5-2010

A New Screening Methodology for Mixture Experiments

Maria Weese
University of Tennessee - Knoxville, mweese@utk.edu

To the Graduate Council:

I am submitting herewith a dissertation written by Maria Weese entitled "A New Screening Methodology for Mixture Experiments." I have examined the final electronic copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Business Administration.

Mary G. Leitnaker, Major Professor

We have read this dissertation and recommend its acceptance:

Robert Mee, Russell Zaretzki, Charles Moore

Accepted for the Council:
Dixie L. Thompson

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

To the Graduate Council:

I am submitting herewith a dissertation written by Maria L. Weese entitled "A New Screening Methodology for Mixture Experiments". I have examined the final electronic copy of this thesis for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Business Administration.

Mary G. Leitnaker, Major Professor

We have read this thesis
and recommend its acceptance:

Robert Mee

Russell Zartezki

Charles Moore

Accepted for the Council:

Carolyn Hodges

Vice Provost
and Dean of Graduate Studies

A New Screening Methodology for Mixture Experiments

A Dissertation Presented for the

Doctor of Philosophy

Degree

The University of Tennessee, Knoxville

Maria L. Weese

May 2010

Acknowledgements

I first wish to thank Mary Leitnaker for all of her guidance patience, and friendship during this research process. I would also like to thank Brandon and my parents for all of their patience and support during this time. I wish to thank Dr. Bobby Mee and Dr. Russell Zartezki for their invaluable comments, help, and input into this research. Lastly, I would like to thank Google for providing Gmail Chat so that even though my old office mates Arto Armagan and David Edwards are far away we can still waste time chatting about nothing.

Abstract

Many materials we use in daily life are comprised of a mixture; plastics, gasoline, food, medicine, etc. Mixture experiments, where factors are proportions of components and the response depends only on the relative proportions of the components, are an integral part of product development and improvement. However, when the number of components is large and there are complex constraints, experimentation can be a daunting task. We study screening methods in a mixture setting using the framework of the Cox mixture model [1]. We exploit the easy interpretation of the parameters in the Cox mixture model and develop methods for screening in a mixture setting. We present specific methods for adding a component, removing a component and a general method for screening a subset of components in mixtures with complex constraints. The variances of our parameter estimates are comparable with the typically used Scheffé model variances and our methods provide a reduced run size for screening experiments with mixtures containing a large number of components.

We then further extend the new screening methods by using Evolutionary Operation (EVOP) developed by Box and Draper [2]. EVOP methods use small movement in a subset of process parameters and replication to reveal effects out of the process noise. Mixture experiments inherently have small movements (since the proportions can only range from zero to unity) and the effects have large variances. We update the EVOP methods by using sequential testing of effects opposed to the confidence interval method originally proposed by Box and Draper. We show that the sequential testing approach as compared with a fixed sample size reduced the required sample size as much as 50% with all other testing parameters held constant. We present two methods for adding a component and a general screening method using a graphical sequential t-test and provide R-code to reproduce the limits for the test.

Contents

List of Tables

List of Figures

Chapter 1

Mixture Experiments

A mixture experiment is an experiment where the responses are assumed to depend only on the relative proportion of the components in the mixture and not on the amount of the mixture. For example, the response might be the texture of a fish patty, the leaching ability of nuclear waste glass, or the color of paint. In each of the above examples, the response depends not on the amount of the substance, but the proportions of the ingredients. In the most basic mixture experiment, the q components in the mixture satisfy the following constraints:

$$0 \leq x_i \leq 1 \tag{1.1}$$

$$\sum_{i=1}^{q} x_i = 1 \tag{1.2}$$

Hence the proportion of each component must be between 0 and 1 and the proportions of the q components in the mixture must sum to unity. The factor space for an experiment with constraints (1.1) and (1.2) is a $q-1$-dimensional simplex which might include the edges and the interior in the experimental region.

Many give credit to Quenouille for his treatment of this type of experiment in his 1953 book [3]. P.J. Claringbold [4] in was the first to utilize a simplex design in an experimental situation. However, in 1958 Henry Scheffè formalized the theory in his paper "Experiments with Mixtures" [5]. In this paper Scheffè formally introduced the Simplex-Lattice design

and the corresponding Scheffè canonical polynomial model. Scheffè defines a $[q, m]$ lattice where q is the number of components in the mixture and m is the degree of the polynomial to be fit to the design. In a Simplex Lattice design there are $\binom{m+q-1}{m}$ design points. The proportions used for each factor have $m+1$ equally spaced values from 0 to 1 of $x_i = 0, 1/m, 2/m, ..., 1$. There is a one-to-one correspondence of the design points to the parameters in the polynomial. For example in a $[q, 1]$ simplex there are $\binom{1+q-1}{1}$ or q design points. The Scheffè polynomial for a $[q, 1]$ simplex lattice if of the form:

$$\eta = \sum_{i=1}^{q} \gamma_i x_i \tag{1.3}$$

owing to the substitution $x_q = 1 - \sum_{i=1}^{q} x_i$ into the standard polynomial form:

$$\eta_1(x) = \beta_o + \sum_{i=1}^{q} \beta_i x_i$$

The coefficients in this polynomial have a one-to-one correspondence with the points in the design. As mentioned above, there are q design points in a $[q, 1]$ lattice design thus for a three component mixture, there are three design points and in the corresponding polynomial, three parameters to be estimated. This correspondence allows for the coefficients to be estimated by means of least squares regression or simple linear combinations of averages of the response values at each design point. The coefficients in equation (1.3) are interpreted as the height of the response over the i^{th} pure component vertex.

For mixtures where the response is expected to take on a non-linear form the Scheffè second order polynomial takes the form:

$$\eta = \sum_{i=1}^{q} \gamma_i x_i + \sum_{i=1}^{q-1} \sum_{j=i+1}^{q} \gamma_{ij} x_{ij}$$

In this canonical form, the pure quadratic terms are combined with the two factor quadratic terms owing to the substitution $x_i^2 = x_i(1 - \sum_{j=1, j\neq i}^{q} x_j)$ in addition to the substitution used in (1.3). With these substitutions, the degree of the polynomial is not changed, and the number of terms is $\binom{m+q-1}{m}$ keeping the one-to-one correspondence of design points and parameters in the model. In mixture experiments, the interaction terms

in the model are referred to as non-linear blending terms. The non-linear blending terms, responses to binary or ternary mixtures, can be interpreted as being either a synergistic (having a positive excess over the linear blend response) effect or an antagonistic (having a negative excess below the linear blending response) effect. These interpretations of binary and ternary mixture terms are widely used in describing effects of components on the properties of a mixture.

Gorman and Hinman [6] extended the Scheffé polynomials up to the third and fourth degree and are the first to note a caution regarding Scheffè polynomials when the experimental region is small. And as Cornell [7] states, Lambrakis [8] generalizes Scheffè's polynomials to order m.

Scheffè [9] continued his pioneering work in the field with the introduction of the Simplex Centroid design. This design contains $2^q - 1$ mixtures containing q pure component blends, $\frac{q}{2}$ binary blends with equal proportions, and $\frac{q}{3}$ ternary blends with equal proportions up to the q-nary mixture with equal proportions. The Simplex Centroid "involves observations on mixtures consisting of every (non-empty) subset of the q components, but only on mixtures in which the components present appear in equal proportions"[9]. Similarly with the Simplex-Lattice designs, the Simplex-Centroid design has a one-to-one correspondence with the Scheffè polynomial and the coefficients can be estimated utilizing linear combinations of the responses at each of the design points. In this 1963 work, Scheffè also introduces the inclusion of process variables in to the mixture experiment as a factorial experiment.

Many, including Scheffè, recognized some of the shortcomings of the Simplex-Lattice and Simplex-Centroid designs. In his first paper, Scheffè already realizes there will be situations where not all of the components are allowed to vary from zero to unity. He introduces pseudocomponents for a case where one of the components has an upper bound. However, the interpretation of the coefficients in his polynomial becomes non-trivial. Additionally, the lower-order designs do not include many if any interior points in the design region. In some cases, the only viable experimentation is with the interior points. Draper and Lawrence [10][11] recognize this deficiency and consider designs for three and four factors where all the design points are interior points. Their primary goal of experimentation is to fit a response surface model to the design space. These points are found by using design

criterion proposed by Box and Draper [12].

Kurotori [13] recognized that in addition to upper bounds, components sometimes will require lower bounds, i.e., the component must be present in the mixture. McLean and Anderson [14] also saw the need for a way to experiment with only complete mixtures and were the first to introduce an algorithm to search for vertices of a constrained region. Some of the vertices become the design points in an Extreme Vertices design for constrained mixture experiments. Similarly, Lambrakis [15] introduced a design where each design point has all of the components present in each mixture. Thompson and Myers [16] also introduce a design to estimate the response surface of a mixture using an ellipsoidal region which is centered at a point of "maximum interest" to the experimenter. In these types of designs, the ellipsoid is determined by the experimenters.

Lambrakis [8] suggests a generalization of Scheffè's mixture problem where he defines major and minor mixture components, these designs are referred to as multiple-lattice designs. He also generalized Scheffè's canonical polynomial to analyze these multiple lattice designs [17]. In all of the cases mentioned above, all of the analyses (if it was mentioned at all) were performed with canonical polynomials either by the Scheffè method or one similar to it. The first alternative model to the canonical polynomial is introduced by Becker [18].

Becker introduces a homogenous model of degree one for fitting mixtures to aid in the interpretation when there is an inert or additive component in the mixture. In a subsequent paper, Becker [19] discusses regression procedures for mixture variables. Various other model forms, such as the log contrast model developed by Aitchison and Bacone-Shone [20] and a model including inverse terms introduced by Draper and St. John [21] to account for possible extreme changes in the response as a component approaches zero. Cox [1] presents an alternative to Scheffè's polynomial where the parameters represent the changes in the slope of the response surface by comparing the changes in the response value at a standard to the response value at other points in the simplex. The Cox polynomial model will be discussed in detail in a later section.

Despite the vast amount of literature regarding mixture experiments, the specific topic of screening in a mixture setting has only been the topic of a few journal articles, one section in *Experiments with Mixtures* [22] and glossed over in *Experimental Design for Formulation* [23]. All of the literature on screening concepts for mixture experiments is

structured on finding large linear effects using the first-degree Scheffè polynomial, which directly follows the definition of a factorial screening design.

In a factorial screening design, the goal is to determine which of a large number of factors are the important ones as defined by Box, Hunter and Hunter (p. 173) [24]. Once the important factors are found, then follow up experiments can be conducted. Wu and Hamada (p. 390) [25] and Montgomery (p. 303) [26] state that a first order model should be used to analyze a screening experiment and the goal is to identify the most important factors out of many based on their first-order effect estimates.

Snee and Marquardt [27] are the first to discuss screening concepts with mixture experiments and state that the only differences in the screening philosophy for mixture variables verses ordinary independent variables is in the "concept of screening as applied to mixtures and the manner in which the designs are constructed". They directly translate the philosophy of searching for large linear effects in a screening experiment by utilizing design points lying on what they define as component axes. In this manner they are able to effectively study "directions through the factor space where the response is constant, or nearly so" in order find the important mixture variables. They recommend using the Scheffè canonical polynomial to model the linear blending of the mixture and then testing the effects in contrasts to determine which effects are equal or not or which sums of contrasts are equal. In this way they can then reduce the number of components to work with when doing follow up experiments. They have directly adapted the screening philosophy from independent variables to the mixture setting. Even though they discuss the Cox [1] polynomial and adopt the concept of the effect direction to find design points, they model and evaluate the effects using the first order Scheffè polynomial. Snee and Marquardt's designs are for six or more mixture components. They do mention that if you have five or fewer components, it would be better to collect enough data to fit a quadratic model.

Snee and Marquardt go on to suggest Simplex Screening designs for when the components can be varied over the total composition range or when the experimental region can be expressed in pseudocompoents. These designs include all pure component blends, the centroid, interior points on the component effect axes, and in cases where it is suspected that a total absence of a component's presence will have a large effect on the response, they add what they call end effect points. These end effect points are the midpoint of the edge

where the component of interest is not present in the mixture, but the other components are. The designs then either consist of $2q+1$ points without the end effects or $3q+1$ points including the end effect points. These designs have fewer points when $q > 5$ than Scheffè's simplex lattice designs since their goal is to give enough design points to estimate large linear effects.

Sung Park [28] expands upon Snee and Marquardt's suggestions of testing various linear contrasts to determine large linear effects by "proposing a criterion to decide which set of contrasts should be adopted to provide the best performance of the fitted response surface over some region of experimental interest in the sense of MSE". Park states that in a mixture screening experiment, the functional relationship is assumed to be linear. Park also mentions Cox's polynomial as an alternative form of a mixture model but does not elaborate any further.

Finally, in a 1990 article, Gergory Piepel [29] develops screening designs for constrained mixture experiments derived from classical screening designs. Piepel states that "The purpose of a mixture screening design is to determine which components have significant effects on the response. This is typically accomplished by conducting a screening experiment that supports fitting the first-order Scheffè canonical polynomial mixture model and then using the fitted model to estimate the component effects." Piepel is assuming that the non-linear blending effects will be small relative to the linear blending effects for the active components. Peipel does discuss Cox component effects, but only in relation to assessing the performance of mixture screening designs generated in his paper.

One of the few pieces of literature found where the Cox model coefficients were utilized is a paper by Greg Piepel [30]. Piepel et. al. makes use of the Cox model and associated design points to assess curvature and interaction mixture experiments by developing Mixture Interaction Plots. The authors define an interaction in terms of mixture components by utilizing an interaction type plot that would typically be seen in a factorial experiment. Piepel et. al. extend concepts that Cox originally proposed in his 1971 paper, but did not discuss extensively, i.e. the interaction terms and associated design points that can be estimated by his model and associated designs. In a factorial experiment, a 2^2 experiment can be used to asses the linear and interaction effects between two independent factors. And to assess curvature, a CCD design is commonly used. Extending the architecture of

these factorial points to the mixture setting by adding and subtracting $+\Delta_1$ or $-\delta_2$ and Δ_2 or $-\delta_1$ where 1 and 2 are the components that are to have the interaction estimated to the mixture setting, the authors use the Cox effect directions with which to add and subtract these amounts and the standard mixture to estimate the interaction.

