

EXPERIMENTOS COM MISTURA PARA OTIMIZAÇÃO DE PROCESSOS: UMA APLICAÇÃO COM RESPOSTAS NÃO NORMAIS

Antonio Fernando de Castro Vieira *

Departamento de Engenharia Industrial
Pontifícia Univ. Católica do Rio de Janeiro (PUC-Rio)
Rio de Janeiro – RJ
afcvc@ind.puc-rio.br

Luiz Henrique Abreu Dal Bello

Instituto de Pesquisa e Desenvolvimento (IPD)
Centro Tecnológico do Exército (CTEx)
Rio de Janeiro – RJ
henrique@dalbello.com.br

* *Corresponding author* / autor para quem as correspondências devem ser encaminhadas

Recebido em 10/2005; aceito em 04/2006 após 1 revisão
Received October 2005; accepted April 2006 after one revision

Resumo

Neste artigo apresenta-se uma metodologia para planejamento de experimentos com mistura, ilustrada com o caso real do estudo de um misto químico que constitui um subsistema de um mecanismo de retardo para acionamento de um motor foguete. O objetivo do estudo é saber qual a proporção dos componentes da mistura que propicia um tempo de queima especificado em projeto. Para isto foi ajustado um modelo de regressão linear com respostas normais, o qual foi considerado inadequado, uma vez que houve indicação de que a variância da resposta não é constante. Recorreu-se então a modelos ajustados pelo método da Quase-Verossimilhança. Com o modelo desenvolvido, foi possível determinar a proporção dos componentes que otimiza o processo. Para a otimização do processo, além de um tempo de queima específico, foi possível considerar a minimização da variância da previsão deste tempo.

Palavras-chave: experimentos com mistura; quase-verossimilhança; otimização.

Abstract

In this article a methodology for the design of mixture experiments is presented, illustrated with a real case study of a chemical compound, that constitutes a subsystem of a delay mechanism of a rocket motor. The aim of the study is to find the proportion of each mixture component in order to attend the specified burn time. For this, a linear regression model with normal responses was adjusted, and next considered inadequate, because there was an indication that the response variance is not constant. Therefore, the models were adjusted using the Quasi-Likelihood method. With the developed model, it was possible to find the proportion of each component that optimizes the process. For the process optimization, beside a specified burn time, it was possible to considerer the minimization of the prediction variance of this time.

Keywords: mixture experiments; quasi-likelihood; optimization.

1. Introdução

Muitos produtos são feitos com a mistura de vários componentes. Tintas, alimentos, produtos químicos e farmacêuticos são alguns exemplos. Para tais produtos, o interesse é determinar qual é a proporção dos componentes que conduz a um resultado desejado em termos de uma variável que caracteriza a qualidade do produto. Quando não se sabe de antemão qual é a proporção ideal de cada componente, devemos realizar experimentos. Nesses experimentos são arbitradas várias combinações de proporções dos componentes e então são observados os valores correspondentes da característica de qualidade. Estes valores são denominados respostas do experimento.

Em experimentos envolvendo mistura, a soma das proporções dos componentes é sempre igual a 1. Ao modificar a formulação no sentido de alterar as propriedades de uma determinada mistura em estudo, as novas proporções devem continuar obedecendo a essa restrição. Assim sendo, no planejamento dos experimentos com mistura devem ser adotados projetos que considerem esta restrição.

Com base nos resultados do experimento, procura-se estabelecer a relação entre a variável resposta e as variáveis que representam as proporções de cada componente. Para tanto são utilizados modelos polinomiais. Na Seção 2.1 são apresentados os polinômios recomendados para experimentos com mistura.

Neste artigo é apresentado um estudo de um misto químico que constitui um subsistema de um mecanismo de retardo para acionamento de um motor foguete (Dal Bello, 2005). O objetivo do estudo é saber qual é a proporção de cada um dos componentes que propicia um tempo de queima (resposta) de 11 segundos e minimiza a variância da resposta.

No caso do misto de retardo, além da restrição de que a soma das proporções dos componentes é sempre igual a 1, são consideradas outras restrições nas proporções dos três componentes. Por motivos técnicos, cada componente possui um limite superior e inferior em sua proporção, o que torna mais complexa a escolha dos pontos experimentais no espaço dos componentes da mistura (Seção 2.2). Nesse caso, para a escolha dos pontos experimentais, é recomendável a utilização dos projetos de experimentos gerados segundo algum critério de otimização (Seção 2.3).

Em experimentos com mistura tem-se usado sistematicamente o modelo linear com distribuição da resposta normal. Cabe registrar que, em todas as referências analisadas, os autores do presente artigo só encontraram aplicações que consideram a resposta com distribuição normal. Para o caso do misto, como será visto na Seção 3.1, o modelo com distribuição da resposta normal mostrou-se inadequado, uma vez que há indicação de que a variância da resposta não é constante. Para a estimativa e inferência dos parâmetros foi então ajustado um modelo que considera o método de quase-verossimilhança (Seção 3.3), que é uma extensão dos modelos lineares generalizados (Seção 3.2). Na Seção 4 são apresentadas as conclusões.

Cornell (1990) é a principal referência sobre a teoria de experimentos com mistura. Nela pode-se encontrar uma exposição abrangente e detalhada. Myers & Montgomery (2002) dedicam dois capítulos a experimentos com mistura, constituindo uma boa introdução ao assunto. Kowalski *et al.* (2002) e Prescott (2004) analisam a modelagem de experimentos com mistura para o caso em que, além dos componentes da mistura, há outros fatores que afetam a característica de qualidade produto (caso mistura-processo). Para o caso em que as

estimativas dos coeficientes são altamente correlacionadas, Prescott *et al.* (2002) propõem um modelo quadrático alternativo ao modelo quadrático geralmente utilizado em experimentos com mistura (ver seção 2.1). Goldfarb *et al.* (2003) e Goldfarb *et al.* (2004b) consideram o caso em que, além dos componentes da mistura, há fatores que não podem ser controlados no processo produtivo (variáveis de ruído), embora possam ser controladas em experimentos de laboratório. Goldfarb *et al.* (2004a) propõem um método gráfico (gráfico de dispersão de variância) para avaliar projetos de experimentos utilizados no caso mistura-processo.

2. Modelos para Componentes de Mistura

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

$$\sum_{i=1}^q x_i = 1 \text{ e } x_i \geq 0, \quad i = 1, \dots, q \quad (1)$$

As restrições (1) são mostradas graficamente na Figura 1, para o caso de dois e três componentes. A região factível da mistura de dois componentes é representada por um segmento de reta e para o caso de três componentes é representada por um triângulo.

