

UNIVERSIDADE FEDERAL DE SANTA CATARINA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE PRODUÇÃO

**IDENTIFICAÇÃO DE EFEITOS DE DISPERSÃO EM EXPERIMENTOS
COM POUCAS REPLICAÇÕES**

TESE DE DOUTORADO

VIVIANE LEITE DIAS DE MATTOS

Florianópolis, SC - Brasil.
2004

UNIVERSIDADE FEDERAL DE SANTA CATARINA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE PRODUÇÃO

IDENTIFICAÇÃO DE EFEITOS DE DISPERSÃO EM EXPERIMENTOS
COM POUCAS REPLICAÇÕES

VIVIANE LEITE DIAS DE MATTOS

TESE A SER APRESENTADA AO CURSO DE DOUTORADO DO PROGRAMA DE
PÓS-GRADUAÇÃO EM ENGENHARIA DE PRODUÇÃO DA UNIVERSIDADE
FEDERAL DE SANTA CATARINA COMO REQUISITO PARCIAL À OBTENÇÃO DO
TÍTULO DE
DOUTOR EM ENGENHARIA DE PRODUÇÃO

ORIENTADOR: Dr. PEDRO ALBERTO BARBETTA
CO-ORIENTADOR: ROBERT WAYNE SAMOHYL, Ph.D.

Florianópolis, SC - Brasil.
2004

UNIVERSIDADE FEDERAL DE SANTA CATARINA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE PRODUÇÃO

IDENTIFICAÇÃO DE EFEITOS DE DISPERSÃO EM EXPERIMENTOS
COM POUCAS REPLICAÇÕES

Esta tese foi julgada e aprovada para a obtenção do grau de Doutora em Engenharia de Produção no Programa de Pós-Graduação em Engenharia de Produção da Universidade Federal de Santa Catarina.

Florianópolis, 20 de agosto de 2004

Prof. Edson Pacheco Paladini, Dr.
Coordenador do Programa de Pós-Graduação em Engenharia de Produção

BANCA EXAMINADORA:

Prof. Pedro Alberto Barbeta, Dr.
Universidade Federal de Santa Catarina
Orientador

Prof. Robert Wayne Samohyl, Ph.D.
Universidade Federal de Santa Catarina
Co-Orientador

Prof. Paulo José Ogliari, Dr.
Universidade Federal de Santa Catarina
Examinador/Moderador

Prof. Dalton Francisco de Andrade, Ph.D.
Universidade Federal de Santa Catarina
Examinador

Prof. José Duarte Ribeiro, Dr.
Universidade Federal do Rio Grande do Sul
Examinador

Prof.a Clarice G. B. Demétrio, Dra.
Universidade de São Paulo
Examinadora

M435i Mattos, Viviane Leite Dias de
Identificação de efeitos de dispersão em experimentos
com poucas replicações / Viviane Leite Dias de Mattos. –
Florianópolis: UFSC, 2004.
100f.

Tese (doutorado) - Universidade Federal de Santa
Catarina, Programa de Pós – Graduação em Engenharia de
Produção, Florianópolis , BR - SC,
2004. Orientador: Barbetta, Pedro Alberto.
Co-orientador: Samohyl, Robert Wayne.

1. Estatística – projeto de experimentos – efeito de
dispersão. I . Barbetta, Pedro Alberto. II . Samohyl, Robert
Wayne. III. Título.

A Lil, Waldo, Marcelo e Luciano, que acompanharam zelosamente todas as etapas deste trabalho.

AGRADECIMENTOS

Sintetizar agradecimentos a todas as pessoas que, de uma maneira ou de outra, contribuíram para a realização deste trabalho, mostrou ser uma tarefa difícil. Gostaria de deixar registrada aqui minha imensa gratidão e meu reconhecimento a todos que contribuíram, direta ou indiretamente, para o êxito deste trabalho.

Um agradecimento especial aos professores Pedro Alberto Barbetta e Robert Wayne Samohyl pela amizade, atenção, incentivo e, principalmente, carinho, dedicação, paciência e sabedoria com que dirigiram minha caminhada.

Agradeço ainda:

aos meus familiares, o apoio constante;

aos professores Dalton Andrade e João Neiva de Figueiredo, as sugestões por ocasião do exame de qualificação;

à Universidade Católica de Pelotas, o incentivo moral e a ajuda financeira e, em especial, à professora Clarisse Siqueira Coelho, diretora da Escola de Educação;

à Fundação de Ciência e Tecnologia de Santa Catarina, o apoio financeiro;

aos colegas, Maria Luisa Cañas Martins, Ricardo Curi Terra, Graçaliz Pereira Dimuro, Marilton Sanchotene Aguiar, Renata Sander Reiser, Antonio Carlos Rocha e Vera Lúcia Marques de Figueiredo, a receptividade, o apoio ou as sugestões em algumas etapas deste trabalho;

à Alexandre A. Ferreira, a valiosa colaboração na implementação dos programas utilizados nas simulações;

aos colegas do Programa de Pós-Graduação em Engenharia de Produção da Universidade Federal de Santa Catarina, em especial, Rubson Rocha e Andréa Konrath, pela oportunidade de convívio;

aos professores do Programa de Pós-Graduação em Engenharia de Produção da Universidade Federal de Santa Catarina, os ensinamentos oferecidos;

aos funcionários da Universidade Católica de Pelotas e do Programa de Pós-Graduação em Engenharia de Produção da Universidade Federal de Santa Catarina, em especial, Aduino Scaloni, a solicitude com que atenderam os meus vários pedidos.

RESUMO

MATTOS, V.L.D. Identificação de efeitos de dispersão em experimentos com poucas replicações. Tese de Doutorado. Programa de Pós-Graduação em Engenharia de Produção e Sistemas. Universidade Federal de Santa Catarina, 2004.

O presente trabalho elabora uma estratégia capaz de identificar os efeitos de dispersão em experimentos fatoriais do tipo 2^K , com poucas replicações. O estudo começou com a seleção dos principais métodos para identificar efeitos de dispersão, em experimentos replicados e não-replicados. Os métodos selecionados utilizam estatísticas, baseadas em quocientes entre médias (aritméticas ou geométricas) de variâncias amostrais, resíduos quadráticos ou resíduos quadráticos modificados. Foram feitas algumas considerações sobre a utilização de resíduos quadráticos, empregados nos métodos para identificar efeitos de dispersão em experimentos não-replicados, estendendo-os à situação de experimentos com poucas replicações. Matematicamente, foi possível concluir que, nos métodos que empregam quociente de médias aritméticas, é equivalente usar todos os resíduos quadráticos ou suas médias aritméticas em cada ponto experimental. Entretanto, nos métodos que empregam quociente entre médias geométricas, simulações de Monte Carlo permitiram concluir que é melhor utilizar médias dos resíduos quadráticos. Em outras simulações de Monte Carlo realizadas para avaliar o desempenho dos métodos em termos de proporções de identificação correta de efeitos de dispersão e proporções de identificação falsa de efeitos de dispersão, os resultados sugerem que o desempenho de um método é mais influenciado pelo tipo de média empregado na estimativa do efeito: média aritmética ou média geométrica, principalmente se o experimento analisado apresentar apenas um efeito de dispersão. Os resultados também indicam que os métodos que utilizam resíduos quadráticos apresentam melhor desempenho, desde que o modelo de locação tenha sido corretamente identificado. Baseando-se nestes resultados, foi proposta uma estratégia iterativa para identificar efeitos de dispersão em experimentos do tipo 2^K , com poucas replicações, que também pode ser estendida para projetos experimentais fracionados do tipo 2^{K-p} . Ajusta-se, primeiramente, um modelo de locação, testando algumas eventuais transformações matemáticas. Se este modelo for de boa qualidade, trabalha-se com métodos que usam resíduos quadráticos e, em caso contrário, com métodos que usam variâncias amostrais. Primeiramente, usam-se métodos que empregam médias aritméticas e, se for detectado mais de um efeito de dispersão, métodos que empregam médias geométricas. Complementa-se a estratégia propondo um processo iterativo para estimar os efeitos de locação e dispersão, baseado no método dos mínimos quadrados ponderados. A estratégia foi aplicada em um experimento realizado com o objetivo de avaliar a influência da utilização da adição cinza de casca de arroz em concretos de alto desempenho, relacionando-a com vários parâmetros do seu processo produtivo, obtendo-se bons resultados.

ABSTRACT

MATTOS, V.L.D. Identification of dispersion effects in experiments with few replications. Doctorate Thesis. Programa de Pos-Graduação em Engenharia de Produção e Sistemas. Universidade Federal de Santa Catarina, 2004.

This thesis presents a strategy to identify dispersion effects in 2^k factorial experiments, with few replications. The study was begun by selecting the principal methods for identifying dispersion effects, in replicated and unreplicated factorial experiments. The methods selected used statistics with quotients between means (arithmetical or geometrical) of sample variances, squared residuals or modified squared residuals. Some considerations were carried out on the utilization of squared residuals, used in methods for identifying dispersion effects in unreplicated experiments. These methods are adapted for designs of experiments with few replications. Mathematics make it possible to search for a conclusion, about methods with arithmetical means quotients, that uses squared residuals and is equivalent to its arithmetical means at each experimental point. However, by Monte Carlo simulations, the methods that use geometrical means show better results than the square residuals means. The results of other Monte Carlo simulations, done to assess the performance method to find the proportions of correct identification of dispersion effects and proportions of false identification of dispersion effects, suggest that the methods performance is more influenced by the mean type used to estimate the effect: arithmetical mean or geometrical mean. The influence is greater if the experiment presents only one dispersion effect. The results also show that the methods using squared residuals have better performance if the location model is correctly identified. This thesis presents an iterative strategy to identify dispersion effects in 2^k factorial experiments, with few replications. It can be used in 2^{k-p} factorial fractionated experiments. First, a model location is adjusted. Some mathematical transformations are tested. If the model is of good quality, methods with arithmetical means of squared residuals are used. If the model is of doubtful quality, methods with arithmetical means of sample variances are used. In both cases, if more than one dispersion effect is detected, equivalent methods with the geometrical mean are carried out. The strategy is complemented by an iterative process to estimate location and dispersion effects using reweighted least square. The strategy was applied to an experiment to study the behavior of concrete containing rice husk ash. Good results were obtained.

LISTA DE FIGURAS

FIGURA 1.1	Sistema produtivo.....	20
FIGURA 1.2	Representação gráfica de efeitos de locação e efeitos de dispersão de fatores de um experimento fictício (Adaptado de Barbeta <i>et al.</i> , 1999).....	22
FIGURA 2.1	Representação gráfica de efeitos de fatores em experimento fictício (Adaptada de ZAR, 1999, p.240).....	32
FIGURA 2.2	Representação gráfica do efeito de fatores sem interação e com interação em experimento fictício (Adaptadas de ZAR, 1999, p.241).....	33
FIGURA 2.3	Projeto central composto (Adaptado de CATEN, 1995, p.64).....	35
FIGURA 2.4	Diagrama de Pareto para um exemplo fictício.....	39
FIGURA 2.5	Gráfico normal dos efeitos para um exemplo fictício.....	40
FIGURA 2.6	Gráfico do efeito do fator <i>B</i> para um exemplo fictício.....	40
FIGURA 2.7	Representação gráfica da função perda de Taguchi, onde LIE e LSE representam, respectivamente, os limites de especificação inferior e superior. (Adaptado de FIOD NETO, 1997, p.25).....	48
FIGURA 2.8	Funções de perda quadrática em termos do tipo de característica de qualidade.....	49
FIGURA 2.9	Gráfico linear para arranjo ortogonal L8 de Taguchi (FIOD NETO, 1997, p.46), onde <i>A</i> , <i>B</i> , <i>C</i> e <i>D</i> são os fatores manipulados	52
FIGURA 2.10	Matriz experimental de um arranjo-produto ou arranjo-cruzado (Adaptado de CATEN, 1995), onde <i>A</i> , <i>B</i> e <i>C</i> são fatores de controle e <i>D</i> e <i>E</i> são fatores de ruído. CE representa condição experimental	53
FIGURA 2.11	Representação gráfica dos valores reais das observações (a) e dos valores transformados (b). (Adaptado de ZAR, 1999, p.355).....	63
FIGURA 2.12	Função logarítmica.....	64
FIGURA 3.1	Fluxograma do programa computacional utilizado na simulação realizada para comparar o desempenho dos métodos investigados.....	81
FIGURA 4.1	Gráfico do efeito do fator <i>G</i> (método utilizado) sobre o valor médio da proporção de identificação correta de <i>todos</i> os efeitos de dispersão.	94
FIGURA 4.2	Gráficos dos efeitos das interações do fator <i>G</i> (método utilizado) com os fatores <i>A</i> , <i>B</i> , <i>C</i> , <i>D</i> , <i>E</i> e <i>F</i> sobre as proporções de identificação correta de <i>todos</i> os efeitos de dispersão.....	95
FIGURA 4.3	Gráfico do efeito do fator <i>G</i> (método utilizado) sobre a proporção de identificação falsa de <i>algum</i> efeito de dispersão.....	100
FIGURA 4.4	Gráfico do efeito das interações do fator <i>G</i> com os fatores <i>C</i> , <i>D</i> , <i>E</i> e <i>F</i> sobre a proporção de identificação falsa de <i>algum</i> efeito de dispersão	101
FIGURA 5.1	Fluxograma do programa computacional utilizado na simulação realizada para avaliar o desempenho dos métodos com a utilização de resíduos quadráticos individuais ou médios.....	109
FIGURA 5.2	Efeito do fator <i>L</i> (tratamento do resíduo quadrático) nas variáveis (a) <i>PIC</i> e (b) <i>PIF</i>	112
FIGURA 5.3	Fluxograma do programa computacional utilizado na simulação realizada para comparar o desempenho dos métodos investigados que empregam resíduos quadráticos e variâncias amostrais.....	115

FIGURA 5.4	Proporção média, por método, de identificação correta de <i>todos</i> os efeitos de dispersão nas condições experimentais simuladas.....	117
FIGURA 5.5	Efeito da interação do fator <i>G</i> (método) com o fator <i>C</i> (efeitos de dispersão) sobre a variável resposta <i>PIC</i> , nas condições experimentais simuladas.....	117
FIGURA 5.6	Proporção média, por método, de identificação falsa de <i>algum</i> os efeitos de dispersão nas condições experimentais simuladas.....	119
FIGURA 5.7	Efeito da interação do fator <i>G</i> (método) com o fator <i>C</i> (efeitos de dispersão) sobre a variável resposta <i>PIF</i> , nas condições experimentais simuladas.....	119
FIGURA 6.1	Fluxograma da estratégia para identificar possíveis efeitos de dispersão.....	125
FIGURA 6.2	Valores preditos e resíduos (a) e gráfico de probabilidade normal (b) dos resíduos do modelo descrito na equação 6.1.....	135
FIGURA 6.3	Valores preditos e resíduos (a) e gráfico de probabilidade normal (b) dos resíduos do modelo descrito na equação 6.2.....	136
FIGURA 6.4	Gráfico de probabilidade normal de efeitos de dispersão, construídos a partir de estimativas encontradas pelo método <i>R</i>	137
FIGURA 6.5	Gráfico de probabilidade normal de efeitos de dispersão, construídos a partir de estimativas encontradas pelo método <i>S</i>	138
FIGURA 6.6	Efeito principal do fator <i>A</i> sobre a resistência à compressão.....	139
FIGURA 6.7	Efeito principal do fator <i>D</i> sobre a resistência à compressão.....	139
FIGURA 6.8	Efeito principal do fator <i>E</i> sobre a resistência à compressão.....	139
FIGURA 6.9	Efeito da interação entre o fator <i>A</i> e o fator <i>E</i> sobre a resistência à compressão.....	140
FIGURA 6.10	Efeito da interação entre o fator <i>B</i> e o fator <i>C</i> sobre a resistência à compressão.....	140
FIGURA 6.11	Efeito da interação entre os fatores <i>ACD</i> sobre a resistência à compressão.....	141
FIGURA 6.12	Efeito principal do fator <i>A</i> sobre a variabilidade da resistência à compressão.....	142
FIGURA 6.13	Efeito principal do fator <i>B</i> sobre a variabilidade da resistência à compressão.....	142
FIGURA 6.14	Efeito da interação entre os fatores <i>A</i> e <i>B</i> sobre a variabilidade da resistência à compressão.....	142

LISTA DE TABELAS

TABELA 2.1	Níveis dos fatores e interações de um projeto fatorial do tipo 2^3	32
TABELA 2.2	Projeto fatorial fracionado do tipo 2^{3-1}	34
TABELA 2.3	Algumas características dos delineamentos propostos por Taguchi	50
TABELA 2.4	Projeto fatorial completo do tipo 2^3	51
TABELA 2.5	Projeto fatorial do tipo 2^3 para um projeto $L8$	51
TABELA 2.6	Delineamento de um projeto fatorial do tipo 2^5	52
TABELA 4.1	Proporção média de identificação correta de efeitos de dispersão, por método e por fator, nas condições experimentais simuladas	92
TABELA 4.2	Média e desvio-padrão, por método, de proporções de identificação correta de <i>todos</i> os efeitos de dispersão nas condições experimentais simuladas	92
TABELA 4.3	Análise do efeito do fator G (método utilizado) e suas interações com os demais fatores na variável transformada $\arcsen\sqrt{PIC}$	94
TABELA 4.4	Coeficientes utilizados nos contrastes	97
TABELA 4.5	Análise do efeito dos contrastes aplicados ao fator G (método utilizado) na variável transformada $\arcsen\sqrt{PIC}$	97
TABELA 4.6	Proporção média de identificação falsa de efeitos de dispersão, por fator, em 32 condições experimentais simuladas	98
TABELA 4.7	Média e desvio-padrão, por método, das proporções de identificação falsa de <i>algum</i> efeito de dispersão em 32 condições experimentais simuladas	99
TABELA 4.8	Análise do efeito do fator G (método utilizado) e suas interações com os demais fatores na variável transformada $\arcsen\sqrt{PIF}$	100
TABELA 4.9	Análise do efeito dos contrastes aplicados ao fator G (método utilizado) na variável transformada $\arcsen\sqrt{PIF}$	102
TABELA 5.1	Proporções médias de identificação correta de <i>todos</i> os efeitos de dispersão e de identificação falsa de <i>algum</i> efeito de dispersão nas condições experimentais simuladas	110
TABELA 5.2	Análise do efeito do fator L (tratamento do resíduo quadrático) e suas interações com os demais fatores na variável transformada $\arcsen\sqrt{PIC}$	111
TABELA 5.3	Análise do efeito do fator L (tratamento do resíduo quadrático) e suas interações com os demais fatores na variável transformada $\arcsen\sqrt{PIF}$	112
TABELA 5.4	Proporção média, por método, de identificação correta de <i>todos</i> os efeitos de dispersão nas condições experimentais simuladas	116
TABELA 5.5	Análise do efeito do fator G (tratamento do resíduo quadrático) e suas interações com os fatores C e F na variável transformada $\arcsen\sqrt{PIC}$	117
TABELA 5.6	Coeficientes utilizados nos contrastes	118
TABELA 5.7	Resultados da ANOVA aplicado aos contrastes	118
TABELA 5.8	Proporção média de identificação falsa de <i>algum</i> efeito de dispersão	118
TABELA 5.10	Análise do efeito do fator G (tratamento do resíduo quadrático) e suas interações com os fatores C e F na variável transformada	

	$\arcsen\sqrt{PIF}$	119
TABELA 5.10	Resultados da ANOVA aplicado aos contrastes.....	120
TABELA 6.1	Traços utilizados para confecção dos concretos estudados.....	129
TABELA 6.2	Teor de aditivo superplastificante empregado e abatimentos obtidos...	129
TABELA 6.3	Descrição dos fatores utilizados no experimento.....	132
TABELA 6.4	Teste F aplicado aos dados originais para identificação dos efeitos de locação.....	134
TABELA A1	PIC's obtidas no estudo de simulação realizado para comparar o desempenho dos métodos investigados.....	154
TABELA A2	PIF's obtidas no estudo de simulação realizado para comparar o desempenho dos métodos investigados.....	155
TABELA A3	PIC's obtidas no estudo de simulação realizado para avaliar o desempenho dos métodos H e HM quando são utilizados resíduos quadráticos individuais ou médios.....	156
TABELA A4	PIF's obtidas no estudo de simulação realizado para avaliar o desempenho dos métodos H e HM quando são utilizados resíduos quadráticos individuais ou médios.....	157
TABELA A5	PIC's obtidas no estudo de simulação que compara métodos que usam variâncias amostrais com métodos que usam resíduos quadráticos.....	158
TABELA A6	PIF's obtidas no estudo de simulação que compara métodos que usam variâncias amostrais com métodos que usam resíduos quadráticos.....	158
TABELA A7	Resistência à compressão (MPa) de corpos-de-prova de concreto.....	159

LISTA DE QUADROS

QUADRO 2.1	Desenvolvimento de uma ANOVA para avaliação da significância de fatores e interações em um projeto fatorial com dois fatores	38
QUADRO 2.2	Resultados da ANOVA para avaliação da qualidade de um modelo de regressão.....	43
QUADRO 3.1	Descrição de fatores e níveis utilizados no programa computacional utilizado na simulação realizada para comparar o desempenho dos métodos investigados.....	79
QUADRO 5.1	Descrição de fatores e níveis utilizados no estudo de simulação realizado para avaliar o desempenho dos métodos H e HM quando são utilizados resíduos quadráticos individuais ou médios.....	108
QUADRO 5.2	Descrição de fatores e níveis utilizados no estudo de simulação que compara métodos que usam variâncias amostrais com métodos que usam resíduos quadráticos.....	114

LISTA DOS PRINCIPAIS SÍMBOLOS E SIGLAS

Letras Latinas

A, B, C, D, E, F, G, L	Fatores manipulados em experimento
a	a -ésimo nível do fator A ($a=1, 2, \dots, A'$)
b	b -ésimo nível do fator B ($b=1, 2, \dots, B'$)
BH	Método de Bergman e Hynén
BN	Estratégia de Breneman e Nair
BM	Método de Box e Meyer
CE	Condição experimental
d	Afastamento máximo de y em relação ao valor alvo, que deve estar de acordo com os limites de especificação
D_k^S	Estimativa do efeito de dispersão do k -ésimo fator ou interação pelo método S
D_k^R	Estimativa do efeito de dispersão do k -ésimo fator ou interação pelo método R
D_k^H	Estimativa do efeito de dispersão do k -ésimo fator ou interação pelo método H
D_k^{HM}	Estimativa do efeito de dispersão do k -ésimo fator ou interação pelo método HM
$D_k^{BM(0)}$	Estimativa do efeito do fator ou interação do k -ésimo fator ou interação pelo método BM
D_k^{BM}	Estimativa do efeito de dispersão do k -ésimo fator ou interação pelo método BM
$D_k^{BH(0)}$	Estimativa do efeito de dispersão do k -ésimo fator ou interação pelo método BH
D_k^{BH}	Estimativa do efeito de dispersão do k -ésimo fator ou interação pelo método BH
$D_k^{BM(1)}$	Estatística D_k^{BM} calculada com a média dos resíduos quadráticos
$D_k^{BM(2)}$	Estatística D_k^{BM} calculada com os resíduos quadráticos individuais
$D_k^{H(1)}$	Estatística D_k^H calculada com a média dos resíduos quadráticos
$D_k^{H(2)}$	Estatística D_k^H calculada com os resíduos quadráticos individuais
$D \subset \mathfrak{R}^k$	Região experimental
\bar{D}_q	Média amostral das estatísticas que medem efeitos de dispersão na q -ésima amostra
\mathcal{D}^0	Conjunto inicial de efeitos de dispersão ativos
\mathcal{D}	Conjunto de efeitos de dispersão ativos
e	Erro experimental
f	Função que liga os fatores ao valor esperado de y
F	Estatística com distribuição de Snedecor-Fisher
gl	Graus de liberdade
h	Função que liga os fatores à variância de y
h_n	Coefficiente de correção
H	Método de Harvey
HM	Método de Harvey modificado
k'	Constante de proporcionalidade da função perda
K	Quantidade de fatores de controle manipulados em um experimento

