X4^*X6$.

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 new 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 2.7, 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 2.7.

Figure 2.7 Changing the Estimability

→ Model	
Main Effects Interactions V RSM Cross	Powers 🗸 Remove Term
Name	Estimability
X5	If Possible
X6	If Possible
X7	If Possible
X8	If Possible
хэ	If Possible Necessary
X10	If Possible If Possible
X11	If Possible
X12	If Possible
Design Generation	
Group runs into random blocks of size:	2
Number of Runs:	
O Minimum 1	
O Default16	
O User Specified 16	
Make Design	

- 4 The desired number of runs is eight so type 8 in the User Specified edit box in the Number of Runs section.
- 5 Click the red triangle on the Custom Design title bar and select **Simulate Responses**, as shown in Figure 2.8.

Figure 2.8 Simulating Responses



6 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.

Chapter 2

24 **Examples Using the Custom Designer** Creating Screening Experiments

🔛 Cı																
Cu Desig	👫 Model			X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	¥
Criter	Simula	te Responses	1	1	-1	1	-1	-1	-1	1	1	1	-1	1	-1	4.47
💌 Mo	Effects	Y	2	1	-1	1	-1	1	1	-1	-1	-1	-1	-1	1	28.08
	Intercept	17	3	1	1	1	1	1	1	1	-1	1	1	1	1	18.68
	X1	-4	4	-1	1	-1	-1	-1	1	-1	1	1	1	-1	1	-16.38
	X2	-3	5	-1	-1	-1	-1	1	-1	1	-1	1	1	-1	-1	41.96
Co 🗩	X3	4	6	1	1	-1	1	-1	-1	1	-1	-1	-1	-1	1	29.14
⊿ X'	X4 X5	-9	7	-1	-1	1	1	-1	-1	-1	-1	-1	1	1	1	12.26
🚄 X:	X6	-10	8	-1	1	-1	1	1	1	-1	1	-1	-1	1	-1	14.85
📕 X:	X7	8														
📕 X4	X8	-6	_													
A X:	X9	-7	_													
	X10	8														
X 8	X11	-1														
X X8	X12	-3	_													
	Error Std.															
📕 X'	Apply															
🖌 X1																

Figure 2.9 Simulated Responses and Design Table

7 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

🖬 Ci	ustom Desig	yn														
♥ Cu Desig	m mouer			X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	Y
Criter	🔻 Simula	te Response	es 1	1	-1	1	-1	-1	-1	1	1	1	-1	1	-1	43.67
💌 Sc	Effects	<u> </u>	2	1	-1	1	-1	1	1	-1	-1	-1	-1	-1	1	35.97
💌 Mo	in toroopt	100	3	1	1	1	1	1	1	1	-1	1	1	1	1	6.56
	X1	0	4	-1	1	-1	-1	-1	1	-1	1	1	1	-1	1	-8.79
	X2	-7	5	-1	-1	-1	-1	1	-1	1	-1	1	1	-1	-1	-1.25
	X3 X4	0	6	1	1	-1	1	-1	-1	1	-1	-1	-1	-1	1	53.2
💌 Co	115	0	7	-1	-1	1	1	-1	-1	-1	-1	-1	1	1	1	54.68
📕 X'	X6	0	8	-1	1	-1	1	1	1	-1	1	-1	-1	1	-1	43.45
A X:	X7	0														
	X8	0														
X	X9	0														
X	X10	9														
🚄 X:	X11 X12	0														
📕 X8	Error Std	0.5														
		0.5														
X 1	Apply															
🚄 X1																

To identify active factors using stepwise regression:

- 1 To run the Model script in the design table, click the red triangle beside Model and select Run Script.
- 2 Change the **Personality** in the Model Specification window from **Standard Least Squares** to **Stepwise**.
- 3 Click Run Model on the Fit Model dialog.
- 4 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.

💌 Stepwise	e Fit						
Response: Y							
	e Regressi	on Con	trol				
Prob to Ente Prob to Leas Direction: Fo	ve 0.100	Enter Remo	All ove All				
Go Stop	Step M	ake Mode	J				
Current	Estimates						
SSE	DFE	MSE	RSqua	re F	Square Ad	j C	p AIC
2.0412625	5 0.	4082525	0.99	81	0.9973	3	4.92698
LockEntere	d Paramete	er Est	timate	nDF	SS	"F Ratio"	"Prob>F"
	Intercept	99	.98125	1	0	0.000	1.0000
	X1		0	1	0.440104	1.099	0.3536
	X2	-6	.99625	1	391.5801	959.162	0.0000
	Х3		0	1	0.348004	0.822	0.4158
	X4		0	1	0.238004	0.528	0.5077
	X5		0	1	1.162813	5.295	0.0828
	X6		0	1	0.352838	0.836	0.4123
	X7		0	1	0.019012	0.038	0.8557
	X8		0	1	1.161347	5.279	0.0832
	X9		0	1	0.161704	0.344	0.5890
	X10	9	.23375	1	682.0971	1670.773	0.0000
	X11		0	1	0.000113	0.000	0.9889
	X12		0	1	0.204347	0.445	0.5412
✓ Step History							
Step	Parameter	Action	"Sig P	rob"	Seq SS	RSquare	Ср
1	X10	Entered	0.	0180	682.0971	0.6341	
2	X2	Entered	0	0000	391.5801	0.9981	

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 🖌 R	emove	Add Factor 🖌 R	emove Add N Factors	: 1		
Continuous		Name	Role	Changes	Values	
Categorical	•	X1	Continuous	Easy	-1	1
Blocking	2 runs per block	🖌 📈 X2	Continuous	Easy	-1	1
Covariate	3 runs per block	🖌 🖌 🖌 🖌	Continuous	Easy	-1	1
Mixture	4 runs per block	5 📲 X4	Blocking	Easy	1	
Constant	8 runs per block	Specify Factors				
Uncontrolled	Other		ng the Add Factor butto	n. Double cli	rk on a factor	
		name or level to edit it	-	1. 500010 01		
		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 Des	ign					
Responses						
Factors						
Add Factor 🛛 🗸	Remove Add N	Factors	1			
Name	Role	Change	s Values			
🚄 X1	Continuous	Easy	-1	1		
🚄 X2	Continuous	Easy	-1	1		
4 X3	Continuous	Easy	-1	1		
🞿 X4	Blocking	Easy	1	2	3	
• Model	or Constraints					
Design Gen	eration					
	into random blocks	of size:	2			
Number of Run	s: 5					
 Minimum Default 	9					
User Specifi						

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 i	is 24	Gives	Eight	Blocks
-------------	--------	-----------	-------	-------	-------	--------

6					
Custom Desi	ign				
Responses					
Factors					
Add Factor 🛛 🗸	Remove Add N	Factors 1			
Name	Role	Changes Values	:		
A X1	Continuous	Easy -1		1	
🔺 X2	Continuous	Easy -1		1	
🔺 🛛 🖌	Continuous	Easy -1		1	
4 X4	Blocking	Easy 1 2	3 4	5 6	7 8
• Model					
Design Gene	eration				
	into random blocks	s of size:	2		
Number of Runs	s:				
🔘 Minimum	5				
🔿 Default	9				
 User Specifi 	ed 24				
Make Design					

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 Desi	gn						
•	Responses							
÷	Factors							
	Add Factor 🖌	Remove Add	d N Factors	1				
	Name	Role	Changes	Values				
	🚄 X1	Continuou	us Easy	-1		1		
	🚄 X2	Continuou	us Easy	-1		1		
	⊿ ХЗ	Continuou		-1		1		
	 X4	Blocking	Easy	1 2	2 3	}	4	5
•	Define Facto	r Constrain	ts					
)	Model							
	Main Effects	nteractions 🗸	RSM Cr	oss Pow	ers 🗸	Remov	/e Term	
	Name			Estimabilit	у			
	Intercept			Necessar	у			
	X1			Necessar	У			
	X2			Necessar	У			
	Х3			Necessar	У			
	X4			Necessar	У			
	X1*X2			Necessar	У			
	X1*X3			Necessar	У			
	X2*X3			Necessar	У			
`	Design Gene	ration						
	Group runs i	nto random blo	cks of size:	:	2			
					_			
	Number of Runs	s:						
	O Default	24						
	 Derault User Specific 		_					
	Make Design							

Click **Make Design**, then click the disclosure button (\diamondsuit on Windows/Linux and \blacktriangleright v on the Macintosh) 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

🕈 Relativ	e Variano	e of Co	efficients
Significand	e Level		0.050
Signal to N	1.000		
Effect	Variance	Power	
Intercept	0.070	0.802	
X1	0.078	0.763	
X2	0.106	0.638	
X3	0.083	0.736	
X4 1	0.099	0.668	
X4 2	0.086	0.723	
X4 3	0.085	0.731	
X4 4	0.081	0.747	
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. 68.

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 lf 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 lf 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 lf 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 is a four-run design with the factor settings shown in Figure 2.17.

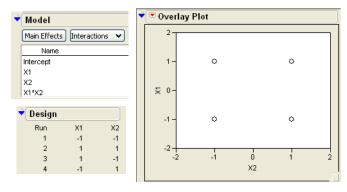


Figure 2.17 Two Continuous Factors with Interaction

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 Const	raints			
÷	Model				
	Main Effects Interaction:	s 🖌 RSM C	ross Pow	ers 🗙 Remove Term	
	Name			Estimability	
	Intercept			Necessary	
	X1			Necessary	
	X2			Necessary	
	X1*X2			Necessary	

- 5 Choose 2nd from the Powers button in the Model panel. This adds two quadratic terms.
- 6 Select the two quadratic terms (X1*X1 and X2*X2) 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

Add Factor 🖌 R	emove Add N Factor:	s 1	
Name	Role	Changes	Values
A X1	Continuous	Easy	-1 1
A X2	Continuous	Easy	-1 1
Define Factor O	Constraints		
Model			
Main Effects Inter	actions 🗙 RSM	Cross Pow	ers 🗸 Remove Term
Main Effects Inter Name	actions 🖌 RSM (Cross Pow	ers 💙 Remove Term Estimability
	actions 🗙 RSM (Cross Pow	
Name	actions 💙 (RSM) (Cross Pow	Estimability
Name Intercept	ractions 💙 (RSM) (Cross Pow	Estimability Necessary
Name Intercept X1	actions 💙 (RSM) (Cross Pow	Estimability Necessary Necessary Necessary Necessary
Name Intercept X1 X2	actions 🗸 (R5M) (Cross Pow	Estimability Necessary Necessary Necessary

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.

- 7 Enter 5 into the User Specified edit box in the Number of Runs section of the Design Generation panel.
- 8 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

♥ D	esign		
1	Run	X1	Х2
	1	-1	-1
	2	1	1
	3	1	-1
	4	-1	1
	5	0	0

- 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.

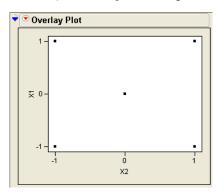


Figure 2.21 Overlay Plot of Five-run Bayesian D-Optimal Design

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).

31

Creating Response Surface Experiments

Figure 2.22 Adding a Factor and a Quadratic Term

Add Factor 🐱	▼ Model
Continuous	Main Effects Interactions 🗸 RSM Cross Powers 🗸 Remove Term
Categorical 🗟 🕨	Name 2nd Notify
Blocking 🕨 🕨	Intercept 3rd V sary
Covariate	X1 4th sary
Mixture	X1*X1 Sth
Constant	Str
Uncontrolled	

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

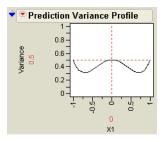
🕈 Design Genera	tion	
Group runs into	random blocks of size:	2
Number of Runs: Minimum Default User Specified Make Design	3 6 6	

5 Click the disclosure button (♠ ♠ on Windows/Linux 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

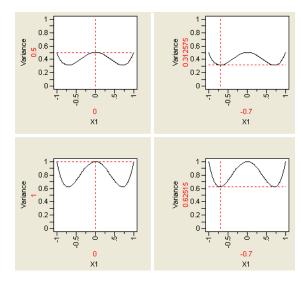


The prediction variance is relative to the error variance. When the relative prediction variance is one, the absolute variance is equal to the error variance of the regression model. More detail on the Prediction Variance Profiler is in "Understanding Design Evaluation," p. 64.

- 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: Control-click (Command-click on the Mac) 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.

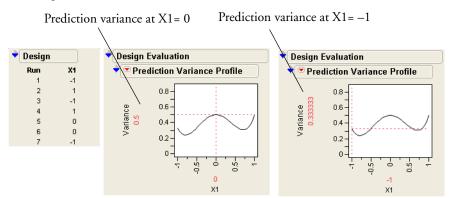


Figure 2.26 Sample Size of Seven for the One-Factor Quadratic Model

Introducing I-Optimal Designs for Response Surface Modeling

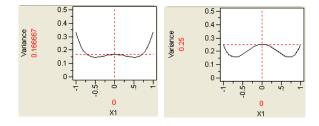
The custom designer generates designs using a mathematical optimality criterion. All the designs in this chapter so far have been *D*-Optimal designs. *D*-Optimal designs are most appropriate for screening experiments because the optimality criterion focuses on precise estimates of the coefficients. 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.





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/Linux and ▶ ♥ on the Macintosh) to open the **Prediction Variance Profile**, as shown on the left in Figure 2.27.
- 8 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.
- 7 The design is shown in the Design panel (the left in Figure 2.28).
- 8 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.

36 Examples Using the Custom Designer

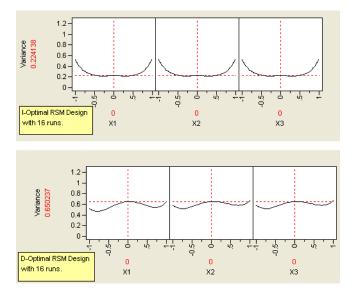
I-Optim	al Design]		D-Optim	al Desigi	n	
Run	X1	X2	ХЗ	Run	X1	X2	XЗ
1	1	0	0	1	-1	1	1
2	-1	1	1	2	-1	-1	-1
3	-1	0	0	3	-1	1	-1
4	0	-1	0	4	0	1	1
5	1	-1	-1	5	-1	0	1
6	-1	-1	1	6	-1	-1	1
7	1	1	-1	7	1	1	1
8	0	0	-1	8	-1	-1	0
9	0	0	1	9	-1	1	0
10	-1	-1	-1	10	-1	0	-1
11	0	1	0	11	1	0	0
12	0	0	0	12	0	0	-1
13	1	-1	1	13	1	-1	-1
14	-1	1	-1	14	0	-1	0
15	1	1	1	15	1	-1	1
16	0	0	0	16	1	1	-1

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

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



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 temper-

ature 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 Doney, 1992).

The response surface designer (described in "Response Surface Designs," p. 115) 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

Factors					
Add Factor 🗸	Remove Add N Factors	1			
Name	Role	Changes	Values		
A X1	Continuous	Easy	-1	1	
A X2	Continuous	Easy	-1	1	
at X3	Blocking	Easy	1		

3 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

▼ 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

4 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.

Creating Response Surface Experiments

Figure 2.32 Blocking Factor Now Shows Three Levels

Add Factor 🖌 Re	move 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 C	onstraints				
Model					
Main Effects Intera	ictions 🗸 RSM C	ross Power	s 🗸 Remov	e Term	
Naura -			Estimability		
Name					
Intercept			Necessary		
		I	,		
Intercept		l	Necessary		
Intercept X1			Necessary Necessary		
Intercept X1 X2			Necessary Necessary Necessary		
Intercept X1 X2 X3			Necessary Necessary Necessary Necessary		

- 5 Click Make Design.
- 6 In the Output Options, select Sort Right to Left from the Run Order list.
- 7 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:

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

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

🕈 I-Optim	al Design			D-Optim	al Design		
Run	X1	X2	X3	Run	X1	X2	X 3
1	1	-1	1	1	-1	-1	1
2	-1	0	1	2	0	-1	1
3	0	0	1	3	1	0	1
4	0	1	1	4	-1	1	1
5	0	-1	2	5	1	-1	2
6	0	0	2	6	0	0	2
7	-1	1	2	7	-1	1	2
8	1	1	2	8	1	1	2
9	-1	-1	3	9	-1	-1	3
10	0	0	3	10	1	-1	3
11	1	0	3	11	-1	0	3
12	0	1	3	12	0	1	3

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 below.

N

Examples

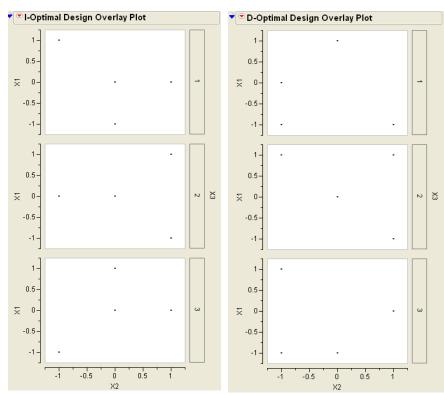


Figure 2.34 Plots of I-Optimal (left) and D-Optimal (right) Design Points by Block.

Either of the above designs 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. 53, 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		I	Design Diagnostics	
I Optimal Design			D Optimal Design	
D Efficiency	50.85382		D Efficiency	54.98822
G Efficiency	76.47609		G Efficiency	78.12922
A Efficiency	37.51264		A Efficiency	32.5789
Average Variance of Prediction	0.49973		Average Variance of Prediction	0.59188
Design Creation Time (seconds)	0.083333		Design Creation Time (seconds)	0.083333

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 folder in the sample data. The custom design panels should now look like those shown in Figure 2.36.

•	Custom Design							
- 🔶	Responses							
	Add Response 🖌 Remo	ove Numl	ber of Response:	s				
	Response Name		Goal	Lower Limit	Upper L	.imit	Importance	
	Damping		None	NA	NA		NA	_
	optional item							
- 🔶	Factors							
	Add Factor 🖌 Remove	Add N Fac	ctors 1					
	Name	Role	Changes	Values				
	CuSO4	Mixture	Easy	0.2		0.8		
	⊿ Na2S2O3	Mixture	Easy	0.2		0.8		
	🚄 Giyoxal	Mixture	Easy	0		0.6		
	✓ Wavelength	Categorica	al Easy	L1	L2		L3	

Figure 2.36 Mixture Experiment Response Panel and Factors Panel

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 Des	sign						
Responses							
Factors							
Add Factor 🗸	Remove Add N F	actors [1				
Name CuSO4	Role	Change					
ACUSO4	Mixture Mixture	Easy Easy	0.2			0.8 0.8	
Giyoxal	Mixture	Easy Easy	0.2			0.6	
✓ Wavelength	Categorical	Easy	L1		L2	L3	
errarolongar	outogonour	Luoy				120	
Define Fact	tor Constraints]					
Model							
Main Effects	Interactions 🗸 🔿	ross Po	wers 🗸	Cohoff	e Cubic	Remove Term	
		ross Po			e Cubic	Kellove Terlin	
Name			Estima	,			
CuSO4			Necessa				<u> </u>
Na2S2O3			Necessa	· ·			
Glyoxal CuSO4*Na2S20	~~		Necessa Necessa	· · ·			
CuSO4*Na2S2C CuSO4*Glyoxal			Necessa	·			
CuSO4*Wavele			Necessa	· · ·			
Na2S2O3*Glyo:	-		Necessa	· ·			
Na2S2O3*Wav			Necessa				-
Design Gen	eration						
	into random blocks o	f size:	2				
	Into Fundom biocito o	1 8120.					
lumber of Run	s:						
	12						
 Default 	24						
O User Specif	fied 18						
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.

42 **Examples Using the Custom Designer** Creating Mixture Experiments

Figure 2.38 Mixture Experiment Design Table

💌 Custom Design	♦ ●					
Design Custom Design	•	CuS04	Na2S2O3	Glyoxal	Wavelength	Damping
Criterion D Optimal	1	0.8	0.2	0	L3	•
Model	2	0.8	0.2	0	L1	•
DOE Dialog	3	0.2	0.5	0.3	L1	•
	4	0.8	0.2	0	L2	•
	5	0.5	0.2	0.3	L3	•
Columns (5/0)	6	0.2	0.2	0.6	L3	•
⊿ CuSO4 ≭ ⊿ Na2S2O3 ≭	7	0.5	0.5	0	L3	•
A Glyoxal 🛠	8	0.2	0.8	0	L3	•
wavelength *	9	0.5	0.2	0.3	L1	•
A Damping 🛣	10	0.2	0.2	0.6	L2	
	11	0.5	0.5	0	L1	•
Rows	12	0.5	0.2	0.3	L2	•
All rows 18	13	0.2	0.8	0	L1	•
Selected 0	14	0.2	0.8	0	L2	•
Excluded 0	15	0.2	0.2	0.6	L1	
Hidden 0	16	0.5	0.5	0	L2	
Labelled 0	17	0.2	0.5	0.3	L3	
	18	0.2	0.5	0.3	L2	

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					
Design Generation	Run	CuSO4	Na2S2O3	Glyoxal	Wavelength	Damping
Group runs into random blocks of size: 2	1	0.2	0.8	0	Ľ1	• •
	2	0.8	0.2	0	L1	
Number of Runs:	3	0.8	0.2	0	L2	
O Minimum 5	4	0.2	0.2	0.6	L2	
O Default 6	5	0.2	0.8	0	L3	
 User Specified 10 	6	0.2	0.2	0.6	L1	
	7	0.8	0.2	0	L3	
Make Design	8	0.2	0.8	0	L2	
	9	0.2	0.2	0.6	L3	
	10	0.2	0.2	0.6	L3	

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, 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.

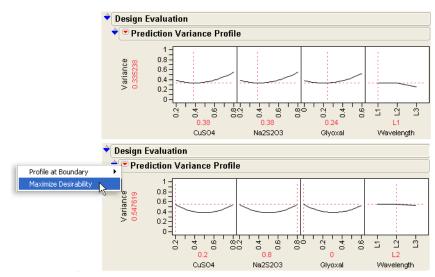
The values of the first three ingredients sum to one because they are mixture ingredients. If you vary one of the values, the others adjust to keep the sum constant.

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 L2, 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

•	Custom Design					
- 🗎	Responses					
	Add Response 🖌 Remov	e Number o	f Responses)		
	Response Name	G	bal	Lower Limit	Upper Limit	Importance
	Taste	M	aximize	1	10	
*	Factors Add Factor V Remove	Add N Factor		Values		
	Name	Role	Changes	-		
	A Cocoa	Mixture	Easy	0.1	0.2	
	⊿ Sugar	Mixture	Easy	0	0.1	
	A Flour	Mixture	Easy	0.2	0.3	
	🚄 Butter	Mixture	Easy	0.1	0.2	
	⊿Milk ⊿Eggs	Mixture	Easy	0.25	0.3	5

- 5 Next, click **Continue**.
- 6 Open the Constraints pane and click Add Constraint twice.
- 7 Enter the constraints as shown at the top in Figure 2.42.

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.

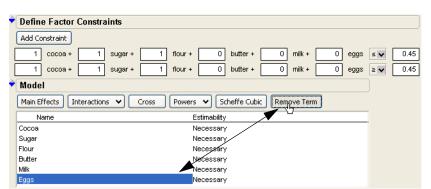


Figure 2.42 Constraints to Define the Double Mixture Experiment

- 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 inFigure 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

Custom Design	◆ ●							_
Design Custom Design	•	Cocoa	Sugar	Flour	Butter	Milk	Eggs	Taste
Criterion D Optimal	1	0.20	0.05	0.20	0.10	0.25	0.20	9.30
▼ Model ▼ Constraint	2	0.20	0.05	0.20	0.10	0.35	0.10	6.66
 DOE Dialog 	3	0.20	0.00	0.25	0.20	0.25	0.10	7.61
	4	0.15	0.00	0.30	0.15	0.35	0.05	6.38
Columns (7/0)	5	0.15	0.00	0.30	0.15	0.35	0.05	5.25
🚄 Cocoa 🛠 🔺 Sugar 🛠	6	0.10	0.15	0.20	0.20	0.25	0.10	5.77
🚄 Suyar 🛪 ⊿ Flour 🛠	7	0.10	0.15	0.20	0.20	0.30	0.05	5.02
🖌 Butter \star	8	0.10	0.15	0.20	0.10	0.35	0.10	5.47
🚄 Milk 🛠	9	0.10	0.05	0.30	0.10	0.25	0.20	4.14
🚄 Eggs 🛠	10	0.10	0.05	0.30	0.10	0.25	0.20	5.81
🚄 Taste 🛠								
	/	$\overline{}$			L		$\overline{}$	1
	Each run	sums	to 0.4	5 (45%	6)	Each 1	run sùi	ns to (

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 clicking Add Factors 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

÷	Factors					
	Add Factor 👻 Re	mov	e Add N Factors	: 1		
	Continuous		Role	Changes	Values	
	Categorical	۲	Continuous	Easy	-1	1
	Blocking	•	Continuous	Easy	-1	1
	Covariate 📐					
	Mixture 🧏					
	Constant					
A	Uncontrolled		Add Factor butto	n. Double clia	ck on a factor	

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

Figure 2.45 Design with Fixed Covariate

name age	2	OK Cancel]
sex height			
vveight			

- 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 Const	raints			
▼ Model				
Main Effects Interaction	is 🖌 RSM C	ross Powe	ers 🗙 Remove Term	
Name		3	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.

Chapter 2

48 Examples Using the Custom Designer

Special-Purpose Uses of the Custom Designer

Figure 2.47 Design Table for Covariate Example

💌 Custom Design	♦∖्♥					ĺ
Design Custom Design	•	calories	sleep	weight	Y	
Criterion D Optimal	1	-1	-1	95	•	I
 Screening 	2	1	1	123	•	Ī
▼ Model	3	1	-1	74	•	İ
Columns (4/0)	4	1	-1	145	•	Î
<pre></pre>	5	-1	-1	64	•	Ī
∠a sleep ★	6	1	1	84	•	İ
veight *	7	1	1	128	•	Ī
🖌 Y 🕺	8	1	-1	79	•	Ī
(9	1	1	112	•	
Rows	10	1	1	107	•	I
All rows 40	11	1	1	67	•	Ī
Selected 0	12	-1	1	98	•	İ
Excluded 0	13	-1	-1	105	•	İ
Hidden 0 Labelled 0	14	-1	1	95	•	İ

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 **Mode**l in the data table and select **Run Script**, as shown on the left in Figure 2.48.

Custom Design	🌓 🔨 💌					1	🖊 🏝 Model Specificati	ion
Design Custom Design		calories	sleep	weight	Y		Select Columns	Pick Role Variables Personality:
Criterion D Optimal	1	-1	-1	95	•		🚄 calories	Y Aweight Emphasis:
 Screening 	2	1	1	123	•		📕 📶 sleep	optional
Model	3	1	-1	74	•		/ weight	
Run Script	4	1	-1	145	•		⊿ Y	Weight optional Numeric Help
	5	-1	-1	64	•			
Delete	6	1	1	84	•			Freq optional Numeric Remove
👅 sieeb 🛧	7	1	4	400				By optional
								Construct Model Effects
								Calories
								Add calories
								Cross calories*sleep
								Next
								Nest

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 Model**.

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.

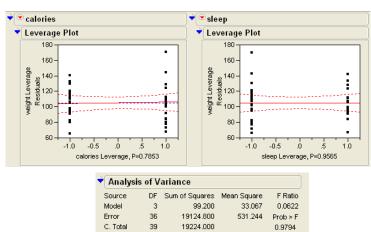


Figure 2.49 Analysis to Check That Weight is Independent of Calories and Sleep

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.

N

Examples

50 Examples Using the Custom Designer

Special-Purpose Uses of the Custom Designer

- 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

•	Custom Design						
Ì	Responses						
	Add Response 🖌 Rema	ve N Respons	es				
	Response Name	Goal		Lower Limit	Upper Li	mit	Importance
	thickness	Maxi	mize	10			
•	Factors						
	Add Factor 🖌 Remove	Add N Factors	1				
	Name	Role	Changes	Values			
	⊿m1	Mixture	Easy	0		1	
	⊿ m2	Mixture	Easy	0		1	
	⊿ m3	Mixture	Easy	0		1	
	dextrusion rate	Continuous	Hard	-1		1	
	a 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

Model							
Main Effects	Interactions	▼ Cross Powers ▼ Scheffe Cubic Remove Term					
Name	2nd	Estinability					
m1	3rd VS	Necessary					
m2	4th	Necessary					
m3	5th	Necessary					
m1*m2	Jan	Necessary					
m1*m3		Necessary					
m1*extrusion r		Necessary					
m1*temperatur	e	Necessary					
m2*m3		Necessary					
m2*extrusion r	ate	Necessary					
m2*temperatur	e	Necessary					
m3*extrusion r	ate	Necessary					
m3*temperatur		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).

Design Generation	
Number of Whole Plots	7
Number of Runs:	
	13
O Default O User Specified	24 28
Make Design	

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

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

Desig	jn					
Run	Whole Plots	m1	m2	m3	extrusion rate	temperature
1	1	0	0	1	-1	1
2	1	0	1	0	-1	1
3	1	0.9	0.1	0	-1	1
4	1	0.5	0	0.5	-1	1
5	2	1	0	0	1	1
6	2	0	0.4	0.6	1	1
7	2	0	1	0	1	1
8	2	0.5	0	0.5	1	1
9	3	0	1	0	1	-1
10	3	0	0	1	1	-1
11	3	0	0.5	0.5	1	-1
12	3	0.8	0.2	0	1	-1
13	4	0	0	1	-1	-1
				0	4	4

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.

Chapter 2

52 **Examples Using the Custom Designer**

Special-Purpose Uses of the Custom Designer

Figure 2.54 The Vinyl Data Design Table

💌 vinyl data		♦							
Design Custom De	sign	$\overline{}$	Whole Plots	m1	m2	m3	extrusion rate	temperature	thickness
Criterion D Opt	timal	1	1	0	1	0	-1	-1	4
 Model 		2	1	1	0	0	-1	-1	10
		3	1	0	0	1	-1	-1	3
Columns (7/0)		4	1	0.5	0	0.5	-1	-1	9
💼 Whole Plots \star		5	2	0.5	0.5	0	-1	1	8
⊿ m1 ≭ ⊿ m2 ≭		6	2	0	1	0	-1	1	5
⊿m2 木 ⊿m3 ★		7	2	1	0	0	-1	1	6
🔺 extrusion rate 🗶		8	2	0	0	1	-1	1	2
🖌 temperature 🗶		9	3	0	1	0	1	-1	7
🚄 thickness 🗶		10	3	0	0.5	0.5	1	-1	9
		11	3	0	0	1	1	-1	14
Rows		12	3	1	0	0	1	-1	12
All rows	28	13	4	0	1	0	1	1	5
Selected	0	14	4	0	0	1	1	1	2
Excluded	0	15	4	1	0	0	1	1	6
Hidden	0	16	4	0.5	0	0.5	1	1	7
Labelled	0	17	5	1	0	0	1	1	11

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 the *JMP Statistics and Graphics Guide*.

Figure 2.55 Define the Model in the Fit Model Dialog

Select Columns	Pick Role Variables	Personality: Standard Least Squares V
m1 m2 m3 extrusion rate temperature thickness	Y ■ Interfaces optional Optional Numeric Freq optional Numeric By optional Optional Whole Plots& Ram 28 R58 Mixture Macros m18 R58 Mixture Macros m1*n2 Macros m1*extrusion rat m1*extrusion rat m2*m3 Transform m2*extrusion rat	e

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

Figure 2.56 Split Plot Analysis Results

Summary of Fit						
RSquare	0.85681					
RSquare Adj	0.742258					
Root Mean Square Error	1.97493					
Mean of Response	6.892857					
Observations (or Sum Wgts)	28					
Parameter Estimates						
Term	Estimate	Std Error	DFDen	t Ratio	Prob> t	
m1	9.4464782	0.989604	7.99	9.55	<.0001*	
m2	5.378571	1.120267	10.35	4.80	0.0007*	
m3	5.5546291	1.030955	8.786	5.39	0.0005*	
m1*m2	-6.458845	5.83806	13.36	-1.11	0.2881	
m1*m3	8.5532871	5.607183	13.15	1.53	0.1508	
m1*extrusion rate	-0.647611	0.958111	7.261	-0.68	0.5200	
m1*temperature	-1.8151	0.961061	7.328	-1.89	0.0990	
m2*m3	10.575333	5.226599	13.04	2.02	0.0640	
m2*extrusion rate	0.2944425	0.997579	8.162	0.30	0.7752	
m2*temperature	1.2044413	0.993319	8.065	1.21	0.2596	
m3*extrusion rate	2.6611621	0.979087	7.733	2.72	0.0272*	
m3*temperature	-1.498647	0.97904	7.731	-1.53	0.1657	
extrusion rate*temperature	-1.866921	0.740748	2.858	-2.52	0.0903	
Random Effect Predictions						
REML Variance Component Estimates						
Random Effect Var Ratio	Var Compon	ent Std	Error 95	% Lowe	r 95% Upper	Pct of To
Whole Plots 0.6350069	2.476	748 2.971	1536 -	3.34671	8.3002091	38.8
Residual	3.9003	483 1.589	3101 2	.007348	7 10.607283	61.1

Technical Discussion

This section provides more technical information about *I*-, *D*-, and Bayesian *D*-Optimal designs in JMP.

