JOHN A. CORNELL

EXPERIMENTS WITH MIXTURES

DESIGNS, MODELS, AND THE ANALYSIS OF MIXTURE DATA

WILEY SERIES IN PROBABILITY AND MATHEMATICAL STATISTICS

Experiments With Mixtures

DESIGNS, MODELS, AND THE ANALYSIS OF MIXTURE DATA

John A. Cornell
University of Florida

JOHN WILEY & SONS New York · Chichester · Brisbane · Toronto

Library of Congress Cataloging in Publication Data:

Cornell, John A 1941–
Experiments with mixtures.

(Wiley series in probability and mathematical statistics)
Includes bibliographies and index.
1. Experimental design. 2. Mixtures—Statistical methods. I. Title.

QA279.C66 519.5′352 80-22153
ISBN 0-471-07916-2

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

Preface

The primary purpose of this book is to present the fundamental concepts in the design and analysis of experiments with mixtures. The book focuses on the most frequently used statistical techniques and methods for designing, modeling, and analyzing mixture data, as claimed in the literature, and includes appropriate computing formulas and completely worked out examples of each method. Most of the numerical examples were taken from real research situations.

The book is written for anyone who is engaged in planning or performing experiments with mixtures. In particular, research scientists and technicians in the chemical industries, whether or not trained in statistical methods, should benefit from the many examples that are chemical in nature. Several examples have been taken from research activities conducted in areas of food technology, while some examples were provided by research entomologists. Persons who are engaged in applied research in universities, principally from such departments as chemical engineering, chemistry, and statistics, as well as scientists in areas of agriculture such as food science, entomology, and nematology, should find the methods that are presented to be relevant and useful in their research. As a textbook on the subject of mixture experiments, the contents could serve quite nicely as a one-semester course in most applied curricula, or perhaps could supplement the coverage of a two-semester sequence of regression and response surface methodology.

Since this is the first edition, it has been necessary to exercise considerable selectivity in the choice of topics covered. Hence no claim is made that the coverage is exhaustive in either scope or depth. However, it is my feeling that the reader who works through the numerical examples in the middle five chapters (Chapters 2–6) and answers the questions listed at the end of these chapters will achieve a high level working knowledge of the tools that are used by most of the practitioners today who are involved in solving mixture problems.

The mathematical prerequisites have been kept to a minimum. Summation notation is used throughout and some background knowledge of the use of matrices is helpful. A review of matrix algebra is presented in Chapter 7 along with a discussion of the method of least squares for obtaining the parameter estimates in polynomial models. Chapter 7 could also serve as a refresher to readers who wish to review some of the fundamental ideas on the use of matrices in regression analysis. The matrix material has been placed at the end of the book so that the reader with an adequate knowledge of matrices may begin with the subject of mixture experiments in Chapter 1. Almost all of the computations throughout the book were performed on the APL system 360.

The first chapter introduces the subject of mixture experiments with several examples. Some general remarks on response surface methodology are made. An historical perspective of the relevant literature which presented most of the statistical research on mixture experiments is listed. Chapter 2 introduces the original mixture problem where the Scheffé lattice designs and associated polynomial models are applicable. Several numerical examples are provided that help to illustrate the fitting of the polynomial models to samples of mixture data that were collected at the points of the simplex-lattice and simplex-centroid designs.

In Chapter 3 a transformation is made from the system of the dependent mixture components to a system of independent variables. With the independent variables, standard regression procedures are suggested not only for the designing of the experimental runs but for the fitting of model forms as well. The idea of isolating the experimentation to a subregion of interest inside the simplex space where the region may be ellipsoidal or cuboidal in shape is also considered. Process variables, such as cooking time and cooking temperature in the preparation of fish patties, are introduced. Different types of model forms used to measure the influence the process variables could have on the blending characteristics of the components in mixture experiments are presented and discussed.

How the placing of additional constraints on the component proportions can affect the design configuration and the usual interpretation of the model parameters is considered in Chapter 4. Experimental design configurations for use in covering the restricted region of the simplex are mentioned, as are several types of polynomial model forms used for depicting the surface characteristics. Pseudocomponents are introduced, and the use of pseudocomponents rather than the original components is seen to simplify the steps in the design construction and the fitting of models when lower bound constraints are placed on the original component proportions. Some discussion on the design strategy when some or all of the component proportions are subjected to both upper and lower bound constraints is presented along with some suggested modifications that need to be made to interpret the model coefficients in these highly constrained problems. Grouping the components by categories is also studied.

Chapter 5 presents many techniques that are used in the analysis of mixture data. Testing the form of the fitted model, model reduction procedures, and the screening of unimportant components are just some of the topics covered. Investigating the shape characteristics of the surface by the measuring of the slopes of the surface along the component axes is discussed. Combining lattice designs in the mixture components with factorial arrangements of process variables is illustrated with data from a fish patty experiment in which patties that were made from three species of fish are prepared and processed by the three cooking and processing factors.