K'	Quantidade de fatores ou interações possíveis ($k = 1, \dots, K'$)
\mathcal{L}	Conjunto de efeitos de locação
\mathcal{L}_E	Conjunto de efeitos de locação para o modelo expandido
L	Função perda
LIE	Limite de especificação inferior
LSE	Limite de especificação superior
m	m -ésima replicação ($m = 1, 2, \dots, M$)
MQO	Mínimos quadrados ordinários
MQG	Mínimos quadrados generalizados
MQP	Mínimos quadrados ponderados
MLE	Método de máxima verossimilhança
mEM	<i>menor-é-melhor</i>
MEM	<i>maior-é-melhor</i>
mod	Resto da divisão
n	n -ésima condição experimental ($n = 1, 2, \dots, N$)
\mathcal{N}	Distribuição normal
N	Tamanho da amostra
NEM	<i>nominal-é-melhor</i>
$n(k+)$	Condições experimentais do nível superior do k -ésimo fator ou interação
$n(k-)$	Condições experimentais do nível inferior do k -ésimo fator ou interação
p	Grau de fracionamento
$P_{\max.}$	Perda máxima devido a um item não-conforme
P	Proporção
PIC	Proporção de identificação correta de <i>todos</i> os efeitos de dispersão
PIF	Proporção de identificação falsa de <i>algum</i> efeito de dispersão
q	q -ésima amostra ($q = 1, 2, \dots, Q$)
QIC_k	Quantidade de identificações corretas do k -ésimo fator ou interação, por condição experimental
QIF_k	Quantidade de identificações falsas do k -ésimo fator ou interação, por condição experimental
QS	Quantidade de amostras simuladas
QIC	Quantidade de identificações corretas de <i>todos</i> os efeitos de dispersão, por condição experimental
QIF	Quantidade de identificações falsas de <i>algum</i> efeito de dispersão, por condição experimental
QM	Quadrado médio
R	Modelo restrito
R^2	Coefficiente de determinação
$R^2_{\text{corrigido}}$	Coefficiente de determinação corrigido
r^2	Resíduo quadrático
\tilde{r}^2	Resíduo quadrático modificado
$\overline{r^2}$	Média de resíduos quadráticos
$\overline{\tilde{r}^2}$	Média de resíduos quadráticos modificados
RMLE	Método de máxima verossimilhança restrita
S	Modelo saturado
SQ	Soma dos quadrados
s^2	Variância das observações do experimento

S/R	Razão sinal-ruído
$S_{D_q}^2$	Variância amostral das estatísticas que medem efeitos de dispersão na q -ésima amostra
U	Distribuição uniforme
u_0	Semente de um algoritmo gerador de números aleatórios
u_i	i -ésimo número aleatório $U(0;1)$, sendo $(i=1,2,\dots,I)$
valor-p	Probabilidade de significância de um teste
\mathbf{x}	Vetor de variáveis independentes ou fatores controlados
x_1, x_2, \dots, x_I	Variáveis independentes ou fatores manipulados em um experimento
\mathbf{X}	Matriz de planejamento de um experimento
y	Observação, variável dependente ou resposta
\hat{y}_n	Valor predito para a n -ésima condição experimental
\bar{y}_n	Valor médio para a n -ésima condição experimental
\bar{y}	Média das observações do experimento
$\bar{\mathbf{Y}}$	Vetor que contém as médias de cada condição experimental
$\hat{\mathbf{Y}}_{predito}$	Vetor que contém os valores preditos em cada condição experimental
\mathbf{z}	Vetor de variáveis independentes ou fatores controlados
z	Escore normal padronizado
z_n	n -ésimo número aleatório $N(0;1)$, sendo $(n = 1,2,\dots,N)$

Letras gregas

β	Vetor de parâmetros do modelo matemático do valor esperado
χ^2	Variável aleatória com distribuição Qui-quadrado
δ	Multiplicador um algoritmo gerador de números aleatórios
ε_{nm}	Erro experimental na m -ésima replicação da n -ésima condição experimental
ϕ, ψ, φ	Funções de ligação
γ	Incremento um algoritmo gerador de números aleatórios
η	Estimativa de um efeito
λ	Coefficiente da transformação Box-Cox
μ_y ou $E(y)$	Valor esperado da resposta y
μ_u	Média da distribuição $U(0;1)$
v	Quantidade de parâmetros usados em um modelo
o	Divisor um algoritmo gerador de números aleatórios
θ	Vetor de parâmetros da equação da variância
ρ	Coefficiente um algoritmo gerador de números aleatórios
σ_y	Desvio-padrão da resposta y
σ_y^2	Variância da resposta y
σ_u^2	Variância da distribuição $U(0;1)$
τ	Valor alvo de uma característica de qualidade
*	Operador de interação

SUMÁRIO

Capítulo 1

INTRODUÇÃO	19
1.1 APRESENTAÇÃO DO TEMA.....	19
1.2 O PROBLEMA DA PESQUISA.....	21
1.3 OBJETIVOS.....	23
1.3.1 Objetivo geral.....	23
1.3.2 Objetivos específicos.....	23
1.4 JUSTIFICATIVA.....	23
1.5 IMPORTÂNCIA.....	25
1.6 DELIMITAÇÕES.....	28
1.7 ESTRUTURA DO TRABALHO.....	28

Capítulo 2

REVISÃO DE LITERATURA	30
2.1 EXPERIMENTAÇÃO ESTATÍSTICA: UMA ABORDAGEM CLÁSSICA.....	30
2.1.1 Planejamento de experimentos.....	31
2.1.2 Análise de experimentos.....	36
2.1.3 Superfície de resposta.....	41
2.2 O MÉTODO DE TAGUCHI.....	45
2.2.1 A função perda de Taguchi.....	47
2.2.2 Experimentos sob a ótica de Taguchi.....	49
2.2.3 Considerações sobre o trabalho desenvolvido por Taguchi.....	55
2.3 EFEITOS DE LOCAÇÃO E DISPERSÃO.....	58
2.4 TRANSFORMAÇÕES MATEMÁTICAS.....	62
2.5 EFEITOS DE DISPERSÃO.....	64
2.5.1 Experimentos replicados.....	65
2.5.1.1 Método S.....	66
2.5.1.2 Método R.....	66
2.5.1.3 Considerações sobre os métodos S e R.....	67
2.5.2 Experimentos não-replicados.....	68
2.5.2.1 Método H.....	68
2.5.2.2 Método BM.....	70
2.5.2.3 Método BH.....	72
2.5.2.4 Estratégia BN.....	73

Capítulo 3

METODOLOGIA	76
3.1 CARACTERÍSTICAS DO ESTUDO DE SIMULAÇÃO.....	76
3.2 DESCRIÇÃO DO ESTUDO DE SIMULAÇÃO.....	77
3.2.1 Projeto experimental estudado.....	77
3.2.2 Projeto experimental para a simulação.....	78
3.3 DESCRIÇÃO DO PROGRAMA COMPUTACIONAL.....	80
3.3.1 Primeira etapa: planejamento do experimento.....	80

3.3.2 Segunda etapa: geração de amostras.....	82
3.3.3 Terceira etapa: identificação de fatores com efeitos de dispersão.....	83
3.3.4 Quarta etapa: estimativa das proporções de identificação correta e de identificação falsa.....	87

Capítulo 4

ANÁLISE DOS RESULTADOS DA SIMULAÇÃO.....	90
4.1 DESCRIÇÃO DA ANÁLISE.....	90
4.2 IDENTIFICAÇÃO CORRETA.....	91
4.2.1 Estatística descritiva.....	91
4.2.2 Análise da variância.....	93
4.2.3 Contrastes.....	96
4.3 IDENTIFICAÇÃO FALSA.....	98
4.3.1 Estatística descritiva.....	98
4.3.2 Análise da variância.....	99
4.3.3 Contrastes.....	101
4.4 CONSIDERAÇÕES SOBRE OS RESULTADOS.....	102

Capítulo 5

SIMULAÇÕES COMPLEMENTARES.....	105
5.1 RESÍDUOS QUADRÁTICOS INDIVIDUAIS E MÉDIOS.....	105
5.1.1 Descrição do estudo de simulação.....	107
5.1.2 Resultados obtidos no estudo de simulação.....	110
5.2 VARIÂNCIAS AMOSTRAIS E RESÍDUOS QUADRÁTICOS.....	113
5.2.1 Descrição do estudo de simulação.....	113
5.2.2 Resultados do estudo de simulação.....	116
5.2.2.1 Descrição do estudo.....	116
5.2.2.2 Identificações corretas.....	116
5.2.2.3 Identificações falsas.....	118
5.2.2.4 Considerações sobre os resultados.....	120

Capítulo 6

ESTRATÉGIA PARA IDENTIFICAR EFEITOS DE DISPERSÃO.....	122
6.1 CONSIDERAÇÕES SOBRE OS ESTUDOS DE SIMULAÇÃO.....	122
6.2 A ESTRATÉGIA.....	123
6.3 ESTUDO DE CASO.....	125
6.3.1 O produto avaliado.....	126
6.3.2 Possíveis causas da variabilidade da resistência à compressão do concreto.....	131
6.3.3 O experimento.....	131
6.3.4 Análise dos resultados.....	133
6.3.5 Discussão e considerações sobre os resultados.....	138
6.3.6 Considerações sobre a estratégia utilizada.....	143

Capítulo 7

CONCLUSÕES E RECOMENDAÇÕES.....	144
7.1 CONCLUSÕES.....	144
7.2 FUTURAS PESQUISAS.....	146

REFERÊNCIAS BIBLIOGRÁFICAS.....	147
--	------------

APÊNDICE.....	157
----------------------	------------

Capítulo 1

INTRODUÇÃO

1.1 APRESENTAÇÃO DO TEMA

Existem indícios do homem já valorizar o melhor desde a Antigüidade. Entretanto, somente nas primeiras décadas do século passado, a busca da qualidade começa a ser realizada em processos produtivos. Nos anos 30, surgem as conhecidas cartas de controle desenvolvidas pelo Dr. W. A. Shewhart, da *Bell Laboratories*, fazendo parte do chamado controle de qualidade *on-line*, por serem técnicas de monitoramento de processos. Atribui-se a disseminação de seu uso na indústria, assim como o desenvolvimento de vários métodos estatísticos utilizados com essa finalidade, à segunda guerra mundial. Tais técnicas de monitoramento de processos industriais, acompanhadas do conseqüente desenvolvimento dos métodos estatísticos, são consideradas, por alguns, como um recurso fundamental utilizado por países vitoriosos, como os Estados Unidos e a Inglaterra. Entretanto, o Japão, um dos perdedores, assimilou rapidamente esses conceitos e, com a importante colaboração de Deming, Juran e Ishikawa, conseguiu reverter seus cenários industriais de produtos baratos, mas de má qualidade, iniciando, nos anos 50, a era da qualidade do produto japonês, com reconhecimento mundial.

Desde então, o crescente aumento da competição industrial fez as indústrias buscarem a melhoria da qualidade com a redução de custos. Os cientistas iniciam assim, a pesquisa pelo aprimoramento dos métodos estatísticos por elas empregados. Começam a ser criados, aprimorados e difundidos os métodos de controle de qualidade *off-line*, entre os quais os de Taguchi, que sugerem a utilização de algumas técnicas específicas para planejamento e análise de experimentos.

Nas últimas décadas, utilizar a experimentação estatística na indústria tem sido valorizado, sendo ela empregada em estudos preliminares de programas de qualidade, visando, principalmente, ao conhecimento de um sistema de engenharia, por possibilitar o estudo simultâneo dos efeitos individuais de vários fatores sobre as características funcionais de um produto e os efeitos das interações entre fatores. É considerada muito eficiente, quando

conduzida sob um enfoque clássico, com a incorporação das idéias de Taguchi que relacionam qualidade com variabilidade.

Os sistemas de engenharia apresentam quatro componentes: *entradas*, *fatores de controle*, *fatores de ruído* e *saídas*, convenientemente esquematizados na figura 1.1. As *entradas* correspondem aos insumos alimentadores do processo. Os *fatores de controle* são aqueles que interferem no processo e podem ser manipulados pelo engenheiro, enquanto os *fatores de ruído*, embora interferindo no processo, não podem ou são muito caros ou difíceis de serem controlados na linha de produção. As *saídas*, freqüentemente denominadas de *respostas*, correspondem às características funcionais de um produto, definidoras do seu desempenho.

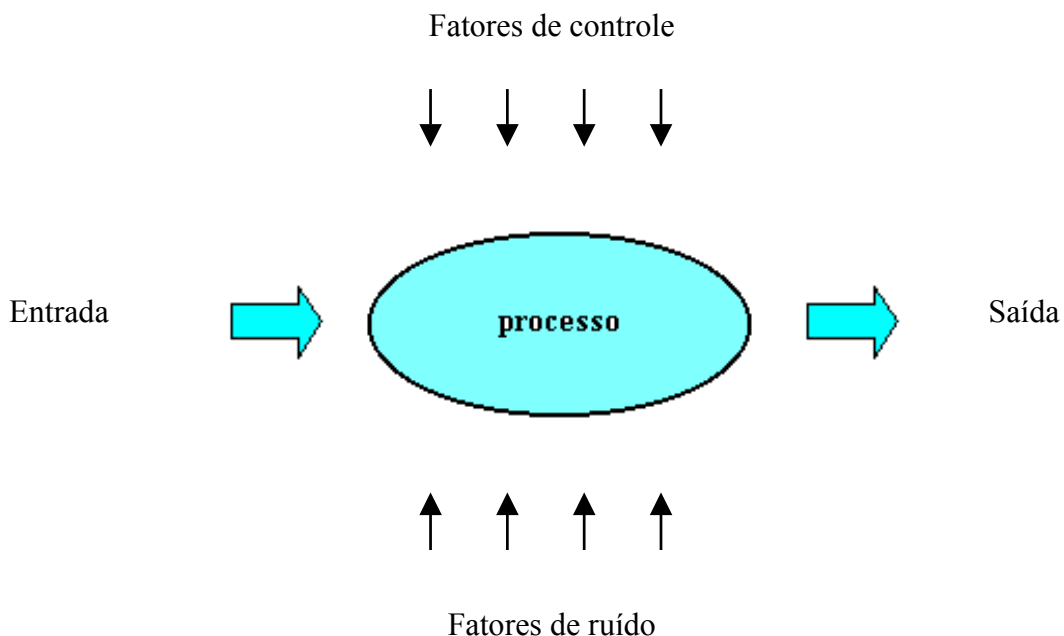


FIGURA 1.1 – Sistema produtivo

De acordo com a filosofia de Taguchi, a utilização de experimentos permite a concentração dos esforços para obter qualidade nas etapas iniciais de desenvolvimento de um projeto. Nelas, denominadas pelo autor de *projeto de parâmetros* e *projeto de tolerâncias*, são definidos, respectivamente, os valores-alvo das características funcionais de um produto ou processo e a tolerância permitida para essas medidas-alvo, enfocadas no projeto. Isto possibilita o “fazer certo na primeira vez” e, certamente, resulta em redução de custos por eliminar perdas e retrabalho.

Segundo Box *et al.* (1986b), uma das causas mais importantes do êxito das empresas japonesas em alcançar alta qualidade com custos compatíveis, são os inúmeros programas experimentais implementados durante a fase de projeto e desenvolvimento de um produto.

1.2 O PROBLEMA DA PESQUISA

A busca pela melhoria da qualidade de produtos e processos é uma tarefa freqüentemente executada pelo engenheiro da qualidade que tem, na experimentação estatística, uma de suas principais ferramentas.

As técnicas clássicas de projeto experimental priorizavam suas ações em fatores controláveis do processo, dando grande importância à média e ao valor-alvo, ou seja, focavam sua análise apenas no valor médio da resposta. Entretanto, o desenvolvimento da Engenharia da Qualidade, alavancado por Taguchi, mostrou que ajustar um valor médio a um valor-alvo é insuficiente para a obtenção dos chamados *sistemas robustos*: sistemas insensíveis a qualquer fonte de variação, pois existe também a necessidade de diminuir a variabilidade em torno do valor-alvo. Dentro desse enfoque, é necessário identificar os fatores que afetam a média (*efeitos de locação*) e os fatores que afetam a variabilidade (*efeitos de dispersão*) da resposta.

Diz-se que um fator afeta a média da resposta e, portanto apresenta *efeito de locação ativo*, quando o valor esperado da resposta for diferente nos níveis estudados desse fator, conforme mostra o gráfico (a) da figura 1.2. Um fator afeta a variabilidade da resposta, apresentando *efeito de dispersão ativo*, quando possui variâncias diferentes nas respostas obtidas nos seus diferentes níveis, como mostra o gráfico (b) da figura 1.2. Um fator pode apresentar, simultaneamente, efeito de locação e de dispersão ativos, caso do fator *C* (gráfico (c) da figura 1.2) e não apresentar efeito ativo, caso do fator *D* (gráfico (d) da mesma figura).

Os efeitos de locação e de dispersão são detectados mais facilmente em *experimentos com replicações* (repetição do ensaio sob as mesmas condições experimentais), nos quais podem ser encontradas, em cada condição experimental, as médias e as variâncias amostrais. Análises de variâncias ou gráficos de probabilidade normal (mais detalhes subseção 2.1.2), aplicados a essas estatísticas permitem identificar efeitos de locação e dispersão.

Entretanto, essas técnicas supõem normalidade e variâncias constantes, o que não ocorre na distribuição de variâncias amostrais. De acordo com Bartlett e Kendall (1946), a utilização do logaritmo neperiano das variâncias amostrais estabiliza as variâncias e contorna o problema da normalidade, se o número de replicações for grande. Esta transformação também é relativamente eficiente quando houver bastante replicações (eficiência superior a

80%, em relação à variância amostral, quando houver, pelo menos, cinco replicações). Surge, então, a seguinte questão: como proceder, quando não for possível ou conveniente utilizar, pelo menos, cinco replicações? Como se comportam os métodos já propostos que utilizam variâncias amostrais para estimar efeitos de dispersão, quando são usadas menos de 5 replicações?

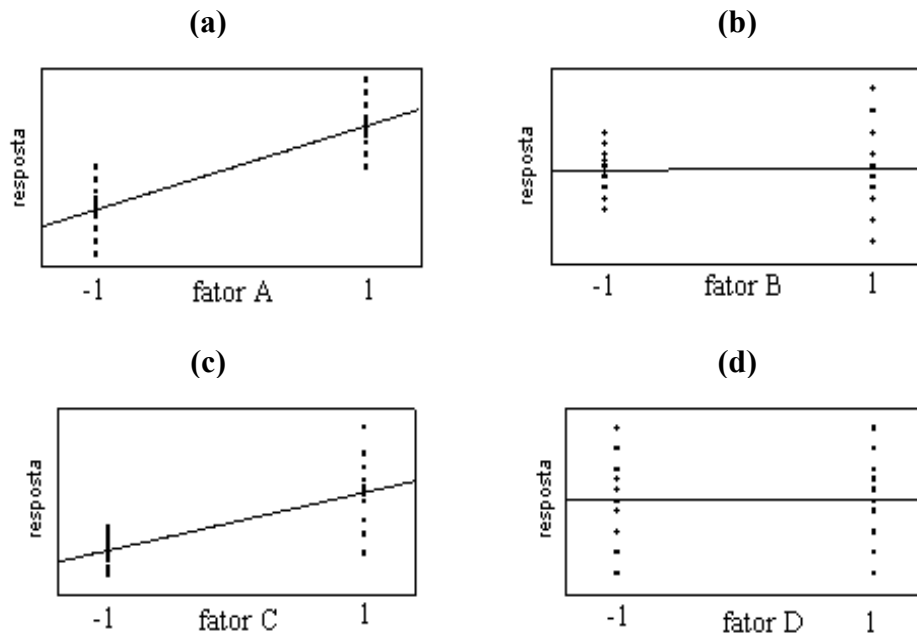


FIGURA 1.2 – Representação gráfica de efeitos de locação e efeitos de dispersão em um experimento fictício (Adaptado de Barbetta *et al.*, 1999)

Alguns métodos já foram desenvolvidos para detectar efeitos de dispersão em experimentos não-replicados. Os métodos mais utilizados baseiam-se em médias aritméticas ou médias geométricas entre resíduos quadráticos. Como se comportam os mesmos métodos, se aplicados à situação de experimentos com poucas replicações? São os resíduos quadráticos melhores do que as variâncias amostrais na estimação de variâncias, quando usadas menos de 5 replicações? Os métodos com aplicação da média aritmética apresentam melhor ou pior desempenho do que aqueles com o uso da média geométrica? Os métodos com a utilização de resíduos quadráticos modificados (BERGMAN e HYNÈN, 1997) se comportam melhor do que aqueles com o emprego de resíduos quadráticos simples?

O presente estudo pretende explorar métodos para identificar efeitos de dispersão em projetos fatoriais dos tipos 2^K ou 2^{K-p} com poucas replicações.

1.3 OBJETIVOS

1.3.1 Objetivo geral

Elaborar uma estratégia capaz de identificar os efeitos de dispersão em experimentos fatoriais do tipo 2^K , com poucas replicações, baseando-se em métodos ou técnicas descritos na literatura.

1.3.2 Objetivos específicos

Visando ao objetivo geral proposto, pretende-se atingir os seguintes objetivos específicos:

- a – selecionar os métodos mais utilizados para identificar efeitos de dispersão, em experimentos replicados e não-replicados;
- b – adaptar os procedimentos descritos na identificação de efeitos de dispersão em experimentos não-replicados para a situação de experimentos replicados;
- c – avaliar, por meio de simulações de Monte Carlo, o desempenho dos principais métodos em experimentos com poucas replicações;
- d – descrever uma estratégia para identificar efeitos de dispersão em experimentos com poucas replicações;
- e – aplicar a estratégia proposta em um experimento.

1.4 JUSTIFICATIVA

Experimentos fatoriais com dois níveis, completos ou fracionados, incluindo vários fatores de diferentes tipos (fatores de controle e fatores de ruído), são muito usados em estudos preliminares de programas de melhoria da qualidade pois, nesta fase, existe grande quantidade de fatores a serem estudados. Aqueles com maior interferência no processo devem ser selecionados para investigação mais detalhada em etapas posteriores.

Na análise dos resultados de um experimento, precisam ser encontrados os fatores que afetam a média e os que afetam a variabilidade. Para otimização do sistema, fatores com efeitos de locação ativos precisam ser ajustados de forma a aproximar ao máximo a resposta média do valor-alvo; fatores com efeitos de dispersão ativos devem ser aferidos de forma a minimizar a variabilidade em torno desse valor-alvo e fatores sem efeitos de locação nem de dispersão devem ser aferidos pelo nível mais econômico.

