Mixture Designs

BIOE 498/598 PJ

Mixtures

Some materials or reagents are **mixtures**, or combinations of different proportions of a fixed number of ingredients.
Examples of mixtures include:

- Polymers
- Metal alloys
- Fabric blends
- Cell culture media

Studying mixtures requires special designs and models.

Defining Mixtures

A mixture has k components, each with proportion x_i, where

$$0.0 \leq x_i \leq 1.0$$

and

$$\sum_{i=1}^{k} x_i = 1.0$$

Defining Mixtures

A mixture has k components, each with proportion x_i, where

$$0.0 \leq x_i \leq 1.0$$

and

$$\sum_{i=1}^{k} x_i = 1.0$$

For example, a three-ingredient mixture ($k = 3$) is defined by

$$\begin{aligned} 0.0 \leq x_1 &\leq 1.0 \\ 0.0 \leq x_2 &\leq 1.0 \\ 0.0 \leq x_3 &\leq 1.0 \\ x_1 + x_2 + x_3 &= 1.0 \end{aligned}$$

Why not factorial designs?

The summation constraint $\sum_i x_i = 1$ prevents us from using factorial designs. Let's imagine a factorial design for a three-ingredient mixture with levels $- = 0.0$ and $+ = 1.0$. The runs are

1	2	3	x_1	x_2	x_3	$\sum_i x_i$
−	−	−	0.0	0.0	0.0	0.0
+	−	−	1.0	0.0	0.0	1.0
−	+	−	0.0	1.0	0.0	1.0
−	−	+	0.0	0.0	1.0	1.0
+	+	−	1.0	1.0	0.0	2.0
+	−	+	1.0	0.0	1.0	2.0
−	+	+	0.0	1.0	1.0	2.0
+	+	+	1.0	1.0	1.0	3.0

Only three of the eight runs satisfy the summation constraint!

Models for Mixture Experiments

The standard first-order model for a factorial experiment with three factors is

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \epsilon$$

This model is not appropriate for mixture experiments because of the summation constraint.

Instead, we can use two alternative models:

1. The slack variable model
2. The Scheffé model

The slack variable model

Let's start with the standard FO model

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \epsilon$$

By the summation constraint, $x_3 = 1 - x_1 - x_2$. Substituting, we see that

$$\begin{aligned} y &= \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 (1 - x_1 - x_2) + \epsilon \\ &= (\beta_0 + \beta_3) + (\beta_1 - \beta_3) x_1 + (\beta_2 - \beta_3) x_2 + \epsilon \\ &= \beta_0' + \beta_1' x_1 + \beta_2' x_2 + \epsilon \end{aligned}$$

The slack variable model

Let's start with the standard FO model

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \epsilon$$

By the summation constraint, $x_3 = 1 - x_1 - x_2$. Substituting, we see that

$$\begin{aligned} y &= \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3(1 - x_1 - x_2) + \epsilon \\ &= (\beta_0 + \beta_3) + (\beta_1 - \beta_3)x_1 + (\beta_2 - \beta_3)x_2 + \epsilon \\ &= \beta_0' + \beta_1' x_1 + \beta_2' x_2 + \epsilon \end{aligned}$$

The effect of factor x_3 is folded into the intercept and the other effects. When using a slack variable model we choose x_3 to be the least important factor, but the effects remain confounded.

The Scheffé model

The Scheffé model takes a different approach. Rather than substitute for x_3, we substitute for the intercept (the 1 multiplied by β_0) using the constraint $x_1 + x_2 + x_3 = 1$:

$$\begin{aligned}
y &= \beta_0(1) + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \epsilon \\
&= \beta_0(x_1 + x_2 + x_3) + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \epsilon \\
&= (\beta_0 + \beta_1)x_1 + (\beta_0 + \beta_2)x_2 + (\beta_0 + \beta_3)x_3 + \epsilon \\
&= \beta_1^* x_1 + \beta_2^* x_2 + \beta_3^* x_3 + \epsilon
\end{aligned}$$