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

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

Chapter 2

D-optimal split plot designs maximize D when

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

where 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 lf Possible terms to force in runs that allow for detecting any inadequacy in the model containing only the Necessary terms. Let 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 (X'X + K). The difference between the criterion for D-Optimality and Bayesian D-Optimality is this constant added to the diagonal elements corresponding to the lf Possible terms in the X'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.
- maximizes this criterion: If *f*'(x) denotes a row of the X matrix corresponding to factor combinations *x*, then

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

where

$$M = \int_{R} f(x) f(x)' dx$$

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 X'X + K where 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:

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

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

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

where M is defined as before.



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
Enter Responses and Factors into the Custom Designer
Describe the Model
Select the Number of Runs
Understanding Design Evaluation
Specify Output Options
Make the JMP Design Table
Creating Random Block Designs
Creating Split Plot Designs
Creating Split-Split Plot Designs
Creating Strip Plot Designs
Special Custom Design Commands
Save Responses and Save Factors
Load Responses and Load Factors
Save Constraints and Load Constraints
Set Random Seed: Setting the Number Generator
Simulate Responses
Save X Matrix: Viewing the Design (X) Matrix in the Log
Optimality Criterion: Changing the Design Criterion (D- or I- Optimality)
Number of Starts: Changing the Number of Random Starts
Sphere Radius: Constraining a Design to a Hypersphere
Disallowed Combinations: Accounting for Factor Level Restrictions 80
Advanced Options for the Custom Designer
Save Script to Script Window
Assigning Column Properties
Define Low and High Values (DOE Coding) for Columns
Set Columns as Factors for Mixture Experiments
Define Response Column Values
Assign Columns a Design Role
Identify Factor Changes Column Property 87
How Custom Designs Work: Behind the Scenes

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.

Figure 3.1 Entering Responses

		Click to enter lower and upper			
		limits and importance weights.			
		4			
To enter one	Responses				
response at a (1)	Add Response 🖌 Remove	Number of Responses			
time, click then	Response Name	Goal Lower Limit Upper Limit Importance			
select a goal type:	Y optional item	Maximize . / . ` `			
Maximize, Match					
Target, Minimize,					
or None.	2	(3)			
	Double-click to edit	Click to change the response goal,			
	the response name, if	if desired.			
	desired.				

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):

60 Building Custom Designs

Creating a Custom Design

- 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 Then, 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 + 4X1 + 5X2 + random noise, where the *random noise* is distributed with mean zero and standard deviation one.

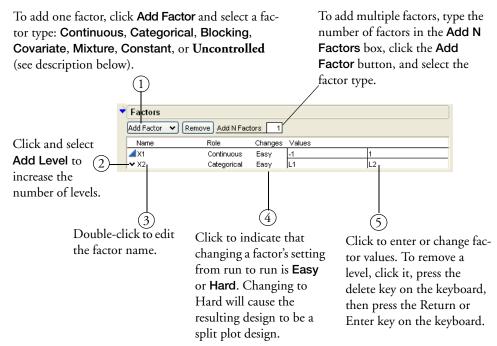
💌 Custom Design		۲			
Design Custom Desi	gn 🗨 🗋		X1	X2	Y
▼Model		1	1	-1	27.32
🕺 Model 🚺		2	1	1	35.1
pa model		3	-1	1	29.06
🕈 Simulate Resp	onses	4	-1	-1	21.8
Effects	Y				
Constant 2	8				
X1	4				
X2	5				
Error Std.	1				

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

How to Enter Factors

To enter factors, follow the steps in Figure 3.3.

Figure 3.3 Entering Factors in a Custom Design



Types of Factors

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

ω

Custom Designs



- **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 Random Block Designs," p. 71, 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 blue disclosure button (♠ ♠ on Windows/Linux 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.
- 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.

)	۲	Custom Design
	٠	Responses
	•	Factors
•	•	Define Factor Constraints
		Add Constraint
	[1 X1 + 1 X2 ≥ ▼ 1

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 3.4. 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 3.4 Add Terms

•	Custom De	sign
•	Responses	
•	Factors	
•	Define Fact	tor Constraints
)	Model	
	Main Effects	Interactions V RSM Cross Powers V Remove Term
ſ	Name	Estimability
- F	Intercept	Necessary
	X1	Necessary
1	X2	Necessary

Table 3.1 summarizes the ways to add specific factor types to the 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.
Add all second-order effects, including two-factor interactions and quadratic effects	Click the RSM button. The design now uses <i>I</i> -Optimal- ity criterion rather than <i>D</i> -Optimality criterion.
Add selected cross product terms	 Highlight the factor names. Highlight term(s) in the model list. Click the Cross button.
Add powers of continuous factors to the model effects	Click the Powers button and select 2nd , 3rd , 4th , or 5th .

Select the Number of Runs

The Design Generation panel (Figure 3.5) 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 3.5 Options for Selecting the Number of Runs

💌 Custom Design	
Responses	
Factors	
Define Factor Constraints	
Model	
Design Generation	
Group runs into random blocks of size:	2
Number of Runs:	
O Minimum 3	
 Default 4 	
O User Specified 4	
Make Design	

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 3.6. JMP always provides the Prediction Variance Profile, but the Prediction Surface Plot only appears if there are two or more variables.

Figure 3.6 Custom Design Diagnostic Tools

🔶 💌 Custom	Design			
Response	ses			
Factors				
Define F	actor Co	nstraints		
Model]			
🔶 Design				
Run	X1	X2		
1	1	1		
2	-1	1		
3	1	-1		
 4	-1	-1		
🔶 Design l	Evaluatio	n		
🔶 💌 Pred	liction Va	ariance F	Profile	
Fraction	on of Des	ign Spac	e Plot	
🔶 💌 Pred	liction Va	ariance S	Surface	
Relativ	ve Variar	nce of Co	efficients	
Alias I	Matrix			
🕨 Desigr	n Diagnos	stics		

The Prediction Variance Profile

The example in Figure 3.7 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.

Building Custom Designs

Creating a Custom Design

66

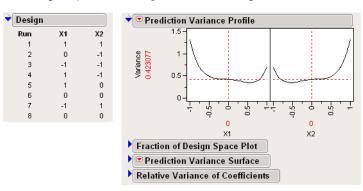


Figure 3.7 A Factor Design Layout For a Response Surface Design with 2 Variables

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.

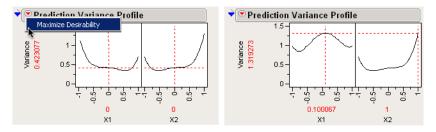
After you run the experiment, collect the data, and fit the model, you can estimate the actual variance of prediction at any setting by multiplying the relative variance of prediction by the mean squared error (MSE) of the least squares fit.

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 3.8. 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 3.8 Find Maximum Prediction Variance

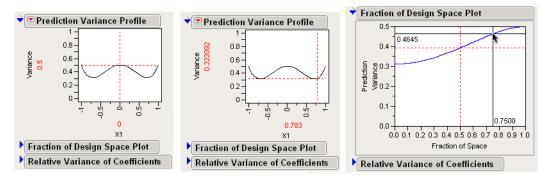


The Fraction of Design Space Plot

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 3.9. 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 or percentage of prediction variance values, ranging from 0 to 100%, and the *Y* axis is the range of prediction variance 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 3.9 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 3.9 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 3.10 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 the high and low variance areas for both factors. Compare this plot to the Prediction Variance Profile shown in Figure 3.8.

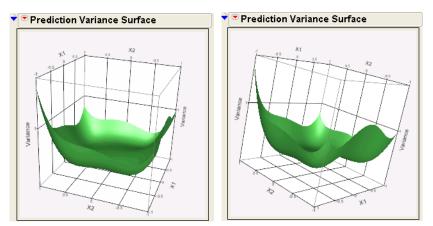


Figure 3.10 Prediction Variance Surface Plot for Two-Factor RSM Model

You can find complete documentation for the Surface Plot platform in *JMP Statistics and Graphics Guide*.

The Relative Variance of Coefficients and Power Table

Before clicking **Make Table** in the custom designer, click the disclosure button ($\diamondsuit \diamondsuit$ on Windows/ Linux and $\blacktriangleright \blacksquare$ on the Macintosh) to open the Relative Variance of Coefficients table.

The Relative Variance of Coefficients table (Figure 3.11) shows the relative variance of all the coefficients for the example RSM custom design (see Figure 3.7). 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 3.11 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 3.11 shows that the resulting power is 0.984 for all the parameters.

▼ Relative Variance of Coefficients					Relative Variance of Coefficients						
Significance Level 0.050				Significand	0.050						
Signal to N	loise Ratio		1.000		Signal to N	loise Ratio		2.000			
Effect	Variance	Power			Effect	Variance	Power				
Intercept	0.423	0.153			Intercept	0.125	0.984				
X1	0.231	0.231			X1	0.125	0.984				
X2	0.231	0.231			X2	0.125	0.984				
X1*X1	0.808	0.106			Х3	0.125	0.984				
X1*X2	0.250	0.218									
X2*X2	0.808	0.106									

The Alias Matrix (Confounding Pattern)

Click the Alias Matrix disclosure button ($\diamond \diamond$ on Windows/Linux and $\blacktriangleright \nabla$ on the Macintosh) to open the alias matrix (Figure 3.12).

The alias matrix only appears for custom designs if all factors are two-level factors. The table shows the aliasing (if any) between the model terms and all the two-factor interactions. It allows you to see the confounding pattern in *D*-Optimal and *I*-Optimal designs.

Figure 3.12 Alias Matrix

🕇 Alias M	atrix									
Effect	12	13	14	15	23	24	25	34	35	45
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
Х3	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

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:

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

where

Creating a Custom Design

- $N_{\rm D}$ is the number of points in the design
- *p* is the number of effects in the model including the intercept
- $\sigma_{\rm M}$ 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 3.13 Custom Design Showing Diagnostics

Model									
Design									
Design Evaluation									
Prediction Variance Profile									
Fraction of Design Space Plot									
Prediction Variance Surface									
Relative Variance of Coefficients									
Design Diagnostics									
l Optimal Design									
D Efficiency	40.54457								
G Efficiency	87.69085								
A Efficiency 28.99299									
Average Variance of Prediction	0.496709								
Design Creation Time (seconds)	0.166667								

Specify Output Options

Output Options		И	Keep the Same
Run Order:	Keep the Same 🗙		Sort Left to Right ^人
Make JMP Table from design plus			Randomize
Number of Center Points:	0		Sort Right to Left
Number of Replicates:	0		Randomize within Blocks
Make Table			
Back			

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

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 3.14 Example Design Table

This area identifies the design type that generated the table. Click **Custom Design** to edit.

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.

DOE Dialog is a script. Click the red trangle icon and select **Run Script** to recreate the DOE Custom Dialog and generate a new design table.

	Custom Design				
_	Design Custom Design		X1	X2	т
	Criterion I Optimal	1	-1	1	-
_	_♥ Model	2	0	0	•
	DOE Dialog	3	1	-1	-
/	Columns (3/0)	4	0.74	1	-
	A X1 *	5	-1	-1	•
/	🖌 X2 🗱	6	1	0	-
	🔺 Y 🗱	7	0	0	-
		8	0	-1	•

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.

Design Generation	tion	
Group runs into	random blocks of size:	4
Number of Runs:		
🔘 Minimum	24	
O Default	48	
 User Specified 	28	
Make Design		

In this example, the Design Generation Panel shown here des-

ignates four runs per block, and the user-specified number of runs (28) indicates there will be seven 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. 49.

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 chang-

ing 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. 87, 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 3.15 shows an example of how factors might be defined for the cheese processing example.

Figure 3.15 Example of Split-Split Response and Factors in Custom Designer Dialog

Custom Design							
Responses							
Add Response 🖌 Rem	ove Number of R	esponses					
Response Name	Goal	Lower Lim	it UpperLimit	Importance			
Consistancy	None	NA	NA	NA			
Factors							
Add Factor 🖌 Remove	Add N Factors	1					
Name	Role	Changes	Values				
⊿ storage 1	Continuous	Very Hard		1			
storage 2	Continuous	Very Har E	asy	1			
dcurds 1	Continuous	Hard H	lard	1			
	Continuous Continuous			1			
dcurds 1			lard 'ery Hard	1 1 1	_		
✓ curds 1 ✓ curds 2	Continuous	Hard		1 1 1 1			
<pre>curds 1 curds 2 curds 3</pre>	Continuous Continuous	Hard Hard		1 1 1 1 1 1			
 curds 1 curds 2 curds 3 curds 4 	Continuous Continuous Continuous	Hard Hard Hard Hard		1 1 1 1 1 1 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 3.16. 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 3.16 lists the run with subplots that are not nested in the whole plots.

Figure 3.16 Example of Strip Split Factors and Design Generation panel in Custom Designer Dialog

÷[💌 Custom Design			🔶 💌 Cust	om Design				
•	Responses					🕈 Desi	gn		
÷	Factors					Run	Whole Plots	Subplots	deposition 1
	Add Factor 🖌 Remove	Add N Factors	1			1	1	1	-1
	Name	Role	Changes	Values		2	1	2	-1 -1
	deposition 1	Continuous	Very Hard		1	4	. 1	4	-1
	deposition 2	Continuous	Very Hard		1	5	2	5	1
	detch 1	Continuous	Hard	-1	1	6	2	6	1
	detch 2	Continuous	Hard	-1	1	7	2	7	1
	detch 3	Continuous	Hard	-1	1	8	2	8	1
	detch 4	Continuous	Hard	-1	1	9	3	9	-1
	🚄 clean 1	Continuous	Hard	-1	1	10	3	10	-1
	🚄 clean 2	Continuous	Easy	-1	1	11	3	11	-1
	🚄 clean 3	Continuous	Easy	-1	1	12	3	12	-1
	Define Factor Const	ainte				13	4	1	1
						14	4	2	1
•	Define Factor Const	raints				15	4	3	1
	Model					16	4	4	1
1	Design Generation					17	5	5	-1
						10	5	7	-1
	Hard to change factors	s can vary indeper	ndently of Ver	y Hard to change fact	ors	20	5	8	-1
	Number of Whole Plots	6				20	6	9	-1
1	Number of Subplots	12				22	6	10	1
	· · ·					23	6	11	1
	lumber of Runs:					24	6	12	1
		13				-			
	O Default	16							
	User Specified	24							
	Make Design								

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 3.17). 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.

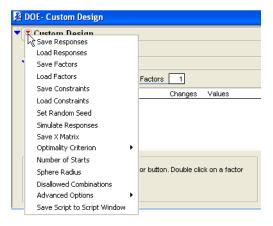


Figure 3.17 Click the Red Triangle Icon to Reveal Commands

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. 59, 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 a column property (noted with an asterisk icon in the column panel) called Design Role that identifies them as DOE factors and lists their type (continuous, categorical, blocking, and so on).
- 4 Save the data table.

Load Responses and Load Factors

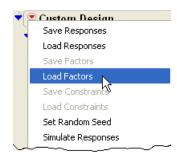
If you have saved responses and factors, you can quickly apply them to your design and avoid retyping this information each time you run an experiment.

Special Custom Design Commands

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 (Figure 3.18) and select Load Responses or Load Factors.

Figure 3.18 Loading Factor Names and Values from a JMP Table



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. 59, for details).
- 3 Click the red triangle icon on the title bar (Figure 3.18) 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 (Figure 3.18) 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 (Figure 3.18) and select **Set Random Seed**.

The window that appears shows the *generating seed* for that design (Figure 3.19). 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 3.19 Setting the Random Seed

ĺ	JMP: Please Enter a Number	
	Choose a positive whole number as a seed for a random starting design.	1084940514
	OK Cancel	

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 3.20) would be,

y = 19 + 2X1 + (-5)X2 + 6X1X2 + random noise,

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

Figure 3.20 Example of a Custom Design with Simulated Responses

🔜 Custom Design					📓 Model 📃 🗖 🔯
 Custom Design Design Custom Design 		X1	X2	Y	Simulate Responses
Criterion D Optimal	1	-1	1	5.38	Intercept 19
 Screening 	2	-1	-1	27.62	X1 2
▼Model	3	-1	-1	26.96	X2 -5
	4	1	-1	20.88	X1*X2 6
Columns (3/0)	5	1	1	21.98	Error Std. 1
🔺 X1 🗶 🔺	6	1	-1	20.74	Apply
🖌 X2 \star 📃	7	-1	1	5.43	
<u> </u>	8	1	1	22	

Save X Matrix: Viewing the Design (X) Matrix in the Log

To create a script and save the design matrix as a table property in a data table, click the red triangle icon in the Custom Design title bar (Figure 3.17) and select **Save X Matrix**. A script is saved as a table property called **Design Matrix**. When you run this script, JMP creates a global matrix called **X** and displays its number of rows in the log. If you do not have the log visible, select **View > Log (Window > Log** on the Macintosh).

Optimality Criterion: Changing the Design Criterion (D- or I- Optimality)

To change the design criterion, click the red triangle icon in the Custom Design title bar (Figure 3.21) and select **Optimality Criterion**, then choose **Make D-Optimal Design** or **Make I-Optimal Design**.

Custom Design Save Responses Load Responses Save Factors Load Factors Save Constraints Load Constraints Set Random Seed Simulate Responses Save X Matrix Optimality Criterion 🕨 🗹 Recommended Number of Starts Make D-Optimal Design Sphere Radius Make I-Optimal Design Disallowed Combinations Advanced Options Save Script to Script Window

Figure 3.21 Changing the Design Criterion

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. 53.

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 3.17) and select **Number of Starts**. When you select this com-

JMP: Please Enter a Number					
Choose a positive whole number for the number of random starting designs.					

mand, the window shown here 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.

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 Log*D* = *c*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. 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 5.2 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			
more than 32	2			

 Table 3.2
 Sample Size and Random Starts

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 sample sizes 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 3.17) 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. 62, and "Creating Random Block Designs," p. 71), 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.
- 2 Click the red triangle icon in the title bar (Figure 3.17) of the designer window and select Disallowed Combinations. Note that this menu item is not available if you have already defined linear inequality constraints.
- 3 Enter a Boolean expression that identifies what you do not want allowed (Figure 3.22). 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 3.22, 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 3.22 Enter a Boolean Expression

F	Enter Expression
	Specify constraint script or expression that returns a non-zero value for infeasible factor combinations.
	Feature==3 & Price==3
	OK Cancel

4 Make the design table. It excludes the combination of factors you specified, as shown in Figure 3.23. Figure 3.23 No Row Contains L3 for Both Price and Feature

🛗 Custom Design						
Custom Design	• •					
Design Custom Design	•	Feature	Price	Y		
Criterion D Optimal	1	L1	L2	•		
▼Model	2	L1	L3	•		
	3	L2	L3	•		
	4	L3	L2	•		
	5	L3	L1	•		
Columns (3/0)	6	L2	L1	•		

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 3.17) 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 3.17) 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 3.17) of the designer window and select Advanced Options > Split Plot Variance Ratio.
- 3 Enter a positive number 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 3.17) 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.

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 3.24 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 3.24, 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 3.24	Column	Properties	Menu in	the	Column	Info	Dialog
	ooranni	roperties	internet in		ooranni		2

🔝 Stretch		
Stretch' in Table 'Bounce Data'	ОК	Column Properties 👻
Column Name Stretch	Capeet	Formula
		Notes
Data Type Numeric ¥	Apply	Range Check
Normal Norma	Help	List Check
Modeling Type Continuous V		Value Labels
Format Best 🗸 Width 8		Value Ordering
Column Properties V		Value Colors
Role Response Limits		Axis
Description of the last		Coding
		Mixture
contour profilers use these values. Click		Row Order Levels
below to key in values.		Spec Limits
Match Target 🗸		Control Limits
Remove		Response Limits
Importance 1		Design Role
Add Desirability		Factor Changes
Lower 350		Sigma
Middle 450		Units
Upper 550		Distribution
		Time Frequency
		Other

All special column properties are covered in the *JMP User Guide*. 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 3.25 shows the coding values for the Temperature variable in the Reactor 8 Runs data table from the Design Experiment Sample Data.

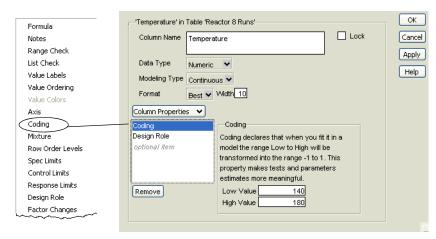


Figure 3.25 Coding Column Property in Column Info Dialog

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 3.26. 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 3.26, where the mixture sum value is 1, the terms are coded as:

$$X_{iL} = (X_{i} - L_{i})/(1 - L)$$
 for the *L* pseudocomponent

 $X_i U = (U_i - X_i)/(U - 1)$ for the U 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, of 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 3.26 Column Info to Create Mixture Column For Analysis

	'CuSO4' in Table 'Donev Mixture Data'	Donev Mixture Data		۲	
Column Properties 👻	Column Name CuSO4	Design Custom Desi	igr 🔺 💂 📄		CuSO4
Formula		Criterion D Optim		1	0.8
Notes	Data Type Numeric 🗸	💌 Model		2	0.44
Range Check			<u> </u>	3	0.2
List Check	Modeling Type Continuous 🗸	Columns (6/1)	•@	4	0.5
Value Labels	Format Best 🗸 Width 10		•@	5	0.2
Value Ordering		📕 Na2S2O3 🛠 🥝	• 🔁	6	0.2
Value Colors	Column Properties 🗸	🚄 Giyoxal 🛠 📀	•@	7	0.8
Axis	Mixture	🆺 Wavelength \star 📀	•@	8	0.56
Coding	Design Role Mixture is specified if the column	A Damping * 📎	•@	9	0.2
Mixture	Factor Changes partisipates in a mixture where it and	🆺 Point Type 🖾		10	0.5
Row Order Levels	optional item other columns add up to a constant	Rows	•.2	11	0.2
Spec Limits	Lower Limit 0.2	All rows	18 . 🔁	12	0.2
Control Limits	Upper Limit 0.8				
Response Limits	Remove Sum of Terms 1				
Design Role	L PseudoComponent Coding				
Factor Changes	U PseudoComponent Coding				
Sigma	(If both checked, it chooses most appropriate)				
Units		Mix	ture 🗸		
Distribution				1	
Time Frequency			Continuous		
Other			Categorical		
/			Blocking		
	Column Properties 👻		Covariate		
	Mixture Design Role		Mixture		
	Design Role Design Role indicates how the column		LUIISLAIIL *		
	Factor Changes is used as a factor in a model for an		Signal		
	optional item experimental design.		Voise		
	Mixture	-	Uncontrolled		
		F	Random Block		
	Remove				

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 3.27, open Bounce Data.jmp from the Design Experiment folder in the sample data installed with JMP.

Figure 3.27 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 in the data grid. The Column Info dialog appears.
- 2 Select Response Limits from the Column Properties drop-down menu.

86 Building Custom Designs

Assigning Column Properties

- 3 Complete the settings in the window, as described in Figure 3.27.
- 4 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.

	ta'	ОК	
Column Name Stretch	Lock	Cancel	Click to select a goal for the response
Data Type Numeric Modeling Type Continuous Format Best V Column Properties V		Apply Help	 variable. For example, if you are in the prediction profiler and want the desired value to be close to zero, select Match Target
Role Response Limits optional item Remove	Response Limits Response Limits are bounds on a response's ange of acceptability. The prediction and contour profilers use these values. Click below to key in values. Match Target V		When you have two responses, enter a number to indicate the amount of weight you want this response to have when JMP computes the over- all desirability
i	Add Desirability Lower 350 . Middle 450 . Upper 550 .		Enter the lower, middle and upper limits as well as the desirability values

Figure 3.27 Define Response Column Values

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 3.28 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 3.28. 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.

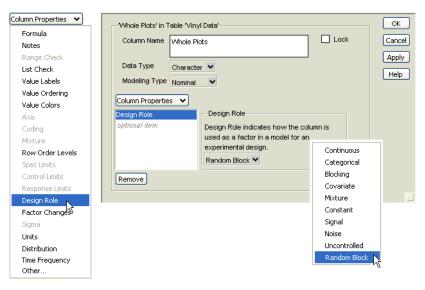


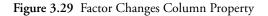
Figure 3.28 Assign a Design Role to a Factor Column

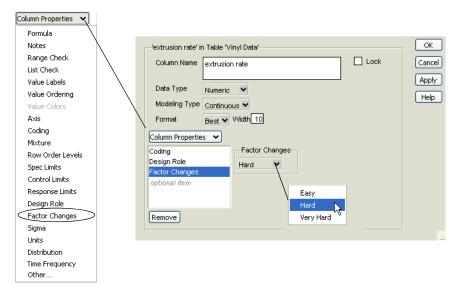
Identify Factor Changes Column Property

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. 72, 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 3.29.
- 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.

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.





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: Changing the Design Criterion (D- or I- Optimality)," p. 78.

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.



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 $2 \ge 2 \ge 2 = 8$ 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
Using Two Continuous Factors and One Categorical Factor
Using Five Continuous Factors93
Create a Fractional Factorial Design
Enter Responses
Enter Factors
Choose a Design
Display and Modify a Design102
Specify Output Options
View the Design Table
Create a Plackett-Burman design
Analysis of Screening Data
Using the Screening Analysis Platform
Using the Fit Model Platform

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.

92 Screening Designs

Screening Design Examples

Figure 4.1 Screening Design Response With Match Target Goal

Responses Add Response 🗸 Remove Number of Responses. Response Name Goal Lower Limit Upper Limit Importance Match Target Depth (In.) optional item Goal Lower Limit Click the Goal box Match Target and choose Match Maximize Match Targ Target Minimize None

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 Mary and John 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

-	Factors			
	Add 2 Continu	ious		
	Add 1 2-Leve	l Categorical		
	Add 1 3-Leve	l Categorical		
	Remove Selected			
	Name	Role	Values	
	IL Operator	Categorical	John	Mary
	Speed	Continuous	3	5
	Current	Continuous	150	165
So	reening Design -Specify Factors			
1	Add a Continuous or C	ategorical factor by cli	cking its button. Double	click
0	on a factor name or le	vel to editit.		
	Continue			

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

Respo	inses		
Facto	rs		
Desig	n List		
hoose a	design	by clicking on its row i	n the list.
Number	Block		Resolution
Of Runs	Size	Design Type	- what is estimable
		Fractional Factorial	3 - Main Effects Only
4			
4 3		Full Factorial	>6 - Full Resolution
	4	Full Factorial Full Factorial	>6 - Full Resolution 5+ - All 2-factor interactions
3	4 2	500	

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.jmp.

Figure 4.4 The Design Data Table

▼Full Factorial	• •					
Design Full Factorial	·	Pattern	Operator	Speed	Current	Depth (in.)
 Screening 	1		John	3	150	-
 Model 	2	+	John	3	165	-
Columns (5/0)	3	-+-	John	5	150	-
Pattern	4	-++	John	5	165	•
Derator *	5	+	Mary	3	150	•
⊿ Speed 🗶	6	+-+	Mary	3	165	-
🚄 Current 🛠	7	++-	Mary	5	150	-
🚄 Depth (in.) 🛠	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

94 Screening Designs

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

Add Response 💙	Remove Number o	f Responses							
Response Nam	ne Goal	Lower Limit	Upper Limit	Importance					
Percent Reacted	Maximize	90	99						
Factors									
Add 1 2-Level Categorical Add 1 3-Level Categorical Remove Selected 1									
Remove Selected									
Name	Role	Values	1						
Name Feed Rate	Continuous	Values 10	15						
Name Feed Rate Catalyst	Continuous Continuous	10 1	2						
Name Feed Rate Catalyst Stir Rate	Continuous Continuous Continuous	10 1 100	2 120						
Name Feed Rate Catalyst Stir Rate Temperature	Continuous Continuous Continuous Continuous	10 1 100 140	2 120 180						
Name Feed Rate Catalyst Stir Rate	Continuous Continuous Continuous	10 1 100	2 120						

- 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 Display and Modify Design panel on the right in Figure 4.6.

Respor	nses			♦ Re	sponses	
Factors	s			♦ Fa	ctors	
Design	List			🔶 🗘 Di	splay and Modify	Design
hoose a	design	by clicking on its row i	n the list.	•	Change Generatir	ng Rules
Number	Block		Resolution	•	Aliasing of Effects	
Of Runs	Size	Design Type	- what is estimable	•	Coded Design	
3		Fractional Factorial	3 - Main Effects Only		•)	
3	4	Fractional Factorial	3 - Main Effects Only	. De	sign Evaluation	
12		Plackett-Burman	3 - Main Effects Only	÷ 0.	tput Options	
16		Fractional Factorial	5 - All 2-factor interactions		• •	
16	8	Fractional Factorial	4 - Some 2-factor interactions	Run	Order:	Randomize
16	4	Fractional Factorial	4 - Some 2-factor interactions	Mak	e JMP Table from desig	n plus
16	2	Fractional Factorial	4 - Some 2-factor interactions	Num	ber of Center Points:	0
32		Full Factorial	>6 - Full Resolution	Num	ber of Replicates:	0
32	16	Full Factorial	5+ - All 2-factor interactions	[m	ke Table	
32	8	Full Factorial	5+ - All 2-factor interactions	IMa	Re Table	
32	4	Full Factorial	4 - Some 2-factor interactions		Back	
32	2	Full Factorial	4 - Some 2-factor interactions			

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

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

Figure 4.7 Output Options for Design Table

🕈 💌 Screening Design	
Responses	
Factors	
Change Generating Rules	
Aliasing of Effects	
Coded Design	Keep the Same
	Sort Left to Right
Design Evaluation	Randomize
	Sort Right to Left
Run Order: Randomize 🗸	Randomize within Blocks
Make JMP Table from design plus	
Number of Center Points: 0	
Number of Replicates: 0	
Make Table	
Back	

10 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.

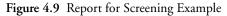
Screening Design Examples

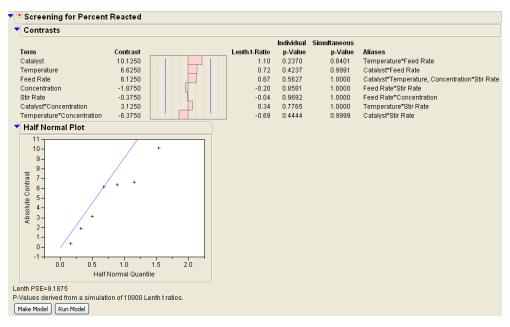
Fractional Factorial	♦ _ ●		Feed		Stir			Percent
Design Fractional Factorial	•	Pattern	Rate	Catalyst	Rate	Temperature	Concentration	Reacted
 Screening 	1	++	10	1	100	180	6	
💌 Model	2	++-	10	1	120	180	3	
Columns (7/0)	3	-++	10	2	100	140	6	
📕 Pattern 🚇	4	-++	10	2	120	140	3	
🖌 Feed Rate 🗴	5	+	15	1	100	140	3	
🚄 Catalyst 🛠	6	+-+-+	15	1	120	140	6	
🚄 Stir Rate 🛠	7	++-+-	15	2	100	180	3	
🥖 Temperature 🛠	8	+++++	15	2	120	180	6	
Concentration * Percent Reacted								

Figure 4.8 JMP Table of Runs for Screening Example

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.

- 11 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.
- 12 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.





Note: Analysis of the screening data is covered in the section "Create a Plackett-Burman design," p. 106 at the end of this chapter.

Create a Fractional Factorial Design

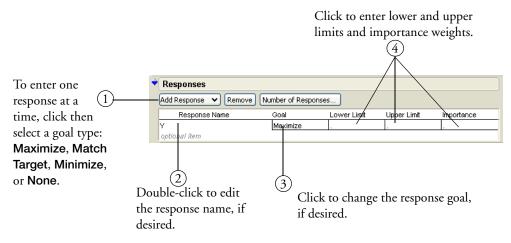
To begin, select **DOE** > Screening Design, or click the Screening Design button on the JMP Starter **DOE** page. Then, see the following sections for each step to create a screening design:

- 1 "Enter Responses," p. 97
- 2 "Enter Factors," p. 98
- 3 "Choose a Design," p. 99
- 4 "Display and Modify a Design," p. 102
- 5 "Specify Output Options," p. 105
- 6 "View the Design Table," p. 106

Enter Responses

To enter responses, follow the steps in Figure 4.10.