Para atingir tais objetivos, é necessário identificar os fatores com efeitos de locação ativos e aqueles com efeitos de dispersão ativos. De acordo com Carrol e Ruppert (1988, p. 8), as técnicas para identificar e analisar fatores com efeitos de locação já foram muito discutidas (vários livros existem sobre o assunto) e apresentam uma teoria consagrada, o mesmo não acontecendo com as técnicas utilizadas no estudo de fatores com efeitos de dispersão.

Na análise dos resultados de um *experimento com replicações*, a identificação dos fatores ou interações com efeitos de locação *ativos* pode ser realizada por meio de técnicas estatísticas formais, como o teste *t* de *Student* e o teste *F* da análise de variância. Em experimentos não-replicados costumam ser utilizados métodos menos formais, como o gráfico de Pareto e o gráfico de probabilidade normal (Montgomery, 1997a, p.318).

Identificar fatores com efeitos de dispersão ativos não é uma tarefa simples. Vários métodos, envolvendo gráficos, testes de significância e construção de modelos, além de estratégias iterativas, já foram propostos para identificar a causa da variabilidade em respostas de experimentos. Eles medem a variabilidade empregando variâncias amostrais ou resíduos quadráticos, podendo ser aplicados em experimentos replicados ou não-replicados, não existindo, entretanto, uma teoria consagrada.

Os métodos existentes para experimentos com replicações são comprovadamente eficientes apenas se existirem muitas replicações, razão pela qual se tornam dispendiosos e, em algumas situações, até inviáveis para o setor industrial, que busca, permanentemente, a redução dos custos de manufatura. Essa meta fez a indústria valorizar os experimentos não-replicados que, embora reduzam custos, são superficiais no estudo da variabilidade. Em tal situação, a avaliação da variabilidade é feita pelo *erro residual* (abrange efeitos dos fatores que não entram no modelo da média) e não pelo *erro puro*, razão pela qual o desempenho dos métodos depende muito da qualidade do modelo da média.

Um experimento realizado pela *National Railway Corporation of Japan* para estudar a força de elasticidade de determinado tipo de solda, tornou-se um caso clássico. Trata-se de um projeto experimental fatorial fracionado e não-replicado, com nove fatores ensaiados em dois níveis, combinados em 16 condições experimentais. Foi analisado em Taguchi e Wu (1980), Box e Meyer (1986a), Bergman e Hynén (1997), Pan (1999), Liao (2000) e Brenneman e Nair (2001). Em cada um desses estudos, diferentes métodos foram usados, identificando conjuntos de fatores com efeitos de dispersão ativos podendo conter de zero até três fatores.

A partir desses estudos, é possível constatar que, embora já existam vários métodos para identificar fatores com efeitos de dispersão em experimentos não-replicados, ainda não

existe uma teoria consagrada, pois diferentes métodos apresentam diferentes resultados, o mesmo acontecendo quando são usadas poucas replicações.

Uma estratégia ou método capaz de possibilitar uma correta identificação dos efeitos de dispersão de forma econômica, poderia ser amplamente utilizado na indústria como poderosa ferramenta para a Engenharia da Qualidade, por permitir a obtenção da melhoria da qualidade de produtos e processos produtivos, minimizando perdas.

O presente trabalho fixa-se em experimentos com poucas replicações, muito pouco explorados na literatura consultada, por considerar que esta situação ameniza alguns dos problemas apontados nos métodos já propostos para a identificação de efeitos de dispersão. Há redução de custos em relação aos experimentos com muitas replicações, tendo em vista os métodos já propostos para esta situação empregarem variâncias amostrais, comprovadamente eficientes para estimar variâncias apenas se forem utilizadas pelo menos cinco replicações (BARTLETT e KENDALL, 1946), enquanto, em relação aos experimentos não-replicados, embora a situação não seja tão econômica, é obtida uma análise mais fidedigna da variabilidade.

1.5 IMPORTÂNCIA

No anos 80, foram divulgados, no mundo ocidental, os trabalhos de Taguchi, que redefiniam qualidade como sendo a perda imposta por um produto à sociedade, mensurando-a pelo afastamento das características funcionais de um produto de seus valores-alvo, além do custo.

Até esta época, para otimizar a característica funcional de um produto, os fatores possíveis de serem controlados em seu processo produtivo deveriam ser aferidos de maneira a aproximar ao máximo seus resultados do valor-alvo. A meta de todo fabricante: o afastamento zero do valor-alvo, seria impossível de ser atingida em função da existência da variabilidade, algo inerente à natureza e, portanto, ao processo produtivo.

Não se podem produzir dois produtos exatamente iguais. Mas, se a variabilidade não pode ser eliminada, pode ser conhecida e controlada: é possível reduzi-la de forma a torná-la pequena e despercebida. Reduzir variabilidade tem um importante significado para a melhoria da qualidade. Montgomery (1997a, p.4-5) afirma ser a qualidade inversamente proporcional à variabilidade e a melhoria da qualidade poder ser obtida com a redução da variabilidade de um produto ou processo.

Para a redução dessa variabilidade, é preciso entender o comportamento do sistema produtivo, o que pode ser feito por meio de experimentos estatisticamente planejados, envolvendo *fatores de controle* e *fatores de ruído*. Os níveis dos *fatores de controle* são facilmente definidos e medidos em laboratório ou na linha de produção, sendo responsáveis pelo afastamento do valor esperado das características de qualidade em relação ao seu valor-alvo. Os *fatores de ruído* (variações ambientais, desgaste de peças, ajuste de máquinas etc) são os grandes causadores da variabilidade e, normalmente, não podem ou são muito caros ou difíceis de serem controlados. Entretanto, em muitas situações, pelo menos em laboratório, é possível estudar, por meio de experimentos, a possível relação dos *fatores de ruído* com os fatores possíveis de serem controlados na linha de produção.

Na análise dos resultados de um experimento, são identificados, entre os *fatores de controle* e os *de ruído*, aqueles com efeitos de dispersão, que interferem diretamente na variabilidade e podem reduzi-la, se forem, pelo menos indiretamente, convenientemente aferidos.

Em 1997, durante a implantação de um controle estatístico de processo em uma olaria de pequeno porte, foram utilizados alguns experimentos extremamente simples para caracterização do funcionamento de um forno utilizado na queima do produto, sendo possível constatar que sua temperatura variava de acordo com o operário responsável por sua alimentação e com o horário de funcionamento (MATTOS, 1997, p. 78-79). Nesse estudo, as interferências foram consideradas inerentes ao processo produtivo, pois o equipamento tinha de funcionar ininterruptamente dia e noite, necessitando de vários operadores, para quem apenas algumas instruções foram repassadas. Embora esses fatores não pudessem ter seus efeitos controlados na linha de produção, poderiam ter sido analisadas as possíveis interações, por meio de experimentos, com os fatores possíveis de serem controlados, de forma a minimizar seus efeitos.

Na construção civil, quando se elabora uma edificação ou uma obra de arte, são utilizados vários materiais e milhares de componentes a serem moldados, encaixados ou combinados. Por isto, muitas vezes, reduzir a variabilidade, é tão ou mais importante do que aproximar o valor médio do valor-alvo da característica focada.

Na região sul do Brasil, a maior parte das edificações é construída com paredes de alvenaria. Existe uma norma, a NBR 7171, que regulamenta as dimensões de blocos cerâmicos utilizados com esta finalidade, entre outras características de qualidade. Para blocos cerâmicos de 6 furos, por exemplo, a norma recomenda para cada peça, a medida de 9cm x 14cm x 19cm, com uma tolerância de ± 3 mm (MATTOS, 1997, p.23). Entretanto, a execução

de uma parede em alvenaria, praticamente, não ficará prejudicada, se todos os blocos apresentarem as dimensões, por exemplo, de 8cm x 12cm x 20cm, diferentes das definidas pela norma. Porém, se existir uma grande variabilidade nas dimensões, não só a execução da parede propriamente dita fica prejudicada, como também a de seu revestimento. Realmente, a irregularidade nas dimensões, no sentido longitudinal e vertical, não tem muitas conseqüências, pois a espessura da junta de assentamento recomendada é bastante larga, aproximadamente 1cm, e incorporaria as irregularidades existentes nas peças. Mas, no sentido da espessura da parede, podem ocorrer problemas na marcação do prumo, cuja possível conseqüência seria a necessidade de revestimento de maior espessura, além de dificultar a execução da alvenaria.

A importância do controle e diminuição da variabilidade também se evidencia, na hora do revestimento de uma parede com produtos cerâmicos, por exemplo. Se as dimensões médias de certo tipo de azulejo forem alteradas, poucas conseqüências serão observadas na qualidade do revestimento. Entretanto, se a variabilidade das dimensões for elevada, o engenheiro certamente enfrentará sérios problemas de estética, durabilidade e trabalhabilidade.

Um engenheiro civil também pode enfrentar problemas graves se uma variabilidade excessiva ocorrer em produtos com finalidade estrutural, como o caso do concreto, que apresenta a homogeneidade como uma de suas principais características de qualidade.

A indústria farmacêutica é um outro setor da economia no qual o controle da variabilidade é extremamente importante, pois os diversos produtos são elaborados a partir de fórmulas com vários componentes, em diferentes graus de concentração ou dosagens. O não-cumprimento das especificações pode comprometer, total ou parcialmente, o desempenho do produto e, em algumas situações, estimular o aparecimento de efeitos colaterais.

Dentro desse novo enfoque, a principal característica de qualidade de um sistema produtivo passa a ser sua robustez. Um sistema produtivo é dito robusto quando consegue ser pouco sensível a fontes de variação, apresentando uma variabilidade muito pequena, muitas vezes não percebida visualmente.

Segundo Tsui (1992), a Engenharia Robusta é uma importante metodologia para melhorar a qualidade, trabalhabilidade e confiabilidade de produtos, a baixo custo. Certamente esta metodologia pode fornecer vantagem competitiva a uma empresa para enfrentar os concorrentes, que não agem como se o bolo fosse grande o bastante para todos.

1.6 DELIMITAÇÕES

Os estudos de simulação realizados restringem-se a algumas condições experimentais, convenientemente definidas no referente à quantidade e à intensidade de fatores com efeitos de locação; à quantidade e à intensidade de fatores com efeitos de dispersão; nível de coincidência de efeitos de locação e dispersão em um mesmo fator e quantidade de replicações, sendo os diferentes parâmetros de simulação organizados segundo um projeto de experimentos.

Há de considerar, também, que a avaliação do desempenho dos métodos para identificar efeitos de dispersão é realizada apenas em projetos fatoriais completos, de três e quatro fatores, ensaiados em dois níveis, podendo, entretanto, ser equivalentes a outros tipos de projetos fatoriais fracionados.

As respostas dos experimentos, representativas das características funcionais de um produto, que definem a sua qualidade ou os parâmetros de um processo produtivo, são consideradas como variáveis aleatórias contínuas com distribuição normal, conforme o usual na literatura sobre o assunto.

O objetivo geral do trabalho é elaborar uma estratégia capaz de identificar os efeitos de dispersão em experimentos fatoriais com poucas replicações. Isto é feito pela avaliação do desempenho entre métodos já propostos, adaptados ou não para a situação de experimentos com poucas replicações. Na comparação, alguns critérios são simplesmente adotados, pois considera-se que a adoção do mesmo critério para avaliar o desempenho de todos os métodos é suficiente. A mensuração do desempenho de determinado método, por exemplo, é feita pela proporção de identificação correta de *todos* os efeitos de dispersão (*PIC*) e pela proporção de identificação falsa de *algum* efeito de dispersão (*PIF*).

1.7 ESTRUTURA DO TRABALHO

O trabalho foi estruturado em sete capítulos, a saber:

Capítulo 1 – apresenta o trabalho desenvolvido, justificando o tema escolhido, além de mostrar sua importância, a definição do problema e dos objetivos atingidos, as delimitações do trabalho e sua estruturação;

Capítulo 2 – trata da revisão de literatura. São feitas algumas considerações gerais e básicas sobre projeto de experimentos, metodologia de superfície de resposta, filosofia de Taguchi e métodos utilizados para a identificação de efeitos de dispersão em experimentos;

Capítulo 3 – descreve o planejamento de um estudo de simulação cujo objetivo é avaliar o desempenho dos métodos selecionados para identificar efeitos de dispersão em experimentos com poucas replicações;

Capítulo 4 – apresenta e discute os resultados encontrados no estudo de simulação descrito no capítulo anterior;

Capítulo 5 – mostra um estudo de simulação realizado para adaptar os métodos propostos para experimentos não-replicados à situação de experimentos com replicações e um outro estudo de simulação realizado para comparar métodos que usam diferentes medidas para estimar a variabilidade;

Capítulo 6 – com base nos resultados apresentados nos capítulos anteriores, elabora uma estratégia para identificar efeitos de dispersão em experimentos com poucas replicações. Também apresenta uma aplicação da estratégia proposta em um estudo de caso real;

Capítulo 7 – finalizando o trabalho, manifesta as conclusões originadas do desenvolvimento desta tese, fazendo algumas recomendações para futuros trabalhos relacionados ao tema.

Complementando a estruturação, é utilizado um apêndice para apresentar as respostas dos experimentos analisados.

Capítulo 2

REVISÃO DE LITERATURA

Este capítulo inicia com a apresentação de algumas idéias gerais e básicas de projeto de experimentos, priorizando o enfoque clássico e evidenciando sua importância para a melhoria da qualidade de produtos e processos produtivos. A seguir, mostra a importância e a utilidade das idéias de Taguchi, comentando algumas técnicas por ele criadas para implementar sua metodologia. Após, trata da importância do estudo da variabilidade, apresentando a nova abordagem para projeto de experimentos que incorpora as idéias de Taguchi. Finalmente, apresenta os métodos mais utilizados para identificar os efeitos de dispersão em experimentos replicados e não-replicados.

2.1 EXPERIMENTAÇÃO ESTATÍSTICA: UMA ABORDAGEM CLÁSSICA

De acordo com Montgomery (1997a, p.17), o responsável pela utilização da estatística em projeto experimental foi Ronald A. Fisher, desenvolvendo a técnica de análise de variância, com primeiras aplicações nas áreas da agricultura e da biologia. Na área industrial, as primeiras aplicações ocorreram nos anos 30 do século XX, mas, somente depois da segunda guerra mundial, este setor começou a utilizar essas técnicas com mais frequência, especialmente na indústria química e na indústria eletrônica.

Nas últimas décadas, em função dos cenários altamente competitivos encontrados pelas indústrias, a experimentação estatística começou a ser valorizada no setor, sendo muito utilizada em estudos preliminares de programas de controle de qualidade e considerada, por alguns (BOX, 1988; BERGMAN e HYNÉN, 1997), como fator-chave para obter a melhoria da qualidade.

De acordo com Montgomery (1997a, p.1), um experimento é um teste ou uma série de testes nos quais as variáveis de entrada de um sistema são manipuladas para serem identificadas as razões das mudanças nas variáveis de saída.

Na indústria, o projeto experimental permite manipular os fatores interferidores em um processo produtivo, fazendo-os variar para avaliar seus efeitos nas características funcionais

de um produto, as quais definem sua qualidade, sendo, segundo Montgomery (1997a, p.9-11), especialmente indicado para:

- a) caracterizar um processo, identificando, entre os fatores investigados, os que afetam a resposta do experimento;
- b) otimizar um processo, encontrando a combinação ótima entre os níveis dos fatores investigados que fornecem um melhor desempenho;
- c) melhorar a capacidade (uniformidade) de um processo, determinando tolerâncias para o sistema e seus componentes.

Na abordagem clássica, o estudo é realizado a partir dos fatores do processo capazes de fazer o valor esperado da variável resposta afastar-se de seu valor-alvo.

2.1.1 Planejamento de experimentos

Existem várias estratégias para conduzir um experimento. Independentemente da estratégia selecionada, é indispensável estarem perfeitamente definidas: a *unidade experimental* (elemento a ser medido ou observado no experimento); a variável analisada ou *resposta* (o que é medido na unidade experimental); como será mensurada e a região experimental investigada, o que pode ser feito pela definição dos *fatores* a serem manipulados e seus diferentes *níveis* adotados.

Na área industrial, entre os projetos mais utilizados está o fatorial. Nesse tipo de projeto, os diversos fatores investigados variam, gerando várias condições experimentais, resultantes das diferentes combinações possíveis entre seus níveis.

Dentre os planejamentos fatoriais, o mais utilizado na área industrial, talvez por necessitar de uma menor quantidade de ensaios, é o do tipo 2^K , que também costuma servir de base para outros tipos de planejamento mais complexos. Um projeto desse tipo envolve K fatores ensaiados em apenas dois níveis, codificados como: -1 (o nível mais baixo ou inferior) e $+1$ (o nível mais alto ou superior), sendo especialmente indicado quando se supõe uma relação linear entre os fatores considerados e a resposta.

Suponha estar sendo planejado um experimento, com três fatores (A , B e C), ensaiados em dois níveis (-1 e $+1$). Num projeto fatorial, os fatores serão *cruzados*, gerando $2^3=8$ condições experimentais. As combinações possíveis entre os níveis dos três fatores são apresentadas nas colunas (3), (4) e (5) na tabela 2.1.

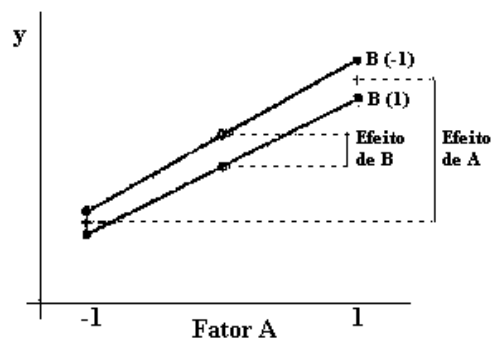
TABELA 2.1 – Níveis dos fatores e interações de um projeto fatorial do tipo 2^3

(1) Condição experimental	(2) <i>I</i>	(3) <i>A</i>	(4) <i>B</i>	(5) <i>C</i>	(6) <i>AB</i>	(7) <i>AC</i>	(8) <i>BC</i>	(9) <i>ABC</i>
1	+1	-1	-1	-1	+1	+1	+1	-1
2	+1	-1	-1	+1	+1	-1	-1	+1
3	+1	-1	+1	-1	-1	+1	-1	+1
4	+1	-1	+1	+1	-1	-1	+1	-1
5	+1	+1	-1	-1	-1	-1	+1	+1
6	+1	+1	-1	+1	-1	+1	-1	-1
7	+1	+1	+1	-1	+1	-1	-1	-1
8	+1	+1	+1	+1	+1	+1	+1	+1

Por meio desse experimento, é possível avaliar o efeito isolado de cada um dos fatores investigados, denominado *efeito principal*, pela diferença produzida na resposta quando o fator muda de um nível para outro.

Também é possível avaliar, e esta é uma das grandes vantagens da utilização do projeto fatorial, o efeito das *interações* entre os fatores. Diz-se existir interação entre fatores quando a diferença na resposta entre os níveis de um fator não é a mesma segundo os níveis de outro ou outros fatores. As colunas (6), (7), (8) e (9) da tabela 2.1 apresentam o sinal a ser adotado no cálculo das interações no projeto fatorial citado anteriormente.

A figura 2.1 mostra a representação gráfica do efeito principal de dois fatores: *A* e *B*, ensaiados em dois níveis. É possível constatar que, embora exista mudança no valor esperado da resposta (*y*) para os níveis de ambos fatores, esta diferença é mais acentuada entre os níveis do fator *A*.

**FIGURA 2.1** – Representação gráfica de efeitos de fatores em experimento fictício (Adaptada de ZAR, 1999, p.240)

Os gráficos (a) e (b) da figura 2.2 descrevem as variações provocadas por dois fatores, também identificados por *A* e *B*, ensaiados em dois níveis. O gráfico (a) mostra não existir interação entre os fatores, pois a variação no valor esperado da resposta nos diferentes níveis do fator *B* é, aproximadamente, a mesma para os diferentes níveis do fator *A*. Já o gráfico (b)

evidencia a existência de interação entre os fatores, pois o valor esperado da resposta no nível inferior do fator A é maior quando o fator B também se encontra em seu nível inferior, o mesmo não acontecendo quando o fator A se encontra em seu nível superior.

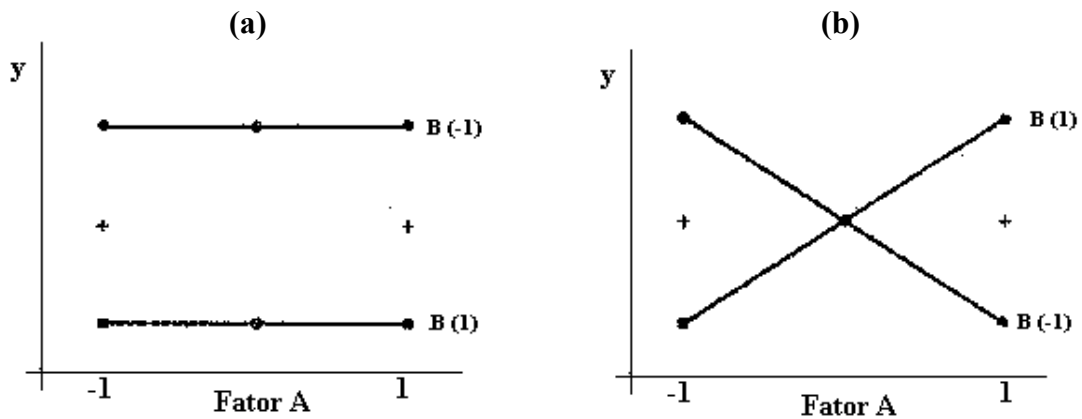


FIGURA 2.2 – Representação gráfica do efeito de fatores sem interação e com interação em experimento fictício (Adaptadas de ZAR, 1999, p.241)

Se a quantidade de fatores investigados no experimento for elevada, mesmo os projetos 2^K apresentam o inconveniente de necessitar de uma quantidade muito grande de ensaios. Por exemplo, se um experimento fatorial considerar cinco fatores ensaiados em apenas dois níveis, o projeto fatorial apresentará $2^5 = 32$ condições experimentais, necessitando de, no mínimo, 32 ensaios.

Em tais casos, ou seja, quando existe uma quantidade relativamente grande de fatores, os projetos fatoriais 2^K podem ser fracionados. A idéia desse tipo de projeto é ensaiar apenas parte das possíveis combinações de níveis, sendo planejado de forma a garantir a possibilidade de se estimarem os efeitos principais e, às vezes, os efeitos das interações entre dois fatores. Esse tipo de projeto é particularmente importante quando se pretende fazer uma *triagem* de fatores para serem usados em estudos posteriores.

Os projetos fracionados, identificados como sendo do tipo 2^{K-p} para fatores ensaiados em dois níveis, utilizam K fatores e p fracionamentos. Por exemplo, se um experimento fatorial fracionado considerar cinco fatores ensaiados em dois níveis, e apresentar grau de fracionamento 2, o projeto fatorial necessitará de, no mínimo, $2^{5-2} = 8$ ensaios. O fracionamento de grau 2 reduz a quantidade mínima de ensaios de 32 para 8, vantajoso em termos de custo, especialmente quando se trata de ensaios demorados, caros ou destrutivos.

Para a construção desse tipo de projeto, são selecionados $K-p$ fatores, cruzados como num projeto fatorial 2^T , sendo $T = K-p$, gerando 2^T condições experimentais. Os p fatores

restantes terão seus níveis distribuídos de acordo com os sinais das interações de mais alto grau dos $K-p$ fatores. O procedimento propicia, entretanto, o aparecimento de *confundimentos* entre os efeitos.