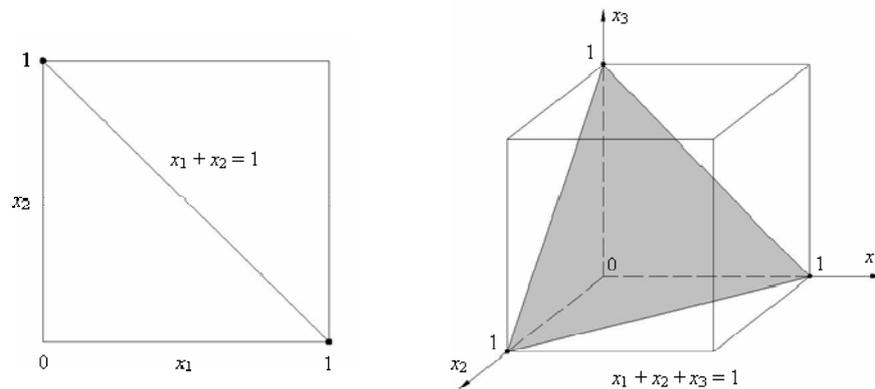


Figura 1 – Espaço fatorial restrito para misturas com 2 e 3 componentes.

Por uma questão prática, no caso de experimentos com misturas de três componentes, a região experimental restrita pode ser representada com a utilização de um sistema de coordenadas trilinear, como mostrado na Figura 2.

Cada lado do triângulo corresponde a uma mistura binária e os vértices dos triângulos correspondem às formulações de componentes puros. No interior do triângulo, estão situadas as possíveis misturas ternárias. Neste caso, são necessárias apenas duas dimensões para representar graficamente o experimento. Como cada componente é representado por um vértice, uma figura geométrica com três vértices e duas dimensões, ou seja, um triângulo equilátero, representa o espaço fatorial restrito de uma mistura ternária.

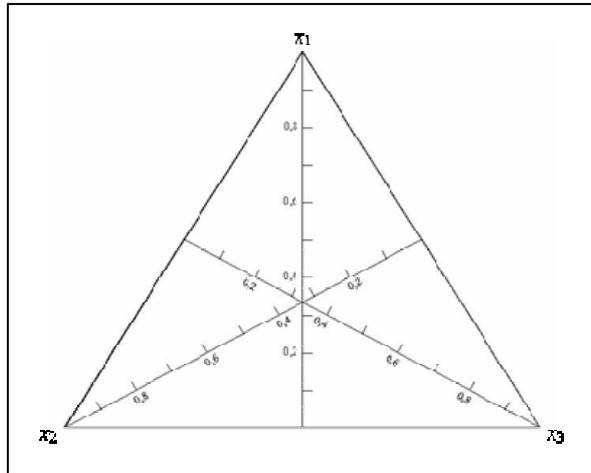


Figura 2 – Sistema de coordenadas trilinear.

2.1 Polinômios canônicos

Os modelos recomendados para experimentos com mistura são os polinômios canônicos de Scheffé (Cornell, 1990). Considerando que a resposta é dada por y , em geral, estes polinômios tomam as seguintes formas:

Modelo Linear:

$$E[y] = \sum_{i=1}^q \beta_i x_i$$

Modelo Quadrático:

$$E[y] = \sum_{i=1}^q \beta_i x_i + \sum_{i < j} \beta_{ij} x_i x_j$$

Modelo Cúbico Especial:

$$E[y] = \sum_{i=1}^q \beta_i x_i + \sum_{i < j} \beta_{ij} x_i x_j + \sum_{i < j < k} \beta_{ijk} x_i x_j x_k$$

Modelo Cúbico Completo:

$$E[y] = \sum_{i=1}^q \beta_i x_i + \sum_{i < j} \beta_{ij} x_i x_j + \sum_{i < j < k} \beta_{ijk} x_i x_j x_k + \sum_{i < j} \beta_{i-j} x_i x_j (x_i - x_j).$$

2.2 Restrições nos componentes da mistura

Em certos experimentos com mistura, pode haver a necessidade de restringir a proporção de um ou mais componentes, que, por motivos técnicos ou práticos, não podem varrer todas as proporções possíveis. Com isso, o novo espaço experimental passa a ser uma sub-região da

região das proporções matematicamente possíveis. Essas restrições dos componentes, que são muito comuns nos casos industriais, podem ser superiores, inferiores ou uma combinação das duas.

Na prática, há casos em que muitos dos componentes de uma mistura possuem simultaneamente limites superiores e inferiores em suas composições. De uma forma geral, essas restrições podem ser representadas da seguinte forma:

$$0 \leq L_i \leq x_i \leq U_i \leq 1, \quad i = 1, 2, \dots, q$$

Com essas restrições, a região factível pode ter forma bem diferente da região factível sem restrições nas proporções dos componentes da mistura. Nesses casos, não é viável recorrer mais aos planejamentos experimentais tradicionais apresentados na literatura, e, muitas vezes, recorre-se a algoritmos computacionais para a escolha da localização dos pontos experimentais, havendo a necessidade de considerar alguns aspectos técnicos, como o número mínimo de pontos necessários para a estimativa dos parâmetros do modelo. Como se sabe, o número de pontos experimentais deve ser, no mínimo, igual ao número de parâmetros do modelo considerado, apresentados na Tabela 1.

Tabela 1 – Número de termos nos polinômios canônicos (Cornell, 1990).

<i>Número de Componentes</i>	<i>Linear</i>	<i>Quadrático</i>	<i>Cúbico Especial</i>	<i>Cúbico Completo</i>
q	q	$q(q+1)/2$	$q(q^2+5)/6$	$q(q+1)(q+2)/6$

2.3 D-Otimização

Pode-se representar o modelo geral para misturas sob a forma matricial.

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad \text{ou} \quad E[\mathbf{y}] = \mathbf{X}\boldsymbol{\beta}$$

Para n observações tem-se que \mathbf{y} e $\boldsymbol{\varepsilon}$ são vetores de dimensão n , \mathbf{X} é uma matriz $n \times p$ e $\boldsymbol{\beta}$ é um vetor de dimensão p , conforme apresentado a seguir.

$$\begin{matrix} \mathbf{y} & & \mathbf{X} & & \boldsymbol{\beta} & & \boldsymbol{\varepsilon} \\ \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_n \end{bmatrix} & = & \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ x_{31} & x_{32} & \cdots & x_{3p} \\ \vdots & \vdots & & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix} & \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \vdots \\ \beta_p \end{bmatrix} & + & \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \vdots \\ \varepsilon_n \end{bmatrix} \end{matrix}$$

Note que a matriz \mathbf{X} não possui a coluna identidade relativa ao termo independente do modelo de regressão linear. Isto ocorre porque os polinômios canônicos em experimentos com misturas não possuem esse termo.

No modelo linear clássico, $\boldsymbol{\varepsilon}$ segue a distribuição normal multivariada, ou seja, $\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \mathbf{I}\sigma^2)$.

