Optimal designs for experiments with mixtures

Stefanie Biedermann

University of Southampton

Joint work with: Rana Khashab and Steve Gilmour

Design of Experiments: New Challenges
CIRM, Marseille, France
2 May 2018

UNIVERSITY OF Southampton

Mixture Experiments

Experiments with mixtures may occur in many areas of scientific research (Cornell, 1990), for example:

- Building construction: Mix cement, sand and water – measure the hardness of the resulting concrete
- Animal husbandry: The weight gain of chicks depends on the mix of energy supplements in their diets
- Food processing: The taste and texture of a cake depend on the mix of its ingredients

MIXTURE EXPERIMENTS

Experiments with mixtures usually investigate how changing the *proportions* of the different components affects the response

- Let $x_1, x_2, \ldots, x_q$ describe the proportions of the q ingredients or mixture components
- Then, $x_r \geq 0$, $r = 1, \ldots, q$, and

$$\sum_{r=1}^{q} x_r = 1$$

- The above constraints are called the 'natural constraints' of a mixture experiment

EXPERIMENTAL REGION

The experimental region for a mixture experiment with natural constraints is a regular simplex

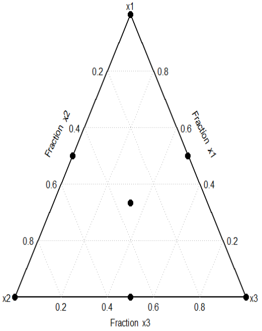

Figure: Experimental region when $q = 3$

Mixture Experiments

Often, there may be additional constraints on the proportions of the mixture components

- The proportions of flour, butter, sugar and eggs in a cake should all be bounded away from zero and one!
- This leads to additional constraints of the form $l_r \leq x_r \leq u_r$, where $l_r > 0$ and $u_r < 1$, $r = 1, \ldots, q$
- The experimental region is then no longer a regular simplex

Experimental Region

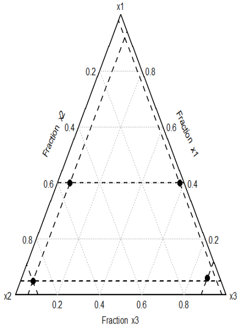

Figure: Experimental region with additional constraints when $q = 3$

DESIGN OF EXPERIMENTS

- A design point $\mathbf{x}$ is a combination of the proportions of the q ingredients, i.e. $\mathbf{x} = (x_1, x_2, \ldots, x_q)$
- A design is then the selection of $\mathbf{x}_1, \ldots, \mathbf{x}_n$ to be used in the experiment
- We will consider approximate designs of the form

$$\xi = \left\{ \begin{matrix} \mathbf{x}_1 & \mathbf{x}_2 & \ldots & \mathbf{x}_m \\ w_1 & w_2 & \ldots & w_m \end{matrix} \right\}; \; 0 < w_i \leq 1, \; \sum_{i=1}^{m} w_i = 1$$

  - $\mathbf{x}_i \in \mathcal{X}$, $i = 1, \ldots, m$, $m \leq n$: support points of ξ.
  - w_i, $i = 1, \ldots, m$: weights (proportions) corresponding to $\mathbf{x}_i$s.

Optimal Design of Experiments

- We consider D-optimal designs, which maximise the determinant, $|M(\xi, \theta)|$, of the information matrix
- To compute $M(\xi, \theta)$, we need to specify the model

MODEL

Assume the responses are modelled as

$$Y_i = \eta_i + \varepsilon_i, \quad i = 1, \ldots, n,$$

where $\varepsilon_i \overset{iid}{\sim} \mathcal{N}(0, \sigma^2)$

- We can then express the mean η_i as a function of the proportions of the ingredients and an unknown parameter vector $\boldsymbol{\theta}$
- $\eta = \phi(x_1, \ldots, x_q, \boldsymbol{\theta})$
- Note that in many applications there is no mechanistic model available, and hence empirical models are used
- For identifiability, the natural constraint, $x_1 + x_2 + \ldots + x_q = 1$, has to be taken into account
- Various functional relationships have been suggested in the literature