Um projeto 2^{3-1} é apresentado na tabela 2.2. Ele é construído a partir de um projeto fatorial do tipo 2^2 , envolvendo os fatores A e B (4 condições experimentais). Os níveis do fator C são alocados segundo o sinal da interação AB , resultando em confundimento do fator C com a interação AB ; do fator A com a interação BC e do fator B com a interação AC (colunas com a mesma seqüência de sinais). Conseqüentemente, um projeto 2^{3-1} só deve ser usado nas situações em que se supõe não haver interações ou quando os efeitos das interações possam ser negligenciados.

TABELA 2.2 - Projeto fatorial fracionado do tipo 2^{3-1}

Condição experimental	I	A	B	C	AB	AC	BC	ABC
1	+1	-1	-1	+1	+1	-1	-1	+1
2	+1	-1	+1	-1	-1	+1	-1	+1
3	+1	+1	-1	-1	-1	-1	+1	+1
4	+1	+1	+1	+1	+1	+1	+1	+1

Quando um projeto fatorial fracionado é delineado dessa forma, diz-se apresentar *Resolução III*. No caso, efeitos principais podem ser confundidos com efeitos de interações de segunda ordem, como ocorreu no exemplo citado anteriormente, assim como efeitos de interações de segunda ordem podem ser confundidos entre si.

Quando existir apenas a possibilidade de confundimento entre efeitos de interações de segunda ordem, o projeto apresenta *Resolução IV*. Se os efeitos principais e os de interações de segunda ordem puderem ser confundidos apenas com efeitos de interações de terceira ordem ou superior, o projeto é dito de *Resolução V*.

Outros tipos de projeto também podem ser utilizados na área industrial, embora os projetos fatoriais com dois níveis, completos ou fracionados, sejam os mais aplicados na área, principalmente em estudos preliminares. Os outros tipos de projetos de experimentos mais citados são:

- a) o projeto fatorial do tipo 3^K , envolvendo K fatores ensaiados em três níveis, usualmente codificados como: -1 , 0 e $+1$, que correspondem aos níveis baixo, intermediário e alto, respectivamente. Este tipo de projeto pode detectar efeitos não-lineares e também pode ser fracionado;

- b) o projeto fatorial denominado misto, considerando dois níveis para alguns de seus fatores e três níveis para outros;
- c) o projeto central composto, originando-se de um projeto fatorial do tipo 2^K , completo ou fracionado, com a adição de pontos centrais e pontos axiais. Num projeto do tipo 2^K ou do tipo 2^{K-p} , supõe-se uma relação linear entre fatores e observações, muitas vezes não sendo o mais adequado. Para ajustar uma equação de grau mais elevado, são necessários mais pontos na região experimental, obtidos pela adição de pontos centrais e pontos axiais (figura 2.3). A parte fatorial, fracionada ou completa, serve, nesse caso, para um estudo preliminar, no qual pode ser ajustado um modelo linear para, posteriormente, ser verificada a necessidade de incorporação de termos de ordem mais elevada. Quando pontos são adicionados ao projeto experimental, considera-se que os K fatores são quantitativos.

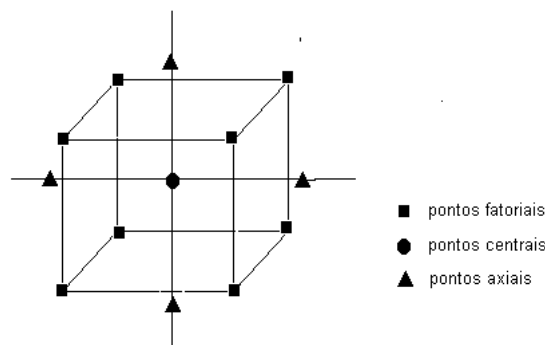


FIGURA 2.3 – Projeto central composto (Adaptado de CATEN, 1995, p.64)

O planejamento do tipo hierárquico também pode ser bastante útil na área industrial quando, por exemplo, comparam-se produtos de diferentes fornecedores, que provêm de lotes diferenciados. De acordo com Zar (1999, p.303), tal tipo de projeto é indicado quando os níveis de um fator são similares, mas não idênticos, nos vários níveis de um outro fator. Também pode ser utilizado, com relativa frequência, o planejamento com dados longitudinais, para avaliar, por exemplo, a velocidade da aquisição de determinadas características físicas e químicas.

Ao planejar um experimento, devem ser também definidos critérios a respeito de aleatoriedade e utilização, ou não, de replicações e de blocos.

Por replicação entende-se a repetição do ensaio, pelo menos mais uma vez, para cada condição experimental considerada no experimento. Por exemplo, se um projeto fatorial considerar 5 fatores ensaiados em dois níveis com duas replicações, necessitará de $2(2^5) = 64$

ensaios. As replicações permitem não apenas fazer estimativas do erro experimental, como também realizar estimativas mais precisas dos valores médios. Segundo Vieira (1999, p.151), do ponto de vista estatístico, é sempre desejável os experimentos terem um grande número de replicações, pois isto aumenta a confiança no resultado.

Por bloco, entende-se uma parcela dos ensaios realizados sob condições mais homogêneas que os demais. Seu emprego tem o objetivo de aumentar a precisão da análise dos resultados do experimento. Um experimento, por exemplo, realizado para avaliar a interferência de determinadas características de qualidade na resistência à compressão de tijolos, utilizando tijolos oriundos de diferentes olarias, pode ser planejado com blocos. Se cada conjunto de tijolos, proveniente de cada uma das olarias consideradas, for considerado como um bloco, é eliminada uma das causas de variabilidade nos resultados.

Já o princípio da aleatoriedade deve ser utilizado na escolha do material experimental e na determinação da ordem de realização dos ensaios para garantir a obtenção de respostas que sejam variáveis aleatórias com distribuições independentes. O procedimento ameniza a probabilidade de ocorrerem valores extremos, além de garantir que alguns fatores não-controlados tenham a mesma probabilidade de interferir nos resultados de diferentes condições experimentais. O princípio da aleatoriedade também pode ser considerado na determinação dos níveis dos fatores utilizados no projeto fatorial, os quais podem ser fixos ou aleatórios.

2.1.2 Análise de experimentos

Várias técnicas estatísticas, tanto gráficas como analíticas, podem ser utilizadas para analisar dados de um experimento. As técnicas formais exigem que os erros sejam variáveis aleatórias independentes, com distribuição normal de valor esperado nulo e variância constante.

É aconselhável a realização de uma análise exploratória dos dados para, como sondagem, avaliar características técnicas, por exemplo: ocorrência de *outliers*. Medidas descritivas de tendência central como a média, e de variabilidade, como o desvio-padrão, podem também ser utilizadas e, do mesmo modo, medidas de distribuição, como os coeficientes de assimetria e de curtose. Nessa fase, são muito úteis o gráfico de probabilidade normal para avaliação da normalidade e o *boxplot* para avaliar a assimetria e a ocorrência de *outliers*. Esse tipo de análise pode ser realizado diretamente nos resíduos.

Após, procuram-se os fatores e interações que interferem de maneira mais acentuada no valor médio das observações. A análise dos dados de um experimento inicia-se, então, com a identificação dos efeitos ativos, ou seja, aqueles fatores com diferentes valores esperados para a resposta em seus diferentes níveis.

Em experimentos com replicações, a identificação desses efeitos pode ser realizada por técnicas estatísticas formais como o teste F da análise da variância ou o teste t de *Student*, o último no caso de dois grupos.

A estatística F tem seu cálculo baseado em estimativas da variância, feitas para cada fator investigado, para cada interação possível, assim como para o erro experimental. O Quadro 2.1, adaptado de Zar (1999, p.242), mostra as várias etapas do cálculo da estatística F , para avaliar a significância dos efeitos de fatores e interações, em um projeto fatorial com dois fatores.

São encontradas as somas dos quadrados para todas as observações (SQ_{Total}), para as condições experimentais (SQ_{CE}) e para cada fator (SQ_A e SQ_B) (expressões 2.1, 2.2, 2.3 e 2.4).

$$SQ_{Total} = \sum_{a=1}^{A'} \sum_{b=1}^{B'} \sum_{m=1}^M y_{abm}^2 - \frac{\left(\sum_{a=1}^{A'} \sum_{b=1}^{B'} \sum_{m=1}^M y_{abm} \right)^2}{A' B' M} \quad (2.1)$$

$$SQ_{CE} = \frac{\sum_{a=1}^{A'} \sum_{b=1}^{B'} \left(\sum_{m=1}^M y_{abm} \right)^2}{M} - \frac{\left(\sum_{a=1}^{A'} \sum_{b=1}^{B'} \sum_{m=1}^M y_{abm} \right)^2}{A' B' M} \quad (2.2)$$

$$SQ_A = \frac{\sum_{a=1}^{A'} \left(\sum_{b=1}^{B'} \sum_{m=1}^M y_{abm} \right)^2}{B' M} - \frac{\left(\sum_{a=1}^{A'} \sum_{b=1}^{B'} \sum_{m=1}^M y_{abm} \right)^2}{A' B' M} \quad (2.3)$$

$$SQ_B = \frac{\sum_{b=1}^{B'} \left(\sum_{a=1}^{A'} \sum_{m=1}^M y_{abm} \right)^2}{A' M} - \frac{\left(\sum_{a=1}^{A'} \sum_{b=1}^{B'} \sum_{m=1}^M y_{abm} \right)^2}{A' B' M} \quad (2.4)$$

em que:

- A e B são fatores investigados no experimento;
- a identifica o a -ésimo nível do fator A ($a=1,2,\dots,A'$);
- b identifica o b -ésimo nível do fator B ($b=1,2,\dots,B'$);

- m identifica a m -ésima replicação ($m=1,2,\dots,M$);
- y_{abm} é a observação na m -ésima replicação do a -ésimo nível do fator A e b -ésimo nível do fator B ;
- CE representa uma condição experimental.

Os quocientes entre estes valores e os respectivos graus de liberdade (gl), mostrados na coluna (4) do Quadro 2.1, fornecem estimativas da variância (QM) devida a cada fonte de variação.

QUADRO 2.1 – Desenvolvimento de uma ANOVA para avaliação da significância de fatores e interações em um projeto fatorial com dois fatores

(1) Fontes de variação	(2) Soma de quadrados (SQ)	(3) Graus de liberdade (gl)	(4) Quadrado médio (QM)	(5) F
Total	SQ_{Total}	$A' B' M - 1$		
Condição experimental	SQ_{CE}	$A' B' - 1$		
A	SQ_A	$A' - 1$	$\frac{SQ_A}{gl_A}$	$\frac{QM_A}{QM_{erro}}$
B	SQ_B	$B' - 1$	$\frac{SQ_B}{gl_B}$	$\frac{QM_B}{QM_{erro}}$
AB	$SQ_{CE} - SQ_A - SQ_B$	$(A' - 1)(B' - 1)$	$\frac{SQ_{AB}}{gl_{AB}}$	$\frac{QM_{AB}}{QM_{erro}}$
Erro	$SQ_{total} - SQ_{CE}$	$A' B' (M - 1)$	$\frac{SQ_{erro}}{gl_{erro}}$	

Fonte: Adaptada de Zar (1999, p.242).

Nota: A e B são os fatores de controle e AB , sua interação. A' e B' são a quantidade de níveis dos fatores A e B , respectivamente, e M , a quantidade de replicações. F é a estatística da distribuição de Snedecor-Fisher.

O quociente entre a estimativa de cada variância encontrada e a estimativa da variância do erro fornecerá a estatística F (coluna (5)), que sob a suposição de normalidade e homocedasticidade, tem distribuição de Snedecor-Fisher com graus de liberdade para o numerador e denominador indicados no Quadro 2.1. Por meio desta distribuição, é possível encontrar o *valor-p*. Usualmente, os fatores e interações considerados com efeitos significativos são aqueles com *valor-p* inferior a 0,05.

Resumindo, a estatística F utilizada na análise da variância é calculada por:

$$F = \frac{SQ_{Fonte\ de\ variac\~ao} / gl_{Fonte\ de\ variac\~ao}}{SQ_{Erro} / gl_{Erro}} \quad (2.5)$$

em que SQ representa soma de quadrados e gl , graus de liberdade.

Em experimentos não-replicados, quando não se conhece a verdadeira variância, é comum utilizar métodos menos formais, como o gráfico de Pareto e o gráfico de probabilidade normal, pois não existe uma estimativa confiável da variância do erro experimental. Nesses processos gráficos, são identificados os efeitos aparentemente significativos.

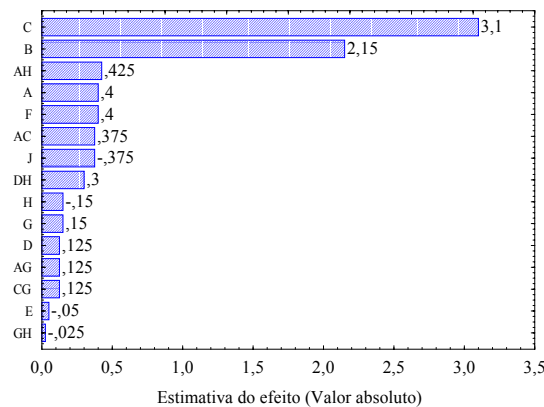


FIGURA 2.4 – Diagrama de Pareto para um exemplo fictício

O diagrama de Pareto é um gráfico em barras horizontais, dispostas em ordem crescente ou decrescente de grandeza. Na situação estudada, cada barra representa um fator ou uma interação, sendo seu comprimento proporcional à intensidade da estimativa do efeito do fator ou da interação que representa. Os fatores ou interações, representados por barras com comprimento diferenciado das demais (maior), são considerados como tendo efeito ativo. A figura 2.4 mostra um diagrama de Pareto, onde é possível identificar os efeitos do fator B e do fator C como ativos.

O gráfico de probabilidade normal é construído num sistema de coordenadas cartesianas, nos quais são plotados pontos que têm, como coordenadas, uma estimativa do efeito (x_k ; $k = 1, 2, \dots, K'$) e seu valor teórico esperado (z_k ; $k = 1, 2, \dots, K'$), de acordo com uma distribuição normal padronizada $[N(0;1)]$. Para traçá-lo, as K' estimativas dos efeitos (x_k) são ordenadas. Após, encontram-se os escores padronizados (z_k), correspondentes aos pontos medianos de K' intervalos consecutivos, definidos a partir da divisão da área sob uma curva $N(0;1)$ em K' partes iguais. Se os efeitos reais forem nulos, haverá uma relação

aproximadamente linear entre suas estimativas: x_k , e os escores padronizados esperados: z_k . A figura 2.5 mostra um gráfico de probabilidade normal, construído a partir dos mesmos dados utilizados na construção do diagrama de Pareto, no qual é evidenciada a existência de dois efeitos de dispersão: fator B e fator C .

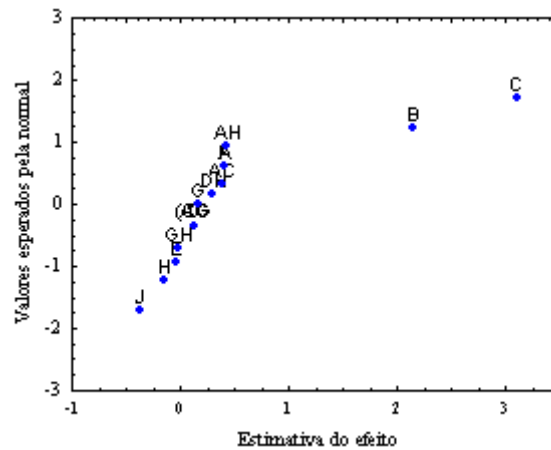


FIGURA 2.5 – Gráfico normal dos efeitos para em exemplo fictício

Se o objetivo da análise é simplesmente identificar os fatores que afastam o valor esperado da resposta do valor-alvo, a análise se encerra. Se for a otimização de um processo, busca-se a combinação entre os níveis de fatores que otimizam a resposta. Os gráficos dos efeitos são especialmente úteis para esta finalidade, pois evidenciam as respostas médias nos diferentes níveis de um fator ou interação. A figura 2.6 mostra um gráfico dessa natureza, em que é possível verificar uma resposta média maior no nível identificado por 1 . Se a característica observada tem melhor desempenho para valores maiores, este seria o nível indicado para otimização.

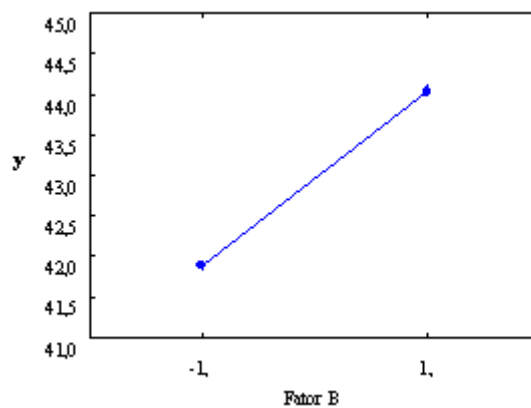


FIGURA 2.6 – Gráfico do efeito do fator B para um exemplo fictício

2.1.3 Superfície de resposta

Em muitas situações, para encontrar a solução ótima de um experimento, é útil a técnica de análise de regressão, pois nem sempre a solução ótima corresponde a um dos níveis pré-estabelecidos no planejamento, principalmente se os fatores usados no experimento do processo forem quantitativos.

A análise de regressão é utilizada, então, para descrever o possível relacionamento entre variáveis, ou seja, para descobrir se há uma relação entre as variáveis de entrada ou independentes e a variável de saída ou dependente, que possa ser descrita por uma equação matemática.

Matematicamente, uma equação de regressão pode ser descrita por:

$$y = \phi(x_1, x_2, \dots, x_l), \quad (2.6)$$

em que y é a variável dependente (variável resposta do experimento); x_1, x_2, \dots, x_l , as variáveis independentes (fatores do experimento) e ϕ , a função de ligação. Como é impossível controlar todas as variáveis independentes que podem influenciar a variável dependente e, até mesmo, a possível ocorrência de erro na medição dessas variáveis sob controle, a Estatística reescreve o modelo como:

$$y = \psi(x_1, x_2, \dots, x_K) + e. \quad (2.7)$$

sendo e o erro experimental. Para a n -ésima observação, tem-se:

$$y_n = \psi(x_{1n}, x_{2n}, \dots, x_{Kn}) + e_n, \quad (2.8)$$

supondo-se que e_n é aproximadamente $N(0; \sigma^2)$. A suposição de os erros apresentarem distribuição normal é necessária para poder ser utilizada a distribuição F para o teste das hipóteses a respeito dos parâmetros do modelo.

Num estudo experimental para aplicar a técnica de regressão, as variáveis independentes, x_1, x_2, \dots, x_K , são os fatores controlados no experimento e a variável dependente, y , sua resposta. Determinar a equação de regressão, a partir dos resultados do experimento, permite prever o valor esperado da resposta, possibilitando encontrar a combinação ótima entre os fatores, a qual faça o valor esperado afastar-se o mínimo possível do valor-alvo.

Portanto, a equação de regressão, também denominada superfície de resposta, pode ser descrita por:

$$E(y) = f(\mathbf{x}, \boldsymbol{\beta}), \quad (2.9)$$

onde $\mathbf{x} \in D \subset \mathfrak{R}^k$, sendo:

- $E(y)$ o valor esperado para a resposta y ;
- f uma função de ligação;
- \mathbf{x} o vetor de variáveis independentes ou fatores controlados;
- $\boldsymbol{\beta}$ o vetor de parâmetros do modelo matemático;
- $D \subset \mathfrak{R}^k$ a região experimental.

Em projetos experimentais do tipo 2^K ou 2^{K-p} , a função f , normalmente utilizada, é a linear. Para identificar-se uma determinada superfície linear, é necessário obter estimativas para os elementos do vetor $\boldsymbol{\beta}$, o que pode ser feito pelo método dos mínimos quadrados. Ele tem como princípio de que a melhor superfície que se ajusta a um conjunto de dados é aquela com a menor soma dos quadrados dos desvios de cada valor em relação ao correspondente valor predito pelo modelo.

Para a realização de inferências (intervalos de confiança e testes estatísticos) sobre parâmetros do modelo, são necessárias algumas suposições sobre a variável resposta:

- (a) para qualquer conjunto de valores \mathbf{x} existe uma população de valores de y com distribuição normal;
- (b) existe homogeneidade entre as variâncias, ou seja, as variâncias de todas as populações de y devem ser semelhantes. Esta característica é denominada de homocedasticidade;
- (c) os valores médios das populações de y para cada conjunto de valores de \mathbf{x} apresentam linearidade, considerando a função f linear;
- (d) os valores de y são obtidos aleatoriamente e são independentes;
- (e) as medidas dos componentes do vetor \mathbf{x} não contêm erros.

A violação das três primeiras suposições, (a), (b) e (c), pode ser contornada pela modificação da métrica da resposta. Se a violação ocorrer apenas em relação à segunda suposição das três, a correção pode ser feita pela utilização de mínimos quadrados ponderados. O cumprimento da suposição (d) pode ser garantido pela forma de obtenção dos dados. Já a suposição (e) dificilmente será cumprida na prática, mas em geral, os erros de medição de \mathbf{x} são tão pequenos em relação aos erros de medição de y que se tornam desprezíveis.

Empregam-se várias técnicas estatísticas para avaliar a qualidade do modelo encontrado. Uma delas é a análise da variância, na qual a adequação é medida pela estatística F que testa a hipótese nula de todos os coeficientes serem nulos. Seu cálculo, descrito no Quadro 2.2 (Adaptado de ZAR, 1999, p.422), permite encontrar o *valor-p*. Se o resultado for pequeno, menor do que o nível de significância adotado, normalmente de $0,05$, considera-se ser o modelo significativo: pelo menos uma das variáveis independentes incluídas no modelo altera o valor esperado da resposta. A maioria dos *softwares* que realiza esse procedimento apresenta os resultados na forma do Quadro 2.2, acrescido de uma coluna com o *valor-p*.

QUADRO 2.2 – Resultados da ANOVA para avaliação da qualidade de um modelo de regressão

(1) Fontes de variação	(2) Soma de quadrados SQ	(3) Graus de liberdade gl	(4) Quadrado médio QM	(5) F
Total	$\sum_{n=1}^N \sum_{m=1}^M (y_{nm} - \bar{y}_n)^2$	$(NM - 1)$		
Regressão	$\sum_{n=1}^N (\hat{y}_n - \bar{y}_n)^2$	$(v - 1)$	$\frac{SQ_{regressão}}{gl_{regressão}}$	$\frac{QM_{regressão}}{QM_{erro}}$
Erro	$SQ_{total} - SQ_{regressão}$	$(NM - v)$	$\frac{SQ_{erro}}{gl_{erro}}$	

Fonte: Adaptada de Zar (1999, p.422).

Obs: Nas expressões utilizadas, N é a quantidade de condições experimentais; M é a quantidade de replicações; y_{nm} é a m -ésima replicação da n -ésima condição experimental; \bar{y}_n é a média das respostas na n -ésima condição experimental; \hat{y}_n é o valor predito para a n -ésima condição experimental e v é a quantidade de parâmetros usados no modelo.

A partir desses resultados, também é possível encontrar o coeficiente de determinação (R^2), que estabelece, percentualmente, o quanto da variação da resposta média de y é explicado pelo modelo, por meio da expressão (ZAR, 1999, p.423):

$$R^2 = \frac{SQ_{regressão}}{SQ_{total}} = 1 - \frac{SQ_{erro}}{SQ_{total}} \quad (2.10)$$

sendo que $SQ_{regressão}$, SQ_{total} e SQ_{erro} estão definidos no Quadro 2.2.

