

IDENTIFICAÇÃO DE EFEITOS DE DISPERSÃO EM EXPERIMENTOS FATORIAIS DOS TIPOS 2^K E 2^{K-P} : UM ESTUDO DE CASO

Viviane Leite Dias de Mattos

Universidade Católica de Pelotas. Programa de Pós-Graduação em Engenharia de Produção/UFSC.
e-mail: viviane@atlas.ucpel.tche.br

Pedro Alberto Barbetta

Universidade Federal de Santa Catarina. Departamento de Informática e Estatística/UFSC.
e-mail: barbetta@inf.ufsc.br

Robert Wayne Samohyl

Universidade Federal de Santa Catarina. Programa de Pós-Graduação em Engenharia de Produção/UFSC.
e-mail: samohyl@eps.ufsc.br

Abstract: This paper comments a strategy for analyzing location and dispersion effects in unreplicated factorial experiments, especially methods for identifying dispersion effects. A case is presented like an unreplicated experiment and like a replicated experiment, and the results are compared.

Key Words: dispersion effects, experimental design, quality improvement.

1 Introdução

Avaliar a variabilidade de um produto ou processo produtivo é fundamental para a obtenção da melhoria da qualidade. Embora Shewhart e Deming já tivessem mostrado a necessidade do entendimento e redução da variabilidade, foi Taguchi quem sugeriu a metodologia de projeto robusto, por meio da qual produtos e processo tornam-se insensitivos a fontes de variação. De acordo com esta metodologia, realizam-se experimentos planejados, procurando-se identificar os fatores que afetam a média ou a locação e os fatores que afetam a variabilidade ou a dispersão do processo. Busca-se encontrar a combinação ótima, que aproxima ao máximo a média do processo ao valor alvo e, ao mesmo tempo, minimiza a variabilidade.

Na indústria, durante as duas últimas décadas, tem sido dada, muita importância à identificação destes efeitos de dispersão em experimentos, principalmente os não replicados por serem mais rápidos e econômicos.

O objetivo deste artigo é apresentar e comentar a estratégia proposta por Brenneman (2000) e Brenneman e Nair (2001) para identificação de efeitos de dispersão em experimentos fatoriais não replicados, ilustrando com um estudo de caso.

2 Efeitos de dispersão

A identificação de efeitos de dispersão iniciou-se em experimentos com replicações, sendo a variabilidade avaliada por meio da variância amostral, calculada em cada ponto experimental. Embora este método apresente a vantagem de não depender de

resultados anteriores, é considerado eficiente apenas em experimentos com mais de 9 replicações (eficiência acima de 90% de acordo com Bartlett e Kendal, 1946), podendo apresentar um alto custo pelo grande número de ensaios requerido. Nos casos de poucas replicações, comum nos experimentos industriais devido ao alto custo das replicações, Barbetta et al. (1999) e Barbetta et al. (2000) apresentam procedimentos alternativos.

Como o custo é um fator extremamente importante, vários autores, tais como: Box e Meyer (1986), Bergman e Hynén (1997) e Wang (1989), entre outros, dedicaram-se a estudar métodos para identificação de efeitos de dispersão em experimentos não replicados, dando maior ênfase aos do tipo 2^{k-p} em função de sua grande aceitação no meio industrial.

Nesta situação a identificação dos efeitos de dispersão é feita a partir dos resíduos, ou seja, o que sobra depois que os efeitos de locação são removidos, apresentando, portanto, o inconveniente de depender da qualidade da modelagem do valor médio. Se esta modelagem, entretanto, é de boa qualidade, este método de identificação de fatores com efeitos de dispersão, teoricamente, conduz a bons resultados.

Entre os vários métodos propostos, talvez os mais divulgados sejam o método H proposto por Harvey (1976), o método BM proposto por Box e Meyer (1986), método W proposto por Wang (1989) e o método BH proposto por Bergman e Hynén (1997). É bastante utilizado também o método HM citado em Brenneman (2000) e Brenneman e Nair (2001).

O método H é uma extensão da proposta de Bartlett e Kendal (1946), relativa a experimentos replicados, com a substituição de variâncias por resíduos e avaliação do efeito de cada fator pela razão entre médias geométricas, com a suposição de modelo log-linear para a dispersão (1).