A estimativa dos parâmetros por mínimos quadrados é $\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\mathbf{y}$ com matriz de variância-covariância $\text{var}(\hat{\boldsymbol{\beta}}) = \sigma^2 (\mathbf{X}'\mathbf{X})^{-1}$. A predição da resposta em algum ponto da composição \mathbf{x} é denotada por $\hat{y}(\mathbf{x})$ e sua variância pode ser obtida da seguinte forma:

$$\text{var}[\hat{y}(\mathbf{x})] = \sigma^2 \mathbf{x}' (\mathbf{X}'\mathbf{X})^{-1} \mathbf{x} \quad (2)$$

onde no lado direito da Equação 2, o vetor \mathbf{x} de dimensão p contém os valores das proporções dos q componentes e os $p - q$ termos do produto cruzado no ponto da composição.

Cornell (1990) descreve quatro critérios de otimização alfabética (A-otimização, D-otimização, G-otimização e V-otimização) para a seleção dos pontos candidatos. Estes critérios consideram a minimização de várias funções de $(\mathbf{X}'\mathbf{X})^{-1}$.

Devido à sua popularidade e por ser o critério adotado pelo *software Design-Expert*, desenvolvido e distribuído pela empresa *Stat-Ease*, foi utilizado no presente trabalho o critério D-otimização. Myers & Montgomery (2002) definem o critério D-otimização utilizando a matriz de momento $\mathbf{M} = \mathbf{X}'\mathbf{X}/n$, onde n é o número de observações do experimento. Segundo estes autores, um projeto de experimento D-ótimo é aquele que faz com que o determinante da matriz de momento, $|\mathbf{M}|$, seja maximizado. Eles demonstram que o determinante da matriz de momento possui a seguinte propriedade

$$|\mathbf{M}| = \frac{|\mathbf{X}'\mathbf{X}|}{n^p} \quad (3)$$

sendo p o número de parâmetros do modelo considerado.

Com isso, e supondo que os erros são normalmente distribuídos, independentes e com variância constante, o determinante $|\mathbf{X}'\mathbf{X}|$ é inversamente proporcional ao quadrado do volume da região de confiança sobre os coeficientes regressores. Quando $|\mathbf{X}'\mathbf{X}|$ é pequeno, significa que o inverso de $|\mathbf{X}'\mathbf{X}|$ é grande, sendo, então, o volume da região de confiança grande e, portanto, a estimativa de $\boldsymbol{\beta}$ não é considerada boa (Myers & Montgomery, 2002, Apêndice 7).

Portanto, o projeto de experimento D-ótimo é aquele que minimiza o volume do elipsóide de confiança sobre $\boldsymbol{\beta}$, o que é conseguido maximizando o determinante $|\mathbf{X}'\mathbf{X}|$. Observando-se a Equação (3), pode-se concluir que maximizar o determinante $|\mathbf{X}'\mathbf{X}|$ é equivalente a maximizar o determinante da matriz de momento, $|\mathbf{M}|$.

3. Experimento do Misto de Retardo

O misto químico em estudo constitui um subsistema de um mecanismo de retardo produzido para acionamento de um motor foguete, o qual encontra-se em fase inicial de pesquisa e desenvolvimento. O objetivo do experimento é encontrar a formulação ideal desse misto químico de forma que o seu tempo de queima (resposta) esteja de acordo com a especificação de projeto, ou seja, um tempo de queima de 11 segundos com tolerância de 1s.

O misto é composto por uma mistura de Zarfesil (x_1), Vidro Moído (x_2) e Nitrocelulose (x_3). Dessa forma, as restrições nas proporções dos três componentes são as seguintes:

$$x_1 + x_2 + x_3 = 1 \quad 0,79 \leq x_1 \leq 0,87 \quad 0,08 \leq x_2 \leq 0,16 \quad 0,05 \leq x_3 \leq 0,07$$

Com essas restrições, foi utilizado o *software Design-Expert* para a escolha dos pontos candidatos e seleção dos pontos experimentais segundo o critério D-otimização. Considerando a sugestão de Myers & Montgomery (2002) para a seleção dos pontos candidatos na região experimental resultante e sabendo que o modelo adotado inicialmente é o quadrático, o software sugere um total de 10 pontos, dos quais 6 são necessários para o ajuste do modelo quadrático e 4 pontos adicionais são para testar a falta de ajuste do modelo. Ademais, 4 dos 10 pontos são replicados para estimar o desvio-padrão da resposta, perfazendo um total de 14 observações. Com isso, o *Design-Expert* gera um projeto de experimento D-Ótimo conforme a Tabela 2. Na Figura 3 tem-se a representação gráfica do projeto de experimento.

Tabela 2 – Experimento do misto de retardo D-Ótimo.

	x_1	x_2	x_3	Tempo (s)
1	0,79	0,16	0,05	19,70
2	0,79	0,16	0,05	19,14
3	0,83	0,12	0,05	4,72
4	0,87	0,08	0,05	1,85
5	0,87	0,08	0,05	1,59
6	0,8075	0,1375	0,055	9,53
7	0,8475	0,0975	0,055	1,90
8	0,79	0,15	0,06	12,03
9	0,86	0,08	0,06	1,04
10	0,79	0,14	0,07	0,77
11	0,79	0,14	0,07	0,89
12	0,82	0,11	0,07	0,23
13	0,85	0,08	0,07	0,25
14	0,85	0,08	0,07	0,51

Considerando a estatística de teste $t_0 = \hat{\beta}_j / \sqrt{\hat{\sigma}^2 C_{jj}}$, onde $\hat{\sigma}^2 = \sum_{i=1}^n (y_i - \hat{y}_i) / (n - p)$ e C_{jj} é o elemento da diagonal de $(\mathbf{X}'\mathbf{X})^{-1}$ correspondente a $\hat{\beta}_j$, e aplicando o teste *t-student* para 95% foi ajustado o seguinte modelo:

$$\hat{y} = 318,85x_1 + 8721,33x_2 - 256135,53x_3 - 12494,31x_1x_2 + 421048,23x_1x_3 + 153766,65x_2x_3 - 174591,55x_1x_3(x_1 - x_3)$$

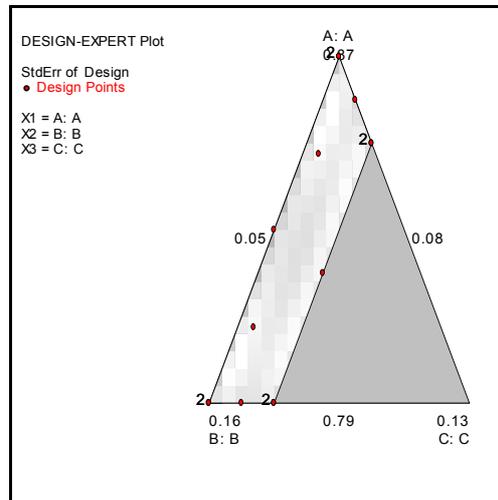


Figura 3 – Experimento do misto de retardo D-Ótimo.