O resultado pode ser corrigido em função da perda dos graus de liberdade com cada termo do modelo ($R^2_{corrigido}$), assumindo a forma (ZAR, 1999, p.423):

$$R^2_{corrigido} = 1 - \frac{QM_{erro}}{QM_{Total}} \quad (2.11)$$

sendo que QM_{total} e QM_{erro} estão definidos no Quadro 2.2.

O resultado do coeficiente de determinação, entretanto, é muito influenciado pela quantidade de variáveis independentes consideradas na equação e pode aumentar se essa quantidade for aumentada sem, necessariamente, o modelo melhorar. Por isto, uma análise gráfica dos resíduos é extremamente útil para complementar a avaliação. Um gráfico de resíduos relacionando estimativas de resíduos com os correspondentes valores preditos de y pode ser utilizado, assim como um gráfico de probabilidade normal dos resíduos.

Quando o experimento apresentar replicações, pode ser realizado o teste de falta de ajuste, que também utiliza a estatística F (ver Montgomery, 1997a, p.568-569). Nesse caso, considera-se a variabilidade devida ao erro decomposta em dois componentes: um deles estimado a partir da variabilidade existente entre as replicações (erro puro) e o outro, a partir da variabilidade devida aos fatores e interações não incluídos no modelo considerado (erro residual). Testa-se a hipótese nula de o modelo ser adequado por não apresentar a variabilidade residual significativamente maior do que a variabilidade devida ao erro puro. Se o resultado é significativo, o teste rejeita o modelo por falta de ajuste.

Existe, também, a possibilidade de ser testado cada um dos coeficientes da equação encontrada, por meio da estatística t (ver Montgomery, 1997a, p.557-560).

Uma possível causa da não-adequação de um modelo pode ser a métrica usada em sua resposta, a qual pode ser alterada por meio de uma transformação matemática para, dentro do possível, amenizar, simultaneamente, as conseqüências do não-cumprimento das suposições teóricas dos modelos lineares. A transformação $\ln y$ é bastante utilizada para modificar a métrica de respostas de experimentos (mais detalhes na seção 2.5).

Também utiliza-se a família de transformações de Box-Cox. Assume-se existir sempre um modelo aditivo com erros normalmente distribuídos e variância constante para determinado valor de y^λ , sendo λ estimado pelo método de máxima verossimilhança. Para determinado modelo, usando as expressões (BOX *et al.*, 1978, p.239):

$$y^{(\lambda)} = \frac{y^\lambda - 1}{\lambda \cdot \bar{y}^{\lambda-1}} \quad \text{para } \lambda > 0 \quad (2.12)$$

e

$$y^{(0)} = \bar{y} \ln y \quad \text{para } \lambda = 0, \quad (2.13)$$

sendo que \bar{y} é a média geométrica de todos os dados. A estimativa de máxima verossimilhança de λ é aquela com menor

soma dos quadrados dos resíduos associados a determinado modelo. O gráfico da transformação Box-Cox, que coloca a soma dos quadrados dos resíduos como função dos valores de λ , pode ser utilizado para identificar esse valor. Se λ for igual a -1 , $-0,5$, 0 , $0,5$ ou 1 , a transformação é equivalente, respectivamente, à transformação: inversa, inverso da raiz quadrada, logarítmica natural, raiz quadrada e identidade (mais detalhes em Box *et al.*, 1978).

Depois de identificado o estimador de máxima verossimilhança de λ , a significância da redução da soma dos quadrados dos resíduos pela utilização da transformação Box-Cox com o valor estimado, pode ser avaliada pelo teste Qui-quadrado. Outros recursos, como o gráfico de probabilidade normal, podem ser usados. Na prática, não é indispensável ser empregado exatamente o valor da estimativa encontrado para o coeficiente λ e, se seus resultados estiverem próximos dos valores: -1 ; $-0,5$; 0 ou $0,5$, a transformação Box-Cox pode ser substituída por alguma das transformações acima citadas.

Após o ajuste do modelo de regressão aos dados experimentais, é possível obter várias informações sobre o comportamento da resposta do sistema em função dos fatores controláveis do processo. Quando se aplica a técnica de regressão às observações obtidas por meio de um experimento, normalmente têm-se várias variáveis independentes (fatores controlados no experimento). Nem sempre todos esses fatores interferem, de forma significativa, na resposta, não necessitando serem considerados na superfície de resposta. As técnicas citadas podem ser utilizadas, em conjunto, para identificar quais fatores do processo têm efeitos de locação e devem ser considerados na equação matemática que representa a superfície de resposta.

2.2 O MÉTODO DE TAGUCHI

Nas últimas décadas, houve uma valorização muito grande da qualidade, razão pela qual vários estudos foram desenvolvidos visando assegurar o bom desempenho de produtos e processos a baixo custo. Vários conceitos, técnicas e métodos foram revisados, redefinidos, aprimorados ou propostos. Foi relevada a importância holística da qualidade que deve iniciar

no planejamento de um produto, preocupando-se, também, com o planejamento do processo, a produção propriamente dita e a assistência técnica.

Nesse período, Taguchi propôs uma metodologia que, de acordo com Tsui (1992), é eficiente e utiliza experimentos estatísticos, entre outras ferramentas, para obter a melhoria da qualidade de produtos e processos sem o aumento de custos. O objetivo dela é tornar os produtos e processos insensíveis a fatores muito difíceis, caros ou impossíveis de serem controlados, como as variações do meio ambiente e alterações no processo produtivo (desgaste de peças, aferição de máquinas etc).

O método, inicialmente difundido e amplamente utilizado na indústria japonesa, foi introduzido, nos anos 80, em várias e importantes indústrias americanas, como *AT&T Bell Laboratories*, *Ford Motors* e *Xerox*, com melhoria significativa de seus produtos e processos, conforme atestam Nair (1992) e muitos outros. De acordo com Kackar e Shoemaker (1986), na indústria eletrônica o método permitiu redução de até 60% na variância de processos.

Entretanto, seu trabalho tem recebido muitas críticas, principalmente de estatísticos, que questionam a sustentação teórica da maior parte dos procedimentos propostos e, de acordo com Nair (1992), mesmo no Japão, suas técnicas não são totalmente aceitas.

Taguchi, porém, é considerado o pioneiro no controle de qualidade *off-line*, por evidenciar a necessidade de considerar que a preocupação com as técnicas para aperfeiçoar a qualidade e a produtividade deve existir desde a concepção do produto, quando se deve procurar maximizar a qualidade, minimizando os custos de fabricação. Nesse tipo de controle, os esforços para a obtenção da qualidade ocorrem durante o projeto, antes da fabricação.

O controle de qualidade *off-line* inicia-se com o projeto do sistema, o qual estabelece os aspectos conceituais do produto e de seu processo produtivo. Isso é feito a partir das características de desempenho, definidoras da qualidade, indicadas pelos possíveis clientes ou consumidores. Então, é desenvolvido um protótipo do produto, com a determinação do tipo de material, peças e componentes e do sistema de montagem a ser utilizado, além da identificação de fatores capazes de afetar seu desempenho. O projeto do sistema requer o conhecimento não apenas das necessidades do consumidor, mas também das diversas técnicas a serem utilizadas no processo de manufatura.

A segunda etapa, denominada projeto de parâmetros, define a combinação entre os níveis dos parâmetros do sistema idealizado, possibilitando a obtenção de produtos e processos pouco sensíveis ou insensíveis a fontes causadoras de variabilidade, fazendo as características de qualidade apresentarem uma variabilidade mínima.

Já a terceira, chamada projeto de tolerâncias, define a amplitude tolerável dos desvios em relação ao valor ideal de cada característica, levando em consideração os custos de fabricação. Ela busca equilíbrio entre a perda para o consumidor pelo comportamento da variabilidade e o aumento dos custos de produção.

Nas duas últimas etapas são muito utilizadas várias ferramentas estatísticas, entre as quais o projeto de experimentos, por viabilizarem melhor conhecimento do sistema de engenharia. É também muito empregada a função perda, por permitir avaliar, simultaneamente, qualidade e custo.

2.2.1 A função perda de Taguchi

A filosofia de Taguchi, de acordo com Taguchi *et al.* (1990, p. 2), enfatiza a redução da variabilidade para a obtenção da melhoria da qualidade, definindo-a como a perda imposta por um produto à sociedade. Chama a atenção para o fato de qualquer variação nas dimensões de uma característica funcional de um produto ou processo envolver custo adicional, representando perda para o consumidor. De acordo com ele, preço e qualidade estão muito relacionados, pois entende que o preço representa uma perda para o consumidor no momento da aquisição do produto, enquanto a falta de qualidade, uma perda adicional durante sua utilização. Usando uma expansão em série de Taylor, Taguchi *et al.* (1990, p.18), justificam a quantificação dessa perda por uma função quadrática definida por:

$$L(y) = k'(y - \tau)^2 \quad \text{para } y \in [\tau - d; \tau + d], \quad (2.14)$$

sendo:

- L a perda monetária ocorrida em função de y ;
- k' uma constante de proporcionalidade;
- y o resultado da medição da característica focalizada;
- τ o valor-alvo da característica focalizada;
- d o afastamento máximo de y em relação ao valor-alvo, que deve estar de acordo com os limites de especificação.

A constante k' de proporcionalidade pode ser obtida pela expressão:

$$k' = \frac{P_{\max.}}{d^2}, \quad (2.15)$$

sendo que P_{\max} representa a perda máxima devido a um item não-conforme.

A expressão (2.14) indica existir um aumento da perda à medida que o valor medido afasta-se do valor-alvo, conferindo à qualidade a idéia de continuidade, de acordo com o visualizado no gráfico da figura 2.7.

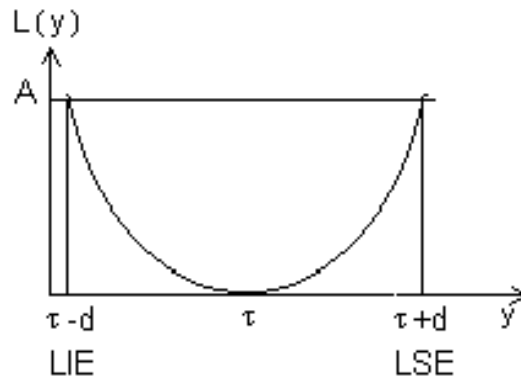


FIGURA 2.7 – Representação gráfica da função perda de Taguchi, onde LIE e LSE representam, respectivamente, os limites de especificação inferior e superior. (Adaptado de FIOD NETO, 1997, p.25)

A obtenção da qualidade é, então, um esforço contínuo a perseguir a idéia de zero defeito no produto final, assim como no processo, não bastando produzir dentro dos limites de especificação. Para ter qualidade, um produto deve ser projetado para ser manufaturado eficientemente, com as características de qualidade mais próximas, dentro do possível, dos valores nominais de seu projeto e, ao mesmo tempo, com a maior resistência possível a fontes de variação. Essa qualidade seguramente fornecerá redução de custos.

A expressão (2.14) quantifica a perda para uma unidade de produto. Entretanto, nos programas de melhoria da qualidade, o objetivo principal é a minimização do valor esperado dessa perda, dada pela expressão (TAGUCHI, 1990, p.38):

$$\mu_L = E\{L(y)\} = k \left[(\mu_y - \tau)^2 + \sigma_y^2 \right], \quad (2.16)$$

sendo que μ_y e σ_y^2 são, respectivamente, o valor esperado e a variância de y .

De acordo com a expressão (2.16), a perda média quadrática depende de dois componentes: um deles representa o afastamento quadrático do valor médio das observações em relação ao valor-alvo $(\mu_y - \tau)^2$, enquanto o outro refere-se à variância de y (σ_y^2), relacionada com o afastamento quadrático dos valores observados em relação à sua média.

A expressão 2.16 refere-se ao caso em que o valor ideal da característica é um valor específico. Nesta situação, denominada *nominal-é-melhor (NEM)*, são utilizadas tolerâncias bilaterais, pois a perda é nula quando o valor médio de y coincide com o valor-alvo, aumentando simetricamente à medida que y se desvia desse alvo (figura 2.8 - b).

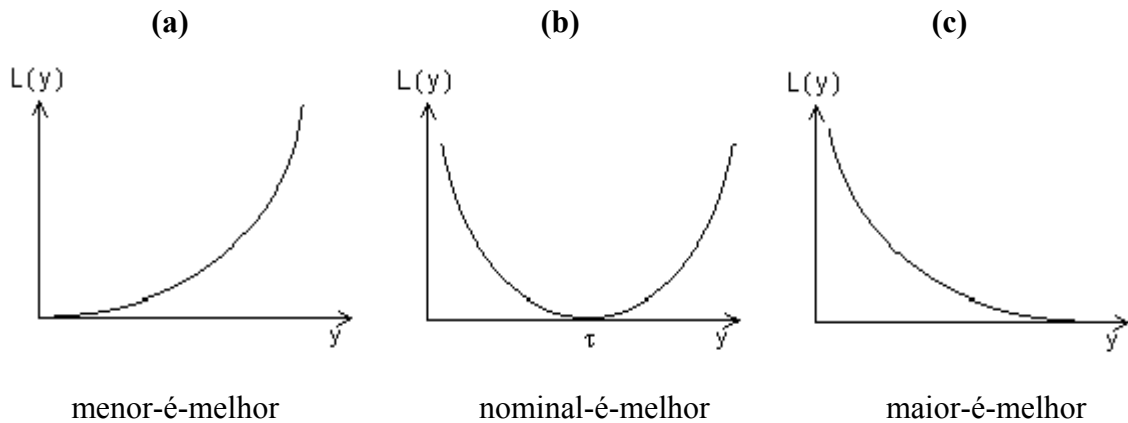


FIGURA 2.8 – Funções de perda quadrática em termos do tipo de característica de qualidade

Além deste, existem mais dois casos. Um deles, correspondente a uma característica do tipo *menor-é-melhor (mEM)*, envolve característica não-negativa, cujo valor ideal é zero (figura 2.8 – a).

O outro caso, denominado *maior-é-melhor (MEM)*, envolve situações em que existe apenas um valor mínimo estabelecido e, quanto maior o valor da característica y , melhor a qualidade (figura 2.8 – c).

2.2.2 Experimentos sob a ótica de Taguchi

Taguchi focou a meta de seu trabalho no projeto de parâmetros, que utiliza experimentos estatisticamente planejados, propondo uma estratégia experimental baseada em experimentos fatoriais fracionados.

Na nova abordagem, os fatores capazes de afetar as características de qualidade são classificados em: fatores de controle, facilmente controlados e manipulados, e fatores de ruído, impossíveis, difíceis ou muito caros para serem controlados. Num processo industrial, os fatores controláveis são os parâmetros do processo que podem afetar as características de qualidade do produto ou processo e serem alterados pelo engenheiro. Os fatores de ruído são todas as variáveis (temperatura, umidade, erro humano, desgaste de peças, imperfeições do

processo,...) suficientes para causar o afastamento dessas características de qualidade de seus valores-alvo, não podendo ou sendo muito difíceis ou caros de serem controlados na linha de produção. Na realização de um experimento em laboratório, entretanto, alguns fatores de ruído podem ser controlados, tornando-se também fatores de um projeto experimental.

A idéia básica da estratégia é identificar, por meio da análise da interação entre fatores de controle e fatores de ruído, os níveis dos fatores controláveis que tornam o sistema: produto e processo, robusto às variações nos fatores de ruído. A redução da variabilidade é obtida mais pela redução do efeito das fontes da variação do que pelo seu controle propriamente dito. Por esse motivo, essa etapa também é denominada *projeto robusto*.

Considerando os dois tipos de fatores possíveis de interferir na resposta de um processo, Taguchi propôs um outro tipo de projeto de experimento, organizado por duas matrizes: uma matriz para os fatores de controle, também chamada de *arranjo interno*, e outra para os fatores de ruído, também chamada de *arranjo externo*. Elas são cruzadas, razão pela qual esse modelo de projeto recebe o nome de *arranjo-produto*. A inclusão da matriz com os fatores de ruído é realizada para estimular a variação na resposta devida aos ruídos, para seus efeitos poderem ser melhor avaliados.

A organização dessas matrizes baseia-se na matriz de um projeto fatorial, completo ou fracionado. A tabela 2.3 apresenta algumas características de 11 adaptações propostas por Taguchi para estes tipos de projeto. O tipo *L8*, por exemplo, necessita de, no mínimo, 8 ensaios e pode investigar até 7 fatores ensaiados em dois níveis.

TABELA 2.3 – Algumas características dos experimentos propostos por Taguchi

Tipo de planeja mento	Quantidade de ensaios	Quantidade máxima de fatores	Quantidade de fatores com			
			2 níveis	3 níveis	4 níveis	5 níveis
<i>L4</i>	4	3	3
<i>L8</i>	8	7	7
<i>L9</i>	9	4	...	4
<i>L12</i>	12	11	11
<i>L16</i>	16	15	15
<i>L'16</i>	16	5	5	...
<i>L18</i>	18	8	1	7
<i>L25</i>	25	6	6
<i>L27</i>	27	13	...	13
<i>L32</i>	32	31	31
<i>L'32</i>	32	10	1	...	9	...

Fonte: *Help do Statistica 6.0*. [Copyright Stat soft Inc.,2001]

Taguchi também desenvolveu alguns instrumentos, ditos facilitadores para auxiliar na organização das planilhas usadas por tais experimentos, entre os quais estão os *gráficos lineares* que relacionam posição da coluna na tabela ou planilha de planejamento, fatores e interações. A tabela 2.4, por exemplo, apresenta o delineamento de um projeto fatorial completo com 3 fatores, ensaiados em 2 níveis, que geram 8 condições experimentais. De acordo com a proposta de Taguchi, esse planejamento pode ser usado no arranjo ortogonal $L8$, investigando até 7 fatores ensaiados em 2 níveis, sendo os fatores organizados conforme a tabela 2.5.

A figura 2.9 mostra um gráfico linear para ser usado na montagem do projeto mostrado na tabela 2.5. Se os fatores A , B e C forem colocados, respectivamente, nas colunas 1, 2 e 4 dessa tabela, as interações AB , AC e BC deverão ficar, respectivamente, nas colunas 3, 5 e 6. A sétima coluna normalmente é usada para um quarto fator ou é considerada como parte do erro experimental. As colunas (3), (5) e (6) também podem ser usadas para representar outros fatores, mas isto leva ao aparecimento de confundimentos, como em qualquer projeto fracionado.

TABELA 2.4 – Projeto fatorial completo do tipo 2^3

Condição Experimental	Fatores							Respostas (y_n)
	(1) A	(2) B	(4) C	(3) (AB)	(5) (AC)	(6) (BC)	(7) (ABC)	
1	+1	+1	+1	+1	+1	+1	+1	y_1
2	+1	+1	-1	+1	-1	-1	-1	y_2
3	+1	-1	+1	-1	+1	-1	-1	y_3
4	+1	-1	-1	-1	-1	+1	+1	y_4
5	-1	+1	+1	-1	-1	+1	-1	y_5
6	-1	+1	-1	-1	+1	-1	+1	y_6
7	-1	-1	+1	+1	-1	-1	+1	y_7
8	-1	-1	-1	+1	+1	+1	-1	y_8

Nota: y_n representa a resposta da n -ésima condição experimental e A , B , e C , os fatores manipulados.

TABELA 2.5 – Projeto fatorial do tipo 2^3 para um projeto $L8$

Condição Experimental	Fatores							Respostas (y_n)
	(1) A	(2) B	(3) (AB)	(4) C	(5) (AC)	(6) (BC)	(7) (ABC)	
1	+1	+1	+1	+1	+1	+1	+1	y_1
2	+1	+1	+1	-1	-1	-1	-1	y_2
3	+1	-1	-1	+1	+1	-1	-1	y_3
4	+1	-1	-1	-1	-1	+1	+1	y_4
5	-1	+1	-1	+1	-1	+1	-1	y_5
6	-1	+1	-1	-1	+1	-1	+1	y_6
7	-1	-1	+1	+1	-1	-1	+1	y_7
8	-1	-1	+1	-1	+1	+1	-1	y_8

Nota: y_n representa a resposta da n -ésima condição experimental e A , B , e C , os fatores manipulados.

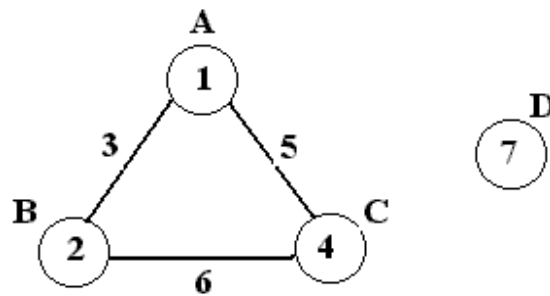


FIGURA 2.9 - Gráfico linear para arranjo ortogonal L8 de Taguchi (FIOD NETO, 1997, p.46), em que *A*, *B*, *C* e *D* são os fatores manipulados

TABELA 2.6 – Delineamento de um projeto fatorial do tipo 2^5

Condição experimental	Fatores					Respostas (y_n)
	A	B	C	D	E	
1	+1	+1	+1	+1	+1	y_1
2	+1	+1	+1	+1	-1	y_2
3	+1	+1	+1	-1	+1	y_3
4	+1	+1	+1	-1	-1	y_4
5	+1	+1	-1	+1	+1	y_5
6	+1	+1	-1	+1	-1	y_6
7	+1	+1	-1	-1	+1	y_7
8	+1	+1	-1	-1	-1	y_8
9	+1	-1	+1	+1	+1	y_9
10	+1	-1	+1	+1	-1	y_{10}
11	+1	-1	+1	-1	+1	y_{11}
12	+1	-1	+1	-1	-1	y_{12}
13	+1	-1	-1	+1	+1	y_{13}
14	+1	-1	-1	+1	-1	y_{14}
15	+1	-1	-1	-1	+1	y_{15}
16	+1	-1	-1	-1	-1	y_{16}
17	-1	+1	+1	+1	+1	y_{17}
18	-1	+1	+1	+1	-1	y_{18}
19	-1	+1	+1	-1	+1	y_{19}
20	-1	+1	+1	-1	-1	y_{20}
21	-1	+1	-1	+1	+1	y_{21}
22	-1	+1	-1	+1	-1	y_{22}
23	-1	+1	-1	-1	+1	y_{23}
24	-1	+1	-1	-1	-1	y_{24}
25	-1	-1	+1	+1	+1	y_{25}
26	-1	-1	+1	+1	-1	y_{26}
27	-1	-1	+1	-1	+1	y_{27}
28	-1	-1	+1	-1	-1	y_{28}
29	-1	-1	-1	+1	+1	y_{29}
30	-1	-1	-1	+1	-1	y_{30}
31	-1	-1	-1	-1	+1	y_{31}
32	-1	-1	-1	-1	-1	y_{32}

Nota: y_n representa a resposta da n -ésima condição experimental e *A*, *B*, *C*, *D* e *E*, os fatores manipulados.

Usando tais modelos de delineamento, Taguchi constrói uma matriz para os fatores de controle e uma matriz para os fatores de ruído, cruzando-as. Por exemplo, um projeto fatorial

completo de um experimento com cinco fatores (A, B, C, D e E), todos ensaiados em 2 níveis (-1 e $+1$), apresenta 32 condições experimentais, organizadas conforme o disposto na tabela 2.6.