In *Experiments with Mixtures* by John Cornell (2000), a section of Chapter 5 is devoted to the concept of screening in a mixture setting. Cornell defines screening as when there are six or more components and "if it becomes necessary to perform experimental runs and decide on the most important components from the sizes of their effects, then the reduction of the number of components so that only the most important components are considered further is known as screening the components." He then states that the "construction of screening designs and the setting up of screening models quite often begin with the Scheffè first-degree model." Cornell recommends setting the ranges of the different components as close as possible to each other so that the relative effects of the components can then be studied by taking the ratio of the effect estimate to its standard error. If the ranges of the components are similar then one can just look at the magnitude of the beta estimates to infer the larger effect. Cornell goes on to define a general component effect since he states that in order to screen out the unimportant components it is necessary to know how to measure the effects of the individual components. He defines the effect of a component as the definition given in Snee and Marquardt and shows their effect testing procedure. Lastly he gives a recommendation that when screening, it is better to combine component proportions than to drop components from the model. This section is followed by a seven-component octane-blending experiment where he illustrates this procedure of combining proportions.

In a Scheffè mixture model the parameters represent the heights of the surface over the i^{th} pure component vertex. If the experimental region is constrained, the parameters are the extrapolated height over the pure component vertex (or the height over the pseudocomponent vertex if possible to use pseudocomponents) even though a pure component was never contained in the experimentation.

When an experimenter adds a component to a mixture it would be useful to be able to directly quantify the changes in the response from the addition of that component. Using the Scheffè mixture model it is impossible to tell whether the response changed

from the addition of the new component or the decrease in one of the other components. Additionally the Scheffè model provides no way to incorporate the current mixture into the model, to account for improvements made by experimentation. Further, there is no clear way to address curvature or the quadratic mixture effects using the Scheffè model.

In a Cox mixture model (discussed further in Chapter 2), the parameters represent the slope between a mixture point x relative to the response at a standard mixture. We believe that in a screening situation in an industrial setting, there will often be some standard or starting point from which to start screening. The Cox model provides a natural way to incorporate this current formulation into the experiment and then to use this standard response as a basis for comparison when components are added to the mixture.

Possibly the most significant disadvantage of screening designs that use the Scheffé model is that these models require fitting (at least) a full linear model. This will require a large number of design points, especially when considering the large amount of replication required to account for the large variability inherent in many mixture applications. And complicated constraints on the design space make the interpretation of individual Scheffé parameters meaningless. In this case the fitted Scheffé model can only really be used to choose large effects or to build a complete model of the surface, not to gain understanding of individual component behavior on the response. It is not practical to assess more than linear effects when the number of mixture components is at all large which, in industrial settings, is most often the case. For smaller simpler mixtures, the current methodology works well. However, for large mixtures with complicated constraints, fitting a model to the entire region is not generally the most practical way to begin experimentation. Further, we assert that it is likely that there will be a current setting, a standard mixture, a reliable starting point in this type of experimentation. We also assert that it will not always be necessary to study the effect of all the components. We illustrate examples of this in Chapter 3.

Thinking about screening an industrial mixture setting does not involve finding the most important components by quantifying the largest linear effects, but by measuring the changes in the response of interest at the new point and comparing that response to the current accepted response. We develop screening methods not for extensive laboratory studies of a mixture and its response properties, but for screening in a production setting,

where studying all components and all properties is often not necessary to answer a specific question. We believe that the framework in the Cox mixture model form is the most useful for screening a mixture in an industrial setting and this will the focus of Chapters 3 and 4.

Chapter 2

The Cox Mixture Model

The Cox [1] polynomial model provides an alternative to the Scheffé model. The Cox model is developed to give individual parameters an interpretation. Owing to the inherent restriction ($\sum x_i = 1$) in mixture experiments Cox places a restriction on the parameter estimates to account for the redundancy in the parameters. The parameters, β_i, in a Cox mixture model represent slopes. Each β_i is the relative change in the measured response at point x to the measured response at a standard mixture we define as s, where

$$s = (s_1, s_2, s_3, ..., s_q). \tag{2.1}$$

Many take s to be the centroid of the region, but that is not necessary. For example, say we add a small amount Δ_q to one of the q components in our standard mixture and we will call that new point x. But we are going to add this small amount in such a manner that the ratio's of the other i components in s remain constant, but the ratio of component q to the other component will change. In this manner, we are able to study what Cornell [22] defines as a component effect. The new point x lies on the ray that connects the vertex of our component q with our original point s (2.1).
In mathematical terms, we define the point x as

$$x_q = s_q + \Delta_q \tag{2.2}$$

and the other $q - 1$ components in the mixture change according to their proportion in s.

$$x_i = s_i - \frac{\Delta_q s_i}{1 - s_q}.$$

To begin to formulate a model for our small experiment, we define a standard first degree polynomial as

$$\eta_1(x) = \beta_o + \sum_{i=1}^{q} \beta_i x_i.$$

The change in the response due to the addition of Δ_q is expressed as the difference between the response at x and s:

$$\Delta\eta_1(x) = \eta_1(x) - \eta_1(s).$$

Substituting (2.1) and (2.2)

$$\Delta\eta_1(x) = \beta_o + \sum_{i=1}^{q-1} \beta_i (s_i - \frac{\Delta_q s_i}{1 - s_q}) + \beta_q (s_q + \Delta_q) - \beta_o - \sum_{i=1}^{q} \beta_i s_i. \tag{2.3}$$

Canceling terms and utilizing the restriction on the parameters,

$$\sum_{i=1}^{q} \beta_i s_i = 0, \tag{2.4}$$

(2.3) reduces to

$$\Delta\eta_1(x) = \beta_q (s_q + \Delta_q) + \sum_{i=1}^{q-1} \beta_i (s_i - \frac{\Delta_q s_i}{1 - s_q})$$

$$= \beta_q s_q + \beta_q \Delta_q + \sum_{i=1}^{q-1} \beta_i s_i - \sum_{i=1}^{q-1} \beta_i s_i \frac{\Delta_q}{1 - s_q}.$$

Combining terms applying the restriction (2.4),

$$\Delta\eta_1(x) = \beta_q \Delta_q - \sum_{i=1}^{q-1} \beta_i s_i \frac{\Delta_q}{1 - s_q}.$$

Adding and subtracting we get

$$\Delta\eta_1(x) = \beta_q\Delta_q - \sum_{i=1}^{q-1}\beta_i s_i\frac{\Delta_q}{1-s_q} + \beta_q s_q\frac{\Delta_q}{1-s_q} - \beta_q s_q\frac{\Delta_i}{1-s_i},$$

and once more applying the constraint (2.4) we get

$$\Delta\eta_1(x) = \beta_q\Delta_q - \sum_{i=1}^{i}\beta_i s_i\frac{\Delta_q}{1-s_q} + \beta_q s_q\frac{\Delta_q}{1-s_q}$$

$$= \frac{\beta_q\Delta_q}{1-s_q} - \sum_{i=1}^{q}\beta_i s_i\frac{\Delta_q}{1-s_q}.$$

Finally we define the change in the expected response as

$$\eta_1(x) = \frac{\beta_q\Delta_q}{1-s_q}. \tag{2.5}$$

In particular this implies that the intercept β_o is the response at the standard point s. An estimate of the effect of component q is calculated as:

$$\beta_q = \frac{\Delta\eta_1(x)(1-s_q)}{\Delta_q},$$

where $\Delta\eta_1(x) = E[y(x) - y(s)]$ and where $y(x)$ and $y(s)$ are the observed responses at x and s. Therefore each β value is the difference in the heights of the surface at x and s weighted by the incremental change Δ_q made in the proportion of component q relative to the amount $(1 - s_q)$.

If the surface is better represented by a quadratic surface Cox's model contains an additional term. The second degree polynomial is defined by

$$\eta_2(x) = \beta_o + \sum_{i=1}^{q}\beta_i x_i + \sum_{i=1}^{q}\sum_{j=1}^{q}\beta_{ij}x_i x_j,$$

where $\beta_{ij} = \beta_{ji}$. To find the change in the response between our points x and s we write the change in the expected response as:

$$\Delta\eta_2(x) = \eta_2(x) - \eta_2(s)$$

or substituting

$$\Delta\eta_2(x) = \sum_{i=1}^{q} \beta_i x_i + \sum_{i=1}^{q}\sum_{j=1}^{q} \beta_{ij} x_i x_j - \sum_{i=1}^{q} \beta_i s_i - \sum_{i=1}^{q}\sum_{j=1}^{q} \beta_{ij} s_i s_j; \tag{2.6}$$

rearranging the terms in (2.6),

$$\Delta\eta_2(x) = \sum_{i=1}^{q} \beta_i x_i - \sum_{i=1}^{q} \beta_i s_i \tag{2.7}$$

$$\sum_{i=1}^{q}\sum_{j=1}^{q} \beta_{ij} x_i x_j - \sum_{i=1}^{q}\sum_{j=1}^{q} \beta_{ij} s_i s_j. \tag{2.7a}$$

Using the introduced restriction (2.4) and a new restriction $\sum_{j=1}^{q} \beta_{ij} s_j = 0$ for $i = 1, .., q$ (Note: $\sum_{i=1}^{q} \beta_{ij} s_i = 0$ for $j = 1, .., q$ as well), we develop the quadratic Cox model. We know that (2.7) is equal to $\Delta\eta_1(x)$ as shown previously. Rearranging (2.7a) we get

$$= \sum_{i=1}^{q-1}\sum_{j=1}^{q-1} \beta_{ij} x_i x_j + 2\sum_{i=1}^{q-1} \beta_{iq} x_i x_q + \beta_{qq} x_q^2 - \sum_{i=1}^{q}\sum_{j=1}^{q} \beta_{ij} s_i s_j.$$

Substituting the relations for x_q (2.2) and x_i (2) and letting

$$a = \frac{\Delta_q}{1 - s_q},$$

is equivalent to

$$= \sum_{i=1}^{q-1}\sum_{j=1}^{q-1} \beta_{ij} s_i s_j (1-a)^2 + 2\sum_{i=1}^{q-1} \beta_{iq} s_i (1-a)(s_q + \Delta_q) + \beta_{qq}(s_q + \Delta_q)^2 - \sum_{i=1}^{q}\sum_{j=1}^{q} \beta_{ij} s_i s_j,$$

$$\begin{aligned} &= \sum_{i=1}^{q-1}\sum_{j=1}^{q-1} \beta_{ij} s_i s_j + (-2a + a^2)\sum_{i=1}^{q-1}\sum_{j=1}^{q-1} \beta_{ij} s_i s_j + 2\sum_{i=1}^{q-1} \beta_{iq} s_i s_q \\ &+ 2(\Delta_q - a s_q - a\Delta_q)\sum_{i=1}^{q-1} \beta_{iq} s_i + \beta_{qq} s_q^2 + \beta_{qq}(2\Delta_q s_q + \Delta_q^2) - \sum_{i=1}^{q}\sum_{j=1}^{q} \beta_{ij} s_i s_j. \end{aligned}$$

$$
\begin{aligned}
&= (-2a + a^2)\sum_{i=1}^{q-1} s_i \sum_{i=1}^{q-1} \beta_{ij} s_j + 2(\Delta - a s_q - a\Delta)\sum_{i=1}^{q-1} \beta_{ij} s_i + \beta_{qq}(2\Delta_q s_q + \Delta^2 \\
&+ (-2a + a^2)\sum_{i=1}^{q-1} \beta_{iq} s_i s_q + (2a - a^2)\sum_{i=1}^{q-1} \beta_{iq} s_i s_q + 2(\Delta_q - a s_q - a\Delta_q)\beta_{qq} s_q \\
&- 2(\Delta_q - a s_q - a\Delta_q)\beta_{qq} s_q
\end{aligned}
$$

$$= \beta_{qq}(2\Delta_q s_q + \Delta^2) - (2a - a^2)\beta_{qq} s_q^2 - 2(\Delta_q - a s_q - a\Delta_q)\beta_{qq} s_q$$

$$= \beta_{qq}[2\Delta_q s_q + \Delta_q^2 - 2a s_q^2 + a^2 s_q^2 - 2\Delta_q s_q + 2\Delta_q a s_q^2 + 2a\Delta_q].$$

And finally,

$$\Delta\eta_2(x) = \beta_{qq}[\Delta_q^2 + 2a\Delta_q s_q + a^2 s_q^2]. \tag{2.8}$$

Substituting the value of a in (2.8) and adding this term onto the linear Cox model we get the following expression for the expected response.

$$\Delta\eta_2(x) = \beta_q \frac{\Delta_q}{1 - s_q} + \beta_{qq} \frac{\Delta_q^2}{(1 - s_q)^2}$$

The Cox polynomials can be fit using constrained least squares where the constraints are the matrix of linear restrictions on the parameters. Or the coefficients can be calculated from the fitted Scheffè model utilizing relations as shown by Cornell [31]. Smith and Beverly [32] provide Fortran code to generate the linear and quadratic Cox model from Scheffè models with up to $q = 10$. Design Expert and JMP 8 software have the ability to fit a Cox Model to a mixture experiment directly using the Scheffè parameter estimates.

We believe that the elegance of the interpretation of the Cox model coefficients gives an advantage to experimenters when screening in an industrial setting. Cox [1] makes the following three points regarding the Scheffè polynomial.

1. If two replicated experiments on the same system have the same expected responses, except for a constant difference between replicates, it is obvious that on fitting Scheffè linear or quadratic models to the separate replicates all of the parameters β_i appear different in the two replicates.

2. It is meaningless to consider the direction and magnitude of the curvature of the response to a particular component, due to the absence of squared terms from the quadratic model.

3. The interpretation of the Scheffé model parameters is in terms of responses to very simple mixtures. The main interest, however, may be the relative behavior of quite complicated mixtures and indeed the region of experimentation may well exclude mixtures with only a few components present.

It should be noted that Cox directly addresses these three shortcomings with his model. First, any constant replicate effect or blocking effect is absorbed into β_0 and does not affect the other parameter estimates. Secondly, the parameter β_i is the slope of the response surface for changes made in the i^{th} component which provides information about the response surface. Cox also includes pure quadratic terms in his model to directly assess the direction and magnitude of the curvature of the response surface. Lastly, by starting with a standard mixture and moving outward, the "design" is not defined by the complexities of a mixture, thereby allowing experimentation on very complicated mixtures.

In light of Cox's points, if the experimental goal is optimization, this interpretation does not matter; Scheffeè's equations map the surface fine. However, if your goal is to study the effects of components in the mixture, this interpretation is hard to understand. As a result, some practitioners have chosen to give up all ability to interpret the components at all and throw it into the "black box" of PLS analysis Kettaneh-Wald [33], Muteki and MacGregor [34], Johansson et. al. [35]. In some of the fore mentioned references, the practitioners recognize that the Cox mixture model is superior to the Scheffè model but they use PLS to analyze the Cox model, too. Maybe this is due to the availability of literature on how to utilize the Cox mixture model for experimentation or analysis. In that respect, we examine a simple example to illustrate the practicality of the Cox mixture model.