3.1 Verificação da adequação do modelo

Uma vez ajustado o modelo de regressão, é necessário verificar se este é adequado para descrever os dados. Para isso, deve-se verificar se as suposições feitas, ou seja, independência, normalidade, valor esperado do erro igual a zero e variância constante, não foram violadas. Para isto, procede-se à análise dos resíduos. O vetor dos valores ajustados é dado por $\hat{\mathbf{y}} = \mathbf{X}\hat{\boldsymbol{\beta}}$, e o vetor dos resíduos é $\mathbf{e} = \mathbf{y} - \hat{\mathbf{y}} = \mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}$, cujo elemento i é e_i .

Myers & Montgomery (2002) afirmam que os resíduos *studentized* são os mais indicados para verificação das suposições de independência, normalidade e de que o valor esperado do erro é igual a zero, ou seja, $E(\varepsilon_i) = 0$. Os resíduos *studentized* são definidos por

$$r_i = \frac{e_i}{\sqrt{\hat{\sigma}^2 (1 - h_{ii})}}$$

onde h_{ii} são elementos da diagonal da matriz “chapéu” $\mathbf{H} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$, que é uma matriz de dimensão $n \times n$. A matriz chapéu é assim chamada pelo fato de ela “por um chapéu” (acento circunflexo) em \mathbf{y} , já que $\hat{\mathbf{y}} = \mathbf{H}\mathbf{y}$. Quando o modelo é correto, $\text{var}(r_i) = 1$ qualquer que seja a localização de \mathbf{x}_i (Myers & Montgomery, 2002, pg. 46).

Para verificar a suposição de independência, tem-se o gráfico de resíduos *studentized* das observações na ordem em que foram realizados os experimentos. Para verificar se o valor esperado do erro é igual a zero, tem-se o gráfico dos resíduos *studentized* versus valores ajustados. Para verificar a suposição de normalidade tem-se o gráfico de probabilidade normal dos resíduos *studentized*. Em nenhum desses três gráficos (não apresentados) houve indicação que essas suposições não devam ser aceitas.

Para verificar se a variância é constante, Cook & Weisberg (1999, pg. 346) propuseram um teste. Os quadrados dos resíduos $e_i^2 = (y_i - \hat{y}_i)^2$ contêm informação sobre a função de

variância. Procede-se então à regressão de e^2 sobre \hat{y} , pelo método dos mínimos quadrados. A estatística de teste (ET) é calculada dividindo-se $SS_R = \sum_{i=1}^n \left(\hat{e}_i^2 - \sum_{i=1}^n e_i^2 / n \right)^2$ pelo fator de escala $2 \left(\sum_{i=1}^n e_i^2 / n \right)$.

$$ET = SS_R / 2 \left(\sum_{i=1}^n e_i^2 / n \right)^2 \quad (4)$$

Cook & Weisberg (1999) asseguram que ET tem distribuição qui-quadrado com um grau de liberdade (número de termos da regressão de e^2 sobre \hat{y}), caso a variância seja constante. Para o caso do misto de retardo, fazendo a regressão de e^2 sobre \hat{y} , utilizando a planilha *Excel*, obtém-se SS_R e o fator de escala. Em seguida, utilizando a Equação 4, tem-se que $ET = 4,83$. Na distribuição qui-quadrado com um grau de liberdade, $ET = 4,83$ corresponde a um P -valor de 0,028. Então, rejeita-se a hipótese de que a variância não aumenta quando a média aumenta. Esta conclusão é confirmada pelo gráfico do valor absoluto dos resíduos *studentized* versus valores ajustados, apresentado na Figura 4. Nele se observa a indicação do crescimento da variância com o aumento da média.

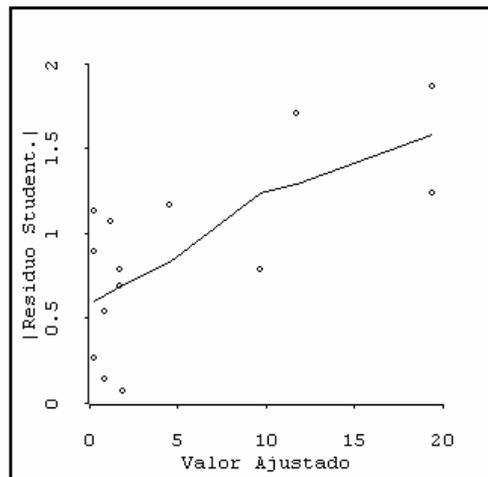


Figura 4 – Gráfico para verificação de não constância da variância.

3.2 Modelos lineares generalizados

No modelo linear clássico, com o método dos mínimos quadrados, três considerações são feitas: i) a resposta tem distribuição normal, ii) a variância da resposta é constante e iii) os efeitos dos fatores sobre a variável resposta se combinam aditivamente. Essas condições são encontradas em muitas aplicações industriais, porém nem sempre isso ocorre, como no presente caso.

Caso um ou mais dos três pressupostos não possam ser considerados, uma alternativa é alterar a escala da variável resposta, transformando-a e usando o modelo linear e o método

dos mínimos quadrados com os dados transformados. A transformação da resposta muitas vezes pode ser satisfatória, podendo-se alcançar os três pressupostos. Entretanto, quando não se consegue isso, deve-se recorrer aos modelos lineares generalizados (MLG). Ademais, Lewis *et al.* (2001a) e (2001b) concluíram que os MLG devem ser preferidos à transformação da resposta, uma vez que, em todos os casos estudados pelos autores, os MLG apresentaram melhor desempenho na estimativa da resposta, pois propiciaram intervalos de confiança menores.

Com os MLG, é possível considerar outras distribuições além da normal, não é necessário considerar a variância constante (ela pode ser função da média) e pode-se conseguir linearidade por meio de uma função que faz a ligação entre a média da variável resposta e o polinômio linear das variáveis independentes. Nos modelos lineares generalizados, usam-se as distribuições da família exponencial e a estimativa dos coeficientes do modelo é obtida com a maximização da função de log-verossimilhança. No caso de distribuição normal, efeitos aditivos e variância constante, a maximização da função de log-verossimilhança leva aos estimadores de mínimos quadrados, de modo que os modelos lineares ordinários podem ser considerados um caso particular dos MLG. Para uma descrição dos MLG, ver McCullagh & Nelder (1989) e Myers *et al.* (2002).

3.2.1 Estrutura dos MLG

Considere um experimento com os dados da Tabela 3, com n respostas independentes y_i .

