

Modelagem e Otimização de Experimentos com Misturas

Edmilson R. Pinto¹
&
Leandro A. Pereira¹

¹UNIVERSIDADE FEDERAL DE UBERLÂNDIA
FACULDADE DE MATEMÁTICA

Projeto de pesquisa apoiado pela Fapemig

Contatos: *edmilson.pinto@ufu.br* e *leandro.ap@ufu.br*

19 e 20 de setembro de 2019



Sumário

- Introdução
- Modelos usuais e desenhos padrão de Scheffé
- Modelos de regressão em experimentos com mistura
- Misturas incluindo variáveis de processo
- Introdução à teoria do planejamento ótimo de experimentos
- Misturas com respostas não normais
- Projeto robusto de experimentos



Introdução

Experimentos com mistura é uma classe especial de experimentos na qual o produto sob investigação é composto de diversos componentes ou ingredientes.

Objetivos dos experimentos com misturas.

- Determinar qual é a proporção dos componentes que conduz a um resultado desejado em termos de uma variável que caracteriza a qualidade do produto;
- As variáveis não são independentes, ou seja, mudando o nível de uma variável muda-se o nível de pelo menos uma das outras variáveis.

Exemplos

Em várias indústrias o processo de fabricação consiste apenas em misturar diversos ingredientes, nas proporções adequadas para dar um produto final com as características desejadas. Você pode dar exemplos de indústrias desse tipo, de preferência na sua cidade?

Modelos usuais

- Sejam $x_i, i = 1, \dots, q$, as variáveis que representam as proporções dos q componentes da mistura. Temos então:

$$\sum_{i=1}^q x_i = 1 \text{ e } x_i \geq 0$$

Para especificar a composição da mistura, só precisamos fixar as proporções de $q - 1$ componentes. A proporção do último componente será sempre o que falta para completar 100%.

Modelos usuais

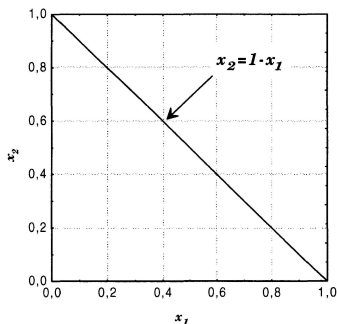
- Considere uma mistura binária (formada por apenas 2 componentes). A equação anterior se resume a:

$$x_1 + x_2 = 1$$

Que representa a equação da reta no plano cartesiano (x_1, x_2) . Todas as misturas dos 2 componentes correspondem a pontos localizados nesta reta. No estudo de misturas, portanto, o espaço experimental fica restrito aos pontos da reta, ou seja, torna-se unidimensional.

Modelos usuais

Figura: O espaço experimental para misturas de dois componentes está limitado aos pontos sobre a reta $x_2 = 1 - x_1$.



Modelos usuais

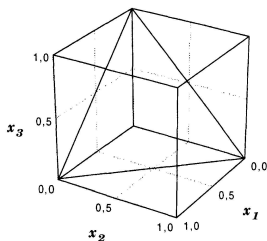
- Caso o sistema seja uma mistura de três componentes, terá de obedecer à restrição

$$x_1 + x_2 + x_3 = 1$$

Que representa que define um triângulo equilátero inscrito no cubo que representa o sistema (x_1, x_2, x_3) . Todas as misturas dos 3 componentes correspondem a pontos localizados nesta plano triangulo.

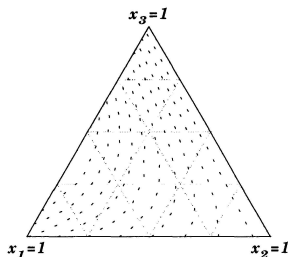
Desenhos padrão de Sheffé

Figura: Espaço experimental para misturas de três componentes. Limita-se aos pontos pertencentes ao triângulo.



Desenhos padrão de Scheffé

Figura: Uma superfície de resposta para todas as possíveis misturas dos componentes 1,2 e 3.



Modelos com 2 variáveis de mistura

O modelo mais simples para uma mistura de dois componentes é o modelo aditivo, ou linear:

$$y_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + e_i$$

onde y_i ($i = 1, \dots, n$) é a resposta observada, β_i são os parâmetros do modelo e e_i são os erros do modelo. A resposta estimada é representada por

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2$$

Sendo $\hat{\beta}_i$ são os parâmetros estimados do modelo. Porém, como x_1 é uma combinação linear de x_2 , não temos como obter as estimativas como nos modelos tradicionais.



Modelos com 2 variáveis de mistura

Sabendo-se que $x_1 + x_2 = 1$, a alternativa é reescrever o modelo da seguinte forma:

$$\hat{y}_i = \hat{\beta}_0(x_1 + x_2) + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2$$

$$\hat{y}_i = (\hat{\beta}_0 + \hat{\beta}_1)x_1 + (\hat{\beta}_0 + \hat{\beta}_2)x_2$$

$$\hat{y}_i = \hat{\beta}_1^* x_1 + \hat{\beta}_2^* x_2$$

Desta forma, o modelo passa a ter apenas dois coeficientes a serem determinados. Este é o modelo padrão de Scheffé.

Modelos com 2 variáveis de mistura

O caso mais simples de um modelo quadrático é dado por:

$$\widehat{y}_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \beta_{12} x_1 x_2$$

considerando $x_1 + x_2 = 1$, $x_1^2 = x_1(1 - x_2)$ e $x_2^2 = x_2(1 - x_1)$, a equação reescrita no modelo padrão de Scheffé, fica:

$$\widehat{y}_i = \beta_1^* x_1 + \beta_2^* x_2 + \beta_{12}^* x_1 x_2$$

sendo $\beta_i^* = \beta_0 + \beta_i + \beta_{ii}$ e $\beta_{12}^* = \beta_{12} - \beta_{11} - \beta_{22}$

Modelos com 3 variáveis de mistura

Modelo de 3 componentes linear:

$$\hat{y}_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3$$

sendo $x_1 + x_2 + x_3 = 1$, no padrão Scheffé:

$$\hat{y}_i = \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3$$

Modelos com 3 variáveis de mistura

Modelo de 3 componentes quadrático:

$$\hat{y}_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \beta_{33} x_3^2 + \beta_{12} x_1 x_2 + \beta_{13} x_1 x_3 + \beta_{23} x_2 x_3$$

Modelo de 3 componentes quadrático, Padrão Scheffé:

$$\hat{y}_i = \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{12} x_1 x_2 + \beta_{13} x_1 x_3 + \beta_{23} x_2 x_3$$

Modelos de regressão de Scheffé para misturas

Sejam x_i as variáveis de mistura ($i = 1, \dots, q$), β 's valores constantes e $Q(\beta, x)$ o preditor linear da resposta de interesse. Abaixo temos os modelos de Scheffé para misturas:

Modelo linear:

$$Q(\beta, x) = \sum_{i=1}^q \beta_i x_i$$

Modelo quadrático:

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

Modelo Cúbico Especial:

$$Q(\beta, x) = \sum_{i=1}^q \beta_i x_i + \sum_{i < j}^q \beta_{ij} x_i x_j + \sum \sum_{i < j < k}^q \beta_{ijk} x_i x_j x_k$$

Modelo Cúbico completo:

$$Q(\beta, x) = \sum_{i=1}^q \beta_i x_i + \sum_{i < j}^q \beta_{ij} x_i x_j + \sum \sum_{i < j < k}^q \beta_{ijk} x_i x_j x_k +$$

$$\sum_{i < j}^q \beta_{i-j} x_i x_j (x_i - x_j)$$

Método de Estimação

Mínimos Quadrados: Consiste em encontrar $\beta = (\beta_1, \beta_2, \beta_3, \dots)$ que minimiza a função $\sum_{i=1}^n e_i^2$ ou ee' em notação matricial, sendo :

$$e' = [e_1 \ e_2 \dots e_n]$$

$$e = y - X\beta$$

$$y' = [y_1 \ y_2 \dots y_n]$$

$$X = \begin{bmatrix} x_{11} & x_{21} & \dots & x_{k1} \\ x_{12} & x_{22} & \dots & x_{k2} \\ \dots & \dots & \dots & \dots \\ x_{1n} & x_{2n} & \dots & x_{kn} \end{bmatrix}$$

Onde os termos de interação são representados como novas variáveis na matriz X.

Método de estimação

Equação final:

$$\hat{\beta} = X'X^{-1}X'y$$

Obtenção das estimativas de MQ no R usando o pacote stats

Estimativa dos parâmetros, teste de significância e medidas de qualidade de ajuste: Comando `lm` (Linear model).

Exemplo: Comando que fornece as estimativas do modelo quadrático com 3 variáveis de mistura:

```
lm(y~-1+x1+x2+x3+x1*x2+x1*x3+x2*x3+x1^2+x2^2+x3^2, data=dados)
```



Superfícies de resposta

Consiste num método de obtenção do modelo e deslocamento, com o objetivo de atingir uma região ótima da superfície investigada. O deslocamento se dá sempre ao longo do caminho de máxima inclinação de um determinado modelo.

Exemplo

Deseja-se investigar o efeito de dois fatores no rendimento de uma reação:

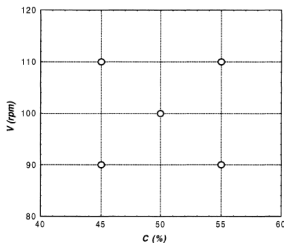
- concentração de um reagente;
- velocidade de agitação.

Já se sabe que o processo vem funcionando há algum tempo com os valores desses fatores fixados em 50% e 100 rpm, respectivamente, e que os rendimentos médios obtidos têm ficado em torno de 68%. Será que podemos obter um rendimento melhor escolhendo outros níveis dos fatores?

Exemplo

Planejamento fatorial:

Figura: Planejamento fatorial com 2 níveis e ponto central.



