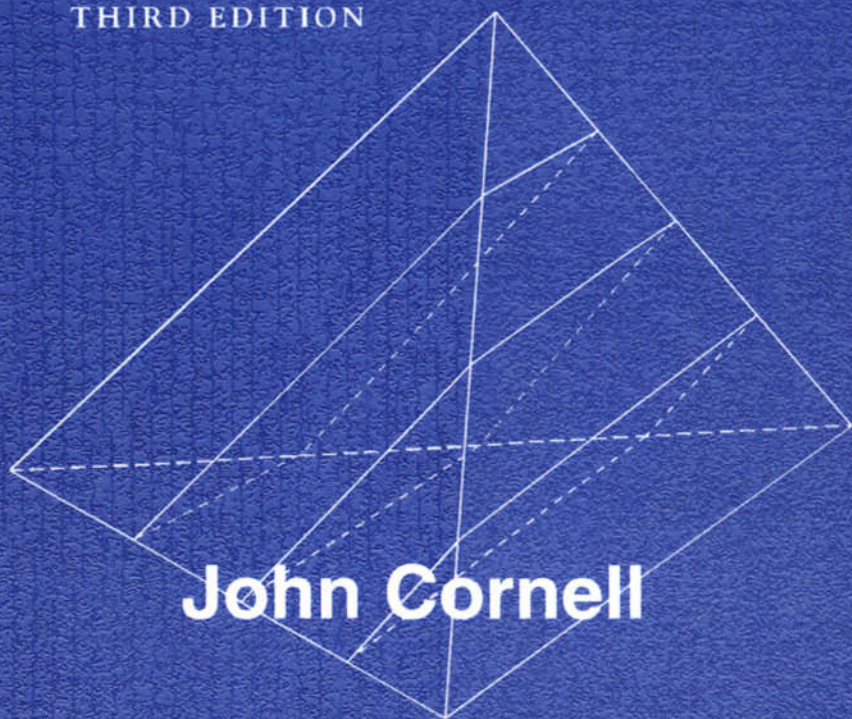


# Experiments with Mixtures

Designs, Models, and the  
Analysis of Mixture Data

THIRD EDITION



**John Cornell**

This page intentionally left blank

## Experiments with Mixtures

**WILEY SERIES IN PROBABILITY AND STATISTICS**

Established by **WALTER A. SHEWHART** and **SAMUEL S. WILKS**

Editors: *Peter Bloomfield, Noel A. C. Cressie, Nicholas I. Fisher,  
Iain M. Johnstone, J. B. Kadane, Louise M. Ryan, David W. Scott,  
Bernard W. Silverman, Adrian F. M. Smith, Jozef L. Teugels;*  
Editors Emeriti: *Vic Barnett, J. Stuart Hunter, David G. Kendall*

A complete list of the titles in this series appears at the end of this volume.

# Experiments with Mixtures

## Designs, Models, and the Analysis of Mixture Data

Third Edition

JOHN A. CORNELL  
University of Florida



A Wiley-Interscience Publication  
JOHN WILEY & SONS, INC.

This book is printed on acid-free paper. ⊗

Copyright © 2002 by John Wiley & Sons, Inc., New York. All rights reserved.

Published simultaneously in Canada.

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, recording, scanning or otherwise, except as permitted under Sections 107 or 108 of the 1976 United States Copyright Act, without either the prior written permission of the Publisher, or authorization through payment of the appropriate per-copy fee to the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923, (978) 750-8400, fax (978) 750-4744. Requests to the Publisher for permission should be addressed to the Permissions Department, John Wiley & Sons, Inc., 605 Third Avenue, New York, NY 10158-0012, (212) 850-6011, fax (212) 850-6008, E-Mail: PERMREQ@WILEY.COM.

For ordering and customer service, call 1-800-CALL-WILEY.

***Library of Congress Cataloging-in-Publication Data:***

Cornell, John A., 1941-

Experiments with mixtures : designs, models, and the analysis of mixture data / John A.

Cornell. -- 3rd ed.

p. cm. -- (Wiley series in probability and statistics. Applied probability and statistics section)

"A Wiley-Interscience publication."

Includes bibliographical references and index.

ISBN 0-471-39367-3 (cloth : alk. paper)

1. Experimental design. 2. Mixtures--Statistical methods. I. Title. II. Wiley series in probability and statistics. Applied probability and statistics.

QA279 .C66 2002

519.5'38--dc21

2001045457

10 9 8 7 6 5 4 3

# Contents

<b>Preface to the Third Edition</b>	<b>xiii</b>
<b>Preface to the Second Edition</b>	<b>xvii</b>
<b>1. Introduction</b>	<b>1</b>
1.1. The Original Mixture Problem, 2	
1.2. General Remarks About Response Surface Methods, 8	
1.3. A Factorial Experiment or a Mixture Experiment?, 12	
1.4. An Historical Perspective, 18	
References and Recommended Reading, 20	
Questions, 20	
<b>2. The Original Mixture Problem: Designs and Models for Exploring the Entire Simplex Factor Space</b>	<b>22</b>
2.1. The Simplex-Lattice Designs, 22	
2.2. The Canonical Polynomials, 25	
2.3. The Polynomial Coefficients as Functions of the Responses at the Points of the Lattices, 30	
2.4. Estimating the Parameters in the $\{q, m\}$ Polynomials, 34	
2.5. Properties of the Estimate of the Response, $\hat{y}(\mathbf{x})$ , 36	
2.6. A Three-Component Yarn Example Using a $\{3, 2\}$ Simplex-Lattice Design, 37	
2.7. The Analysis of Variance Table, 41	
2.8. Analysis of Variance Calculations of the Yarn Elongation Data, 44	
2.9. The Plotting of Individual Residuals, 46	

- 2.10. Testing the Degree of the Fitted Model: A Quadratic Model or Planar Model?, 47
  - 2.11. Some Comments on the Use of Check Points for Testing Model Lack of Fit, 52
  - 2.12. A Numerical Example Illustrating the Use of Check Points for Testing Lack of Fit, 57
  - 2.13. The Simplex-Centroid Design and the Associated Polynomial Model, 60
  - 2.14. An Application of a Four-Component Simplex-Centroid Design. Blending Chemical Pesticides for Control of Mites, 63
  - 2.15. Axial Designs, 66
  - 2.16. Comments on a Comparison Made Between an Augmented Simplex-Centroid Design and a Full Cubic Lattice for Three Components Where Each Design Contains Ten Points, 69
  - 2.17. Reparameterizing Scheffé's Mixture Models to Contain a Constant ( $\beta_0$ ) Term: A Numerical Example, 72
  - 2.18. Questions to Consider at the Planning Stages of a Mixture Experiment, 80
  - 2.19. Summary, 81  
References and Recommended Reading, 82  
Questions, 83
  - Appendix 2A. Least-Squares Estimation Formulas for the Polynomial Coefficients and Their Variances: Matrix Notation, 88
  - Appendix 2B. Cubic and Quartic Polynomials and Formulas for the Estimates of the Coefficients, 92
  - Appendix 2C. The Partitioning of the Sources in the Analysis of Variance Table When Fitting the Scheffé Mixture Models, 94
- 3. The Use of Independent Variables 96**
- 3.1. Transforming from the  $q$  Mixture Components to  $q-1$  Mathematically Independent Variables, 97
  - 3.2. A Numerical Example: Sensory Flavor Rating of Fish Patties, 100
  - 3.3. Defining a Region of Interest Inside the Simplex: An Ellipsoidal Region, 107
  - 3.4. A Numerical Illustration of the Inverse Transformation from the Design Variables to the Mixture Components, 111



- 3.5. Enlarging the Unit Spherical Region of Interest, 113
- 3.6. Some Discussion on Design Strategy When Fitting Response Surfaces, 115
- 3.7. Rotatable Designs, 116
- 3.8. A Second-Order Rotatable Design for a Four-Component System, 119
- 3.9. Defining a Cuboidal Region of Interest in the Mixture System, 122
- 3.10. Summary, 124
  - References and Recommended Reading, 125
  - Questions, 125
- Appendix 3A. An Alternative Transformation from the Mixture Component System to the Independent Variable System, 127
- Appendix 3B. A Form of the Orthogonal Matrix  $T$ , 129