O método BM estuda os efeitos de dispersão por meio de uma estatística, D_k^{BM} , calculada a partir da razão entre médias aritméticas (2), sem fazer referência, entretanto, ao tipo de modelo que descreve a dispersão.

O método W emprega a estatística D_k^W (3), que apresentaria distribuição Qui-quadrado com um grau de liberdade, calculando-a a partir do quociente entre resíduos e estimativas da variância, sob a suposição de normalidade e variância constante, sendo a dispersão descrita por um modelo log-linear.

O método BH realiza uma modificação no método BM, definindo a estatística D_k^{BH} , que apresentaria distribuição F, sendo encontrada a partir de resíduos modificados calculados em função de um modelo expandido para a média (4). Não é feita referência sobre o modelo que descreve a dispersão.

O método HM resulta da junção do método H com o método BH. Aplica-se o método H aos resíduos modificados resultantes de um modelo de locação expandido, como o utilizado no método BH.

$$D_k^H = \frac{1}{N} \log \left(\frac{\prod_{s(k=+1)} r_j^2}{\prod_{s(k=-1)} r_j^2} \right) \quad (1)$$

$$D_k^{BM} = \frac{1}{2} \log \left[\frac{\sum_{s(k=+1)} r_j^2}{\sum_{s(k=-1)} r_j^2} \right] \quad (2)$$

$$D_k^W = \frac{1}{2N} \left(\sum_{s(k+)} \frac{r_j^2}{\sigma^2} - \sum_{s(k-)} \frac{r_j^2}{\sigma^2} \right) \quad (3)$$

$$D_k^{BH} = \frac{\sum_{s(k+)} r^2}{\sum_{s(k-)} r^2} \quad (4)$$

sendo r_j ($j=1, 2, \dots, N$) os resíduos do modelo da média.

3 A estratégia para identificação de efeitos de dispersão

Brenneman (2000) e Brenneman e Nair (2001) realizam inicialmente, a partir de uma reflexão teórica e avaliação de resultados de simulações, uma análise dos métodos supra citados que podem ser utilizados para identificar fatores com efeitos de dispersão, concluindo que:

- o método BM apresenta resultados viesados tanto para modelos de dispersão lineares como não lineares em função da correlação existente entre os resíduos. Este viés é sério mesmo quando são utilizadas replicações e não diminui com o aumento do tamanho do experimento;

- o método BH pode sofrer o mesmo tipo de viés estrutural que o método BM. Além disso, a estatística criada por estes autores só apresentará uma distribuição F se a hipótese nula considerar que todos os efeitos de dispersão são nulos;

- o método W é viesado para modelos de dispersão log-linear, além de não seguir uma distribuição Qui-quadrado quando existe mais de um efeito de dispersão ativo.

- o método H apresenta viés que depende da estrutura do efeito de dispersão ativo e da adequação do modelo, diminuindo com o aumento do tamanho do experimento;

- o método HM corrige em parte o viés do método H por usar o modelo de locação expandido.

Portanto, em todos os métodos não interativos analisados, os autores constataram a existência de um certo grau de viés, que pode ser decorrente da adequação do modelo de locação, mesmo que corretamente especificado, bem como da presença de múltiplos efeitos de dispersão. Estes métodos podem identificar não apenas falsos efeitos de dispersão, como também podem subestimá-los ou superestimá-los.

A partir desta análise crítica, Brenneman (2000) e Brenneman e Nair (2001) sugerem a utilização de métodos interativos para refinar o modelo e estimar os parâmetros de forma mais eficiente, propondo a seguinte estratégia para identificar, em experimentos não replicados, efeitos de locação e efeitos de dispersão que apresentem estrutura não linear:

- adequar um modelo de locação pelo método dos mínimos quadrados ordinários, identificando os efeitos de locação ativos (\mathcal{L});

- se \mathcal{L} apresentar pelo menos $n/2$ variáveis, a identificação dos efeitos de dispersão é difícil e é necessária experimentação adicional;

- se não, usar o método de Harvey, modificado ou não, para identificar efeitos ativos de dispersão (\mathcal{D}) e um método adequado para estimar o modelo;