Exemplo

Figura: Resultado do planejamento fatorial com ponto central. $x_1 = \frac{C-50}{5}$ e $x_2 = \frac{v-100}{10}$

Ensaio	C (%)	v (rpm)	x_1	x_2	y (%)
1	45	90	-1	-1	69
2	55	90	1	-1	59
3	45	110	-1	1	78
4	55	110	1	1	67
5	50	100	0	0	68
6	50	100	0	0	66
7	50	100	0	0	69

Exemplo

Considerando primeiramente que superfície de resposta na região é função linear dos fatores, temos o seguinte modelo de regressão:

$$\hat{y}_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2$$

Ajustando por Mínimos Quadrados, temos:

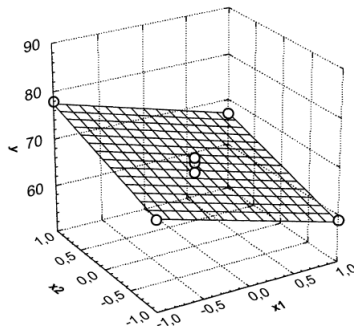
$$\hat{y}_i = 68 - 5.25x_1 + 4.25x_2$$

Sendo os coeficientes significativos segundo o teste de significância

Exemplo

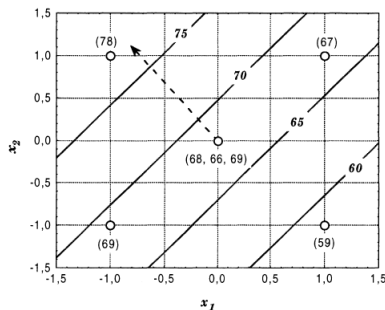
Caso o modelo escolhido seja adequado, temos a superfície de resposta:

Figura: Superfície de resposta (Plano descrito pela equação ajustada).



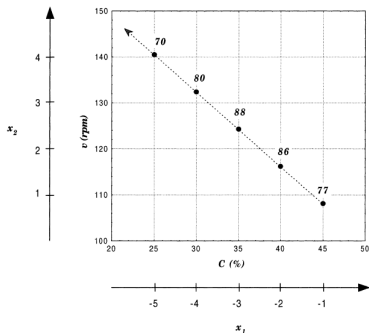
Exemplo

Figura: Curvas de nível do plano descrito pela Equação linear. A linha tracejada é a trajetória de máxima inclinação partindo do ponto central do planejamento. Os valores entre parênteses são as respostas determinadas experimentalmente.



Exemplo

Figura: Resultados dos ensaios realizados na trajetória de máxima inclinação.



Exemplo

Notamos que no ponto sobre coordenada $(-3, 2.43)$ que corresponde a $(35\%, 125)$ em escala original, tivemos a resposta máxima, assim um novo planejamento centrado em $(35\%, 125)$ é realizado

Figura: Resultados de um novo planejamento 2^2 com ponto central

Ensaio	$C(\%)$	$v(rpm)$	x_1	x_2	$y(\%)$
1	30	115	-1	-1	86
2	40	115	1	-1	85
3	30	135	-1	1	78
4	40	135	1	1	84
5	35	125	0	0	90
6	35	125	0	0	88
7	35	125	0	0	89

Exemplo

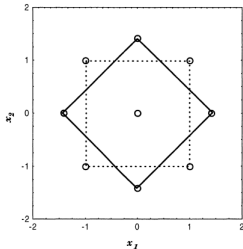
Uma análise da qualidade do ajustamento mostrou que o modelo linear já não é mais adequado para modelar a superfície de resposta. Assim partiu-se para um modelo quadrático:

$$\hat{y}_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \beta_{33} x_3^2 + \beta_{12} x_1 x_2 + \beta_{13} x_1 x_3 + \beta_{23} x_2 x_3$$

Exemplo

Porém este modelo tem 6 parâmetros e temos 5 níveis experimentais. Dessa forma adicionou-se ao experimento pontos adicionais.

Figura: Resultados de um novo planejamento 2^2 com ponto central



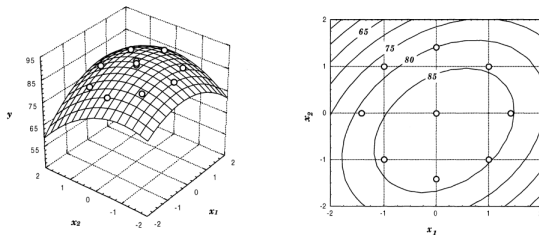
Ensaio	C (%)	v (rpm)	x_1	x_2	y (%)
1	30	115	-1	-1	86
2	40	115	1	-1	85
3	30	135	-1	1	78
4	40	135	1	1	84
5	35	125	0	0	90
6	35	125	0	0	88
7	35	125	0	0	89
8	28	125	$-\sqrt{2}$	0	81
9	35	139	0	$\sqrt{2}$	80
10	42	125	$\sqrt{2}$	0	86
11	35	119	0	$-\sqrt{2}$	87

Exemplo

Modelo ajustado:

$$\hat{y}_i = 89 + 1,51x_1 + 2,36x_2 + 2,81x_1^2 + 2,81x_2^2 + 1,75x_1x_2.$$

Figura: Superfície quadrática descrita pela Equação e suas curvas de nível.



A região contém um ponto de máximo, situado aproximadamente em $(0.15, -0.37)$ ou $(36\%, 121)$, representando um rendimento de 89,6%,

Exemplo

Mistura de 3 componentes: Determinar a composição da membrana que produzisse o maior sinal analítico.

Figura: Estudo de membranas para fabricação de eletrodo seletivo. O sinal é a altura do pico observada.

i	x_1	x_2	x_3	Sinal			$\bar{y}_i$	s_i^2
1	1	0	0	3,2	3,0		3,10	0,020
2	0	1	0	0,5	0,4		0,45	0,005
3	0	0	1	0,4	0,3		0,35	0,005
4	1/2	1/2	0	1,9	1,2	2,0	1,70	0,190
5	1/2	0	1/2	3,9	4,4	4,1	4,13	0,063
6	0	1/2	1/2	0,3	0,3	0,2	0,27	0,003

Exemplo

Modelo quadrático estimado:

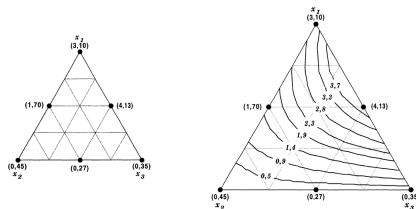
$$\hat{y}_i = 3.10x_1 + 0.45x_2 + 0.35x_3 - 0.3x_1x_2 + 9.62x_1x_3 + -0.52x_2x_3$$

Após selecionar os coeficientes significativos, temos o modelo final:

$$\hat{y}_i = 3.10x_1 + 9.62x_1x_3$$

Exemplo

Figura: Planejamento em rede simplex e sinais analíticos médios observados para as misturas representadas pelos pontos e Curvas de nível.



Pela Equação, o valor máximo do sinal analítico para o tipo de membrana estudado deve ser 4,2 cm, e deve ser obtido com uma membrana contendo somente os componentes 1 e 3, na proporção 66% - 34%.

Experimentos de mistura com Restrição

Em alguns casos, pode ser necessário restringir a proporção de um ou mais componentes, que, por motivos técnicos ou práticos, podem não contemplar todas as proporções possíveis, que correspondem ao intervalo fechado entre 0 e 1. O novo espaço experimental passa a ser uma sub-região do espaço original Os limites superiores e/ou inferiores nas proporções são representados da seguinte forma:

$$0 \leq L_i \leq x_i \leq U_i \leq 1, i = 1, \dots, q$$

Experimentos de mistura com Restrição - Comando Xvert

Com as restrições, o desenho experimental altera, sendo portanto um sub-conjunto no desenho em que não temos restrições.

Comando Xvert (Pacote Mixexp): cria a região experimental com as restrições, fornece o desenho e os pontos.

Experimentos de mistura com Restrição- Exemplo

Formulação de um xampu, no qual Lauril Sulfato (A), Cocamida (B) e Lauramida (C) são os ingredientes do produto. Os três ingredientes compõem 50% do xampu, onde os outros componentes, que foram mantidos fixos em todas as misturas, eram água, perfume e corante.

Resposta: Altura da espuma

Objetivo: formular um produto no qual a altura da espuma excedesse 170,0mm

Restrições:

$$0,4 \leq x_1 \leq 0,6$$

$$0,14 \leq x_2 \leq 0,2$$

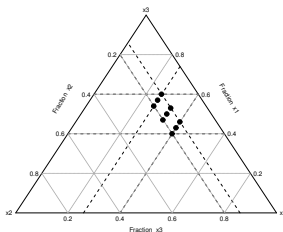
$$0,26 \leq x_3 \leq 0,4$$

Experimentos de mistura com Restrição- Exemplo

Comando:

```
Xvert(nfac = 3, lc = c(0.4, 0.14, 0.26), uc = c(0.6, 0.2, 0.4), ndm = 1,
pseudo = FALSE)
```

Figura: Planejamento em rede simplex com restrições



Experimentos de mistura com Restrição- Exemplo

Tabela:

	x1	x2	x3
1	0.6	0.14	0.26
2	0.4	0.20	0.40
3	0.54	0.20	0.26
4	0.46	0.14	0.40
5	0.53	0.14	0.33
6	0.47	0.20	0.33
7	0.57	0.17	0.26
8	0.43	0.17	0.40
9	0.5	0.17	0.33

Análise de resíduos

Análise de conjunto de técnicas utilizadas para investigar a adequabilidade de um modelo de regressão com base nos resíduos. Nos modelos tradicionais de regressão, definimos o i -ésimo resíduo ($i = 1, \dots, n$) como:

$$e_i = y_i - \hat{y}_i$$

A ideia básica da análise dos resíduos é que, se o modelo for apropriado, os resíduos devem refletir as propriedades impostas pelo termo de erro do modelo. Tais suposições são:

- e_i e e_j são independentes (Independência);
- $e_i \sim N(0, \sigma^2)$ (Homocedasticidade e Normalidade);
- Não existem pontos atípicos (Influentes).

Análise de Resíduos

- Técnicas para verificar Independência:
 - Gráfico resíduos versus a ordem de coleta dos dados;
 - Teste de Durbin-Watson.
- Teste para verificar Homocedasticidade:
 - Gráfico dos resíduos versus valores ajustados;
 - Teste de Breusch-Pagan e Goldfeld-Quandt.
- Teste para verificar Normalidade:
 - Gráfico de probabilidade normal (QQ-plot);
 - Histograma dos resíduos;
 - Testes de Normalidade (Anderson Darling, Kolmogorov-Smirnov, etc).
- Teste para verificar pontos atípicos:
 - Gráfico de Distâncias de Cook (QQ-plot);
 - Gráfico de Leverage;



Técnicas de Seleção de modelos

Existem vários procedimentos para a seleção de modelos de regressão.
Os procedimentos mais conhecidos são

- Backward
- Forward
- Stepwise
- AIC
- BIC
- Outros.