A figura 2.10 apresenta esse mesmo experimento disposto conforme um arranjo-produto, em que A, B e C são fatores de controle e D e E , fatores de ruído. Na figura, as variáveis y_{nm} são os resultados do experimento relativos à característica funcional do produto, variando no sentido vertical de acordo com as condições experimentais definidas a partir dos níveis dos diferentes fatores de controle e, no sentido horizontal, de acordo com as condições experimentais definidas a partir dos níveis dos diferentes fatores de ruído.

Assim, cada uma das quatro observações de cada condição experimental do arranjo interno (fatores de controle) é feita para uma combinação diferente entre os níveis dos fatores de ruído investigados, de forma a mensurar a variabilidade.

arranjo interno (fatores de controle)				arranjo externo (fatores de ruído)				
				+	-	+	-	D
CE	A	B	C	+	+	-	-	
1	+	+	+	y_{11}	y_{12}	y_{13}	y_{14}	
2	+	+	-	y_{21}	y_{22}	y_{23}	y_{24}	
3	+	-	+	y_{31}	y_{32}	y_{33}	y_{34}	
4	+	-	-	y_{41}	y_{42}	y_{43}	y_{44}	
5	-	+	+	y_{51}	y_{52}	y_{53}	y_{54}	
6	-	+	-	y_{61}	y_{62}	y_{63}	y_{64}	
7	-	-	+	y_{71}	y_{72}	y_{73}	y_{74}	
8	-	-	-	y_{81}	y_{82}	y_{83}	y_{84}	

FIGURA 2.10 - Matriz experimental de um arranjo-produto ou arranjo-cruzado (Adaptado de CATEN, 1995), em que A, B e C são fatores de controle e D e E são fatores de ruído. CE representa condição experimental

Uma outra contribuição relevante proposta por Taguchi, para analisar os resultados de um experimento, é a utilização de uma estatística de desempenho, denominada *razão sinal/ruído*. Existem várias fórmulas para encontrar essa estatística, sendo três delas especialmente indicadas para características classificadas como NEM , MEM e mEM . De acordo com Vinning e Myers (1990), as expressões são:

$$\left(\frac{S}{R}\right)_n = 10 \log \left(\frac{\overline{y}_n}{s_n^2} \right), \quad \text{para característica do tipo } NEM, \quad (2.17)$$

$$\left(\frac{S}{R}\right)_n = -10 \log \left[\frac{\sum_{m=1}^M y_{nm}^2}{M} \right], \quad \text{para característica do tipo } mEM, \quad (2.18)$$

$$\left(\frac{S}{R}\right)_n = -10 \log \left[\frac{1}{M} \left(\sum_{m=1}^M \left(\frac{1}{y_{nm}^2} \right) \right) \right], \quad \text{para característica do tipo } MEM, \quad (2.19)$$

sendo:

$\left(\frac{S}{R}\right)_n$ a razão sinal-ruído na n -ésima condição experimental ($n=1, 2, \dots, N$);

\bar{y}_n a média das observações na n -ésima condição experimental, encontrada pela expressão $\bar{y}_n = \left(\sum_{m=1}^M y_{nm} \right) / M$;

s_n^2 a variância das observações na n -ésima condição experimental, encontrada pela expressão $s_n^2 = \left(\sum_{m=1}^M (y_{nm} - \bar{y}_n)^2 \right) / (M - 1)$

Estas medidas, representativas da variabilidade quando existem fatores de ruído presentes, são tratadas como resposta na análise do experimento. Existem várias maneiras de conduzir essa análise. Uma delas consiste em utilizar uma análise de variância, identificando, por meio do teste F , os efeitos significativos. Gráficos de efeitos também podem ser usados para identificar fatores aparentemente significativos.

Para características do tipo NEM , deve-se encontrar primeiro a combinação de níveis de fatores que maximizam a razão sinal/ruído, para depois trazer a resposta média para o alvo, modificando o ajuste de alguns fatores. Fatores incapazes de afetar a razão sinal-ruído, denominados fatores de ajustamento, são utilizados para ajustar o valor esperado da resposta ao alvo. Para características dos tipos mEM e MEM , a combinação ótima de fatores seria encontrada simplesmente pela maximização da razão sinal/ruído.

Além dessas, outras técnicas para analisar dados experimentais foram propostas por Taguchi, algumas delas, consideradas por alguns autores, de qualidade bastante duvidosa. Taguchi também propõe que seja realizado, pelo menos, um experimento confirmatório para avaliar o desempenho da solução encontrada.

2.2.3 Considerações sobre o trabalho desenvolvido por Taguchi

Embora o método de Taguchi seja bastante controverso, elogiado por uns e extremamente criticado por outros, ele foi muito importante para a Engenharia da Qualidade. Talvez ele tenha conseguido, com grande objetividade e precisão, detectar e delimitar alguns sérios problemas da área, não tendo o mesmo sucesso, entretanto, com as técnicas propostas para operacionalizar a busca de solução dos problemas detectados.

Taguchi chamou a atenção para a importância da preocupação com a qualidade iniciar durante a fase de projeto, enfatizando a necessidade de um trabalho preventivo para a obtenção da qualidade com custos reduzidos, em lugar do trabalho simplesmente curativo até então efetivamente desenvolvido.

De acordo com Kackar (1986), qualidade é um conceito complexo e de múltiplas facetas, podendo ser avaliada sob diversos aspectos: desempenho, apresentação, confiabilidade, conformidade, durabilidade, serventia, estética e qualidade percebida. A importância de cada um dos aspectos depende da natureza do produto e das necessidades do consumidor, fazendo seu conceito mudar, dependendo do contexto onde estiver sendo utilizado. Taguchi, ao redefinir qualidade, considerou apenas algumas das dimensões, relacionando seu conceito com a idéia de perda imposta por um produto à sociedade, a qual descreveu matematicamente por uma função quadrática, definida a partir de um valor-alvo, de uma variabilidade e de um custo.

Ao formular a definição, Taguchi consegue relacionar definitivamente qualidade com três idéias extremamente importantes: continuidade, diminuição de variabilidade e redução de custos. Taguchi conseguiu evidenciar que os consumidores desejam produtos de alta qualidade e preços baixos e as indústrias devem conscientizar-se de a qualidade nunca ser suficientemente alta, nem o preço suficientemente baixo, devendo portanto, buscar, de forma contínua, a melhoria de seus produtos e processos, bem como a redução de custos.

Para atingir tais objetivos, torna-se indispensável a diminuição da variabilidade das características consideradas relevantes para a obtenção da qualidade. Entretanto, normalmente, quanto menor a variabilidade, maior o custo do processo produtivo. O ponto de equilíbrio entre uma variabilidade máxima, que satisfaça o consumidor, e um custo mínimo, deve ser encontrado.

A minimização da variabilidade pode ser conseguida, diretamente, com a aferição de fatores controláveis e, indiretamente, com a aferição dos fatores de ruído atuantes num processo produtivo. Os fatores de ruído não podem ser controlados na linha de produção, mas muitos deles podem ser controlados durante um experimento. A experimentação estatística

permite verificar se os fatores de ruído interagem com os fatores controláveis, de forma a serem controlados indiretamente por uma combinação ideal entre fatores controláveis. Fiod Neto (1997) ressalta que, em lugar de remover os ruídos, o que pode ser de difícil execução ou encarecer demais o processo produtivo, os métodos de Taguchi ambicionam remover os efeitos negativos causados pelos ruídos, sem aumentar os custos de produção.

Pela abordagem de Taguchi, durante o projeto de parâmetros, procura-se identificar as causas da variabilidade, tornando o projeto robusto a ela. Isto pode ser feito pela experimentação estatística. Talvez as técnicas propostas por Taguchi para operacionalizar uma experimentação tenham pouca sustentação teórica, como é o caso do delineamento arranjo-produto e da razão sinal/ruído.

Phadke (1992) considera que, como o projeto robusto é uma metodologia usada para melhorar a produtividade durante o projeto, produtos de alta qualidade e baixo custo podem ser obtidos. Considera, ainda, ter o trabalho de Taguchi dado uma nova dimensão à estatística experimental por definir como, economicamente, reduzir a variação de desempenho de um produto e como assegurar que as decisões tomadas durante um experimento de laboratório se comprovem no ambiente de fabricação e consumo. Entretanto, chama a atenção para o fato de o delineamento experimental proposto não contemplar o estudo de interações entre fatores de controle. Considera não existirem regras aptas para garantir sua ausência e, muitas vezes, seus efeitos poderem estar confundidos com efeitos principais.

Box (1992) considera que a importância do trabalho de Taguchi foi evidenciar a necessidade de estudos de robustez no projeto de produtos e processos, salientando existir, embora pareça que a estratégia experimental pretenda encontrar apenas a combinação ótima, um objetivo maior: conhecer o sistema de engenharia. Sugere, ao invés de utilizar a razão sinal/ruído para simplificar a análise estatística, empregar transformações matemáticas que incluam a função logarítmica. Questiona a qualidade e a suposta simplicidade dos gráficos lineares, assim como do projeto arranjo-produto por conduzir a um experimento muito grande.

Kackar (1992) diz que a maior contribuição do trabalho de Taguchi foi expandir o campo de aplicação do projeto experimental clássico, cobrindo um amplo espectro de problemas de engenharia, além de formar um excelente instrumental para diminuir a distância entre engenheiros e estatísticos. Chama a atenção, entretanto, para o fato de o sucesso do projeto de parâmetros depender de dois fatores: a existência de certas interações entre fatores de controle e fatores de ruído e a habilidade do engenheiro em encontrar os fatores envolvidos, não sendo, portanto, universal. Salienta também, no experimento, de acordo com

Taguchi, a resposta dever ser escolhida de tal forma que os efeitos dos fatores de controle sejam aditivos, o que considera irreal.

Myers e Vinning (1992) consideram lamentável que os conceitos apresentados por Taguchi tenham sido ofuscados por seus controversos métodos para modelar e analisar dados. Afirmam ser de vital importância a utilização de fatores de ruído, embora salientem a existência desse conceito já antes de Taguchi. Contudo, Taguchi foi responsável pela demonstração das vantagens de seu uso formal como parte de um projeto de experimentos.

Myers, Khuri e Vinning (1992) julgam que a maior contribuição de Taguchi foi a necessidade de incluir a variabilidade do processo como resposta. Os mesmos autores consideram também importante a formalização da idéia de os parâmetros do processo e os fatores de ruído variarem de forma aleatória, afetando as características de qualidade.

Lucas (1994) chama a atenção para o fato de as interações entre fatores de arranjos externos e entre fatores de arranjos internos não serem analisadas, além de mostrar que os projetos compostos são comparáveis em tamanho com os menores projetos possíveis de Taguchi. Considera, ainda, uma contribuição importante a inclusão de fatores ambientais no experimento.

A abordagem proposta por Taguchi também evidenciou a importância do conhecimento do contexto industrial (produto – manufatura - consumidor) pelo idealizador, operacionalizador e analista do experimento, considerando que este deve ser o resultado de um trabalho integrado entre a Engenharia da Qualidade e o método estatístico.

Outro aspecto importante dessa abordagem é a simplicidade da maior parte das técnicas utilizadas, consideradas de fácil compreensão e utilização, mesmo por pessoas não-especialistas. Talvez tal fato tenha sido a causa da grande aceitação do método proposto que, embora apresentasse falhas, de acordo com Caten (1995, p. 31), era bem melhor do que a maioria das técnicas não-estatísticas empregadas até então, na prática, com a mesma finalidade.

Já Kackar (1986) salienta, no método de Taguchi, ser um fator relevante a importância da variação do desempenho poder ser reduzida pela retirada dos efeitos não-lineares dos parâmetros das características de desempenho, além do fato da utilização de experimentos estatisticamente planejados para identificar grupos de parâmetros que reduzem a variação do desempenho.

O sucesso do método de Taguchi talvez deva ser atribuído muito mais à sua filosofia, organizada, sistematizada e expressa de forma simples e objetiva, do que ao conjunto de técnicas propriamente ditas. O desenvolvimento do conceito de Engenharia Robusta, pela qual

o engenheiro deve concentrar-se na função do que está sendo projetado, foi de vital importância para a Engenharia da Qualidade. Várias empresas reconhecem a eficiência dessa estratégia de otimização do desempenho de produtos e processos. De maneira geral, os engenheiros consideram terem os métodos de Taguchi um impacto significativo e positivo na indústria, pressionada pela globalização, para melhorar a qualidade com redução de custos, embora estatísticos sejam mais cautelosos em sua avaliação.

Por outro lado, de acordo com Myers e Vinning (1992), nos anos 80, apenas uma pequena percentagem de companhias americanas usava efetivamente métodos estatísticos. Taguchi talvez tenha conseguido dar o primeiro passo para reverter a situação, mostrando os benefícios da utilização desses métodos nesse setor da economia.

Seu trabalho certamente foi um marco para a Engenharia da Qualidade por estimular a indústria a priorizar o controle de qualidade *off-line*, com a ampla utilização de técnicas estatísticas, possibilitando o desenvolvimento de técnicas para criação de produtos de alta qualidade e baixo custo, além de processos que conseguem manufaturá-los corretamente logo na primeira vez. Também conseguiu contribuir para o desenvolvimento da estatística, enquanto ciência, dando um novo enfoque à análise de experimentos estatisticamente planejados.

Engenheiros, consultores e estatísticos concordam com o fato de a filosofia de Taguchi certamente tornar uma empresa mais lucrativa e competitiva, conforme a atual necessidade apontada pela globalização do mercado.

“A América levou muito tempo para descobrir sobre o que Taguchi estava falando. Por muitos anos pensou-se que ele estava falando sobre projeto de experimentos, mas nos últimos anos, nós, da Ford, descobrimos que era muito mais do que isto”

comentou NEIL W. RESSLER, vice-presidente de tecnologia de pesquisa e veículo da Ford Motors Co., citado em WILKINS JR. (2000,p.1).

2.3 EFEITOS DE LOCAÇÃO E DISPERSÃO

Avaliar a variabilidade de um produto ou processo produtivo é fundamental para a obtenção da melhoria da qualidade. Esta talvez tenha sido a grande contribuição de Taguchi para a Engenharia da Qualidade: a ênfase dada à minimização da variabilidade, além da centralização do processo. Embora Shewhart e Deming já tivessem salientado a importância do entendimento e redução dessa variabilidade, fizeram-no pela remoção de causas assinaláveis ou causas especiais que interferissem no processo. Foi Taguchi quem introduziu a

necessidade de desenvolvimento de produtos e processos robustos, insensíveis a qualquer fonte de variação.

O novo contexto, ao analisar um experimento, procura identificar os fatores que afetam a média (locação) e aqueles que afetam a variabilidade (dispersão). Na indústria, utilizam-se fatores com efeitos de locação para ajustar o valor médio da característica de qualidade ao alvo; fatores com efeitos de dispersão são usados para diminuir a variabilidade em torno do valor-alvo e os sem efeitos de locação nem efeitos de dispersão são ajustados pelo nível mais econômico.

A situação a ser analisada neste trabalho refere-se a um projeto fatorial com M replicações e fatores mensurados em dois níveis e, cujas observações para as N condições experimentais: $y_{11}, y_{12}, \dots, y_{1M}; y_{21}, y_{22}, \dots, y_{2M}; y_{N1}, y_{N2}, \dots, y_{NM}$, são independentes, normalmente distribuídas em torno da média μ_y e com variância σ_y^2 . Para cada nível de cada fator, têm-se as médias:

$$\mu_{y(k+)} = \frac{2}{N} \sum_{n(k+)} \mu_{y_n} \quad (2.20)$$

e

$$\mu_{y(k-)} = \frac{2}{N} \sum_{n(k-)} \mu_{y_n} \quad (2.21)$$

e as variâncias médias:

$$\sigma_{y(k+)}^2 = \frac{2}{N} \sum_{n(k+)} \sigma_{y_n}^2 \quad (2.22)$$

e

$$\sigma_{y(k-)}^2 = \frac{2}{N} \sum_{n(k-)} \sigma_{y_n}^2 \quad (2.23)$$

sendo:

μ_{y_n} o valor esperado das observações da n -ésima condição experimental ($n = 1, \dots, N$);

$\sigma_{y_n}^2$ a variância das observações da n -ésima condição experimental ($n = 1, \dots, N$);

$n(k+)$ o conjunto das condições experimentais do nível superior do k -ésimo fator;

$n(k-)$ o conjunto das condições experimentais do nível inferior do k -ésimo fator.

Diz-se que, em projetos 2^K , um fator k apresenta efeito de locação quando os valores esperados das observações em seus níveis forem diferentes, ou seja, $\mu_{y(k+)} \neq \mu_{y(k-)}$, e apresenta efeito de dispersão quando as variâncias das observações em seus níveis forem diferentes, ou seja, $\sigma^2_{y(k+)} \neq \sigma^2_{y(k-)}$.

Conforme já mencionado, a identificação dos efeitos de locação e a estimação da superfície de resposta para o valor esperado já foram exaustivamente discutidas na literatura. A significância dos efeitos principais de cada fator, assim como da interação entre os mesmos, pode ser avaliada pela técnica de análise da variância, enquanto a determinação dos coeficientes da equação matemática resultante da modelagem, pode ser obtida pelo método usual de regressão por mínimos quadrados, acompanhado de uma análise gráfica de resíduos. O mesmo não acontece, entretanto, com a identificação dos efeitos de dispersão e estimação da superfície de resposta para a variância, só valorizados a partir da divulgação das idéias de Taguchi.

De acordo com Box (1988), embora os métodos de Taguchi orientem para o uso de algumas técnicas específicas de experimentação e análise de dados, as técnicas clássicas de planejamento e análise de experimentos podem incorporar os conceitos de Taguchi de forma consistente e mais eficiente. Surge, então, uma nova maneira de conduzir a análise de experimentos estatísticos. Na nova abordagem, com os dados de um experimento, procura-se identificar os fatores que afetam a média e os que afetam a variância e, se necessário, descrever as relações dessas medidas com os fatores controláveis do processo produtivo, por meio de equações de regressão.

Suponha, então, que as respostas de um projeto fatorial com replicações, y_{nm} , em que:

K representa a quantidade de fatores manipulados no experimento;

N representa o número de condições experimentais, $N = 2^K$;

n representa a n -ésima condição experimental ($n = 1, \dots, N$);

k representa o k -ésimo fator ou interação ($k = 1, \dots, K'$);

m representa a m -ésima replicação ($m = 1, \dots, M$);

sigam o modelo:

$$y_{nm} = \mu_{y_n} + \sigma_{y_n} \varepsilon_{nm}, \quad (2.24)$$

em que:

μ_{y_n} é o valor esperado para a resposta na n -ésima condição experimental;

- σ_{y_n} é desvio-padrão das resposta na n -ésima condição experimental;
 ε_{nm} é o erro experimental na m -ésima replicação da n -ésima condição experimental
 ($n=1, 2, \dots, N; m=1, 2, \dots, M$).

Os erros experimentais são independentes e identicamente distribuídos por meio de uma distribuição normal com média zero e variância unitária.

Se as respostas ou observações estiverem caracterizadas como uma variável aleatória contínua, as superfícies para μ_y e σ_y^2 podem ser representadas por (BARBETTA, 1998, p.36-37):

$$\mu_y = f(\mathbf{x}, \boldsymbol{\beta}), \quad (2.25)$$

e

$$\sigma_y^2 = h(\mathbf{x}, \boldsymbol{\theta}), \quad (2.26)$$

onde $\mathbf{x} \in D \subset \mathfrak{R}^k$, sendo:

- μ_y o valor esperado de y ;
 σ_y^2 a variância de y ;
 f uma função que liga os fatores ao valor esperado de y ;
 h uma função que liga os fatores à variância de y ;
 \mathbf{x} o vetor de variáveis independentes ou fatores manipulados no experimento;
 $\boldsymbol{\beta}$ o vetor de parâmetros com os efeitos de locação;
 $\boldsymbol{\theta}$ o vetor de parâmetros com os efeitos de dispersão;
 $D \subset \mathfrak{R}^k$ a região experimental.

Normalmente, a média é descrita por uma função linear, também denominada de aditiva, enquanto a variância, por uma função log-linear, também chamada multiplicativa, embora em algumas situações, o modelo aditivo possa ser utilizado para descrição da variância.

Neste trabalho, será enfatizado o uso do modelo multiplicativo para a variância, por ser o mais comum em função das propriedades desta estatística. Além disto, com esse modelo, não existe necessidade de nenhuma restrição relacionada aos parâmetros para encontrar uma

estimativa positiva (BRENNEMAN, 2000), pois a função inversa do logaritmo natural é sempre positiva, não importando seu expoente.

Dentro desse contexto, o valor esperado (μ_{y_n}) e a variância ($\sigma_{y_n}^2$) da resposta na n -ésima condição experimental, empregadas na expressão 2.24, são fornecidas pelas expressões:

$$\mu_{y_n} = \mathbf{x}_n' \boldsymbol{\beta} \quad (2.27)$$

e

$$\sigma_{y_n}^2 = \exp\{\mathbf{z}_n' \boldsymbol{\theta}\}, \quad (2.28)$$

sendo que \mathbf{x}_n é um vetor formado pela n -ésima linha da matriz de planejamento \mathbf{X} , \mathbf{z}_n é um outro vetor formado pela n -ésima linha da matriz de planejamento \mathbf{X} , onde a primeira coluna é formada por elementos iguais a 1 (associados ao parâmetro constante) e as demais colunas, pelos contrastes associados aos fatores e interações. $\boldsymbol{\beta}$ e $\boldsymbol{\theta}$ são os vetores dos parâmetros das equações.

Embora o foco deste trabalho seja apenas a identificação de efeitos de dispersão, é importante especificar o modelo adotado para descrição do valor esperado e da variância para se conhecer o comportamento do que se está pretendendo avaliar.

2.4 TRANSFORMAÇÕES MATEMÁTICAS

No estudo da variabilidade em experimentos, as variâncias são estimadas a partir de resíduos quadráticos ou variâncias amostrais, utilizadas na análise como se fossem as respostas do experimento. As distribuições dessas medidas aproximam-se de uma distribuição qui-quadrado com um grau de liberdade no caso de resíduos quadráticos e $M-1$ graus de liberdade no caso de variâncias amostrais (M = quantidade de replicações). São distribuições assimétricas, apresentando alta probabilidade de ocorrência de valores positivos próximos de zero. Portanto, seus modelos teóricos não cumprem as suposições básicas dos modelos lineares: normalidade, homocedasticidade e linearidade.

Uma maneira de contornar tais problemas consiste em modificar a métrica da resposta por meio de uma transformação matemática, sendo a transformação logarítmica (Bartlett e Kendall, 1946) a mais utilizada nessas situações, pois, além de estabilizar a variância, torna a distribuição aproximadamente normal, além de poder propiciar o aparecimento de uma relação de linearidade entre fatores e resposta.

O gráfico (a) da figura 2.11 mostra a representação gráfica de observações obtidas em um experimento, no qual fica evidente a existência de heterocedasticidade e não-linearidade. Após uma transformação logarítmica, os dados são novamente plotados. O gráfico (b) da mesma figura mostra a correção dos problemas.

Há de se considerar, entretanto, no cálculo de resíduos e variâncias, a possibilidade de ocorrência de valores muito próximos de zero ou até mesmo nulos. Esses resultados, quando transformados pela função logarítmica, podem tornar-se valores discrepantes: negativos e de alta magnitude, denominados *inliers*.