Take a standard mixture s, not necessarily $s_1 = s_2 = s_q$. Suppose an experimenter wishes to add an amount Δ_q to component i and at the same time take away Δ_1 from the first component. In that manner, only the amounts of the first and last components change and the others remain in exactly the same amount (and proportion to each other) as in s. We define this new point $\omega = (s_1 - \Delta_1, s_2, s_3, ..., s_q + \Delta_q)$ where $\Delta_1 = \Delta_q$ so that the points in ω add to unity. Let's derive then the change in the expected response between s and our new point ω using the linear Cox model. The expected change in response can be represented as

$$\Delta\eta_1(\omega) = \eta_1(\omega) - \eta_1(s).$$

Substituting for $\eta_1(s)$ and $\eta_1(\omega)$ we get

$$\Delta\eta_1(\omega) = \beta_o + \beta_1(s_1 - \Delta_1) + \beta_q(s_q + \Delta_q) + \sum_{i=1}^{q-2} \beta_i s_i - \beta_o - \sum_{i=1}^{q} \beta_i s_i$$

or, simplifying,

$$\Delta\eta_1(\omega) = \beta_q\Delta_q - \beta_1\Delta_1 = (\beta_q - \beta_1)\Delta. \tag{2.9}$$

The expression (2.9) illustrates the simple interpretation of the Cox model coefficients. When we add and take away an amount Δ from each of s_i and s_1 the change in the response from s to my new point ω is the difference in the slopes of each individual component weighted by the change Δ. If β_i is much larger than β_1, adding more of component i is going to increase the response, regardless if I take away from component 1. If the difference is close to zero, then neither taking away from one or adding to the other is going to change the response much from the standard s, and the surface is "flat" in that direction. This type of simple, intuitive thinking of the response surface in a mixture experiment is not possible using the Scheffè parametrization.

The β_q and the β_{qq} in the Cox model show the experimenter the shape and direction of the response in the effect direction of component q when moving away from the standard mixture s. In this way the experimenter is allowed a view of the response surface in any direction desired. Additionally, there is not a need to experiment over the entire region if

the interest lies only in a certain component or a few components.

In the following chapters, we develop screening methods for industrial mixture experiments utilizing the simple intuitive interpretation of the Cox mixture model.

Chapter 3

Screening in a Mixture Setting

3.1 Introduction

The Cox model provides the experimenter information about what will happen to a response when an incremental change is made in any direction from some current mixture. This type of insight can not be had using a Scheffé model since there is no direct way to incorporate the current mixture into the model and the parameters of the model do not describe change.

We can illustrate this insight using a simple mixture of three components where the components can vary as $0.4 \leq A \leq 0.7$, $0.1 \leq B \leq 0.4$ and $0.2 \leq C \leq 0.5$. We simulated data for an augmented simplex lattice design (vertices plus interior points) Estimates for the Scheffé (in terms of L-pseudo components)

$$\hat{y}(x) = 8.29A' + 9.62B' + 10.29C' \tag{3.1}$$

where $A' = A - 0.4/(1-0.7)$, $B' = B - 0.1/(1-0.7)$ and, $C' = C - 0.2/(1-0.7)$. Estimates for the Cox model with $s = (0.5, 0.2, 0.3)$ are

$$\hat{y}(x) = 9.40 - 2.89A + 1.56B + 3.78C \tag{3.2}$$

From (3.1) we see that the highest response will be at the pseudo component vertex of component C. This is also readily apparent from the fitted Cox model in (3.2). However, we

know that at s the response to our mixture is 9.40 and that adding some of B, some of C and taking away component A will increase the response. Using the Scheffé model estimates, this type of insight is not as easily had. Some find the fact that the Cox model is specific to a standard mixture a disadvantage, we believe that the ability to directly compare changes made with a current product formulation is a more efficient way to experiment in an industrial setting.

After discussing mixture experimentation with practitioners who conducted industrial experiments, it became obvious that there are a few consistent types of experimentation. The first type of experimentation is adding a component to a mixture in order to enhance a certain property of the current mixture. For example, a practitioner could be investigating adding a component to paint to speed the drying process. It would not be necessary to explore all of the components in the paint plus the additive, but just the effect of the additive and maybe certain other components. In this case, there is a starting point, the current paint formulation and a particular subset of components of interest. Similarly a company could be considering replacing a current component in the paint with a cheaper one. Again, there would be no need to quantify all the individual components in the paint and the new additive, but a need to compare the performance of the new additive to the current additive using the difference or similarity of the responses. Conversely the goal might be to remove a component from a mixture to either change the response properties, make the product cheaper, or change the properties altogether to create a new product. Additionally it might be of interest to screen a subset of components, i.e., the active ingredients in a drug, the expensive ingredients in a mixture, or the salts in a soft drink. These discussions led us to define our concept of screening a mixture in an industrial setting. The concepts are as follows:

1. The results of the each step in experimentation determine the next step in the experimental process; experimentation is sequential.

2. There is a current mixture or standard mixture from which to begin experimentation.

3. The initial experimental goal is not to build a complete model of the response surface, but to understand enough about effects to improve the mixture.

4. There is a large number of mixture components, $q \geq 6$, with any number of constraints.

5. The experimenter wishes to quantify the effect of individual components, either one or a few, but not necessarily all of the components in the mixture.

Using the criteria above, we propose using the Cox mixture model form and corresponding designs points for screening in a mixture setting. The Cox model coefficients have a meaningful interpretation and provide a nice framework to begin experimentation from a standard mixture. Additionally, we contend that it is not necessary to initially collect enough data to model the entire mixture surface.

3.2 Sequential Screening Method, Adding a Component

A subset of our general method (to be discussed) is for a particular situation where the experimenter wishes to add a component to a mixture, either to change or enhance the response properties. The current methodology would have the experimenter build the entire linear Scheffé model of the surface including the new component. However, this experiment could consist of many experimental runs and replicates and possibly include mixtures where the new component is not present. In this situation, especially when the number of components is large, it would be more efficient and economical to quantify the effect of the new component using the method about to be described. We describe a sequential approach, the next step in experimentation is based on the results of the previous steps.

Prior to developing any effect estimates we need to introduce a model for error. We will assume that

$$y_{ij}(x_i) = \mu(x_i) + \epsilon_{ij} \tag{3.3}$$

where ϵ_{ij} are i.i.d. $N(0, \sigma^2)$ and $j = 1, 2, ..r$. This is the error model assumed for all estimates presented in the next two Sections, unless otherwise stated.

3.2.1 Step 1

Let the current mixture be defined as $s = (s_1, s_2, ..., s_q, 0)$ where the $(q+1)^{st}$ component is the new component to be added and its value in the starting mixture is zero. To quantify the effect of the new component, add an amount Δ_{q+1} to the $(q+1)^{st}$ component and decrease the subsequent components by an amount $\Delta_{q+1}s_i$ to make a second design point (including s):

$$x_{q+1} = \Delta_{q+1}$$

$$x_i = s_i - \Delta_{q+1}s_i \quad \text{for } i = 1 \text{ to } q$$

With the two design points (s and x) replicated r times, the effect of the addition of the new component can be determined. Note that each replicate contains an experimental run using mixtures s and x, in a random order within each replicate. Also note that, in the Cox model, any blocking effects are absorbed into the intercept term, which we do not use to evaluate whether the addition has improved the mixture. We could again express the change in the response using the first degree polynomial as

$$\Delta\eta(x) = \eta(x) - \eta(s).$$

Evaluating these responses gives

$$\Delta\eta(x) = \beta_o + \sum_{i=1}^{q+1} \beta_i x_i - \beta_o - \sum_{i=1}^{q} \beta_i s_i.$$

Substituting the design points x and s gives the change in the response as

$$\Delta\eta(x) = \beta_o + \sum_{i=1}^{q} \beta_i(s_i - \Delta_{q+1}s_i) + \beta_{q+1}\Delta_{q+1} - \beta_o - \sum_{i=1}^{q} \beta_i s_i.$$

Simplifying and applying the constraint $\sum_{i=1}^{q} \beta_i s_i = 0$,

$$\Delta\eta(x) = \sum_{i=1}^{q} \beta_i s_i - \sum_{i=1}^{q} \beta_i s_i \Delta_{q+1} + \beta_{q+1}\Delta_{q+1} - \sum_{i=1}^{q} \beta_i s_i$$

$$= -\sum_{i=1}^{q} \beta_i s_i \Delta_{q+1} + \beta_{q+1}\Delta_{q+1}.$$

$$= \beta_{q+1}\Delta_{q+1} - \sum_{i=1}^{q} \beta_i s_i \Delta_{q+1}$$

Once more applying the constraint $\sum_{i=1}^{q} \beta_i s_i = 0$ gives the expression for the change in the expected response from our standard point s and our new point x

$$\Delta\eta(x) = \beta_{q+1}\Delta_{q+1}.$$

The parameter estimate is calculated as,

$$b_{q+1} = \frac{1}{\Delta_{q+1}}[y(x) - y(s)],$$

with variance,

$$Var(b_{q+1}) = \frac{2}{\Delta_{q+1}^2}\frac{\sigma^2}{r}$$

where r is the number of replications done at s and x and $Cov[y(x)y(s)] = 0$.
The parameter estimate is the slope of the response surface relative to the starting point, s.

The results of this addition can lead to several conclusions. If the estimated slope is a statistically significant positive number, our recommendation would be to make the mixture containing Δ_{q+1} of the new component and make another addition. If the addition is the same size as the first addition, then the quadratic effect may be estimated using equation (3.5). However, if the initial estimated slope is not statistically significant, but the variance is large, we would recommend more replications at s and at the new point to be sure the effect is not being masked by the variance of the estimate.

3.2.2 Step 2

If the first addition of Δ_{q+1} shows promise we recommend adding $2\Delta_{q+1}$ to the starting mixture to create a second design point $(x_1, x_2, ..., x_q)$ with components defined by

$$x_{2,q+1} = 2\Delta_{q+1}$$

$$x_i = s_i - 2\Delta_{q+1}s_i \quad \text{for } i = 1 \text{ to } q.$$

Using this new design point, the original standard mixture and the point made from the first addition of the new component, we can estimate the quadratic effect β_{q+1q+1}. To do this, take the point from the single addition and set that as s_{new}. The original mixture, s, becomes the $y(s_{new} - \Delta_{q+1})$ point and the new point made from the $2\Delta_{q+1}$ addition becomes the point $y(s_{new} + \Delta_{q+1})$. The estimate of a quadratic effect of an individual component moved along its Cox direction is shown by Cornell [31] as

$$b_{ii} = (1 - s_i)^2)[\frac{y(s + \Delta_i) + y(s - \Delta_i) - 2y(s)}{2\Delta_i^2} \tag{3.4}$$

were the response values $y()$ are averages. Adapting Cornell's estimate in equation (3.4) for when an experimenter is adding a new component to a mixture as described, the estimate for the quadratic effect of the $(q+1)^{st}$ component is

$$b_{q+1q+1} = \frac{(1 - \Delta_{q+1})^2}{\Delta_{q+1}^2}[y(s_{new} + \Delta_{q+1}) + y(s_{new} - \Delta_{q+1}) - 2y(s_{new})] \tag{3.5}$$

where the response values are averages.

The variance of this estimate is

$$Var(b_{q+1q+1}) = \frac{4(1 - \Delta_{q+1})^4 3\sigma^2}{4r\Delta_{q+1}^4}.$$

(Note the restrictions for this estimate are different from the linear estimate since $s = s_{new}$). This quadratic effect will give the experimenter information about the shape of the response surface in the direction from the edge where the component was zero out to the point where $2\Delta_{q+1}$ is in the mixture. The sign and magnitude of this estimate will allow the experimenter to gain knowledge about the shape of the response surface. Again, if the variance is large, we recommend taking a few more replications at s_{new} and the $2\Delta_{q+1}$ point. At this step in the experimentation, the addition of the new component can be

Table 3.1: Design Points for Adding a Component Example

Design Point	Response Value
0.5, 0.5, 0	35.8
0.4, 0.4, 0.2	32.7
0.30, 0.30, 0.40	29.6

quantified further by using the general screening method (see Section 3.4) to explore the surface around this component. This addition method is especially valuable if the current mixture contains many components (i.e. $q \geq 30$). The variance of this estimate, like all mixture model estimates is fairly large, so a substantial number of replications will be required. This will be discussed further in Section 6.

3.2.3 Example of Adding a Component

We illustrate the interpretation of the Cox parameter estimates using the Strawberry Mite Experimental Data from Cornell (p. 297) [22]. Our intent is to illustrate the interpretation of the parameter estimate when we are experimenting close to a standard mixture. Of course, our method was developed primarily for use with a complex mixture with a large number of components, but this small example suffices to illustrate our method. In the original example, the goal of the experiment was to observe the mite numbers corresponding to a simplex design in order to model the response surface. After that experiment was completed, it was suspected that component x_3 had an additive effect with x_1 and x_2 and eight additional formulations were run to explore this additive relationship. For each mixture, the average number of the mites on five leaves of each of 10 plants was the response. We use this data assuming we are starting with a different problem and experimental goal. We assume that we had a mixture of components x_1 and x_2 and wished to investigate the addition of component x_3. The original example contained the points in Table 3.1 (plus other) design points.

We assume that we start with a mixture of two components with the experimental goal of exploring the response to the addition of component 3 to the first design point in Table 3.1. If we take $s = (0.5, 0.5, 0)$ and make two additions of $\Delta_3 = 0.2$ of x_3 we get design points 2 and 3 in Table 3.1. Using the design points and the responses from the original

data, we can calculate the following slopes.

$$b_3 = [32.7 - 35.8]/0.2 = -15.5$$

relative to $s = (0.5, 0.5, 0)$

$$b_{33} = ((0.8)^2/2(0.2)^2) * [29.6 + 35.8 - 2 * 32.7] = 0$$

relative to $s = (0.4, 0.4, 0.2)$.

Adding x_3 starting from s will decrease the response. The estimate $b_{33} = 0$ shows that the surface is linear in the direction of increasing and decreasing component x_3 with $s = (0.4, 0.4, 0.2)$. We reached the same conclusion as the original experimenters using 3 unique data points and their associated replications. We conclude that there is an additive effect when adding x_3 and holding x_1 and x_2 constant. Now that the necessary information has been gained, the next step in the experimental process can begin. This might be to model the surface in the area of interest, do some confirmatory runs or continue to add x_3 to find an optimum.