Técnicas de Seleção de modelos

- Forward: As variáveis são acrescentadas no modelo uma a uma, sempre escolhendo a de maior relevância. O procedimento para quando a de maior relevância não for significativa;
- Backward: A partir do modelo com todas as variáveis, retira-se uma a uma, sempre escolhendo a de menor relevância. O procedimento para quando a de menor relevância for significativa;
- Stepwise: para cada interação do método Forward, aplica-se uma interação do método Backward;



Técnicas de Seleção de modelos

- AIC: Akaike Information Criterion: $AIC = -2\log(L) + 2k$
- BIC: Bayesian Information Criterion: $BIC = -2\log(L) + k\ln(n)$

onde k é o numero de parâmetros estimados e L_p a função de verossimilhança aplicada nas estimativas dos parâmetros. Ambos Penalizam modelos com muitas variáveis. Valores menores de AIC e BIC são preferíveis.

Misturas incluindo variáveis de processo

Problemas de mistura incluindo variáveis de processo surgem quando em um experimento de mistura a propriedade de interesse é função das proporções dos ingredientes e de outros fatores que não formam parte alguma da mistura, como, por exemplo, tempo e temperatura. Esses fatores, os quais dependem das condições de processo, são chamados de variáveis de processo.

Nestes problemas o objetivo é determinar as proporções dos componentes de mistura, juntamente com as condições de processo, a fim de satisfazer algumas propriedades ou características de interesse envolvendo a variável resposta.

Modelos utilizados

Desta forma, se em adição aos q componentes de mistura $\mathbf{x}^t = (x_1, \dots, x_q)$ existirem r variáveis de processo $\mathbf{z}^t = (z_1, \dots, z_r)$, podemos considerar modelos como:

- $\eta(\mathbf{x}, \mathbf{z}, \boldsymbol{\beta}, \boldsymbol{\gamma}) = \zeta(\mathbf{x}, \boldsymbol{\beta}) + \vartheta(\mathbf{z}, \boldsymbol{\gamma})$ (modelo aditivo)
- $\eta(\mathbf{x}, \mathbf{z}, \boldsymbol{\beta}, \boldsymbol{\gamma}) = \zeta(\mathbf{x}, \boldsymbol{\beta}) \vartheta(\mathbf{z}, \boldsymbol{\gamma})$ (modelos produto cruzado)
- $\eta(\mathbf{x}, \mathbf{z}, \boldsymbol{\beta}, \boldsymbol{\gamma}) = \zeta(\mathbf{x}, \boldsymbol{\beta}) + \nu(\mathbf{x}, \mathbf{z}, \boldsymbol{\beta}, \boldsymbol{\gamma})$, (combinações dos modelos anteriores)

onde $\vartheta(\mathbf{z}, \boldsymbol{\gamma})$ representa o modelo para as variáveis de processo e $\nu(\mathbf{x}, \mathbf{z}, \boldsymbol{\beta}, \boldsymbol{\gamma})$ compreende produtos de termos em $\zeta(\mathbf{x}, \boldsymbol{\beta})$ and $\vartheta(\mathbf{z}, \boldsymbol{\gamma})$.

Exemplo

- Modelo multiplicativo combinado, o qual inclui o modelo cúbico de Scheffé para mistura e o modelo quadrático reduzido para as variáveis de processo.
- Seja uma mistura com três componentes (x_1, x_2, x_3) e duas variáveis de processo (z_1, z_2).
- $\zeta(\mathbf{x}) = \sum_{i=1}^3 \beta_i x_i + \sum_{i=1}^2 \sum_{j>i}^3 \beta_{ij} x_i x_j + \beta_{123} x_1 x_2 x_3 + \sum_{i=1}^2 \sum_{j>i}^3 \gamma_{ij} x_i x_j (x_i - x_j)$
- $\vartheta(\mathbf{z}) = \alpha_0 + \sum_{i=1}^2 \alpha_i z_i + \sum_{i=1}^2 \alpha_{ii} z_i^2 + \alpha_{12} z_1 z_2.$

Modelo multiplicativo combinado

$$\begin{aligned}
 \eta(\mathbf{x}, \mathbf{z}) = & \sum_{i=1}^3 \beta_i^0 x_i + \sum_{i=1}^2 \sum_{j>i}^3 \beta_{ij}^0 x_i x_j + \beta_{123}^0 x_1 x_2 x_3 + \sum_{i=1}^2 \sum_{j>i}^3 \gamma_{ij}^0 x_i x_j (x_i - \\
 & x_j) + \sum_{l=1}^2 \left[\sum_{i=1}^3 \beta_i^l x_i + \sum_{i=1}^2 \sum_{j>i}^3 \beta_{ij}^l x_i x_j + \beta_{123}^l x_1 x_2 x_3 + \right. \\
 & \left. \sum_{i=1}^2 \sum_{j>i}^3 \gamma_{ij}^l x_i x_j (x_i - x_j) \right] z_l + \sum_{l=1}^2 \left[\sum_{i=1}^3 \beta_i^{ll} x_i + \sum_{i=1}^2 \sum_{j>i}^3 \beta_{ij}^{ll} x_i x_j + \right. \\
 & \left. \beta_{123}^{ll} x_1 x_2 x_3 + \sum_{i=1}^2 \sum_{j>i}^3 \gamma_{ij}^{ll} x_i x_j (x_i - x_j) \right] z_l^2 + \left[\sum_{i=1}^3 \beta_i^{12} x_i + \right. \\
 & \left. \sum_{i=1}^2 \sum_{j>i}^3 \beta_{ij}^{12} x_i x_j + \beta_{123}^{12} x_1 x_2 x_3 + \sum_{i=1}^2 \sum_{j>i}^3 \gamma_{ij}^{12} x_i x_j (x_i - x_j) \right] z_1 z_2
 \end{aligned}$$



CNIMAC 2019
Centro Nacional de Modelagem e Otimização em Experimentos
 Universidade Federal de Uberlândia - UFU

Planejamento experimental

Em geral, a metodologia usada para construir desenhos experimentais para mistura envolvendo variáveis de processo é uma combinação de dois desenhos experimentais, sendo um desenho experimental de mistura e o outro sendo um desenho fatorial ou fatorial fracionário para as variáveis de processo.

A seguir um exemplo de um desenho fatorial completo 3^2 cruzado com o desenho simplex-centróide de terceira ordem em três componentes de mistura.

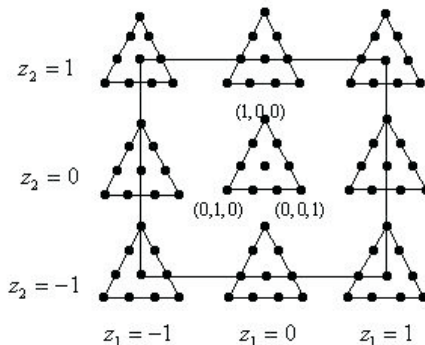


Figura: Fatorial completo 3^2 em duas variáveis de processo cruzado com o desenho simplex-centróide de terceira ordem em três componentes de mistura

Misturas incluindo variáveis de ruído

- Quando algumas variáveis de processo são incontroláveis ou difícil de ser controladas elas deveriam ser tratadas como variáveis de ruído.
- Variáveis de ruído são consideradas como variáveis aleatórias com uma distribuição de probabilidade conhecida. A presença de variáveis de ruído no modelo significa que a variância não pode ser considerada como constante.
- Uma forma de controlar a variabilidade presente no processo é modelando a variância.

Planejamento ótimo de experimentos

- Planejamento ótimo de experimentos é uma poderosa e flexível ferramenta para obter desenhos experimentais eficientes, proporcionando redução no número de ensaios experimentais e, desta forma, reduzindo custos e economizando tempo e dinheiro.
- O desenho ótimo depende do modelo considerado e se o modelo não for apropriado, o desenho experimental também não será apropriado.
- A construção de desenhos ótimos para experimentos com mistura são de fundamental importância, especialmente quando os desenhos tradicionais não puderem ser aplicados, por exemplo, quando temos mistura com restrição, onde a região experimental é parte da região simplex.



Desenho ótimo para modelos clássicos de regressão

- Considere o modelo linear $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$
- $\mathbf{y}^t = (y_1, \dots, y_n)$ é o vetor resposta, $\mathbf{X}$ é uma conhecida matriz de planejamento de ordem $n \times p$ e posto $p \leq n$, $\boldsymbol{\beta}$ é um vetor de parâmetros desconhecidos, e $\boldsymbol{\epsilon}^t = (\epsilon_1, \dots, \epsilon_n)$ é um vetor de erros aleatórios com média zero e matriz de variância-covariância dada por $\sigma^2 \mathbf{I}$, onde $\sigma^2 = \text{Var}(\epsilon_i)$ para $i = 1, \dots, n$ e $\mathbf{I}$ é a matriz identidade.
- O valor da resposta em um ponto $\mathbf{x}^t = (x_1, \dots, x_k)$ na região de interesse $\mathcal{X}$ é dado por $y(\mathbf{x}) = \eta(\mathbf{x}, \boldsymbol{\beta}) + \epsilon$, onde $\eta(\mathbf{x}, \boldsymbol{\beta}) = \mathbf{f}^t(\mathbf{x})\boldsymbol{\beta}$ é um modelo polinomial nas variáveis de controle $x_1, \dots, x_k$.



- Suponha, no caso geral, que temos k unidades experimentais onde n respostas devem ser observadas. O conjunto de vetores $\{\mathbf{x}_i, i = 1, \dots, k\}$ representa uma possível escolha de k unidades experimentais, nas quais n respostas serão observadas.
- O desenho experimental pode ser representado por

$$\xi = \begin{Bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \dots & \mathbf{x}_k \\ p_1 & p_2 & \dots & p_k \end{Bmatrix}$$

com $p_j = \frac{r_j}{n}$, onde r_j é o número de replicações na j -ésima unidade experimental, para $j = 1, \dots, k$.

- As unidades experimentais $\mathbf{x}_1, \dots, \mathbf{x}_k$ são os pontos suporte do experimento e $p_1, \dots, p_k$ são os pesos associados aos pontos suporte.