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.

Create a Fractional Factorial Design

- 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.

Figure 4.11 Entering Factors

		tors			
To enter factors, type the-	Add	1 Continuous			
number of factors and click	Add	1 2-Level Categ	orical		
Add.	Add	1 3-Level Categ	orical		
	Remo	ve Selected			
Highlight the factor and	Nar	ne	Role	Values	
click the Remove Selected	4 X1	1	Continuous	-1	1
button to remove a factor in	,				
the list.					
	Doubl	e-click to	Click to enter fa	ctor values. To re	move a level, click
	edit th	e factor	it, press the dele	te key on your ke	yboard, then press
	name.		the Return or E	nter key on your	keyboard.

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

creening l Factors Choose	-	gn	
Number	Block		Resolution
Of Runs	Size	Design Type	- what is estimable
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 samples 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.

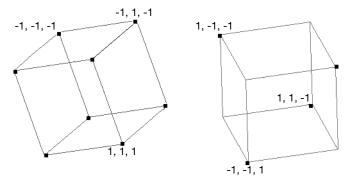
Experiments can therefore be classified by resolution number into three groups:

- 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.

All the fractional factorial designs are *minimum aberration* designs. 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 "Screening Designs," p. 89 of *JMP Statistics and Graphics Guide* for documentation of the Screening platform.

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.

Tuble fill Tuble of Mixed Elever Designs			
Design	Two-Level Factors	Three–Level Factors	
L18 John	1	7	
L18 Chakravarty	3	6	
L18 Hunter	8	4	
L36	11	12	

Table 4.1 Table of Mixed-Level Designs

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	
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.

Create a Fractional Factorial Design

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 Supress Cotter Designs (to uncheck it), as shown to the right.

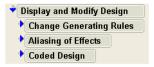
Changing the setting via the red triangle menu applies only to the *current* design. To alter the setting for *all* screening designs:

- 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, click the disclosure buttons ($\diamond \diamond$ on Windows/Linux and $\triangleright \forall$ on the Macintosh) to display the design and show modification options using the Display and Modify Design panel to tailor the design (Figure 4.14).

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, click the disclosure button ($\diamond \diamond$ on Windows/Linux and $\triangleright \bigtriangledown$ on the Macintosh) to reveal the Aliasing of Effects panel. It shows effects and confounding up to two-factor interactions (Figure 4.15).

Display and Mod	lify Design
🕈 Change Gener	rating Rules
Factors Temper	atureConcentration
Feed Rate 🔽	
Catalyst 🔽	
Stir Rate 🔽	
Apply	
Aliasing of Effective	ects
Effects	Aliases
Feed Rate	= Temperature*Concentration
Catalyst	= Stir Rate*Concentration
Stir Rate	= Catalyst*Concentration
Temperature	= Feed Rate*Concentration
Concentration	= Feed Rate*Temperature = Catalyst*Stir Rate
Feed Rate*Catalyst	= Stir Rate*Temperature
Feed Rate*Stir Rate	= Catalyst*Temperature
Coded Design	

Figure 4.15	Generating Ru	iles 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

1			
Aliasing of Effective	ects		
Effects	Aliases		
Feed Rate	= Temperature*Concentration		
Catalyst	= Stir Rate*Concentration		
Stir Rate	= Catalyst*Concentration		
Temperature	= Feed Rate*Concentration		
Concentration	= Feed Rate*Temperature = Catalyst*Stir Rate		
Feed Rate*Catalyst = Stir Rate*Temperature			
Feed Rate*Stir Rate	= Catalyst*Temperature		
x			
Show Confound	ing Pattern		
Output Options			
Run Order:	Randomize 🗸 🗸		
Make JMP Table from c	lesign plus		
Number of Center Points: 0			
Number of Replicates: 0			

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

IO4 Screening Designs

Create a Fractional Factorial Design

Figure 4.17 Enter Order of Confounding in Text Edit Box

JMP: Please Enter a Number	
Show table of aliases to what order?	3
OK Cancel	

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 Pat	tern	◆		
		•	Effect Names	Alias Names
		1	с	= 2 3 5 = 1 4 5
Columns (2/0)		2	1	= 2 3 4 = 4 5
I. Effect Names		3	2	= 1 3 4 = 3 5
I. Alias Names		4	12	= 34 = 135 = 245
		5	3	= 1 2 4 = 2 5
	_	6	13	= 24 = 125 = 345
Rows		7	23	=14=5
All rows	26	8	123	= 4 = 15
Selected	0	9	4	=123=15
Excluded Hidden	0	10	14	= 2 3 = 5
Labelled	0	11	24	=13=125=345
Labored	0	12	124	= 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 Catalyst	Temperature	Concentration
++- -+-+-		-+	+ _
+++ +-+	-+- -++ +	+ - +	- + +
++ +++++	+-+ ++-	-	-
	+++	+	+

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.

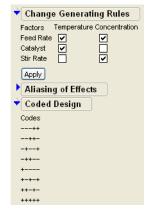
Figure 4.20 Modified Coded Designs and Generating Rules



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

Design Evaluation		
Output Options		Keep the Same
Run Order:	Randomize 🛩	Sort Left to Right
Make JMP Table from design plus		Randomize 😽
Number of Center Points:	0	Sort Right to Left
Number of Replicates:	0	Randomize within Blocks
Make Table		
Back		



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 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

The name of the table is	The column called Pattern shows the pattern of low values
the design type that	denoted "-" and high values denoted "+". Pattern is especially
generated it.	useful as a label variable in plots.

This script \									
allows you to	 Fractional Factorial 	◆ _ ●		Feed		Stir			Percent
	Design Fractional Factorial	•	Pattern	Rate	Catalyst	Rate	Temperature	Concentration	Reacted
screen for active 🔪	` ▼ Screening	1	++	10	1	100	180	6	-
effects.	 Model 	2	++-	10	1	120	180	3	-
	Columns (7/0)	3	-++	10	2	100	140	6	•
	🔒 Pattern 🔎	4	-++	10	2	120	140	3	•
	📕 Feed Rate 🗶	5	+	15	1	100	140	3	•
	🥖 Catalyst 🛠	6	+-+-+	15	1	120	140	6	-
	🥖 Stir Rate 🛠	7	++-+-	15	2	100	180	3	-
	🚄 Temperature 🗙	8	+++++	15	2	120	180	6	-
	🚄 Concentration 🛠								
	Percent Reacted								

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.

🕈 🖻 Screening Design Responses Factors Name Role Values Feed Rate Continuous 10 15 🚄 Catalyst Continuous 1 🚄 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 Block Resolution Of Runs Size Design Type - 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 Fractional Factorial 16 5 - All 2-factor interactions 16 8 Fractional Factorial 4 - Some 2-factor interactions 16 4 Fractional Factorial 4 - Some 2-factor interactions 16 Fractional Factorial 2 4 - Some 2-factor interactions 32 Full Eactorial >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 ontional item Continue Back

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

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.

108 Screening Designs

Analysis of Screening Data

Figure 4.24 No aliasing of the included interactions

Display and Modify De	sign
Change Generating	Rules
Aliasing of Effects	
Effects No effect up to two-way int with any other effect up to t Coded Design Design Evaluation	
Output Options	
Run Order:	Sort Left to Right 🗸 🗸
Make JMP Table from design pl Number of Center Points: Number of Replicates:	us 0
Make Table Back	

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 Placket-Burman Design Table with Results

♥Plackett-Burman Design Plackett-Burman		Pattern	Feed Rate	Catalyst	Stir Rate	Temperature	Concentration	Percent Reacted		Percent Reacted
 Screening 	1	+-	10	1	100	180	3	•	1	69
▼Model	2	+	10	1	120	140	3	•	1	53
Columns (8/0)	3	+-+	10	1	120	140	6	•	Ī	59
Feed Rate *	4	-++	10	2	100	140	6	•		70
Catalyst *	5	-+-++	10	2	100	180	6	•		78
🖌 Stir Rate \star 📃	6	-+++-	10	2	120	180	3	•	Ī	95
🚄 Temperature 🛠 📃	7	++	15	1	100	140	6	•	Ī	63
🖌 Concentration 🛠 📃	8	++-	15	1	100	180	3	•	Ī	61
A Percent Reacted 🗶 🧹	9	+-+++	15	1	120	180	6	•	Ī	42
- Paula	10	++	15	2	100	140	3	•	1	61
■Rows All neuros	11	+++	15	2	120	140	3	•	1	61
All rows 12 Selected 0	12	+++++	15	2	120	180	6	•	1	82

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.

 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

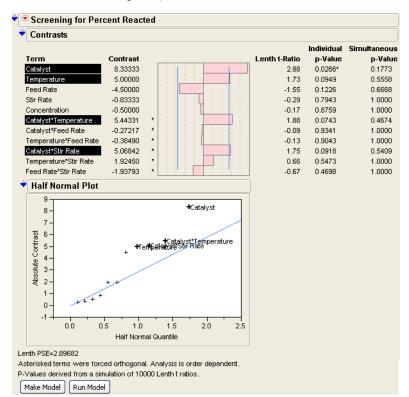
III Screening		
Looking at lots of effects to help deci Select Columns Pattern Feed Rate Catalyst Stir Rate Temperature Concentration Percent Reacted	de which to put in the model. Cast Selected Columns into Roles V Percent Reacted optional numeric X Feed Rate Catalyst Stir Rate Temperature Concentration By optional	Action OK Cancel Remove Recall Help

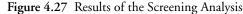
2 Click **OK** to see the Screening platform result shown in Figure 4.27.

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 *Statistics and Graphics Guide* for complete documentation of the Screening analysis platform.

IIO Screening Designs

Analysis of Screening Data





3 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.

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

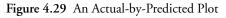
Select Columns	Pick Role Variables Personality: Standard Least Squares
Pattern Feed Rate Catalyst Stir Rate	Y Percent Reacted Emphasis: Effect Leverage
Temperature	Weight optional numeric Help Run Model
Concentration	Freq optional numeric Recall
	By Optional Remove
	Construct Model Effects
	Add Catalyst Temperature
	Cross Catalyst*Temperature
	Catalyst*Stir Rate
	Macros
	Degree 2
	Attributes 💌
	Transform 💌 🦾

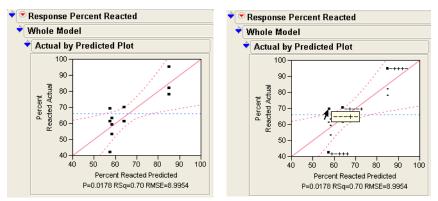
Figure 4.28 Create Fit Model Dialog and Remove Unwanted Effect

The Actual-by-Predicted Plot

The Whole Model actual-by-predicted plot, shown on the left in Figure 4.29, appears at the top of the Fit Model report. You see at a glance that this model fits well. The mean 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.

In this example, the Pattern variable is designated as a label column. To show labels in the plot (on the right in Figure 4.29), shift click to select points of interest, right-click the graph, and select **Row Label**. The pattern variable displayed in the data table serves as the label for each point. Also, you can hover over a point to see the label.





The Scaled Estimates Report

)♦	💌 Roenoneo Porcoi	nt	6
	Regression Reports	۲	
	Estimates	×	
	Effect Screening	►	🗸 Scaled Estimates 📉
	Factor Profiling	×	Normal Plot
	Row Diagnostics	×	Bayes Plot
	Save Columns	Þ	Pareto Plot
	Script	Þ	Pareco Piot

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

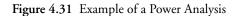
Scaled Estimates	;				
Term	Scaled Estimate	 	Std Error	t Ratio	Prob> t
Intercept	66.166667		2.596739	25.48	<.0001*
Catalyst(1,2)	8.3333333		2.596739	3.21	0.0124*
Temperature(140,180)	5		2.596739	1.93	0.0903
Catalyst*Temperature	5.5		2.596739	2.12	0.0670

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 80% chance of detecting a significant effect for Catalyst.



💌 Cata	alyst(1,2	2)) ♦ 🛡 (Catalyst	(1,2)			
Leverage Plot					• I	.everag	e Plot			
Pow	rer Deta	ails Dialo	g) 🔶	Powe	Details			
Catalys	t(1,2)				Te	st Catalys	t(1,2)			
Click ar	nd Enter 1	,2 or a sec	uence of va	lues for each:	÷	Power	•			
From:	α 0.050	σ 8.995369	õ 8.333333	Number		α	σ	δ	Number/	Power
From: To:	0.050	0.995369	0.3333333	12		0.0500	8.995369	8.333333	12	0.8020
By										\smile
	olve for L	.east Signifi .east Signifi	icant Numbe icant Value Confidence I							
Dor		Cancel		ions of sequence						

Refer to the JMP Statistics and Graphics Guide 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:

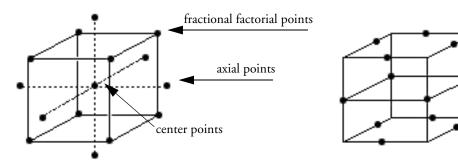
- Center points, for which all the factor values are at the zero (or midrange) value.
- *Axial* (or *star*) *points*, for which all but one factor are set at zero (midrange) and that one factor is set at outer (axial) values.

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.

Central Composite Design

Box-Behnken Design



Contents

A Box-Behnken Design: The Tennis Ball Example
The Prediction Profiler
A Response Surface Plot (Contour Profiler)121
Geometry of a Box-Behnken Design122
Creating a Response Surface Design 123
Enter Responses and Factors124
Choose a Design
Specify Output Options
View the Design Table

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 5.1 shows the completed Response panel and Factors panel.

Figure 5.1 Response and Factors For Bounce Data

💌 Response Surface Des	ign										
Responses											
Add Response 🖌 Remov	e Number of Respon	ses									
Response Name	Goal	Lower Limit	Upper Limit	Importance							
Stretch	Match Target	350	550	1							
▼ Factors											
Name R	ole	Values									
ASilica C	ontinuous	0.7 1.7									
C	ontinuous	1.8 2.8									
Silane C	ontinuous	40 60									

After the response data and factors data are loaded, the Response Surface Design Choice dialog lists the designs in Figure 5.2.

Figure 5.2 Response Surface Design Selection

esponse S Factors —Choose a		2		i Factors Jox-Behnken —Display and Modify Design———		
Number Of Runs		Center Points	Design Type	Output Options Run Order:	Randomize	~
15 16 20		3 2 6	Box-Behnken Central Composite Design CCD-Uniform Precision	Make JMP Table from design plus Number of Center Points: Number of Replicates:	3	
20 23 optional I	6 item	6 9	CCD-Orthogonal Blocks CCD-Orthogonal	Make Table Back		

The Box-Behnken design selected for three effects generates the design table of 15 runs shown in Figure 5.3.

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 5.3).

II8 Response Surface Designs

A Box-Behnken Design: The Tennis Ball Example

💌 Bounce Data	◆ ●					
Design Box-Behnken	•	Pattern	Silica	Sulfur	Silane	Stretch
 Model 	1	0	0.7	1.8	50	570
Columns (5/0)	2	-+0	0.7	2.8	50	285
Pattern	3	+-0	1.7	1.8	50	260
🖌 Silica 🗶	4	++0	1.7	2.8	50	433
🖌 Sulfur 💥	5	0	1.2	1.8	40	422
🥖 Silane 🗱	6	0-+	1.2	1.8	60	351
🚄 Stretch 🛠	7	0+-	1.2	2.8	40	278
	8	0++	1.2	2.8	60	392
Rows	9	-0-	0.7	2.3	40	451
All rows 15	10	+0-	1.7	2.3	40	372
Selected 0	11	-0+	0.7	2.3	60	474
Excluded 0 Hidden 0	12	+0+	1.7	2.3	60	394
Labelled 0	13	000	1.2	2.3	50	398
	14	000	1.2	2.3	50	394
	15	000	1.2	2.3	50	396

Figure 5.3 JMP Table for a Three-Factor Box-Behnken Design

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 Model**.

The standard Fit Model analysis results appear in tables shown in Figure 5.4, 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 5.4 JMP Statistical Reports for a Response Surface Analysis of Bounce Data

💌 Respons	e Stret	tch			
♥ Summar	y of Fit	t			Parameter
RSquare		0.99	9777		Effect Test
RSquare Adj		0.99	9375		
Root Mean S	quare Er	ror 1.98	37461		Source
Mean of Res	ponse	391	.3333		Silica(0.7,1.7)
Observation:	s (or Sur	n Wgts)	15		Silane(40,60)
	of Var	iance			Sulfur(1.8,2.8)
		Sum of			Silane*Silica Sulfur*Silica
Source	DF	Squares I	Mean Square	F Ratio	Sulfur*Silca
Model	9	88453.583	9828.18	2488.146	Silica*Silica
Error	5	19.750	3.95	Prob > F	Silane*Silane
C. Total	14	88473.333		<.0001*	Sulfur*Sulfur
	Fit				Sana Sana
		Sum of		F Ratio	
Source	DF	Squares	Mean Square	0.9792	
Lack Of Fit	3	11.750000	3.91667	7 Prob > F	
Pure Error	2	8.000000	4.0000	0.5411	
Total Error	5	19.750000		Max RSq 0.9999	

Parameter	Estimate	es 📄			
Effect Tests	\$				
			Sum of		
Source	Nparm	DF	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 the chapter "Standard Least Squares: Introduction" of *JMP Statistics and Graphics Guide* for more information about interpretation of the tables in Figure 5.4.

The Response Surface report also has the tables shown in Figure 5.5.

	★ Response Surface	
Summary of the parameter	Coef	
estimates —	Silica(0.7,1.7) Silane(40,60) Sulfur(1.8,2.8) Streto	ch
commuted	Silica(0.7,1.7) 26.5 -0.25 114.5 -40.12	25
		11
Lists the critical values of the	Sulfur(1.8,2.8)35.5 -26.8	75
surface factors and tells the kind	▼ Solution	
of solution (maximum, mini-	Variable Critical Value	
mum, or saddle point). The	Silica(0.7,1.7) 1.7912411	
-	Silane(40,60) 23.424426	
solution for this example is a	Sulfur(1.8,2.8) 2.1986422	
saddle point. The table also	Solution is a SaddlePoint Critical values outside data range	
warns that the critical values	Predicted Value at Solution 360.38388	
given by the solution are outside	Eigenvalues and Eigenvectors	
the range of data values.	Eigenvalue 62.9095 3.2989 -74.9584	
0	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	
C1		т

Figure 5.5 Statistical Reports for a Response Surface Analysis

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.

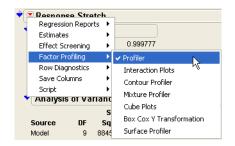
See the chapter "Standard Least Squares: Exploring the Prediction Equation" of *JMP Statistics and Graphics Guide* for details about the response surface analysis tables in Figure 5.5.

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 title bar and select **Factor Profiling > Profiler**, as shown in Figure 5.6.

Figure 5.6 The Profiler

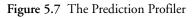


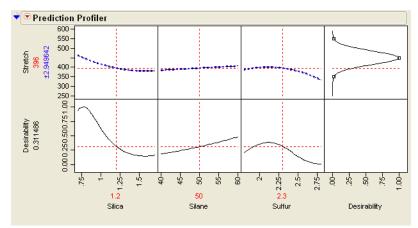
A Box-Behnken Design: The Tennis Ball Example

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 well away 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.





2 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 0.94, Silane from 50 to 46, and Sulfur from 2.3 to 2.1 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.



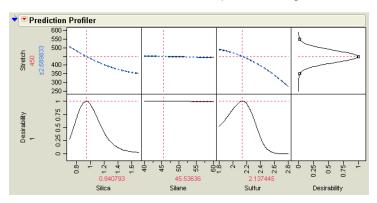


Figure 5.8 Prediction Profiler with Maximum Desirability Set for a Response Surface Analysis

See the JMP Statistics and Graphics Guide for further discussion of the Prediction Profiler.

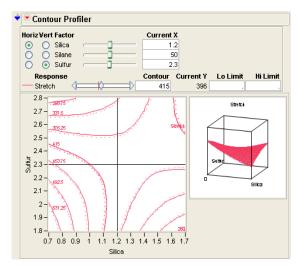
A Response Surface Plot (Contour Profiler)

Another way to look at the response surface is to use the Contour Profiler. Click the red triangle on the Response 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



A Box-Behnken Design: The Tennis Ball Example

Figure 5.10 shows the Contour profile when the Current x values have Lo and Hi limits in effect, which cause shaded regions to show on the contour.

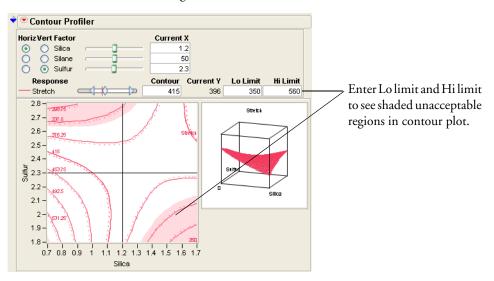


Figure 5.10 Contour Profiler with High and Low Limits

The Prediction Profiler and the Contour Profiler are discussed in more detail in the *JMP Statistics and Graphics Guide*.

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 the *JMP Statistics and Graphics Guide*.

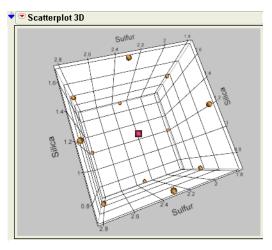


Figure 5.11 Scatterplot 3D Rendition of a Box-Behnken Design for Three Effects

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. 124
- "Choose a Design," p. 124
- "Specify Axial Value (Central Composite Designs Only)," p. 125
- "Specify Output Options," p. 126
- "View the Design Table," p. 126

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. 59

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

Responses	rface Design				
Add Response 🗸	Remove Number o	f Responses			
Response N	ame Goal	Lower Limit	Upper Limit	Importance	
ť	Maxi	inize .			
optional item					
Factors					
Add 1 Continuous					
Add 1 Conti	nuous				
		Values			
Remove Selected)	Values -1	1		
Remove Selected Name X1 X2) Role		1		
Remove Selected Name X1	Role Continuous	-1			
Remove Selected Name X1 X2	Role Continuous Continuous	-1 -1	1		
Remove Selected Name X1 X2 X3	Role Continuous Continuous Continuous	-1 -1	1		
Remove Selected Name X1 X2 X3 ponse Surface De	Role Continuous Continuous Continuous	-1 -1	1		
Remove Selected Name X1 X2 X3 ponse Surface De Specify Factors	Role Continuous Continuous Continuous	-1 -1 -1	1		
Remove Selected Name X1 X2 X3 ponse Surface De Specify Factors Decify desired nurr	Role Continuous Continuous Continuous	-1 -1	1		
Remove Selected Name X1 X2 X3 ponse Surface De Specify Factors	Role Continuous Continuous Continuous	-1 -1 -1	1		

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

Choose :	a Desig	n	
Number	Block	Center	
Of Runs	Size	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
optional .	item		
Continue			
Back			

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

0	Central Composite Desig Display and Modify D		
	Axial Value:	1.000	
	 Rotatable 	1.682	
	 Orthogonal 	1.287	
	 On Face 	1.000	
	 User Specified 		
	Inscribe		

- **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 shrunken 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

Response Surface Design	
3 Factors	
Box-Behnken	
Display and Modify Design	
Output Options	
Run Order:	Randomize 🗙 🗙
Make JMP Table from design plus	
Number of Center Points:	3
Number of Replicates:	0
Make Table	
Back	

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 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 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

			umn called Pa				0		
		all the co	odings with "	+" for h	nigh, "-	-" for l	ow fact	tor, "a" and	d "A" for low
		and higł	n axial values,	and "0	" for n	nidrang	ge. Patt	tern is suit	able to use as
The name of the	table	a label v	ariable in plo	ts becau	ise who	en vou	hover	over a poin	nt in a plot of
is the design type			ors, the patter			•		-	-
generated it.	, that			1					. r
generated it.	\backslash			D			1	1	
				K	uns are	e in a ra	andom	order.	The Y column
This script fits a			۰. ۲					/	
model using the	Box-Beh		·						is for recording
values in the	Design Model	Box-Behnken		Pattern/	X1	X2	X3	Y	experimental
design table.	Moder		2	-0-	-1 0	0	-1 0		results.
8	Columns	(5/0)	3	+-0	1	-1	0		-
	🖺 Pattern	<i>•</i>	4	0	-1	-1	0	•	
	🔺 X1 🛠		5	0-+	0	-1	1	•	
771	X3 *		6	-+0	-1	1	0	•	
There are \leq	<u>4</u> Y *		7	+0-	1	0	-1	•	
three center			8	0+-	0	1	-1	•	
points.	Rows		9	000	0	0	0	•	_
	All rows Selected	15		0	0	-1	-1	•	-
	Excluded		11	++0	1	1	0	•	-
	Hidden	0	12	0++ +0+	1	0	1		-
	Labelled	0		000	0	0	0		
			15		-1	0	1	•	
									f



Full Factorial Designs



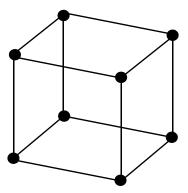
A full factorial design contains all possible combinations of a set of factors. This is the most fool proof design approach, but it is also the most costly in experimental resources. The full factorial designer supports both continuous factors and categorical factors with up to nine levels.

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^{t} 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.



Contents

The Five-Factor Reactor Example 131
Analyze the Reactor Data
Creating a Factorial Design
Enter Responses and Factors
Select Output Options
Make the Table

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. 89). 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 (installed with JMP), open Reactor Response.jmp found in the Design Experiment Sample Data folder.
- 4 Click the red triangle icon on the Full Factorial Design title bar and select Load Factors.
- 5 In the sample data (installed with JMP), open Reactor Factors.jmp found in the Design Experiment folder.

The completed dialog should look like the one shown in Figure 6.1.

Figure 6.1 Full Factorial Example Response and Factors Panels

	Remove Number of Re			
Response Name	Goal	Lower Limit	Upper Limit	Importance
ercent Reacted	Maximize	90	99	1
actors				
ontinuous 🗸 Cate	gorical 🗸 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	
Factorial Design				
Specify Factors				

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.

n	All De stavist De size		
	ull Factorial Design	Keep the Same	
2	2x2x2x2x2 Factorial		
	Output Options	Sort Left to Right	
	Run Order: Randomize V	Randomize	
	Number of Runs: 32	Sort Right to Left	
	Number of Center Points: 0		
	Number of Replicates: 0		
	Make Table		
	Back		

The Five-Factor Reactor Example

7 Click Make Table.

The design data table (Figure 6.2) 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.

Reactor 32 Runs] ♦∖ ऱ							Percent
Design 2x2x2x2x2 Factori	al 🕤 🔪	Pattern	Feed Rate	Catalyst	Stir Rate	Temperature	Concentration	Reacted
▼Fit Model	1		10	1	100	140	3	6
	2	+	10	1	100	140	6	5
Columns (7/0)	3	+-	10	1	100	180	3	6
🔥 Pattern 🖉	4	++	10	1	100	180	6	4
🚄 Feed Rate 🗶	5	+	10	1	120	140	3	6
🖌 Catalyst 🗱	6	+-+	10	1	120	140	6	5
🖌 Stir Rate 🗶 🖌 Temperature 🗶	7	++-	10	1	120	180	3	6
Concentration 🛪	8	+++	10	1	120	180	6	4
Percent Reacted 🗶	9	-+	10	2	100	140	3	E
-	10	-++	10	2	100	140	6	7
Rows	11	-+-+-	10	2	100	180	3	9
	2 12	-+-++	10	2	100	180	6	7
		-++	10	2	120	140	3	6
	-	-++-+	10	2	120	140	6	6
	_		10	2	120	180	3	

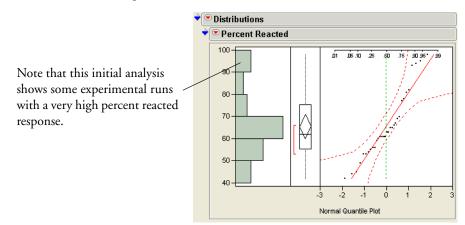
Figure 6.2 Partial Listing of Reactor 32 Runs.jmp from the Sample Data Folder

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 6.3.

Figure 6.3 Distribution of Response Variable for Reactor Data



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 6.4.
- 5 The probability to enter a factor (Prob to Enter) in the model should be 0.05.
- 6 The probability to remove a factor (Prob to Leave) should be 0.1.

Figure 6.4 Stepwise Control Panel



- 7 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 6.4).
- 8 Check the boxes for the main effects of the factors as shown in Figure 6.5.
- 9 Click Go.

Figure 6.5 Starting Model For Stepwise Process

Current E	stimate	es						
SSE	DFE	MSE	RSquare	RSquare	Adj	Ср	AIC	
2642.75	26	101.64423	0.6192	0.5	460	210.64	153.2429	
Lock Entered	Paramete	er		Estimate	nDF	SS	"F Ratio"	"Prob>F
	Intercept	t		65.5	1	0	0.000	1.000
	Feed Rat	te(10,15)		-0.6875	1	15.125	0.149	0.702
	Catalyst((1,2)		9.75	1	3042	29.928	0.000
	Stir Rate	(100,120)		-0.3125	1	3.125	0.031	0.862
	Tempera	ture(140,180	0	5.375	1	924.5	9.095	0.005
	Concent	ration(3,6)		-3.125	1	312.5	3.074	0.091
	Catalyst'	*Feed Rate		0	1	15.125	0.144	0.707
	Stir Rate	*Feed Rate		0	1	4.5	0.043	0.838
	Tempera	ture*Feed Ra	ate	0	1	6.125	0.058	0.811
	Concent	ration*Feed R	Rate	0	1	0.125	0.001	0.972
	Stir Rate	*Catalyst		0	1	6.125	0.058	0.811
	Tempera	iture*Catalyst		0	1	1404.5	28.357	0.000
	Concent	ration*Cataly:	st	0	1	32	0.306	0.584
	Concent	ration*Tempe	rature	0	1	968	14.450	0.000

The mixed stepwise procedure removes insignificant main effects and adds important interactions. The end result is shown in Figure 6.6. Note that the Feed Rate and Stir Rate factors are no longer in the model.

I34Full Factorial Designs

The Five-Factor Reactor Example

Figure 6.6 Model After Mixed Stepwise Regression

Cu	rrent E	stimate	es						
	SSE	DFE	MSE	RSquare	RSquare	Adj	Ср	AIC	
	288.5	26	11.096154	0.9584	0.9	504	5.1781818	82.36669	
Lock I	Entered	Paramete	er		Estimate	nDF	- SS	"F Ratio"	"Prob>F
\checkmark	\checkmark	Intercept			65.5	1	0	0.000	1.0000
		Feed Rat	te(10,15)		0	1	15.125	1.383	0.2506
	✓	Catalyst((1,2)		9.75	2	2 4446.5	200.362	0.0000
		Stir Rate	(100,120)		0	1	3.125	0.274	0.6054
	✓	Tempera	ture(140,180	0	5.375	3	3297	99.043	0.0000
	✓	Concent	ration(3,6)		-3.125	2	2 1280.5	57.700	0.0000
		Catalyst'	Feed Rate		0	2	2 30.25	1.406	0.2647
		Stir Rate	*Feed Rate		0	3	3 22.75	0.656	0.5872
		Tempera	ture*Feed Ra	ate	0	2	2 21.25	0.954	0.3993
		Concent	ration*Feed R	late	0	2	2 15.25	0.670	0.5212
		Stir Rate	*Catalyst		0	2	2 9.25	0.397	0.6763
	✓	Tempera	ture*Catalyst		6.625	1	1404.5	126.575	0.000
		Concent	ration*Cataly:	st	0	1	32	3.119	0.0896
	✓	Concent	ration*Tempe	rature	-5.5	1	968	87.237	0.000

10 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 6.7).

Figure 6.7 Fitting a Prediction Model

👻 💌 Model Specificatior	ı			
Select Columns	Pick Role Variables	Personality:	Standard Least Squ	ares 🗸
Pattern Peed Rate Catalyst	Y Percent Reacted optional	Emphasis:	Effect Leverage	*
Stir Rate Temperature Concentration Percent Reacted	Weight optional Numeric Freq optional Numeric By optional	Help Remove	Run Model	
	Construct Model Effects Add Catalyst Temperature Cross Concentration Temperature*Cate Concentration*Te Macros Degree Attributes Transform No Intercept			

11 Click Run Model to see the analysis for a candidate prediction model (Figure 6.8).

The figure on the left in Figure 6.8 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 6.8 shows a table of model coefficients and their standard errors (labeled Parameter Estimates). All effects selected by the stepwise process are highly significant.