#### **4. Multiple Constraints on the Component Proportions**

132

- 4.1. Lower-Bound Restrictions on Some or All of the Component Proportions, 132
- 4.2. Introducing  $L$ -Pseudocomponents, 133
- 4.3. A Numerical Example of Fitting an  $L$ -Pseudocomponent Model, 136
- 4.4. Upper-Bound Restrictions on Some or All of the Component Proportions, 138
- 4.5. An Example of the Placing of an Upper Bound on a Single Component: The Formulation of a Tropical Beverage, 140
- 4.6. Introducing  $U$ -Pseudocomponents, 144
- 4.7. The Placing of Both Upper and Lower Bounds on the Component Proportions, 149
- 4.8. Formulas for Enumerating the Number of Extreme Vertices, Edges, and Two-Dimensional Faces of the Constrained Region, 156
- 4.9. Some Procedures for Calculating the Coordinates of the Extreme Vertices of a Constrained Region, 160
- 4.10. Multicomponent Constraints, 180
- 4.11. Some Examples of Designs for Constrained Mixture Regions: CONVRT and CONAEV Programs, 182
- 4.12. The Use of Symmetric-Simplex Designs for Fitting Second-Order Models in Constrained Regions, 188

- 4.13. Multiple Lattices for Major and Minor Component Classifications, 191
- 4.14. Categorizing the Mixture Components: An Ellipsoidal Region of Interest, 208
- 4.15. A Numerical Example of a Categorized Component Experiment, 212
- 4.16. Summary, 215
  - References and Recommended Reading, 216
  - Questions, 218
- Appendix 4A. An Orthogonal Matrix for the Categorized-Components Problem, 220
- Appendix 4B. The Relationship Between the Coefficients of the Terms in the Double-Scheffé Model (4.82) and the Interaction Model (4.83), 222

## 5. The Analysis of Mixture Data

223

- 5.1. Techniques Used in the Analysis of Mixture Data, 224
- 5.2. Test Statistics for Testing the Usefulness of the Terms in the Scheffé Polynomials, 227
- 5.3. Model Reduction, 233
- 5.4. An Example of Reducing the System from Three to Two Components, 236
- 5.5. A Criterion for Selecting Subsets of the Terms in the Scheffé Models, 238
- 5.6. A Numerical Example Illustrating the Integrated Mean-Square Error Criterion, 241
- 5.7. Screening Components, 244
- 5.8. A Seven-Component Octane-Blending Experiment: An Exercise in Model Reduction, 248
- 5.9. Other Techniques Used to Measure Component Effects, 254
- 5.10. The Slope of the Response Surface Along the Component Axes, 265
- 5.11. A Numerical Example Illustrating the Slope Calculations for a Three-Component System: Studying the Flavor Surface Where Peanut Meal Is Considered a Substitute for Beef in Sandwich Patties, 268
- 5.12. Leverage and the Hat Matrix, 271
- 5.13. A Three-Component Propellant Example, 273
- 5.14. Summary, 276
  - References and Recommended Reading, 277

Questions, 278

Appendix 5A. The Derivation of the Moments of the Simplex Region, 284

**6. Other Mixture Model Forms**

**286**

- 6.1. The Inclusion of Inverse Terms in the Scheffé Polynomials, 286
- 6.2. Fitting Gasoline Octane Numbers Using Inverse Terms in the Model, 289
- 6.3. An Alternative Model Form for Modeling the Additive Blending Effect of One Component in a Multicomponent System, 291
- 6.4. A Biological Example on the Linear Effect of a Powder Pesticide in Combination with Two Liquid Pesticides Used for Suppressing Mite Population Numbers, 296
- 6.5. Other Models That Are Homogeneous of Degree One, 300
- 6.6. The Use of Ratios of Components, 305
- 6.7. Cox's Mixture Polynomials: Measuring Component Effects, 309
- 6.8. An Example Illustrating the Fits of Cox's Model and Scheffé's Polynomial, 314
- 6.9. Log Contrast Models, 320
- 6.10. A Numerical Example Illustrating the Testing of Inactivity and Additivity Effects of the Components in a Three-Component System Using Log Contrast Models, 323
- 6.11. Octane Blending Models, 326
- 6.12. A Numerical Example Illustrating the Calculations Required for Obtaining the Research and Motor Octane Prediction Equations for a Group of Blends, 329
- 6.13. Fitting a Slack-Variable Model, 333
- 6.14. A Numerical Example Illustrating the Fits of Different Reduced Slack-Variable Models: Tint Strength of a House Paint, 337
- 6.15. Summary, 343
  - References and Recommended Reading, 344
  - Questions, 346
- Appendix 6A. The Form of the Multiplier Matrix  $B_2$  for Expressing the Parameters in Cox's Quadratic Model as Functions of the Parameters in Scheffé's Model, 349
- Appendix 6B. Estimation Equations for Coefficients That Are Subject to Linear Restrictions, 352

<b>7. The Inclusion of Process Variables in Mixture Experiments</b>	<b>354</b>
7.1. Designs Consisting of Simplex-Lattices and Factorial Arrangements, 354	
7.2. A Numerical Example of a Fish Patty Experiment: Studying Blends of Three Fish Species Prepared with Three Processing Factors, 361	
7.3. Testing the Component Blending Properties and the Effects of the Process Variables When the Set of Mixture Blends Is Embedded in the Processing Conditions, 365	
7.4. A Numerical Example of a Three-Component by Two-Process Variable Split-Plot Experiment: Fitting a Quadratic Mixture Model in the Presence of Interacting Process Variables, 376	
7.5. A Reparameterization of the Combined Model Form for Measuring the Effects of the Process Variables: An Example of Model Reduction, 382	
7.6. The Use of Fractional Designs in the Process Variables, 387	
7.7. A Numerical Example of the Fit of a Combined Model to Data Collected on Fractions of the Fish Patty Experimental Design, 397	
7.8. Computer-Aided Fractionation of Lattice Designs, 400	
7.9. Mixture-Amount Experiments, 403	
7.10. Process Variables and $q-1$ Mixture-Related Variables, 419	
7.11. A Numerical Example Involving Three Mixture Components and One Process Variable, 420	
7.12. Questions Raised and Recommendations Made When Fitting a Combined Model Containing Mixture Components and Other Variables, 424	
7.13. Summary, 429	
References and Recommended Reading, 431	
Questions, 432	
Appendix 7A. A Generalized Least-Squares Solution for Fitting the Mixed Model in the Mixture Components and Process Variables to Data from a Split-Plot Experiment, 435	

**8. Additional Topics****438**

- 8.1. Block Designs for Mixture Experiments, 438
- 8.2. Symmetric-Simplex Block Designs for Fitting the Scheffé Second-Order Model, 441
- 8.3. An Example of Orthogonal Blocking Using a Symmetric-Simplex Design, 444
- 8.4. Constructing Orthogonal Blocks Using Latin Squares, 447
- 8.5. Weighted Versus Unweighted Least-Squares Estimates of the Parameters in the Scheffé Models, 454
- 8.6. Some Comments on Design Criteria and Some Results Using the ACED Program, 460
- 8.7. Constant Prediction Variance on Concentric Triangles for Three-Component Systems, 469
- 8.8. Altering the Terms in the Scheffé-Type Models to Improve the Accuracy and/or Stability of the Coefficient Estimates, 476
- 8.9. Collinearity Problems Resulting from Performing Experiments in Highly Constrained Regions, 482
- 8.10. A Numerical Example Illustrating the Fitting of Segmented Scheffé Models to Freezing-Point Data from a Two-Component System, 489
- 8.11. Biplot Displays for Multiple Responses, 496
- 8.12. A Five-Response Plastics-Compounding Example, 500
- 8.13. Optimizing Several Responses Simultaneously, 504
- 8.14. Recalling the Three-Component Propellant Example of Section 5.13, 507
- 8.15. Summary, 510  
References and Recommended Reading, 512  
Questions, 515
- Appendix 8A. The Modified *L*-Pseudocomponent Model and the Centered and Scaled Intercept Model, 516
- Appendix 8B. Expressing the Coefficients in the Scheffé Quadratic Model As Functions of the Coefficients in the *L*-Pseudocomponent, Modified *L*-Pseudocomponent, and Centered and Scaled Intercept Models, 518

<b>9. Matrix Algebra, Least Squares, and the Analysis of Variance</b>	<b>521</b>
9.1. Matrix Algebra, 521	
9.2. Some Fundamental Definitions, 522	
9.3. A Review of Least Squares, 524	
9.4. The Analysis of Variance, 527	
9.5. A Numerical Example: Modeling the Texture of Fish Patties, 528	
9.6. The Adjusted Multiple Correlation Coefficient, 532	
9.7. The Press Statistic and Studentized Residuals, 532	
9.8. Testing Hypotheses About the Form of the Model: Tests of Significance, 534	
References and Recommended Reading, 536	
<b>10. Data Sets from Mixture Experiments with Partial Solutions</b>	<b>538</b>
10.1. Experiment One: Fruit Punch Experiment, 538	
10.2. Experiment Two: Chick Feeding Experiment, 543	
10.3. Experiment Three: Concrete Batches, 550	
10.4. Experiment Four: Surface Resistivity of Paper Coatings, 557	
10.5. Experiments Five, Six, and Seven: Estimating Solubilities of Multisolvent Systems, 559	
References and Recommended Reading, 568	
Appendix 10A. CONVRT Program Listing for Calculating the Coordinates of the Extreme Vertices of a Constrained Region, 569	
Appendix 10B. CONAEV Program Listing for Calculating the Coordinates of the Centroids (Approximate) of the Boundaries of a Constrained Region, 576	
Appendix 10C. Listings of Subroutines Called by CONVRT and CONAEV, 581	
<b>Bibliography and Index of Authors</b>	<b>589</b>
<b>Answers to Selected Questions</b>	<b>604</b>
<b>Appendix</b>	<b>637</b>
<b>Index</b>	<b>643</b>