- Para o desenho experimental ξ , o estimador de mínimos quadrados ponderados para β é dado por $\hat{\beta} = (X^t P X)^{-1} X^t P y$, onde a matriz diagonal $P = n \text{ Diag}\{p_1, \dots, p_k\}$.
- Da literatura, sabe-se que a matriz de variância-covariância de $\hat{\beta}$ is é dada por $\text{Cov}(\hat{\beta}) = \sigma^2 (X^t P X)^{-1}$, o qual é a inversa da matriz de informação de Fisher.
- Na teoria de planejamento ótimo de experimento a matrix de informação de Fisher é de fundamental importância.
- Defina a matriz experimental como $M(\xi) = X^t P X$. Desta forma, o desenho ótimo pode ser obtido a partir de $M(\xi)$, de acordo com algum critério de otimalidade.



Observações

- No contexto da teoria de regressão linear, se assumirmos normalidade dos erros, podemos obter um elipsóide de confiança para β , com uma dada confiança e para uma dada soma de quadrados de resíduos, que depende da matriz de planejamento X (e, por conseguinte dos pontos do desenho x), da seguinte forma: $\{\beta; (\beta - \hat{\beta})^t M^{-1}(x)(\beta - \hat{\beta}) \leq C\}$.
- Pode ser mostrado que o conteúdo desse elipsóide é proporcional a $|M(x)|^{-\frac{1}{2}}$, desta forma, se maximizarmos $\ln |M(x)|$ minimizaremos o volume desse elipsóide e teremos o critério D-ótimo (D de determinante)

Exemplo - Atkinson et al. (2007)

Considere o modelo de regressão linear simples $\eta(x, \beta) = \beta_0 + \beta_1 x$ e sejam os seguintes desenhos experimentais

$$\xi_1 = \left\{ \begin{array}{cc} -1 & 1 \\ \frac{6}{8} & \frac{2}{8} \end{array} \right\}$$

e

$$\xi_2 = \left\{ \begin{array}{cccccc} -1 & -0.90 & -0.85 & -0.80 & -0.75 & 1 \\ \frac{2}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} \end{array} \right\}$$

Exemplo - Atkinson et al. (2007)

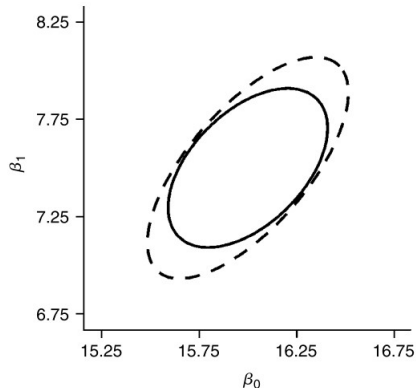


Figura: Elipses de confiança para β_0 e β_1 . ξ_1 (—) e ξ_2 (- - -)

Dois critérios de otimalidade

- **Critério da D-otimalidade**

Um experimento ξ^* é D-ótimo se $\ln |\mathbf{M}(\xi^*)| = \max_{\xi} \ln |\mathbf{M}(\xi)|$.

- A variância do valor predito em $\mathbf{x}$, $\hat{y}(\mathbf{x}) = \mathbf{f}^t(\mathbf{x})\hat{\boldsymbol{\beta}}$, é dada por $\text{Var}(\hat{y}(\mathbf{x})) = \sigma^2 \mathbf{f}^t(\mathbf{x})\mathbf{M}^{-1}(\xi)\mathbf{f}(\mathbf{x})$.

- Defina a variância padronizada como: $d(\mathbf{x}, \xi) = \mathbf{f}^t(\mathbf{x})\mathbf{M}^{-1}(\xi)\mathbf{f}(\mathbf{x})$.

- **Critério da G-otimalidade**

Um experimento ξ^* é G-ótimo se $d(\mathbf{x}, \xi^*) = \min_{\xi} \max_{\mathbf{x}} d(\mathbf{x}, \xi)$.

- A G-otimalidade está relacionada com as previsões que podem ser obtidas para o valor esperado da variável resposta. O critério da G-otimalidade consiste em determinar o experimento para o qual a variância da previsão menos precisa, em toda região experimental será mínima.



Teorema Geral da Equivalência

Todas as condições, relacionadas abaixo, envolvendo o experimento ξ^* são equivalentes.

1. O experimento ξ^* é D-ótimo.
2. $d(\mathbf{x}, \xi^*) \leq p, \forall \mathbf{x} \in \mathcal{X}$.
3. $\max_{\mathbf{x}} d(\mathbf{x}, \xi^*) = p$
4. $d(\mathbf{x}, \xi^*)$ atinge o máximo nos pontos de suporte de ξ^* .
5. O experimento ξ^* é G-ótimo.

Obs.: p é o número de parâmetros do modelo.

Exemplo 1 - Regressão linear simples

- Seja o modelo de regressão linear simples, cuja parte sistemática é dada pela expressão: $\eta(x, \beta) = \beta_0 + \beta_1 x$, onde $x \in [-1, 1]$.
- Desta forma, tomando o desenho: $\xi = \begin{Bmatrix} x_1 & x_2 \\ p_1 & p_2 \end{Bmatrix}$,
 $-1 \leq x_1, x_2 \leq 1, 0 \leq p_1, p_2 \leq 1$
- A matriz de informação será:

$$\begin{aligned} M(\xi) &= \begin{pmatrix} 1 & 1 \\ x_1 & x_2 \end{pmatrix} \begin{pmatrix} p_1 & 0 \\ 0 & p_2 \end{pmatrix} \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \end{pmatrix} = \\ &= \begin{pmatrix} p_1 + p_2 & p_1 x_1 + p_2 x_2 \\ p_1 x_1 + p_2 x_2 & p_1 x_1^2 + p_2 x_2^2 \end{pmatrix} \end{aligned}$$



Desta forma, $|\mathbf{M}(\xi)| = (p_1 x_1^2 + p_2 x_2^2) - (p_1 x_1 + p_2 x_2)^2$. Assim, queremos encontrar x_1, x_2, p_1, p_2 de modo que a função critério $\phi(\mathbf{M}(\xi)) = \ln |\mathbf{M}(\xi)|$ sujeita a $p_1 + p_2 = 1, 0 \leq p_1, p_2 \leq 1, -1 \leq x_1, x_2 \leq 1$, seja maximizada. Isto é,

$$\begin{array}{ll} \max \ln |\mathbf{M}(\xi)| & \\ \text{Sujeito a} & \left\{ \begin{array}{l} p_1 + p_2 = 1.0 \\ -1 \leq x_1 \leq 1 \\ -1 \leq x_2 \leq 1 \\ 0 \leq p_1 \leq 1 \\ 0 \leq p_2 \leq 1 \end{array} \right. \end{array}$$

Exemplo 2

Seja o modelo de regressão quadrática homocedástico tal que $\eta(\mathbf{x}, \xi) = \beta_0 + \beta_1 x + \beta_2 x^2$, onde $x \in [-1, 1]$. A região experimental é $\mathcal{X} = [-1, 1]$, $\mathbf{f}^t(x) = (1, x, x^2)$, $\boldsymbol{\beta}^t = (\beta_0, \beta_1, \beta_2)$. Para este modelo o desenho ótimo é

$$\xi^* = \left\{ \begin{array}{ccc} -1 & 0 & 1 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{array} \right\}$$

Gráfico da variância padronizada

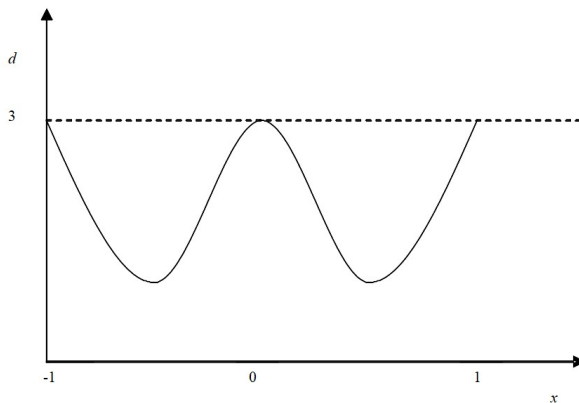


Figura: Gráfico da variância padronizada versus os valores de x

CNIMAC 2019
2019 - 10ª Conferência Nacional em Modelagem e Otimização
10 a 12 de Setembro de 2019 - Universidade Federal de Uberlândia - Uberlândia, MG

Exemplo 3 - Desenho ótimo para mistura com restrição

Exemplo retirado de Atkinson et al. (2007). Considere três variáveis de mistura x_1 , x_2 and x_3 e suponha que o modelo é o polinomial canônico de segunda ordem, dado por: $\eta(\mathbf{x}, \boldsymbol{\beta}) = \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{12} x_1 x_2 + \beta_{13} x_1 x_3 + \beta_{23} x_2 x_3$. Para as variáveis de mistura existem as seguintes restrições: $0.2 \leq x_1 \leq 0.7$, $0.1 \leq x_2 \leq 0.6$ and $0.1 \leq x_3 \leq 0.6$. O desenho D-ótimo é:

$$\xi^* = \left\{ \begin{array}{ccccc} (0.7, 0.1, 0.2) & (0.2, 0.6, 0.2) & (0.7, 0.2, 0.1) & (0.2, 0.2, 0.6) & (0.3, 0.6, 0.1) \\ 0.10 & 0.10 & 0.10 & 0.10 & 0.10 \\ (0.3, 0.1, 0.6) & (0.2, 0.4, 0.4) & (0.5, 0.1, 0.4) & (0.5, 0.4, 0.1) & (0.4, 0.3, 0.3) \\ 0.10 & 0.11 & 0.11 & 0.11 & 0.07 \end{array} \right\}$$



Desenho ótimo para mistura com restrição

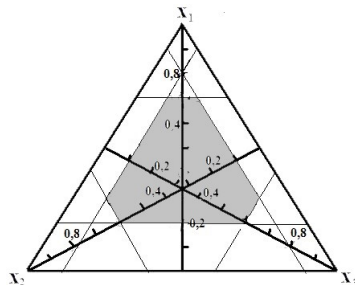


Figura: Região experimental para mistura com restrição

Desenho ótimo no R usando Mixexp

Desenho D-ótimo usando o pacote Mixexp do R

- Criando a região experimental com restrição

$$RegExp <- -Xvert(nfac = 3, lc = c(0.2, 0.1, 0.1), uc = c(0.7, 0.6, 0.6), ndm = 1, pseudo = FALSE)$$
- Obtendo o desenho D-ótimo aproximado.