SCHEFFÉ POLYNOMIALS

Scheffé (1958) modified standard polynomial models to incorporate the natural constraints

- In a polynomial of degree 1, $\eta = \beta_0 + \sum_{r=1}^{q} \beta_r x_r$, the intercept is multiplied by $x_1 + x_2 + \ldots + x_q = 1$
- Simplification yields the Scheffé polynomial of degree 1

$$\eta = \sum_{r=1}^{q} \beta_r^* x_r$$

where $\beta_r^* = \beta_0 + \beta_r$

Scheffé Polynomials

Similar arguments and the fact that

$$x_r^2 = x_r\left(1 - \sum_{s=1,\, s\neq r}^{q} x_s\right)$$

yield a polynomial of degree 2 as

$$\eta = \sum_{r=1}^{q} \beta_r^{**} x_r + \sum_{r=1}^{q-1} \sum_{s=r+1}^{q} \beta_{rs}^{*} x_r x_s$$

and so on

Ratio Models (Cornell, 1990)

First and second order ratio models are given by

$$\eta = \beta_0 + \sum_{r=1}^{q-1} \beta_r \left(\frac{x_r}{x_q} \right)$$

and

$$\eta = \beta_0 + \sum_{r=1}^{q-1} \beta_r \left(\frac{x_r}{x_q} \right) + \sum_{r=1}^{q-1} \beta_{rr} \left(\frac{x_r}{x_q} \right)^2 + \sum_{r=1}^{q-2} \sum_{s=r+1}^{q-1} \beta_{rs} \left(\frac{x_r}{x_q} \right) \left(\frac{x_s}{x_q} \right),$$

respectively, where the proportion x_q for the denominator is usually chosen such that its range, $u_q - l_q$, is smallest

General Blending Models

Brown, Donev and Bissett (2015) propose a broader, more flexible, class of models called General Blending Models (GBM) where

$$\eta = \sum_{r=1}^{q} \beta_r x_r + \sum_{r=1}^{q} \sum_{s=r}^{q} \beta_{rs} \left(\frac{x_r}{x_r + x_s} \right)^{k_{rs}} \left(\frac{x_s}{x_r + x_s} \right)^{k_{sr}} (x_r + x_s)^{m_{rs}}$$

Here, the powers k_{rs} and k_{sr} have possible values of 0.5, 1, 1.5, 2, 2.5 or 3, and the powers m_{rs} may attain values 0, 1, 2 or 3

These models often fit well, but have a large number of parameters to be estimated and long computing times for model fitting

General Blending Models

Brown, Donev and Bissett (2015) propose a broader, more flexible, class of models called General Blending Models (GBM) where

$$\eta = \sum_{r=1}^{q} \beta_r x_r + \sum_{r=1}^{q} \sum_{s=r}^{q} \beta_{rs} \left(\frac{x_r}{x_r + x_s} \right)^{k_{rs}} \left(\frac{x_s}{x_r + x_s} \right)^{k_{sr}} (x_r + x_s)^{m_{rs}}$$

Here, the powers k_{rs} and k_{sr} have possible values of 0.5, 1, 1.5, 2, 2.5 or 3, and the powers m_{rs} may attain values 0, 1, 2 or 3

These models often fit well, but have a large number of parameters to be estimated and long computing times for model fitting

Research Question

Can we find a similarly flexible class of models to fit data from experiments with mixtures, which avoids these drawbacks?

FRACTIONAL POLYNOMIAL MODELS

Royston and Altman (1994) propose Fractional Polynomial Models (FPM) where the first and second order models, respectively, are given by

$$\eta = \beta_0 + \sum_{r=1}^{q} \beta_r x_r^{(\alpha_r)}$$

and

$$\eta = \beta_0 + \sum_{r=1}^{q} \beta_r x_r^{(\alpha_r)} + \sum_{r=1}^{q} \beta_{rr} x_r^{2(\alpha_r)} + \sum_{r=1}^{q-1} \sum_{s=r+1}^{q} \beta_{rs} x_r^{(\alpha_r)} x_s^{(\alpha_s)}$$

For better interpretability, the powers α_r are often restricted to a set such as $\{-3, -2, -1, -0.5, -1/3, 0, 1/3, 0.5, 1, 2, 3\}$, where $x_r^{(\alpha_r)} = \log(x_r)$ if $\alpha_r = 0$

Fractional Polynomial Models

How can we modify FPMs, so they can be applied to experiments with mixtures?