## Preface to the Third Edition

When Wiley asked if I would consider putting together a third edition, I admit to having mixed feelings. Part of me wondered, “Is it really necessary to add more material to an already large collection of tools?” During the past ten years presenting short courses on mixture designs and mixture data analysis, I would generally cover only the first seven chapters of the book over a  $2\frac{1}{2}$  to 3-day period of time. Yet another part of me responded, “I believe there has been quite a lot of information on mixture experiments written since the printing of the second edition 11 years ago. And then it hit me: “If I dedicate myself to improving the current product (second edition) by offering additional features (such as designs and techniques of data analysis) that the competition (other texts that discuss mixture experiments in one or two chapters at most) will not match, isn’t this the philosophy behind what we mean when talking about product improvement?”

In this third edition, little new material has been added to the first four chapters. In Chapter 1, the chronological ordering of the statistical literature on mixtures has been extended to the new millennium spanning a 47-year period of time since the first mention of mixture experiments in 1953. In Chapter 1 mention is made of a strategy that does not work, which is to set up the mixture experiment in only  $q-1$  of the  $q$  components and proceed as if the  $q-1$  components were the only ones that are part of the mixture. Chapter 2 on the original mixture problem received many simple insertions throughout the text along with a new Figure 2.6 and a new Table 2.9 showing the output of DESIGN-EXPERT. A new Appendix 2C discusses how the partitioning of the sources of variation in the analysis of variance table takes form when fitting the Scheffé-type mixture models. Chapter 3 on the use of independent variables is the same as in the second edition while Chapter 4 on multiple constraints on the component proportions received a new section entitled, “Allowing the Major Component Proportions to Vary: Mixtures of Mixtures,” featuring the inter- and intra-category blending of resins in a photoresist formulation discussed by Cornell and Ramsey (1998). A second appendix also has been added to Chapter 4.

Chapter 5 has been expanded by showing the type of computer software output that has become available. Output in the form of the listing of residuals, studentized residuals, Cook's distance measure, and outlier  $t$  values are just some of the useful model-diagnostic measures offered either by DESIGN-EXPERT, Minitab, and/or other software packages. Two new sections, "Leverage and the Hat Matrix" and "A Three-component Propellant Example," finish up Chapter 5. Added to Chapter 6 is a section entitled "Fitting a Slack-Variable Model" along with a numerical example illustrating the fits of different reduced slack-variable models and a warning about fitting a slack-variable model.

Chapter 7 has been expanded by the addition of a section: "Questions Raised and Recommendations Made When Fitting a Combined Model Containing Mixture Components and Other Variables." An appendix has been added that discusses the generalized least-squares solution for fitting the mixed model in the mixture components and process variables to data from a split-plot experiment. This type of combined mixture component-process variable experiment has gained considerable attention in the engineering and statistical communities since the publication of the second edition.

Chapter 8 has received the most attention in this third edition, as one might expect, by the addition of four sections. Biplot displays for looking at multiple response data are discussed. Biplots are so named because both mixture formulation and multiple response information are displayed in a single plot. A five-response, plastics-compounding example is presented, illustrating the ease in applying this graphical technique. An attempt at multiple response optimization through the use of the desirability function or by the overlaying of contour plots is the topic of the final two sections of Chapter 8. A modified  $L$ -pseudocomponent transformation/model and a centered and scaled intercept model now make up a new appendix. A section on the press statistic and on studentized residuals has been added to Chapter 9, while Chapter 10 remains the same as in the second edition. An updated Bibliography now contains more than 200 listings taken from the mixtures and related literature.

Much of what was written in the preface to the second edition about its coverage of mixture designs and models holds true for this third edition. Chapters 2 and 4 remain the lengthiest in terms of number of pages and are the most important. As a classroom textbook for a one-semester graduate-level course, I would suggest Chapters 1-7 as basic material and then select from Chapters 8-10 according to need. If used as a reference or for self-study, Chapters 1 and 2 and 4-7 provide the necessary tools for dealing with almost any type of mixture problem.

In putting together the material for this third edition, I am indebted to many people, some of whom I do not know but who offered suggestions on a proposal I submitted on specific topics to be included in the third edition. Friends who are retired from industry such as John Gorman, formerly of AMOCO, Wendell F. Smith, Jr., formerly of Eastman Kodak Company, and Gregory Piepel, presently with Battelle Pacific Northwest Laboratories, will forever remain part of what I write about mixture experiments because of the discussions we've had and the input they've provided as well as the inspiration they've had on me over the years. Others



who attended the sessions at national meetings and provided questions that fueled my research efforts, to you I extend a heartfelt “thank you.”

**JOHN A. CORNELL**

*Gainesville, Florida  
October 2001*

This page intentionally left blank

# Preface to the Second Edition

Since the publication of the first edition 9 years ago, approximately 35 papers have appeared in statistical literature that dealt with the construction of mixture designs and/or the discussion of methods for analyzing mixture data. Many of the paper topics had foundations set previously such as improved ways (algorithms) to generate designs for constrained mixture regions, alternative methods for measuring component effects in highly constrained mixture problems, and new designs and models for including process variables and/or the total amount of the mixture in these types of experiments.

As an active reviewer of papers as well as being an active participant in onsite seminars for various industries and short courses sponsored by the American Society for Quality Control, I have become increasingly aware of the types of mixture experiments that dominate the different areas of application, particularly industrial applications. Furthermore, during the past 22 years I have had the good fortune of working in an academic environment which has afforded the opportunity to fuel my research in a manner that blends theory with practice. As a result, this second edition contains a considerable amount of additional material, both new and not so new.

A new section in Chapter 1 compares a factorial experiment to a mixture experiment. Chapter 2 discusses the method of using check points for testing the adequacy of the fitted model, along with a numerical example which illustrates the technique. Rewriting the Scheffé-type mixture models by deleting one of the linear blending terms and inserting a constant term so that the model form conforms to the requirements of some of the more popular least-squares model fitting programs is also included. Chapter 2 ends with questions that one might ask while planning to conduct a mixture experiment.

Chapter 3 once again addresses the use of independent variables with only slight changes. Additional comments regard the settings of the mixture components upon transforming from the design settings in the independent variables. A new formula is given for calculating the radius of the largest sphere, centered at the point of main

interest, that will fit inside the simplex region. The four sections on the inclusion of process variables have been removed from Chapter 3 and placed in Chapter 7, which is now a chapter devoted entirely to the inclusion of process variables in mixture experiments.

The assigning of multiple constraints on the component proportions, Chapter 4, has experienced the greatest revision. Previously, only lower-bound pseudocomponents, now called *L*-pseudocomponents, were considered. In this edition, upper-bound pseudocomponents, called *U*-pseudocomponents, are introduced and used for detecting inconsistent constraints. It is shown how the set of inconsistent constraints can be adjusted to make the set of lower and upper bounds consistent. A formula is given for enumerating the number of extreme vertices, edges, faces, and so on of a constrained region based on the set of consistent constraints. This then leads to calculating the coordinates of the extreme vertices of the constrained region using *U*-pseudocomponents. A section on design strategy for fitting the Scheffé quadratic model over a constrained region has been added along with a section containing several examples of constrained mixture experiments. This section on examples of constrained mixture experiments aids users in checking the results obtained with their software packages, against those obtained by this author using the two programs, CONVRT and CONAEV, whose steps are now listed in Appendices A, B, and C of Chapter 10.

The analysis of mixture data, Chapter 5, now includes the plotting of the response trace for measuring changes in the response brought about by changing the proportion of a single component at a time. Chapter 6, on other mixture model forms, contains a new section on measuring additivity and interaction in the component blending properties by fitting log contrast models.

Chapters 7, 8, and 10 are new additions. Chapter 7 is devoted entirely to the inclusion of process variables and considers the total amount of the mixture as well. Combining lattice designs with factorial arrangements is taken from Chapters 3 and 5 of the first edition and is illustrated using the fish patty experiment previously found in Chapter 5. Testing the component blending properties and the effects of the process variables when one set of variables is embedded in the other set (mixture components embedded in the process variables and vice versa) is also illustrated. The use of fractional factorial designs in the process variables as well as fractionating the lattice designs in the mixture components through a computer-aided approach are two new sections. Designs and models for mixture-amount experiments close out Chapter 7.