$$Dotimo <- -optFederov(x1 + x2 + x3 + x1 : x2 + x1 : x3 + x2 : x3 - 1, RegExp, approximate = TRUE)$$

Região experimental

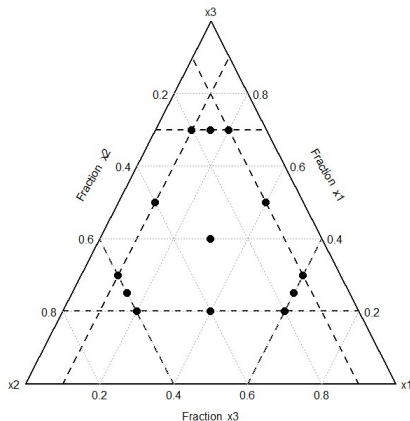


Figura: Região experimental para mistura com restrição

Desenho D-ótimo no R - OptFedorov

	\$design				
	Proportion	x1	x2	x3	dimen
1	0.1008	0.7	0.1	0.2	0
2	0.1008	0.2	0.6	0.2	0
3	0.1008	0.7	0.2	0.1	0
4	0.1008	0.3	0.6	0.1	0
5	0.1008	0.3	0.1	0.6	0
6	0.1008	0.2	0.2	0.6	0
7	0.1083	0.2	0.4	0.4	1
9	0.1083	0.5	0.1	0.4	1
11	0.1083	0.5	0.4	0.1	1
13	0.0703	0.4	0.3	0.3	2

Figura: Desenho D-ótimo para mistura com restrição

Desenho ótimo no R usando função D.otimo.mix

- A função `D.otimo.mix` foi criada por Pinto e Pereira (2010) para obtenção de desenhos ótimos de mistura com restrição.
- `Dotimo <- -d.otimo.mix(c(npont, nvar, npar), c(xi), c(wi), c(modelo), c(li), c(ls))`
- onde: *npont*: número de pontos iniciais; *nvar*: número de variáveis de mistura; *npar*: número de parâmetros no modelo; *xi* e *wi* desenho inicial; *modelo*: modelo proposto; *li*: limite inferior; *ls*: limite superior.
- `Dotimo <- -d.otimo.mix(c(13, 3, 6), c(0.6, 0.3, 0.1, 0.2, 0.6, 0.2, 0.7, 0.2, 0.1, 0.4, 0.3, 0.3, 0.3, 0.4, 0.3, 0.2, 0.3, 0.5), c(0.076, 0.076, 0.076, 0.076, 0.076, 0.076, 0.076, 0.076, 0.076, 0.076, 0.076, 0.076, 0.088), c("x1", "x2", "x3", "x1 * x2", "x1 * x3", "x2 * x3"), c(0.2, 0.1, 0.1), c(0.7, 0.6, 0.6))`

Modelos Lineares Generalizados (MLG)

- Definido por uma função de distribuição de probabilidade membro da família exponencial para a variável resposta. (Componente aleatório com $E(Y) = \mu$).
- A classe dos modelos lineares generalizados é dada por $f(y; \theta; \phi) = \exp\{\frac{1}{\phi}[y\theta - b(\theta)] - c(y, \phi)\}$.
- Um conjunto de covariáveis independentes descrevendo a estrutura linear do modelo. (Componente sistemático: $\eta = X\beta$)
- Um função de ligação entre a média da variável resposta e a estrutura linear. (função de ligação: $g(\mu) = \eta$).
- Nos MLG, temos que $E(Y) = \mu = b'(\theta)$ e $Var(Y) = \phi b''(\theta) = \phi V(\mu)$.

Modelos Lineares Generalizados

Exemplos

- Modelo normal:** $f(y; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left\{ \frac{-1}{2\sigma^2} (y - \mu)^2 \right\} = \exp \left\{ \frac{1}{\sigma^2} \left[y\mu - \frac{\mu^2}{2} \right] - \frac{1}{2} \log(2\pi\sigma^2) - \frac{y^2}{2\sigma^2} \right\}$. Desta forma, $\theta = \mu$, $\phi = \sigma^2$, $b(\theta) = \frac{\theta^2}{2}$, $V(\mu) = 1$
- Modelo de Poisson:**
 $f(y; \mu) = \frac{\mu^y e^{-\mu}}{y!} = \exp \left\{ \frac{1}{1} (y \log \mu - \mu) - \log y! \right\}$. Desta forma, $\theta = \log \mu$, $\phi = 1$, $b(\theta) = \mu = e^\theta$, $V(\mu) = \mu$.
- Outros modelos: Binomial, Gama, Normal inversa, Binomial negativa, etc.



Algoritmo de estimação dos MLG

A estimativa de β é obtida usando o seguinte processo.

I. Enquanto a convergência não for atingida, obtenha

$$\beta^{(m+1)} = \left(X^t W^{(m)} X \right)^{-1} X^t W^{(m)} z^m$$

II. Quando a convergência for atingida, pare. O valor corrente de β é o parâmetro estimado.

- Um critério de convergência pode ser: $\|\beta^{(m+1)} - \beta^{(m)}\| < \epsilon$, $\|\cdot\|$ é a norma de um vetor e $\epsilon > 0$
- z é um vetor $n \times 1$ e $W = \text{Diag}\{w_i\}$ é uma matriz diagonal $n \times n$
- Inicie com os valores $\mu_i^{(1)} = y_i$ e obtenha as estimativas $\eta_i^{(1)} = g(y_i)$ e $w_i^{(1)} = \frac{1}{V(y_i)[g'(y_i)]^2}$ para $i = 1, \dots, n$.

Algoritmo de estimação dos MLG

Em cada passo m , a obtenção dos valores de η_i , μ_i , z_i e w_i , para $i = 1, \dots, n$, é dada por:

- $\eta_i^{(m)} = \sum_{j=1}^p x_{ij} \beta_j^{(m)}$
- $\mu_i^{(m)} = g^{-1} \left(\eta_i^{(m)} \right)$
- $z_i^{(m)} = \eta_i^{(m)} + \left(y_i - \mu_i^{(m)} \right) g' \left(\mu_i^{(m)} \right)$
- $w_i^{(m)} = \frac{1}{V \left(\mu_i^{(m)} \right) \left[g' \left(\mu_i^{(m)} \right) \right]^2}$

Estimação de MLG no R

- No software R a estimação de parâmetros do MLG é obtida usando a função `glm`. Existem vários argumentos que podem ser considerados na função `glm`. O mais simples é:
- `glm(formula, family, data)`
- Para maiores detalhes veja documentação no R.

Comentários sobre desenhos ótimos

- No caso de MLG (modelo não linear), a matriz de informação $\mathbf{M}(\xi)$ depende dos parâmetros do modelo.
- Caso contínuo: $\mathbf{M}(\xi, \theta) = \int_{\mathcal{X}} w(\theta, \mathbf{x}) \mathbf{f}(\mathbf{x}) \mathbf{f}^t(\mathbf{x}) d\xi(\mathbf{x})$
- Caso discreto: $\mathbf{M}(\xi, \theta) = \mathbf{X}^t \mathbf{W} \mathbf{P} \mathbf{X}$
- O teorema da equivalência continua válido.
- Teoria Bayesiana: a função critério D-ótimo é $\phi_{\theta}(\xi) = E_{\theta} \{\ln |\mathbf{M}(\theta, \xi)|\}$ e deve ser dada uma priori para θ .
- $d(\xi, \theta) = \max_{\mathcal{X}} w(\theta, \mathbf{x}) \mathbf{f}^t(\mathbf{x}) [\mathbf{M}(\theta, \xi)]^{-1} \mathbf{f}(\mathbf{x})$



Exemplo - Modelo dados binários

- Seja $Y \sim \text{Bin}(m, \pi)$
- $\eta(x, \beta) = \beta_0 + \beta_1 x$
- Função de ligação logística: $g(\pi) = \ln\left(\frac{\pi}{1-\pi}\right)$.
- Suponha $\beta = \{(\beta_0, \beta_1); \beta_0 \in [-1, 3] \text{ e } \beta_1 \in [1, 3]\}$
- Considere as seguintes prioris discretas
- $P_1 = \begin{Bmatrix} (-1, 1) & (-1, 3) & (3, 1) & (3, 3) \\ 0.25 & 0.25 & 0.25 & 0.25 \end{Bmatrix}$
- $P_2 = \begin{Bmatrix} (-1, 1) & (-1, 2) & (-1, 3) & (1, 1) & (1, 2) & (1, 3) & (3, 1) & (3, 2) & (3, 3) \\ 0.11 & 0.11 & 0.11 & 0.11 & 0.12 & 0.11 & 0.11 & 0.11 & 0.11 \end{Bmatrix}$



Desenhos D-ótimos Bayesianos

- Desenho D-ótimo Bayesiano com respeito à priori P_1

- $\xi_1^* = \begin{Bmatrix} -4.1097 & -1.4327 & -0.3363 & 0.9438 \\ 0.1255 & 0.2595 & 0.3387 & 0.2763 \end{Bmatrix}$

- Desenho D-ótimo Bayesiano com respeito à priori P_2

- $\xi_2^* = \begin{Bmatrix} -3.6640 & -1.6280 & -0.6407 & -0.1210 & 0.8347 \\ 0.0603 & 0.2778 & 0.1972 & 0.1814 & 0,2833 \end{Bmatrix}$

Exemplo - Mistura com dados de contagem

Este exemplo diz respeito a um projeto relacionado a uma disciplina envolvendo planejamento de experimentos e modelagem estatística, elaborado por professores e alunos de uma universidade no Canadá. Os alunos executaram o experimento planejado, coletaram os dados, propuseram uma estratégia de modelagem e conduziram a análise desses dados. O Experimento realizado é relacionado a um problema cuja resposta é uma contagem e, portanto, o modelo normal não é adequado.