Recall that in the original experiment the mites were studied on five leaves of 10 plants for each mixture and the response is the average mite count per treatment combination. In this example the number of components was very small, but when the component number is large the run savings using our method vs. building a linear model (as the current methods define) in order to quantify the effect of adding a component is going to be significant. Our method allows for ample replication and the flexibility to do a small number of runs when the number of components is large.

3.2.4 Adding a Component using a Pre-Mix

In the adding a component method just described, we assumed that for each replicate the experimenter would be making a new mixture of s and a new mixture of x. Another way to perform this type of experimentation is to start with a pre-mix of the standard mixture s, then use smaller samples of the pre-mix for each replicate. For one replicate the experimenter would obtain a suitable amount of the pre-mixed standard s, measure the

response for that smaller amount, then add an amount Δ_{q+1} of the new component to this same sample of s and again measure the response. In this manner, the experimenter would not have to make a new mixture of s for each replicate, but simply make an addition to an already made mixture, saving experimental time and increasing the precision of the effect estimate.

The error structure defined in equation (3.3) would change if the experiment was performed using a pre-mix. We can think of the error associated with the observed y as two pieces, one due to constructing the pre-mix and another piece due to experimental error (e.g., measurement, adding a component). We can then rewrite the error model as

$$y_{ij}(x_i) = \mu(x_i) + \epsilon'' + \epsilon'_{ij}. \tag{3.6}$$

where $\epsilon'' + \epsilon'_{ij} \sim N(0, \sigma_\epsilon{\prime\prime}^2 + \sigma_\epsilon{\prime}^2)$. If the experiment is performed as descried above, there is no need to consider the component due to constructing the standard, ϵ''. Thus the error is potentially reduced, assuming $\sigma^2 > \sigma_\epsilon{\prime}^2$ Hence, the precision for estimating the effect of the addition is increased, i.e. σ^2 in equation (3.3) is replaced by $\sigma_\epsilon{\prime}^2$.

3.3 Removing a Component Sequential Screening Method

Another unique screening situation in a mixture setting is when an experimenter wants to quantify the effect of removing a component. The goal might be to completely remove the component from the mixture, to change the response properties, to lessen the amount of an expensive component, or to remove one of the inert ingredients in a mixture for savings. To perform this experiment we define $s = (s_1, s_2, ..., s_q)$ as the current formulation and set $\Delta_q = s_q$. Note that we are assuming that s_q is a small part of the total mixture; and with a mixture containing many components this will typically be the case. Using the same method as in adding a component, the second design point $x = (x_1, x_2, ..., x_q$ is defined as

$$x_q = 0$$

$$x_i = s_i + \frac{\Delta_q s_i}{(1 - s_q)} \quad \text{for } i = 1 \text{ to } q - 1$$

The parameter estimate with restriction $\sum_{i=1}^{q} \beta_i s_i = 0$ for q components is

$$b_q = \frac{(1 - s_q)}{\Delta_q}[y(s) - y(x)],$$

with variance:

$$Var(b_q) = \frac{2(1 - s_q)^2}{\Delta_q^2}\frac{\sigma^2}{r} \tag{3.7}$$

Again, when the number of mixture components is large, this method of determining the effect of removing a component on the response will have a significant run savings over fitting a Scheffé linear model to the entire surface to assess the component effect.

Similarly as with the adding a component method, we assumed that the experimenter will make a new mixture of s and x for each replicate. However, it is possible to perform the removing a component method using a pre-mix as described in Section 3.2.4 to gain precision when estimating the effect of removing a component. In this case however, the pre-mix would be the mixture x which does not contain component q. Then for each replicate, the experimenter could add component q to the x mixture. If this were done, then the error model is defined as equation (3.6) and the σ^2 in equation (3.7) should be replaced by $\sigma_\epsilon I^2$.

3.4 General Screening Method

We will again use the Cox mixture model framework to estimate individual parameter effects. Like our previous two methods, the general screening method is built around a starting mixture defined as $s = (s_1, ..., s_q)$. To efficiently explore the mixture component properties for a subset of components, we propose simple additions and subtractions of an amount Δ around the starting mixture s. Our method is the start of sequential experimentation, the starting experiment and a number of subsequent experimental steps.

When experimenting in a mixture setting there will always be some confounding, as moving one component necessitates moving the rest of the components. By nature of mixture experimentation and especially when utilizing Scheffé model parameters (particularly in a constrained region) it is hard to determine whether the change in the response was due to an increase in one component or the corresponding decrease in another component.

Knowing which movement changed the response (increase in one or decrease in the others) leads to the next step in experimentation. Our strategy begins with deliberate movement in two components, this way there is no confusion as to which component changes caused the changes in the response.

The general screening method assumes that a subset of components are of interest. That subset can be chosen either around a particular component of interest or by selecting a group of components (i.e., the active ingredients or the drying agents) in a large mixture. We start at the current formulation s and add and subtract amount Δ to pairs of components, $(s_1, s_2, ..., s_{i-1}, s_i + \Delta, s_{i+1}, ..., s_{q-1}, s_q - \Delta)$. We define each replicate of this experiment as containing the standard mixture and each of the mixtures after the additions and subtractions have been made, with in each replicate the run order should be completely randomized. Again, there is no need to consider a blocking effect since the estimates are differences.

The change in the response when moving from s to x can be expressed using a first order polynomial function:

$$\Delta\eta_1(x) = \eta_1(x) - \eta_1(s).$$

Substituting for $\eta_1(s)$ and $\eta_1(x)$ we get

$$\Delta\eta_1(x) = \beta_o + \beta_q(s_q - \Delta) + \beta_i(s_i + \Delta) + \sum_{j=1}^{q-2} \beta_j s_j - \beta_o - \sum_{i=1}^{q} \beta_i s_i.$$

Then simplifying we obtain an expression for the change in the response:

$$\Delta\eta_1(x) = \beta_i\Delta - \beta_q\Delta = (\beta_i - \beta_q)\Delta$$

$$\Delta\eta(x) = E[y(x) - y(s)] = (\beta_i - \beta_q)\Delta.$$

With a simple calculation the estimate D_{iq} is calculated as

$$D_{iq} = (y(x) - y(s))/\Delta \tag{3.8}$$

Initially, like most screening methods, we estimate the linear effects, but we will use the

first-order Cox mixture model form (2.5). The variance of the estimate depends on Δ and r and takes the form:

$$Var(D_{iq}) = \frac{2}{\Delta^2}\frac{\sigma^2}{r}.$$

These estimates give simple interpretable meaning to the small movements made in components i and q. If D_{iq} is positive, then adding component i is going to increase the response when some of component q is taken away. If D_{iq} is negative, then removing component q and increasing component i will increase the response. If D_{iq} is close to zero, then either adding component i and taking away component q will not change the response.

Suppose that we wish to experiment on a mixture with many components. Also suppose that the experimenter is interested in studying behavior for a particular subset of components. One way to study this would be to fit the Scheffé linear model and, using the estimated response, calculate the response trace of that component over a range. The Cox effect direction is widely used for this type of analysis because of the useful interpretation of the parameters in the Cox model. However, using this method, many mixtures have to be tested to estimate the Scheffé model and this method assumes that the Scheffé model generated is adequate. Additionally, using the current screening methods, all components must be moved to obtain each design point, which is inefficient if the goal is to study a particular component or grouping of components. Another option would be to fit a slack-variable model using the inert or inactive ingredients as the slack variable. This approach suffers from the similar issues already stated regarding the interpretation of the Scheffé parameter estimates, especially when the region is highly constrained. Cornell [22] also shows that it does matter to the final model form, which components are selected to be the "slack" variables.

We also make the assumption that in most industrial mixture experiments there is a logical starting mixture, a standard mixture or a current formulation. There is no way to incorporate this mixture into the Scheffé model or the slack variable model. It seems logical that if there is a current formulation then all subsequent formulations will be compared to the response at the current standard.

For example, say we have a mixture with $q = 9$ components and the current mixture setting is at s. In addition, suppose we want to gain information about how certain

components (for example, the active ingredients) affect the response. In this example, assume that the active ingredients are components 5, 6, 7, 8 and 9. The next step depends on the goal of the experiment and the bounds on the components. For this example, assume that component 9 is the most expensive, so a decrease in that component would be advantageous. The set of runs we recommend is the following:

$$
\begin{aligned}
s &= (s_1, s_2, s_3, s_4, s_5, s_6, s_7, s_8, s_9) \\
x_1 &= (s_1, s_2, s_3, s_4, s_5 + \Delta, s_6, s_7, s_8, s_9 - \Delta) \\
x_2 &= (s_1, s_2, s_3, s_4, s_5, s_6 + \Delta, s_7, s_8, s_9 - \Delta) \\
x_3 &= (s_1, s_2, s_3, s_4, s_5, s_6, s_7 + \Delta, s_8, s_9 - \Delta) \\
x_4 &= (s_1, s_2, s_3, s_4, s_5, s_6, s_7, s_8 + \Delta, s_9 - \Delta)
\end{aligned}
$$

Using s and each other design point the following estimates can be directly calculated from the responses at each point and the response at s using the Cox linear mixture model.

$$
\begin{aligned}
y(x_1) - y(s) &= (b_5 - b_9)\Delta = D_{59}\Delta \\
y(x_2) - y(s) &= (b_6 - b_9)\Delta = D_{69}\Delta \\
y(x_3) - y(s) &= (b_7 - b_9)\Delta = D_{79}\Delta \\
y(x_4) - y(s) &= (b_8 - b_9)\Delta = D_{89}\Delta
\end{aligned}
$$

The size of Δ does not have to be constant between points; it can depend on the individual component constraints. However, it does need to be constant within points so that unity is maintained. Of course it could have been the case that at one point the component was at its upper bound so the design point would have to be $s_1, s_2, s_3, s_4, s_5, s_6 - \Delta, s_7, s_8, s_9 + \Delta$ and the corresponding estimate would be $y(x_2) - y(c) = (b_9 - b_6)\Delta$. Using the interpretations listed above, the experimenter now has interpretable coefficients to

explain how changes in components affect the response. For example, if estimate D_{59} is large and positive, then increasing component 5 while decreasing component 9 will increase the response. If D_{59} was negative then increasing component 5 while decreasing component 9 will decrease the response. Lastly, if D_{59} is small, then the changes in the components will not appreciably change the response. Although two components are completely confounded in their movements, there is no question about movement in which components caused the change in the response. No other design or analysis method in mixture experimentation can provide that insight. Even if each of the five components was moved an amount Δ along its Cox Effect direction, the other four components would have to be moved to create new point and that would require at least six design points. We believe the same information (which components affect the response, by how much and in which direction) is gained in the five points presented with one less experimental run.

We now return to our small example. Assume that estimates D_{59} and D_{79} were the largest, and they were positive, meaning that increasing components 5 and 7 while decreasing component 9 will increase the response. But which one is better? This will have to be worked out with a follow-up run. A good starting point would be the largest of the estimates. A comparison of the variances would also be a good indication about which component's addition would be more promising. At this point we believe that the experimenter will have enough insight to decide the next set of runs. For this next set of runs, any number of model choices, design choices can be used, depending on the experimental goal. Define the starting point as the best point from the initial experiment and then a steepest accent type method can be used to move to the desired response. With the knowledge gained about the effects of the individual components on the response, this type of methodology is possible.

An advantage of using this method is the clear interpretation of the component effects on the response. Another advantage is the flexibility this method allows for experimenting with some or all of the mixture components. Additionally, there is a large run savings over fitting a full Scheffé linear model.

This method can also be particularly useful if an experimenter wished to quantify the effects of a set of active ingredients, e.g., analgesics in a particular cold medicine. This could easily be done without moving other active ingredients, by pairing the active

ingredients movements with an inert ingredient. This way the movements are only in the active ingredients (assuming the binding properties are not a response in this case). This method is also particularly useful when the number of components is large, (e.g., $q \geq 30$) as it allows for any number of experimental runs. Building a model of the surface and running replications would require many design points, whereas if the goal is to study a particular part of the mixture, our method could provide the same information in fewer runs.

3.4.1 Simple Example of General Screening Method

We will illustrate the general screening method using the Strawberry Mite Data from Cornell (p. 297) [22] . Again, our intent is to illustrate the interpretation when we are experimenting near a standard. Even though we developed this method for a mixture with a large number of components and complex constraints, we will use this smaller example to illustrate the use of our methodology. Recall the goal of our first experiment was to specifically investigate the effect of x_3 with x_1 and x_2. In the original example, the experimenters added eight additional runs to a simplex design. For this example we use the response function generated by the original analysis to simulate the response values at our design points. Recall the response is the average number of mites on a plant after treatment. The fitted response function is

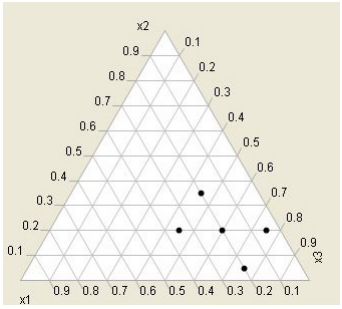
$$\hat{y}(x) = 49.70x_1 + 83.42x_2 + 20.21x_3 - 119.14x_1x_2 - 2.55x_1x_3 + 1.24x_2x_3 - 232.51x_1x_2x_3. \quad (3.9)$$

For this example we chose $s = (0.2, 0.2, 0.6)$, $\Delta = 0.15$, and simulated the corresponding responses for each design point using the model (3.9) with $N(0, 1)$ error added. For further discussion on the choice of Δ see Section 3.2. Table 3.2 is a table of the general screening method design points and a ternary plot of the points in Table 3.2 is shown in Figure 3.1. Using the response values, the estimates are:

Table 3.2: Design Points for General Screening Method Example

Design Point	1	2	3	Response Value
s	0.2	0.2	0.6	38.0
x_1	0.35	0.2	0.45	39.6
x_2	0.2	0.35	0.45	47.4
x_3	0.05	0.2	0.75	35.2
x_4	0.2	0.05	0.75	28.6

Figure 3.1: Design Points for General Screening Method for Strawberry Mite Example



$$D_{13} = [39.6 - 38.0]/0.15 = 1.6$$

$$D_{23} = [47.4 - 38.0]/0.15 = 9.4$$

$$D_{31} = [35.2 - 38.0]/0.15 = -2.8$$

$$D_{32} = [28.6 - 38.0]/0.15 = -9.4$$

From these four points we see that moving in the direction of decreased component 2 and increased component 3 will decrease the response the quickest. However, since increasing component 1 and decreasing component 3 might also decrease the response, that relationship can be further investigated, too.