Tabela 3 – Dados para o modelo.

x_1	x_2	...	x_k	y
x_{11}	x_{12}	...	x_{1k}	y_1
x_{21}	x_{22}	...	x_{2k}	y_2
.
.
.
x_{n1}	x_{n2}	...	x_{nk}	y_n

Temos então que:

1. Sejam y_1, y_2, \dots, y_n as variáveis de resposta com médias $\mu_1, \mu_2, \dots, \mu_n$.
2. A distribuição de probabilidade de y_i é um dos membros da família exponencial.
3. A porção sistemática do modelo é composta pelas variáveis de regressão x_1, x_2, \dots, x_k .
4. O modelo é construído com um preditor linear $\eta = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k$. Cabe registrar que uma variável x_i pode representar um termo de um grau mais elevado. Por exemplo $x_1^2 x_m$.
5. A função de ligação $g(\mu_i)$ faz a ligação entre a média μ_i e o preditor linear. A função de ligação define a forma com que os efeitos sistemáticos de x_1, x_2, \dots, x_k são transmitidos para a média. $\eta_i = g(\mu_i) = (\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik})$.

6. Para os membros da família exponencial, a variância da resposta tem a seguinte expressão: $\text{var}(y_i) = \phi V(\mu_i)$. Onde $V(\mu_i)$ é a parte da variância da resposta y que depende da média e ϕ é o parâmetro de dispersão que não depende da média e é constante.

A relação da função de variância $V(\mu_i)$ com a média pode ser representada por uma função de potência, ou seja, $V(\mu_i) = \mu_i^t$.

Para testar a significância dos coeficientes, Atkinson & Riani (2000), Lindsey (1997) e McCullagh & Nelder (1989) recomendam usar a função desvio (*deviance*).

A função desvio está para o método dos MLG como a soma dos quadrados dos resíduos está para o método dos mínimos quadrados. A função desvio de um modelo qualquer é definida como sendo o desvio deste modelo em relação ao modelo saturado, conforme a definição:

$D = -2 \ln[L_{Mod} / L_{Sat}]$, onde L_{Mod} é a função de verossimilhança do modelo em questão e L_{Sat} é a função de verossimilhança do modelo saturado, que é o modelo para o qual os valores ajustados $\hat{\mu}_i$ são iguais às respostas observadas y_i . Com isso,

$$D(y_i, \hat{\mu}_i) = -2 \ln[L(y_i, \hat{\mu}_i) / L(y_i, y_i)] = -2 [\ln L(y_i, \hat{\mu}_i) - \ln L(y_i, y_i)]$$

Lindsey (1997, pg. 214) propõe o procedimento de análise da função desvio para testar a significância dos coeficientes. Este procedimento está para os MLG assim como a análise de variância está para os modelos baseados no método dos mínimos quadrados.

3.3 Quase-verossimilhança

Muitas vezes, os membros da família exponencial não são distribuições adequadas para representar a variável resposta. Como foi visto, para o caso do misto de retardo, a variância possui uma tendência de crescimento com o aumento da média. Portanto, considerando, em princípio, $t = 1$, ou seja, $V(\mu_i) = \mu_i$, tem-se que a distribuição correspondente da família exponencial é a distribuição de Poisson. Entretanto, a variável aleatória de Poisson é discreta, o que não é o caso do misto de retardo, cuja resposta é uma variável contínua. Ademais, para o caso da distribuição de Poisson o parâmetro de dispersão ϕ é igual a 1, enquanto que, para o presente caso, como será visto mais adiante, $\hat{\phi} = 0,023$.

Para os casos em que um dos membros da família exponencial não são distribuições adequadas para representar a variável resposta há um método que estima os coeficientes maximizando a função de quase-verossimilhança (QV). No método da QV, define-se apenas a relação da variância da resposta com a média da resposta, não sendo necessário definir a distribuição de probabilidade. McCullagh & Nelder (1989, pg. 325) definem a função de QV como a seguir:

Para uma observação $Q_i(y_i, \mu_i) = \int_{y_i}^{\mu_i} \frac{y_i - u}{\phi V(u)} du$.

Para n observações $Q(\mathbf{y}, \boldsymbol{\mu}) = \sum_{i=1}^n Q_i(y_i, \mu_i) = \sum_{i=1}^n \int_{y_i}^{\mu_i} \frac{y_i - u}{\phi V(u)} du$.

Com $V(\mu_i) = \mu_i$, temos $Q(\mathbf{y}, \boldsymbol{\mu}) = \sum_{i=1}^n [y_i \ln(\mu_i) - \mu_i]$.

Omite-se o parâmetro ϕ porque não influi na estimativa dos coeficientes, sendo estimado após o ajuste do modelo.

Como foi visto, a relação da função de variância com a média pode ser representada por uma função de potência $V(\mu_i) = \mu_i^t$. Vieira (2004) demonstra que, para $t = 0, 1, 2$ e 3 , as funções de QV são iguais, respectivamente, às funções de log-verossimilhança (LV) das distribuições normal, de Poisson, gama e normal inversa, e que, para $V(\mu) = \mu(1-\mu)$, que é a função de variância da binomial, a função de QV corresponde a função de LV da binomial. Portanto, a maximização da função de QV, para estas funções de variância, produz as mesmas estimativas dos coeficientes que a função de LV nos MLG, podendo-se então usar o mesmo algoritmo de estimativa dos coeficientes que nos MLG, e conseqüentemente, qualquer *software* comercial que tenha a opção dos MLG. Para o caso do misto, a distribuição correspondente é a de Poisson. Usando esta opção nos MLG, obtemos as estimativas dos coeficientes.

Para testar a significância dos coeficientes Davison (2003) recomenda a estatística *quasi-deviance*. A *quasi-deviance* está para a modelagem pela função de quase-verossimilhança como a função desvio para os MLG. Por analogia, a *quasi-deviance* de um modelo qualquer é definida como sendo o desvio deste modelo em relação ao modelo saturado:

$$D_i(y_i, \hat{\mu}_i) = -2\phi[Q_i(y_i, \hat{\mu}_i) - Q_i(y_i, y_i)] = -2\phi[Q_i(y_i, \hat{\mu}_i)] = 2 \int_{\mu_i}^{y_i} \frac{y_i - u}{V(u)} du$$

onde $Q_i(y_i, \hat{\mu}_i)$ é a função quase-verossimilhança do modelo em questão e $Q_i(y_i, y_i)$ é a função de quase-verossimilhança do modelo saturado.

Para $V(u) = u$ tem-se:

$$D_i(y_i, \hat{\mu}_i) = 2 \int_{\mu_i}^{y_i} \frac{y_i - u}{u} du = 2[y_i \ln(y_i/\mu_i) - (y_i - \mu_i)]$$

Para n observações tem-se: $D(\mathbf{y}, \hat{\boldsymbol{\mu}}) = \sum_{i=1}^n D_i(y_i, \hat{\mu}_i)$.