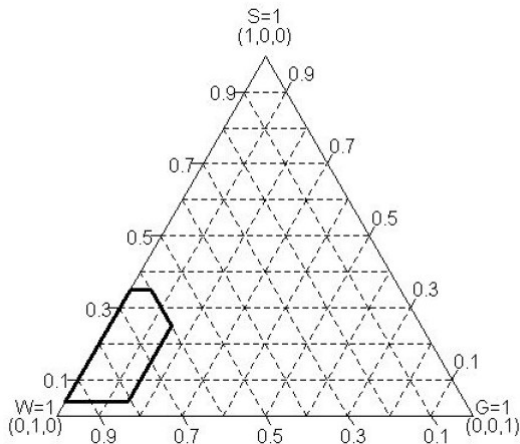
O Experimento

- O objetivo do experimento foi encontrar uma mistura ótima de água (W), sabão (S) e glicerina (G) que produzisse o maior número de bolhas de sabão.
- Foram consideradas duas variáveis de processo: tipo de água (N_W): Spring (-1) e Tap (1) e tipo de sabão (N_S): Joy (-1) e Ivory (1)
- Os alunos misturaram as 48 soluções de água, sabão e glicerina com o tipo de água apropriado e marca de sabão, cada solução com um volume de 100 ml.
- O experimento foi realizado em uma tarde e nas mesmas condições de umidade e temperatura. Para cada solução foram realizadas 5 repetições e o número de bolhas obtido foi anotado.
- Não houve replicações. A mistura foi a mesma para todas as repetições do processo.

Restrições na mistura

- De acordo com manuais, foram consideradas restrições para as variáveis de mistura.
- As restrições de mistura consideradas foram: $0.04 \leq S \leq 0.35$, $0.60 \leq W \leq 0.98$ e $0 \leq G \leq 0.15$.
- O desenho nas variáveis de processo foi um fatorial completo 2^2 .
- Desenhos ótimos não foram considerados.

Região experimental - Mistura com dados de contagem



Desenho experimental - Mistura com dados de contagem

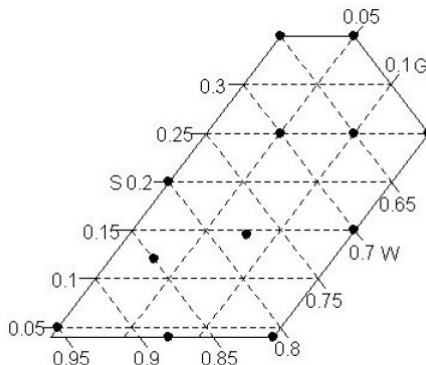


Figura: Desenho experimental de mistura com restrição

conjunto de dados - Mistura com dados de contagem

Proportions			Soap Type	Water Type	Response Number of Bubbles	Proportions			Soap Type	Water Type	Response Number of Bubbles
Water	Soap	Glycerin				Water	Soap	Glycerin			
0.6	0.35	0.05	-1	-1	119	0.6	0.35	0.05	1	-1	21
0.6	0.25	0.15	-1	-1	63	0.6	0.25	0.15	1	-1	13
0.65	0.35	0	-1	-1	39	0.65	0.35	0	1	-1	12
0.65	0.25	0.1	-1	-1	60	0.65	0.25	0.1	1	-1	29
0.75	0.25	0.05	-1	-1	64	0.75	0.25	0.05	1	-1	20
0.7	0.15	0.15	-1	-1	54	0.7	0.15	0.15	1	-1	16
0.775	0.145	0.08	-1	-1	44	0.775	0.145	0.08	1	-1	22
0.8	0.2	0	-1	-1	51	0.8	0.2	0	1	-1	19
0.81	0.04	0.15	-1	-1	12	0.81	0.04	0.15	1	-1	15
0.85	0.12	0.03	-1	-1	64	0.85	0.12	0.03	1	-1	8
0.88	0.04	0.08	-1	-1	10	0.88	0.04	0.08	1	-1	3
0.95	0.05	0	-1	-1	20	0.95	0.05	0	1	-1	15
0.6	0.35	0.05	-1	1	71	0.6	0.35	0.05	1	1	18
0.6	0.25	0.15	-1	1	68	0.6	0.25	0.15	1	1	17
0.65	0.35	0	-1	1	53	0.65	0.35	0	1	1	5
0.65	0.25	0.1	-1	1	60	0.65	0.25	0.1	1	1	13
0.75	0.25	0.05	-1	1	77	0.75	0.25	0.05	1	1	22
0.7	0.15	0.15	-1	1	46	0.7	0.15	0.15	1	1	17
0.775	0.145	0.08	-1	1	56	0.775	0.145	0.08	1	1	15
0.8	0.2	0	-1	1	68	0.8	0.2	0	1	1	11
0.81	0.04	0.15	-1	1	20	0.81	0.04	0.15	1	1	21
0.85	0.12	0.03	-1	1	40	0.85	0.12	0.03	1	1	17
0.88	0.04	0.08	-1	1	23	0.88	0.04	0.08	1	1	3
0.95	0.05	0	-1	1	14	0.95	0.05	0	1	1	15

Quantidade de bolhas em relação às variáveis de ruído

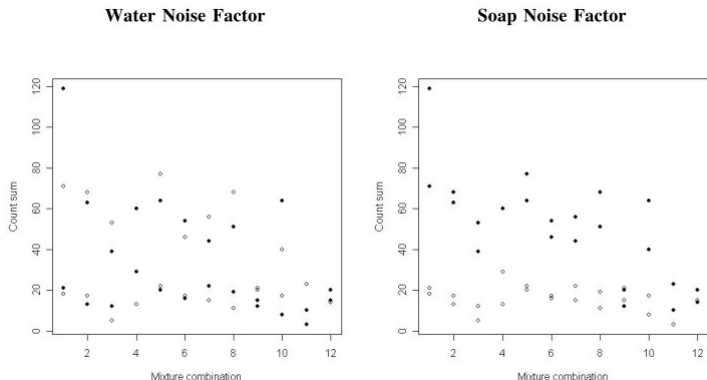


Figura: Figura à esquerda (tipo de água): Círculos sólidos = Spring, Círculos abertos = Tap. Figura à direita (tipo de sabão): Círculos sólidos = Joy, Círculos abertos = Ivory.

Modelo considerado

- $Y \sim \text{Poisson}(\mu)$
- $\log(\mu) = \eta = \mathbf{X}\boldsymbol{\beta}$
- Preditor linear inicial considerado

$$\begin{aligned} \eta = & \beta_1 S + \beta_2 W + \beta_3 G + \beta_4 SW + \beta_5 SG + \beta_6 WG + \\ & + (\beta_7 S + \beta_8 W + \beta_9 G + \beta_{10} SW + \beta_{11} SG + \beta_{12} WG) N_W + \\ & + (\beta_7 S + \beta_8 W + \beta_9 G + \beta_{10} SW + \beta_{11} SG + \beta_{12} WG) N_S + \\ & + (\beta_7 S + \beta_8 W + \beta_9 G + \beta_{10} SW + \beta_{11} SG + \beta_{12} WG) N_W N_S \end{aligned}$$

Modelo final obtido

- O modelo final obtido foi:
- $\log(\hat{\mu}) = -3.4S + 2.4W + 17.8G + 13.97SW - 5.3SG - 17.6WG + 2.3SN_S + -0.2WN_S + 47.3GN_S - 5.2SWN_S - 69.7SGN_S$
- O tipo de água, N_W , foi não significativo.
- Realizado um processo de otimização para maximização de $\log(\hat{\mu})$, considerada a restrição de mistura e as restrições nas variáveis de mistura, obteve-se, para cada um dos tipos de sabão os valores:
- Sabão Ivory: Mistura ótima (0.31, 0.69, 0) com máximo de bolhas = 4.3.
- Sabão Joy: Mistura ótima (0.31, 0.60, 0.9) com máximo de bolhas = 16.8.

Gráficos de contorno

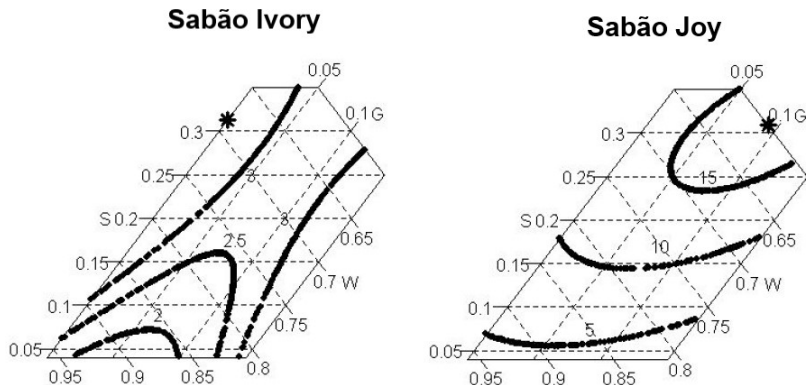


Figura: Figura à esquerda: Sabão Ivory. Figura à direita: Sabão Joy. Asterisco = máxima predição.

Projeto Robusto de Experimentos

- É uma metodologia para melhoria da qualidade em um produto ou processo.
- Tem como objetivos: 1) garantir que a média da resposta de saída esteja no nível ou destino desejado; 2) garantir que a variabilidade em torno desse valor-alvo seja a menor possível.
- Obtenha modelos para média e para a variância.
- Encontre fatores no modelo da variância a fim de minimizar a variabilidade.
- Encontre fatores no modelo da média a fim de ajustá-la a um valor alvo.



CNMAC 2019
XXX Congresso Nacional de Matemática Aplicada e Computacional
2019 - 2020

Projeto Robusto de Experimentos - O método delta

O método delta permite encontrar modelos para média e variância em problemas de mistura com variáveis de ruído.