Chapter 8 reviews orthogonal blocking strategies, comparing the estimates of the coefficients in the Scheffé models obtained using weighted and unweighted least-squares formulas, the generation of optimal designs with the ACED algorithm, and a technique that provides a prediction equation that possesses constant variance of prediction on concentric triangles for three-component systems. Reparameterizing the Scheffé-mixture models to models containing a constant term so that the terms can be centered and standardized is suggested as a remedy for improving the accuracy of the calculating formulas for the model coefficient estimates. Collinearity problems that arise from the fitting of Scheffé's models to data from

highly constrained regions are also addressed. Chapter 8 ends with a method for fitting segmented Scheffé models to freezing-point data that is collected from a binary system and exhibits a eutectic point. The method is illustrated with a numerical example.

Chapter 9 contains a review of matrix algebra, the method of least squares, and setting up an analysis of variance table. Chapter 10 is a collection of real data sets with partial solutions provided. The example data sets offer the reader an opportunity to work on problems with data sets that are larger in size than those provided in the exercises at the end of Chapters 2–8. An updated Bibliography contains more than 150 entries taken from the mixture literature and also recent related literature.

This expanded second edition is much more complete in its coverage of mixture designs and models particularly in the area of constrained mixture regions. Many of the new topics covered, particularly in Chapters 2 and 4, are important because most arose from questions that were asked during short-course discussions and plant visits. And while Chapters 2 and 4 are the lengthiest in terms of number of pages, in my opinion, they are the most important because they present designs and model fitting exercises for exploring the mixture surface *over the entire simplex region* and *over constrained subregions of the simplex*, respectively. As a classroom textbook for a one-semester graduate-level course, I would suggest Chapters 1–7 as basic material and then select from Chapters 8–10 according to need. If used as a reference or for self-study, Chapters 1–2 and 4–7 provide the necessary tools for dealing with almost any type of mixture problem.

In putting together the material for this second edition, I am indebted to many people. Heading the list of friends from industry are John Gorman who is now retired from AMOCO, Wendell Smith of Eastman Kodak Company, and Gregory Piepel of Battelle Pacific Northwest Laboratories. Others who attended the ASQC short courses on mixtures or who invited me to their particular companies to share a common interest in solving problems with mixtures are too numerous to mention here and to them I extend a heart warmed “thank you.” I wish also to express my sincere appreciation to Ms. Pamela Somerville for her excellent typing.

JOHN A. CORNELL

Gainesville, Florida  
April 1990

This page intentionally left blank

## Experiments with Mixtures

This page intentionally left blank



## CHAPTER 1

# Introduction

Many products are formed by mixing together two or more ingredients. Some examples are:

1. *Cake formulations* using baking powder, shortening, flour, sugar, and water.
2. *Building construction concrete* formed by mixing sand, water, and one or more types of cement.
3. *Railroad flares*, which are the product of blending together proportions of magnesium, sodium nitrate, strontium nitrate, and binder.
4. *Fruit punch* consisting of juices from watermelon, pineapple, and orange.
5. *Photographic film* coating made by blending silver halide, a coupler, two coupler solvents, and two stabilizers.
6. *Tobacco blends* consisting of flue-cured tobacco, burley, turkish blend, and processed tobacco.

In each of cases 1–6, one or more properties of each product are generally of interest to the manufacturer or experimenter who is responsible for mixing the ingredients. Such properties are (1) the fluffiness of the cake or the layer appearance of the cake where the fluffiness or layer appearance is related to the ingredient proportions; (2) the hardness or compression strength [measured in pounds per square inch (psi)] of the concrete, where the hardness is a function of the percentages of cement, sand, and water in the mix; (3) the illumination in foot-candles and the duration of the illumination of the flares; (4) the fruitiness flavor of the punch, which depends on the percentages of watermelon, pineapple, and orange that are present in the punch; (5) the color stability of the photographs when exposed to heat and light; and (6) the flavor and aroma of the tobacco blend. In every case, the measured property of the final product depends on the percentages or proportions of the individual ingredients that are present in the formulation.

Another reason for mixing together ingredients in blending experiments is to see whether there exist blends of two or more ingredients that produce more desirable product properties than are obtainable with the single ingredients

individually. For example, let us imagine we have three different gasoline stocks, labeled  $A$ ,  $B$ , and  $C$ , and that we are interested in comparing the antiknock quality of the three stocks, singly and in combination. In particular we would like to know whether there are combinations of the stocks, such as a 50% : 50% blend of  $A : B$ , or a 33% : 33% : 33% blend of  $A : B : C$ , or a 25% : 75% blend of  $B : C$ , that yield a higher antiknock rating than is obtained from using  $A$  alone or from using  $B$  alone, or  $C$  alone. If so, we would no doubt select the particular blend of two or more gasoline stocks that produces the highest rating, assuming of course that all other variables such as the cost and availability of the blending ingredients remain fixed.

In each of cases 1–6 listed above, it is assumed that the properties of interest are functionally related to the product composition and that, by varying the composition through the changing of ingredient proportions, the properties of the product will vary or change also. From an experimental standpoint, often the reason for studying the functional relationship between the measured property or the measured response (such as the strength of the concrete) and the controllable variables (which in the concrete case are the proportions of the ingredients cement, sand, and water) is (i) to determine whether some combination of the ingredients can be considered best in some sense, or (ii) to simply gain a better understanding of the overall system by studying the roles played by the different ingredients. The best ingredient combination for the concrete would be the combination that produced the absolutely hardest concrete without incurring an increase in the cost of the concrete batch. In an attempt to determine the best combination of ingredients (or combinations if more than one blend produces concrete samples having approximately equally high strengths), often one resorts to trial and error. Other attempts resemble “scattergun” procedures, where a large number of combinations of the ingredients are tried. Such scattergun procedures can require large expenditures in terms of time and cost of experimentation and in most cases better methods can be employed. Procedures used in screening out unimportant mixture ingredients are discussed in Section 5.7. Before we discuss some methods that have been developed for studying functional relationships and that are referred to as response surface methods, we introduce the original mixture problem.

## 1.1. THE ORIGINAL MIXTURE PROBLEM

To formulate our thinking about experiments involving mixtures, we simplify the gasoline-blending example mentioned earlier by considering only two gasoline stocks, which we label fuels  $A$  and  $B$ . Instead of discussing the antiknock rating, let us assume that the response of interest is the mileage obtained by driving a test car with the fuel, where the mileage is recorded in units of the *average number of miles traveled per gallon*. It is known ahead of time that fuel  $A$  normally yields 13 miles per gallon and fuel  $B$  normally yields only 7 miles per gallon. If the car is tested with each fuel separately by driving with 1 gallon of fuel  $A$  and then with 1 gallon

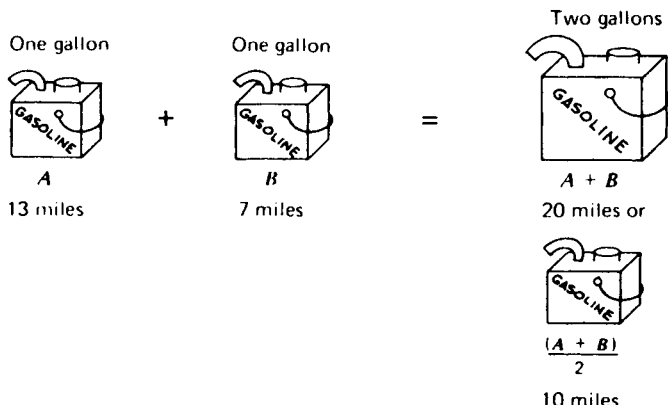


Figure 1.1. Summing the miles per gallon of fuels *A* and *B*.

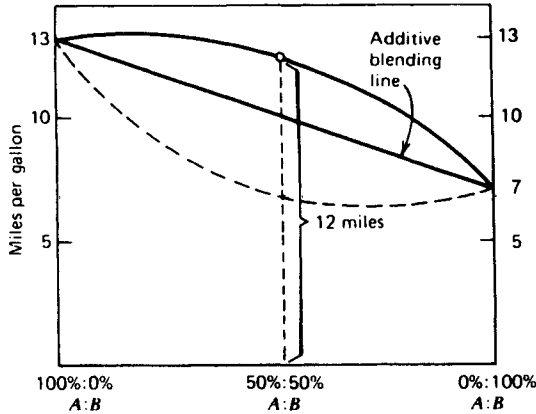
of *B*, we would expect to travel  $13 + 7 = 20$  miles on the 2 gallons or, equivalently, we expect to average  $20/2 = 10$  miles per gallon (Figure 1.1). The question we should like to answer therefore is: “If we combine or blend the two fuels and drive the same test car, is there a blend of *A* and *B* such as a 50% : 50% blend or a 33% : 67% blend of *A* : *B* that yields a higher average number of miles per gallon than the 10 miles per gallon that was obtained by simply averaging the mileage of *A* and of *B*?”