The Scheffé model

The Scheffé model takes a different approach. Rather than substitute for x_3, we substitute for the intercept (the 1 multiplied by β_0) using the constraint $x_1 + x_2 + x_3 = 1$:

$$\begin{aligned} y &= \beta_0(1) + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \epsilon \\ &= \beta_0(x_1 + x_2 + x_3) + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \epsilon \\ &= (\beta_0 + \beta_1)x_1 + (\beta_0 + \beta_2)x_2 + (\beta_0 + \beta_3)x_3 + \epsilon \\ &= \beta_1^* x_1 + \beta_2^* x_2 + \beta_3^* x_3 + \epsilon \end{aligned}$$

The effects in a Scheffé model are clean, but they have a special interpretation. The coefficient β_i^* is the expected response for a pure mixture with $x_i = 1.0$ and all other ingredients set to zero.

The Simplex-Lattice Design (SLD)

A Simplex-Lattice Design SLD$\{k, m\}$ studies mixtures with k ingredients set at $m + 1$ equally-spaced levels. The design uses all combinations of the ingredient levels.

The Simplex-Lattice Design (SLD)

A Simplex-Lattice Design SLD$\{k, m\}$ studies mixtures with k ingredients set at $m + 1$ equally-spaced levels. The design uses all combinations of the ingredient levels.

SLD$\{3, 1\}$: Levels 0, 1

(1,0,0) (0,1,0) (0,0,1)

The Simplex-Lattice Design (SLD)

A Simplex-Lattice Design SLD$\{k, m\}$ studies mixtures with k ingredients set at $m + 1$ equally-spaced levels. The design uses all combinations of the ingredient levels.

SLD$\{3, 1\}$: Levels 0, 1

(1,0,0) (0,1,0) (0,0,1)

SLD$\{3, 2\}$: Levels 0, $\frac{1}{2}$, 1

$(\frac{1}{2},\frac{1}{2},0)$ $(\frac{1}{2},0,\frac{1}{2})$ $(0,\frac{1}{2},\frac{1}{2})$ + SLD$\{3, 1\}$

The Simplex-Lattice Design (SLD)

A Simplex-Lattice Design SLD$\{k, m\}$ studies mixtures with k ingredients set at $m + 1$ equally-spaced levels. The design uses all combinations of the ingredient levels.

SLD$\{3, 1\}$: Levels 0, 1

$$(1,0,0) \quad (0,1,0) \quad (0,0,1)$$

SLD$\{3, 2\}$: Levels 0, $\frac{1}{2}$, 1

$$(\tfrac{1}{2},\tfrac{1}{2},0) \quad (\tfrac{1}{2},0,\tfrac{1}{2}) \quad (0,\tfrac{1}{2},\tfrac{1}{2}) \quad + \text{SLD}\{3, 1\}$$

SLD$\{3, 3\}$: Levels 0, $\frac{1}{3}$, $\frac{2}{3}$, 1

$$(\tfrac{1}{3},\tfrac{2}{3},0)\ (\tfrac{2}{3},\tfrac{1}{3},0) \quad (\tfrac{1}{3},0,\tfrac{2}{3})\ (\tfrac{2}{3},0,\tfrac{1}{3}) \quad (0,\tfrac{1}{3},\tfrac{2}{3})\ (0,\tfrac{2}{3},\tfrac{1}{3}) \quad (\tfrac{1}{3},\tfrac{1}{3},\tfrac{1}{3}) + \text{SLD}\{3, 1\}$$

The Simplex-Lattice Design (SLD)

A Simplex-Lattice Design SLD$\{k, m\}$ studies mixtures with k ingredients set at $m + 1$ equally-spaced levels. The design uses all combinations of the ingredient levels.