Considerando aproximação de primeira ordem para série de Taylor, o método delta é obtido da seguinte forma. Sejam $Z_1, \dots, Z_r$ variáveis aleatórias com médias $\mu_1, \dots, \mu_r$ e matriz de variância e covariância Σ_z . Defina $\mathbf{Z}^t = (Z_1, \dots, Z_r)$ e $\boldsymbol{\mu}^t = (\mu_1, \dots, \mu_r)$. Para nosso problema de experimentos com mistura, também consideramos $\mathbf{x}$ como o vetor dos componentes de mistura. Suponha que exista uma função diferenciável $\eta(\mathbf{x}, \mathbf{Z}, \boldsymbol{\beta}, \boldsymbol{\gamma})$ para qual nós queremos uma estimativa da média e da variância. Considere $\eta'(\mathbf{x}, \boldsymbol{\mu}, \boldsymbol{\beta}, \boldsymbol{\gamma}) = \left. \frac{\partial}{\partial \mathbf{z}} \eta(\mathbf{x}, \mathbf{z}, \boldsymbol{\beta}, \boldsymbol{\gamma}) \right|_{\mathbf{z}=\boldsymbol{\mu}}$



Assim, a expansão de série de Taylor de primeira ordem da função η em torno de μ é

$$\eta(\mathbf{x}, \mathbf{z}, \boldsymbol{\beta}, \boldsymbol{\gamma}) = \eta(\mathbf{x}, \boldsymbol{\mu}, \boldsymbol{\beta}, \boldsymbol{\gamma}) + (\mathbf{z} - \boldsymbol{\mu})^t \boldsymbol{\eta}'(\mathbf{x}, \boldsymbol{\mu}, \boldsymbol{\beta}, \boldsymbol{\gamma}) + \mathcal{R},$$

Desta forma, como $\eta(\mathbf{x}, \mathbf{Z}, \boldsymbol{\beta}, \boldsymbol{\gamma})$ é uma variável aleatória, podemos obter

$$E [\eta(\mathbf{x}, \mathbf{Z}, \boldsymbol{\beta}, \boldsymbol{\gamma})] \approx \eta(\mathbf{x}, \boldsymbol{\mu}, \boldsymbol{\beta}, \boldsymbol{\gamma})$$

and

$$Var [\eta(\mathbf{x}, \mathbf{Z}, \boldsymbol{\beta}, \boldsymbol{\gamma})] \approx [\boldsymbol{\eta}'(\mathbf{x}, \boldsymbol{\mu}, \boldsymbol{\beta}, \boldsymbol{\gamma})]^t \boldsymbol{\Sigma}_{\mathbf{z}} \boldsymbol{\eta}'(\mathbf{x}, \boldsymbol{\mu}, \boldsymbol{\beta}, \boldsymbol{\gamma}).$$

Exemplo - Problema da mistura de pão (Naes et al., 1998)

Problema clássico de mistura de pão que consiste de um experimento com três ingredientes de mistura e duas variáveis de ruído e tem como objetivo investigar o valor da qualidade final da farinha, composta por diferentes misturas de farinha de trigo, para a produção de pão. Os tipos de farinha de trigo consideradas foram duas farinhas norueguesas Tjalve (x_1) e Folke (x_2) e uma americana Hard Red Spring (x_3). As variáveis de ruído foram: tempo de mistura (z_1) e tempo de descanso da massa (z_2). A variável resposta considerada foi o volume do pão após assado, com valor alvo de 530 ml.

Além da restrição de mistura, as restrições nas variáveis são: $0.25 \leq x_1 \leq 1.0$; $0 \leq x_2 \leq 0.75$ and $0 \leq x_3 \leq 0.75$. Para as variáveis de ruído, foram consideradas três situações para o tempo de mistura: 5, 15 e 25 minutos e para as situações de descanso de massa, os tempos considerados foram: 35, 47.5, and 60 minutos. Um desenho fatorial completo 3^2 foi usado para as variáveis de ruído e 10 corridas experimentais correspondendo a um desenho lattice-simplex foi replicado em cada uma das nove combinações dos tempos de mistura e de descanso da massa, de modo que o desenho experimental completo envolveu 90 corridas experimentais. As variáveis de ruído foram codificadas em -1 , 0 e 1 .



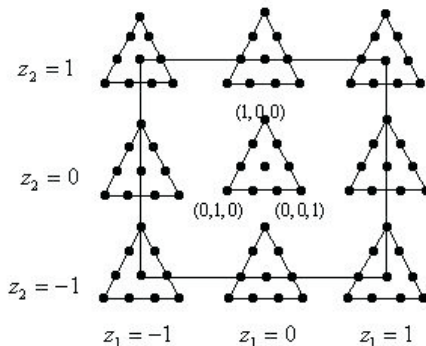


Figura: Fatorial completo 3^2 em duas variáveis de processo cruzado com o desenho simplex-centróide de terceira ordem em três componentes de mistura

Tabela: Volume de pão para 10 tipos de farinha e 9 combinações de variáveis de ruído

n ^o	Design factors			Noise Factors									
				z ₁	-1	0	1	-1	0	1	-1	0	1
	x ₁	x ₂	x ₃	z ₂	-1	-1	-1	0	0	0	1	1	1
1	0.25	0.75	0.00		378.89	396.67	392.22	445.56	452.22	487.78	457.22	500.56	472.78
2	0.50	0.50	0.00		388.89	423.33	416.11	460.00	488.89	475.78	472.78	478.00	506.11
3	0.75	0.25	0.00		426.11	483.33	389.44	474.44	514.44	462.78	506.67	591.67	522.22
4	1.00	0.00	0.00		386.11	459.11	423.33	458.33	506.11	514.44	545.56	522.22	551.11
5	0.25	0.50	0.25		417.78	437.22	444.56	484.44	490.00	495.00	497.78	531.11	577.78
6	0.50	0.25	0.25		389.44	447.22	415.00	490.89	528.89	507.78	517.78	567.22	538.33
7	0.75	0.00	0.25		448.33	459.44	455.56	436.00	535.00	552.22	507.44	578.89	590.00
8	0.25	0.25	0.50		413.89	485.56	462.22	483.89	529.44	540.00	565.00	598.89	580.56
9	0.50	0.00	0.50		415.56	514.44	437.78	493.89	583.33	578.89	524.44	694.44	640.00
10	0.25	0.00	0.75		432.78	498.33	517.22	474.44	568.33	579.44	541.11	638.89	638.89

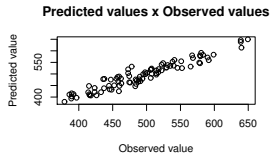
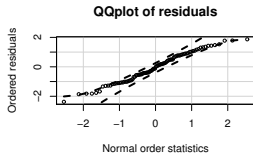
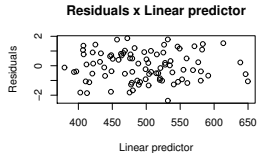
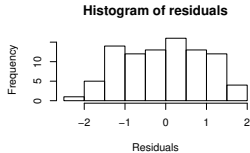
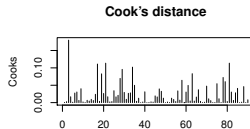
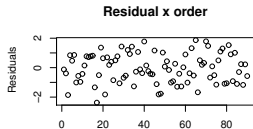
Modelo inicial

$$\begin{aligned}
 \eta(\mathbf{x}, \mathbf{z}, \boldsymbol{\beta}, \boldsymbol{\gamma}) = & \beta_1^0 x_1 + \beta_2^0 x_2 + \beta_3^0 x_3 + \beta_{12}^0 x_1 x_2 + \beta_{13}^0 x_1 x_3 + \gamma_{12}^0 x_1 x_2 (x_1 - x_2) + \\
 & \gamma_{13}^0 x_1 x_3 (x_1 - x_3) + \{\beta_1^1 x_1 + \beta_2^1 x_2 + \beta_3^1 x_3 + \beta_{12}^1 x_1 x_2 + \beta_{13}^1 x_1 x_3 + \\
 & \gamma_{12}^1 x_1 x_2 (x_1 - x_2) + \gamma_{13}^1 x_1 x_3 (x_1 - x_3)\} z_1 + \{\beta_1^2 x_1 + \beta_2^2 x_2 + \\
 & \beta_3^2 x_3 + \beta_{12}^2 x_1 x_2 + \beta_{13}^2 x_1 x_3 + \gamma_{12}^2 x_1 x_2 (x_1 - x_2) + \\
 & \gamma_{13}^2 x_1 x_3 (x_1 - x_3)\} z_2 + \{\beta_1^{11} x_1 + \beta_2^{11} x_2 + \beta_3^{11} x_3 + \beta_{12}^{11} x_1 x_2 + \\
 & \beta_{13}^{11} x_1 x_3 + \gamma_{12}^{11} x_1 x_2 (x_1 - x_2) + \gamma_{13}^{11} x_1 x_3 (x_1 - x_3)\} z_1^2
 \end{aligned} \quad (1)$$

Análise estatística

Table: Coeficientes de regressão para o modelo da média via método de mínimos quadrados ordinários

Terms	Estimate	Std. Error	t value	p-value
x_1	484.624	6.363	76.161	0.0000
x_2	474.875	13.369	35.521	0.0000
x_3	436.381	64.837	6.730	0.0000
$x_1 x_3$	468.313	164.234	2.851	0.0057
$x_1 x_2 (x_1 - x_2)$	375.341	94.623	3.397	0.0002
$x_1 x_3 (x_1 - x_3)$	-403.031	199.679	-2.018	0.0473
$x_1 z_1$	16.768	5.452	3.076	0.0029
$x_3 z_1$	51.876	8.406	6.171	0.0000
$x_1 x_2 (x_1 - x_2) z_1$	-144.553	60.706	-2.381	0.0199
$x_1 z_2$	54.933	6.703	8.195	0.0000
$x_2 z_2$	42.504	8.470	5.018	0.0000
$x_1 x_3 z_2$	188.762	25.167	7.500	0.0000
$x_1 x_3 (x_1 - x_3) z_2$	-202.822	61.681	-3.288	0.0016
$x_2 z_1^2$	-52.644	14.972	-3.516	0.0008
$x_3 z_1^2$	164.077	79.249	2.070	0.0420
$x_1 x_3 z_1^2$	-600.046	199.173	-3.013	0.0036
$x_1 x_2 (x_1 - x_2) z_1^2$	-440.721	109.730	-4.016	0.0001
$x_1 x_3 (x_1 - x_3) z_1^2$	525.480	244.486	2.149	0.0349



$$\hat{\eta}(\mathbf{x}, \mathbf{z}, \hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\gamma}}) = c_1 + c_2 z_1 + c_3 z_2 + c_4 z_1^2,$$

$$c_1 = 484.62x_1 + 474.88x_2 + 436.38x_3 + 468.31x_1x_3 + 375.34x_1x_2(x_1 - x_2) - 403.03x_1x_3(x_1 - x_3),$$

$$c_2 = 16.77x_1 + 51.88x_3 - 144.55x_1x_2(x_1 - x_2),$$

$$c_3 = 54.93x_1 + 42.50x_2 + 188.76x_1x_3 - 202.82x_1x_3(x_1 - x_3)$$

$$c_4 = -52.64x_2 + 164.08x_3 - 600.05x_1x_3 - 440.72x_1x_2(x_1 - x_2) + 525.48x_1x_3(x_1 - x_3).$$

Note que, c_1 , c_2 , c_3 and c_4 são constantes em relação a $\mathbf{z}$.