To answer this question, an experiment is performed that consists of driving the test car containing a 50% : 50% blend of fuels *A* and *B*. A trial consists of driving the car with 2 gallons of fuel until the fuel is used up. Five trials were performed with the same car and the average mileage was calculated to be 12.0 miles per gallon. (See Table 1.1.)

The average number of miles per gallon for the blend is 12.0 and is higher than the simple average mileage of the two fuels, which was 10 miles per gallon. Thus fuels *A* and *B* are said to be complementary to each other when blended together. If the average mileage for all blends of *A* and *B* is higher than the simple average of the two, this phenomenon might be depicted by the solid curve in Figure 1.2. If the mileage figure per gallon is strictly additive, that is, if the 50% : 50% blend resulted

Table 1.1. Average Mileage for Each of Five Trials

Trial	Mileage from Two Gallons of 50% : 50% Blended Fuel	Average Mileage per Gallon
1	24.6	12.30
2	23.3	11.65
3	24.3	12.15
4	23.1	11.55
5	24.7	<u>12.35</u>
	Overall average	12.00



**Figure 1.2.** Plotting the mileage of the 50% : 50% blend of fuels *A* and *B*. The formula for the additive blending line is  $\text{mileage} = [(13 \text{ miles} \times A) + (7 \text{ miles} \times B)]/100\%$ .

in exactly 10 miles per gallon, or if a 33% : 67% blend of *A* : *B* resulted in  $[(13 \text{ miles} \times 33\%) + (7 \text{ miles} \times 67\%)]/100\% = 9$  miles, and if this additivity property is true of all possible *A* : *B* blends, then this additive mileage property is represented by the straight line connecting the mileage values 13 and 7 for fuels *A* and *B*, respectively, in Figure 1.2. If the average mileage for blends is lower than the simple average mileage, this is represented by the dashed curve.

**Definition.** In the general mixture problem, the measured response is assumed to depend only on the proportions of the ingredients present in the mixture and not on the amount of the mixture. (We shall modify this definition later in Chapter 7 when discussing mixture-amount experiments.)

In the fuel mileage example, the measured response was the average number of miles traveled per gallon and by putting the amount of fuel on a per-gallon basis, we made the mileage dependent only on the proportions of the two fuels in the blends and not on the quantity of fuel used. Experimenters often satisfy the definitional requirement by fixing the total amount of the ingredients to be the same value for all blends. On a slightly different note, if in the measurement of crop yields due to various mixtures of fertilizers the amount of fertilizer applied to the plots is allowed to vary, then the amount can greatly affect the yield. If we fix the amount of applied fertilizer to be constant and the same on all plots, however, then the fertilizer trials can be considered a legitimate mixture problem because the crop yield per plot would then be a function only of the ingredient proportions.

The distinguishing feature of the mixture problem is that the independent or controllable factors (fuels *A* and *B*, or the fertilizer ingredients) *represent proportionate amounts* of the mixture rather than unrestrained amounts where the proportions are by volume, by weight, or by mole fraction. The proportions are

nonnegative, and, if expressed as fractions of the mixture, they must sum to unity, especially if they are the only ingredients to be studied comprising the mixtures. If the sum of the component proportions is less than unity, for example, if the sum is equal to 0.80 because 0.20 of the blend is held fixed, and we wish to work only with the variable proportions summing to 0.80, then the variable proportions are rewritten as scaled fractions so that the scaled fractions sum to unity. Clearly, if we let  $q$  represent the number of ingredients (or constituents) in the system under study and if we represent the proportion of the  $i$ th constituent in the mixture by  $x_i$ , then

$$x_i \geq 0, \quad i = 1, 2, \dots, q \quad (1.1)$$

and

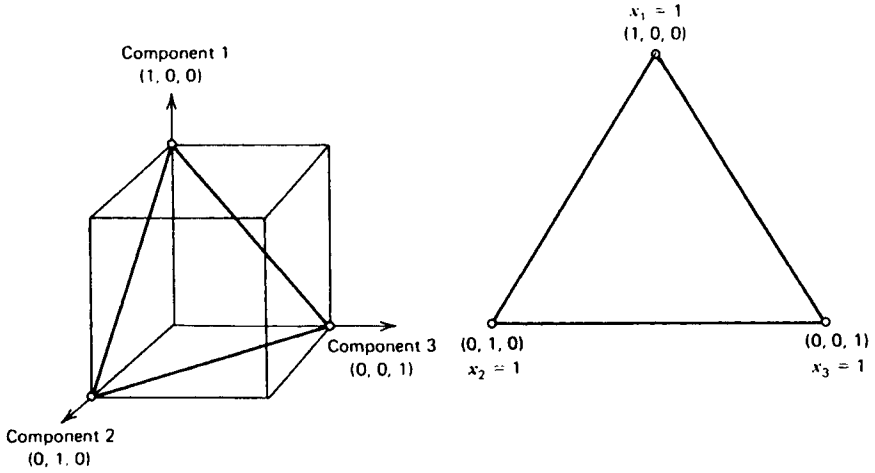
$$\sum_{i=1}^q x_i = x_1 + x_2 + \dots + x_q = 1.0 \quad (1.2)$$

According to Eq. (1.2), the sum of the nonnegative component proportions or fractions is unity. This latter condition (1.2) will be the fundamental restriction assigned to the proportions comprising the mixture experiment.

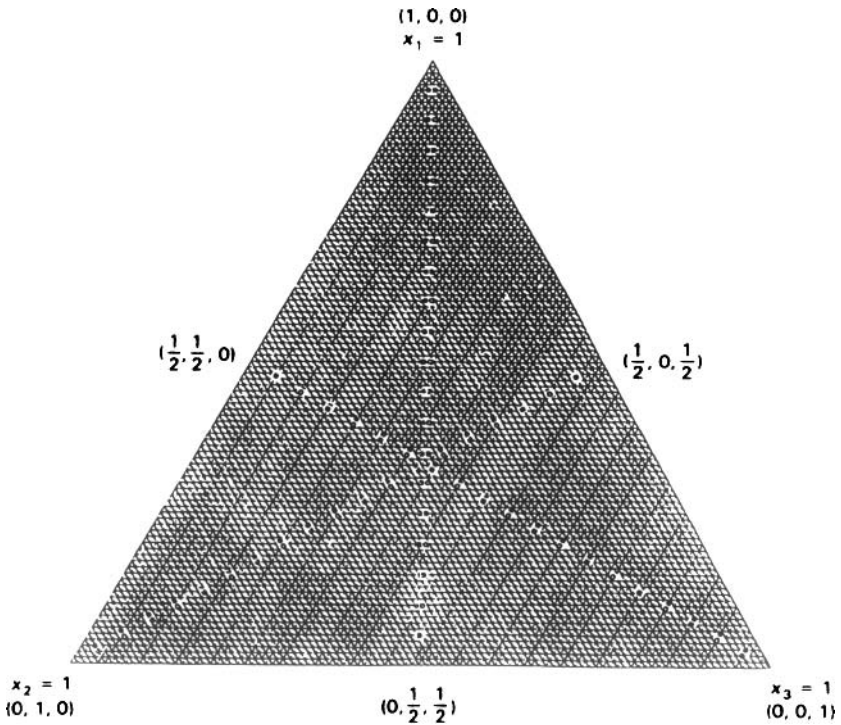
Satisfying the restrictions in Eqs. (1.1) and (1.2) means only that a mixture composition will be formed by adding together nonnegative quantities. Actually, since in Eq. (1.2) an individual proportion  $x_i$  could be unity, a mixture could be a single ingredient or constituent. Such a mixture is called a *pure* mixture or a “*single-component*” mixture. Single-component mixtures are used mainly as a benchmark or as a standard against which multicomponent blends are compared. We discuss the use of single-component mixtures in the next chapter when introducing the simplex-lattice designs, but in Chapters 3–6 nearly all of the design points selected will require most of the constituents to be simultaneously present in the blends. Hereafter we shall call the  $x_i$ ,  $i = 1, 2, \dots, q$ , satisfying Eqs. (1.1) and (1.2) the *components* of the mixture.

By virtue of the constraints on the  $x_i$  shown in Eqs. (1.1) and (1.2) the geometric description of the factor space containing the  $q$  components consists of all points on or inside the boundaries (vertices, edges, faces, etc.) of a regular  $(q-1)$ -dimensional simplex. For  $q=2$  components, the simplex factor space is a straight line, represented by the horizontal axis in Figure 1.2. Each blend of the two fuels  $A$  and  $B$  is represented by a point on the line or axis.