Estimate Std Error t Ratio Prob>|t|

111.23

16.56

9.13

-5.31

11.25

-9.34

< 0001*

<.0001*

 $< 0001^{*}$

<.0001*

<.0001*

<.0001*

0.588859

0.588859

0.588859

0.588859

0.588859

0.588859

65.5

9.75

5.375

-3.125

6.625

-5.5

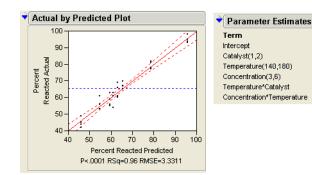


Figure 6.8 A	Actual by Predicted Plo	t and Prediction Param	eter Estimates Table
--------------	-------------------------	------------------------	----------------------

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**, as shown in Figure 6.9.

Figure 6.9 Selecting the Profiler

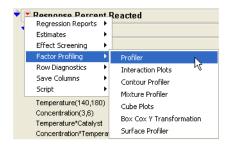
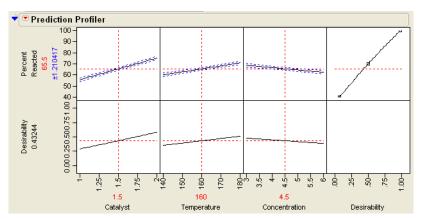


Figure 6.10 shows the profiler's initial display. The Prediction Profiler is discussed in more detail in the chapter "Response Surface Designs," p. 115, and in the chapter "Standard Least Squares: Exploring the Prediction Equation" of *JMP Statistics and Graphics Guide*.

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.

136 Full Factorial Designs

Creating a Factorial Design

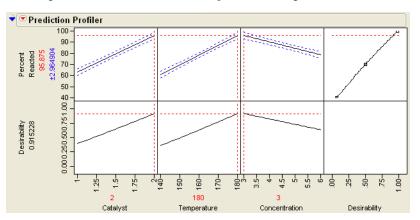


Figure 6.11 Viewing the Prediction Profiles at the Optimum Settings

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 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. 136
- "Select Output Options," p. 137
- "Make the Table," p. 138

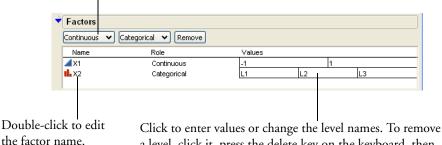
Enter Responses and Factors

The steps for entering responses are outlined in "Enter Responses and Factors into the Custom Designer," p. 59

The steps for entering factors in a full factorial design are unique to this design. To add factors, see Figure 6.12.

Figure 6.12 Entering Factors in a Full Factorial Design

To enter factors, click either the **Continuous** button or the Categorical button and select a factor type, level 2 - 9.



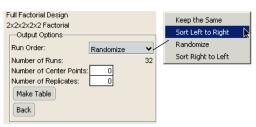
a level, click it, press the delete key on the keyboard, then press the Return or Enter key on the keyboard.

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 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.

Add additional points to the data table with these options:

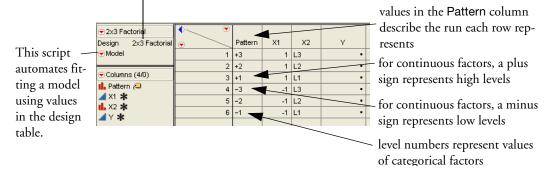
- Number of Center Points Specifies additional runs placed at the center of each continuous factor's range.
- Number of Replicates Specify the number of times to replicate the entire design, including centerpoints. Type the number of times you want to replicate the design in the associated text box.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

The name of the table is the design type that generated it.



Chapter **7** Mixture Designs



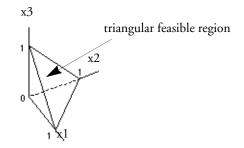
The mixture designer supports experiments with factors that are ingredients in a mixture. You can choose among several classical mixture design approaches, such as simplex, extreme vertices, and lattice. For the extreme vertices approach you can supply a set of linear inequality constraints limiting 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.



Contents

Mixture Design Types
The Optimal Mixture Design141
The Simplex Centroid Design142
Creating the Design
Simplex Centroid Design Examples143
The Simplex Lattice Design
The Extreme Vertices Design
Creating the Design147
An Extreme Vertices Example with Range Constraints
An Extreme Vertices Example with Linear Constraints
Extreme Vertices Method: How It Works 151
The ABCD Design 151
The ABCD Design. 151 Creating Ternary Plots 152
Creating Ternary Plots
Creating Ternary Plots 152 Fitting Mixture Designs 153 Whole Model Tests and Analysis of Variance Reports 154 Understanding Response Surface Reports 155 A Chemical Mixture Example 155 Create the Design 155
Creating Ternary Plots 152 Fitting Mixture Designs 153 Whole Model Tests and Analysis of Variance Reports 154 Understanding Response Surface Reports 155 A Chemical Mixture Example 155 Create the Design 155 Analyze the Mixture Model 157
Creating Ternary Plots152Fitting Mixture Designs153Whole Model Tests and Analysis of Variance Reports154Understanding Response Surface Reports155A Chemical Mixture Example155Create the Design155Analyze the Mixture Model157The Prediction Profiler158

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 7.1. 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 7.1 Mixture Design Selection Dialog

- Choose Mixture E	Decign Tumo
	Jesigii Type
Optimal	Create a design tailored to meet specific requirements.
Simplex Centroid	Run each ingredient without mixing, then mix equal K proportions of K ingredients at a time to the specified limit.
Simplex Lattice	Triangular grid. Number Specify number of levels per factor: of Levels 5
Extreme Vertices	Find the vertices of the simplex. Then Degree add the mid-points of the edges and 2 averages of vertices to the specified degree. 2 Linear Constraint Add linear constraints on the relative proportions
ABCD Design Back	A mixture design for factor screening.

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. 59.
- 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.

The Simplex Centroid Design

Adding Effects to the Model

Initially, the Model panel lists only the main effects corresponding to the factors you entered, as shown in Figure 7.2.

Figure 7.2 The Model Panel

÷	Model	
	Main Effects Interactions 🗙 Cross Powers 🗸	Scheffe Cubic Remove Term
	Name	Estimability
	X1	Necessary
	X2	Necessary
	X3	Necessary

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. 59.
- 3 After you enter responses and factors, click **Continue**.
- 4 Enter the number of ingredients in the box labeled K. JMP will create runs for each ingredient with-

out 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 7.3.

Figure 7.3 Example of Factor Settings and Output Options

	ire Desigi	n					
Response	onses						
🔸 Facto	rs						
Mixture Des	sign						
3Factors							
🔶 Facto	r Setting	s					
Run	X1	X2	X3				
1	1.00000	0.00000	0.00000				
2	0.00000	1.00000	0.00000				
3	0.00000	0.00000	1.00000				
4	0.50000	0.50000	0.00000				
5	0.50000	0.00000	0.50000			Г	
6	0.00000	0.50000	0.50000				Keep the Same
7	0.33333	0.33333	0.33333				Sort Left to Right
Display	and Modify	Design					Randomize
Output Op	tions					▶	Sort Right to Left
Run Order	r:		Randomize	e	*		Randomize within Blocks
Make JMP	Table from	design plus					
Number of	f Replicates	:	0				
Make Ta	ble Back	1					
make ra	Dack	,					

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 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 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 7.4) 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.

I44Mixture Designs

The Simplex Centroid Design

Ru	n 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	з	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

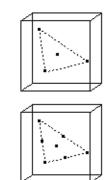


Figure 7.4	Three-Factor	Simplex (Centroid I	Designs of	Degrees 1	and 2

To generate the two sets of runs in Figure 7.4:

- 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 7.5.
- 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 7.5.

Figure 7.5 Create Simplex Centroid Designs of Degrees 1 and 2

Factors Facto	r Settings	5		3Factors	r Setting	6	
Run	X1	Х2	ХЗ	Run	X1	X2	X
1	1.00000	0.00000	0.00000	1	1.00000	0.00000	0.0000
2	0.00000	1.00000	0.00000	2	0.00000	1.00000	0.0000
3	0.00000	0.00000	1.00000	3	0.00000	0.00000	1.0000
4	0.33333	0.33333	0.33333	4	0.50000	0.50000	0.0000
				5	0.50000	0.00000	0.5000
				6	0.00000	0.50000	0.5000
				7	0.33333	0.33333	0.3333

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 7.6.

 Simplex Centr 	oid	◆ ♥						
Design Simpl	lex Centroid	•	X1	X2	Х3	X4	X5	Y
💌 Model		1	0.33333333	0.33333333	0	0	0.33333333	•
		2	0	0	0.5	0.5	0	•
Columns (6/0))	3	0.25	0.25	0	0.25	0.25	•
🚄 X1 🗶		4	0.25	0	0.25	0.25	0.25	
🔺 X2 🛠		5	0.2	0.2	0.2	0.2	0.2	•
⊿ X3 ≭ ⊿ X4 ≭		6	0	0	1	0	0	•
🚄 X4 🛪		7	0.5	0.5	0	0	0	-
AY *		8	0	0	0.33333333	0.33333333	0.33333333	-
<u> </u>		9	0	0	0	0	1	-
Rows		10	0	0.33333333	0.33333333	0	0.33333333	-
All rows	31	11	0	0.5	0	0.5	0	•
Selected	0	12	0.25	0.25	0.25	0.25	0	•
Excluded	0	13	0.25	0.25	0.25	0	0.25	•
Hidden	0	14	0	1	0	0	0	-
Labelled	0	15	0.33333333	0	0	0.33333333	0.33333333	•
		16	0	0	0	1	0	-
		17	0.33333333	0.33333333	0.33333333	0	0	-
		10	0 00000000	0	0.00000000	0 00000000		_

Figure 7.6 Partial Listing of Factor Settings for Five-Factor Simplex Centroid Design

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. 59.
- 3 Click Continue.
- 4 Specify the number of levels you want in the Mixture Design Type dialog (Figure 7.1) and click **Simplex Lattice**.

Figure 7.7 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.

146 Mixture Designs

The Simplex Lattice Design

Figure 7.7 Three-Factor Simplex Lattice Designs for I	Factor Levels 3, 4, and 5
---	---------------------------

acto	r Setting	5		1	Facto	r Setting	s		Facto	r Setting:	5	
Run	X1	X2	X3		Run	X1	X2	X3	Run	X1	X2	x
1	0.00000	0.00000	1.00000		1	0.00000	0.00000	1.00000	1	0.00000	0.00000	1.0000
2	0.00000	0.33333	0.66667		2	0.00000	0.25000	0.75000	2	0.00000	0.20000	0.8000
3	0.00000	0.66667	0.33333		3	0.00000	0.50000	0.50000	3	0.00000	0.40000	0.6000
4	0.00000	1.00000	0.00000		4	0.00000	0.75000	0.25000	4	0.00000	0.60000	0.400
5	0.33333	0.00000	0.66667		5	0.00000	1.00000	0.00000	5	0.00000	0.80000	0.2000
6	0.33333	0.33333	0.33333		6	0.25000	0.00000	0.75000	6	0.00000	1.00000	0.0000
7	0.33333	0.66667	0.00000		7	0.25000	0.25000	0.50000	7	0.20000	0.00000	0.8000
8	0.66667	0.00000	0.33333		8	0.25000	0.50000	0.25000	8	0.20000	0.20000	0.6000
9	0.66667	0.33333	0.00000		9	0.25000	0.75000	0.00000	9	0.20000	0.40000	0.400
10	1.00000	0.00000	0.00000		10	0.50000	0.00000	0.50000	10	0.20000	0.60000	0.200
	_				11	0.50000	0.25000	0.25000	11	0.20000	0.80000	0.000
ſ	4	-1			12	0.50000	0.50000	0.00000	12	0.40000	0.00000	0.600
					13	0.75000	0.00000	0.25000	13	0.40000	0.20000	0.400
		_			14	0.75000	0.25000	0.00000	14	0.40000	0.40000	0.200
		•II			15	1.00000	0.00000	0.00000	15	0.40000	0.60000	0.000
		\rightarrow							16	0.60000	0.00000	0.400
						Æ			17	0.60000	0.20000	0.200
									18	0.60000	0.40000	0.000
							.		19	0.80000	0.00000	0.200
									20	0.80000	0.20000	0.000
							-+		21	1.00000	0.00000	0.000



Figure 7.8 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 7.8 JMP Design Table for Simplex Lattice with Five Variables, Order (Degree) 3

5Factors						5						
	r Setting	•					17	0.00000	0.66667	0.00000	0.00000	0.33333
Facio							18	0.00000	0.66667	0.00000	0.33333	0.00000
Run	X1	X2	X3	X4	X5		19	0.00000	0.66667	0.33333	0.00000	0.00000
1	0.00000	0.00000	0.00000	0.00000	1.00000		20	0.00000	1.00000	0.00000	0.00000	0.00000
2	0.00000	0.00000	0.00000	0.33333	0.66667		21	0.33333	0.00000	0.00000	0.00000	0.66667
3	0.00000	0.00000	0.00000	0.66667	0.33333		22	0.33333	0.00000	0.00000	0.33333	0.33333
4	0.00000	0.00000	0.00000	1.00000	0.00000		23	0.33333	0.00000	0.00000	0.66667	0.00000
5	0.00000	0.00000	0.33333	0.00000	0.66667		24	0.33333	0.00000	0.33333	0.00000	0.33333
6	0.00000	0.00000	0.33333	0.33333	0.33333		25	0.33333	0.00000	0.33333	0.33333	0.00000
7	0.00000	0.00000	0.33333	0.66667	0.00000		26	0.33333	0.00000	0.66667	0.00000	0.00000
8	0.00000	0.00000	0.66667	0.00000	0.33333		27	0.33333	0.33333	0.00000	0.00000	0.33333
9	0.00000	0.00000	0.66667	0.33333	0.00000		28	0.33333	0.33333	0.00000	0.33333	0.00000
10	0.00000	0.00000	1.00000	0.00000	0.00000		29	0.33333	0.33333	0.33333	0.00000	0.00000
11	0.00000	0.33333	0.00000	0.00000	0.66667		30	0.33333	0.66667	0.00000	0.00000	0.00000
12	0.00000	0.33333	0.00000	0.33333	0.33333		31	0.66667	0.00000	0.00000	0.00000	0.33333
13	0.00000	0.33333	0.00000	0.66667	0.00000		32	0.66667	0.00000	0.00000	0.33333	0.00000
14	0.00000	0.33333	0.33333	0.00000	0.33333		33	0.66667	0.00000	0.33333	0.00000	0.00000
15	0.00000	0.33333	0.33333	0.33333	0.00000		34	0.66667	0.33333	0.00000	0.00000	0.00000
16	0.00000	0.33333	0.66667	0.00000	0.00000		35	1.00000	0.00000	0.00000	0.00000	0.00000
					1							

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. 59. Remember to enter the upper and lower limits in the factors panel (see Figure 7.9).
- 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.

The Extreme Vertices Design

- 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 7.9, 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 7.9.
- 3 Click Continue.
- 4 Enter '4' in the **Degree** text box (Figure 7.9).

Figure 7.9 Example of Five-factor Extreme Vertices

Responses			
Factors			
Name	Role	Values	
⊿ X1	Mixture	0.05	0.25
⊿ X2	Mixture	0.1	0.3
4 X3	Mixture	0.1	0.3
∠X4	Mixture	0.1	0.4
4 X5	Mixture	0.05	0.25
Optim Simplex Ci	entroid Run each ingr proportions of	gn tailored to meet specif edient without mixing, the Kingredients at a time to	n mix equal K the specified limit. 4
Simplex Co	entroid Run each ingr proportions of attice Triangular grid Specify numbe	edient without mixing, the K ingredients at a time to I. er of levels per factor: o	n mix equal K the specified limit. 4 Number f Levels 3
Simplex Ci	entroid Run each ingr proportions of attice Triangular grid Specify number Vertices Find the vertice add the mid-p	edient without mixing, the K ingredients at a time to a of levels per factor: es of the simplex. Then oints of the edges and vitices to the specified de	n mix equal K the specified limit. 4 Number f Levels 3 Degree 4 gree,
Simplex Co	entroid Run each ingr proportions of attice Triangular grid Specify number Vertices Find the vertice add the mid-p	edient without mixing, the K ingredients at a time to ar of levels per factor: es of the simplex. Then oints of the edges and writces to the specified den aint Add linear constraint	n mix equal K the specified limit. 4 Number f Levels 3 Degree 4

- 5 Click Exreme Vertices.
- 6 Select Sort Left to Right from the Run Order menu.

7 Click **Make Table.** Figure 7.10 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.

 Extreme Vertices 	◆ ●						
Design Extreme Vertices	•	X1	X2	Х3	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	0.26666667	0.36666667	0.21666667	
⊿ X1 🗱	4	0.05	0.1	0.3	0.4	0.15	
🚄 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 ≭ ⊿ Y ≭	8	0.05	0.2	0.2	0.4	0.15	
4 ' T	9	0.05	0.2	0.3	0.4	0.05	
Rows	10	0.05	0.25	0.1	0.4	0.2	
All rows 116	11	0.05	0.25	0.25	0.4	0.05	
Selected 0	12	0.05	0.26666667	0.1	0.36666667	0.21666667	
Excluded 0	13	0.05	0.26666667	0.26666667	0.36666667	0.05	
Hidden 0	14	0.05	0.3	0.1	0.4	0.15	
Labelled 0	15	0.05	0.3	0.15	0.4	0.1	
	16	0.05	0.3	0.2	0.4	0.05	
	47	0.05	0.4	0.05	0.25	0.05	

Figure 7.10 JMP Design Table for Extreme Vertices with Range Constraints

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 **Exreme 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 7.11) 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 7.11 JMP Design Table for 10-Run Subset of the 116 Current Runs

 Extreme Vertices 	◆ _ ▼						
Design Extreme Vertices		X1	X2	XЗ	X4	X5	Y
▼Model	1	0.05	0.1	0.3	0.4	0.15	-
Columns (6/0)	2	0.05	0.3	0.2	0.4	0.05	-
X1 *	3	0.05	0.3	0.3	0.1	0.25	-
∠ ×1 ★ X2 ★	4	0.05	0.3	0.1	0.3	0.25	-
▲ X3 ★	5	0.15	0.1	0.1	0.4	0.25	-
A X4 *	6	0.25	0.1	0.3	0.1	0.25	-
⊿ X5 🗱	7	0.25	0.1	0.3	0.3	0.05	-
🔺 Y \star	8	0.25	0.2	0.1	0.4	0.05	-
	9	0.25	0.3	0.1	0.1	0.25	-
Rows	10	0.25	0.3	0.3	0.1	0.05	-
All rows 10							
Selected 0							

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 Peipel (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 Peipel Example

$X1 \ge 0.1$	$90 \leq 85*X1 + 90*X2 + 100*X3$
$X1 \leq 0.5$	$85*X1 + 90*X2 + 100*X3 \le 95$.4 $\le 0.7*X1 + X3$
$X2 \ge 0.1$	$.4 \le 0.7$ $\times 1 + 1.5$
$X2 \leq 0.7$	
$X3 \leq 0.7$	

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 7.12.
- 4 Click the **Extreme Vertices** button.
- 5 Change the run order to **Sort Right to Left**, and keep the sample size at 13 to see the 13-run factor settings like those shown on the right in Figure 7.12.
- 6 Click Make Table.

Figure 7.12 Constraints

Extreme Vertices Find the vertices of the simplex. Then Degree add the mid-points of the edges and		r Setting	s		
averages of vertices to the specified degree.	Run	X1	X2	Х3	
	1	0.50000	0.10000	0.40000	
Linear Constraint) Add linear constraints on the relative proportions of ingredients. Click once for each constraint.	2	0.50000	0.25000	0.25000	
of ingredients. Click once for each constraint.	3	0.26667	0.10000	0.63333	
ABCD Design A mixture design for factor screening.	4	0.10000	0.35000	0.55000	
	5	0.10000	0.57000	0.33000	
Back	6	0.33333	0.50000	0.16667	
	7	0.30000	0.31167	0.38833	
85 X1 + 90 X2 + 100 X3 ≥= ♥ 90	8	0.10000	0.46000	0.44000	
85 X1 + 90 X2 + 100 X3 <= ¥ 95	9	0.50000	0.17500	0.32500	
	10	0.38333	0.10000	0.51667	
0.7 X1 + 0 X2 + 1 X3 ≥= ♥ 0.4	11	0.41667	0.37500	0.20833	
	12	0.18333	0.22500	0.59167	
	13	0.21667	0.53500	0.24833	
	Display a	and Modify I	Design		
	Output Op	tions			
	Run Order	:	Sor	t Right to L	eft 🗸
	D Optimal :	Subset			
	Choose de	esired samp	ole size: 13	3	
	Alaka Ta		Cuture C	-	
	Make Ta	able Find	Subset E	lack	

This example is best understood by viewing the design as a ternary plot, as shown at the end of this chapter, in Figure 7.14. 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. 152, 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 Peipel (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 Peipel (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. 59.
- 3 After you enter responses and factors, click Continue.
- 4 Click the **ABCD** 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 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.

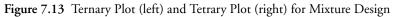
7 Specify Number of Replicates. The number of replicates is the number of times to replicate the entire design, including centerpoints. Type the number of times you want to replicate the design in

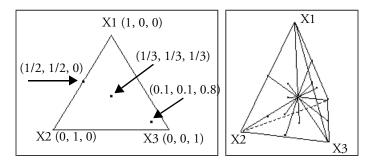
the associated text box. One replicate doubles the number of runs.

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 7.14 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 7.13 illustrates a ternary plot.





The Peipel (1979) example referenced in "An Extreme Vertices Example with Linear Constraints," p. 150 is best understood by the ternary plot shown in Figure 7.14.

To view a mixture design as a ternary plot:

- 1 Create the Peipel mixture data as shown previously, or open the table called Peipel.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 Peipel data had additional constraints, entered as linear constraints for the extreme vertices design. Each constraint is a line, drawn by the line graphics tool on the ternary plot. 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.

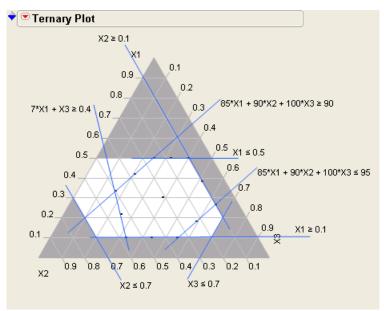


Figure 7.14 Diagram of Ternary Plot Showing Peipel Example Constraints

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 X1*X1)
- includes all the cross terms (such as X1*X2)

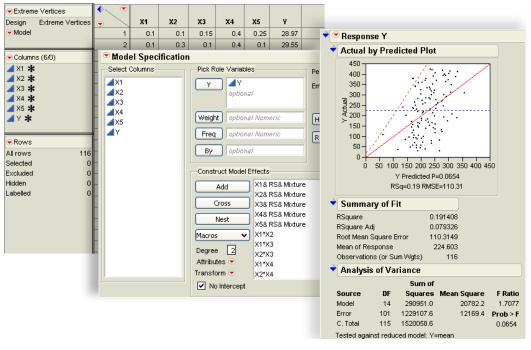
To fit a model:

- 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 Model on the Fit Model dialog.

154 **Mixture Designs** Fitting Mixture Designs

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. Your values may be different from those shown below.

Figure 7.15 Fitting a Mixture Design



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.

The contour plot feature of these reports does not fold to handle mixtures. If you want a contour plot of the surface, you can do any of the following:

- Save the model prediction formula and then create a ternary plot with **Graph > Ternary Plot**.
- Refit the model using a full response surface that omits the last factor.
- Create a separate contour plot with **Graph > Contour Plot**, and add points to make the plot less granular.

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: $0.409 \le x1 \le 0.849$, $0 \le x2 \le 0.252$, and $0.151 \le x3 \le 0.274$.

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

Add Response 🖌 Re	move NRe	esponses			
Response Name		Goal	Lower Limit	Upper Limit	Importance
Y		Maximize			
optional item					
Factors					
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 Exreme Vertices.

A Chemical Mixture Example

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

Extreme Vertices	• •				
Design Extreme Vertices	•	p1	p2	p3	Y
▼ Model	1	0.474	0.252	0.274	•
Columna (40)	2	0.726	0	0.274	•
Columns (4/0)	3	0.849	0	0.151	•
/ p1 *	4	0.597	0.252	0.151	•
_ p2 ↓ _ p3 ★	5	0.6	0.126	0.274	•
Y *	6	0.6615	0.126	0.2125	•
	7	0.7875	0	0.2125	•
■Rows	8	0.723	0.126	0.151	•
All rows 9	9	0.5355	0.252	0.2125	•

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.

8 Run the LabelMixturePoints.jsl to see the results inFigure 7.18, and highlight the vertex points and the interior point as shown.

▼Extreme Vertices	•					
Design Extreme Vertices	•	p1	p2	р3	Y	Point Type
▼ Model		0.474	0.252	0.274	•	Vertex
▼ Ternary Plot		0.726	0	0.274	•	Vertex
		0.849	0	0.151	•	Vertex
Columns (5/0)	• 4	0.597	0.252	0.151	•	Vertex
🚄 p1 🗶 🔺 p2 🗶	• 5	0.6	0.126	0.274	•	Edge
2 ↓ 2 p3 ★	- 6	0.6615	0.126	0.2125	•	Interior
∠ Y ★	• 7	0.7875	0	0.2125	•	Edge
📕 Point Type 🚑	• 8	0.723	0.126	0.151	•	Edge
Rows	• 9	0.5355	0.252	0.2125	•	Edge
All rows 9						
Selected 5						

Figure 7.18 Identify Vertices and Center Point with Sample Data Script

- 9 Select Edit > Copy, to copy the selected rows to the clipboard.
- 10 Select Rows > Add Rows and enter 5 as the number of rows to add
- 11 Click the At End radio button on the dialog, then click OK.
- 12 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).

▼ plasticizer	•							Pred
Design Extreme Vertices	•	~	p1	p2	р3	Y	Point Type	Formula Y
▼ Model	1.1	1	0.474	0.252	0.274	12	Vertex	10.8923077
 Ternary Plot 	1.1	2	0.726	0	0.274	4	Vertex	5.09230769
Columna (00)		3	0.849	0	0.151	8	Vertex	7.49230769
Columns (6/0)		4	0.597	0.252	0.151	10	Vertex	11.2923077
▲ p1 米 ▲ p2 米		5	0.6	0.126	0.274	13	Edge	13.0307692
▲ p2 木 ▲ p3 ★		6	0.6615	0.126	0.2125	18	Interior	19.2692308
∡ Y *		7	0.7875	0	0.2125	12	Edge	11.8307692
📕 Point Type 🔎		8	0.723	0.126	0.151	14	Edge	14.4307692
🚄 Pred Formula Y 🐥		9	0.5355	0.252	0.2125	16	Edge	16.6307692
_		10	0.474	0.252	0.274	10	Vertex	10.8923077
Rows		11	0.726	0	0.274	6	Vertex	5.09230769
All rows 14 Selected 0		12	0.849	0	0.151	7	Vertex	7.49230769
Excluded 0		13	0.597	0.252	0.151	10	Vertex	11.2923077
Hidden 0		14	0.6615	0.126	0.2125	21	Interior	19.2692308
_abelled 0								

Figure 7.19 Plasticizer.jmp Data Table from the Sample Data Library

Analyze the Mixture Model

Use the Cornell plasticizer data from the Sample Data library (Figure 7.19) to run the mixture model:

- 1 Right-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 Model 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
```

Your values may be different from those shown here.

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).

158 Mixture Designs

A Chemical Mixture Example

Figure 7.20 Mixture Response Surface Analysis

🕈 💌 Response Y	
Actual by Predicted Plot	
Summary of Fit	
Analysis of Variance	
Lack Of Fit	
Parameter Estimates	
Effect Tests	
Effect Details	
Coef	
p1(Mixture) p2(Mixture) p3(Mixture)	Y
p1(Mixture) 0 317.36341 1464.3298	-50.14653
p2(Mixture) . 0 1846.2177	-282.1982
p3(Mixture) 0	-911.6484
▼ Solution	
Variable Critical Value	
p1(Mixture) 0.635052	
p2(Mixture) 0.1556806	
p3(Mixture) 0.2092675	
Solution is a Maximum	
Assuming the following mixture sum: p1+p2+p3=1	
Predicted Value at Solution 19.570299	
Eigenvalues and Eigenvectors	
Eigenvalue -146.550 -3164.00	
p1(Mixture) 0.75052 0.66085	
n2(Mixture) -0.66085 0.75052	

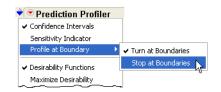
The Prediction Profiler

The report contains a prediction profiler.

3 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.



- 4 To limit the visible profile curves to bounds that use all three variables, use the **Stop at Boundaries** command from the menu on the Prediction Profiler title bar.
- 5 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.
- 6 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

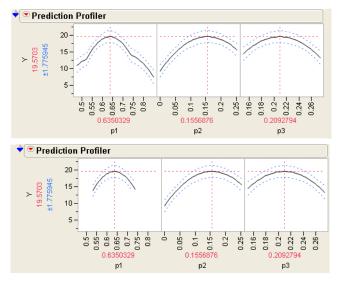
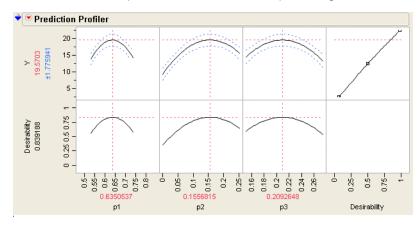


Figure 7.22 Maximum Desirability in Profiler for Mixture Analysis Example



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.
- 7 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.

Current X

0.6615

0.126

0.2125

Current Y

12.5 19.269231

0.1

Contour

p1

0.9

Lo Limit

0.474

0.151

0

Lo Limit

Hi Limit 0.849

0.252

0.274

Hi Limit

Figure 7.23 Mixture Profiler for Plasticizer Example

■ Response Y
■ Mixture Profiler

0 🔿 🔿 p1

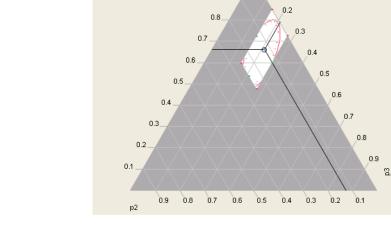
О́О́О́р3

Response

R Factor

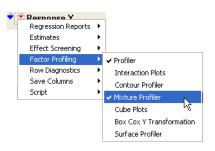
ΤL

🔿 💽 🔿 p2



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:



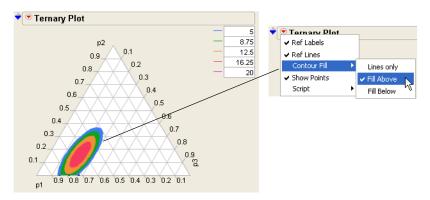
- 1 Choose Graph > Ternary Plot.
- 2 Specify plot variables (p1, p2, and p3) and click X, Plottting, as shown in Figure 7.24. To identify the contour variable (the prediction equation), select Pred Formula Y and click the Contour Formula button. The contour variable must have a prediction formula to 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

Plot how components of a mixture ad	ld up	
Select Columns	Cast Selected Columns into Roles	Action —
4p1	X, Plotting	ОК
p2 p3	p2	Cancel
Y	optional numeric	Concor
Leoint Type	Contour Formula Pred Formula Y	
Pred Formula Y		Remove
	By optional	Recall
	Column values must not be negative	Help
		<u> </u>

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, 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

r				-		H
Ŀ		_		_	1	н
U				_	1	u
L	-				. 11	н
ш		-	-	-	. 11	н
Ľ	-				14	۲

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
Create an Example Choice Experiment
Analyze the Example Choice Experiment
Design a Choice Experiment Using Prior Information
Administer the Survey and Analyze Results
Initial Choice Platform Analysis174
Find Unit Cost and Trade Off Costs with the Profiler

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.

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	_X_
	2	80 GB	1.5 GHz	4 hours	\$1,200	
2	1	40 GB	1.5 GHz	4 hours	\$1,500	
	2	80 GB	2.0 GHz	4 hours	\$1,200	_X_
3	1	40 GB	2.0 GHz	4 hours	\$1,200	_X_
	2	80 GB	2.0 GHz	6 hours	\$1,500	
4	1	40 GB	2.0 GHz	4 hours	\$1,000	_X_
	2	80 GB	1.5 GHz	6 hours	\$1,200	
5	1	40 GB	1.5 GHz	6 hours	\$1,000	_X_
	2	40 GB	2.0 GHz	4 hours	\$1,500	
6	1	40 GB	2.0 GHz	6 hours	\$1,200	_X_
	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	_X_
8	1	40 GB	1.5 GHz	4 hours	\$1,200	
	2	80 GB	2.0 GHz	4 hours	\$1,000	_X_

Table 8.1	Hypothetical	Choice Survey	Results from	a single Respor	ndent, Subject ID 2
-----------	--------------	---------------	--------------	-----------------	---------------------

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
--------------	--

	move Add N Factors 1			
Name	Role	Attribute Level	s	
 Disk Size 	Categorical	40 GB		80 GB
Speed	Categorical	1.5 GHz	1.5 GHz	
Battery Life	Categorical	4 Hrs	4 Hrs	
 Price 	Categorical	\$1,500	\$1,200	\$1,000
pecify Attributes d a Continuous or Cat an attribute name or le continue	egorical attribute by clicking its			

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	
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 Create an Example Choice Experiment

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 De	sign					
Attributes						
Model						
Design						
Choice Set	Disk Size	Speed	Battery Life	Price		
1	40 GB	1.5 GHz	6 Hrs	\$1,00		
1	80 GB	1.5 GHz	4 Hrs	\$1,20		
2	40 GB	1.5 GHz	4 Hrs	\$1,50		
2	80 GB	2.0 GHz	4 Hrs	\$1,20		
3	40 GB	2.0 GHz	4 Hrs	\$1,20		
3	80 GB	2.0 GHz	6 Hrs	\$1,50		
4	40 GB	2.0 GHz	4 Hrs	\$1,00		
4	80 GB	1.5 GHz	6 Hrs	\$1,20		
5	40 GB	1.5 GHz	6 Hrs	\$1,00		
5	40 GB	2.0 GHz	4 Hrs	\$1,50		
6	40 GB	2.0 GHz	6 Hrs	\$1,20		
6	80 GB	1.5 GHz	4 Hrs	\$1,50		
7	40 GB	2.0 GHz	6 Hrs	\$1,50		
7	80 GB	1.5 GHz	4 Hrs	\$1,00		
8	40 GB	1.5 GHz	4 Hrs	\$1,20		
8	80 GB	2.0 GHz	4 Hrs	\$1,00		
 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.

Choice Profiles		◆ _ ●	Choice	Response				
Design (Discrete Choice	•	Set	Indicator	Disk Size	Speed	Battery Life	Price
💌 Choice		1	1	•	40 GB	1.5 GHz	6 Hrs	\$1,000
		2	1	-	80 GB	1.5 GHz	4 Hrs	\$1,200
		3	2	•	40 GB	1.5 GHz	4 Hrs	\$1,500
		4	2	•	80 GB	2.0 GHz	4 Hrs	\$1,200
		5	3	•	40 GB	2.0 GHz	4 Hrs	\$1,200
Columns (6/0) Choice Set Response Indicator		6	3	•	80 GB	2.0 GHz	6 Hrs	\$1,500
		7	4	•	40 GB	2.0 GHz	4 Hrs	\$1,000
Disk Size	- ·		4	•	80 GB	1.5 GHz	6 Hrs	\$1,200
L Speed		9	5	•	40 GB	1.5 GHz	6 Hrs	\$1,000
🔥 Battery Lif	e	10	5	•	40 GB	2.0 GHz	4 Hrs	\$1,500
🔥 Price		11	6	•	40 GB	2.0 GHz	6 Hrs	\$1,200
		12	6	•	80 GB	1.5 GHz	4 Hrs	\$1,500
		13	7	•	40 GB	2.0 GHz	6 Hrs	\$1,500
Rows		14	7	•	80 GB	1.5 GHz	4 Hrs	\$1,000
All rows	16	15	8	•	40 GB	1.5 GHz	4 Hrs	\$1,200
Selected Excluded	0 0	16	8	•	80 GB	2.0 GHz	4 Hrs	\$1,000
Exclusion	0							

Figure 8.4	JMP Data	Table for	Preliminary	Laptop	Choice Survey
0	J				

Analyze the Example Choice Experiment

Once a survey deign 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**).

Note: When you create a survey data table, the default name of the table is Choice Profiles. A Choice analysis is saved with the table as a script that references the default data table name. This example table is saved as Laptop Design.jmp in the Design of Experiment folder of the Sample data. Note in Figure 8.5 that the script was modified to reference the table named Laptop Design.

170 Discrete Choice Designs

Analyze the Example Choice Experiment

Price \$1,000
61,000
61,200
61,500
61,200
\$1,200
61,500
\$1,000
61,200
61,000
61,500
61,200
61,500
61,500
61,000
61,200
61,000
7
1

Figure 8.5	JMP Table with Survey Choice Sets and Responses
------------	---

4 Click the Choice script in the Laptop Design data table and select **Run** to analyze that data with the Choice modelling 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 Chapter 30, "Choice Modeling," in the JMP Statistics and Graphics User Guide.

Figure 8.6 Choice Model Fitting Dialog

Choice Dialog		
◆ Profile Data		Run Model
Select Data Table Laptor	o Design	
Select Columns	Pick Role Variables	Help
Choice Set	Profile ID	Remove
Response Indicator Disk Size Speed Battery Life Price	Grouping Choice Set	✓ Firth Bias-adjusted Estimates
	Speed Cross Battery Life Price Macros V Degree 2 Transform V	

- 5 Click Run Model on the Choice model fitting dialog.
- 6 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:

- 7 Right click on the Parameter Estimates report and choose **Make into Data Table** from the menu, as shown.
- 8 In the new Untitled data table, create a new column and call it Var.
- 9 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.
- 10 In the Formula Editor, click the Std Error column in the column list and click the exponent button ((x)) 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

Parameter Estima	ates						Std	
Term				•	Term	Estimate	Error	Var
	Estimate	Std Error		1	Disk Size[40 GB]	0.07842	0.4799/	0.2303
Disk Size[40 GB] 0	0.07842569	0.4799842066	Columns (4/1)	2	Speed[1.5 GHz]	-0.23007	0.5309	0.2819
	0.23007537	0.5309583378	L Term	3	Battery Life[4 Hrs]	-0.02064	0.528/2	0.2790
	0.02064181	0.5282175127	A Estimate	4	Price[\$1,000]	1.02768	0.76/33	0.5826
	1.02768703 0.00408286 43.502	0.7633468524 0.6836310541 Table Style	✓ Std Error ✓ Var ♣	5	Price[\$1,200]	0.00408	0.6836	0.4673
-2*LogLikelihood -2*Firth LogLikelihood Converged in Gradient Firth Bias-adjusted Estir	3.5023 -2.067	Columns Sort by Column Make into Data Table Make Combined Data Table		I		Std E	rrof ²	

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.

I72 Discrete Choice Designs

Design a Choice Experiment Using Prior Information

- 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

				•	Choice Desi	gn							
cop	y and pa	aste to			Attributes								
ente	enter information				Name	R	ole		Attribute	e Levels			
					✓ Disk Size	С	ategorical		40 GB		80 GB		
					✓ Speed		ategorical		1.5 GHz			2.0 G	Hz
					➤ Battery Life		ategorical		4 Hrs			6 Hrs	1
		Std			✓ Price	С	ategorical		\$1,000		\$1,200		\$1,500
	Estimate	Error	Var	+	Model								
	0.0784	0.4800	0.2304		DOE Mode	l Controls							
	-0.2301	0.5310	0.2819		Prior Spec	ification							
	-0.0206	0.5282	0.2790			Ignore prior specifications. Generate the Utility Neutral design.							
	1.0277	0.7633	0.5827		Prior Me		٦.			.g			
	0.0041	0.6836	0.4674		Effect Prior Mean								
					Disk Size	0.07							
		\			Speed	-0.2	-						
		1			Battery Life	-0.2	-						
Ente	er mean		/		Price 1	1.02	-						
			/		Price 2	0.00							
estii	nates		/		🔲 Ignore pric	or variance. C	∋enerate t	he local design	for the pri	or mean.			
			/		🕈 Prior Va	riance Ma	trix						
					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			
Ent	er varia	nce			Price 1				0.583	0.000			
					Price 2					0.467			
esti	mates			•	Design Gene	ration							

- 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).

- Choice Design
 Attributes
 Model
 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
- Figure 8.9 Design Specifications for final Laptop Survey

5 Click Make Design to see a design similar to the one shown in Figure 8.10.

Figure 8.10 Design Runs for Two Choice Surveys

esi	yn				
Surve	y Choice Se			Battery Life	Price
	1		1.5 GHz		\$1,200
	1		2.0 GHz		\$1,000
			2.0 GHz		\$1,200
			1.5 GHz		\$1,000
			2.0 GHz		\$1,500
			1.5 GHz		\$1,200
			2.0 GHz		\$1,500
			1.5 GHz		\$1,000
			1.5 GHz		\$1,000
			2.0 GHz		\$1,200
			2.0 GHz 1.5 GHz		\$1,200
			2.0 GHz		\$1,500
			1.5 GHz		\$1,200
			1.5 GHz		\$1,000 \$1,500
			2.0 GHz		
		0 0000	2.0 GHZ	onrs	\$1,200

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 is 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. As shown previously, the default script accesses the default data table created by the Choice designer named Choice Profiles. If necessary, change the data table name in the script to the name you used for the final design table. Note in Figure 8.11, the data table name is changed to Laptop Results.

1 To continue with this example, open the table called Laptop Results, saved in the Design Experi-

I74Discrete Choice Designs

Administer the Survey and Analyze Results

ment folder of the Sample Data.

Figure 8.11	Partial Listing of Final Survey with Results
-------------	--

Choice Response Batterv 포 Laptop Results Respondent Survey Disk Size Speed Price Design Discrete Choic Set Indicator Life Choice 22 2 2 11 0 80 GB 1.5 GHz \$1,000 4 Hrs 圮 Run Script 23 2 2 12 0 40 GB 1.5 GHz 4 Hrs \$1,200 Edit 24 2 2 12 1 80 GB 2.0 GHz 6 Hrs \$1,000 Delete 25 2 2 13 1 40 GB 2.0 GHz 6 Hrs \$1,200 26 2 2 13 0 80 GB 1.5 GHz 4 Hrs \$1,500 2 💌 Columns (8/0) 27 2 14 0 40 GB 1.5 GHz 6 Hrs \$1,200 🔥 Respondent 28 2 2 14 1 80 GB 2.0 GHz 4 Hrs \$1,500 🔥 Survey 29 2 2 40 GB 2.0 GHz 4 Hrs 15 1 \$1,200 🆺 Choice Set 30 2 2 15 0 80 GB 1.5 GHz 6 Hrs \$1,500 📕 Response Indicator 31 2 2 1 40 GB 2.0 GHz 4 Hrs 16 \$1,000 🆺 Disk Sjize , Speed 32 2 2 16 0 80 GB 1.5 GHz 6 Hrs \$1,200 🆺 Battery Life 33 3 1 1 80 GB 1.5 GHz 4 Hrs \$1,200 1 📕 Price 34 3 1 1 0 40 GB 2.0 GHz 6 Hrs \$1,000 35 3 1 2 0 40 GB 2.0 GHz 4 Hrs \$1,200 1 36 3 2 1 80 GB 1.5 GHz 6 Hrs \$1,000 Choice(Profile DataTable("Laptop Results"), Profile Grouping(:Respondent, :Survey, :Choice Set), Profile ID(:Response Indicator), Profile Effects(:Disk Size, :Speed, :Battery Life, :Price Launch Dialog

Initial Choice Platform Analysis

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 8.12.

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.

٠	Choice Dialog		
*	Profile Data	p Results Pick Role Variables Profile ID Grouping th, Respondent th, Survey th, Choice Set Construct Model Effects Add Disk Size	Run Model Help Remove V Firth Bias-adjusted Estimates
		Add Disk Size Cross Battery Life Nest Price	

Figure 8.12 Choice Model Fitting Dialog to Analyze the Laptop Survey

3 Click Run Model 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 8.13.

The results are clear. The most significant attribute is Speed (Prob < .0001). Disk Size and Price are also significant, but Battery Life is not.

Figure 8.13 Initial Analysis of the Final Laptop Survey

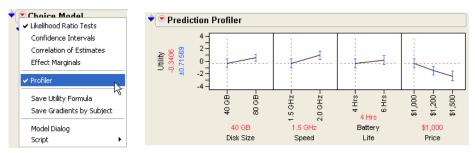
🕈 💌 Choice M	lodel					
🕈 Paramete	er Estimates					
Term	Est	imate	Std Error			
Disk Size[40	GB] -0.454	65897	0.1566686541			
Speed[1.5 GH	iz] -0.664	62895	0.1726760880			
Battery Life[4	Hrs] -0.259	66701	0.1559370832			
Price[\$1,000]	1.038	35088	0.3103813337			
Price[\$1,200]	-0.086	93113	0.2103135075			
AICc	77.	855675	5			
-2*LoqLikelih	ood 67	044864	1			
-2*Firth LogL	ikelihood 50.	3				
Converged in	Gradient					
Firth Bias-adjusted Estimates						
Effect Lik	elihood Rat		sts			
	L-R					
Source	ChiSquare	DF	Prob>ChiSq			
Disk Size	10.045	1	0.0015*			
Speed	20.904	1	≺.0001*			
Battery Life	3.036	1	0.0814			
Price	15.154	2	0.0005*			

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.

4 Select **Profiler** from the menu on the Choice Model title bar to see the Prediction Profiler in Figure 8.14.

Figure 8.14 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.

5 To find the unit utility cost, move the trace for price to and note how the Utility value changes.

Compare the Utility values in Figure 8.14 and Figure 8.15. 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.

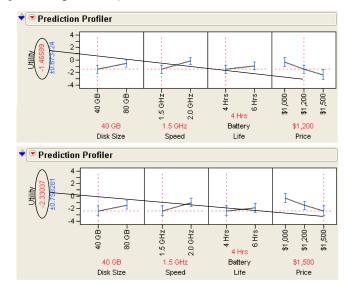


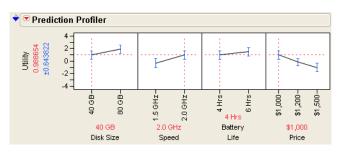
Figure 8.15 Compare Change in Utility Over Price

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 8.13).

6 In the Prediction Profiler, set Price to its lowest value and change Speed to its higher value.

You can see in Figure 8.16 that the Utility value changes from the original value shown in Figure 8.14 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*\$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 8.16 Change Speed in Profiler and Note Utility Value



This simple Choice survey and its analysis shows how this kind if information can be used to help manufacturers and retailers identify important product attributes and assign values to them. 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



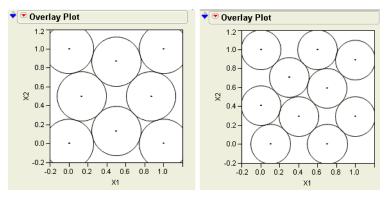
Space-filling designs are useful in situations where run-to-run variability is of far less concern than the form of the model. Sensitivity studies of computer simulations is one such situation. For this case, and any mechanistic or deterministic modeling problem, any variability is small enough to be 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:

- The first objective is to prevent replicate points by spreading the design points out to the maximum distance possible between any two points.
- The second objective is to 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.



Contents

Introduction to Space-Filling Designs 181
Sphere-Packing Designs
Creating a Sphere-Packing Design
Visualizing the Sphere-Packing Design183
Latin Hypercube Designs184
Creating a Latin Hypercube Design184
Visualizing the Latin Hypercube Design 185
Uniform Designs
Comparing Sphere-Packing, Latin Hypercube, and Uniform Methods
Minimum Potential Designs
Maximum Entropy Designs
Gaussian Process IMSE Optimal Designs192
Borehole Model: A Sphere-Packing Example192
Create the Sphere-Packing Design for the Borehole Data
Guidelines for the Analysis of Deterministic Data
Results of the Borehole Experiment 195

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

To use the Sphere-Packing method:

- 1 Select DOE > Space Filling Design.
- 2 Enter responses and factors. (See "Enter Responses and Factors into the Custom Designer," p. 59.)
- 3 Alter the factor level values, if necessary. For example, Figure 9.1 shows the two existing factors, X1 and X2, with values that range from 0 to 1 (instead of the default -1 to 1).

182 Space-Filling Designs

Sphere-Packing Designs

4 Click **Continue** in the dialog.

Figure 9.1 Space-Filling Dialog for Two Factors

Response Name				
	Goal	Lower Limit	Upper Limit	Importance
Y	Maximize			
Factors				
Add 1 Contin	iuous	Chai	nge Values	to 0 and 1
Remove Selected			/	
Name	Role	Values		
🚄 X1	Continuous	0	1	
⊿ X2	Continuous	0	1	
Specify Factors				

5 In the design specification dialog, specify a sample size (Number of Runs). Figure 9.2 shows a sample size of eight.

Figure 9.2 Space-Filling Design Dialog

_	- Space Filling Design Method	e
1	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	

6 Click Sphere Packing.

JMP creates the design and displays the design runs and the design diagnostics. Figure 9.3 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 9.3 Sphere-Packing Design Settings and Design Diagnostics

			Minimum	Nearest
-			7	
Run	ScaledX1	ScaledX2	Distance	Point
1	1.00000	1.00000	0.518	3
2	0.13398	0,50000	0.518	7
3	0.50000	0.86603	0.518	5
4	1.00000	0.00000	0.518	6
5	0.86602	0.50000	0.518	3
6	8.50000	0.13397	0.518	4
7/	0.00000	0.00000	0.518	6
/8	0.00000	1.00000	0.518	3
discrepa	ncy = 0.0469	9		
Design	Table ——			
Make Tal	e Back			

7 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 9.3 as the radius for the circles. Using the table you just created:

- 1 Select Graph > Overlay Plot.
- 2 Specify X1 as X and X2 as Y, then click OK.
- 3 Adjust the frame size so that the frame is square by right-clicking the plot and selecting Size/Scale > Size to Isometric.
- 4 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:

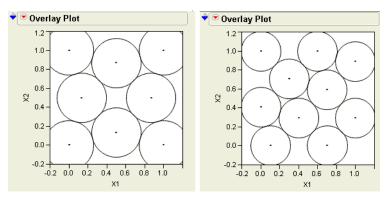
For Each Row(Circle({:X1, :X2}, 0.518/2))

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 9.4. This plot shows the efficient way JMP packs the design points.

5 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 9.4. 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 9.4 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. 59.)
- 3 Alter the factor level values, if necessary. For example, Figure 9.5 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 9.5 Space-Filling Dialog for Four Factors

C	emove N Responses	Lower Limit	l han av Limit	Importance
Response Name	Goal Maximize		Upper Limit	Importance
Factors		· ·		
Add 2 Continuou	IS			
Remove Selected				
Name AX1	Role Continuous	Values	8	
▲ X1 ▲ X2	Continuous	1	8	
A2 AX3	Continuous	1	8	
⊿ X4	Continuous	1	8	
ace Filling Design				

- 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 9.2). Factor settings and design diagnostics results appear similar to those in Figure 9.6, 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.

4 Factors						
	g Latin Hype					
Facto	r Settings	\$				
Run	X1	X2	X3	X4		
1	7.00000	4.00000	2.00000	8.00000		
2	2.00000	6.00000	1.00000	5.00000		
3	5.00000	2.00000	3.00000	1.00000		
4	6.00000	8.00000	4.00000	3.00000		
5	4.00000	7.00000	7.00000	7.00000		
6	1.00000	5.00000	6.00000	2.00000		
7	8.00000	3.00000	8.00000	4.00000		
8	3.00000	1.00000	5.00000	6.00000		
🕈 Desig	n Diagno:	stics				
♦ Desig	n Diagno	stics			Minimum	Nearest
♥ Desig Run	n Diagno: ScaledX1	stics ScaledX2	ScaledX3	ScaledX4	Minimum Distance	Nearest Point
	Ŭ		ScaledX3 0.14286	ScaledX4 1.00000		
Run	ScaledX1	ScaledX2			Distance	Point
Run 1	ScaledX1 0.85714	ScaledX2 0.42857	0.14286	1.00000	Distance 0.881	Point 8
Run 1 2	ScaledX1 0.85714 0.14286	ScaledX2 0.42857 0.71429	0.14286 0.00000	1.00000 0.57143	Distance 0.881 0.821	Point 8 4
Run 1 2 3	ScaledX1 0.85714 0.14286 0.57143	ScaledX2 0.42857 0.71429 0.14286	0.14286 0.00000 0.28571	1.00000 0.57143 0.00000	Distance 0.881 0.821 0.833	Point 8 4 8
Run 1 2 3 4	ScaledX1 0.85714 0.14286 0.57143 0.71429	ScaledX2 0.42857 0.71429 0.14286 1.00000	0.14286 0.00000 0.28571 0.42857	1.00000 0.57143 0.00000 0.28571	Distance 0.881 0.821 0.833 0.782	Point 8 4 8 5
Run 1 2 3 4 5	ScaledX1 0.85714 0.14286 0.57143 0.71429 0.42857	ScaledX2 0.42857 0.71429 0.14286 1.00000 0.85714	0.14286 0.00000 0.28571 0.42857 0.85714	1.00000 0.57143 0.00000 0.28571 0.85714	Distance 0.881 0.821 0.833 0.782 0.782	Point 8 4 8 5 4 3 5
Run 1 2 3 4 5 6	ScaledX1 0.85714 0.14286 0.57143 0.71429 0.42857 0.00000	ScaledX2 0.42857 0.71429 0.14286 1.00000 0.85714 0.57143	0.14286 0.00000 0.28571 0.42857 0.85714 0.71429	1.00000 0.57143 0.00000 0.28571 0.85714 0.14286	Distance 0.881 0.821 0.833 0.782 0.782 0.782 0.845	Point 8 4 8 5 4 3

Figure 9.6 Latin Hypercube Design for Four Factors and Eight Runs with Eight Levels

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 9.7.

Figure 9.7 Latin Hypercube Design with two Factors and Eight Runs

Space	Filling	Latin Hype	rcube							
⇒ Fa	actor	Settings	;	+	Desig	jn Diagno	ostics			
F	Run	X1	X2		Run	ScaledX1	ScaledX2	Minimum Distance	Nearest Point	
	1	1.00000	0.85714		1	1.00000	0.85714	0.452	2	
	2	0.57143	0.71429		2	0.57143	0.71429	0.404	6	
	3	0.42857	0.28571		3	0.42857	0.28571	0.404	4	
	4	0.14286	0.57143		4	0.14286	0.57143	0.404	3	
	5	0.00000	0.14286		5	0.00000	0.14286	0.452	3	
	6	0.28571	1.00000		6	0.28571	1.00000	0.404	2	The minimum
	7	0.85714	0.42857		7	0.85714	0.42857	0.404	>2	
	8	0.71429	0.00000		8	0.71429	0.00000	0.404	3	distance
					discrepa	incy = 0.009	92			

- 6 Click Make Table.
- 7 Select Graph > Overlay Plot.
- 8 Specify X1 as X and X2 as Y, then click OK.
- 9 Right-click the plot and select **Size/Scale > Size to Isometric** to adjust the frame size so that the frame is square.

10 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 9.7). 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 9.8. This plot shows the efficient way JMP packs the design points.

11 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 9.8. 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.

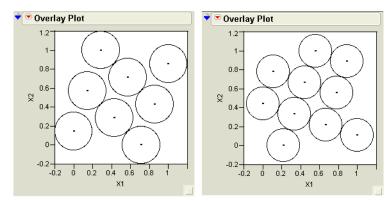


Figure 9.8 Comparison of Latin Hypercube Designs with Eight Runs (left) and 10 Runs (right)

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. 59.)
- 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 9.9.

Note: The emphasis of the Uniform design method is not to spread out the points. The minimum distances in Figure 9.9 vary substantially.

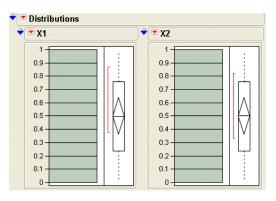
Figure 9.9 Factor Settings and Diagnostics for Uniform Space-Filling Designs with Eight Runs

•	Factor	Settings	;	-	Desig	ın Diagno	stics		
	Run	X1	X2		Run	ScaledX1	ScaledX2	Minimum Distance	Nearest Point
	1	0.43532	0.93541		1	0.43532	0.93541	0.284	3
	2	0.56091	0.31491		2	0.56091	0.31491	0.352	6
	3	0.69276	0.81486		3	0.69276	0.81486	0.277	7
	4	0.18514	0.69286		4	0.18514	0.69286	0.285	5
	5	0.06460	0.43504		5	0.06460	0.43504	0.285	4
	6	0.30959	0.06908		6	0.30959	0.06908	0.352	2
	7	0.81410	0.56557		7	0.81410	0.56557	0.277	3
	8	0.92607	0.19080		8	0.92607	0.19080	0.386	2
					discrepa	ncy = 0.004	6		

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 9.10 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 mini-

mum distance from each point to its closest neighbor. The discrepancy value is the integrated difference between the design points and the uniform distribution.

Figure 9.11 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 Hyper-cube 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.

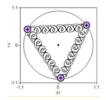
Figure 9.11 Comparison of Diagnostics for Three Eight-Run Space-Filling Methods

Design Diagnostics - Sphere Packing		
Run ScaledX1 ScaledX2 Minimum Distance	Nearest Point	
1 0.00000 1.00000 0.518		
2 0.50000 0.86603 0.518	2	
3 0.86603 0.50000 0.518	4	
4 1.00000 1.00000 0.518	4	Subara
5 1.00000 0.00000 0.518	6	Sphere-
6 0.50000 0.13397 0.518	5	Packing
7 0.00000 0.00000 0.518	8	0
8 0.13397 0.50000 0.518	7	
0 0.0000 0.0000 0.000		
discrepancy = 0.0469		
Design Diagnostics - Latin Hypercube		
Run ScaledX1 ScaledX2 Minimum Distance	Nearest Point	
1 0.57143 0.71429 0.404	6	
2 0.85714 0.42857 0.404	1	
3 1.00000 0.85714 0.452	1	т.
4 0.00000 0.14286 0.452	5	Latin
5 0.42857 0.28571 0.404	7	Hypercube
6 0.28571 1.00000 0.404	1	rijpereuse
7 0.14286 0.57143 0.404	5	
8 0.71429 0.00000 0.404	5	
discrepancy = 0.0092		
Design Diagnostics - Uniform		
	Nearest Point	
1 0.43909 0.68509 0.276	6	
2 0.69031 0.93093 0.351	1	
3 0.07393 0.19080 0.346	6	TT .C
4 0.81410 0.56585 0.282	7	Uniform
5 0.56496 0.06460 0.44	7	
6 0.31491 0.43909 0.276	1	
7 0.93092 0.30959 0.282	4	
8 0.18696 0.81304 0.283	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.



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 Enter responses and factors. (See "Enter Responses and Factors into the Custom Designer," p. 59.)
- 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 9.12), 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 9.12) and the design diagnostics.

Figure 9.12 Space-Filling Methods and Design Diagnostics for Minimum Potential Design

mber of Runs: 12						Minimum	Neares
Sphere Packing	Optimal Packing of Spheres	Run	ScaledX1	ScaledX2	ScaledX3	Distance	Poir
sprioro r deferig	Inside of a Cube.	1	0.97320	0.62495	0.60234	0.526	
Latin Hypercube	Latin HyperCube with Optimal Spacing	2	0.61675	0.98461	0.46100	0.526	
Eddiningporcabo	East Hyperodue wat oparital opacing	3	0.72434	0.24477	0.86677	0.526	
Uniform	Uniform Design	4	0.02680	0.37506	0.39766	0.526	
		5	0.21410	0.36946	0.88887	0.526	1
Minimum Potential	Minimum energy designs in a spherical region.	6	0.57870	0.74619	0.92801	0.526	
Maximum Entropy	Maximum entropy designs for a Gaussian process	7	0.85240	0.17328	0.36192	0.526	
Haxinan chaopy	waxing in encopy designs for a Gaussian process	8	0.14760	0.82672	0.63809	0.526	
aussian Process IMSE Optimal	Integrated mean square error optimal designs for a Gaussian	9	0.42130	0.25381	0.07199	0.526	
	process	10	0.27566	0.75524	0.13322	0.526	1
Back		11	0.78590	0.63053	0.11113	0.526	1
		12	0.38325	0.01539	0.53900	0.526	
		discrep	ancy = 0.011	1			
		Design	Table				

7 Click Make Table.

You can see the spherical symmetry of the Minimum Potential design using the Scatterplot 3D graphics platform.

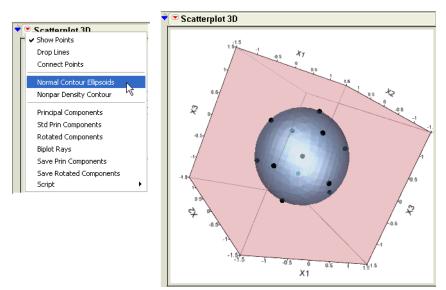
1 After you make the JMP design table, choose the Graph > Scatterplot 3D command.

Maximum Entropy Designs

- 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 9.13, double-click on the axes to scale them, and use the **Normal Contour Ellipsoids** option from the menu in the Scatterplot 3D title bar.

Now it is easy to see the points spread evenly on the surface of the ellipsoid.

Figure 9.13 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 its projections onto the coordinate axes are uniform.

However, as the example at the top in Figure 9.14 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. Note that this design appears to leave a hole in the bottom right of the overlay plot.

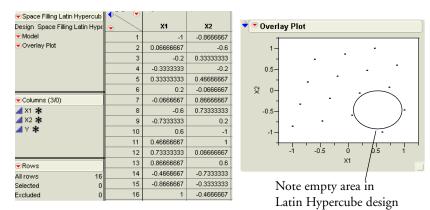
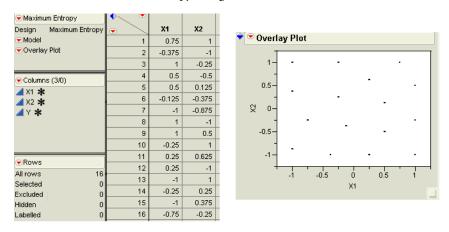


Figure 9.14 Two-factor Latin Hypercube Design

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. The two-factor Maximum Entropy design shown in Figure 9.15 covers the region better than the Latin hypercube design in Figure 9.14. The space-filling property improves as the number of runs increases without bound.

Figure 9.15 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

$$R_{ij} = \exp\left(-\sum_{k} \theta_k (x_{ik} - x_{jk})^2\right)$$

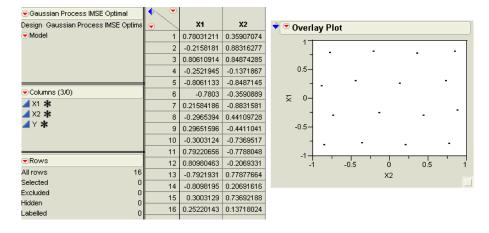
is the correlation of response values at two different design points, x_i and x_j . Computationally, these designs maximize $|\mathbf{R}|$, the determinant of the correlation matrix of the sample. When x_i and x_j are far apart, then \mathbf{R}_{ij} approaches zero. When x_i and x_j are close together, then \mathbf{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 9.14). The table and overlay plot in Figure 9.16 show a Gaussian IMSE optimal design. You can see that the design provides uniform coverage of the factor region.

Figure 9.16 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 m³/year and is determined by the 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_{μ} = transmissivity of upper aquifier, 63,070 to 115,600 m²/year

 H_u = potentiometric head of upper aquifier, 990 to 1100 m

 T_l = transmissivity of lower aquifier, 63.1 to 116 m²/year

 H_l = potentiometric head of lower aquifier, 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 DOE folder, open Borehole Factors.jmp from the Design Experiment folder to load the factors (Figure 9.17).

Figure 9.17 Factors Panel with Factor Values Loaded for Borehole Example

Factors				
Add 1 Continu	ous			
Remove Selected				
Name	Role	Values		
🚄 log10 Rw	Continuous	-1.3	-0.82	
🚄 log10 R	Continuous	2	4.7	
🚄 Tu	Continuous	63070	115600	
🚄 TI	Continuous	63.1	116	
🚄 Hu	Continuous	990	1110	
⊿ HI	Continuous	700	820	
⊿ L	Continuous	1120	1680	
⊿ Kw	Continuous	9855	12045	

Note: The logarithm of r and r_w are used in the following discussion.

- 4 Click Continue.
- 5 Specify a sample size of 32 runs, as shown in Figure 9.18.

Figure 9.18 Space-Filling Design Method Panel Showing 32 Runs

Space Filling Design 2 Factors	
Space Filling Design Methods	
Choose Sample Size 32	
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	

- 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 (Figure 9.19) 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.

Baushala Cubaus Baskiau	♦ ∖										
Borehole Sphere Packing	· \		1	T	T 1				Rec.		