SLD$\{3, 1\}$: Levels 0, 1

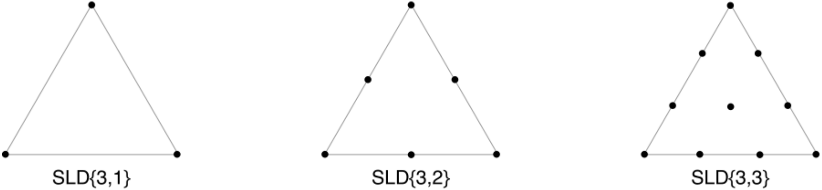
$$(1,0,0) \quad (0,1,0) \quad (0,0,1)$$

SLD$\{3, 2\}$: Levels 0, $\frac{1}{2}$, 1

$$(\tfrac{1}{2},\tfrac{1}{2},0) \quad (\tfrac{1}{2},0,\tfrac{1}{2}) \quad (0,\tfrac{1}{2},\tfrac{1}{2}) \quad + \text{SLD}\{3,1\}$$

SLD$\{3, 3\}$: Levels 0, $\frac{1}{3}$, $\frac{2}{3}$, 1

$$(\tfrac{1}{3},\tfrac{2}{3},0)\ (\tfrac{2}{3},\tfrac{1}{3},0) \quad (\tfrac{1}{3},0,\tfrac{2}{3})\ (\tfrac{2}{3},0,\tfrac{1}{3}) \quad (0,\tfrac{1}{3},\tfrac{2}{3})\ (0,\tfrac{2}{3},\tfrac{1}{3}) \quad (\tfrac{1}{3},\tfrac{1}{3},\tfrac{1}{3}) + \text{SLD}\{3,1\}$$



The Simplex-Centroid Design (SCD)

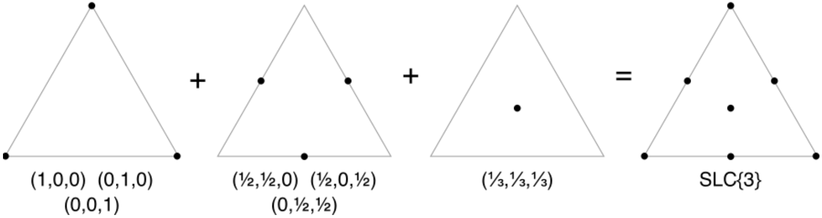
A downside of the SLD$\{k, m\}$ design is there are no interior points until $m \geq 3$. An alternative is the SCD$\{k\}$ design which includes

- All k pure mixtures: (1,0,. . .,0)
- All binary combinations at $\frac{1}{2}$: ($\frac{1}{2}$,$\frac{1}{2}$,0,. . .,0)
- All trinary combinations at $\frac{1}{3}$: ($\frac{1}{3}$,$\frac{1}{3}$,$\frac{1}{3}$,0,. . .,0)
- $\vdots$
- The single k-nary mixture: ($\frac{1}{k}$,$\frac{1}{k}$,. . .,$\frac{1}{k}$)

The Simplex-Centroid Design (SCD)

A downside of the SLD$\{k, m\}$ design is there are no interior points until $m \geq 3$. An alternative is the SCD$\{k\}$ design which includes

- All k pure mixtures: (1,0,...,0)
- All binary combinations at $\frac{1}{2}$: $(\frac{1}{2},\frac{1}{2},0,\ldots,0)$
- All trinary combinations at $\frac{1}{3}$: $(\frac{1}{3},\frac{1}{3},\frac{1}{3},0,\ldots,0)$
- $\vdots$
- The single k-nary mixture: $(\frac{1}{k},\frac{1}{k},\ldots,\frac{1}{k})$



Choosing the SLD or SCD

- The SLD$\{k, m\}$ model allows fitting an m^{th} order model (1 = first order, 2 = quadratic). The SCD$\{k\}$ can fit a k-order model with up to a k-way interaction term.
- In a SCD model, no single ingredient is run at a proportion $\geq \frac{1}{2}$.
- Given the same number of runs, the SLD has better coverage of the boundary of the simplex while the SCD has better coverage of the interior.