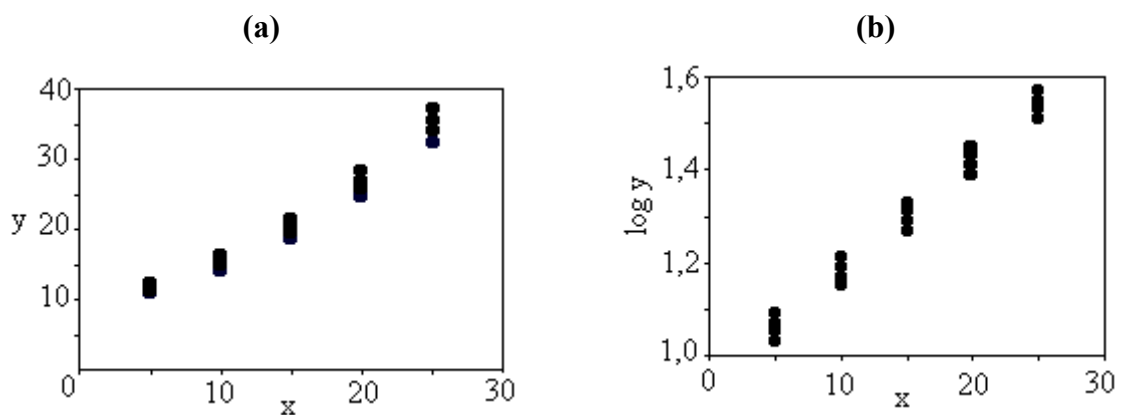


FIGURA 2.11 – Representação gráfica dos valores reais das observações (a) e dos valores transformados (b). (Adaptado de ZAR, 1999, p.355)

De acordo com Barbetta (1998), alguns autores, nesses casos, sugerem a utilização de alguns artificios, como adicionar uma constante aos resultados ou, simplesmente, eliminar os *inliers*. Em experimentos, porém, o último procedimento compromete a estrutura do projeto utilizado.

Vinning e Bohn (1998) sugerem a adição de uma unidade antes da aplicação da função logarítmica para se trabalhar sempre com valores positivos. Barbetta (1998) propõe uma outra transformação, denominada transformação logarítmica modificada, com a adição de uma constante, também antes da aplicação do logaritmo. De acordo com este autor, a transformação se mostrou mais eficiente que a transformação logarítmica padrão, principalmente para experimentos com poucas replicações, e é especialmente indicada para uso em experimentos com menos de cinco replicações.

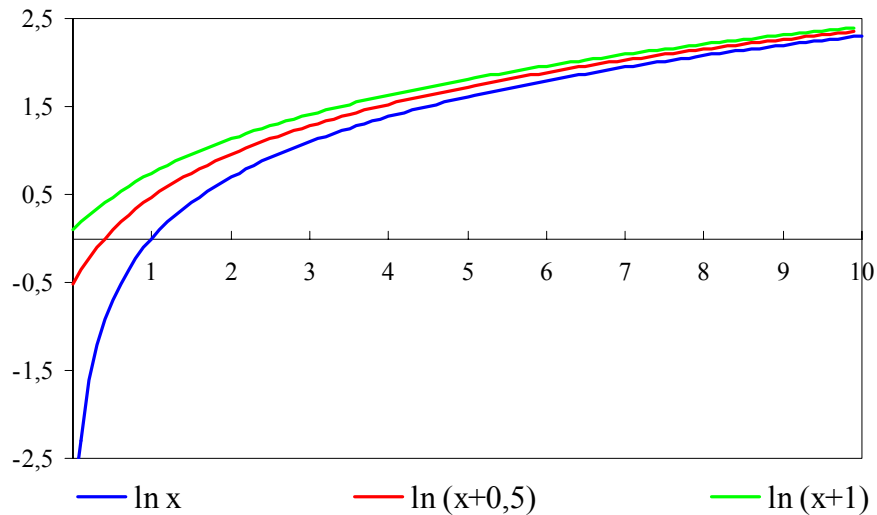


FIGURA 2.12 – Função logarítmica

A figura 2.12 mostra o comportamento da função logarítmica agindo sobre valores de x , $x+0,5$ e $x+1$, evidenciando que, ao adicionar uma constante ao valor da resposta considerada, diminui a probabilidade de aparecerem valores discrepantes devido à transformação.

Outras transformações como, $|y|$, $1/y$, \sqrt{y} , $y^{2/3}$, são, às vezes, utilizadas.

2.5 EFEITOS DE DISPERSÃO

Durante as duas últimas décadas, houve um enorme interesse no desenvolvimento de técnicas capazes de permitir a identificação de fatores com efeitos de dispersão, ou seja, fatores que afetam a variabilidade das observações, em decorrência da influência de fatores não controlados no processo produtivo (ruídos). Tais métodos se dividem em dois grupos, conforme o experimento seja replicado ou não.

Se o experimento apresentar replicações, a identificação dos fatores é relativamente simples, embora, na maioria dos casos, a quantidade de ensaios requerida seja consideravelmente grande. O mesmo não acontece se o experimento não apresentar replicações. Embora sejam necessários menos ensaios, a identificação se torna bastante difícil.

Inicialmente, a identificação de efeitos de dispersão foi proposta para experimentos com replicações (BARTLETT e KENDALL, 1946; NAIR e PREGIBON, 1988). Entretanto, em função da grande quantidade de ensaios requerida, pesquisadores concentraram seus esforços no desenvolvimento de métodos que possibilitassem a identificação em experimentos

não-replicados. Box e Meyer (1986a) foram os primeiros a propor um método para identificar efeitos de dispersão em experimentos fatoriais do tipo 2^{K-p} , o que foi muito bem aceito no meio industrial pela considerável redução de custos, relacionando-o com a etapa relativa ao projeto de parâmetros. Também são considerados importantes os métodos de Harvey (1976), Bergman e Hynén (1997) e Brenneman e Nair (2001). Observa-se que o trabalho de Harvey (1976) não foi realizado com o propósito de identificar efeitos de dispersão, mas corrigir a heterocedasticidade na análise de regressão.

Alguns dos métodos, propostos inicialmente para projetos do tipo 2^{K-p} , já foram generalizados para outros tipos de experimentos.

2.5.1 Experimentos replicados

Neste tipo de experimento, para avaliar efeitos de dispersão, são utilizadas as variâncias amostrais calculadas em cada condição experimental ensaiada.

Considere um experimento fatorial com replicações e fatores mensurados em dois níveis (projetos experimentais do tipo 2^K), onde:

$$N = 2^K, \quad (2.29)$$

$$\bar{y}_n = \frac{\sum_{m=1}^M y_{nm}}{M}, \quad (2.30)$$

$$s_n^2 = \frac{\sum_{m=1}^M (y_{nm} - \bar{y}_n)^2}{M - 1}, \quad (2.31)$$

sendo:

\bar{y}_n média das observações na n -ésima condição experimental;

s_n^2 variância amostral das observações na n -ésima condição experimental;

Considera-se ser a resposta desse experimento, y_{nm} , uma variável aleatória contínua com distribuição normal, seguindo o modelo apresentado pela equação 2.24. Admite-se também que o valor esperado dessa variável siga um modelo aditivo (equação 2.27), enquanto sua variância, um modelo multiplicativo (equação 2.28), conforme já mencionado.

2.5.1.1 Método S

Este método baseia-se nos mínimos quadrados ordinários, aplicado no logaritmo das variâncias amostrais. A utilização do logaritmo dessas estatísticas para estimar variabilidade foi proposta por Bartlett e Kendall (1946). Em projetos fatoriais do tipo 2^K , o efeito de dispersão (θ_k) do k -ésimo fator ou interação pode ser estimado pela estatística D_k^S , encontrada pela expressão (NAIR e PREGIBON, 1988):

$$D_k^S = \log \left(\frac{\prod_{n(k+)} s_n^2}{\prod_{n(k-)} s_n^2} \right)^{1/N} = \frac{1}{N} \left(\sum_{n(k+)} \log s_n^2 - \sum_{n(k-)} \log s_n^2 \right), \quad (2.32)$$

De acordo com os autores, esta estatística é um estimador de máxima verossimilhança do k -ésimo efeito de dispersão, θ_k (k -ésimo elemento do vetor θ do modelo 2.28), quando o modelo da variância for saturado, isto é, quando inclui todos os fatores e todas as interações, razão pela qual foi identificado por *S* (*Saturated*).

2.5.1.2 Método R

É uma extensão, para experimentos com replicações, do método de Box e Meyer (1986a), a ser apresentado na subseção seguinte, proposto para a situação de experimentos não-replicados e sob a suposição do efeito da esparsidade (poucos fatores ou interações apresentando efeito de dispersão ativo). Em experimentos fatoriais do tipo 2^K , o efeito de dispersão (θ_k) do k -ésimo fator ou interação pode ser calculado pela estatística D_k^R , que utiliza a expressão (NAIR e PREGIBON, 1988):

$$D_k^R = \frac{1}{2} \log \left(\frac{\sum_{n(k+)} s_n^2}{\sum_{n(k-)} s_n^2} \right) = \frac{1}{2} \left(\log \sum_{n(k+)} s_n^2 - \log \sum_{n(k-)} s_n^2 \right), \quad (2.33)$$

De acordo com Nair e Pregibon (1988), esta estatística é o estimador de máxima verossimilhança do k -ésimo efeito de dispersão, θ_k , quando o modelo da variância for restrito, isto é, não incluir todos os fatores e todas as interações, razão pela qual foi identificado por *R* (*Restricted*).

2.5.1.3 Considerações sobre os métodos *S* e *R*

Nair e Pregibon (1988) analisaram algumas características das estatísticas utilizadas por estes métodos em relação a viés, variância, poder e eficiência relativa. Neste estudo, os autores consideram que um estimador é dito não-viesado quando se distribui simetricamente em torno do parâmetro estimado. A análise em relação à variância, que pode ser obtida analiticamente, foi executada pela avaliação de sua relação com a forma de distribuição dos erros, assim como com o modelo estrutural. A comparação entre os métodos também foi feita pelo poder dos testes associados aos dois estimadores, assim como pela eficiência relativa, definida como o quociente entre o quadrado médio dos erros para vários modelos, projetos e graus de liberdade e estimadores de máxima verossimilhança do parâmetro. Eles concluíram que:

- o estimador D_k^S distribui-se simetricamente em torno de θ_k ($k = 2, \dots, K'$), razão pela qual não é viesado;
- o estimador D_k^R normalmente é viesado, embora ocorram exceções. Se existirem dois fatores: A e B , com efeitos de dispersão não-nulos, o estimador do efeito de dispersão da interação AB será viesado;
- a variância de D_k^S depende do tipo de distribuição do erro, enquanto a de D_k^R também depende da estrutura do modelo;
- se D_k^R for não-viesado e a quantidade e a intensidade de efeitos pequena, este estimador é mais eficiente do que D_k^S ;
- quando D_k^R apresentar um grande viés ou quando apresentar uma grande quantidade de fatores com considerável efeito de dispersão, é menos eficiente do que D_k^S ;
- D_k^S pode não ser eficiente quando existem poucos efeitos não-nulos, especialmente em grandes experimentos;
- D_k^R pode ser muito eficiente sob a suposição do efeito da esparsidade, não sendo recomendado para experimentos preliminares que normalmente envolvem uma quantidade grande de fatores;
- de maneira geral, o estimador D_k^R apresenta pior desempenho na identificação de efeitos de dispersão do que o estimador D_k^S , podendo não apenas perder efeitos importantes como identificar falsos efeitos.

2.5.2 Experimentos não-replicados

Neste tipo de experimento, a avaliação de efeitos de dispersão é realizada por meio de resíduos quadráticos, calculados em cada condição experimental ensaiada.

Seja o modelo:

$$y_n = \mu_{y_n} + \sigma_{y_n} \varepsilon_n, \quad (2.34)$$

em que:

y_n é a observação na n -ésima condição experimental;

μ_{y_n} é o valor esperado para a resposta na n -ésima condição experimental e satisfaz o modelo 2.27;

σ_{y_n} é a variância das resposta na n -ésima condição experimental e satisfaz o modelo 2.28;

ε_n é o erro experimental na n -ésima condição experimental com as suposições anteriormente enunciadas.

Considere um experimento fatorial não-replicado do tipo 2^K , em que:

$$N = 2^K, \quad (2.35)$$

$$\hat{y}_n = \mathbf{x}'_n \hat{\boldsymbol{\beta}}, \quad (2.36)$$

$$r_n^2 = (y_n - \hat{y}_n)^2, \quad (2.37)$$

sendo:

\hat{y}_n valor predito na n -ésima condição experimental;

\mathbf{x}_n vetor formado pela n -ésima linha da matriz de planejamento \mathbf{X} ;

$\hat{\boldsymbol{\beta}}$ estimativa do vetor de parâmetros $\boldsymbol{\beta}$ do modelo 2.27;

r_n^2 resíduo quadrático na n -ésima condição experimental.

2.5.2.1 Método H

O método é apenas uma extensão da idéia de Bartlett e Kendall (1946), relativa a experimentos replicados, com a substituição das variâncias amostrais por resíduos quadráticos. Com a suposição de modelo log-linear, a avaliação desse modelo é feita a partir

da razão entre médias geométricas. Embora seja denominado de método H , de Harvey, não foi proposto por este autor. Sua estratégia sobre heteroscedasticidade em problemas econométricos foi estendida para a área de identificação de efeitos de dispersão. A estratégia, apresentada em Harvey (1976), em um contexto de análise de regressão, foi denominada de procedimento em duas etapas (PDE): na primeira, são estimados os efeitos de locação pelo método dos mínimos quadrados ordinários (MQO) e, na segunda, os efeitos de dispersão com a utilização de resíduos quadráticos.

A estatística que avalia a dispersão em experimentos do tipo 2^K , por meio desse método, D_k^H , utiliza a seguinte expressão (BRENNEMAN e NAIR, 2001) para avaliar o efeito do k -ésimo fator (θ_k):

$$D_k^H = \frac{1}{N} \left(\sum_{n(k+)} \log r_n^2 - \sum_{n(k-)} \log r_n^2 \right) = \log \left(\frac{\prod_{n(k+)} r_n^2}{\prod_{n(k-)} r_n^2} \right)^{1/N}, \quad (2.38)$$

em que r_n^2 é o resíduo quadrático na n -ésima condição experimental.

Quando a variância segue o modelo log-linear, o mais lógico seria encontrar estimativas dos efeitos de dispersão por meio de medidas que também apresentassem modelo log-linear, ou seja, utilizar o método H que trabalha com o logaritmo de resíduos quadráticos. Entretanto, de acordo com Brenneman e Nair (2001), esse estimador é viesado e o viés está relacionado à estrutura do modelo da variância, definido a partir dos efeitos de dispersão; à estrutura do modelo de locação, definido a partir dos efeitos de locação, assim como à adequação do modelo, diminuindo com o aumento do tamanho do experimento. Entretanto, a estimativa do efeito de dispersão de um fator A , por esse método, será viesada, se o fator B e a interação AB apresentarem efeitos de dispersão ativos.

De acordo com Brenneman e Nair (2001), é possível contornar o último com a utilização de resíduos modificados, calculados sobre um modelo da média expandido. Para calcular o efeito de dispersão do k -ésimo fator, é definido um novo conjunto de fatores e interações que entram no modelo da média. O conjunto é composto pelos elementos identificados inicialmente pelo método dos mínimos quadrados ordinários (MQO), pelo k -ésimo fator e por todas as interações do k -ésimo fator com os elementos identificados como tendo efeito de locação. Por exemplo, suponha-se ter sido identificado o seguinte conjunto de efeitos significativos no modelo da média: $\mathcal{L} = \{I, A, B\}$, sendo I associado à média global e A e B , dois fatores do experimento. Para calcular o efeito de dispersão do fator C , os resíduos são calculados em relação a um modelo com o seguinte conjunto de efeitos: $\mathcal{L}_E = \{I, A, B, C$,

$AC, BC\}$. Se for utilizado esse tipo de resíduo, que diminui o viés, o método é referenciado como *HM* (Harvey modificado).

Há de se considerar também que, se avaliados com rigor, esses métodos falham quando pelo menos um dos resíduos é zero, embora existam artifícios matemáticos para contornar o problema.

2.5.2.2 Método *BM*

Box e Meyer (1986a) apresentam um método para identificar efeitos de dispersão em experimentos fatoriais fracionados não-replicados, mensurados em dois níveis, baseado no princípio da esparsidade, por considerar que grandes efeitos ocorrem em função de um pequeno número de fatores testados.

Para identificá-los, os autores sugerem a possibilidade de calcular, para cada fator e interação, as variâncias amostrais das observações, separadamente, para o nível mais alto e para o nível mais baixo, encontrando a razão entre elas.

$$D_k^{BM(0)} = \log \left[\frac{s_{k+}^2}{s_{k-}^2} \right] \quad (2.39)$$

sendo:

- $D_k^{BM(0)}$ uma estimativa do efeito do fator ou interação do k -ésimo fator ou interação (parâmetro θ_k);
- s_{k+}^2 a variância das observações no nível superior do k -ésimo fator ou interação;
- s_{k-}^2 a variância das observações no nível inferior do k -ésimo fator ou interação.

Embora as razões das variâncias apresentem distribuição F , com a transformação logarítmica passam a ter distribuição aproximadamente normal. Para serem interpretados, os resultados podem ser plotados em um diagrama de pontos.

Para o método *BM* poder realmente detectar efeitos de dispersão, é necessário serem nulos todos os efeitos de locação pois, caso contrário, podem interferir nos resultados dos efeitos de dispersão que ficam superestimados. Para contornar o problema, os autores sugerem que o procedimento seja executado a partir de resíduos, ou seja, o que sobra depois da remoção dos efeitos de locação pelo método dos mínimos quadrados. Nesse caso, a identificação dos efeitos de dispersão seguiria a seguinte metodologia:

- (a) identificar os fatores com efeitos de locação e modelar o valor esperado;
- (b) calcular os resíduos ($r_n = y_n - \hat{y}_n$), encontrando a diferença entre cada resposta (y_n) e o respectivo valor predito (\hat{y}_n);
- (c) estimar as variâncias dos resíduos, separadamente, para o nível mais alto e para o nível mais baixo de cada fator e interação;
- (d) encontrar o logaritmo da razão entre as estimativas das variâncias para cada fator e plotá-los em um diagrama de pontos, interpretando os resultados.

O logaritmo do quociente entre a soma dos quadrados dos resíduos quadráticos foi, posteriormente, identificado como estatística D_k^{BM} e, para o k -ésimo fator, a estimativa do efeito de dispersão (θ_k) pode ser encontrado pela expressão (adaptada de BRENNEMAN e NAIR, 2001):

$$D_k^{BM} = \frac{1}{2} \left(\log \sum_{n(k+)} r_n^2 - \log \sum_{n(k-)} r_n^2 \right) = \frac{1}{2} \log \left[\frac{\sum_{n(k+)} r_n^2}{\sum_{n(k-)} r_n^2} \right], \quad (2.40)$$

Montgomery (1997a) complementou o método, sugerindo interpretar os resultados da estatística pelo gráfico de probabilidade normal pois, se a variância dos erros experimentais associados ao nível mais alto de um fator não for diferente da variância dos erros experimentais associados ao nível mais baixo do mesmo fator, então o logaritmo dessa razão tem distribuição normal e os pontos, quando plotados num gráfico desse tipo, se distribuirão em torno de uma linha reta.

Box e Meyer (1986a) ainda comentam que, depois de os modelos para o valor esperado e para a variância terem sido identificados, estimativas mais eficientes podem ser encontradas pelo refinamento dos modelos obtidos. Modelos para os efeitos de locação podem ser obtidos pelo método dos mínimos quadrados generalizados (MQG) ou, mais especificamente, método dos mínimos quadrados ponderados (MQP), sendo os pesos definidos a partir de estimativas das variâncias (peso = inverso da variância), obtidas por meio de um modelo de regressão para a variância.

Se o processo for realizado várias vezes, de forma iterativa, os resultados tendem a convergir para um ponto estacionário, conforme atestam Hartley e Jayatillake (1973) citado em Box e Meyer (1986a). As estimativas dos efeitos de locação e dispersão gerados por esse processo são assintoticamente iguais às estimativas de máxima verossimilhança.

Ao apresentar o método, entretanto, os autores salientam que a estratégia proposta pode ser utilizada em um estágio preliminar para identificar, de forma econômica, poucos fatores com efeitos de locação e com efeitos de dispersão, os quais devem ser estudados com muito cuidado em outros experimentos, reconhecendo que a estratégia proposta apresenta limitações.

Brenneman e Nair (2001) chamam a atenção para o fato de os autores não fazerem referência à estrutura do modelo utilizado para descrever a variância: aditiva (linear) ou multiplicativa (log-linear), sobre a qual o método foi desenvolvido. Tais autores, entretanto, após analisar as propriedades desta estatística, concluem apresentar esse método resultados viesados, tanto para modelos multiplicativos como aditivos, acontecendo não só pela correlação existente entre os resíduos, como também pela estrutura do modelo. Consideram que o viés estrutural é sério (a estimativa do efeito de dispersão de um fator A será viesada se o fator B e a interação AB apresentarem efeitos de dispersão ativos), mesmo que sejam utilizadas replicações e o tamanho do experimento seja grande, dependendo muito da adequação do modelo de locação.

2.5.2.3 Método BH

Bergman e Hynén (1997) descreveram o método por eles proposto de três maneiras. A mais simples encontra a estatística do método por meio de resíduos quadráticos modificados, já comentados na seção 2.5.2.1. Os resíduos são extraídos de um modelo expandido da média, de forma a eliminar a dependência entre os resíduos do numerador e do denominador. Nesse caso, a estatística que estima o efeito de dispersão do k -ésimo fator, $D_k^{BH(0)}$, assume uma forma semelhante à expressão utilizada pelo método BM e é dada por:

$$D_k^{BH(0)} = \frac{\sum_{n(k+)} \tilde{r}_n^2}{\sum_{n(k-)} \tilde{r}_n^2}, \quad (2.41)$$

onde \tilde{r}_n^2 é o resíduo quadrático modificado na n -ésima condição experimental.

Este procedimento de identificação de efeitos de dispersão permite, então, os dois grupos de resíduos serem independentes, o que, tecnicamente, é bastante vantajoso. Segundo seus autores, o método também apresenta a vantagem de, se os efeitos de dispersão forem corretamente identificados, a estatística proposta apresentará distribuição $F(t, t)$, na qual t é o

número de fatores ou interações identificados como apresentando efeito de dispersão, o que diminui a subjetividade do método.

Na prática, entretanto, é usual aplicar uma transformação logarítmica a tais resultados para seguirem uma distribuição aproximadamente normal, sendo os efeitos ativos identificados pelo gráfico de probabilidade normal. Nesse caso, a estatística que estima o efeito de dispersão do k -ésimo fator, D_k^{BH} , é dada por:

$$D_k^{BH} = \frac{1}{2} \left(\log \sum_{n(k+)} \tilde{r}_n^2 - \log \sum_{n(k-)} \tilde{r}_n^2 \right) \quad (2.42)$$

Assim como Box e Meyer (1986a), os autores salientam a possibilidade de o método ser utilizado em um estágio preliminar para identificar fatores com efeitos de dispersão, a serem estudados com muito cuidado em experimentos posteriores, reconhecendo que a estratégia necessita ser mais investigada.

Brenneman e Nair (2001), por meio de contra-exemplos, mostram que nem sempre a estatística apresentada possui uma distribuição exata F , fato verdadeiro tanto para o modelo aditivo como para o modelo multiplicativo da variância. Essa estatística só apresentará distribuição F se a hipótese nula considerar nulos todos os outros efeitos de dispersão.