3.4.2 The General Method using a Pre-Mix

The general method presented assumes that for each experimental run a new mixture would be fabricated. However, the general method can also be performed with a pre-mix, similar

to adding a component. For the general method, the pre-mix would be dependent upon the particular experiment. For example, assume we are experimenting with two active ingredients, s_1 and s_2, and one inert component s_q. We perform this experiment by adding Δ of components s_1 and s_2 and taking Δ away from ingredient s_q. This experiment would involve three unique mixtures: s, x_1 and x_2. If a pre-mix of the mixture $s_1, s_2, ..., s_q - \Delta$ is made ahead of time, then instead of making a new mixture for every experimental run, the experimenter can use samples of the pre-mix and make additions. The experimenter would construct the standard mixture s, from the pre-mix by adding Δ of component q. To construct the second mixture (x_1) where $s_1 + \Delta$ and $s_q - \Delta$, the experimenter would add Δ of component 1 to the pre-mix. For the third mixture (x_2) where $s_2 + \Delta$ and $s_q - \Delta$, the experimenter would add Δ of component 2. In this manner, the experimenter could save time and have increased precision for testing the D_{iq} effects. Assuming a partition of error variation as in equation (3.6),

$$Var(D_{iq}) = \frac{2}{\Delta^2}\frac{{\sigma_\epsilon}'^2}{r}.$$

3.5 The Choice of Δ

All of the variances presented depend on the size of Δ. The choice of Δ depends on several factors. First, Δ directly effects the variance of the estimates. The size of Δ is limited by possible constraints on individual components. Δ does not have to be the same for each separate run in the general screening method. However, it must be the same within each run to ensure that each mixture sums to unity. The choice of Δ is going to depend on many factors including engineering and scientific knowledge, but clearly choosing the maximum Δ possible for each run will result in smaller variance. The influence of Δ can be offset by the number of replicates done for each mixture. Since the method three methods presented provide a run savings over standard screening procedures, having an increase in the number of replicates is not as costly as running a standard screening design.

3.6 Conclusions

We have introduced a new approach for screening in a mixture setting. Our general screening method provides flexibility to study a subset of mixture components that is not available with the current screening methods. Because our method does not initially require fitting a full model to the surface, it can be utilized in any number of situations for any number of components. By not building a model of the entire surface we give the experimenter more runs for replication, which we have shown are necessary for the large variances inherent in parameter estimation in mixture experiments.

We have also explored two subset methods of the general screening method, adding and removing a mixture components. By choosing s and Δ accordingly, an experimenter can directly study the effects of a single component on the response to a mixture. We have given simple formulas for parameter estimates and variances for each of these cases and provided examples where these methods might be usefully employed.

We have shown that the variances of our estimates similar to the corresponding Scheffé estimate for the same value of n. However, the run savings we introduce by not having to build a model of the surface allows the practitioner more experimental runs for replication.

The methods introduced in this chapter have the practitioner in mind. These methods were developed considering complex, industrial mixtures with many components, and how to best begin experimentation with these large complex mixtures. Our methods are robust to complicated constraints, highly constrained mixtures and large amount of components. Our method does not require advanced knowledge of statistical methods or specialized statistical software to begin experimentation; it can be undertaken by practitioners at any level.

A final advantage to our approach is the ease with which it lends itself to the use of sequential methods of experimentation. As we will see in the next chapter, our methods fit very well with established sequential methods of experimentation.

Chapter 4

EVOP Method for General Screening Method

4.1 Evolutionary Operation

Evolutionary Operation (EVOP) is a statistical method for process improvement developed by George Box and Norman Draper [2]. The purpose of EVOP is to perform online, ongoing experimentation on an operating process by making small movements in a few factors to gain process understanding. Box and Draper refer to this process as "tuning". EVOP gets its name due to the sequential nature of this type of experimentation whereby the process "evolves" to its optimum. EVOP is designed for situations where changes in the individual variables are virtually undetectable in individual runs. In EVOP many replications of a design can be made. After enough cycles have been made to distinguish between the noise and the signals, conclusions are made and the next phase of experimentation is begun. Typically, the next phase will have different factor levels and possibly different factors, based on the knowledge gained from the first phase. In this aspect, the experimentation is evolutionary. Because this type of experimentation is done on a operating process, only small movements are made to ensure that product produced during EVOP time is of good quality. The authors state that "[t]o discover effects buried in noise we must improve the signal to noise ratio." [2] This can be done by either increasing the signal or decreasing "effective noise" level. Because EVOP is an online experimental procedure, increasing

the "signal" could produce poor quality, so the authors instead propose a decrease in the "noise". This is accomplished by taking advantage of the fact that sample averages have a reduced variance.

Box and Draper develop EVOP schemes for two and three variables based on simple full factorial designs. Even though an operating process has many variables, the authors recommend only experimenting with a few factors at one time. The experimental process is initiated by selecting the experimental variables and their corresponding levels. The levels are chosen such that production can continue during the time when EVOP is run, but changes are made such that eventually the effect of those movements can be quantified. Each replication of the factorial design is termed a cycle; however the replications are not done in a completely randomized manner. The full factorial, possibly including a center point is run in each cycle. Many cycles (replications) may be run in order to determine if there is a significant effect or not. Each time multiple cycles are completed in this manner, it is called a phase.

Box and Draper envision EVOP as a philosophical approach to plant operation and once started should never be stopped. Since the EVOP process is for ongoing plant operation, the authors have made "ease of use" an important factor in the choice of their methods [2]. In the original EVOP papers [36], the authors provide worksheets that can be used easily by any operator and without computer access (of course this research was completed when computers were not so widely available). For instance the authors use a simple range calculation to estimate the standard error of the effects as opposed to the summing of the squared differences and they do not use a t-distribution to test effects. They give the reason that typically they "do not use very small sample sizes" so that the normal distribution is appropriate. Additionally this assumption of normality makes the hand calculations simple. Lastly, the authors choose to present a confidence interval method for determining large effects based on the current phase standard deviations (as estimated by the range calculation previously mentioned). The authors originally wished to use a sequential testing method but decided against it because the calculations were too intensive for hand calculation ("[k]eeping tally on a host of sequential tests is laborious..." [2]). In addition, the authors thought that the choice of δ, the difference the test is designed to detect was impossible in an industrial setting. EVOP does not assume that just one

response is being studied and for each response, a separate sheet needs to be maintained during the phase. Therefore the task of performing the sequential testing by hand on multiple responses while deciding on how large a difference is large is not practical for quick hand calculations. However, with the advent of available computing, the sequential testing procedure is more appealing for use today.

Additionally, since this method is for use in ongoing plant operation, a check of the most favorable operating conditions, or a reference can be built into the cycles. Depending on the goal of the EVOP, or if a direction of improvement is known, the reference point can be one of the factorial points, or located in the center of the experimental region. Calculations are also made to determine the average process output during this cycle and the overall difference from the reference condition and the factorial conditions in order to evaluate product quality during this period of operation. It should be noted that "no attempt need be made to fit a response function formally" [2] during an EVOP procedure. The goal of the EVOP procedure is not to build a working model of the process but to gain information about individual process factor levels.

4.2 Similarities between EVOP, Mixture Experimentation and Screening

As mentioned, during periods of evolutionary operation small movements in a few factors are made repeatedly in order to systematically reduce the noise and strengthen the effect signal. In industrial mixture experimentation we assert that there is only capability to make small movements in mixture components since we are assuming the the number of mixture components is large. Additionally we assume that active ingredients (the ingredients of interest to be defined later) only make up a small fraction of the total. It is also inherent in mixture experimentation that there is large variation. In EVOP, the factor movements are kept small on purpose, but in experimenting with mixtures, the component movements are inherently small. Like an EVOP procedure then, the mixture effects have to be tested against large variation and therefore require many replications to separate the "signal" from the "noise". Many examples in Cornell's book [22] on mixture experiments use very high replication in order to assess mixture component effects.

EVOP is a procedure for examining a few process variables in order to understand the effects that these variables have on the process outputs. Like screening it is not necessarily for process optimization or model construction. In fact the estimation from an EVOP procedure only includes linear effects, and if possible, detection of curvature. In that sense EVOP is similar screening for effects. Specifically the EVOP methodology searches for effects in a sequential manner. Our description of screening in a mixture setting involves a subset of components, and similarly EVOP involves only a small number of process factors. With those similarities in mind, we develop a new method for practitioners to screen subsets of components in a mixture using an EVOP methodology of sequential experimentation.

4.3 Sequential Significance Tests

As mentioned, originally Box and Draper had considered sequential significance testing in their EVOP method. However, since they were developing this methodology for use in a time period where computers were not widely available they decided that the calculations for sequential significance testing were too intensive. Another drawback they mention is the choice of δ, the amount of change in the response that would be considered significant. This amount, δ, must be specified prior to beginning experimentation. In an online process situation with many variables to keep track of, and with all of the calculations for analysis being performed by hand, their decision was clearly justified. However, at this point there is no reason why sequential testing cannot be readily applied. We will provide a brief background for sequential significance testing and then demonstrate how to use this technique for adding a component and the general screening method.

Sequential hypothesis testing is based on the Neyman Person test criterion. The framework of the sequential test procedure we will use was first developed as the sequential probability ratio test by A. Wald [37]. Davies [38] gives a nice overview of sequential testing in Chapter 3 of his book Design and Analysis of Industrial Experiments. In sequential hypothesis testing "the number of observations is not fixed in advance but the test is applied to the accumulating data after each observation, the experiment being terminated as soon as a decision between alternate hypotheses can be made with the desired degree of certainty" [38]. The sequential test procedure computes the likelihood ratio with respect

to the null and alternative hypotheses and compares it against limits based on the associated risks of making the correct and incorrect test conclusions. To begin a sequential test procedure hypothesis test, we assume a simple hypothesis test such as:

$$H_0 : \theta = \theta_0 \tag{4.1}$$

$$H_1 : \theta = \theta_1$$

where, for example, θ could be the mean, μ and where σ is known. In general, however, the likelihood ratio is the likelihood of the sample when $\theta = \theta_0$ compared to the likelihood of the sample when $\theta = \theta_1$. If we took one observation, then the likelihood ratio would be

$$\lambda = \frac{f_0(y_1)}{f_1(y_1)} \tag{4.2}$$

If the ratio is large the null hypothesis is chosen. If the ratio is small the alternative hypothesis is chosen. In a sequential testing scheme (developed by Wald [37]) the null hypothesis is accepted if λ is greater than some value λ_0 and the alternative is accepted if λ is less than λ_1. If λ is between λ_1 and λ_0, then another observation is taken. We would like to choose the values λ_0 and λ_1 such that the risk of choosing H_0 when H_1 is true (type II error) is no larger than β and the risk of choosing H_1 when H_0 is true (type I error) will not exceed some value α. Wald [37] shows that bounds are

$$\lambda_0 \geq \frac{(1-\alpha)}{\beta}$$

and

$$\lambda_1 \leq \frac{\alpha}{(1-\beta)}.$$

If we choose $\lambda_o = \frac{1-\alpha}{\beta}$ and $\lambda_1 = \frac{\alpha}{1-\beta}$, we do not get exactly α and β as specified. However Wald [37] shows that the resulting α and β are not appreciably different than specified. Using a function of the likelihood ratio, this type of testing can be readily performed. Suppose the data follow a normal distribution with $\theta = \mu$ and σ is known. The likelihood ratio is a function of the sum of the observations. For example, if the hypothesis

are written as in equation (4.1) where $\theta = \mu$ then the likelihood ratio can be written as

$$\lambda = \frac{exp(-\sum_i (y_i - \mu_0)^2/2\sigma^2)}{exp(-\sum_i (y_i - \mu_1)^2/2\sigma^2)}$$

Solving for $\sum y_i$, which can be directly calculated from the data gathered, the following expression is obtained.

$$\sum y_i = \frac{-\sigma^2 ln\lambda}{\mu_1 - \mu_0} + \frac{n}{2}(\mu_1 + \mu_0) \tag{4.3}$$

Substituting the values of λ_1 and λ_0 into equation (4.3) the boundaries for the hypothesis test are obtained. The experimenter would then compare the sum of the data values to the boundary points and as long as the sum fell between the values of λ_1 and λ_0 the experimenter would sample another point. When the running sum fell outside either boundary he or she would make the appropriate conclusions. These calculations are easily displayed graphically, which for the practitioner can be a very useful visual tool.

The sequential testing procedure is simple and elegant and easy for any practitioner to use. However, one would have to know the value of σ. While this might be appropriate in some cases, we feel that when experimenting, the variation during experimentation could be markedly different from a known value. In this case assuming a value for σ could actually hinder the experimental process by requiring more samples to reach a conclusion. Also note, that the limits for each individual test when a normal distribution is assumed will change for each unique situation, based on each assumed value of σ. Asking a practitioner to provide a value for σ is also vague. For what duration of time? Over what conditions? These decisions could drastically influence the results of the experiment. For these reasons, we will not use the likelihood ratio described above, but will assume that σ is estimated by s, the sample standard deviation. The likelihood ratio will now have to be calculated using the t-distribution.