Vieira (2004) demonstra que, para $t = 0, 1, 2$ e 3 , e para $V(\mu) = \mu(1-\mu)$, as estatísticas *quasi-deviance* são iguais, respectivamente, às funções desvio das distribuições normal, de Poisson, gama, normal inversa e binomial, também fornecidas por qualquer *software* comercial que tenha a opção de MLG.

Considerando inicialmente o modelo quadrático, obtivemos as estimativas dos coeficientes. Observamos que a predição do tempo em um dos pontos experimentais é negativa, o que indica que o modelo quadrático não é adequado para este caso. Com isso, vamos adicionar ao modelo quadrático, seqüencialmente, os termos do modelo cúbico.

A Tabela 4 mostra os valores da *quasi-deviance* com a adição seqüencial de cada um dos termos $x_1x_2x_3$, $x_1x_2(x_1 - x_2)$, $x_1x_3(x_1 - x_3)$ e $x_2x_3(x_2 - x_3)$ ao modelo quadrático.

Tabela 4 – Quasi-Deviance.

Termo Adicionado	Quasi-Deviance
$x_1x_2x_3$	0,824
$x_1x_2(x_1 - x_2)$	2,035
$x_1x_3(x_1 - x_3)$	0,163
$x_2x_3(x_2 - x_3)$	2,316

Será escolhido, portanto, o termo $x_1x_3(x_1 - x_3)$ para ser adicionado ao modelo quadrático, já que este é o que proporciona a menor *quasi-deviance*, conforme apresentado na Tabela 4.

Procede-se então com uma análise de *quasi-deviance* para a adição de novos termos ao modelo base, o qual possui sete termos ($x_1, x_2, x_3, x_1x_2, x_1x_3, x_2x_3$ e $x_1x_3(x_1 - x_3)$).

Considere o Modelo 1 como sendo o saturado, o Modelo 2 com $p + 1$ parâmetros e o Modelo 3, aninhado no Modelo 2, com p parâmetros. Então, a diferença entre as *quasi-deviance* dos modelos 3 e 2 é $D_3(y_i, \hat{\mu}_i^{(3)}) - D_2(y_i, \hat{\mu}_i^{(2)})$.

Para testar o termo adicional do modelo maior Davison (2003, pg. 515) recomenda o teste da distribuição F . A estatística de teste é: $F_0 = (D_3(y_i, \hat{\mu}_i^{(3)}) - D_2(y_i, \hat{\mu}_i^{(2)})) / \hat{\phi}_2$, sendo que $\hat{\phi}_2$ é a estimativa parâmetro de dispersão do modelo 2:

$$\hat{\phi}_2 = \frac{1}{(n-p)} \sum_{i=1}^n \frac{(y_i - \hat{\mu}_i^{(2)})^2}{V(\hat{\mu}_i^{(2)})} = \frac{X^2}{(n-p)}$$

onde X^2 é a estatística generalizada de Pearson e, para o presente caso, $V(\hat{\mu}_i^{(2)}) = \hat{\mu}_i^{(2)}$.

Portanto, para aceitar o Modelo 2 como sendo correto, comparamos F_0 com a distribuição $F_{1,n-p}$.

Os resultados da análise de *quasi-deviance* são apresentados na Tabela 5.

Tabela 5 – Análise de Quasi-Deviance.

Termo Adicionado	Quasi-Deviance	Diferença na Quasi-Deviance	$\hat{\phi}_2$	F_0	P-valor
$x_1x_3(x_1 - x_3)$	0,163				
$x_1x_2x_3$	0,136	0,027	0,023	1,17	0,32
$x_1x_2(x_1 - x_2)$	0,156	0,007	0,026	0,27	0,62
$x_2x_3(x_2 - x_3)$	0,136	0,027	0,023	1,17	0,32

Analisando a coluna “P-valor” da Tabela 5, conclui-se que os demais termos cúbicos não são significativos quando adicionados ao modelo base, já que todos os P-valores são maiores que 0,05.

Com isso, temos o modelo:

$$\hat{y} = 305,89x_1 + 8444,77x_2 - 242540,70x_3 - 12023,23x_1x_2 + 399292,58x_1x_3 + 144387,77x_2x_3 - 166047,61x_1x_3(x_1 - x_3)$$

3.3.1 Verificação da adequação do modelo

Para verificar a adequação do modelo pelo o método de QV, são utilizados os resíduos. Para cada resposta y_i pode-se definir a *quasi-deviance* $d_i = D_i(y_i, \hat{\mu}_i)$, que é usada como medida de discrepância. Cada resposta i contribui com uma quantidade d_i , de tal modo que $\sum_1^n d_i = D(y_i, \hat{\mu}_i)$. Define-se então o resíduo *quasi-deviance* correspondente a cada resposta:

$$r_{Di} = \text{sinal}(y_i - \hat{\mu}_i) \sqrt{d_i}.$$

McCullagh & Nelder (1998, pg. 397) recomendam os resíduos *quasi-deviance studentized* para verificar a adequação do modelo, os quais são definidos como sendo

$$r_i = \frac{r_{Di}}{\sqrt{\hat{\phi}(1-h_{ii})}}.$$

onde $\hat{\phi}$ é a estimativa do parâmetro de dispersão e h_{ii} é o i -ésimo elemento da diagonal da matriz chapéu \mathbf{H} , que é dada por: $\mathbf{H} = \mathbf{W}^{-\frac{1}{2}} \mathbf{X} (\mathbf{X}' \mathbf{W} \mathbf{X})^{-1} \mathbf{X}' \mathbf{W}^{-\frac{1}{2}}$, onde a matriz \mathbf{W} é uma matriz diagonal, com os elementos da diagonal principal dados por: $w_i = [1/V(\mu_i)] [d\mu_i/d\eta_i]^2$.

No presente caso, temos que $\hat{\phi} = 0,023$ e $V(\mu_i) = \mu_i$, $d\mu_i/d\eta_i = 1$, pois a função de ligação é a identidade. Logo, $w_i = 1/\mu_i$.

A adequação do modelo e o comportamento dos resíduos com valores extremos podem ser observados com o gráfico de probabilidade normal dos resíduos *quasi-deviance studentized* (Lee & Nelder, 1998, pg. 98). Para o caso em questão, é mostrado na Figura 5.

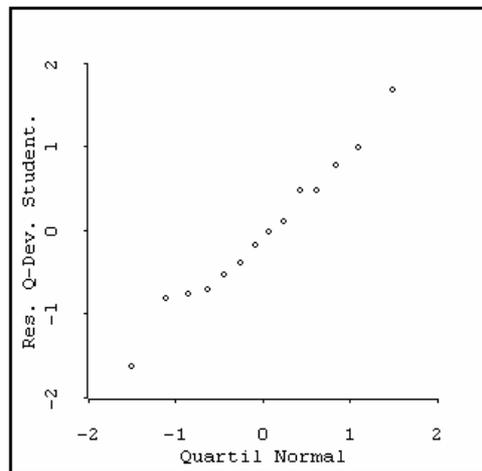


Figura 5 – Gráfico de probabilidade normal.