With three components ( $q=3$ ), the simplex factor space is an equilateral triangle, and for  $q=4$  the simplex is a tetrahedron. Figure 1.3 presents the factor space for the three components 1, 2, and 3, whose proportions are denoted by  $x_1$ ,  $x_2$ , and  $x_3$ . The coordinate system used for the values of the  $x_i$ ,  $i = 1, 2, \dots, q$ , is called a *simplex coordinate system*. With three components, for example, the coordinates can be plotted on triangular graph paper that has lines parallel to the three sides of an equilateral triangle; see Figure 1.4. In Figures 1.3 and 1.4, we see that the vertices of the simplex or triangle represent the single-component mixtures



**Figure 1.3.** Three-component simplex region. All experimental points must lie on or inside the triangle whose equation is  $x_1 + x_2 + x_3 = 1$ .



**Figure 1.4.** Triangular coordinate paper.

and are denoted by  $x_i = 1, x_j = 0$  for  $i, j = 1, 2, \text{ and } 3, i \neq j$ . The interior points of the triangle represent mixtures in which none of the three components is absent; that is,  $x_1 > 0, x_2 > 0, \text{ and } x_3 > 0$ . The centroid of the triangle corresponds to the mixture with equal proportions  $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$  from each of the components. Figure 1.5 is the tetrahedron for the four components whose proportions are  $x_1, x_2, x_3, \text{ and } x_4$ .

Frequently situations exist where some of the proportions  $x_i$  are not allowed to vary from 0 to 1.0. Instead, some, or possibly all, of the component proportions are restricted by either a lower bound and/or an upper bound. In the case of component  $i$ , these constraints might be written as

$$0 \leq L_i \leq x_i \leq U_i \leq 1.0, \quad 1 \leq i \leq q$$

where  $L_i$  is the lower bound and  $U_i$  is the upper bound. As an example, in the production of a commercial laundry bleach that is to be used for removing ink stains and that is comprised of the constituents bromine ( $x_1$ ), dilute HCl ( $x_2$ ), and hypochlorite powder ( $x_3$ ), to be effective the bleach must contain in solution all three of the constituents. This means that each  $L_i > 0$  and each  $U_i < 1.0$ . To be more exact, it might be necessary to require  $x_2$  (dilute HCl) to take values in the interval  $0.05 \leq x_2 \leq 0.09$ , which therefore forces  $L_2 = 0.05$  and  $U_2 = 0.09$ . Furthermore, the value of  $L_2 = 0.05$  forces  $U_1$  and  $U_3$  to be at most equal to  $1 - L_2 = 0.95$ . Such mixtures in which all of the components are present in nonzero proportions are called *complete mixtures*.

Finally, in the chapters that follow, we shall try to keep in mind the following as being the *potential goals of a mixture experiment*. We shall try to model the dependence of the response variable (or variables if more than one response is of

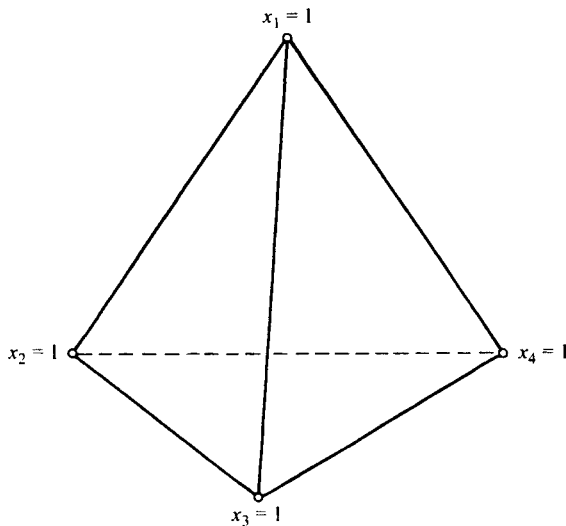


Figure 1.5. Four-component tetrahedron.

interest) on the relative proportions of the components with some form of mathematical equation so that:

1. The influence on the response of each component singly and in combination with the other components can be measured. If this is done successfully, those components having the least effect or felt to be less active might be “screened” out, leaving us with only those components having the greatest effect on the response (Chapters 2 and 5).
2. Predictions of the response to any mixture or combination of the component proportions can be made.
3. Identifying mixtures or blends of the components that yield desirable values of the response.

## 1.2. GENERAL REMARKS ABOUT RESPONSE SURFACE METHODS

In much of the experimental work involving multicomponent mixtures, the emphasis is on studying the physical characteristics, such as the shape or the highest point, of the measured response surface. For example, let us assume that we are making a fruit punch by blending proportions of orange juice ( $x_1$ ), pineapple juice ( $x_2$ ), and grapefruit juice ( $x_3$ ). The response of interest is the fruitiness flavor of the punch quantified on a 1–9 scale as 1 = not fruity, 5 = average, 9 = extremely fruity. If the measured response or flavor rating in this case to any blend of the juices can be represented by the perpendicular height directly above the blend whose coordinates are located inside or on the boundaries of the triangle, then the locus of the flavor values for all one-, two-, and three-juice blends can be visualized as a surface above the triangle. One such surface, which is assumed to be continuous for all possible juice blends, is presented in Figure 1.6 and the contour plot of the estimated flavor surface is presented in Figure 1.7. Geometrically, each contour curve in Figure 1.7 is a projection onto the three-component triangle of a cross section of the flavor surface made by a plane, parallel to the triangle, cutting through the surface at a particular height. The heights of the cutting or intersecting planes that generated the contour curves in Figure 1.7 range from 5.8 up to 6.6.

The main considerations connected with the exploration of the response surface over the simplex region are (1) the choice of a proper model to approximate the surface over the region of interest, (2) the testing of the adequacy of the model in representing the response surface, and (3) a suitable design for collecting observations, fitting the model, and testing the adequacy of fit. To this end, we shall assume that there exists some functional relationship

$$\eta = \phi(x_1, x_2, \dots, x_q) \quad (1.3)$$

that, in theory, exactly describes the surface. We shall write the quantity  $\eta$  to denote the response value that is dependent on the proportions  $x_1, x_2, \dots, x_q$  of the



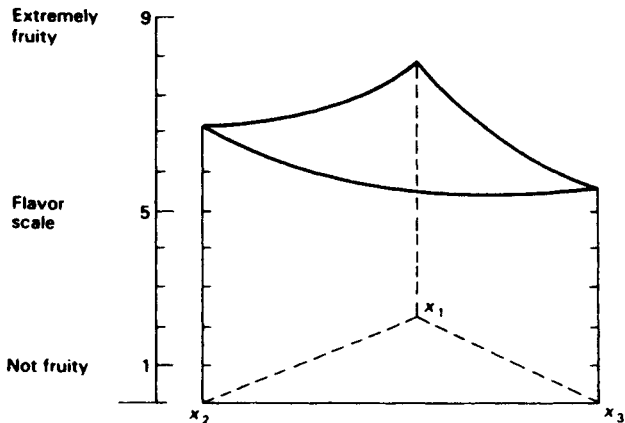


Figure 1.6. Fruitiness flavor surface of fruit punch.

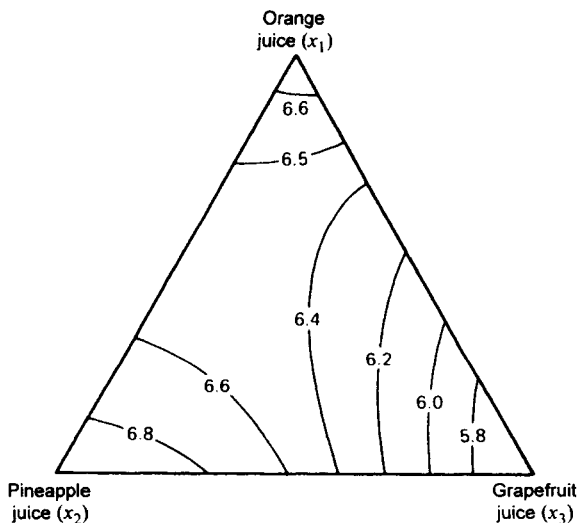


Figure 1.7. Contours of constant fruitiness flavor of the fruit punch surface.

components. One very basic assumption that we are making here is that the response surface, represented by the function  $\phi$ , is depicted to be a continuous function in the  $x_i, i = 1, 2, \dots, q$ . This assumption might be questionable with some systems, for example, a gaseous system whose catalytic reactions break down with the addition or deletion of components. For these systems, model forms other than the standard polynomial equations that we shall work with initially will need to be considered. Chapter 6 presents equations containing inverse terms for the purpose of modeling discontinuities of this type in the surface.