Design Space Filling Sphere	• \	log10 Rw	log10 R	Tu	TI	Hu	Н	L	Kw	Y	
Notes Morris, Max D., Mitch	1	-0.82	2.08350495	100093.363	63.1	990	702.135866	1120	10435.9957	190.882715	
Model A simple demonstrati	2	-0.82	4.29978574	65770.4926	63.3546815	1109.98831	756.151977	1120.86273	9855	219.778636	
♥ Fit Model	3	-0.8208292	2	63070	78.7453851	990	712.582595	1680	12045	141.643972	Γ
	4	-1.3	2.5218531	63070	63.1	994.937857	700	1671.2256	9855.06927	13.7105889	Γ
Columns (15/0)	5	-1.2818251	2.75621225	63070	63.1003347	1047.52096	820	1614.19327	12044.7589	14.5449173	
/ log10 Rw 🗶	6	-1.3	2	115373.914	115.939627	1065.21039	820	1680	11674.0186	13.4385116	
🚄 log10 R ¥ 🚄 Tu ≭	7	-1.2526827	4.7	98480.1565	116	990	715.81919	1679.96528	9855	15.7654872	:
▲ TI ★	8	-1.2751193	4.7	63135.7733	110.110022	1110	820	1120	12045	27.5470676	
📕 Hu 🗱	9	-1.3	2	63070	116	1110	705.262054	1588.51039	10713.7698	21.5295326	
H *	10	-1.3	4.61421766	115600	63.1	1014.12022	820	1468.66543	10475.1989	10.9048754	
🖌 L 🛠	11	-0.82	4.7	115600	115.338111	1110	762.574254	1155.59827	12045	257.234604	
🚄 Kw 🗶	12	-0.82	4.7	65797.2718	112.983532	1078.19028	700	1639.76411	11183.6565	184.016172	
▲ Y 告 *	13	-0.8956276	4.7	113453.352	116	990	820	1680	12045	61.5248634	
🚄 log y 🛟 🚄 R 🐥	14	-0.82	4.68538198	96624.5656	63.1	1110	813.067911	1676.07222	11954.9597	149.964109	
⊿ Rw ∯	15	-0.82	2.0480977	63070	114.731053	1110	820	1677.27628	11861.2706	146.90971	
∠ true model 🖶	16	-1.2262786	3.57799341	113944.663	116	1098.45537	820	1120	9855	27.1109034	
🖌 prediction formula 🕂	17	-1.3	4.6516991	115600	116	1110	700.000759	1680	12045	23.172246	
🚄 prediction bias 🕂	18	-0.8371702	4.7	115600	64.555923	991.55747	700	1680	11141.5642	126.820106	
	40	0.00	4.00000204	70000.0404	CD 4	000 744 407	000 000740	44.00	40046	4 07 6004	

Figure 9.19 Borehole Sample Data

Guidelines for the Analysis of Deterministic Data

It is important to remember that deterministic data has 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 9.20, 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 9.20 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.

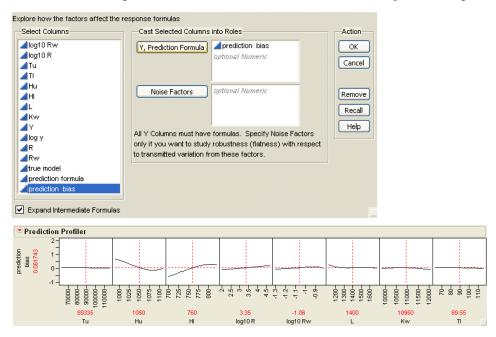


Figure 9.20 Profiler Dialog and Profile of the Prediction Bias in the Borehole Sphere-Packing Data

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 9.21. Note that the maximum bias is 1.826 and the minimum is -0.684 (the range is 2.51).

Figure 9.21 Distribution of the Prediction Bias

🕈 💌 Distributions) 🕈 🔍 Quan	tiles	
♥ prediction bias			maximum	1.826
2-		99.5%		1.826
21	.	97.5%		1.826
	1 1	90.0%		0.459
1.5-	1 1	75.0%	quartile	0.077
	1 1	50.0%	median	-0.029
1 -		25.0%	quartile	-0.146
n	1 1	10.0%		-0.451
0.5	:	2.5%		-0.684
		0.5%		-0.684
0	I 🛱 I	0.0%	minimum	-0.684
-0.5				
	:			
-1 -				

The top plot in Figure 9.22 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.

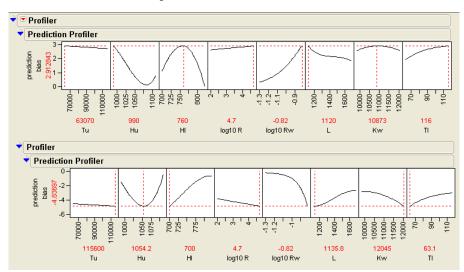


Figure 9.22 Prediction Plots showing Maximum and Minimum Bias Over Factor Domains

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

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, as shown below.

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 non-linear design tool.

DOE Menu

JMP Starter

	Click Category:	Experimental Design. Define factors and design a table of experimental runs.
DOE Analyze Graph Tools	File Basic	Custom Design Create a design tailored to meet specific requirements.
 ☑ Custom Design ☑ Screening Design 	Model Multivariate Reliability	Screening Design Sift through many factors to find the few that have the most effect.
 Response Surface Design Full Factorial Design 	Graph Surface	Response Surface Design Find the best response allowing quadratic effects (curvature).
	Measure Control	Choice Design Find the combination of attribute levels that your customers like the most Conjoint analysis.
Choice Design Space Filling Design	DOE Tables	Nonlinear Design Create an optimal design for models that are nonlinear in the parameters.
Xonlinear Design	SAS	Space Filling Design Designs for computer simulation modeling.
🗃 Augment Design		Full Factorial Design Generate all possible combinations of the specified factor settings.
Sample Size and Power		Taguchi Arrays Make inner and outer arrays from signal and noise factors.
		Mixture Design Optimize a recipe for a mixture of several ingredients.
		Add more runs to an existing data table. Replicate, add centerpoints, fold over or add model terms.
		Sample Size and Power Plot any two of the power to detect an effect, the sample size, and the effect size given the third. Or compute one given the other two.

Contents

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 10.1) contains a column (Model (x))whose values are formed by a formula with a poor guess of the parameter values.

Figure 10.1 Chemical Kinetics.jmp

 Chemical Kinetics 		•			
Notes Meyers, R.H.	. (1986	•	Velocity (y)	Concentration	Model (x)
		1	0.0773895	0.417	0.04
Columns (3/0)		2	0.0688714	0.417	0.04
Velocity (y)		3	0.0819351	0.417	0.04
Concentration		4	0.0737034	0.833	0.04
Model (x) 🕂		5	0.0738753	0.833	0.04
		6	0.0712396	0.833	0.04
Rows		7	0.065042	1.67	0.04
All rows	13	8	0.0547667	1.67	0.04
Selected	13	9	0.0497128	3.75	0.04
Excluded	õ	10	0.0642727	3.75	0.04
Hidden	0	11	0.0613005	6.25	0.04
Labelled	0	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 10.2). Note that the formula given by Model (X) shows in the launch dialog.

202 Nonlinear Designs

Examples of Nonlinear Designs

Figure 10.	2 Initial	Nonlinear	Analysis	Launch	Dialog
------------	-----------	-----------	----------	--------	--------

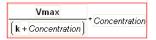
Fitting parameters in formula of Prec	ictor column to X column	
Select Columns	Cast Selected Columns into Roles	Action
Velocity (y)	Y, Response Velocity (y)	ОК
Model (x)	X, Predictor Formula	Cancel
	Group	
· · · · · ·	Weight optional Numeric	Remove
	Freq optional Numeric	Recall
Model Library	Loss optional Numeric	Help
	By optional	
	X Predictor column must have formula	
	Formulas	
	Predictor Parameter((Vmax = 0.04, k = 0), (Vmax / (k + :Concentration)) *:Concentration))	DN

- 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 10.3).

Figure 10.3 Nonlinear Fit Results

▼ 💌 Nonlinear Fit	
Response: Velocity (y), Predictor: Model (x)	
Control Panel	
Converged in Gradient	
Go Criterion Current Stop Limit Iteration 3 60 Stop Obj Change 1.6172054e-8 1e-15 Step Relative Gradient 2.064974e-6 0.000000 Gradient 1.8284546e-7 0.000001	The Save Estimates button adds the new fitted parameter values in the Model (x) column in the Chemical Kinetics.jmp data table.
Parameter Current Value Lock Vmax 0.0569558362 k -0.113473173	The Confidence Limits button computes confidence intervals used to create a nonlinear design.
Save Estimates Confidence Limits Convergence Criterion 0.00001 Goal SSE for CL 0.0011800361	These ranges are the intervals for Vmax and K. They are asymptotically normal. Use these limits to create a nonlinear
▼ Solution	
SSE DFE MSE RMSE 0.0008192453 11 7.4477e-5 0.00863	design in JMP.
Parameter Estimate ApproxStdErr Lower CL Upper CL Vmax 0.0569558362 0.00325866 0.0509514 0.06419848 k -0.113473173 0.02829458 -0.1671998 -0.0445499 Solved By: Analytic NR	\sum
Plot	

- 6 Click the **Confidence Limits** button to produce confidence intervals.
- 7 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:



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.

- 1 With the Chemical Kinetics.jmp data table open, select DOE > Nonlinear Design.
- 2 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 10.4 shows the completed dialog.

Figure 10.4 Initial Nonlinear Design Launch Dialog

Create an optimal design for models t	hat are nonlinear in the parameters	
Select Columns	Cast Selected Columns into Roles	Action
Velocity (y)	Y, Response Velocity (y)	ОК
Model (x)	X, Predictor Formula	Cancel
	Weight optional Numeric	
	X Predictor column must have formula	Remove
		Recall
		Help

3 Click **OK** to see the completed Design panels for factors and parameters, as shown in Figure 10.5. **Figure 10.5** Nonlinear Design Panels for Factors and Parameters

Factors			
Name	Role	Values	
Concentration	Continuous	0.417	6.25
optional item			
Parameters			
Name	Distribution	Values	
Vmax	Normal	0.0284779180988913	0.0854337542966738
k	Normal	-0.0567365862990143	-0.170209758897043
Design Generati	on		
umber of Runs:	15		

Note that in Chemical Kinetics.jmp (Figure 10.1), 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:

4 Change the factor range for Concentration to a broader interval—from 0.1 to 7 (Figure 10.6).

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.

5 Look back at the analysis report inFigure 10.3 and locate the upper and lower confidence limits for Vmax and k in the Solution table. Change the values for Vmax and k to correspond to those limits, as shown in Figure 10.6.

Now you have described the current level of uncertainty of the two parameters.

Figure 10.6 Change Values for Factor and Parameters

Factors				
Name	Role	Values		
Concentration	Continuous	0.1	7	
optional item				
Parameters				
Parameters Name	Distribution	Values		
	Distribution	Values 0.05	0.064	

6 If necessary, type the desired number of runs (15) into the text box.

Use commands from the menu on the Nonlinear Design title bar to get the best possible design:

- 7 Select Number of Starts from the menu on the title bar and enter 100 in the text box.
- 8 Select Advanced Options > Number of Monte Carlo Samples and enter 2 in the text box.
- 9 Click **Make Design** to preview the design (Figure 10.7). Your results may differ slightly than those shown for the additional runs.

Figure 10.7 Selecting the Number of Runs

	🕈 Desig	🕈 Design		
	Run	Concentration	Velocity (y	
	1	0.417	0.07739	
Generation	2	0.417	0.068871	
ber of Runs (counting 13 included runs): 15	3	0.417	0.081935	
2 Design	4	0.7	0.073703	
esign	5	0.7	0.073875	
	6	0.7	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	0.132626		
	15	0.132626		

10 Click Make Table.

This creates a new JMP design table (Figure 10.8) 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.

	Nonlinear Design	•			
	Design Nonlinear Design	•	Concentration	Velocity (y)	Model (x)
	▼ Model	1	0.417	0.0773895	0.0782487
-Output Options-		2	0.417	0.0688714	0.0782487
Make Table		3	0.417	0.0819351	0.0782487
		4	0.7	0.0737034	0.0679748
Augment Table	Columns (3/0)	5	0.7	0.0738753	0.0679748
Back	Concentration *	6	0.7	0.0712396	0.0679748
	Velocity (y) *	7	1.67	0.065042	0.061108
	🚄 Model (x) 🖶	8	1.67	0.0547667	0.061108
		9	3.75	0.0497128	0.0587330
		10	3.75	0.0642727	0.0587330
	Rows	11	6.25	0.0613005	0.0580090
	All rows 15 Selected 0	12	6.25	0.0643576	0.0580090
	Excluded 0	13	6.25	0.0393892	0.0580090
	Hidden 0	14	0.13262625	•	0.3943929
	Labelled 0	15	0.13262625	-	0.3943929

Figure 10.8 Making a Table with the Nonlinear Designer

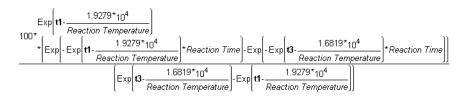
The new runs use the wider interval of allowed concentration, which leads to more precise estimates of k and Vmax.

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 has 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. The formula looks like this:



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 10.9.

206 Nonlinear Designs

Examples of Nonlinear Designs

Figure 10.9 Nonlinear Design launch Dialog

Create an optimal design for models Select Columns Reaction Time Reaction Temperature yield model observed yield	Cast Selected Columns into Roles Y, Response Image: Column Selected Vield X, Predictor Formula Image: Vield model	Action OK Cancel
	Weight optional Numeric X Predictor column must have formula	Remove Recall Help

- 4 Click **OK** to see the nonlinear design Factors and Parameters panels in Figure 10.10.
- 5 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.
- 6 Change the values of the parameter t1 to 25 and 50, and t3 to 30 and 35.
- 7 Click on the Distribution of each parameter and select **Uniform** from the menu to change the distribution from the default **Normal** (see Figure 10.10).
- 8 Change the number of runs to 12 in the Design Generation panel.

Figure 10.10 Change Factor Values, Parameter Distributions, and Number of Runs

•	Nonlinear Design				
)	Factors				
[Name	Role	Values		
ľ	Reaction Temperature	Continuous	510	540	
	Reaction Time	Continuous	0.1	0.3	
)	Parameters				
ſ	Name	Distribution	Values		
ľ	t1	Uniform	25	50	
	t3	North	30	35	
)*	Design Generation	Uniform Normal 12 Lognormal Exponential			

9 Click Make Design, then Make Table. Your results should look similar to those in Figure 10.11. Figure 10.11 Design Table

▼Nonlinear Design		• •	Reaction	Reaction		
Design Nonlinear De	sign	•	Temperature	Time	observed yield	yield model
Model		1	540.00	0.3	•	58.093235
	_	2	540.00	0.3	-	58.093235
💌 Columns (4/0)		3	540.00	0.3	-	58.093235
🥖 Reaction Temperature 🕇	k	4	540.00	0.3	-	58.093235
🚄 Reaction Time 🛠		5	531.45	0.1	-	39.7338437
d observed yield *		6	540.00	0.3	-	58.093235
🥖 yield model 🖶		7	539.68	0.1	-	54.5975641
Rows	_	8	539.68	0.1		54.5975641
All rows	12	9	510.00	0.1	•	11.2590804
Selected	12	10	539.68	0.1		54.5975641
Excluded	ŏ	11	531.44	0.1	•	39.7192502
Hidden	0	12	510.00	0.1	•	11.2590804

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 10.12).

11 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 10.12).

Figure 10.12 Reaction Kinetics.jmp

Reaction Kinetics	• •				
Design Nonlinear Design	•	Reaction Temperature	Reaction Time	observed yield	yield model
▼ Model	1	540	0.3	57	58.093235
	2	540	0.11	56	57.3016426
Columns (4/0)	3	540	0.3	59	58.093235
Reaction Temperature * Reaction Time *	4	530	0.2	53	54.9142081
A observed yield *	5	510	0.1	12	11.2590804
🔺 vield model 🐥	6	540	0.3	57	58.093235
	7	540	0.3	59	58.093235
Rows	8	510	0.1	10	11.2590804
All rows 12	9	517	0.1	18	17.8546023
Selected 0	10	530	0.21	56	55.9838913
Excluded 0	11	540	0.11	56	57.3016426
Hidden 0	12	540	0.3	59	58.093235
Labelled 0					

First, examine the design region with an overlay plot.

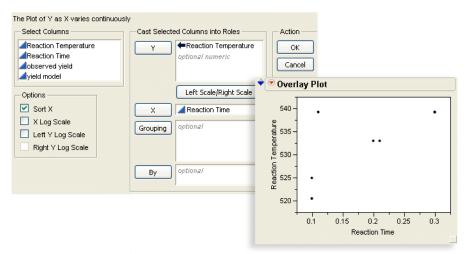
12 Selecting Graph > Overlay Plot.

13 Select Reaction Temperature and click Y

14 Select Reaction Time and click X as shown in the Overlay Plot launch dialog in Figure 10.13.

15 Click **OK** to see the overlay plot in Figure 10.13.

Figure 10.13 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 Select observed yield and click Y, Response.

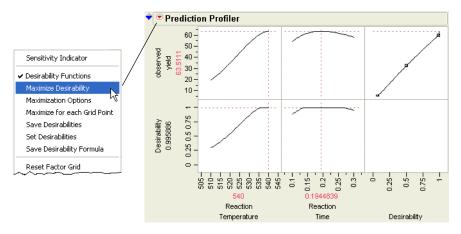
18 Select yield model and click the X, Predictor Formula, then click OK.

19 Click Go on the Nonlinear control panel.

- 20 Now, choose Profilers > Profiler from the red triangle menu on the Nonlinear Fit title bar.
- 21 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 10.14 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 *JMP User Guide*). 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. 208
- "Set Up Factors and Parameters in the Nonlinear Design Dialog," p. 209
- "Enter the Number of Runs and Preview the Design," p. 210
- "Make Table or Augment the Table," p. 211

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 10.15).

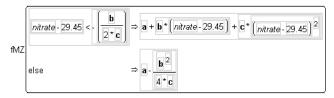
5 Click **OK** on the launch dialog to see the Nonlinear Design DOE panels. **Figure 10.15** Identify Response (Y) and the Column with the Nonlinear Formula (X)

Create an optimal design for models	that are poplinear in the parameters	
Select Columns	Cast Selected Columns into Roles	Action
✓ yield ✓ linear dquad	X, Predictor Formula	Cancel
	Weight optional numeric X Predictor column must have formula	Remove
		Recall Help

Set Up Factors and Parameters in the Nonlinear Design Dialog

First, look at the formula for quad, shown in Figure 10.16, and notice there are three parameters. These parameters show in the Parameters panel of the Nonlinear design dialog, with initial parameter values.

Figure 10.16 Formula for quad has Parameters a, b, and c



Use Figure 10.17 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 10.17.

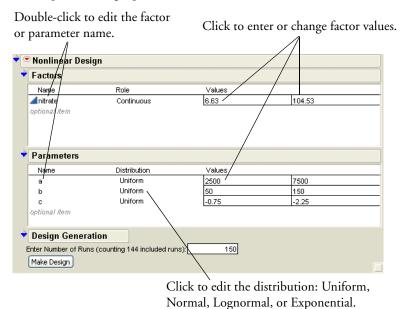


Figure 10.17 Example of Setting Up Factors and Parameters

Enter the Number of Runs and Preview the Design

- 1 The Design Generation panel shows a default number of runs needed to perform the experiment. This example needs one run for each observation (144) plus three additional runs for the three parameters giving 147 as a total minimum number of runs. Enter 147 in the Number of Runs edit box.
- 2 Click **Make Design** before creating the data table to preview the design. Figure 10.18 shows a partial listing of the design.

🕈 💌 Nonli	near Desi	ign				
🕈 Desig	jn		\sim	Run	nitrate	vield
Run	nitrate	yield		130	24.64	9405.75
1	9.59	4713		131	23.08	6278.25
2	7.51	1049.45		132	44.2	8368.39
3	8	1537.89		133	63.26	7379.55
4	13.93	5116.24		134	46	8120.8
5	29.1	8835.98		135	16.56	4771.33
6	32.91	9849.72		136	25.45	6980.81
7	52.68	9588.95		137	41.53	8616.87
8	64.24	9474.68		138	61.18	9250.9
9	23.41	6650.39		139	82.49	7967.53
10	20.09	7783.35		140	13.35	5158.1
11	38.36	9485.95		141	33.33	7627.44
12	47.4	9138.61		142	29.45	7863.4
13	68.23	8079.86		143	81.61	7780.94
14	24.58	2843.44		144	73.03	8151.96
15	18.56	6237.16		145	104.53	
16	26.32	7961.7		146	104.53	
17	30.98	8691.12		147	104.53	
18	26.4	_ 8906.83				

Figure 10.18 Example Preview Design

Make Table or Augment the Table

3 The last step is to click either **Make Table** or **Augment Table**. The **Make Table** command creates a new table (Figure 10.19) with all runs included. The **Augment Table** command adds the new runs to the existing table.

Figure 10.19 Partial Listing of an Example Nonlinear Design Table

Nonlinear	Design	♦ ●			
Design N	lonlinear Design		nitrate	yield	quad
💌 Model		1	9.59	4713	2422.3706
		2	7.51	1049.45	2083.9546
		3	8	1537.89	2164.84625
Columns (•••	4	13.93	5116.24	3086.6944
🚄 nitrate 🗙	ĸ	5	29.1	8835.98	4964.81625
🚄 yield 🗚 ⊿ quad 🐥		6	32.91	9849.72	5328.0426
🚈 quau 🐨		7	52.68	9588.95	6513.55065
Rows		8	64.24	9474.68	6666.66667
All rows	147 🔺	9	23.41	6650.39	4341.2776
Selected	0	10	20.09	7783.35	3932.5856
Excluded	0 —	. 11	38.36	9485.95	5771,91785

Advanced Options for the Nonlinear Designer

For advanced users, the Nonlinear Designer has the two additional options, as shown in Figure 10.20. 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

212 Nonlinear Designs

Advanced Options for the Nonlinear Designer

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 10.20 Advanced Options for the Nonlinear Designer

🕈 💌 Nonlinear Design		
Save Responses		
Load Responses		
Save Factors	le	Values
Load Factors	ontinuous	6.63
Save Constraints		
Load Constraints		
Simulate Responses		
 Number of Starts 		
Advanced Options	Number of Mor	nte Carlo Samples
M	N Monte Carlo	Spheres

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.

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 **11** 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
Taguchi Design Example
Analyze the Data
Creating a Taguchi Design
Detail the Response and Add Factors219
Choose Inner and Outer Array Designs 220
Display Coded Design221
Make the Design Table

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. The table below lists the recommended performance statistics.

Goal	S/N Ratio Formula
nominal is best	$\frac{S}{N} = 10\log\left(\frac{\overline{\gamma}^2}{s^2}\right)$
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.

The signal and noise factors in the Byrne Taguchi Data for this example appear in the table below.

Factor Name Type Levels Con		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
Adhesive	control	3	percent adhesive
Time	noise	2	the conditioning time
Temperature	noise	2	temperature
Humidity	noise	2	the relative humidity

Table 11.2

216 Taguchi Designs

Taguchi Design Example

To start this example:

- 1 Select **DOE > Taguchi**.
- 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 11.1.

Figure 11.1 Response, and Signal and Noise Factors for the Byrne-Taguchi Example

Y Larger Is Better I I I Factors Image: Isenter Values Image: Isenter Image: Isenter	oortance
Name Role Values Interfer Signal 1 2 3 Wall Signal 1 2 3 Wall Signal 1 2 3 Adhesive Signal 1 2 3 Adhesive Signal 1 2 3 Time Noise L1 L2 Humidity Noise L1 L2	
Interfer Signal 1 2 3 Wall Signal 1 2 3 Depth Signal 1 2 3 Adhesive Signal 1 2 3 Adhesive Signal 1 2 3 Time Noise L1 L2 1 Temperature Noise L1 L2 L	
Wall Signal 1 2 3 Depth Signal 1 2 3 Achesive Signal 1 2 3 Achesive Signal 1 2 3 Time Noise L1 L2 1 Temperature Noise L1 L2 1 Humidity Noise L1 L2 1	
Lepth Signal 1 2 3 Adhesive Signal 1 2 3 Time Noise L1 L2 Temperature Noise L1 L2 Humidity Noise L1 L2 actors L1 L2	
Adhesive Signal 1 2 3 Image: Time Noise L1 L2 1 Image: Temperature Noise L1 L2 1 Image: Humidity Noise L1 L2 1 Image: Additional state Noise L1 L2 1 Image: Additional state Noise L1 L2 1	
Line Noise L1 L2 Langerature Noise L1 L2 Lumidity Noise L1 L2	
L Temperature Noise L L L L L L L L L L L L L L L L L	
Humidity Noise L1 L2	
actors	
actors	
iner Array Outer Array	
Number Design Name	
9 L9 - Taguchi L4	
27 L27 - Taguchi L8	
31 Full Factorial	

- 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.

7 Click **Make Table** to create the design table shown in Figure 11.2.

	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-	•	•	•	-	•	•	•	•	•	•

Figure 11.2 Taguchi Design Before Data Entry

Now suppose the pull-off adhesive force measures are collected and entered into the columns containing missing data, as shown in Figure 11.3. 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.

8 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 11.3 shows the completed design.

Figure 11.3 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:

$$-10Log10 \quad \left[Mean \left[\frac{1}{(\mathsf{Y}---)^2}, \frac{1}{(\mathsf{Y}--+)^2}, \frac{1}{(\mathsf{Y}-+-)^2}, \frac{1}{(\mathsf{Y}-++^2)^2}, \frac{1}{(\mathsf{Y}+--)^2}, \frac{1}{(\mathsf{Y}+--)^2}, \frac{1}{(\mathsf{Y}++-)^2}, \frac{1}{(\mathsf{Y}+++)^2} \right] \right]$$

This expression is large when all of the individual *y* values are small.

Analyze the Data

The data in Byrne Taguchi Data.jmp are now ready to analyze. The goal of the analysis is to find factor settings that maximize both the mean and the signal-to-noise ratio.

1 Click the red triangle icon next to Model on the upper left of the data table and select Run Script.

The Model script produces the Fit Model dialog shown in Figure 11.4.

The Fit Model dialog that appears automatically has the appropriate effects. It includes the main effects of the four signal factors. The two responses are the mean (Mean Y) and signal-to-noise ratio (SN Ratio Y) over the outer array.

	Figure 11.4	Fit Model Dialog for Taguchi Data
--	-------------	-----------------------------------

Select Columns	Pick Role Variables	Personality:	Standard Least Squares	~
Interfer Vall Depth Adhesive Pattern Y Y+ Y+ Y+ Y+ Y+ Y+ Y+ Y+ Y+ Y+ Y+ SN Ratio Y	Y Mean Y SN Ratio Y optional numeric Weight optional numeric Freq optional numeric By optional Construct Model Effects Add Interfer Vvail Depth Adhesive Adhesive Degree 2	Emphasis: Help Recall Remove	Effect Leverage	
	Adhesive			

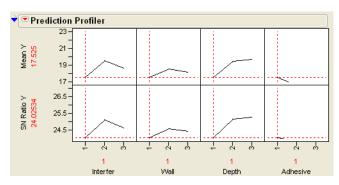
2 Click Run Model on the Fit Model dialog.

The prediction profiler is a quick way to find settings that give the highest signal-to-noise ratio for this experiment.

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 11.5) indicate that different settings of the first three factors would increase SN Ratio Y.

Figure 11.5 The Prediction Profiler



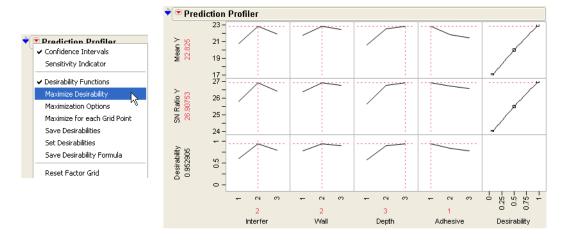
4 To find optimal settings, click the red triangle on the Prediction Profiler title bar and select **Desirability Functions**. This adds the row of traces and a column of function settings to the pro-

filer, as shown in Figure 11.6. The default desirability functions are set to larger-is-better, which is what you want in this experiment. See the chapter "Standard Least Squares: Perspectives on the Estimates" of *JMP Statistics and Graphics Guide* 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 L1 to L2 (see Figure 11.5 and Figure 11.6). The Depth setting changed from L1 to L3. The settings for Adhesive did not change. These new settings increased the signal-to-noise ratio from 24.0253 to 26.9075.

Figure 11.6 Best Factor Settings for Byrne Taguchi Data



Creating a Taguchi Design

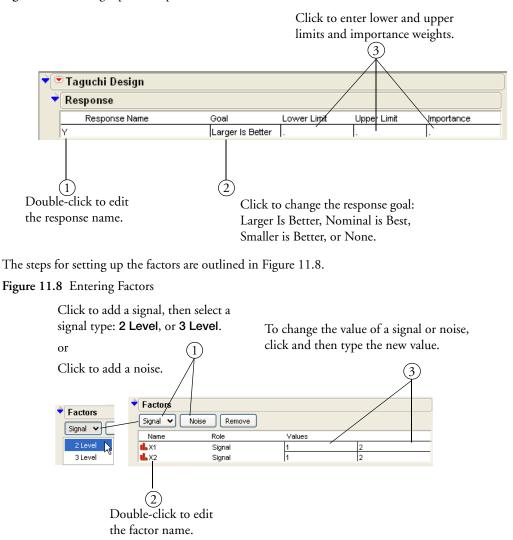
To start a Taguchi design, select **DOE** >**Taguchi Arrays**, or click the **Taguchi Arrays** button on the JMP Starter **DOE** page. Then, follow the steps below:

- "Detail the Response and Add Factors," p. 219
- "Choose Inner and Outer Array Designs," p. 220
- "Display Coded Design," p. 221
- "Make the Design Table," p. 221

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 11.7. For information on importance weights and lower and upper limits, see "Understanding Response Importance Weights," p. 60.

Figure 11.7 Setting Up the Response



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 11.9 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.



Factors —Choose Inner and Outer Array Designs—
Inner Array
Design Name
L12
L16
L32
L64
L128
Continue
Back

After you click **Continue**, a dialog appears that asks to you specify how many times you want to perform each inner array run. Specify one (1) for this example.

Display Coded Design

After you select a design type, click the disclosure buttons ($\diamondsuit \diamondsuit$ on Windows/Linux and $\blacktriangleright \neg$ on the Macintosh) to display the Coded Design (Figure 11.10).

Figure 11.10 Coding for Eight Factor L12 Design

Taguchi Design
8Factors
Taguchi Array L12
Design Description and Display
♦ Coded Design
Codes
+++
+++
-+-++
-++-++-+
-+++-
+-++++
+-+-+++-
++++-+
++++
++-+
+++-
Make Table
Back

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. 104.

Make the Design Table

When you click **Make Table**, a table similar to that shown in Figure 11.11 appears. In the data table, each row represents a run. In the values for the Pattern variable, plus signs designate high levels and minus signs represent low levels.

222

Taguchi Designs Creating a Taguchi Design

💌 Taguchi Array L12	◆													
Design Taguchi Array	•	X1	X2	X3	X4	X5	X6	X7	X8	Pattern	run 1	run 2	Mean	SN Ratio
▼ Model	1	1	1	1	1	1	1	1	1		•	•	•	•
Columns (13/0)	2	1	1	1	1	1	2	2	2	+++	•	-	•	-
L X1 *	3	1	1	2	2	2	1	1	1	+++	•	•	•	•
L X2 *	4	1	2	1	2	2	1	2	2	-+-+++	•	•	•	•
🔥 X3 🛠	5	1	2	2	1	2	2	1	2	-++-++-+	•	•	•	•
🔥 X4 🛠	6	1	2	2	2	1	2	2	1	-+++-++-	•	•	•	-
🔥 X5 🛠	7	2	1	2	2	1	1	2	2	+-++-++	•	•	•	•
📕 X6 🛠	8	2	1	2	1	2	2	2	1	+-+-+++-	•	•	•	•
L X7 *	9	2	1	1	2	2	2	1	2	++++-+	•	•	•	•
🔥 X8 🛠	10	2	2	2	1	1	1	1	2	++++	•	•	•	•
/ run 1	11	2	2	1	2	1	2	1	1	++-+	•	•	•	•
run 2	12	2	2	1	1	2	1	2	1	+++-	•	•	•	-
📕 Mean 🖶														
⊿ SN Ratio 🕂														

Figure 11.11	Taguchi De	sign Table for	· Eight Factor	L12 Design
0	0	0	0	0

Chapter 12

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
Analyze the Augmented Design
Creating an Augmented Design
Replicate a Design
Add Center Points
Creating a Foldover Design
Adding Axial Points
Adding New Runs and Terms
Special Augment Design Commands241
Save the Design (X) Matrix
Modify the Design Criterion (D- or I- Optimality)242
Select the Number of Random Starts243
Specify the Sphere Radius Value
Disallow Factor Combinations243

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 12.1.

Note: You can check **Group New Runs into Separate Blocks** 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 12.1 Augment Design Dialog for the Reactor Example

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					

4 Now click Augment on the Augment Design dialog to see the display in Figure 12.2.

This model shown in Figure 12.2 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.

226 Augmented Designs

A D-Optimal Augmentation of the Reactor Example

Figure 12.2 Initial Augmented Model

actu	rs					
Name		Role		Changes Va	lues	
Feed		Cont	inuous	Easy 10		15
🖉 Catal			inuous	Easy 1		2
🖉 Stir F			inuous	Easy 100		120
	perature		inuous	Easy 140		180
	entration		inuous	Easy 3		6
	new runs inf	to separate	block			
Const	traints					
Mode	.l					
1ain Ef	fects Intera	actions 🗸	RSM C	ross Powers	Remove Term	
,						
	t e ature tration t*Stir Rate	חר				
Stir Rate Tempera Concen Catalyst Catalyst	t e ature tration t*Stir Rate t*Concentratio	חס				
Stir Rate Tempera Concen Catalyst Catalyst	t e ature tration t*Stir Rate	on Catalyst	Stir Rate	Temperature	• Concentration	Percent Reacted
Stir Rate Temper Concern Catalyst Catalyst Facto	t e ature tration t*Stir Rate t*Concentration r Design		Stir Rate 100	Temperature 18(
Stir Rate Tempera Concern Catalyst Catalyst Facto Run 1 2	e ature tration t*Stir Rate t*Concentratio r Design Feed Rate	Catalyst		•) 6	Reacted
Stir Rate Tempera Concern Catalyst Catalyst Facto Run 1	t e ature tration t*Stir Rate t*Concentratio r Design Feed Rate 10	Catalyst 1	100 120 100	- 180) 6) 3	Reacted 44
tir Rate emperation catalyst catalyst Facto Run 1 2 3 4	t e ature tration t*Stir Rate t*Concentration r Design Feed Rate 10 10 10 10	Catalyst 1 1 2 2	100 120 100 120	- 180 180 140 140) 6) 3) 6) 3	Reacted 44 66 70 54
tir Rate empersioncem atalyst atalyst atalyst atalyst acto Run 1 2 3 4 5	t e ature tration t*Stir Rate t*Concentratic tr Design Feed Rate 10 10 10 10	Catalyst 1 2 2 1	100 120 100 120 100	180 180 140 140 140) 6) 3) 6) 3	Reacted 44 66 70 54 53
Stir Rate Tempers Concern Catalyst Catalyst Facto Run 1 2 3 4 5 6	t e ature tration t*Stir Rate t*Concentratio tr Design Feed Rate 10 10 10 10 15 15	Catalyst 1 2 2 1 1	100 120 100 120 100 120	180 180 140 140 140 140 140) 6) 3) 6) 3) 3	Reacted 44 66 70 54 53 55
tir Rate empera concernicatalyst catalyst Facto Run 1 2 3 4 5 6 7	t e ature tration t*Str Rate t*Concentration tr Design 10 10 10 10 15 15	Catalyst 1 2 2 1 1 2	100 120 100 120 100 120 100	180 180 140 140 140 140 140 140 180) 6) 3) 6) 3) 3) 6) 3	Reacted 44 66 70 54 53 55 93
tir Rate concern catalyst catalyst Facto Run 1 2 3 4 5 6	t e ature tration t*Stir Rate t*Concentratio tr Design Feed Rate 10 10 10 10 15 15	Catalyst 1 2 2 1 1	100 120 100 120 100 120	180 180 140 140 140 140 140) 6) 3) 6) 3) 3) 6) 3	Reacted 44 66 70 54 53 55
tir Rate empera- concent catalyst catalyst Facto Run 1 2 3 4 5 6 7 8	t e ature tration t*Str Rate t*Concentration tr Design 10 10 10 10 15 15	Catalyst 1 1 2 2 1 1 1 2 2 2 2	100 120 100 120 100 120 100	180 180 140 140 140 140 140 140 180) 6) 3) 6) 3) 3) 6) 3	Reacted 44 66 70 54 53 55 93