Nesse gráfico, pode-se observar que não há pontos muito fora do alinhamento, não havendo, portanto, a indicação de que o modelo seja inadequado.

A função de ligação pode ser verificada com o gráfico dos resíduos *quasi-deviance studentized* versus valores ajustados (Lee & Nelder, 1998, pg. 99). Esse gráfico é mostrado na Figura 6.

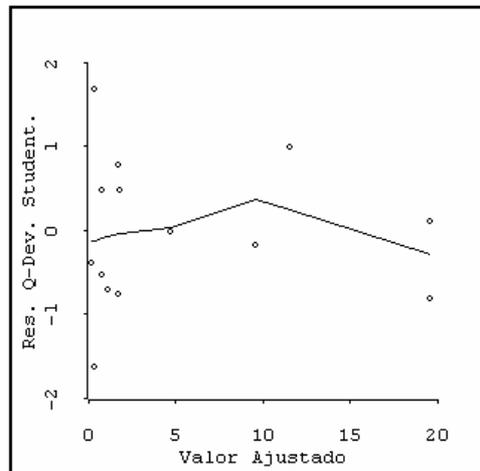


Figura 6 – Resíduo *quasi-deviance studentized* versus valor ajustado.

Pode-se observar que os resíduos, aleatoriamente distribuídos, não apresentam uma tendência de crescimento ou decrescimento em relação à média. A linha resultante do amortecimento (*lowess*) fica situada nas proximidades da reta horizontal de ordenada zero, não havendo indicação de que a função de ligação não é correta.

A função de variância pode ser verificada com o gráfico do valor absoluto dos resíduos *quasi-deviance studentized* versus valores ajustados (Lee & Nelder, 1998, pg. 99). Esse gráfico é mostrado na Figura 7.

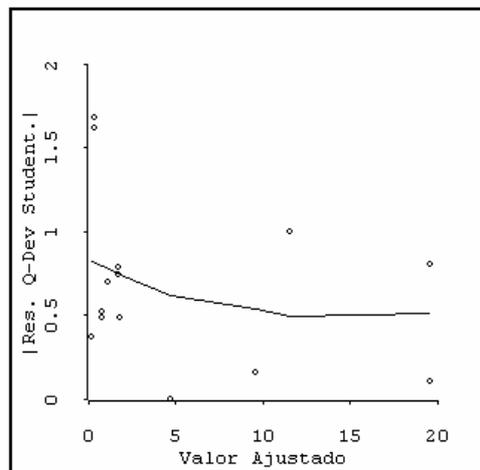


Figura 7 – $|\text{Resíduo } quasi\text{-deviance } studentized|$ versus valor ajustado.

No gráfico da Figura 7 pode-se observar que o valor absoluto dos resíduos *quasi-deviance studentized* não apresenta tendência de crescimento em relação à média, não havendo, portanto, a indicação de função de variância incorreta.

Com isso, chega-se à conclusão de que o modelo ajustado pelo método da quase-verossimilhança, é adequado para representar a resposta no caso do experimento do misto de retardo.

3.3.2 Otimização da resposta

Como se sabe, no experimento do misto de retardo, é desejável a obtenção de um valor de 11s para a resposta, já que a especificação de projeto determina que o tempo de retardo do mecanismo deve ser de $11,0 \pm 1,0$ s. Várias formulações podem resultar em uma média de futuras previsões da resposta de 11s. Por conseguinte, um objetivo desejável é minimizar a variância de uma futura resposta entre as formulações que resultam com 11s de previsão.

Supondo, um ponto qualquer $\mathbf{x}'_0 = [x_1 \ x_2 \ x_3 \ x_1x_2 \ x_1x_3 \ x_2x_3 \ x_1x_3(x_1 - x_3)]$ no espaço dos componentes da mistura, pode-se mostrar que a variância de uma futura previsão é (ver p.ex. Davison, 2003, pg. 511)

$$\text{var}[y(\mathbf{x}_0)] = \text{var}\{E[y(\mathbf{x}_0)]\} + E\{\text{var}[y(\mathbf{x}_0)]\}. \quad (5)$$

Tem-se que:

$$\text{var}\{E[y(\mathbf{x}_0)]\} = \text{var}[\hat{y}(\mathbf{x}_0)] = \text{var}(\mathbf{x}'_0\hat{\boldsymbol{\beta}}) = \mathbf{x}'_0 \text{var}(\hat{\boldsymbol{\beta}})\mathbf{x}_0 \quad (6)$$

Davison (2003), pg. 513, demonstra que, se a função de variância é a correta, tem-se que:

$$\text{var}(\hat{\boldsymbol{\beta}}) = \phi(\mathbf{X}'\mathbf{W}\mathbf{X})^{-1} \quad (7)$$

Substituindo (7) em (6) obtém-se:

$$\text{var}\{E[y(\mathbf{x}_0)]\} = \mathbf{x}'_0\phi(\mathbf{X}'\mathbf{W}\mathbf{X})^{-1}\mathbf{x}_0 \quad (8)$$

No modelo escolhido, tem-se que:

$$\text{var}[y(\mathbf{x}_0)] = \phi\mu(\mathbf{x}_0)$$

Então, tem-se que:

$$E\{\text{var}[y(\mathbf{x}_0)]\} = E[\phi\mu(\mathbf{x}_0)] = \phi\mathbf{x}'_0\boldsymbol{\beta} \quad (9)$$

Substituindo (8) e (9) em (5) obtém-se:

$$\text{var}[y(\mathbf{x}_0)] = \phi\mathbf{x}'_0(\mathbf{X}'\mathbf{W}\mathbf{X})^{-1}\mathbf{x}_0 + \phi\mathbf{x}'_0\boldsymbol{\beta}$$

Podemos então formular o problema da seguinte forma:

$$\min \text{var}[y(\mathbf{x}_0)] = \phi \mathbf{x}'_0 (\mathbf{X}'\mathbf{W}\mathbf{X})^{-1} \mathbf{x}_0 + \phi \mathbf{x}'_0 \boldsymbol{\beta}$$

sujeito a:

$$E[y(\mathbf{x}_0)] = \mathbf{x}'_0 \boldsymbol{\beta} = 11$$

$$x_1 + x_2 + x_3 = 1$$

$$0,79 \leq x_1 \leq 0,87$$

$$0,08 \leq x_2 \leq 0,16$$

$$0,05 \leq x_3 \leq 0,07$$

Utilizando uma rotina de busca exaustiva, na região das restrições dos componentes da mistura, codificada na linguagem do MATLAB, foi encontrada a seguinte solução:

$$x_1 = 0,8051 \quad x_2 = 0,1412 \quad x_3 = 0,0537$$

$$\text{var}[y(x_1, x_2, x_3)] = 0,3035$$

$$E[y(x_1, x_2, x_3)] = 11,0003$$

A estimativa do desvio-padrão de uma futura resposta é então igual a $\hat{\sigma}_0 = \sqrt{0,3035} = 0,5509$. Sabe-se que respostas fora do intervalo 11 ± 1 ou $11 \pm 1,82 \hat{\sigma}_0$ da média resultam em produtos fora da especificação.