- para um efeito k identificado como inerte, mas equivalente ao efeito ativo da interação entre dois outros efeitos d_1 e d_2 , incluí-lo no grupo de efeitos ativos considerando $k = d_1 \cdot d_2$;

- usando \mathcal{L} e \mathcal{D} nos modelos de locação e dispersão, respectivamente, repetir o procedimento empregando Máxima Verossimilhança Restrita. Descarte qualquer efeito de dispersão que agora foi identificado como inerte e refine o submodelo usando Máxima Verossimilhança Restrita. O método dos mínimos quadrados generalizados também pode ser utilizado para refinamento do modelo.

4 Estudo de caso

Com o objetivo de ilustrar a estratégia proposta para identificação e estimação de efeitos de dispersão, com modelo de estrutura não linear, foram utilizados dados de um experimento que buscava otimizar o tempo de resposta de uma central de comutação telefônica para serviços especializados, tais como: auxílio à lista, hora certa, teledespertador automático etc, em relação aos fatores: número de troncos (A), número de filtros (B), número de canais (C) e número de atendentes (D) (Silva, 2002). Embora o projeto experimental utilizado tenha sido do tipo 2^k com 4 replicações (Tabela 1), foi considerada apenas a primeira resposta (Y_1) para que se pudesse aplicar a estratégia interativa proposta, que se refere a experimentos não replicados, sendo os fatores com efeitos de dispersão identificados pelos métodos H e HM.

Tabela 1 – Projeto experimental, média e desvio-padrão da resposta relacionada à característica de qualidade analisada.

Exp.	Fatores				Interações						Respostas					
	A	B	C	D	AB	AC	AD	BC	BD	CD	Y1	Y2	Y3	Y4	M	S
1	-1	-1	-1	-1	1	1	1	1	1	1	51,414	51,576	51,330	51,443	51,441	,102
2	-1	-1	-1	1	1	1	-1	1	-1	-1	42,221	42,310	42,224	42,164	42,230	,060
3	-1	-1	1	-1	1	-1	1	-1	1	-1	51,414	51,576	51,330	51,443	51,441	,102
4	-1	-1	1	1	1	-1	-1	-1	-1	1	42,221	42,310	42,224	42,164	42,230	,060
5	-1	1	-1	-1	-1	1	1	-1	-1	1	51,414	51,576	51,330	51,443	51,441	,102
6	-1	1	-1	1	-1	1	-1	-1	1	-1	42,240	42,185	42,213	42,044	42,171	,087
7	-1	1	1	-1	-1	-1	1	1	-1	-1	51,414	51,576	51,330	51,443	51,441	,102
8	-1	1	1	1	-1	-1	-1	1	1	1	42,240	42,185	42,213	42,044	42,171	,087
9	1	-1	-1	-1	-1	-1	-1	1	1	1	66,575	66,869	66,594	66,809	66,712	,149
10	1	-1	-1	1	-1	-1	1	1	-1	-1	49,622	50,229	49,951	49,930	49,933	,248
11	1	-1	1	-1	-1	1	-1	-1	1	-1	66,575	66,869	66,594	66,809	66,712	,149
12	1	-1	1	1	-1	1	1	-1	-1	1	49,622	50,229	49,951	49,930	49,933	,248
13	1	1	-1	-1	1	-1	-1	-1	-1	1	66,809	66,869	66,881	66,535	66,774	,162
14	1	1	-1	1	1	-1	1	-1	1	-1	49,051	49,419	49,408	49,399	49,319	,179
15	1	1	1	-1	1	1	-1	1	-1	-1	66,809	66,869	66,881	66,535	66,774	,162
16	1	1	1	1	1	1	1	1	1	1	49,051	49,419	49,408	49,399	49,319	,179

OBS: A =tronco; B =filtro; C =canal; D =atendente.

A análise inicial destes dados, por meio do método ANOVA, mostrou que no conjunto dos fatores com efeitos de locação (\mathcal{L}) estão os fatores A, B e D, além de I (referente à média geral). A análise subsequente dos resíduos deste modelo pelo mesmo método dos mínimos quadrados ordinários, transformados pela função logarítmica, identifica I e D para a composição do conjunto dos fatores com efeitos de dispersão (\mathcal{D}). Logo, $\mathcal{L} = \{I, A, D, AD \text{ e } BD\}$ e $\mathcal{D} = \{I, D\}$.