To continue with the augmented reactor design:

5 Choose **2nd** from the Interactions menu as shown in Figure 12.3. This adds all the two-factor interactions to the model. The Minimum number of runs given for the specified model is 16, as shown in the Design Generation text edit box.

Figure 12.3 Augmented Model with All Two-Factor Interactions

Constraints		
Model		
Main Effects	Interactions 💙 R5M Cross Powers 💙 Remove Term	
Intercept	2nd N	
Feed Rate	3rd V	
Catalyst	4th	
Stir Rate	Sth	
Temperature	501	
Concentration		
Catalyst*Stir Re	te	
Catalyst*Conce	ntration	
Feed Rate*Cate	lyst	
Feed Rate*Stir	Rate	
Feed Rate*Terr	perature	
Feed Rate*Con	centration	
Catalyst*Tempe	rature	
Stir Rate*Temp	arature	
Stir Rate*Conc	Intration	
Temperature*C	oncentration	

JMP now computes *D*-optimally augmented factor settings, similar to the design shown in Figure 12.4. Figure 12.4 D-Optimally Augmented Factor Settings

Desig	n					
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	1	120	140	3	
10	15	1	100	180	3	
11	10	2	100	180	3	
12	10	1	100	140	3	
13	10	1	120	140	6	
14	15	2	120	140	6	
15	15	2	100	140	3	
16	15	2	120	180	3	

Note: The resulting design is a function of an initial random number seed. To reproduce the exact factor settings table in Figure 12.4, (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 12.5 shows the random number (12834729) used to generate the runs in Figure 12.4.

Figure 12.5 Specifying a Random Number

Ý [Auamont Desian		
1	Save Responses		
	Load Responses		
	Save Factors	Choose a positive whole number as a seed for a random starting design. [12834729	
1	Load Factors		
	Save Constraints		
	Load Constraints		
	Set Random Seed		
	Simulate Responses 🕏		
	Show Diagnostics		
	Save X Matrix		
	Optimality Criterion		
	Number of Starts		
	Sphere Radius		
	Disallowed Combinations		
	Advanced Options		

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.

228 Augmented Designs

A D-Optimal Augmentation of the Reactor Example

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 12.6.
- 3 In the Column Info dialog that appears, change the response limit to **Maximize**, as shown on the right in Figure 12.6.

Figure 12.6 Change the Response Limits Column Property for the Percent Reacted Column

	-Percent Reacted	" in Table 'Reactor Augment Data'
j1i	Column Name	Percent Reacted
Columns (6/1)		
🚄 Feed Rate \star	Data Type	Numeric 🗸
A Catalyst 🛪	Modeling Type	Continuous 🗸
🚄 Temperature 🛠	Format	Best 🗸 Width 8
Concentration * Percent Reacted	Column Properti	es 🗸
Response Limits	Response Limits	Response Limits
× 0	optional item	Response Limits are bounds on a response's
		range of acceptability. The prediction and
		Maximize Niese values. Click
		Match Target
	Remove	Minimize
	- Comove	None
		importance 1

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 12.7.

Figure 12.7 Completed Augmented Experiment (Reactor Augment Data.jmp)

Reactor Augment Data]♦ ●						Percent
Design Augmented Design		Feed Rate	Catalyst	Stir Rate	Temperature	Concentration	Reacted
Model	1	10	1	100	180	6	44
Run Script 📐 🚃	2	10	1	120	180	3	66
Edit	3	10	2	100	140	6	70
Delete	4	10	2	120	140	3	54
📕 Stir Rate 🔺	5	15	1	100	140	3	53
Temperature 🗶	6	15	1	120	140	6	55
Concentration *	7	15	2	100	180	3	93
Percent Reacted 💥	8	15	2	120	180	6	82
	9	15	2	120	180	3	98
Rows	10	15	2	120	140	6	65
All rows 16	11	10	2	100	140	3	63
Selected 0	1 12	10	1	120	180	6	49
	13	15	1	100	140	6	63
Hidden (15	1	100	180	3	61
Labelled 0		10	1	120	140	6	59
	16	10	2	100	180	3	94

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 12.8.

Select Columns	Pick Role Variables Personality: Stepwise
Feed Rate Catalyst Stir Rate Temperature Concentration Percent Reacted	Y Percent Reacted Weight optional Numeric Freq optional Numeric By optional
	Construct Model Effects
	Add Feed Rate Cross Stir Rate Nest Concentration Macros Catalyst*Stir Rate Degree 2 Attributes Feed Rate*Catalyst Transform Feed Rate*Temperature

Figure 12.8 Fit Model Dialog for Stepwise Regression on Generated Model

6 When you click **Run Model**, the stepwise regression control panel appears. Click the check boxes for all the main effect terms.

Important: Choose **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 12.9.

Figure 12.9 Initial Stepwise Model

🕈 💌 Stepwis	e Fit							
Response: Per	cent Reac	ted						
▼ Stepwis	e Regre	ssion Con	trol					
Prob to Ente	er 0.05	50 r	Enter All					
Prob to Lea	ve 0.10	00						
Direction: M	ixed 1	ະ ເ	Remove All	J				
	Combine	~						
		(
Go Stop	Step	Make Mode	!)					
🔶 Current	Estimate	es						
SSE	DFE	MSE	RSquare	RSquar	e Adj	Ср	AIC	
1257.1058	10	125.71058	0.6933	0	.5399		81.82366	
Lock Entered	Parame	ter	E	stimate	nDF	SS	"F Ratio"	"Prob>F"
V	Intercept	t		66.8125	1	0	0.000	1.0000
	Feed Ra	te(10,15)		4.4375	1	315.0625	2.506	0.1445
	Catalysti	(1,2)	10.	1153846	1	1520.198	12.093	0.0059
	Stir Rate	(100,120)	-0.	3653846	1	1.983516	0.016	0.9025
	Tempera	ature(140,180)) 6.1	1538462	1	555.6264	4.420	0.0618
	Concent	ration(3,6)	-1.	7884615	1	41.58173	0.331	0.5779
	Catalyst	*Stir Rate		0	1	346.6875	3.427	0.0972
	Catalyst	*Concentratio	n	0	1	46.44643	0.345	0.5712
	Feed Ra	te*Catalyst		0	1	2.548951	0.018	0.8954
	Feed Ra	te*Stir Rate		0	1	0.064103	0.000	0.9834
	Feed Ra	te*Temperatu	ire	0	1	0.730769	0.005	0.9439
	Feed Ra	te*Concentra	tion	0	1	0.016484	0.000	0.9916
	Catalyst	*Temperature		0	1	667.5208	10.190	0.0110
	Stir Rate	*Temperature	в	0	1	22.5625	0.164	0.6945
	Stir Rate	*Concentratio	on	0	1	2.571429	0.018	0.8950
	Tempera	ature*Concen	tration	0	1	631.1429	9.074	0.0147

230 Augmented Designs

A D-Optimal Augmentation of the Reactor Example

7 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 12.10, 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 12.10 Completed Stepwise Model

Stepwise								
esponse: Perc	ent Reacted							
Stepwise	e Regressi	on Con	trol					
Prob to Ente	r 0.050	ſ	Enter All					
Prob to Leav	/e 0.100			1				
Direction: Mi	xed 🗸	l	Remove All	J				
	Combine	~						
			-					
Go Stop	Step M	ake Mode	IJ					
Current I	Estimates							
SSE	DFE	MSE	RSquare	RSquar	e Adj	Ср	AIC	
14	8	1.75	0.9966	0	.9936		13.8635	
Lock Entered	Parameter		E	stimate	nDF	55	"F Ratio"	'Prob>i
	Intercept		65.	1730769	1	0	0.000	1.000
	Feed Rate(1	0,15)		0	1	1.821862	1.047	0.340
	Catalyst(1,2)	10.	1153846	2	2086.357	596.102	0.000
	Stir Rate(10	0,120)	-0.	3653846	2	120.7143	34.490	0.000
	Temperature	e(140,180) 6.1	1538462	4	2113.795	301.971	0.000
	Concentratio	on(3,6)	-1.	7884615	2	600.625	171.607	0.000
	Catalyst*Stir	Rate		0	1	0.519231	0.270	0.619
	Catalyst*Co	ncentratio	n	0	1	0.069231	0.035	0.857
	Feed Rate*C	atalyst		0	2	4.370813	1.362	0.325
	Feed Rate*S	tir Rate		0	2	1.885965	0.467	0.647
	Feed Rate*T	emperatu	re	0	2	2.552632	0.669	0.546
	Feed Rate*C	Concentra	tion	0	2	1.838346	0.453	0.655
	Catalyst*Ter	nperature	6.1	7307692	1	566.1593	323.520	0.000
	Stir Rate*Te	mperature	9 2.8	2692308	1	118.7308	67.846	0.000
	Stir Rate*Co	ncentratio	n	0	1	1.969231	1.146	0.319
	Temperature			5576923	1	559.0433	319,453	0.000

- 8 After Stepwise is finished, click **Make Model** on the Stepwise control panel to generate this reduced model, as shown in Figure 12.11.
- 9 Click **Run Model** and fit the reduced model to do additional diagnostic work, make predictions, and find the optimal factor settings.

Figure 12.11 New Prediction Model Dialog

Select Columns	Pick Role Variables Personality: Standard Least Squares	s •
▲ Feed Rate ▲ Catalyst ▲ Stir Rate	Y APercent Reacted optional Emphasis: Effect Screening	•
Temperature Concentration	Weight optional Numeric Help Run Model	
Percent Reacted	Freq optional Numeric Remove	
	By optional	
	Construct Model Effects	
	Add Catalyst Stir Rate	
	Cross Temperature	
	Nest Concentration Catalyst*Temperature	
	Macros V Stir Rate*Temperature	
	Degree 2 Temperature*Concentration	
	Attributes 💌	

The Analysis of Variance and Lack of Fit Tests in Figure 12.12, indicate a highly significant regression model with no evidence of Lack of Fit.

Figure 12.12 Prediction Model Analysis of Variance and Lack of Fit Tests

Respons	e Perc	ent React:	ed	
Actual b	y Pred	icted Plot		
Summar	y of Fi	t		
Analysis	of Var	iance		
		Sum of		
Source	DF	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			
		Sum o	f	F Ratio
Source	DF	Squares	s Mean Square	0.2157
Lack Of Fit	6	5.500000	0.91667	Prob > I
Pure Error	2	8.500000) 4.25000	0.9394
Total Error	8	14.000000)	Max RS 0.997

The Sorted Parameter Estimates table in Figure 12.13, 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. 89, had ambiguous results.

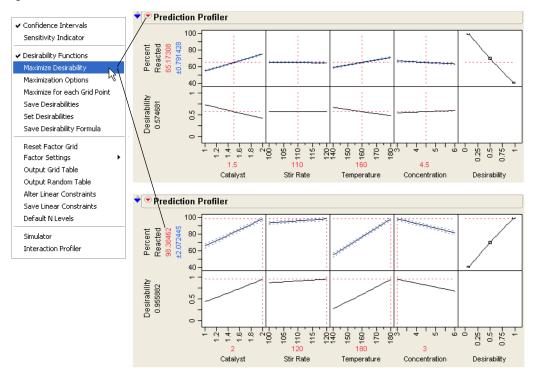
A D-Optimal Augmentation of the Reactor Example

Sorted Parameter Estimates										
Term	Estimate	Std Error	t Ratio		Prob> t					
Catalyst(1,2)	10.115385	0.343203	29.47		<.0001*					
Catalyst*Temperature	6.1730769	0.343203	17.99		<.0001*					
Temperature*Concentration	-6.557692	0.3669	-17.87		<.0001*					
Temperature(140,180)	6.1153846	0.343203	17.82		<.0001*					
Stir Rate*Temperature	2.8269231	0.343203	8.24		<.0001*					
Concentration(3,6)	-1.788462	0.3669	-4.87		0.0012*					
Stir Rate(100,120)	-0.365385	0.343203	-1.06		0.3181					

10 Chose Maximimize Desirability from the menu on the Prediction Profiler title bar.

The prediction profile plot at the bottom in Figure 12.14, 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 12.14 Maximum Percent Reacted



To summarize, compare the analysis of 16 runs with the analyses of reactor data from previous chapters:

- "Screening Designs," p. 89, 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. 129, 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. 233.

Add Centerpoints adds center points. See "Add Center Points," p. 235.

Fold Over creates a foldover design. See "Creating a Foldover Design," p. 236.

- Add Axial adds axial points together with center points to transform a screening design to a response surface design. See "Adding Axial Points," p. 237.
- 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. 238.

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 12.15).

Figure 12.15 Identify Response and Factors

Add more runs to an existing data ta terms.	able. Replicate, add centerpoints, fold over or add	l model
Select Columns	Cast Selected Columns into Roles	Action
Pattern Feed Rate Catalyst Stir Rate	Y, Response	OK Cancel
Temperature Concentration Percent Reacted	X, Factor Feed Rate Catalyst Stir Rate Temperature	Remove Recall Help

5 Click **OK** to see the Augment Design panel shown in Figure 12.16.

Creating an Augmented Design

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 12.16 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 s Augmentation Choices Replicate Add Center	· 	Add Axial A	ugment	

7 Click the **Replicate** button to see the dialog shown on the left in Figure 12.17. Enter the number of times you want JMP to perform each run, then click **OK**.

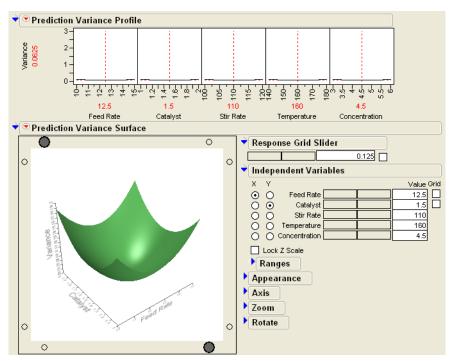
Note: Entering 2 specifies that you want each run to appear twice in the resulting design. This is the same as one replicate (Figure 12.17).

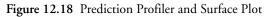
8 View the design, shown on the right in Figure 12.17.

Figure 12.17 Reactor Data Design Augmented With Two Replicates

		or Design					
mber times to perform each run.	Run	Feed Rate	Catalyst	Stir Rate	Temperature	Concentration	Percen Reacted
	1	10	- 1	100	- 180	6	4
OK Cancel	2	10	1	120	180	3	6
	3	10	2	100	140	6	7
	4	10	2	120	140	3	5
	5	15	1	100	140	3	5
	6	15	1	120	140	6	5
	7	15	2	100	180	3	9
	8	15	2	120	180	6	8
	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	
	🕨 🖻 Pre	ediction Va	riance P	rofile			
	🔶 🔍 Pre	ediction Va	riance S	urface			
	Make To Make To 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 12.18.





10 Click Make Table to produce the design table shown in Figure 12.19.

Figure 12.19 The Replicated Design

 Augmented Design 		• •						
Design Augmented Des	ign		Feed Rate	Catalyst	Stir Rate	Temperature	Concentration	Percent Reacted
 Model 		1	10	1	100	180	6	44
	_	2	10	1	120	180	3	66
Columns (6/0)		3	10	2	100	140	6	70
🚄 Feed Rate 苯		4	10	2	120	140	3	54
🚄 Catalyst 🗶		5	15	1	100	140	3	53
🚄 Stir Rate 🗱		6	15	1	120	140	6	55
Temperature * Concentration *		7	15	2	100	180	3	93
Percent Reacted 🗶		8	15	2	120	180	6	82
		9	10	1	100	180	6	•
Rows		10	10	1	120	180	3	-
All rows	16	11	10	2	100	140	6	-
Selected	0	12	10	2	120	140	3	-
Excluded	0	13	15	1	100	140	3	•
Hidden	0	14	15	1	120	140	6	•
Labelled	0	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.

236Augmented Designs

Creating an Augmented Design

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 12.15) 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 12.16).
- 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 12.20.

The table shows two center points appended to the end of the design.

Figure 12.20 Design with Two Center Points Added

Augmented Design	• •						
Design Augmented Design	•	Feed Rate	Catalyst	Stir Rate	Temperature	Concentration	Percent Reacted
▼Model	1	10	1	100	180	6	44
	2	10	1	120	180	3	66
Columns (6/0)	3	10	2	100	140	6	70
🚄 Feed Rate 🗶	4	10	2	120	140	3	54
🚄 Catalyst 🛠 ⊿ Stir Rate ★	5	15	1	100	140	3	53
Temperature 🗶	6	15	1	120	140	6	55
Concentration *	7	15	2	100	180	3	93
A Percent Reacted 🗶	8	15	2	120	180	6	82
	9	12.5	1.5	110	160	4.5	•
Rows	10	12.5	1.5	110	160	4.5	•
40							

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 12.15) and click **OK**.
- 4 Check the box to the left of **Group New Runs into Separate Block** (Figure 12.16). 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 12.21 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 12.21 Listing of a Foldover Design On All Factors

 Augmented Design 	_ ◆ ⊂ ♥							
Design Augmented Desig	an 🕤 🧹	Feed Rate	Catalyst	Stir Rate	Temperature	Concentration	Block	Percent Reacted
▼Model	1	10	1	100	180	6	1	44
Caluara (TAR)	2	10	1	120	180	3	1	66
Columns (7/0)	3	10	2	100	140	6	1	70
✓ Feed Rate ★ ✓ Catalyst ★	4	10	2	120	140	3	1	54
A Stir Rate 🗶	5	15	1	100	140	3	1	53
Temperature 🗶	6	15	1	120	140	6	1	55
Concentration 💥	7	15	2	100	180	3	1	93
I Block 🗱	8	15	2	120	180	6	1	82
🚄 Percent Reacted 🛠	9	15	2	120	140	3	2	-
	10	15	2	100	140	6	2	•
■ Rows	11	15	1	120	180	3	2	•
	16 12	15	1	100	180	6	2	•
Selected Excluded	0 13	10	2	120	180	6	2	•
Hidden	0 14	10	2	100	180	3	2	•
Labelled	0 15	10	1	120	140	6	2	•
	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 12.15) 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 12.16).
- 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 12.22 Entering Axial Values

Axial values dialog	
Please supply an axial value.	1
Please supply the number of center points desired.	2
ОК	
Cancel	//

7 Click **Make Table**. The design table appears. Figure 12.23 shows a table augmented with two center points and two axial points for five variables.

238 Augmented Designs

Creating an Augmented Design

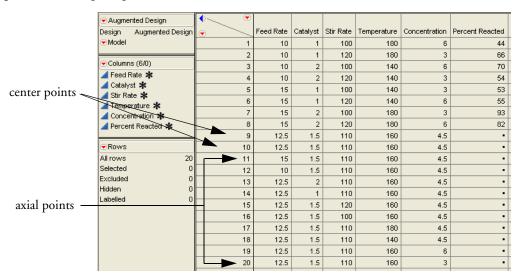


Figure 12.23 Design Augmented With Two Center and Ten Axial Points

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:

- 1 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.
- 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 12.15) 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** (not used in this example).
- 5 Click the Augment button. The original number of runs (Figure 12.24) appear in the Factor Design panel.

Figure 12.24 Viewing the Existing Design

	0		0	U			
Augn	nent Desig	n					
Facto	rs						
Name	•	Role		Changes Valu	ies		
Feed	Feed Rate Continuous		Easy 10		15		
Catal	·		inuous	Easy 1		2	
Stir F			inuous	Easy 100		120	
	perature		inuous	Easy 140 Easy 3		180 6	
_	entration		inuous	Easy 3		lp	
	new runs inf traints	to separate	block				
\geq							
Mode							
Main Ef	fects Intera	actions 👻		ross Powers	Remove Term		
Intercep	ot						
Feed Ra	ate						
Catalys							
Stir Rat							
Temper							
Concen							
	t*StirRate						1
	r Design]	
	Feed Rate	Catalyst	Stir Rate	Temperature	Concentration	Percent Reacted	
1 2	10 10	1	100 120	180 180	6 3	44 66	
3	10	2	120	140	5	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 14	15 15	1	100 100	140 180	6	63 61	
14	15	1	100	180	з 6	59	
15	10	2	120	140	3	59 94	
	n Generat		.50	100			
	mber of Runs		6 included r	uns):	24		
Make D		(counting)	o molaada f	unio).			
L'Hake D	osigit						

12 Augmented Designs

- 6 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.
- 7 Click the **Make Design** button. The new number of runs (Figure 12.25) appear in the Design panel.

Augmented Designs

Creating an Augmented Design

Figure 12.25 24 Total Runs

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	10	1	100	140	3	
18	10	1	100	140	3	
19	10	1	100	140	3	
20	10	1	100	140	3	
21	10	1	100	140	3	
22	10	1	100	140	3	
23	10	1	100	140	3	
24	10	1	100	140	3	
💌 Pre	diction Va	riance P	rofile			
🔍 Pre	diction Va	riance S	urface			
Relat	ive Varian	ce of Co	efficients			
Alias	Matrix					
-Output						
	<u> </u>					
Make Ta						

- 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 12.26) with the additional runs.

 Augmented Design 	◆ ●							
Design Augmented Design		Feed Rate	Catalyst	Stir Rate	Temperature	Concentration	Percent Reacted	
Model	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	
💌 Columns (6/0)	8	15	2	120	180	6	82	
🚄 Feed Rate 🗚	9	15	2	120	180	3	98	
🚄 Catalyst 🗶	10	15	2	120	140	6	65	
🚄 Stir Rate 苯	11	10	2	100	140	3	63	
A Temperature 🗱	12	10	1	120	180	6	49	
Percent Reacted *	13	15	1	100	140	6	63	
Z Percent Neacted 🦡	14	15	1	100	180	3	61	
	15	10	1	120	140	6	59	
	16	10	2	100	180	3	94	
	17	10	1	100	140	3	•	٦
	18	10	1	100	140	3	•	
	19	10	1	100	140	3	•	
Rows	20	10	1	100	140	3	•	additional
All rows 24	21	10	1	100	140	3	•	runs
Selected 0	22	10	1	100	140	3	•	1 4110
Excluded 0	23	10	1	100	140	3	•	
Hidden 0	24	10	1	100	140	3		
I shalled 0								

Figure 12.26 The Augmented Design Table with New Runs

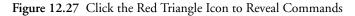
Technical Note: The Augment designer does not support designs that have terms whose estimability has been set to If Possible instead of Necessary, as is done in some screening designs that have fewer runs than terms.

Special Augment Design Commands

After you select **DOE > Augment Design** and identify factors and responses, the window in Figure 12.27 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. 74. The following sections describe commands found in this menu that are specific to augment designs.

242 Augmented Designs

Special Augment Design Commands



	Auamont Nocian						
	h© Save Responses						
	Load Responses			01	1.1-1		
	Save Factors		Changes Values			100	
	Load Factors		nuous	Easy	44	98	98
	Save Constraints						
	Load Constraints						
	Set Random Seed						
[Simulate Responses		⊃k				
	Save X Matrix						
	Optimality Criterion	•	old Over	Add Axial	Augment		
	Number of Starts		010 0 101	Had Hold	Hagmone		
88682	Sphere Radius						
	Disallowed Combinations						
	Advanced Options	•					

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 12.27) and select **Save X Matrix**. The script is saved as a table property called **Design Matrix**. When you run this script, JMP creates a global matrix called X 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 12.28) 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 12.28 Change the Optimality Criterion

💌 Auamont Docian					
Save Responses					
Load Responses					
Save Factors		Changes			
Load Factors	nuous nuous	Easy Easy	10	2	
Save Constraints	nuous	Easy	100	120	
Load Constraints	nuous	Easy	140	180	
Set Random Seed	nuous	Easy	3	6	
Simulate Responses	⊃k				
Save X Matrix					
Optimality Criterion	💽 🗸 Recom	mended	hent		
Number of Starts	Make [0-Optimal Des			
Sphere Radius	Make I	-Optimal Des	ign		-
Disallowed Combinations Advanced Options	•				

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 12.28) and select **Number of Starts**. The window in Figure 12.29 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 12.29 Changing the Number of Starts

Choose a positive whole number for the number of random starting designs.	1
OK Cancel	

For additional information on the number of starts, see "Why Change the Number of Starts?," p. 79.

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 12.27) 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. 80.

Chapter 13

Prospective Power and Sample Size



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, JMP computes the third.

Using the **Sample Size and Power** command when doing a prospective analysis helps 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 give JMP an estimate of the group means and sample sizes in a data table as well as an estimate of the within-group standard deviation (σ).

The sample size and power computations determine how large a sample is needed to be reasonably likely that an experiment or sample will yield a significant result, given that the true effect size is at least a certain size. It requires that you enter any two of three quantities, difference to detect, sample size, and power, and computes the third for the following cases:

- difference between one sample's mean and a hypothesized value
- difference between two samples means
- differences in the means among k samples
- difference between a standard deviation and a hypothesized value
- difference between 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 a number-of-parameters adjustment is specified.

Contents

Prospective Power Analysis 247
One-Sample and Two-Sample Means248
Single-Sample Mean 249
Power and Sample Size Animation for a Single Sample
Two-Sample Means
k-Sample Means
One Sample Standard Deviation
One-Sample and Two-Sample Proportions254
Counts per Unit
Sigma Quality Level

Prospective Power Analysis

The following five values have an important relationship in a statistical test on means:

- **Alpha** Alpha is the significance level that prevents declaring a zero effect significant more than alpha portion of the time.
- **Error Standard Deviation** Error Standard Deviation is the unexplained random variation around the means. An estimate of the Error Standard Deviation could be the root mean square error (RMSE) from a previous model fit.
- **Sample Size** Sample size is how many experimental units (runs, or samples) are involved in the experiment.
- **Power** Power is the probability of declaring a significant result.
- **Effect Size** Effect size is how different the means are from each other or from the hypothesized value.

The **Sample Size and Power** command in JMP helps estimate in advance either the sample size needed, power expected, or the effect size expected in the experimental situation where there is a single mean or proportion comparison, a two sample or proportion comparison, when comparing k sample means, or when comparing counts per unit.

When you select **DOE** > **Sample Size and Power**, the panel shown in Figure 13.1 appears with button selections for experimental situations. The following sections describe each of these selections and explains how to enter estimated parameter values and the desired computation.

Figure 13.1 Sample Size and Power Choices

▼ Sample Size	Sample Size				
Prospective Power and Sample Size	Prospective Power and Sample Size Calculations				
Select Situation for Sample Size or Power calculation					
One Sample Mean	Sample Size for testing a mean in a single sample				
Two Sample Means	Testing that the means are different across 2 samples				
k Sample Means	Testing that the means are different across k samples				
One Sample Standard Deviation	Sample Size for detecting a change in the standard deviation.				
One Sample Proportion	Sample Size for testing a proportion in a single sample				
Two Sample Proportions	Sample Size for testing a proportion across 2 samples				
Counts per Unit	Sample Size for detecting change in count per unit, e.g. DPU (defects per unit)				
Sigma Quality Level	Calculator for a popular index in terms of defects per opportunity.				

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 13.1), the Power and Sample Size dialog in Figure 13.2 appears and asks for the anticipated experimental values. As an example, consider the two-sample situation.

Figure 13.2 Initial Power and Sample Size Dialogs for Single Mean (left) and Two Means (right)

Sample Size	★ Sample Size
One Mean	Two Means
Testing if one mean is different from the hypothesized value.	Testing if two means are different from each other.
Alpha 0.050	Alpha 0.050
Std Dev .	Std Dev .
Extra Parameters 0	Extra Parameters 0
Supply two values to determine the third.	Supply two values to determine the third.
Enter one value to see a plot of the other two.	Enter one value to see a plot of the other two.
Difference to detect	Difference to detect
Sample Size	Sample Size .
Power .	Power .
Continue	Sample Size is the total sample size; per group would be n/2
	Continue
Back	
Animaking Carink	Back
Animation Script	

The dialogs are the same except the One Mean dialog has a button at the bottom that accesses an animation script.

The choice in the initial Power and Sample Size dialog always requires values for Alpha and the error standard deviation (Std Dev), and one or two of the other three values: Difference to detect, Sample Size, and Power. The power and sample size platform then calculates the missing item. If there are two unspecified fields, the power and sample size platform constructs a plot that shows the relationship between those 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 Power and Sample Size dialog asks for these values:

- **Alpha** Alpha is the significance level, usually 0.05. This implies willingness to accept (if the true difference between groups is zero) that 5% (alpha) of the time a significant difference will be incorrectly declared.
- **Std Dev** Std Dev (error standard deviation) is the true residual error. 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** 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 the 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—two parameters for the extra main effects, and three parameters for the interactions.

In practice, it isn't very important what values you enter here unless the experiment is in a range where there is very few degrees of freedom for error.

- **Difference 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** Sample size is the total number of observations (runs, experimental units, or samples). Sample size is not the number per group, but the total over all groups. Computed sample size numbers can have fractional values, which you need to adjust to real units. This is usually done by increasing the estimated sample size to the smallest number evenly divisible by the number of groups.
- **Power** Power is the probability of getting a statistic that will be declared statistically significant. Bigger power is better, but the cost is higher in sample size. Power is equal to alpha when the specified effect size is zero. You should go for powers of at least 0.90 or 0.95 if you can afford it. If an experiment requires considerable effort, plan so that the experimental design has the power to detect a sizable effect, when there is one.
- Continue Evaluates at the entered values.
- **Back** Goes back to the previous dialog.
- **Animation Script** The **Animation Script** button runs a JSL script that displays an interactive plot showing power or sample size. See the section, "Power and Sample Size Animation for a Single Sample," p. 250, for an illustration of this animation script.

Single-Sample Mean

Suppose there is a single sample and the goal is to detect a difference of 1.5 where the error standard deviation is 2, as shown in the left-hand dialog in Figure 13.3. To calculate the power when the sample size is 20, leave Power missing in the dialog and click **Continue**. The dialog on the right in Figure 13.3, shows the power is calculated to be 0.0.8888478174 (round to 0.89).

Figure 13.3 A One-Sample Example

▼ Sample Size	♦ Sample Size
One Mean	One Mean
Testing if one mean is different from the hypothesized value.	Testing if one mean is different from the hypothesized value.
Alpha 0.050	Alpha 0.050
Std Dev 2	Std Dev 2
Extra Parameters 0	Extra Parameters 0
Supply two values to determine the third.	Supply two values to determine the third.
Enter one value to see a plot of the other two.	Enter one value to see a plot of the other two.
Difference to detect 1.5	Difference to detect 1.5
Sample Size 20	Sample Size 20
Power .	Power 0.8888478174
Continue	Continue
Back	Back
Animation Script	Animation Script

One-Sample and Two-Sample Means

To see a plot of the relationship of power and sample size, leave both Sample Size and Power empty in the dialog and click **Continue**.

The plots in Figure 13.4, 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 13.4 shows using the crosshair tool to illustrate the example in Figure 13.3.

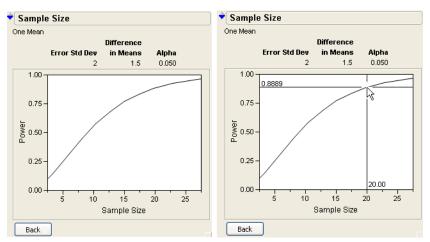


Figure 13.4 A One-Sample Example Plot

When only Sample Size is specified (Figure 13.5) and Difference to Detect and Power are empty, a plot of power by difference appears.

Figure 13.5 Plot of Power by Difference to Detect for a Given Sample Size

▼ Sample Size	★ Sample Size
One Mean	One Mean
Testing if one mean is different from the hypothesized value.	Error Std Dev Sample Size Alpha
Alpha 0.050	2 20 0.050
Std Dev 2	1.00
Extra Parameters 0	0.75-
Supply two values to determine the third.	
Enter one value to see a plot of the other two.	8 0.50 -
Difference to detect	
Sample Size 20	0.25-
Power .	
Continue	0.00
Back	0.0 0.5 1.0 1.5 2.0 2.5
	Difference
Animation Script	
	Back

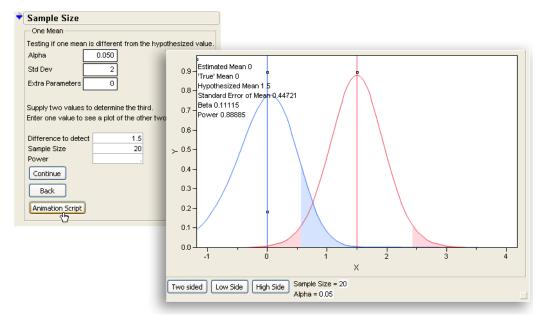
Power and Sample Size Animation for a Single Sample

The **Animation Script** button on the Power and Sample Size dialog for the single mean displays an interactive plot that illustrates the effect that changing the sample size has on power. The example in Figure 13.6 shows a Sample Size of 20, Alpha is 0.05, and the Difference to Detect is set to 1.5. The

initial animation shows a normal curve positioned with mean at zero (representing the estimated mean and the true mean), and another with mean at 1.5 (the difference to be detected). The probability of committing a Type II error (not detecting a difference when there is a difference), often represented as β in literature, is shaded in blue on this plot.