Considere $E(Z_i) = \mu_i$ e $Var(Z_i) = \sigma_i^2$, para $i = 1, 2$. obtemos

$$E(Y) \approx c_1 + c_2\mu_1 + c_3\mu_2 + c_4\mu_1^2,$$

$$Var(Y) \approx (c_2 + 2c_4\mu_1)^2\sigma_1^2 + c_3^2\sigma_2^2 + \sigma^2.$$

Para nosso modelo $\hat{\sigma}^2 = 58.36$.

Seguindo as idéias de Taguchi para melhoramento da qualidade, desde que o valor alvo para a variável resposta é (volume de pão após assado) é 530 ml, nós temos que resolver o seguinte problema de minimização.

$$\text{Min } \text{Var}(Y)$$

$$\text{Subject to } \left\{ \begin{array}{l} E(Y) = 530.0 \\ x_1 + x_2 + x_3 = 1.0 \\ 0.25 \leq x_1 \leq 1.0 \\ 0.0 \leq x_2 \leq 0.75 \\ 0.0 \leq x_3 \leq 0.75 \end{array} \right.$$



Tabela: Valores ótimos de mistura com seus respectivos valores de $Var(Y)$ e coeficiente de variação para alguns cenários envolvendo as variáveis de ruído.

Uncoded noise variables		Coded noise variables		Optimum (x_1, x_2, x_3)	$Var(Y)$ estimated	Coef. of variation
Mixing Time [†] (μ_m, σ_m^2)	Proofing Time [‡] (μ_p, σ_p^2)	Z_1 (μ_1, σ_1^2)	Z_2 (μ_2, σ_2^2)			
(10.0, 6.25)	(47.50, 9.766)	(-0.50, 0.0625)	(0.00, 0.0625)	(0.25,0.10,0.65)	604.91	0.046
(12.5, 6.25)	(44.375, 9.766)	(-0.25, 0.0625)	(-0.25, 0.0625)	(0.25,0.10,0.65)	480.79	0.041
(12.5, 6.25)	(53.75, 9.766)	(-0.25, 0.0625)	(0.50, 0.0625)	(0.27,0.00,0.73)	643.55	0.048
(12.5, 25.0)	(44.375, 39.063)	(-0.25, 0.25)	(-0.25, 0.25)	(0.25,0.10,0.65)	1748.08	0.078
(12.5, 56.25)	(44.375, 87.891)	(-0.25, 0.5625)	(-0.25, 0.5625)	(0.25,0.10,0.65)	3860.23	0.117
(15.0, 6.25)	(53.75, 9.766)	(0.00, 0.0625)	(0.50, 0.0625)	(0.61,0.36,0.03)	216.34	0.028
(15.0, 6.25)	(41.25, 9.766)	(0.00, 0.0625)	(-0.50, 0.0625)	(0.44,0.10,0.46)	396.81	0.038
(15.0, 6.25)	(47.50, 9.766)	(0.00, 0.0625)	(0.00, 0.0625)	(0.62,0.17,0.21)	260.33	0.030
(15.0, 25.0)	(47.50, 39.063)	(0.00, 0.250)	(0.00, 0.250)	(0.62,0.17,0.21)	866.23	0.056
(15.0, 56.25)	(47.50, 87.891)	(0.00, 0.5625)	(0.00, 0.5625)	(0.62,0.17,0.21)	1876.07	0.082
(20.0, 6.25)	(47.50, 9.766)	(0.50, 0.0625)	(0.00, 0.0625)	(0.25,0.30,0.45)	234.32	0.029
(20.0, 6.25)	(53.75, 9.766)	(0.50, 0.0625)	(0.50, 0.0625)	(0.26,0.43,0.31)	211.24	0.027
(20.0, 25.0)	(53.75, 39.063)	(0.50, 0.250)	(0.50, 0.250)	(0.26,0.43,0.31)	669.87	0.049
(20.0, 56.25)	(53.75, 87.891)	(0.50, 0.5625)	(0.50, 0.5625)	(0.26,0.43,0.31)	1434.25	0.071

[†]Mixing time is a random variable with mean μ_m and variance σ_m^2 .

[‡]Proofing time is a random variable with mean μ_p and variance σ_p^2 .

Modelagem conjunta da média e dispersão - MCMD

- Consiste em encontrar modelos conjuntos para média e dispersão.
- Usa a quasi-verossimilhança estendida como um critério de ajustamento dos modelos.
- A quasi -verossimilhança estendida é dada por:

$$Q^+ = \frac{-1}{2} \sum_{i=1}^n \left(\frac{d_i^*}{\phi_i} + \ln\{2\pi\phi_i V(y_i)\} \right)$$

Resumo da modelagem conjunta da média e dispersão

Tabela: Summary of the JMMD

Component	Mean model	Dispersion model [†]
Response variable	y_i	d_i^*
Mean	μ_i	ϕ_i
Variance	$\phi_i V(\mu_i)$	$2\phi_i^2$
Link function	$\eta_i = \varphi(\mu_i)$	$\xi_i = \ln(\phi_i)$
Linear predictor	$\eta_i = \mathbf{f}^t(\mathbf{t}_i)\boldsymbol{\beta}$	$\xi_i = \mathbf{g}^t(\mathbf{u}_i)\boldsymbol{\gamma}$
Deviance component	$d_i = 2 \int_{\mu_i}^{y_i} \frac{y_i - t}{V(t)} dt$	$d_{d_i} = 2 \left\{ -\ln \left(\frac{d_i^*}{\phi_i} \right) + \frac{(d_i^* - \phi_i)}{\phi_i} \right\}$
Prior weight	$\frac{1}{\phi_i}$	$(1 - h_i)/2$

[†]For the dispersion model we are assuming a gamma model with logarithmic link function

$$h_i \text{ é o elemento da diagonal de } \mathbf{H} = \mathbf{W}^{\frac{1}{2}} \mathbf{X} (\mathbf{X}^t \mathbf{W} \mathbf{X})^{-1} \mathbf{X} \mathbf{W}^{\frac{1}{2}} \text{ e } d_i^* = \frac{d_i}{1-h_i}.$$

Pseudo algoritmo para a MCMD

1. **Modelo da média** - Dados uma distribuição de probabilidade para a variável resposta, o preditor linear e uma função de ligação. Calcule $\beta^t = (\beta_1, \dots, \beta_p)$ usando o algoritmo para ajustamento dos MLG.
2. Com o valor encontrado para β no passo 1, calcule $d^{*t} = (d_1^*, \dots, d_n^*)$.
3. **Modelo para dispersão** - Considerando d^* como resposta e assumindo um modelo Gamma com ligação logarítmica, calcule $\gamma^t = (\gamma_1, \dots, \gamma_q)$ usando o algoritmo de ajustamento dos MLG.
4. Com o valor γ obtido no passo 3, calcule $\phi^t = (\phi_1, \dots, \phi_n)$.
5. Com o valor d^* encontrado no passo 2 e o valor de ϕ obtido no passo 4, calcule $Q^+(\mu, \phi; y)$ e use esta medida para verificar a convergência do modelo conjunto. Se a convergência é obtida, pare. Os correntes β e γ são os parâmetros finais do modelo. Caso contrário, como o valor de ϕ , atualizado no passo 4, calcule os novos pesos $w_1, \dots, w_n$ para o modelo da média e vá para o passo 1.

Projetos de pesquisa

- Seleção de variáveis para modelos conjuntos da média e dispersão (artigo enviado para publicação)
- Estimação e seleção de modelos para experimentos de mistura (artigo em fase final de preparação)
- Planejamento ótimo de experimentos para modelagem conjunta da média e dispersão (Pinto e Ponce de Leon (2014)).
- Extensão do planejamento ótimo de experimentos na MCMD para experimentos com mistura e programa em R (ainda não realizado).



References

- Atkinson, A. C., Donev, A. N. , Tobias, R. D. (2007). *Optimum Experimental Designs, with SAS*, Oxford University Press, New York.
- Cornell, J.A. (2002). *Experiments with mixtures: Designs, Models, and the Analysis of Mixture Data.*, 3rd ed.. Wiley, New York.
- Granato, D., Calado, V., Pinto, E.R. (2019). Optimization of food processes using mixture experiments - some applications. In Sevda an Singh (Eds.), *Mathematical and Statistical Applications in Food Engineering*, Chapter 3. CRC Press (in press).
- Lawson, J., Willden, C. (2016). Mixture Experiments in R Using mixexp. *Journal of Statistical Software*, v. 72, pp. 1-20.

References

- McCullagh, P., Nelder, J.A. (1989). *Generalized Linear Models. 2nd ed.*, Chapman and Hall, London.
- Myers, R. H., Montgomery, D. C. (2002). *Response Surface Methodology: Process and Product Optimization Using Designed Experiments. 2nd edition*, John Wiley & Sons, New York.
- Naes, T., Faergestad, E. M., Cornell, J. A. (1998). A comparison of methods for analyzing data from a three component mixture experiment in the presence of variation created by two process variables, *Chenometrics and Intelligence Laboratory Systems*, v. 41, 221-235.
- Pinto, E. R., Pereira, L. A. (2010). Função no R para obtenção do desenho D-ótimo em modelos de misturas com restrição. 19^o SINAPE. *Associação Brasileira de Estatística*.

References

- Pinto, E. R., Ponce de Leon, A. (2006). *Planejamento Ótimo de Experimentos*. 17^o SINAPE - Associação Brasileira de Estatística, São Paulo.
- Pinto, E. R., Ponce de Leon, A. (2006). Modelagem conjunta da média e dispersão de Nelder e Lee como alternativa aos métodos de Taguchi. *Pesquisa Operacional*, v.26, n. 2, 203-224.
- Pinto, E. R., Ponce de Leon, A. (2014). On Bayesian D-optimal design criteria and the General Equivalence Theorem in joint generalized linear models for the mean and dispersion. *Brazilian Journal of Probability and Statistics*, v. 28, n. 4, 483-491.
- Steiner, S. H., Hamada, M., White, B. J. G., Kutsyy, V., Mosesova, S., Salloum, G. (2007). A bubble mixture experiment project for use in an advanced design of experiments class. *Journal of Statistics Education*, v. 15, n. 1, 1-20.

Obrigado!