Para confirmar a composição do conjunto \mathcal{D} foi traçado o gráfico de probabilidade normal a partir dos efeitos dos fatores e respectivas interações, medidos inicialmente pelo método H e, posteriormente, pelo método HM (Figura 1). De acordo com estes gráficos parece claro que, em ambos os casos, os fatores A e D, assim como a interação entre eles, AD, são ativos e, portanto, $\mathcal{D} = \{I, A, D \text{ e } AD\}$.

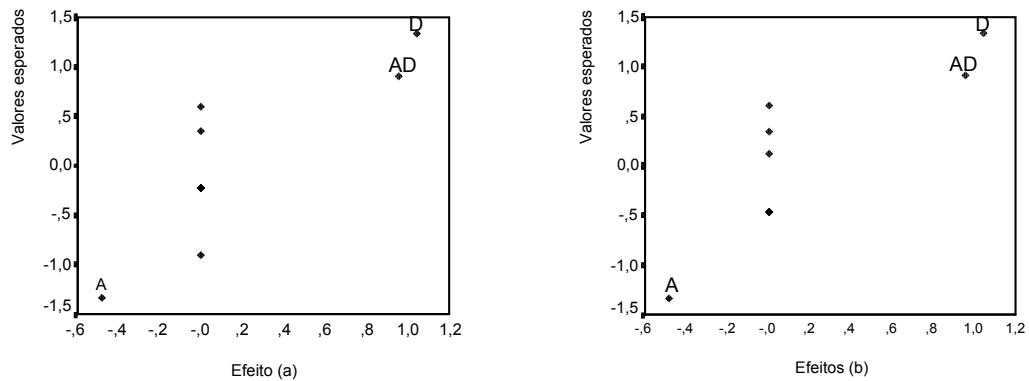


Figura 1 – Gráfico de probabilidade normal para efeitos de dispersão medidos pelos métodos H (a) e HM (b).

A modelagem do logaritmo do quadrado dos resíduos ($\log r^2$) pelo método dos mínimos quadrados ordinários permitiu fazer estimativas de σ^2 para cada condição experimental. Os valores de $(1/\sigma^2)$ são utilizados como pesos em uma nova modelagem do valor médio pelo método dos mínimos quadrados generalizados (MQG). Os resultados desta segunda análise modificam parcialmente a composição do conjunto \mathcal{L} , resultando em $\mathcal{L} = \{I, A, B, D, AD, BD \text{ e } AB\}$ e confirmam a composição do conjunto \mathcal{D} , isto é, $\mathcal{D} = \{I, A, D \text{ e } AD\}$, tanto pelo método H como pelo método HM que apresentam resultados muito semelhantes. Se o processo for repetido mais uma vez, estes resultados são novamente confirmados. Entretanto, ao aplicar MQG para a modelagem do valor médio, são utilizados mais de $n/2$ elementos, o que, segundo os autores, torna a modelagem da dispersão difícil.

Para avaliar qualidade dos resultados encontrados pela estratégia proposta foi feita, então, uma nova análise pelo método de Bartlett e Kendall, considerando as 4 replicações efetivamente realizadas no projeto experimental analisado, sendo a variabilidade avaliada pelo logaritmo da variância.

Os resultados mostrados na Figura 2 evidenciam que parecem ser ativos apenas os efeitos do fator A e da interação AD. Considerando que este método não apresenta viés na presença de replicações, é possível que o fator D, identificado pelos métodos anteriores, não apresente efeitos de dispersão.

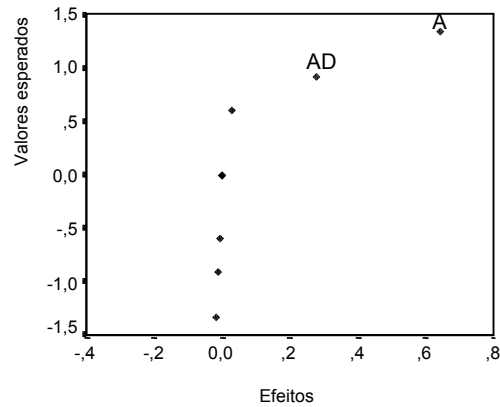


Figura 2 – Gráfico de probabilidade normal para efeitos de dispersão medido pelo método de Bartlett-Kendall.