4.3.1 Barnard's Sequential t-test

The framework for this test is the same as above. The practitioner will choose acceptable risk levels α and β and the hypotheses that will be tested. However, in this case the

practitioner will also have to choose the number of standard deviations, δ, that would be considered a significant change in the response. The set of hypotheses for a sequential test developed by Barnard [39] are similar to a standard t-test, which can be written:

$$H_0 : \mu = \mu_0$$

$$H_1 : \mu = \mu_0 + \delta\sigma$$

The above test is designed to detect a difference in the mean in a specified direction; it is one sided. In this case, the difference which is determined to be important is defined by the number of standard deviations. However, we do not have to assume a value for σ, we just need to specify a value for the number of standard deviations δ that would be important for this experiment to detect. In fact the hypothesis for this test can be written as $\delta = 0$ versus $\delta = \delta_1$ if μ_0 is taken to be zero. This is the case if the difference in the samples are tested. By stating the test this way, the likelihood ratio becomes the ratio of a non-central t-distribution and a t-distribution. Barnard [39] and Wetherill and Glazebrook [40] define the likelihood ratio as

$$\lambda = \frac{f(t_n|\delta_1)}{f(t_n|\delta = 0)} \tag{4.4}$$

The t_n in the ratio above is the ratio of the sample mean divided by the sample standard deviation. The boundaries for the test decisions are defined by $\lambda_0 = \beta/(1-\alpha)$ and $\lambda_1 = (1-\beta)/\alpha$ where again the values of α and β are defined by the user. Wetherill and Glazebrook [40] have shown that the likelihood ratio in equation (4.4) can be reduced to the form:

$$\lambda = exp\left\{\frac{\delta^2(n-u_n^2)}{2}\right\} Hh_{n-1}(-\delta u_n)/Hh_{n-1}(0) \tag{4.5}$$

where

$$Hh_n(x) = \int_0^\infty \frac{y^n}{n!} e^{-\frac{1}{2}(y+x)^2} dy \tag{4.6}$$

and

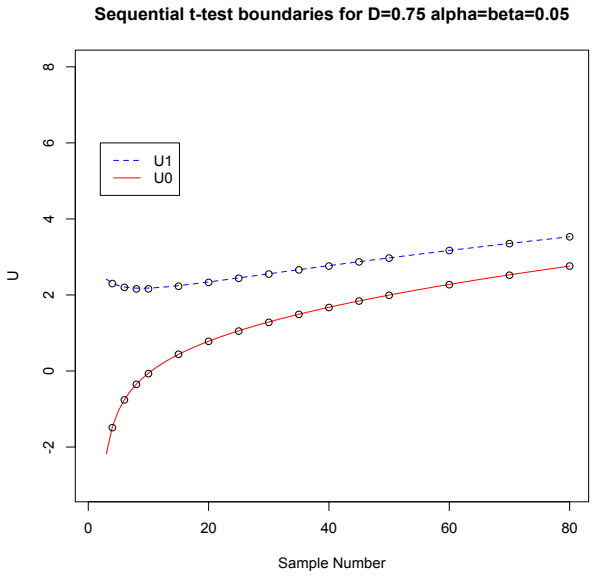
$$u_n = \frac{\sum_{i=1}^{n}(y_i - \mu_0)}{\sqrt{\sum_{i=1}^{n}(y_i - \mu_0)^2}}.$$

The test statistic, u_n, is readily calculable as each observation is taken. The limits for each sequential test are found by finding the roots of equation (4.5) when the appropriate values for λ_1 and λ_0 are subtracted. The limits depend on the values of α, β and δ, which must be specified. The boundaries, U_0 and U_1, determine the next step in the sequential testing procedure. If the statistic u_n falls in the interval from $U_0 < u_n < U_1$ then take another sample. If the value u_n is less than or is equal to the limit U_0 ($u_n \leq U_0$) then the null hypothesis is accepted and there has been no change in the mean (i.e., $\delta = 0$). If the value of u_n is equal to greater than the value of U_1 ($u_n \geq U_1$) then the alternative hypothesis is accepted and there has been an increase in the mean.

Davies [38] gives the limits U_0 and U_1 for different values of α, β, δ and n in the appendix of his book. The inclusion of the integral in equation (4.6) makes the calculations complicated (especially in 1956). However, with advent of readily available software, the limits were easily reproduced using code in the `R` enviroment. It should be noted that unlike the limits for the likelihood assuming a normal distribution, there is a minimum number of samples that need to be gathered in order to apply the sequential t-test. The boundaries on this test are not linear, like the boundaries on the normal sequential test. Figure 4.1 shows the boundaries for the sequential t-test as lines and the open circles are the bounds that Davies has listed in his tables. With the small program in `R` we are able to compute the boundaries. Although the boundaries are only shown for n=80, there is no limit on the potential size of the sample. However, since we will be using this method for an experimental situation, we assume a large sample will not be necessary to make a conclusion. A practitioner may perform this test using a graph similar to the one in Figure 4.1. Figures for $\delta = 1, 1.5, 2$ are shown the appendix.

If desired, one could test for changes in the mean for values of δ less than 0.75. However, we only explore values of δ of 0.75, 1, 1.5, and 2. We feel that in most experimental situations, these values will be adequate. Additionally, in all examples presented we choose values of α and β of 0.05 In practice when analyzing multiple movements in components

Figure 4.1: Boundaries U_0 and U_1 for a sequential t-test and values from Davies table where $\delta = 0.75$ and $\alpha = \beta = 0.5$



it might be advantageous to reduce the values to control for multiple testing errors.

4.4 Applying the Sequential t-test to Adding a Component

In Chapter 3 we introduced a method for adding a component to a mixture. This method of screening in a mixture setting fits nicely into the sequential testing procedure. Recall, when screening in a mixture setting, we assume there is a standard mixture defined as $s = (s_1, s_2, ..., s_q, 0)$. Using this method we wish to study the effect of adding an amount Δ_{q+1} of a new component to s such that the new mixture x is defined as $x = x_i - \Delta_{q+1}s_i$ and $x_{q+1} = \Delta_{q+1}$. Recall, then the estimate for the effect of adding component $q+1$ to s is given by

$$b_{q+1} = \frac{1}{\Delta_{q+1}}[y(x) - y(s)]. \tag{4.7}$$

Note that the parameter estimate using the Cox mixture model is just a weighted difference between the average responses at the new point x and s. We do not assume that the variation is known, for the reasons stated earlier. To use an EVOP procedure for this type of screening the experimenter will take an observation at s, then take an observation at the new mixture x and calculate the difference between the two. That difference will be the test statistic used in the sequential t-test. This implies that μ_0 in the null hypothesis above is zero so that the hypotheses are

$$H_0 : \mu = \mu_0 = 0$$

$$H_1 : \mu = \mu_0 + \delta\sigma = \delta\sigma$$

The experimenter will predetermine the values for δ, α and β; then the procedure for completing this test is as follows:

1. Take an observation at the standard mixture s, $y(s)$.

2. Add Δ_{q+1} of the new component and take an observation at the mixture x, $y(x)$.

3. Calculate the difference in the responses $y(x) - y(s)$ and the statistic $T_n = \sum_i [y(x) - y(s)]$.

4. Calculate the statistic $S_n = \sqrt{\sum_i [y(x) - y(s)]^2}$.

5. For each set of observations (over the minimum amount required) plot the statistic $U_n = T_n/S_n$ to determine whether a change has been made or not.

6. If $U_0 < U_n < U_1$ repeat steps 1-5.

When the path of U_n crosses either one of the boundaries U_0 or U_1 a decision regarding whether or not adding component $q+1$ to the current mixture made the desired change. Figures 4.2 and 4.3 show each possible outcome. The experimental question to be answered by using this procedure is, if an amount Δ_{q+1} of a new component is added to the current formulation, does the response increase or decrease by the desired amount (set by choosing δ)? The possible outcomes of the test procedure above are: (1) the desired change was not observed, (2) a change was observed after many replications, or (3) it took few replications. If the addition did not result in the desired change to y, it is easy to generate the boundaries for smaller values of δ to see if the addition results in a smaller change. If the addition has an effect, but not the magnitude hoped for, an increase in the amount of the component could be considered. The procedure could be repeated for a different amount, Δ, a different new component for comparison and so on. The procedure above is very flexible. Additionally after the procedure is completed and the desired outcome is reached, the effect size can be readily calculated using equation (4.7) which is just the difference in the final average responses $y(x)$ and $y(s)$ weighted by the change Δ. Lastly, since we are studying differences, a block effect need not be considered.

Davies [38] reports minimum sample sizes for values of δ. For the four choices of δ recommended, 0.75, 1, 1.5 and 2 the corresponding minimums are 5, 5, 4, and 3, respectively, for values of $\alpha = \beta = 0.05$. The error rates are also easily adjusted depending on the specific experiment. This statistical procedure can be completed by any practitioner, even with limited statistical knowledge. The test can be done graphically or numerically and can easily applied to multiple responses or components. It is well known that utilizing sequential testing procedures can reduce the number of observations required to make a

Figure 4.2: Sample sequential t-test where true distribution is $N(1,1)$ and $\delta = 1$ and $\alpha = \beta = 0.5$

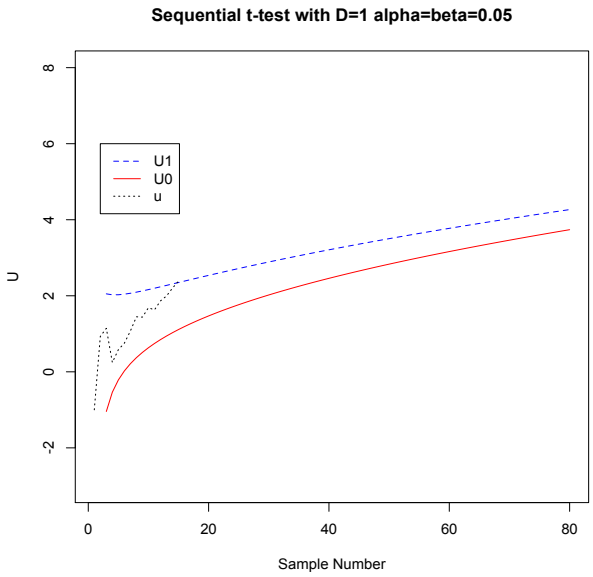
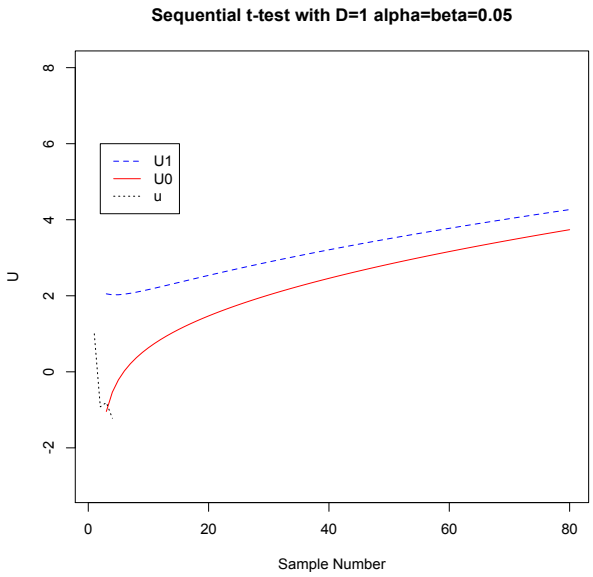


Figure 4.3: Sample sequential t-test where true distribution is $N(0,1)$ and $\delta = 1$ and $\alpha = \beta = 0.5$



decision [38] [37] [39] [41]. In fact, Wald [37] states that using the sequential probability ratio test will reduce the number of observations required by 50%.

The variances of estimates for mixture experiments are large and in general an experimenter will have to do many replications to reach a conclusion. Using a sequential procedure like the one above produces a dramatic runs savings over performing a classical screening experiment. In some cases, it might be preferable to run the classical screening experiments. Although not shown here, this method translates directly to the removing a component screening method described in chapter 3.

4.5 Sequential t-test Simulation of Adding a Component

As stated, sequential testing can dramatically reduce the number of samples required to reach a decision as compared to using fixed sample sizes. To obtain the Average Sample Numbers (ASN) to reach a conclusion for these tests, simulation or an approximation is required, as there is no closed form for the expectation of the sample size when using the t-distribution. Davies [38] provides estimates of these properties in Tables L.1-L.8. We use simulation. Table 4.1 shows the sample size calculation for a test for a difference in means for the considered values of δ. Table 4.2 shows simulation results for testing for a difference of δ in means using the sequential t-test procedure described above when the true distribution was $N(\delta, 1)$. Remember, samples of this size are required for each of the responses $y(s)$ and $y(x)$. The most dramatic differences are at the smaller values of δ but there is still a runs savings of $n = 2$ when $\delta = 2$. Table 4.2 also lists the percent correct decisions out of 10,000 simulations and the maximum sample size required to reach a decision. The error rates are all smaller than the specified error rates of $\alpha = \beta = 0.5$. In Davies table, he lists minimum sample sizes for each error rate combination. These were used in running the simulations and follow the recommendations stated above. Simulations were also run when the true distribution was set as $N(0, 1)$; note that in all cases a conclusion was reached with fewer samples than when the true distribution was $N(\delta, 1)$. All simulations follow the suggested minimum samples sizes and average sample numbers estimated in Davies's tables.

By using the results of the simulations, we are able to look at a histogram of each

Table 4.1: Fixed Sample Size Calculations for a one-sided two-sample t-test with $\alpha = \beta = 0.05$

δ	n
0.75	39.18
1	22.35
1.5	10.37
2	6.231

Table 4.2: 10,000 trials for the sequential t-test from $N(\delta, 1)$, $\alpha = \beta = 0.05$

δ	ASN	Number of Correct Decisions	Percent Correct	Maximum n in Simulation
0.75	13.95	9671	96.71	80
1	9.22	9808	98.08	51
1.5	5.82	9897	98.87	40
2	4.53	9941	99.41	26

Table 4.3: 10,000 trials for the sequential t-test from $N(0, 1)$, $\alpha = \beta = 0.05$

δ	ASN	Number of Incorrect Decisions	Percent Incorrect	Maximum n in Simulation
0.75	11.75	410	4.1	69
1	7.21	334	3.34	42
1.5	3.86	319	3.19	28
2	2.96	305	3.05	16

Figure 4.4: Histogram of sample sizes to reach a decision for 10,000 simulations with $\delta = 0.75$ and the true distribution is $N(0.75, 1)$

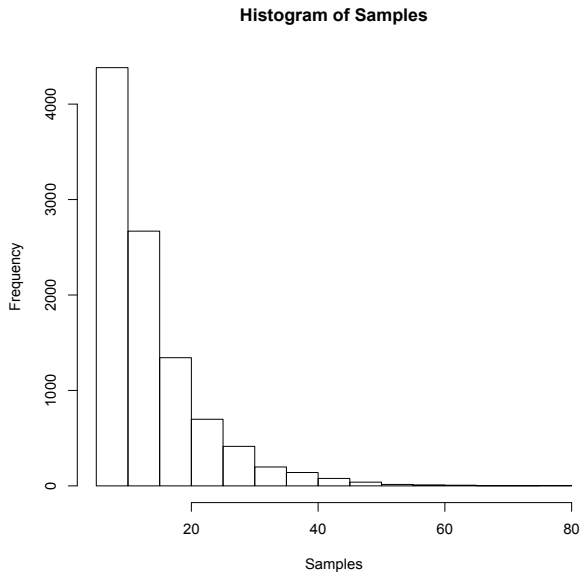


Figure 4.5: Histogram of sample sizes to reach a decision for 10,000 simulations with $\delta = 1$ and the true distribution is $N(1, 1)$