The problem of associating the shape of the response surface with the ingredient composition centers around determining the mathematical equation that adequately

represents the function  $\phi(\cdot)$  in Eq. (1.3). In general, polynomial functions are used to represent  $\phi(x_1, \dots, x_q)$ , the justification being that one can expand  $\phi(x_1, \dots, x_q)$  using a Taylor series, and thus a polynomial can be used also as an approximation. Normally a low-degree polynomial such as the first-degree polynomial

$$\eta = \beta_0 + \sum_{i=1}^q \beta_i x_i \tag{1.4}$$

or the second-degree polynomial

$$\eta = \beta_0 + \sum_{i=1}^q \beta_i x_i + \sum_{i \leq j}^q \sum_{i \leq j}^q \beta_{ij} x_i x_j \tag{1.5}$$

is the kind of model we believe to represent the surface. The low-degree polynomial equations are more conveniently handled than the higher-degree equations because the lower-degree polynomials contain fewer terms and therefore require fewer observed response values in order to estimate the parameters (the  $\beta$ 's) in the equation. On those occasions when a very complicated system is being studied such as shown in Figure 1.8, we may feel the need to use a third-degree equation or some special form of a cubic or third-degree equation (especially when even a transformation of the data values does not simplify the system). Most of the time, however, we shall try to be successful with at most the second-degree model.

Figure 1.8 shows contours of equal dielectric constant lines in the system  $\text{Pb}(\text{Co}_{1/3}\text{Nb}_{2/3})\text{O}_3\text{--PbTiO}_3\text{--PbZrO}_3$  as estimated with a third-degree polynomial equation. As seen from the contours, the dielectric constants in the system increase with increasing proportion of  $\text{Pb}(\text{Co}_{1/3}\text{Nb}_{2/3})\text{O}_3$  up to about 80% : 20% of  $\text{Pb}(\text{Co}_{1/3}\text{Nb}_{2/3})\text{O}_3$ . Near the center of the system is a steep cliff that appears to drop

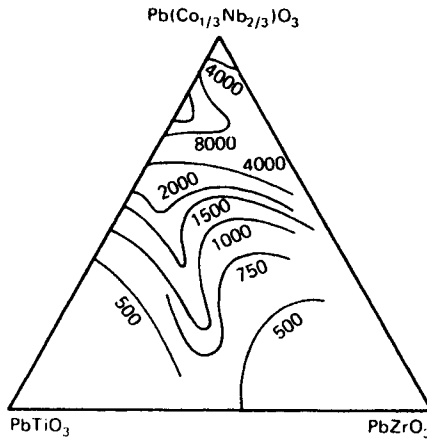


Figure 1.8. Equal dielectric constant lines in system  $\text{Pb}(\text{Co}_{1/3}\text{Nb}_{2/3})\text{O}_3\text{--PbTiO}_3\text{--PbZrO}_3$ .

off in the directions of pure  $\text{PbTiO}_3$  and pure  $\text{PbZrO}_3$ . Contour plots as in Figures 1.7 and 1.8 are extremely helpful when studying a three-component system.

While observing the response  $\eta$  during an experimental program consisting of  $N$  trials, it is natural to assume that the observed value that we denote by  $y_u$  for the  $u$ th trial ( $u = 1, 2, \dots, N$ ) varies about a mean of  $\eta_u$  with a common variance  $\sigma^2$  for all  $u = 1, 2, \dots, N$ . The observed value contains additive experimental error  $\varepsilon_u$ :

$$y_u = \eta_u + \varepsilon_u, \quad 1 \leq u \leq N \quad (1.6)$$

The experimental errors  $\varepsilon_u$  are assumed to be uncorrelated and identically distributed with zero mean and common variance  $\sigma^2$ . These properties of the errors are defined, using an expectation operator  $E(\cdot)$ , as

$$E(\varepsilon_u) = 0, \quad E(\varepsilon_u^2) = \sigma^2, \quad E(\varepsilon_u \varepsilon_{u'}) = 0, \\ u \neq u', \quad u, u' = 1, 2, \dots, N$$

and, therefore, the expected value for the observed value  $y_u$  is  $E(y_u) = \eta_u$ , for all  $u = 1, 2, \dots, N$ .

In order to approximate the functional relationship  $\eta = \phi(x_1, x_2, \dots, x_q)$  with a polynomial or with any other form of model equation, some preselected number of experimental runs are performed at various predetermined combinations of the proportions of the  $q$  components. This *set* of combinations of the proportions (or blends of the ingredients) is referred to as the *experimental design*. Once the  $N$  observations are collected, the parameters in the model are estimated by the *method of least squares*.

As an example, let us suppose we have  $q = 2$  components so that coupled with the structure of  $y_u$  in Eq. (1.6) we may write

$$y_u = (\beta_1 x_1 + \beta_2 x_2)_u + \varepsilon_u \quad (1.7)$$

The absence of the parameter  $\beta_0$  in Eq. (1.7) is due to the restriction  $x_1 + x_2 = 1$ . We shall discuss the derivation of the model (1.7) form in Section 2.2. With some number  $N \geq 2$  of observations collected on  $y_u$ , we can obtain the estimates  $b_1$  and  $b_2$  of the parameters  $\beta_1$  and  $\beta_2$ , respectively. If it is decided that the parameter estimates  $b_1$  and  $b_2$  are satisfactory in the sense that they are nonzero and therefore they relay information about the system we are modeling, then the unknown parameters in Eq. (17) are replaced by their respective estimates to give

$$\hat{y} = b_1 x_1 + b_2 x_2 \quad (1.8)$$

where  $\hat{y}$  (read "y hat") denotes the predicted or estimated value of  $\eta$  for given values of  $x_1$  and  $x_2$ . Of course, before any predictions are made with Eq. (1.8), we must determine that the prediction equation (1.8) does an adequate job of fitting the observed data. We discuss ways of testing the adequacy of empirical models fitted to data in Chapters 2 and 5.

The properties of the polynomials used to estimate the response function depend to a large extent on the specific program of experiments that we have called the experimental design. The experimental design also defines the range of interest of the experimenter with respect to the proportions used for each of the components. This is because the design may cover the entire simplex factor space if the experimenter's interest is with all the values of  $x_i$  ranging from 0 to 1.0 for all  $i = 1, 2, \dots, q$ , or the design might cover only a subportion or smaller subspace within the simplex. This latter situation comes up in practice when additional constraints in the form of upper and/or lower bounds are placed on the component proportions, or, perhaps, when the experimenter is interested only in a group of mixtures that are located in some small region inside the simplex. Both of these cases are discussed in Chapters 4 and 3, respectively.

### 1.3. A FACTORIAL EXPERIMENT OR A MIXTURE EXPERIMENT?

One of the reviewers of the first edition of this text commented that somewhere in the introductory chapter the following question should be addressed: "Why is a book on mixture experiments necessary?" This question no doubt would be followed by a second question: "Can't we perform other types of experiments, such as factorial experiments, and get the same kind of information from them that we get from mixture experiments?"

To answer the second question first, we shall repeat the definition of a mixture experiment given earlier in Section 1.1 and also define a factorial experiment. We shall then present a glass coating example to illustrate the difference between running a factorial experiment and performing a mixture experiment and at the same time address the question of why we need to perform mixture experiments.

**Definition of Mixture Experiment.** This is an experiment in which the response is assumed to depend only on the relative proportions of the ingredients present in the mixture and not on the amount of the mixture. [A mixture-amount experiment (Section 7.9) is an experiment where the amount of the mixture varies as well and the response depends not only on the relative ingredient proportions but also on the total amount of the ingredients.] In a mixture experiment then, if the total amount is held constant and the value of the response changes when changes are made in the relative proportions of those ingredients making up the mixture, then the behavior of the response is said to be a function of the joint blending property of the ingredients in the mixture.

**Definition of Factorial Experiment.** A factorial experiment studies the effect on some observable quantity (the response) of varying two or more factors, such as temperature and source of raw material. A series of values or levels of each factor is chosen, and certain combinations of the levels of the factors are tested. In a complete factorial design, all combinations of the levels of all the factors are tested.

The objective of a factorial experiment is to measure the change in the response when changing the level of each factor while holding the levels of the other factors fixed as well as when changing the levels of two or more factors simultaneously. Such changes in the response are called the main effects of the factors and interaction effects between the factors.

**An Example.** A company manufactures a liquid material that results in a thin film for coating window glass. When applied to the glass, the coating film provides a barrier to ultraviolet rays and also reduces glare from the sun. Initially the coating is a liquid consisting of water ( $W$ ), which is combined with three active solid ingredients: a polymer ( $P$ ), a coupling agent ( $CA$ ), and a lubricant ( $L$ ). The film is formed by extruding the liquid material at high temperatures.

An experiment is to be conducted consisting of several different combinations of  $W$ ,  $P$ ,  $CA$ , and  $L$ , where a single combination comprises a single coating. The thickness of the film coating is determined by how much water and active solids are mixed. The objective of the overall experiment is to determine the combination of  $W$ ,  $P$ ,  $CA$ , and  $L$  that is most effective in terms of reducing light penetration. Light penetration is measured by first taking a bright colored cloth affixed to the coated glass and exposing the cloth to light, through the glass, for a fixed period of time. The color of the cloth after exposure is then compared to an unexposed piece of cloth of the same color and the degree of fading that has occurred is recorded. A low percent fade value is considered to be desirable.