Alterations made to some of the terms of the Scheffé polynomials as well as the suggestion to use nonpolynomial models to model certain types of phenomena is the theme of Chapter six. Models that are homogeneous of degree one are shown to model additive component effects better than the polynomials. The use of ratios of the components as terms in the model is suggested particularly when relationships between the component proportions are more meaningful than the actual fraction of the mixture each component represents. Standard orthogonal designs, such as standard factorial arrangements, that can be used with independent variables are shown to be useful when working with ratios. Cox's polynomial, which is used for measuring the components effects, is compared with the several forms of Scheffé's polynomials. Two classes of octane-blending models are presented at the end of the chapter, and data are provided to help illustrate the numerical computations that are required to set up the prediction equations. Chapter 6 ends with a list of topics that were not covered here but which hopefully will be discussed in a future edition.

I am extremely grateful to many friends for their help in compiling the material for this work. In particular, I am indebted to Drs. John W. Gorman and R. Lymann Ott, Professors Irving John Good (Virginia Polytechnic Institute and State University) and Andre I. Khuri (University of Florida), with whom I have had the pleasure of working on research problems in mixtures, and to Dr. Hubert M. Hill (Tennessee Eastman Company), who introduced me to the subject of mixture experiments in the middle 1960s. I am very much indebted also to Professor J. Stuart Hunter (Princeton University), who reviewed the initial drafts of this book; his many thoughtful and detailed comments on the style and content were instrumental in its organization. I wish to thank the many authors and various publishers for permission to reproduce their papers and tables and to thank the staff at John Wiley and Sons, particularly Beatrice Shube, for her encouragement in completing this work. Finally, I would like to express my sincere appreciation to Donna Alexander for her excellent typing of the final manuscript.

JOHN A. CORNELL

University of Florida
Gainesville, Florida
August 1980

Contents

CHAPTER 1

Introduction

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

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.

In each of cases 1–4, one or more product properties is of interest to the manufacturer or experimenter who is responsible for mixing the ingredients. Properties such as (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 the strength (measured in psi's) of the concrete, where the hardness is a function of the percentages of cement and sand and water in the mix; (3) the illumination in foot-candles and the duration of the illumination of the flares; and (4) the fruitiness flavor of the punch which will depend on the percentages of watermelon, pineapple, and orange that are present in the punch. In each of examples 1–4, the measured property of the final product depended on each of the individual ingredients being present in the formulation.

Another reason for mixing together ingredients in blending experiments is to see if there exist blends of two or more ingredients that produce more desirable product properties than is 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 if 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 $B:C$, which yields a higher antiknock rating than is obtained from using A alone or from using B alone, or C alone. If so, we would probably select the particular blend of two or more gasoline stocks that produces the highest rating, assuming of course that all other factors such as the cost and availability of the blending ingredients remain fixed.

In each of cases 1–4 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 this case are the proportions of the ingredients of cement to sand to water) is to try to determine if some combination of the ingredients can be considered best in some sense. The best ingredient combination for the concrete would be the combination that produced the absolutely strongest concrete without increasing cost. In an attempt to determine the best combination of the 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. The 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 unimportant mixture ingredients are discussed in Section 5.7. Before we discuss some methods that have been developed for studying functional relations and are referred to as response surface methods, we introduce the general mixture problem.

1.1. THE GENERAL 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 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

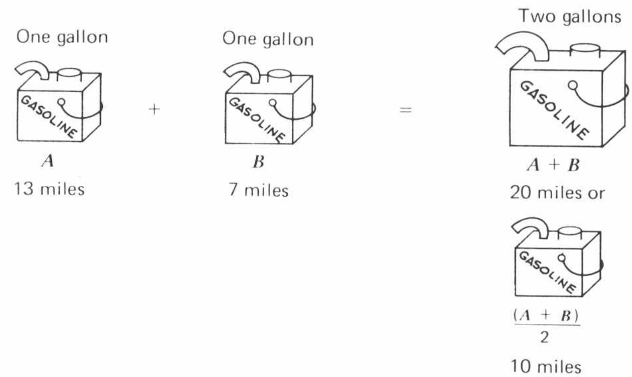

FIGURE 1.1. Summing the miles per gallon of fuels *A* and *B*.

each fuel separately by driving with 1 gallon of fuel *A* and then with 1 gallon of *B*, we would expect to drive $13 + 7 = 20$ 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 *A* and *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 miles per gallon and is higher than the simple average mileage of the two fuels

TABLE 1.1. The 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	12.35
	Overall average	12.00