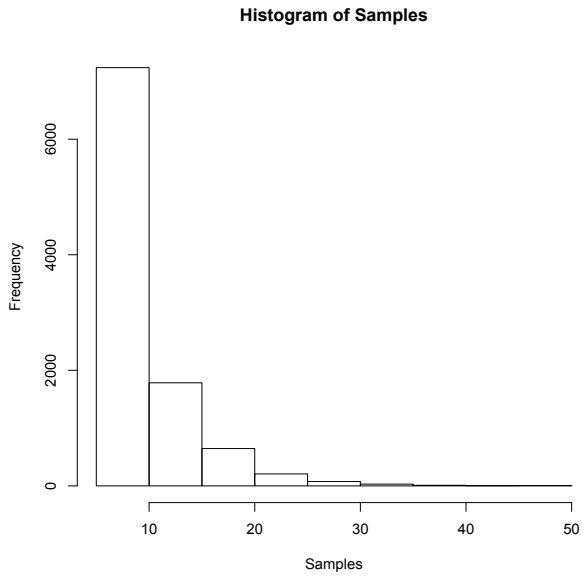


Figure 4.6: Histogram of sample sizes to reach a decision for 10,000 simulations with $\delta = 1.5$ and the true distribution is $N(1.5, 1)$

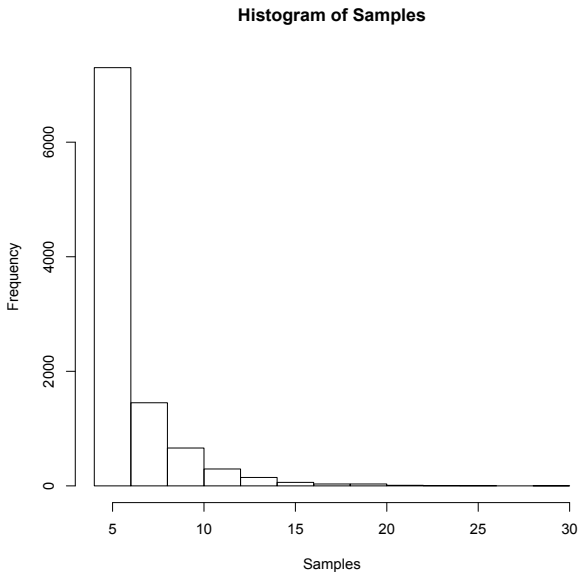


Figure 4.7: Histogram of sample sizes to reach a decision for 10,000 simulations with $\delta = 2$ and the true distribution is $N(2, 1)$

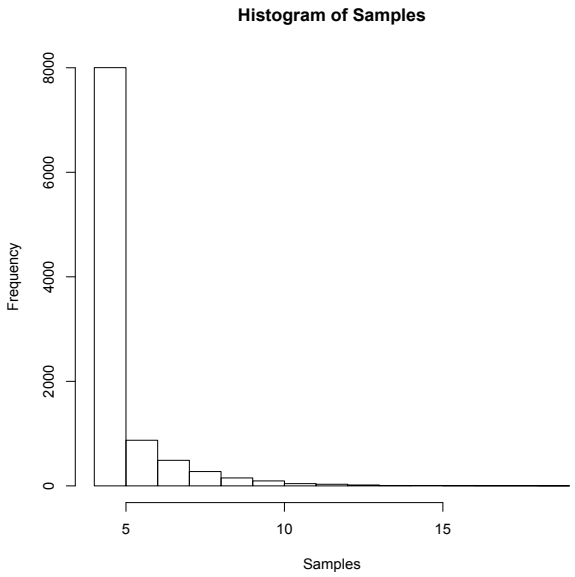


Table 4.4: 10,000 trials for the sequential t-test from $N(\delta, 1)$, $\alpha = \beta = 0.05$ where the sample size was truncated

δ	ASN	Truncated Sample Size	Number of Correct Decisions	Percent Correct
0.75	13.54	40	9570	95.70
1	9.10	22	9502	95.02
1.5	6.20	12	9544	95.44
2	4.45	9	9638	96.38

set of 10,000 runs and recommend some additional stopping points based on the average sample size and shape of the histograms; see Figures 4.2-5. The histograms provide an estimate of potential stopping points for experimenters so that the do not have to attain the possible maximum sample numbers listed in Tables 4.2 and 4.3. For example, looking at the histogram for $\delta = 1$, the majority of the samples are less than $n = 25$ when the true distribution is $N(1, 1)$. We can re-run the simulations above incorporating this cutoff to show that the error rates are unaffected by this number. Practically, we would not want an experimenter to run $n = 50$ (50 for each group) in order to reach a conclusion. Using the histograms as a rough guide we determined the maximum sample size for each value of δ. Table 4.4 lists the maximum run size to maintain the error rate when sampling from a $N(\delta, 1)$ distribution and when the sample reaches the cutoff value a decision is made to accept the null hypothesis. The sample sizes listed in the table are the minimum values at which the simulations maintained the specified error rates. Specifically this means that experimenters can use these numbers as guidelines for sample size cutoffs when using the sequential testing with the screening methods introduced.

4.6 Applying the Sequential t-test EVOP Method to the General Screening Method

Recall the General Screening method introduced in Chapter 3 where the experimental goal is to gain information about the component effects in a particular subset of components in a mixture. The standard mixture is defined as $s = (s_1, s_2, ..., s_q)$. In the general method a subset of components are selected and additions and subtractions of size Δ are made so

that unity is maintained in each mixture.

We will extend this previous method to use sequential testing and the EVOP theme of sequential experimentation. We revisit the example with $q = 9$ components introduced in Chapter 3. We had originally defined components 5-9 as the active ingredients and experimented to reduce component 9, since that was the most expensive ingredient in the mixture. We will now introduce another possible way to use the general screening method. We said that our nine component mixture contained five active ingredients. Now assume that we really just want to study the effects of three of the most expensive active ingredients (7, 8, 9) on the response. If this is our experimental goal, it is not reasonable to move more than one active component at once. We propose moving one active component and one inert ingredient. We define an ingredient as an active component if the component is a small proportion of the whole mixture. For example, there can be active ingredients in cold medicine, an antihistamine, an analgesic, and a decongestant but they do not compose the entire pill or liquid that the person ingests. We also assert that there can be many groups of active ingredients in a mixture; for example in paint there are active drying agents, coloring agents, etc. Each ingredient in those groups composes a small part of the entire product. We define the inert ingredients as the components of the whole that have little to no effect on the intended response but do affect the total product. For example, in a pill, often times corn starch is used to bind the ingredients in to pill form. By making small adjustments in the corn starch the ability to bind into pill form might be affected, while not affecting the effectiveness of the medicine. We also contend that the inert ingredients comprise a larger proportion of the entire mixture, (such as water in a laundry detergent) so taking away a small amount of the inert ingredient is not detrimental to the product. Depending on the response being studied the definition of inert and active might be different.

Assume that ingredient 4 is inert in large proportion in our mixture so we will move it in conjunction with the active ingredients in our nine component mixture. Depending on the scientific knowledge, the direction of movement can vary from component to component, but assume again that we are looking for cheaper alternatives and therefore we wish to study the effects of the most expensive active ingredients. The goal then is to look for the active ingredient with the smallest effect on the response and continue the experimental process from there. The experimental points studied will be

$$
\begin{aligned}
s = &\quad (s_1, s_2, s_3, s_4, s_5, s_6, s_7, s_8, s_9) \\
x_1 = &\quad (s_1, s_2, s_3, s_4 - \Delta, s_5, s_6, s_7 + \Delta, s_8, s_9) \\
x_2 = &\quad (s_1, s_2, s_3, s_4 - \Delta, s_5, s_6, s_7, s_8 + \Delta, s_9) \\
x_3 = &\quad (s_1, s_2, s_3, s_4 - \Delta, s_5, s_6, s_7, s_8, s_9 + \Delta).
\end{aligned}
$$

Recall that Δ need not be equal between points x_1, x_2, x_3, but within each design point it needs to be the same to maintain unity in each mixture. The experimenter will have chosen the values of δ, α and β prior to beginning experimentation. Then the following steps will be taken.

1. Take an observation at the standard mixture s, $y(s)$.
2. Subtract Δ from the inert and add the amount Δ to the first active ingredient. Calculate $y(x_1) - y(s)$ and the statistic $T_1 = \sum_i [y(x_1) - y(s)]$.
3. Calculate the statistic $S_1 = \sqrt{\sum_i [y(x_1) - y(s)]^2}$.
4. For each set of observations plot the statistic $U_1 = T_1/S_1$.
5. Subtract Δ from the inert and add the amount Δ to the active next ingredient. Calculate $y(x_2) - y(s)$ and the statistic $T_2 = \sum_i [y(x_2) - y(s)]$.
6. Calculate the statistic $S_2 = \sqrt{\sum_i [y(x_2) - y(s)]^2}$.
7. For each set of observations plot the statistic $U_2 = T_2/S_2$.
8. Subtract Δ from the inert and add the amount Δ to the active last ingredient. Calculate $y(x_3) - y(s)$ and the statistic $T_3 = \sum_i [y(x_3) - y(s)]$.
9. Calculate the statistic $S_3 = \sqrt{\sum_i [y(x_3) - y(s)]^2}$.
10. For each set of observations plot the statistic $U_3 = T_3/S_3$.

11. Repeat steps until conclusions are reached. Recall there is a minimum sample size to make a decision for different values of δ.

Recall that the estimate generated from the general screening method (3.8) is the weighted change in the response between the standard mixture s and the new mixture x_i. By testing the difference in the responses above we are testing to see if there is a specified (by setting δ) change in the response when moving from the standard mixture s to the new mixture x_i when an amount Δ is added or subtracted from an active ingredient and an inert ingredient. The experimental question is can we increase or decrease the response by δ standard deviations by adding or subtracting an amount Δ of certain active ingredients. Possible outcomes of the procedure above are that certain active ingredients moved by an amount Δ will not have an effect on the response or certain ingredients will have an effect on the response in less replication than others. This could indicate that the effect of those components are smaller or larger and can only be compared if the amount Δ is the same between each experimental point. To calculate the effects D_{iq} from the procedure above, the differences calculated by the final averages are weighted by Δ. This will give the slope of the response surface relative to the standard mixture s to the mixture x_i and then the effect sizes can be judged even if different amounts Δ were used for the different active ingredients between the experimental points. Lastly, since we are studying differences, a block effect need not be considered.

As long as α, β and δ are the same for each component of interest, it would be possible to plot the three statistics U_i on the same plot with the corresponding boundaries. This can easily be done utilizing simple computer code. The procedure would proceed until the desired conclusions are reached. The run savings by using the sequential test would be great considering that each design point will not even require the same amount of replication to reach a conclusion. It could be that the experimenters choose to stop the process even after just one component show the desired result. There is an enormous amount of flexibility using this experimental procedure. It could easily be done such that within some design points the signs of delta are reversed due to the amount of the active ingredient in the standard mixture. Like the EVOP method, observations would be taken at the standard mixture, then at each of the new mixtures, in a cyclic manner until conclusions are reached.

Since the goal of the type of experimentation is not to build a model, when one of the tests reaches a conclusion there would be no reason to consider that component any longer. Therefore as the procedure continued, the simpler it would get. If code is written ahead of time, and the tests are performed graphically, then minimal statistical knowledge is required to carry out and analyze this type of experiment. The procedure above could be used for any number of components; we only show three in the example. However, we developed the procedure in order to experiment on a small subset of components in a complex mixture with a large number of components. Again, we view this as an initial step in the experimental process; the results of this experiment would then lead to the next set of experimental runs and so on until the desired result is achieved. It would also be possible to use multiple responses. However, that would increase the number of charts that would be required. As with any experimental procedure, if multiple responses are being studied with multiple components, the error rates are readily changed to make the tests more stringent. As mentioned this screening procedure is very flexible.

4.7 Conclusions

We have built on the concepts of screening in a mixture setting introduced in Chapter 3 by utilizing the EVOP framework. The sequential EVOP methods are developed for situations where the movements made in factors were intentionally small and by nature the variance is large as in the case of complex mixtures. In the original EVOP methods, sequential significance testing was avoided because it was too complicated for hand calculations. However, in 2010, the availability of computing is so widespread, this is not a concern. Therefore we use the sequential t-test to determine the stopping points in our sequential step experimentation. It has been shown that using a sequential method will greatly reduce the run size required to make a conclusion. The drawback is that the experimenter must decide how large a change is considered significant at the outset of experimentation. When working with mixtures in a lab setting, having deliberate experimental objectives like very resonable. Additionally because of the availability of computing, it would be very easy to plot the same experimental data for a few different values of the desired change (δ) after the experimental runs are complete.

Utilizing the sequential testing procedure outlined will dramatically reduce the run size in mixture experimentation. As stated sequential testing reduces the number of samples required to make a decision by as much as 50 percent. Simulations run match Davies [38] predicated ASN sample numbers and specified error rates and we have shown that we can find a truncated sample size such that, if that certain sample size is obtained, the experimenter can conclude there has been no change with no effect on the overall test error rate. Note that number will be different for different error rates. We have only shown the values when α=β=0.05.

After completing the sequential tests, the averages obtained during the tests may be used to calculate the mixture effects weighted by Δ, which might be necessary if different values of Δ were used in the General Screening method. These methods were designed to make the process of screening in a mixture setting easier for practitioners working with complex mixtures containing a large number of components. The initial goal of our methods is not to build a working model of the mixture response, but to gain information about the individual component effects. There is no current method that allows easy experimentation on a subset of mixture components, nor a method that permits movement in only two mixture components at a time. By defining inert and active ingredients, we further reduce the confounding in the General Screening method by only moving one active ingredient at a time.

Chapter 5

Conclusions and Next Steps

5.1 Conclusions

We have introduced a new screening philosophy and method for mixture experiments. We have shown that for particular screening situations that our methods will produce a runs savings, are intuitive and can be run by a practitioner. We expand those screening methods and concepts and utilized the similarities between mixture experiments and Evolutionary Operation to developed a sequential method of experimentation.

Our screening philosophy is based on the concept of mixture experimentation developed by D.R. Cox in a mixture model form that gives practical interpretation to its coefficients. Recall, this mixture model form is the basis for the effect estimates we use with our screening methods. When adding, removing or manipulating mixture components the coefficients are the slope of the response surface relative to a standard mixture. Many have criticized the Cox model for this fact since it requires a different model for each different standard mixture. Our screening methods revolve around the concept of this standard or current mixture. In industry we assume there is always a current product or formulation with which to begin experimentation.

We then extend Cox's model of mixture experimentation to specific screening situations in mixture experimentation, adding a new component, removing a component and a general screening method designed to study effects of single components and reduce confounding in mixture experimentation. We also define the concepts of screening in a mixture setting

with which we develop these methods. These concepts are (1) screening is sequential, (2) there is always a standard mixture, (3) the goal is not to build a model of the response surface, (4) there is a large number of components and, (5) the desire is to understand effects of individual components. In the current screening methods for mixture experiments the most common approach is to build a linear first order Scheffé mixture model, which requires many design points and movements in components that may not be of interest to the experimenter. Our methods, even when using the same amount of replication, will produce a runs savings over the current methods due to the fact that our method focus on a subset of components.