In making up the various combinations of water ( $W$ ), polymer ( $P$ ), coupling agent ( $CA$ ), and lubricant ( $L$ ) to produce the different coatings, we present the following strategies:

*Strategy A.* Each bath of liquid coating material is mixed in a 250-mL beaker. Exactly 225 g of  $W$  is mixed with 25 g of active solids, fixing the ratio of water to active solids at 9 : 1 for all combinations. With the active solids, the amounts of  $P$ ,  $CA$ , and  $L$  are varied as follows:

$P$ (g)	$CA$ (g)	$L$ (g)	Total Active Solids (g)	Percentage of Active Solids		
				$P$	$CA$	$L$
20	4	1	25	80	16	4
22	2	1	25	88	8	4
22	2.5	0.5	25	88	10	2
22.5	2	0.5	25	90	8	2

The amounts of water (225 g) and active solids ( $P + CA + L = 25$  g) are held fixed with all combinations. The amounts of  $P$ ,  $CA$ , and  $L$  are varied, but since  $P + CA + L = 25$  g, this is a *three-component mixture experiment*. Here the interest is in studying how different proportions (or percentages) of  $P$ ,  $CA$ ,

and *L* when combined with water in a 9:1 ratio of water to active solids influence the degree of light penetration.

*Strategy B.* Each batch of liquid coating is mixed in a 250-mL beaker but the amount of water is fixed at 200 g while the amount of active solids is varied between 25 and 43 g. The combinations are as follows:

Water (g)	Active Solids (g)	Total Weight (g)	Ratio
			Water : Active Solids
200	25	225	8.90 : 1.1
200	28.5	228.5	8.75 : 1.25
200	35.3	235.3	8.50 : 1.5
200	42.4	242.4	8.25 : 1.75

Within each fixed amount of active solids, if the percentages of *P*, *CA*, and *L* are varied as in Strategy A, then this is a three-component mixture-amount experiment. If, on the other hand, only one blend of *P*, *CA*, and *L* is studied, say, *P* = 80%, *CA* = 16%, and *L* = 4% within each of the four amounts of active solids, then we have a single-factor experiment with active solids at four levels. The objective is to measure the effect of changing the ratio of water to active solids (or of changing the amount of active solids while holding the amount of water fixed) on the degree of light penetration.

*Strategy C.* Two levels of water (225 and 270 g) and two levels of active solids (25 and 30 g) are to be tested. The percentages of *P*, *CA*, and *L* are fixed at 80%, 16%, and 4%, respectively, with each level of active solids. The combinations are as follows:

Water (g)	Active Solids (g)	Total Amount (g)	Ratio
			Water : Active Solids
225	25	250	9.0 : 1.0
270	25	295	9.15 : 0.85
225	30	255	8.8 : 1.2
270	30	300	9.0 : 1.0

This is a  $2 \times 2$  factorial experiment. The interest is in measuring how changes in the degree of light penetration are influenced by changing or increasing the level of water and by changing the level of active solids.

The difference between a mixture experiment and a factorial experiment is illustrated by comparing Strategies A and C. With Strategy A, the total amount (or weight in grams) of active solids is fixed at 25 g so that when combined with 225 mL of water, the total amount of liquid coating to be sprayed on the glass is

fixed at 250 g. The four combinations of  $P$ ,  $CA$ , and  $L$  differ only in the relative proportions of each of the three ingredients. Thus changes in the degree of light penetration arising from the four different combinations of  $P$ ,  $CA$ , and  $L$  do not occur as a result of varying the amount ( $P + CA + L$ ) of active solids or the amount of coating sprayed on the glass but rather by changing the proportions or percentages of  $P$ ,  $CA$ , and  $L$  in the active solids portion of the blend.

In Strategy C, we have a factorial experiment. For this case the active solids portion of the blend is varied from 25 to 30 g and the percentages of  $P$ ,  $CA$ , and  $L$  are fixed at 80, 16, and 4, respectively. The amount of water is varied from 225 to 275 g. Thus the total amount of coating liquid that is applied to the glass varies from 250 to 300 g. While it is suspected that the degree of light penetration will change from one combination of water and active solids to the next owing to the varying total amounts, we are primarily interested in finding out how changing the *amount of active solids*, with fixed percentages of  $P$ ,  $CA$ , and  $L$ , affects the degree of light penetration at each level of water and vice versa. Hence the latter experimental strategy emphasizes the amount of coating that is applied by varying the amounts of water and active solids, while the former experiment (Strategy A) fixes the total amount that is applied and investigates the changes in the response of interest that are effected by changing the ingredient proportions within each blend.

### 1.3.1. A Warning Regarding a Strategy that Doesn't Work

A question that often comes up is: “Isn't it possible to perform a factorial experiment in the proportions or amounts of only  $q-1$  of the components and let the  $q$ th component take up the slack?” As an example, suppose with Strategy A we elect to perform a  $2 \times 2$  factorial experiment in the amounts of  $P$  and  $CA$  (and let  $L$  make up the remainder of the 25 g) but ignore  $L$ . Such a strategy might be depicted as follows:

$P$ (g)	$CA$ (g)	$L$ (g)	Total	Response	$z_1$	$z_2$
			Active Solids (g)			
20	2	3	25	$y_1$	-1	-1
22	2	1	25	$y_2$	1	-1
20	3	2	25	$y_3$	-1	1
22	3	0	25	$y_4$	1	1

Next, suppose we define the coded variables  $z_1$  and  $z_2$  as  $z_1 = (P-21)/1$  and  $z_2 = (CA - 2.5)/0.5$ , so that when  $P = 20$  or 22 then  $z_1 = -1$  or 1, and when  $CA = 2$  or 3 then  $z_2 = -1$  or 1. When fitted to the response values  $y_1, y_2, y_3$ , and  $y_4$  above, the first-degree model in  $z_1$  and  $z_2$ ,  $E(y) = \beta_0 + \beta_1 z_1 + \beta_2 z_2$ , becomes

$$\hat{y}(z_1, z_2) = b_0 + b_1 z_1 + b_2 z_2 \quad (1.9)$$

where  $b_0 = (y_1 + y_2 + y_3 + y_4)/4$  is an estimate of the overall mean response, which is represented by  $\beta_0$  in the model for  $E(y)$ . The coefficients  $b_1 = (y_2 + y_4 - y_1 - y_3)/4$

Table 1.2. Chronological Listing of Selected Statistical Literature on Mixtures from 1953 to the Present

Year	Authors	Year	Authors	Year	Authors
1953	Quenouille	1975	Cornell	1984	Aitchison and Bacon-Shone
1955	Claringbold*		Cornell and Ott		Chick and Piepel
1958	Scheffé*		Laake		Cornell and Gorman
1959	Quenouille		<u>Snee</u>		Crosier*
1961	Scheffé		Mendieta, Linssen, and Doornbos		St. John
	John and Gorman	1976	Snee and Marquardt		Ying-nan
1962	Gorman and Hinman*	1977	Cornell	1985	Yonchev
	Wagner and Gorman		Draper and <u>St. John</u>		Aitchison
1963	Kenworthy		<u>Galil and Kiefer</u>		Cornell
	Scheffé*		Hare and Brown		Darroch and Waller
	Wagner and Gorman		Saxena and Nigam		Gorman and Cornell
1964	Myers	1978	Becker		Hare
	Uranisi		Cornell and Gorman		Piepel and Cornell
1965	Bounds, Kurotori, and Cruise		Park		Snee
	<u>Draper and Lawrence*</u>		Vuchkov, Yonchev, and Damgaliev	1986	Ying-nan
1966	Cruise	1979	Cornell and Khuri		Cain and Price
	Box and Gardiner		Cornell		Cornell
	Gorman		Goel and Nigam		Crosier
	Kurotori*		Hare	1987	Zhu, Hu, and Chen
	McLean and Anderson*		<u>Snee*</u>		Hoerl
1967	Diamond	1980	<u>Goel</u>		Piepel and Cornell
	Drew				Sahrman, Piepel, and Cornell
					Ying-nan
					<u>Zhu, Hu, and Chen</u>
					Cornell
					Draper et al.
					Diuneveld, Smilde, and Doornbos
					Mikaeili
					Murthy and Murty
					Prescott et al.
					Vining, Cornell, and Myers
					Derringer
					Lewis et al.
					Montgomery and Voth
					Piepel and Cornell
					Smith and Cornell
					Cornell
					Heinsman and Montgomery
					Hilgers and Bauer
					Chen, Li, and Jackson
					Murthy and Manga