Como não se conhece a distribuição de probabilidade da resposta, não se pode estimar a probabilidade de um ponto fora da especificação. Entretanto, por experiência própria e, acreditamos, dos usuários de estatística aplicada, pode-se afirmar que a probabilidade de respostas além de $11 \pm 1,82 \hat{\sigma}_0$ é não desprezível, qualquer que seja a distribuição de probabilidade da resposta na prática.

Portanto, há evidências de que a solução atual resultaria em um processo produtivo de qualidade inadequada. É necessário continuar com novos experimentos.

4. Conclusões

Foi apresentado um estudo de um misto químico, identificado como um caso típico de experimento com mistura. Nesse tipo de experimento tem-se usado sistematicamente o modelo linear clássico, baseado no método dos mínimos quadrados. Entretanto, para o caso do misto, este modelo foi considerado inadequado, uma vez que há indicação de que a variância da resposta não é constante, aumentando quando a média aumenta, o que viola um dos pressupostos do modelo linear clássico.

Foi então apresentado um modelo onde a estimativa e a inferência dos parâmetros são baseados no método de quase-verossimilhança. No método de quase-verossimilhança não é necessário definir a distribuição de probabilidade da resposta, para tanto basta definir a relação entre a média e a variância, e a forma com que os efeitos sistemáticos das variáveis regressoras são transmitidos para a média.

Os autores enfatizam que este é o único caso, por eles conhecido, que considera um experimento com mistura com resposta não-normal.

Com o modelo, foi possível determinar a proporção de cada um dos três componentes do misto, de modo que, a previsão do tempo de retardo tenha variância mínima e valor esperado igual a 11s. Entretanto, houve evidências de que a solução atual resultaria em um processo produtivo de qualidade inadequada, sendo, portanto, necessário continuar com novos experimentos.

Para futuro estudo, recomenda-se o estreitamento das restrições nas proporções dos componentes do misto, visto que as restrições anteriores proporcionaram um projeto com pontos experimentais, cujas respostas resultaram no intervalo de 0,23s até 19,70s. Tal estreitamento visa obter um projeto com pontos experimentais mais próximos à solução desejada, de modo a reduzir o intervalo da resposta observada.

Agradecimentos

Os autores agradecem aos revisores anônimos pelos úteis comentários e sugestões.

Referências Bibliográficas

- (1) Atkinson, A.C. & Riani, M. (2000). *Robust Diagnostic Regression Analysis*. Springer-Verlag, New York, NY.
- (2) Cornell, J.A. (1990). *Experiments with Mixtures: Designs, Models and the Analysis of Mixture Data*. Second edition, John Wiley & Sons, New York, NY.
- (3) Cook, R.D. & Weisberg, S. (1999). *Applied Regression Including Computing and Graphics*. John Wiley & Sons, New York, NY.
- (4) Dal Bello, L.H.A. (2005). Experimentos com Mistura: Uma Aplicação com Respostas Não-Normais. Dissertação de Mestrado, PUC-Rio, Departamento de Engenharia Industrial.
- (5) Davison, A.C. (2003). *Statistical Models*. Cambridge University Press, Cambridge, UK.
- (6) Goldfarb, H.B.; Borror, C.M. & Montgomery, D.C. (2003). Mixture-Process Variable Experiments with Noise Variables. *Journal of Quality Technology*, **35**, 393-405.
- (7) Goldfarb, H.B.; Borror, C.M.; Montgomery, D.C. & Anderson-Cook, C.M. (2004a). Three-Dimensional Variance Dispersion Graphics for Mixture-Process Experiments. *Journal of Quality Technology*, **36**, 109-124.
- (8) Goldfarb, H.B.; Borror, C.M.; Montgomery, D.C. & Anderson-Cook, C.M. (2004b). Evaluating Mixture-Process with Control and Noise Variables. *Journal of Quality Technology*, **36**, 245-262.
- (9) Kowalski, S.M.; Cornell, J.A. & Vining, G.G. (2002). Split-Plot Designs and Estimation Methods for Mixture Experiments with Process Variables. *Technometrics*, **44**, 72-79.
- (10) Lee, Y. & Nelder, J.A. (1998). Generalized Linear Models for the Analysis of Quality Improvement Experiments. *The Canadian Journal of Statistics*, **26**, 95-105.

- (11) Lewis, S.L.; Montgomery, D.C. & Myers, R.H. (2001a). Examples of Designed Experiments with Nonnormal Responses. *Journal of Quality Technology*, **33**, 265-278.
- (12) Lewis, S.L.; Montgomery, D.C. & Myers, R.H. (2001b). Confidence Interval Coverage for Designed Experiments Analysed With GLMs. *Journal of Quality Technology*, **33**, 279-292.
- (13) Lindsey, J.K. (1997). *Applying Generalized Linear Models*. Springer-Verlag, New York, NY.
- (14) McCullagh, P. & Nelder, J.A. (1989). *Generalized Linear Models*. Chapman-Hall, London, UK.
- (15) Myers, R.H. & Montgomery, D.C. (1997). A Tutorial on Generalized Linear Models. *Journal of Quality Technology*, **29**, 274-291.
- (16) Myers, R.H. & Montgomery, D.C. (2002). *Response Surface Methodology*. Second edition, John Wiley & Sons, New York, NY.
- (17) Myers, R.H.; Montgomery, D.C. & Vining, G.G. (2002). *Generalized Linear Models with Applications in Engineering and the Sciences*. John Wiley & Sons, New York, NY.
- (18) Prescott, P. (2004). Modelling in Mixture Experiments Including Interactions with Process Variables. *Quality Technology & Quantitative Management*, **1**, 87-103.
- (19) Prescott, P.; Dean, A.M.; Draper, N.R. & Lewis, S.M. (2002). Mixture Experiments: ILL-Conditioning and Quadratic Model Specification. *Technometrics*, **44**, 260-268.
- (20) Vieira, A.F.C. (2004). Análise da Média e Dispersão em Experimentos Fatoriais não Replicados para Otimização de Processos Industriais. Tese de Doutorado, PUC-Rio, Departamento de Engenharia Industrial.