5 Considerações Finais

A metodologia de análise de efeitos de dispersão proposta por Brenneman (2000) e Brenneman e Nair (2001) está baseada inicialmente no método de mínimos quadrados ordinários para identificar efeitos de locação e nos métodos H e MH para identificar efeitos de dispersão. Como alguns efeitos falsos podem ser detectados, existe necessidade de utilização do método dos mínimos quadrados generalizados de forma interativa. No estudo de caso realizado, foram encontrados os mesmos resultados para os métodos de identificação de efeitos de dispersão, embora tenham discordado dos resultados encontrados pelo método de Bartlett-Kendall, empregado com a mesma finalidade, mas para a situação de experimentos replicados.

Há concordância entre os resultados dos métodos H e HM, mas ambos podem estar apresentando um certo grau de viés, pois o modelo de locação apresenta mais de $n/2$ elementos. O resultado do método de Bartlett-Kendall para experimentos replicados não foi igual aos resultados encontrados anteriormente, embora bastante parecido.

A partir do exposto, pode-se concluir que identificar efeitos de dispersão é uma tarefa extremamente difícil, principalmente em experimentos não replicados. Os resultados encontrados mostraram que ainda não existe uma teoria consagrada a respeito, pois os métodos e estratégias existentes podem apresentar um certo grau de viés sob determinadas condições. Salienta-se, ainda, que na área industrial, projetos de experimentos do tipo 2^k ou 2^{k-p} , não replicados ou com poucas replicações, têm grande aceitação por apresentarem custo relativamente baixo, sendo muito utilizados como ponto de partida de estudos mais detalhados.

6 Referências Bibliográficas

- Barbetta, P.A. (1998). “Construção de modelos para médias e variâncias na otimização experimental de produtos e processos”. Tese (Engenharia de Produção). Florianópolis.
- Barbetta, P.A.; Ribeiro, J.L.D.; Samohyl, R.W. (1998). “Uma nova fórmula de modelar variâncias em experimentos com poucas replicações”. Enegep98. Rio de Janeiro.
- Bartlett, M.S.; Kendall, D.G. (1946). “The Statistical Analysis of Variance-Heterogeneity and the Logarithmic Transformation.” *Journal of the Royal Statistics Society, Ser. B*, 8, 128-138.
- Bergman, B.; Hynén, A. (1997). “Dispersion Effects From Unreplicated Designs in 2^{k-p} Series”. *Technometrics*, 39, 191-198.

- Box, G.E.P.; Meyer, R.D. (1986). “*Dispersion Effects From Fractional Design.*” Technometrics, 28, 19-27.
- Brenneman, W.A. (2000). “*Inference for Location and Dispersion Effects in Unreplicated Factorial Experiments.*” Tese. Universidade de Michigan. 108p.
- Brenneman, W.A.; Nair, V.N. (2001). “*Methods for Identifying Dispersion Effects in Unreplicated Factorial Experiments: A Critical Analysis and proposed Strategies.*” Technometrics, 43, 388-405.
- Harvey, A. C. (1976). “*Estimating Regression Models with Multiplicative Heterocedasticity*”. Econometrica, Vol. 44, N°3, 461-465.
- Silva, M.V.D.(2002). “*Avaliação de Desempenho de uma Plataforma de Comutação Telefônica para Serviços Especializados de Atendimento ao Cliente*”. Dissertação (Ciências da Computação). Florianópolis.101p.

7 Agradecimentos

Este trabalho foi apoiado pelo Núcleo de Normalização e Qualimetria (NNQ) do Programa de Pós-Graduação em Engenharia de Produção (PPGEP) da Universidade Federal de Santa Catarina (UFSC) – <http://www.qualimetria.ufsc.br> e pela Escola de Educação da Universidade Católica de Pelotas. – <http://www.ucpel.tche.br>.