Noticing the similarities between mixture experimentation and the processes on which EVOP methods are applied, we extend the methods above into an EVOP-type experimentation. Specifically, by introducing a way to use sequential t-testing we further reduce the number of experimental runs required to reach a conclusion. Through simulation we have shown the average sample numbers for sequential testing are less than the fixed sample sizes for the same testing parameters. We have provided `R` code to build the testing boundaries and shown how easy it is to implement these tests graphically. We use the difference in averages to test the difference when adding a component or screening active ingredients, show how the numbers can be weighted by the amount of the components moved Δ and calculate the effects introduced in Chapter 3. We also introduced the concept of inert and active ingredients.

We have presented methods for screening in a mixture setting that are simple to carry out and relevant to the problems facing practitioners in industry. We have given a method for studying individual mixture components, which is not present in the current body of literature. We provide flexible methods that do not require a model to be constructed upon initial experimentation and provides practical interpretation to the effects of adding, subtracting, or moving a component by an amount Δ. Our general screening method can be used with the concepts of active and inert ingredients on a subset of active ingredients as the Chapter 3 example demonstrated. The use of the EVOP method with our screening method allows for a further reduced run size by implementing sequential testing techniques. We believe that the methods introduced give practitioners some new options when screening experiments are used in a mixture setting.

Bibliography

[1] D. Cox, "A Note on Polynomial Response Functions for Mixtures," *Biometrika*, vol. 58, no. 1, pp. 155–159, 1971.

[2] N. R. D. George E.P. Box, *Evolutionary Operation A Statistical Method for Process Improvment.* John Wiley and Sons, Inc., 1969.

[3] M. Quenouille, *The Design and Analysis of Experiments.* London, England: Charles Griffin, 1953.

[4] P. Claringbold, "Use of the Simplex Design in the Study of Joint Action of Related Hormones," *Biometrics*, vol. 11, no. 2, pp. 174–185, 1955.

[5] H. Scheffe, "Experiments with Mixtures," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 20, no. 2, pp. 344–360, 1958.

[6] J. Gorman and J. Hinman, "Simplex Lattice Designs for Multicomponent Systems," *Technometrics*, vol. 4, no. 4, pp. 463–487, 1962.

[7] J. A. Cornell, "Experiments with mixtures: A Review," *Technometrics*, vol. 15, no. 3, pp. 437–455, 1973.

[8] D. Lambrakis, "Experiments with Mixtures: an Alternative to the Simplex-Lattice Design," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 30, no. 1, pp. 123–136, 1968.

[9] H. Scheffe, "The Simplex-Centroid Design for Experiments with Mixtures," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 25, no. 2, pp. 235–263, 1963.

[10] N. R. Draper and W. Lawrence, "Mixture Designs for Three Factors," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 27, no. 3, pp. 450–465, 1965.

[11] N. R. Draper and W. Lawrence, "Mixture Designs for Four Factors," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 27, no. 3, pp. 473–478, 1965.

[12] G. Box and N. R. Draper, "A Basis for the Selection of a Response Surface Design," *Journal of the American Statistical Association*, vol. 54, no. 287, pp. 622–654, 1959.

[13] I. Kurotori, "Experiments with Mixtures of Components Having Lower Bounds," *Industrial Quality Control*, vol. 22, pp. 592–596, 1966.

[14] R. McLean and V. Anderson, "Extreme Verticies Design of Mxiture Experiments," *Technometrics*, vol. 8, no. 3, pp. 447–454, 1966.

[15] D. Lambrakis, "Experiments with p-component Mixtures," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 30, no. 1, pp. 137–144, 1968.

[16] W. O. Thompson and R. H. Myers, "Response Surface Designs for Experiments with Mixtures," *Technometrics*, vol. 10, no. 4, pp. 739–756, 1968.

[17] D. Lambrakis, "Estimated Regression Function of the Multiple Lattice Design," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 31, no. 2, pp. 276–284, 1969.

[18] N. Becker, "Models for the Response of a Mixture," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 30, no. 2, pp. 349–358, 1968.

[19] N. Becker, "Regression Problems with the Predictor Variables are Proportions," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 31, no. 1, pp. 107–112, 1969.

[20] J. Aitchison and J.Bacon-Shone, "Log Contrast Models for Experiments with Mixtures," *Biometrika*, vol. 71, no. 2, pp. 323–330, 1984.

[21] N. R. Draper and R. C. St.John, "A Mixtures Model with Inverse Terms," *Technometrics*, vol. 19, no. 1, pp. 37–49, 1977.

[22] J. A. Cornell, *Experiments with Mixtures: Designs, Models and the Analysis of Mixture Data*. New York, NY: John Wiley and Sons, Inc., third ed., 2002.

[23] W. F. Smith, *Experimental Design for Formulation.* Alexandria, VA: ASA SIAM Series on Statistics and Applied Probability, 2005.

[24] J. S. H. George E.P. Box and W. G. Hunter, *Statistics for Experimenters: Design, Innovation and Discovery.* Hoboken, New Jersey: John Wiley and Sons, Inc., second ed., 2005.

[25] C. J. Wu and M. Hamada, *Experiments: Planning, Analysis, and Parameter Design Optimization.* New York, NY: John Wiley and Sons, Inc., 2000.

[26] D. C. Montgomery, *Design and Analysis of Experiments.* New York, NY: John Wiley and Sons, Inc., 5th ed., 2000.

[27] R. D. Snee and D. W. Marquardt, "Screening Concepts and Designs for Experiments with Mixtures," *Technometrics*, vol. 18, no. 1, pp. 19–29, 1976.

[28] S. H. Park, "Selecting Contrasts among Parameters in Scheffe's Mixture Models: Screening Components and Model Reduction," *Technometrics*, vol. 20, no. 3, pp. 273–279, 1978.

[29] G. F. Piepel, "Screening Designs for Constrained Mixture Experiments Derived from Classical Screening Designs," *Journal of Quality Technology*, vol. 22, no. 1, pp. 23–33, 1990.

[30] J. M. S. Greg F. Piepel, Ruel D. Hicks and J. L. Loeppky, "Methods for Assessing Curvature and Interaction in Mixture Experiments," *Technometrics*, vol. 44, no. 2, pp. 161–172, 2002.

[31] J. A. Cornell, "Some Comments on Designs for Cox's Mixture Polynomial," *Technometrics*, vol. 17, no. 1, pp. 25–35, 1975.

[32] W. F. Smith and T. A. Beverly, "Generating Linear and Quadratic Cox Mixture Models," *Journal of Qualtiy Technology*, vol. 29, no. 2, pp. 211–224, 1997.

[33] N. Kettaneh-Wald, "Analysis of Mixture Data with Partial Least Squares," *Chemometrics and Intellignet Laboratory Systems*, vol. 14, pp. 57–69, 1992.

[34] K. Muteki and J. F. MacGregor, "Sequential Design of Mixture Experiments for the Development of New Products," *Journal of Chemometrics*, vol. 21, pp. 496–505, 2007.

[35] E. J. Lennart Eriksson and C. Wikstrom, "Mixture design-desisgn generation, PLS analysis and Model Usage," *Chemometrics and Intelligent Laboratory Systems*, vol. 43, pp. 1–24, 1998.

[36] G. E. Box and J. S. Hunter, "Condensed Calculations for Evolutionary Operation Programs," *Technometrics*, vol. 1, no. 1, pp. 77–95, 1959.

[37] A. Wald, "Sequential Tests of Hypotheses," *The Annals of Mathematical Statistics*, vol. 16, no. 2, pp. 117–186, 1945.

[38] O. L. Davies, ed., *The Design and Analysis of Industrial Experiments.* Oliver and Boyd, 1956.

[39] G. A. Barnrad, "Sequential Tests in Industrial Statistics," *Supplement to the Journal of the Royal Statistical Society*, vol. 8, no. 1, pp. 1–26, 1946.

[40] G. B. Wetherill and K. D. Glazebrook, *Sequential Methods in Statistics.* Chapman and Hall, third ed., 1966.

[41] S. Rushton, "On a Sequential t-Test," *Biometrika*, vol. 37, pp. 326–333, 1950.

mariabib

Appendix

This is the R code for the sequential t-test boundaries for specified values of δ, α and β. Reference: http://www.biostat.wustl.edu/archives/html/s-news/2002-03/msg00167.html

```
    integrand=function(y, n, x){
            n=n-1
            ans=((y^n) * exp(-0.5*(x+y)^2))


            }


    Hh=function(n, x, lolim=0, hilim=Inf){

    a=integrate(f=integrand, lower=lolim, upper=hilim,
subdivisions=1000, n=n, x=x)
    int=a$value*(1/gamma((n-1)+1))

        }



    Lik=function(U, n, delta, lhs){
        lik=lhs-(exp(-0.5*(delta^2)*(n-U^2))*Hh(n, x=-delta*U)/Hh
(n,x=0))
        lik


        }
    U0=U1=0
    SampNo=1:maxn
    #n=5
    for(n in minn:maxn){
        m=uniroot(Lik, lower=bot, upper=top, n=n, delta=delta,
lhs=(1-beta)/alpha)
        U1[n]=m$root


        k=uniroot(Lik, lower=bot, upper=top, n=n, delta=delta,
lhs=beta/(1-alpha))
        U0[n]=k$root
```

Figure 1: Boundaries U_0 and U_1 for a sequential t-test and values from Davies table where $\delta = 1$ and $\alpha = \beta = 0.5$

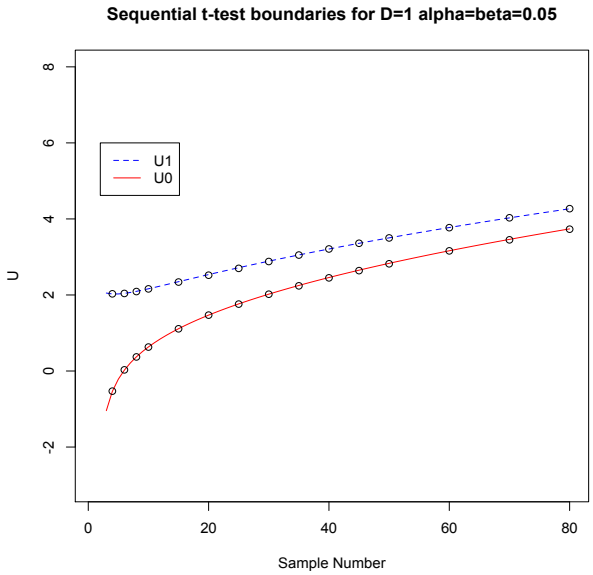


Figure 2: Boundaries U_0 and U_1 for a sequential t-test and values from Davies table where $\delta = 1.5$ and $\alpha = \beta = 0.5$

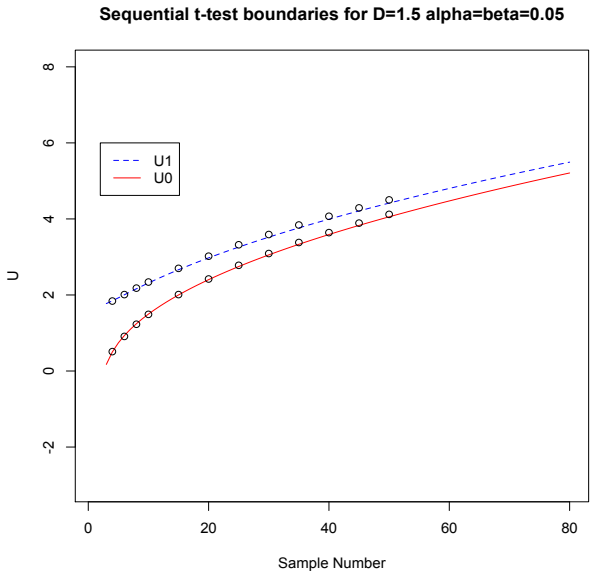
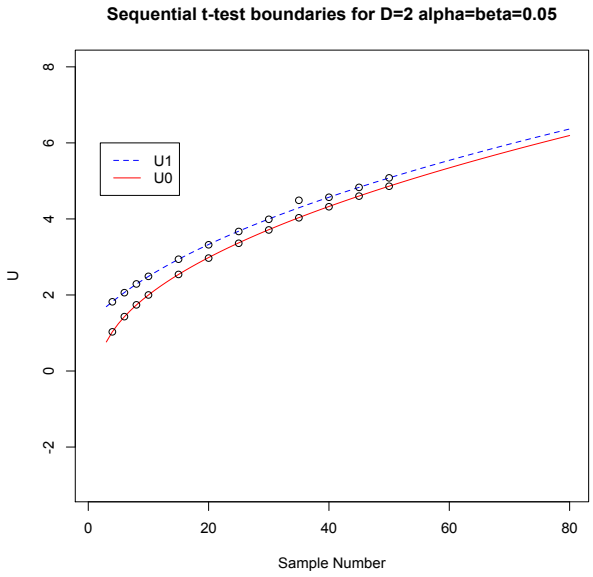


Figure 3: Boundaries U_0 and U_1 for a sequential t-test and values from Davies table where $\delta = 2$ and $\alpha = \beta = 0.5$



Vita

Maria Weese is a Ph.D candidate and graduate teaching associate in Statistics, Operations and Management Science at the University of Tennessee. She received a B.S. degree in Chemical Engineering with a minor in Chemistry from Virginia Polytechnic Institute and State University in 2001. Maria then worked for three years as a process improvement and production engineer at Celanese Acetate in Narrows, VA. In 2004 Maria returned to school to pursue her M.S. in statistics at the University of Tennessee. From 2004 through 2006 Maria worked as a consultant in the Statistical Consulting Center where she assisted students, faculty and staff in statistical analyses. In 2006 Maria became a graduate teaching associate at the University of Tennessee and began work on her Ph.D.. Maria has taught four semesters of Introduction to Statistics and three semesters of Statistical Process Control while obtaining her Ph.D.. In 2008 Maria was named the Associate Director of the Institute for Statistical Engineering where she developed and taught custom statistical workshops for two international corporations. Maria has presented research at the 2009 Joint Statistical Meetings, where she was awarded the SPES Outstanding Poster Presentation, and the 2009 Fall Technical Conference. Maria is a recipient of the 2008 Fall Technical Conference Student Travel award and in 2009 Maria received the College of Business Administration ESPN scholarship to fund future research. After obtaining a Ph.D., Maria will move to Oxford, OH where she will serve as a Visiting Assistant Professor at Miami University of Ohio.