- Similar to the ratio model, we can map x_r, the proportion of the rth component, to the ratio

$$x_r \longmapsto \left(\frac{x_r}{x_q}\right), \quad r = 1, \ldots, q-1$$

- This approach requires that $x_1, \ldots, x_q > 0$

Modified Fractional Polynomial Models

The Modified Fractional Polynomial (MFP) Models of degree 1 and 2, respectively, are then

$$\eta = \beta_0 + \sum_{r=1}^{q-1} \beta_r \left(\frac{x_r}{x_q}\right)^{\alpha_r}$$

and

$$\eta = \beta_0 + \sum_{r=1}^{q-1} \beta_r \left(\frac{x_r}{x_q}\right)^{\alpha_r} + \sum_{r=1}^{q-1} \beta_{rr} \left(\frac{x_r}{x_q}\right)^{2\alpha_r} + \sum_{r=1}^{q-2} \sum_{s=r+1}^{q-1} \beta_{rs} \left(\frac{x_r}{x_q}\right)^{\alpha_r} \left(\frac{x_s}{x_q}\right)^{\alpha_s}$$

where the case $\alpha_r = 0$ corresponds to the logarithmic transformation

MODIFIED FRACTIONAL POLYNOMIAL MODELS

- For better interpretability of the results, we can restrict the powers α_r to a set such as $\{-3, -2, -1, -0.5, -1/3, 0, 1/3, 0.5, 1, 2, 3\}$
- Simplified versions of the MFP Models, with fewer parameters to estimate, can be defined by requiring all exponents α_r to be equal, i.e. $\alpha_1 = \ldots = \alpha_{q-1} = \alpha$
- The choice of the denominator x_q is generally not obvious
- We propose to fit the MPF Model for each possible denominator x_r, $r = 1, \ldots, q$, in turn, and to select the model that provides the best fit

Model Comparisons

So far so good, but are the MFP Models actually useful?

- The candidate models for comparison are the Scheffé polynomials, the ratio models and the General Blending Models
- We will use R^2, AIC and BIC for model comparisons
- For MFP and ratio models, we will use the denominator that provides the best fit

Chick Feeding Example

- Cornell (1990) provides data of a chick feeding experiment
- The chicks were fed purified diets consisting of energy supplements that contain protein (x_1), fat (x_2) and carbohydrate (x_3)
- The proportions of these components are constrained by

$$0.05 \leq x_1 \leq 0.40, \quad 0.02 \leq x_2 \leq 0.89, \quad 0.06 \leq x_3 \leq 0.86$$

- The response is gain in body weight

CHICK FEEDING EXAMPLE

Model	R^2	AIC	BIC
Scheffé (degree 2)	0.9760	225.34	237.15
Ratio (degree 2) divided by x_1	0.9624	267.23	277.04
GBM	0.9936	232.49	242.31
MFP (degree 2, different α_r) divided by x_3	0.9997	205.57	218.18

Table: Model comparison for the chick feeding data

Manufacturing Example

- Box and Draper (2007) provide data of a manufacturing experiment
- The response is the burning rate of the experiment
- The mixture components are fuel (x_1), oxidizer (x_2) and binder (x_3)
- The proportions of these components are constrained by

$$0.1 \leq x_r \leq 0.99, \quad r = 1, 2, 3$$

Manufacturing Example (1st Degree Models)

Model	R^2	AIC	BIC
Scheffé (degree 1)	0.61282	143.09	145.95
Ratio (degree 1)	0.21980	143.70	146.53
MFP (degree 1, same α_r)	0.68824	135.36	139.12
MFP (degree 1, different α_r)	0.68995	133.45	136.67

Table: Model comparison for the manufacturing data (first degree models)

Manufacturing Example (2nd Degree Models)