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 13.6 Example of Animation Script to Illustrate Power



Two-Sample Means

Choose Two Sample Means from the Power and Sample Size options (see Figure 13.1). The dialogs work similarly for one and two samples; the Difference to Detect is the difference between two means. Suppose the standard deviation is 2 (as before), the desired detectable difference is 1.5, and the sample size is 30 (15 per group).

Leave Power blank and click **Continue** to see the power calculation, 0.509347, as shown in the dialog on the left in Figure 13.7. This result is considerably lower than in the single sample (0.89) because each mean has only half the sample size. The comparison is between two random samples instead of one sample and an hypothesized mean.

To have a greater power requires a larger sample. To find out how large, leave both Sample Size and Power blank and examine the resulting plot, shown on the right in Figure 13.7. The crosshair tool estimates that a sample size of about 78 is needed to obtain a power of 0.9.

🕈 Sample Size Sample Size Two Means Two Means Difference Testing if two means are different from each other Error Std Dev in Means Alpha Alpha 0.050 0.050 2 1.5 Std Dev 2 1.00 0.9000 Extra Parameters 0 hý 0.75 Supply two values to determine the third Enter one value to see a plot of the other two. Ъоме D.50 Difference to detect 1.5 Sample Size 30 0.5093479534 Power 0.25 Sample Size is the total sample size; per group would be n/2 77.50 Continue 0.00 40 50 60 70 80 90 100 10 20 30 Back Sample Size Back

Figure 13.7 Plot of Power by Sample Size to Detect for a Given Difference

k-Sample Means

The **k-Sample Means** situation can compare up to 10 means. The next example considers a situation where 4 levels of means are expected to be about 10 to 13, and the standard deviation is 0.9. When a sample size of 16 is entered the power calculation is 0.95, as shown in the dialog on the left in Figure 13.8.

If both Sample Size and Power are left blank, the power and sample size calculations produce the power curve shown on the right in Figure 13.8. 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.

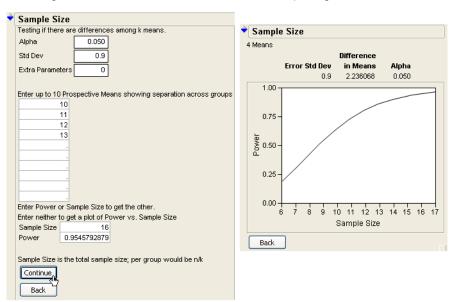


Figure 13.8 Prospective Power for k-Means and Plot of Power by Sample Size

One Sample Standard Deviation

The **One-Sample Standard Deviation** choice on the Power and Sample Size dialog (Figure 13.1) determines sample size for detection of a change in standard deviation. The usual purpose of this option is to compute a large enough sample size to guarantee that the risk of accepting a false hypothesis (β) is small.

In the dialog, specify

- Alpha Alpha is the significance level, usually 0.05. This implies willingness to accept (if the true difference between standard deviation and the hypothesized standard deviation is zero) that 5% (alpha) of the time a significant difference will be incorrectly declared.
- Hypothesized Standard Deviation Hypothesized or baseline standard deviation to which the sample standard deviation is compared.
- Alternative Standard Deviation Select Larger or Smaller from the menu to indicate the direction of the change you want to detect.

In the lower part of the dialog you enter two of the items and the Power and Sample Size calculations determines the third.

The examples throughout the rest of 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 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 and

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).

The hypothesis to be tested is:

Ho: $v_1 = v_2$, where v_1 is the hypothesized variance, and $v_2 = v_1$ plus the difference to detect.

In JMP, the computations use the standard deviation. The hypothesis is stated:

Ho: $s_1 = s_2$, where s_1 is the hypothesized standard deviation, and $s_2 = s_1$ plus the difference to detect.

In this example the hypothesized standard deviation, s_1 , is 10 (the square root 100) and s_2 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 known hypothesized standard deviation of 10, with an alpha of 0.05 and power of 0.99.

Enter these items as shown in Figure 13.9. When you click **Continue**, the computed result shows that you need a sample size of 171 (sample size rounded up the next whole number).

If you select **Smaller** from the Alternative Standard Deviation menu (you want to detect a change to a smaller standard deviation), enter a negative amount in the Difference to Detect box.

Figure 13.9 Dialog To Compare Single-Direction One-Sample Standard Deviation

Sample Size	★ Sample Size
One Sample Standard Deviation	One Sample Standard Deviation
Sample Size for detecting a difference in the standard deviation.	Sample Size for detecting a difference in the standard deviation
Alpha 0.050	Alpha 0.050
Hypothesized Standard Deviation 10	Hypothesized Standard Deviation 10
Alternative Standard Deviation Larger 🗸	Alternative Standard Deviation Larger 🗸
Supply two values to determine the third.	Supply two values to determine the third.
Difference to detect 2 4499	Difference to detect 2.4499
Sample Size	Sample Size 171
Power 0.99	Power 0.99
Continue	Continue
	Back
Back	

One-Sample and Two-Sample Proportions

The dialogs and computations to test power and sample sizes for proportions are similar to those for testing sample means. You enter a baseline **Proportion** (known or hypothesized proportion) and choose an **Alpha** level. Then, for the one-sample proportion case, enter the **Sample Size** to obtain the estimated **Power** or enter the **Power** to obtain the estimated **Sample Size**. You can also see a plot of **Sample Size** versus **Power** by leaving the two fields blank and clicking **Continue**. For the two-sample proportion case, either the two sample sizes or the desired **Power** must be entered. (See Figure 13.10 and Figure 13.11). The sampling distribution for proportions is actually binomial, but the computations to determine sample size and test proportions use exact methods described in Agresti and Coull (1998).

One Sample Proportion

The **One Sample Proportion** choice on the Power and Sample Size dialog (Figure 13.10) determines either a difference in proportion, the power, or the sample size, for detection of a change in proportion in a single sample from a known or hypothesized proportion. 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.

Often, there is a known proportion of defects (or other proportion of interest) and you want 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 given increase 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 true proportion and p_0 is the null proportion to test against.

In the top portion of the dialog, specify Alpha and the following quantity:

Proportion True proportion, which could be known or hypothesized.

In the bottom portion of the dialog, enter two of the following quantities to see the third, or a single quantity to see a plot of the other two.

Null Proportion Proportion to test against.

Sample Size Desired sample size, or blank for computation.

Power Desired power, or blank for computation.

For example, suppose an assembly line has a known proportion of defects of 0.3 and you want to know the power to detect when the proportion is different from 0.5 in a sample of 25. That is, you want to be able to detect a change of 0.2 in the proportion of defects, for an alpha level of 0.05 in a sample size of 25.

Click Continue to see the Power computation of approximately 0.7.

One-Sample and Two-Sample Proportions

Figure 13.10 Power and Sample Dialog for One-Sample Proportions

Sample Size]
One Proportion		
Testing if one proportion is different	from the hypothesized value.	
Alpha 0.050	Ho: P = Po	
Proportion 0.3	110. F - F0	
Enter one value to see a plot of the	other two.	
Null Proportion 0.5		
Sample Size 25		
Power .		
Continue Back	Null Proportion 0 Sample Size 2 Power 0.704478525	5
	Continue Back	

Two Sample Proportions

The **Two Sample Proportions** choice computes the power or sample size 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 p1 and p2 are the true proportions from two populations, and D_0 is the hypothesized difference in proportions.

If you enter any two of the three quantities (Null Difference in Proportion, Sample Size 1 and Sample Size 2, or Power), the third is computed.

As an example, suppose Proportion 1 is 0.2, Proportion 2 is 0.4. You want to know the power for a sample size of 50 (25 in each group) to detect a difference of 0.175.

In the Two Proportions dialog, enter:

- 1 0.2 as Proportion 1.
- 2 0.4 as Proportion 2.
- 3 Enter 0.175 as the Null difference in Proportion.
- 4 Enter 25 as Sample Size 1 and 25 as Sample Size 2 (for a total sample size of 50).
- 5 Leave Power blank.

The completed dialog should look like the one on the left in Figure 13.11.

6 Click **Continue** to see the right dialog in Figure 13.11, which shows the power calculation of approximately 0.83.

Suppose you want to know the sample size needed to obtain a power of 0.9, that detects the same difference (0.175) and given the same observed proportions (0.2 and 0.4).

- 7 Enter the power you want (0.9) and leave the two sample size edit boxes blank.
- 8 Click **Continue** to see the results on the lower right in Figure 13.11.

Equal sample sizes are computed (rounded to the next whole number). In this example, you need 30 in each group to obtain a power of 0.9 when detecting a difference in proportions of 0.175.

Note: The computations for finding two sample sizes may take a little time so be patient.

Figure 13.11 Difference Between Two Proportions

Ŷ	Sample Size	\	Sample Size
	Two Proportions		-Two Proportions
	Testing if two proportions are different from each other. Alpha 0.050 Proportion 1 0.2 Proportion 2 0.4		Testing if two proportions are different from each other. Alpha 0.050 Ho: P1 - P2 = ∆o Proportion 1 0.2 Ho: P1 - P2 = ∆o Proportion 2 0.4
	Supply two of (difference, sample sizes, power) to determine the third. When entering sample sizes, enter a value for both groups.		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 0.175 Sample Size 1 25 Sample Size 2 25 Power .		Null Difference in Proportion 0.175 Sample Size 1 25 Sample Size 2 25 Power 0.8343803193
	Continue Back		Back
4	Comula Sia	1	amula Siaa
1	Sample Size	· · ·	ample Size
	Two Proportions Testing if two proportions are different from each other. Alpha 0.050 Proportion 1 0.2 Proportion 2 0.4	Те Д	Two Proportions esting if two proportions are different from each other. Jpha 0.050 roportion 1 0.2 Ho: P1 - P2 = Δo roportion 2 0.4
	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 0.175 Sample Size 1 . Sample Size 2	N S	upply two of (difference, sample sizes, power) to determine the third. An en entering sample sizes, enter a value for both groups. ull Difference in Proportion 0.175 ample Size 1 30 ample Size 2 30
	Power 0.9 Continue Back	P	Continuen Back

Counts per Unit

The **Counts per Unit** selection from the Power and Sample size selection list calculates sample size for the Poisson-distributed counts typical when you can 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, the area and count are assumed to be 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 dialog, 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 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 in those units, rounded to the next whole number.

As an example, consider a wafer manufacturing process with a target of 4 defects per wafer and you want to verify that a new process meets that target within a difference of 2 defects per wafer. In the Counts per Unit dialog:

- 9 Enter an Alpha of 0.1 to be the chance of failing the test if the new process is as good as the target.
- 10 Enter 4 as the Baseline Counts per Unit, indicating the target of 4 defects per wafer.
- 11 Enter 2 as the Difference to detect.
- 12 Enter a power of 0.9, which is the chance of detecting a change larger than 2 (6 defects per wafer). In this kind of situation, alpha is sometimes called the *producer's risk* and beta is called the *consumer's risk*.
- 13 Click **Continue** to see the results in Figure 13.12, showing a computed sample size of 9 (rounded to the next whole number).

The process meets the target if there are less than 54 defects (6 defects per wafer in a sample of 9 wafers).

Figure 13.12 Dialog For Counts Per Unit Example

Sample Size) 🔶	Sample Size
Counts per Unit	ſ	Counts per Unit
Detecting change in count per unit, e.g. DPU (defects per unit)		Detecting change in count per unit, e.g. DPU (defects per unit)
Alpha 0.100		Alpha 0.100
Baseline Count per Unit 4		Baseline Count per Unit 4
Supply two values to determine the third.		Supply two values to determine the third.
Difference to detect 2		Difference to detect 2
Sample Size		Sample Size 9
Power 0.9		Power 0.9
Using normal approximations		Using normal approximations
Continue		Continue
Back		Back
DOLK		Dduk

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 results in a '6-sigma process. The computation that gives the Sigma Quality Level statistic is

Sigma Quality Level = NormalQuantile(1 - defects/opportunities) + 1.5

Select **DOE > Sample Size and Power** and then click the **Sigma Quality Level** button to see the Sample Size dialog:

Sample Size		
Sigma Quality Level		
NormalQuantile(1-defects/opportunities)+1.5		
Supply two values to determine the third.		
Number of Defects		
Number of Opportunities		
Sigma Quality Level .		
Continue		
Back		

Enter any two of the three quantities on the dialog:

- number of defects
- number of opportunities
- Sigma quality level

When you click **Continue**, the sigma quality calculator computes the missing quantity.

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 Then, click the Sigma Quality Level button.
- 3 Enter 50 for the Number of Defects and 1,000,000 as the Number of Opportunities, as shown in the window to the left in Figure 13.13.
- 4 Click **Continue**. The results, as shown in the window on the right in Figure 13.13, are a Sigma quality level for 5.39 defects in 1,000,000 opportunities.

Figure 13.13 Sigma Quality Level Example 1

Sample Size	▼ Sample Size
Sigma Quality Level	Sigma Quality Level
NormalQuantile(1-defects/opportunities)+1.5	NormalQuantile(1-defects/opportunities)+1.5
Supply two values to determine the third.	Supply two values to determine the third.
Number of Defects 50	Number of Defects 50
Number of Opportunities 1000000	Number of Opportunities 1000000
Sigma Quality Level .	Sigma Quality Level 5.3905918864
Continue	Continue
Back	Back

If you want to know how many defects reduce the Sigma Quality Level to 'six-sigma' for 1,000,000 opportunities, enter 6 as the Sigma Quality Level and leave the Number of Defects blank (dialog on the left in Figure 13.14). The computation (dialog on the right in Figure 13.14) shows that the Number of Defects cannot be more than approximately 3.4.

Figure 13.14 Sigma Quality Level Example 2

▼ Sample Size	▼ Sample Size
Sigma Quality Level	Sigma Quality Level
NormalQuantile(1-defects/opportunities)+1.5	NormalQuantile(1-defects/opportunities)+1.5
Supply two values to determine the third.	Supply two values to determine the third.
Number of Defects Number of Opportunities Sigma Guality Level 6 Continue Back	Number of Defects 3.3976731247 Number of Opportunities 1000000 Sigma Quality Level 6 Continue Back

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 Sudfianto, 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
- 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.
- 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.

- 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 Comunication," 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

Α

ABCD designs 151 aberration designs 100 acceptable values See lower limits and upper limits Actual-by-Predicted plots 111 adding center points in augment designs 235-237 factors 37, 61 linear constraints 147 runs and terms in augment designs 238 additional runs 71, 106, 126, 137 advanced options (nonlinear designer) 211 algorithms, coordinate exchange 88 alias matrix 69 explanation 18 showing 17 aliasing effects 102 Alpha 247–248 animation scripts 249 Anova reports 154 assigning importance (of responses) 59, 97, 220 assigning importance of responses 75 attribute, discrete choice designs 165 augment nonlinear design table 211 Augmented Designs 223–243 add centerpoints 233, 235 add new runs and terms 233, 238 adding axial points 233, 237 disallow factor combinations 243 D-optimal 238 extending experiments 233 foldover 233, 236 modify D- or I- optimality criterion 242 replicate 233 saved desisgn (X) matrix 242 select number of random starts 243 special commands 241 specify sphere radius value 243 axial

points 115

scaling, central composite designs 125

В

balanced designs 64 Bayesian D-optimal designs 28, 53–54 Bayesian I-optimal designs 54 Big Class.jmp 46 block sizes 25 Borehole Factors.jmp 193 Borehole Sphere Packing.jmp 194 Bounce Data.jmp 117 Bounce Factors.jmp 117 Box-Behnken designs 115, 117, 125 *See also* Response Surface designs Byrne Taguchi Factors.jmp 216–217

С

categorical factors 62, 98 CCD See central composite designs center points augment designs 235-237 central composite designs 115 response surface designs 115 simplex centroid designs 142 central composite designs 115, 125 See also response surface designs centroid points 147 Chakravarty 101 Chemical Kinetics.jmp 201 chemical mixture, examples 155 choice designs example 163 purpose 163 choice set, discrete choice designs 165 coding, column property 83 coefficients, relative variance 68 column properties coding 83

constrained state 76 design roles 75, 86 factors for mixture experiments 84 high and low values 83 response limits 85 combinations, disallowing 80, 243 CONAEV method 151 confounding 103 resolution numbers 100 confounding pattern 69, 103 constraints adding 147 disallowing combinations 80, 243 entering 76 linear 150 loading 76 saving 76 continuous factors 62, 98 contour plots 155 control factors 213, 215 CONVRT method 151 coordinate exchange algorithms 88 Corn.jmp 208 Cotter designs 101 counts per unit (power and sample size) 258 covariate factors 46 creating factors tables 76 criterion, optimality 78, 242 custom design commands 74 custom designs data tables 71 Design Generation panel 7, 64 factors, defining 59 flexible block size 25 how they work 88 models, describing 63 random block design 71 screening 17 special commands 74 split plot 72 split-split plot 73 strip plot 74

D

defaults number of random starts 78 defects 258 D-efficiencies 39 describing models 63, 142 design

matrix table properties 78, 242 resolutions 99 roles 75 design diagnostics table 187 Design of Experiments. See DOE design roles 86 designs ABCD 151 aberration 100 augment 233 balanced 64 Bayesian D-optimal 28, 54 Bayesian I-optimal 54 Box-Behnken 115, 117, 125 central composite 115 discrete choice 163 foldover 236 fractional factorial 99 full factorial 89, 129, 136 full factorials 99 Gaussian process IMSE optimal space filling 192 I-optimal 54 Latin hypercube space filling design 184–186 maximum entropy space filling 190 minimum aberration 100 minimum potential space filling 189 mixed-level 101 mixture 139, 155 nonlinear 201 orthogonal screening designs 99 screening experiments 139 surface designs 125 orthogonal arrays 101 Plackett-Burman 100 random block 71 replicating 233-234 response surface 115 saturated 64 screening 89 simplex centroid 142 simplex centroids 142 simplex lattice 145 space filling ??–197 sphere packing space filling 181 split plot 72 split-split plot 73 strip plot 74 uniform precision 125 uniform space filling 186

desirability functions 120, 218 maximizing 120 values 60, 97 desirability traces 120 desirability values, as column properties 85 determinant formula 79 diagnostics for custom designs 77, 242 Difference to Detect option 248–249, 251 disallowing combinations 80, 243 discrete choice designs analysis 169, 173 attribute 165 attributes panel 167 choice set 165 definition 163 design generation panel 167 dialog 166 example 165–166, 168 prior information 171 profile 165 survey 166 DOE (Design of Experiments) coding 83 simple examples 91 DOE Sphere Radius 243 Donev Mixture Factors.jmp 40 D-optimal augmentation 238 designs 28, 53-54 optimality criteria 78, 242

Ε

effect aliasing 103 eigenvalue 119 eigenvector 119 nonestimable 99 orthogonal 125, 213 size 247 sparsity 89, 100-101 efficiencies D, G, and A 69 eigenvalue of effect 119 eigenvector of effect 119 equivalent solutions 88 error standard deviation 247-248 error variance 32 excluding factor combinations 80 extra parameters 248

extreme vertices 147 algorithms 151 finding optimal subsets 148 range constraints 148

F

factor combinations, disallowing 80 Factor Profiling option 119, 121, 135, 158, 218 factorial designs fractional 99 full 89, 99, 129, 136 three level 101 factors adding 37, 61 categorical 62, 98 column 86 continuous 62, 98 control factors 213, 215 covariate 46 for mixture experiments 84 key factors 89 loading 75 nonmixture 40 saving 75 tables, creating 76 false negatives 101 finding optimal subsets (extreme vertices) 148 fitting mixture designs 153 fixed covariate factors 46 fixed covariates example 46 folding the calculation 155 foldover designs 236 fraction of design space plot 67 fractional factorial designs 99 Full Factorial Designs 129–138 design generation panel 64 examples 131 full factorial designs 99, 129 functions, desirability 120, 218

G

Gaussian process IMSE optimal design 179, 192 global optimum 79 goal types 59, 97 matching targets 60, 98 minimizing and maximizing 60, 98 Group New Runs into Separates Block option 236

Η

hidden intercepts 154 hyperspheres 80, 243

identifying key factors 89 importance of responses 59, 75, 97, 220 entering 85 industrial experimention 163 inner arrays, inner designs 213, 215 **Inscribe** option 125 interactions 101 high-order 100 intercepts, hidden 154 I-optimal designs 53–54 optimality criteria 78, 242 I-optimal designs 54

J

JSL (JMP Scripting Language) animation scripts 249 augmented designs 228 sphere radius 243

Κ

Keep the Same 70 k-Sample Means (power and sample size) 252

L

L18 Chakravarty 101 L18 Hunter 101 L18 John 101 L36 101 L9-Taguchi 216 Label column 106, 127 larger-the-better formulas (LTB) 217 Latin hypercube space filling design 184–186 limits, responses 85 linear constraints 147, 150 loading constraints 76 factors 75 responses 75 local optimum 79

Μ

matching target goals 59–60, 85, 97–98 matrix, alias 69 matrix, design 78, 242 maximize responses 59, 97 maximizing desirability 120 goals 60, 97 maximum entropy space filling design 179, 190 means, one- and two-sample (power and sample size) 248 methods CONAEV 151 minimize responses 59, 97 minimizing goals 60, 98 minimum aberration designs 100 minimum potential design 189 minimum potential space filling design 179 mixed-level designs 101 Mixture Designs 139–161, 163 ABCD design 151 compared to screening designs 139 definition 139 examples 155-157 extreme vertices 147 fitting 153 linear constraints 150 optimal 141 optimal subsets 148 response surfaces 158 simplex centroids 142 simplex lattice 145 ternary plots 152 with nonmixture factors 40 mixture designs examples 155 mixture experiments (column property) 84 mixture, column properties 84 Model script Model Specification dialog 109, 118 models custom design 63 describing 63, 142

Ν

N factors, adding 61 N Monte Carlo Spheres (nonlinear design option) 212 N Responses 59 N Responses button 97 no-intercept model 84 noise factors 213, 215 nonestimable effect 99 nonlinear design advanced options 211 augment table 204, 211 create using prior parameter estimates 203 creating 208 creating with no prior data 205 distribution of parameters 206 fitting to find prior parameter estimates 201 introduction 199 launch dialog 203, 208–209 number of Monte Carlo samples 204 number of starts 204 overlay plot of design points 207 using prior parameter estimates 203 Nonlinear Designs 199–212 advanced options 211 creating 208 examples 201 nonmixture factors in mixture designs 40 Number of Monte Carlo Samples (nonlinear design option) 212 number of runs 15, 64 screening designs 99 number of starts 78, 242

0

On Face option 125 one-sample and two-sample means (power and sample size) 248 one-sample and two-sample proportion (power and sample size) 254 one-sample variance (power and sample size) 253 optimal determinant formula 79 optimal subsets (mixture designs) 148 Optimality Criterion 78, 242 optimality criterion 78 order of runs 70, 106, 126, 137, 143, 148, 151 orthogonal array designs 101, 213 orthogonal designs screening designs 99 screening experiments 139 surface designs 125 Orthogonal option 125 outer arrays, outer designs 213, 215

Ρ

parameters, extra 248 Pattern column 93, 106, 127, 138 patterns, confounding 103 performance statistics 215 Placket-Burman design creating 106 Plackett-Burman designs 100 Plastifactors.jmp 155 plots Actual-by-Predicted III Scatterplot 3D 122 ternary 139, 152 points axial 115 center See center points centroid 147 Poisson-distributed counts 258 potential terms (DOE) 28 power and sample size calculations 245–260 analyses 247 animation 250 counts per unit 258 in statistical tests on means 247 k-sample means 252 one-sample and two sample proportions 254 one-sample mean 248–249 one-sample variance 253, 255-256 Sigma quality level 259 two-sample means 251 prediction profilers 32, 35 traces 120 variances 32, 125 prediction formulas, saving 157 prediction profiler 119 prediction variance profile plot 31 prediction variance surface plot 67 primary terms (Bayesian D-optimal design) 28 profile, discrete choice designs 165 profilers mixture response surface 158 prediction profilers 35, 119 prediction variance plot 31 prediction Variance profiler 32 properties, columns 83 proportions (power and sample size) 254 prospective power and sample size 245 prospective power and sample size calculations (see power and sample size) 260

pseudocomponent (mixture column property) 84 pure error 235

Q

quadratic model 31-69

R

radius, sphere 80, 243 random block designs 71 random seed, setting 77 random starts 78 Randomize within Blocks 70 randomizing runs 70 range constraints 147 Reaction Kinetics Start.jmp 205 Reaction Kinetics.jmp 207 Reactor 8 Runs.jmp 225 Reactor Factors.jmp 131–132 Reactor Response.jmp 131 regressor columns 101 relative proportions See mixture designs relative variance of coefficients 68 relative variance of coefficients and power table 28 replicating designs 234 replicating designs 233 reponses loading 75 requesting additional runs 71, 106, 126, 137 rescaling designs 125 resolution numbers 100 resolutions of designs 99 response limits, column property 85 Response Surface Designs 115-127 examples 117-122 introduction 123 purpose 115 reports 118 with blocking factors 36 with categorical factors 36 response surface effects 155 Response Surface Methodology (RSM) 123 response surfaces mixture designs 158 responses custom designs 59, 97, 124 desirability values 60, 97

goals 59, 97 goals, desirability functions 120 lower limits 60, 97 saving 75 upper limits 60, 97 responses, multiple 85 RMSE 111, 134 robust engineering 213 roles, design 86 Rotatable option 125 RSM (Response Surface Methodology) 123 runs additional 71, 106, 126, 137 order they appear in table 70, 106, 126, 137, 143, 148, 151 requesting additional 71, 106, 126, 137 screening designs 99

S

sample means (power and sample size) 248 sample sizes example comparing single-direction one-sample variances 253, 255 example with counts per unit 258 one and two sample means 249 prospective power analysis 247 screening designs 129 saturated designs 64 saving constraints 76 factors 75 prediction formulas 157 responses 75 X Matrix 78, 242 scaling axial 125 designs 125 Scatterplot 3D Box-Behnken designs 122 Scheffé polynomial 153 screening designs 89 custom designs 17 design types 99 dialogs 94, 106 examples 93-126 scripts animation 249 generating the analysis model 109, 118 Model script *See* Model table property scripting See JSL

Set Random Seed command 77 sigma quality level (power and sample size) 259 signal factors 213 signal-to-noise ratios 213 simplex 139 centroid designs 142 lattice designs 145 single-sample means (power and sample size) 249 solutions, equivalent 88 Sort Left to Right options 70 Sort Right to Left option 70 Space Filling Designs ??–197 borehole problem, sphere packing example 193 Gaussian process IMSE optimal 179, 192 Latin hypercube 179, 184–186 maximum entropy 179, 190 minimum potential 179, 189 sphere packing 179, 181 uniform 179, 186 sparsity, effect 89, 100-101 sphere packing design 179, 181 sphere radius 80, 243 split plot designs 62, 72 split-split plot designs 73 standard deviation, error 247 star points 115 starts, number of 78, 242 starts, random 78 statistics, performance 215 strip plot designs 74 subsets, finding optimal 148 supersaturated designs 21 surface designs See response surface designs survey, discrete choice designs 166

Т

tables factors table, creating 76 making in custom designs 71 Taguchi designs 213–219 description 213 methods 213 target values 60, 85, 98 ternary plots 139, 152, 155 three-dimensional scatterplot Box-Behnken designs 122 traces, desirability 120 trade-off in screening designs 99 tutorial examples augment designs 225-233 custom design 1–13 DOE 91 flexible block size 25 full factorial designs 131 mixture designs 155–157 response surface designs 117, 121-122 screening designs 93 two-level categorical 92 two-level fractional factorials 99 two-level full factorials 99 two-sample and one-sample means 251 two-sample means (power and sample size) 248 two-sample proportion (power and sample size) 254

U

Uniform (space filling design) 179 uniform (space filling design) 186 uniform precision designs 125 **User Defined** option 125

V

values high and low for columns (in DOE) 83 target 60, 98 variance of prediction 125 variance, error and prediction 32 vertices extreme 147 extreme, finding optimal subsets 148

W-Z

weight, importance of response 59, 75, 86, 97, 220 whole model test 154 X Matrix, saving 78, 242 XVERT method 151 Index