Model	R^2	AIC	BIC
Scheffé (degree 2)	0.9797	132.16	137.15
Ratio (degree 2)	0.8290	126.95	131.91
GBM	0.9798	127.19	132.14
MFP (degree 2, different α_r)	0.9897	120.00	125.67

Table: Model comparison for the manufacturing data (second degree models)

Conclusion (Modelling)

- The MFP Models provide a good fit to historical data sets (better than competing models)
- They were much quicker to fit than the GBMs
- The MFP Models require additional constraints where the proportions $x_1, \ldots, x_q$ are all positive

- Often, the aim of the experiment is finding the optimal expected response/corresponding setting of the factors
  $\hookrightarrow$ Which model is best for achieving this purpose?
  $\hookrightarrow$ Investigation needed
- How should experiments with mixtures be designed, i.e. which combinations of proportions should be used in the experiment to 'best' answer the research question?

Designing for MFP Models

- To find optimal designs for the MFP Models we have to deal with the problem of irregularly shaped design spaces
- We used two different approaches

Finding a candidate set

- The XVERT algorithm (Snee and Marquardt, 1974) finds the extreme vertices of the experimental region
- We can use convex combinations of these to fill up the design space, e.g. starting with face centroids and the overall centroid
- We can then use an optimisation algorithm on the resulting candidate set

Finding a candidate set

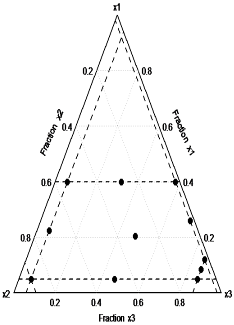

Figure: Illustration of finding a candidate set

Constrained Optimisation

- Alternatively, we can incorporate the constraints on the design space into the optimisation routine, e.g. in ‘constrOptim’ in R
- We found *D*-optimal designs for the different MFP models based on the chick feeding example (3 mixture components)
- In this relatively simple setting, both methods worked well
- For the candidate set approach, we needed large candidate sets (fine ‘grids’) to ensure the required points were included

AN OPTIMAL DESIGN

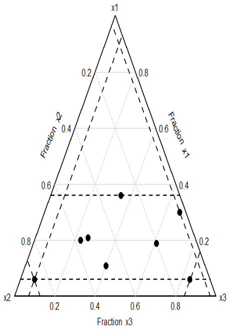

Figure: *D*-optimal design for the 2nd order MFP model (with different exponents) based on the chick feeding example.

Future Work

Open problems for future research:

- How robust are optimal designs with respect to misspecifications of the model parameters, the form of the model and the choice of component in the denominator?
- Often, the aim of the experiment is finding the optimal setting of the factors, so we should consider different optimality criteria in this case
- We could include process variables in the model, and in the design problem
- For more complicated settings, more sophisticated algorithms may be needed

Thank you!

References (1)

- Box, G. E. P. and Draper, N. R. (2007). *Response surfaces, mixtures, and ridge analyses, volume 649*, John Wiley & Sons.
- Brown, L., Donev, A., and Bissett, A. (2015). General blending models for data from mixture experiments, *Technometrics*, **57**, 449-456.
- Cornell, J. A. (1990). *Experiments with Mixtures: Designs, Models, and The Analysis of Mixture Data*, John Wiley & Sons.
- Gilmour, S. G. and Trinca, L. A. (2005). Fractional polynomial response surface models. *Journal of agricultural, biological, and environmental statistics*, **10**, 50-60.
- Khashab, R. (2018). Optimal design of experiments with mixtures. PhD Thesis, University of Southampton.

References (2)

- Khashab, R., Gilmour, S. and Biedermann, S. (2018). Fractional polynomial models for constrained mixture experiments, working paper.
- Royston, P. and Altman, D. G. (1994). Regression using fractional polynomials of continuous covariates: Parsimonious parametric modelling, *Journal of the Royal Statistical Society, Series C*, **43**, 429-467.
- Scheffé, H. (1958). Experiments with mixtures, *Journal of the Royal Statistical Society, Series B*, **20**, 344-360.
- Snee, R.D. and Marquardt, D.W. (1974). Extreme vertices designs for linear mixture models. *Technometrics*, **16**, 399-408.