Salienta-se que, se o modelo de locação expandido coincide com o modelo de locação, o método BH coincide com o método BM e isto faz o primeiro (BH) sofrer o mesmo tipo de viés estrutural do segundo (BM), ou seja, quando se testa a existência de efeito de dispersão em um fator, considera-se todos os demais não possuírem efeito de dispersão, o que é uma suposição muito forte. Brenneman e Nair (2001), confirmam não desaparecer esse tipo de viés com o aumento do tamanho do experimento.

2.5.2.4 Estratégia BN

Brenneman e Nair (2001), a partir de uma análise crítica dos métodos mais empregados para identificar efeitos de dispersão em experimentos não-replicados, concluíram que, embora todos os métodos não-iterativos analisados sofram um certo grau de viés, alguns apresentam problemas mais sérios do que outros. Baseados em uma análise realizada por meio de métodos analíticos e estudos de simulação, eles propuseram duas estratégias iterativas para identificar e estimar efeitos de dispersão, uma delas supondo modelo multiplicativo para a variância.

Os autores consideram que, se a variância apresenta modelo multiplicativo, o procedimento mais adequado para analisar efeitos de dispersão é o uso do logaritmo dos resíduos quadráticos (método H). Entretanto, se efeitos de dispersão existem, a remoção dos efeitos de locação fica prejudicada. Isto pode ser resolvido pela utilização do método HM que utiliza resíduos modificados. Esse procedimento, entretanto, além de necessitar de modelos de locação diferentes para a avaliação de cada efeito de dispersão, pode, eventualmente, levar a um modelo expandido muito grande, que forneça resíduos quadráticos nulos.

Um outro problema do método HM ocorre quando um efeito de dispersão A é estimado ao serem ativos os efeitos de dispersão de B e AB : o efeito não-ativo de A pode ser detectado, assim como o efeito ativo de A pode não ser detectado. Contorna-se o primeiro problema com a utilização do método de máxima verossimilhança (MV) ou método de máxima verossimilhança restrita (MVR). O segundo problema é resolvido com uma ampliação do conjunto de fatores com efeito de dispersão, acrescentando aos fatores identificados inicialmente, os fatores resultantes de sua interação. Por exemplo, se B e AB foram identificados como ativos, A deve ser incluído no conjunto porque $B*(A*B)=A$, sendo o símbolo $*$, o operador de interação.

Em experimentos não-replicados, a estratégia proposta por Brenneman e Nair (2001), para identificar efeitos de locação e efeitos de dispersão com estrutura multiplicativa é a seguinte:

- a) ajustar um modelo de locação pelo método dos mínimos quadrados ordinários (MQO), identificando o conjunto \mathcal{L} de efeitos de locação;
- b) se \mathcal{L} contém um conjunto fechado \mathcal{L}' (diz-se que um conjunto é fechado quando as interações entre os elementos que pertencem ao conjunto também lhe pertencem) com $N/2$ ou mais elementos, então a identificação dos efeitos de dispersão é difícil, sendo necessário aumentar o número de ensaios (alterar o projeto experimental);
- c) se não for assim, usar o método de Harvey modificado para estimar os efeitos de dispersão e identificar os ativos (por exemplo, por meio do gráfico de probabilidade normal).

Seja \mathcal{D}^0 o conjunto de efeitos de dispersão ativos identificados na etapa (c):

- d) se um dos efeitos k não foi identificado como ativo, mas corresponde à interação de dois outros efeitos k_1 e k_2 , identificados como ativos, tal que $k = k_1 * k_2$, incluí-lo no conjunto dos efeitos ativos (\mathcal{D}^0);

- e) com os efeitos de dispersão \mathcal{D}^0 , construir uma equação para a variância usando o modelo log-linear $[\log(\sigma_n^2) = z'_n \theta]$, no qual as estimativas dos elementos de θ podem ser obtidas por MQO, visto que a transformação logarítmica estabiliza a variância;
- f) refazer o modelo de locação usando mínimos quadrados generalizados (MQG), atribuindo a cada observação um peso fornecido pelo inverso da variância predita pelo modelo definido no passo (e);
- g) novos resíduos e novos efeitos de dispersão são calculados. Efeitos de dispersão não identificados como ativos poderão ser eliminados e as estimativas refeitas;
- h) os passos (e), (f) e (g) poderão ser repetidos até o processo se estabilizar. Alguns autores sugerem apenas uma ou duas iterações, devido ao problema da convergência na presença de valores discrepantes.

Essa estratégia sugere, então, utilizar um método iterativo para refinar o modelo e estimar os parâmetros de forma mais eficiente. Os estudos de simulação realizados por Brenneman e Nair (2001) mostraram que a estratégia funciona bem quando se verifica o princípio da esparsidade. A estratégia utilizada superestima o modelo inicial de dispersão para acomodar situações nas quais o método pode apresentar viés. Os métodos de mínimos quadrados ponderados ou de máxima verossimilhança subsequentes eliminam falsos efeitos identificados inicialmente.

METODOLOGIA

Este capítulo descreve a metodologia utilizada no estudo de simulação realizado para comparar os métodos usados para identificar efeitos de dispersão.

3.1 CARACTERÍSTICAS DO ESTUDO DE SIMULAÇÃO

O presente estudo faz uma avaliação do comportamento de seis métodos propostos para a identificação de efeitos de dispersão: *R*, *S*, *H*, *BM*, *HM* e *BH*, na situação de experimentos com poucas replicações. Eles avaliam a existência de efeitos de dispersão por meio de médias aritméticas ou geométricas entre variâncias amostrais, resíduos quadráticos ou resíduos quadráticos modificados.

O estudo foi realizado seguindo a metodologia de planejamento de experimentos e simulações de Monte Carlo, não só pela facilidade de compreender suas idéias básicas e operacionalizar os diversos procedimentos, como pela possibilidade de avaliar resultados obtidos sob diversas condições experimentais simuladas.

Uma simulação tenta reproduzir um sistema num ambiente controlado para observar seu comportamento, possibilitando que algumas de suas características possam ser alteradas, para serem observadas as conseqüências dessas mudanças, sem o sistema sob investigação sofrer qualquer perturbação ou até mesmo, realmente existir.

O desenvolvimento da Engenharia da Computação permitiu ampla utilização desse tipo de estudo em várias áreas da ciência. Para a Estatística, são especialmente importantes as simulações envolvendo variáveis aleatórias, por possibilitar a avaliação do desempenho de novos métodos, a comparação do comportamento de diferentes métodos, a geração de amostras pseudo-aleatórias, a seleção de amostras aleatórias etc.

Além de serem mais fáceis e, na maioria das vezes, também mais rápidas do que os métodos analíticos, as simulações não requerem tantas simplificações da situação real. A construção de modelos eficientes requer, entretanto, um profundo conhecimento do sistema investigado e uma certa habilidade.

As simulações de variáveis aleatórias deram origem aos chamados Métodos de Monte Carlo, que necessitam de um gerador de números aleatórios equiprováveis. O nome do método originou-se de uma cidade localizada no Principado de Mônaco, a cidade de Monte Carlo, famosa por seus cassinos com grande quantidade de roletas, um dos mecanismos mais simples de geração de números aleatórios.

3.2 DESCRIÇÃO DO ESTUDO DE SIMULAÇÃO

3.2.1 Projeto experimental estudado

O estudo foca um experimento com quatro fatores ($K=4$): x_1 , x_2 , x_3 e x_4 , ensaiados a dois níveis (-1 e $+1$), segundo um projeto fatorial completo do tipo 2^4 , com replicações, contendo efeitos de locação e efeitos de dispersão. Contudo, esse projeto simulado pode representar, em termos práticos, por exemplo, um projeto do tipo 2^{5-1} ou do tipo 2^{6-2} , com os cuidados naturais na análise dos confusões entre efeitos.

Os resultados do experimento foram gerados por uma distribuição normal com média μ_y , descrita por uma função aditiva ou linear, e variância σ_y^2 , descrita por uma função multiplicativa ou log-linear. Para o estudo em foco, um projeto 2^4 , se forem considerados todos os efeitos principais e interações possíveis, a média e a variância são descritas, respectivamente, por:

$$\mu_y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_{12} x_1 x_2 + \beta_{13} x_1 x_3 + \beta_{14} x_1 x_4 + \beta_{23} x_2 x_3 + \beta_{24} x_2 x_4 + \beta_{34} x_3 x_4 + \beta_{123} x_1 x_2 x_3 + \beta_{124} x_1 x_2 x_4 + \beta_{134} x_1 x_3 x_4 + \beta_{234} x_2 x_3 x_4 + \beta_{1234} x_1 x_2 x_3 x_4 \quad (3.1)$$

e

$$\sigma_y^2 = \exp \left\{ \begin{array}{l} \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \theta_3 x_3 + \theta_4 x_4 + \theta_{12} x_1 x_2 + \theta_{13} x_1 x_3 + \theta_{14} x_1 x_4 + \theta_{23} x_2 x_3 + \theta_{24} x_2 x_4 \\ \theta_{34} x_3 x_4 + \theta_{123} x_1 x_2 x_3 + \theta_{124} x_1 x_2 x_4 + \theta_{134} x_1 x_3 x_4 + \theta_{234} x_2 x_3 x_4 + \theta_{1234} x_1 x_2 x_3 x_4 \end{array} \right\}, \quad (3.2)$$

sendo β_0 , β_1 , β_2, \dots , β_{1234} e θ_0 , θ_1, \dots , θ_{1234} , os coeficientes das equações, que representam, respectivamente, os efeitos de locação e os efeitos de dispersão. Na maioria dos casos, vários desses valores são nulos, principalmente quando se referem a interações. O projeto simulado adotou apenas alguns efeitos não nulos, conforme será descrito em 3.2.2.

3.2.2 Projeto experimental para a simulação

Conforme já mencionado, nas simulações do tipo Monte Carlo, as amostras são geradas a partir de parâmetros especificados. Entretanto, para o estudo não se restringir a uma única condição especificada, são consideradas várias alternativas experimentais, como: diferente quantidade de efeitos de locação, de efeitos de dispersão e de replicações; sempre variando conforme um projeto experimental adequado. Pela pesquisa bibliográfica realizada sobre o assunto, suspeita-se que seis fatores podem alterar o desempenho dos métodos em estudo. Tais fatores, descritos a seguir, são codificados por A , B , C , D , E e F e sintetizados no quadro 3.1.

O fator A é utilizado para manipular a quantidade de fatores com efeitos de locação. Em um dos níveis (-1), considera-se que dois fatores possuem efeitos principais de locação,

enquanto no outro (+I), três fatores com efeitos principais e duas interações. As equações das médias (μ_y) utilizadas para a geração das amostras são, respectivamente:

$$\mu_y = \beta_1 x_1 + \beta_2 x_2 \quad (\text{Fator } A \text{ no nível } -I) \quad (3.3)$$

e

$$\mu_y = \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{12} x_1 x_2 + \beta_{13} x_1 x_3 \quad (\text{Fator } A \text{ no nível } +I). \quad (3.4)$$

O fator *B* é utilizado para definir a intensidade dos efeitos de locação. No nível $-I$, a intensidade adotada para os efeitos principais é 2, enquanto, no nível $+I$, é 4. Nos dois níveis considera-se que os efeitos das interações apresentam metade da intensidade dos efeitos principais.

O fator *C* é usado para manipular a quantidade de fatores com efeitos de dispersão. Em um dos níveis ($-I$) é considerado que apenas um fator apresenta efeito de dispersão, enquanto no outro ($+I$), dois fatores. As equações da variância (σ_y^2) utilizadas para a geração das amostras são, respectivamente:

$$\sigma_y^2 = \exp\{\theta_k x_k\} \quad (\text{Fator } C \text{ no nível } -1) \quad (3.5)$$

e

$$\sigma_y^2 = \exp\{\theta_k x_k + \theta_2 x_2\} \quad (\text{Fator } C \text{ no nível } +1) \quad (3.6)$$

em que o subíndice *k* é 1 ou 4, conforme descrito no fator *D*.

QUADRO 3.1 – Descrição de fatores e níveis utilizados no programa computacional utilizado na simulação realizada para comparar o desempenho dos métodos investigados

Fator	Descrição	Nível (-1)	Nível (+1)
-------	-----------	------------	------------

<i>A</i>	Fatores com efeitos de locação	β_1 e β_2 (dois efeitos principais)	$\beta_1, \beta_2, \beta_3, \beta_{12}$ e β_{13} (três efeitos principais e duas interações)
<i>B</i>	Intensidade dos efeitos de locação	$\beta_i = 2$ e $\beta_{ij} = 1$ para i e j iguais aos índices especificados no fator A $\beta_i = 0$ e $\beta_{ij} = 0$ para i e j diferentes dos índices especificados no fator A	$\beta_i = 4$ e $\beta_{ij} = 2$ para i e j iguais aos índices especificados no fator A $\beta_i = 0$ e $\beta_{ij} = 0$ para i e j diferentes dos índices especificados no fator A
<i>C</i>	Fatores com efeitos de dispersão	θ_k (um efeito de dispersão), sendo $\theta_2 = 0$	θ_k e $\theta_2 = 0,643$ (dois efeitos de dispersão)
<i>D</i>	Nível de coincidência entre efeito de locação e dispersão	$\theta_k = \theta_1$	$\theta_k = \theta_4$
<i>E</i>	Intensidade dos efeitos de dispersão	$\theta_k = 0,549$ e $\theta_i = 0$ para $i \neq k$ e $i \neq 2$	$\theta_k = 0,896$ e $\theta_i = 0$ para $i \neq k$ e $i \neq 2$
<i>F</i>	Quantidade de replicações	2	4

O fator *D* refere-se ao nível de coincidência entre efeitos de locação e de dispersão. Quando o fator *C* está no nível -1 , o experimento apresenta apenas um fator com efeitos de dispersão e, neste caso, se $k = 1$ (equação 3.5), o fator com efeito de dispersão apresenta também efeito de locação; se $k = 4$, o fator com efeito de dispersão não apresenta efeito de locação. Quando o fator *C* está no nível $+1$, o experimento apresenta dois fatores com efeitos de dispersão. Neste caso, se $k = 1$ (equação 3.6), todos os fatores que apresentam efeito de dispersão (x_1 e x_2) também apresentam efeito de locação; se $k = 4$, apenas um dos fatores (x_2) apresenta simultaneamente efeito de dispersão e efeito de locação, enquanto que o outro (x_4), apenas efeito de dispersão.

O fator *E* é utilizado para definir a intensidade dos efeitos de dispersão. Em um dos níveis (-1) considera-se que o efeito principal do k -ésimo fator é $0,549$, enquanto no outro, o efeito principal do mesmo fator é $0,849$. No caso de existirem dois fatores com efeitos de dispersão, o efeito principal do fator x_2 é considerado igual a $0,643$. Os valores $0,549$, $0,643$ e

0,849 foram escolhidos por indicar uma alteração no valor da variância, de um nível para outro, de três, quatro e seis vezes, respectivamente.

O fator F trata da quantidade de replicações: duas ou quatro.

Os níveis dos seis fatores do projeto de simulação são combinados segundo um projeto fatorial fracionado 2^{6-1} , isto é, seis fatores (A, B, C, D, E e F), ensaiados a dois níveis (-1 e $+1$), com grau de fracionamento 1 , originando uma matriz de planejamento com 32 combinações entre os níveis dos cinco primeiros fatores. O fracionamento foi feito considerando $F = ABCDE$.

3.3 DESCRIÇÃO DO PROGRAMA COMPUTACIONAL

Para executar a primeira fase do estudo, cria-se um programa computacional desenvolvido no *proc IML (Interativa Matrix Language)* do *software* estatístico *SAS*, versão 8.0 (*Statistical Analysis System for windows*), possuidor de uma linguagem de programação com recursos de álgebra matricial, ferramenta básica computacional para o tratamento de espaços vetoriais, que está apresentado na página [www](http://www.sas.com). .

O programa, desenvolvido de acordo com o projeto lógico apresentado na figura 3.1, tem como objetivo encontrar as proporções de identificação correta e de identificação falsa de efeitos de dispersão encontrados pelos métodos R, S, H, BM, HM e BH , sendo desenvolvido em quatro etapas, descritas a seguir.

3.3.1 Primeira etapa: planejamento do experimento

Consiste na construção da matriz do projeto de simulação para os fatores de A a F do experimento das condições de simulação, segundo um projeto 2^{6-1} , com 32 condições experimentais.

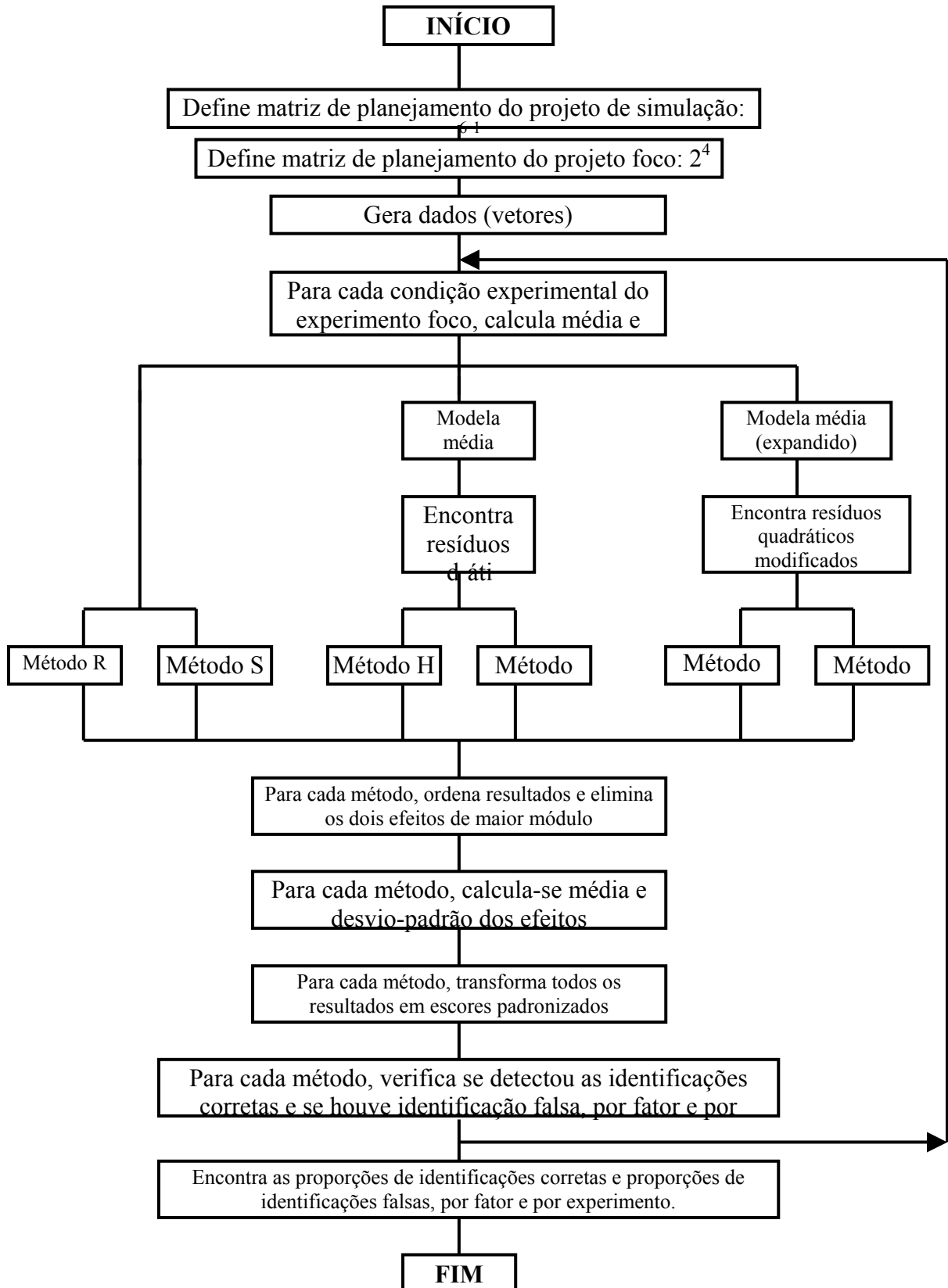


FIGURA 3.1 - Fluxograma do programa computacional utilizado na simulação realizada para comparar o desempenho dos métodos investigados

3.3.2 Segunda etapa: geração de amostras

Esta etapa se inicia com a construção da matriz de planejamento do projeto foco do estudo, um projeto do tipo 2^4 com 16 condições experimentais. Os elementos da matriz são utilizados na construção de diferentes equações da média e da variância. Para gerar as respostas dos experimentos são acrescentados números aleatórios provindos de uma distribuição normal com média zero e variância unitária $[N(0;1)]$.

As possíveis combinações entre os níveis dos fatores A e B do projeto de simulação 2^{6-1} definem as equações geradoras do valor médio (equações 3.3 e 3.4), enquanto os valores de x_1 , x_2 , x_3 e x_4 , que aparecem nessas equações, dependem dos níveis assumidos por tais fatores em cada condição experimental do projeto foco 2^4 .

As possíveis combinações entre os níveis dos fatores C , D e E do projeto de simulação 2^{6-1} definem as equações geradoras da variância (equações 3.5 e 3.6), enquanto os valores de x_1 , x_2 , x_3 e x_4 , existentes nas equações, dependem dos níveis assumidos pelos fatores em cada condição experimental do projeto foco 2^4 .

A partir destas equações, para cada condição experimental do projeto de simulação 2^{6-1} , são gerados de acordo com uma distribuição normal, vetores de dimensão $N \times 1$, com as respostas dos experimentos. Eles são agrupados em matrizes de dimensão $N \times M$, nas quais N representa a quantidade de condições experimentais e M , a quantidade de replicações. Nessas matrizes, a m -ésima replicação da n -ésima condição experimental (y_{nm}) é dada por:

$$y_{nm} = \mu_{y_n} + \sigma_{y_n} \cdot e_{nm} \quad (3.7)$$

onde μ_{y_n} e σ_{y_n} são, respectivamente, o valor esperado e o desvio-padrão da n -ésima condição experimental e e_{nm} , o m -ésimo erro experimental da n -ésima condição experimental gerado por uma distribuição $N(0, 1)$.

Usam-se os valores simulados para encontrar proporções de identificação correta e proporções de identificação falsa de efeitos de dispersão (mais detalhes na subseção 3.3.4). Como as variáveis-resposta em estudo são proporções, suas variâncias são dadas por:

$$Var(\hat{P}) = \frac{P(1-P)}{Q} \leq \frac{1}{4Q}, \quad (3.8)$$

em que P representa proporção que se deseja estimar e Q , o tamanho da amostra que, no presente estudo, corresponde à quantidade de amostras simuladas.

Nesse estudo de simulação, para cada uma das 32 situações experimentais, foram simulados 5.000 experimentos do tipo 2^4 . Assim, o cálculo de identificações corretas e falsas

baseou-se em amostras de tamanho $Q = 5.000$. Em conseqüência, o valor máximo das variâncias amostrais é $0,00005$. Essa quantidade de amostras simuladas possibilita estimar valores de probabilidade de identificação correta e de identificação falsa de efeitos de dispersão com um erro máximo de $0,0138$, associados ao nível de confiança de $0,95$.

3.3.3 Terceira etapa: identificação de fatores com efeitos de dispersão

O procedimento descrito a seguir é realizado em cada um dos 5.000 projetos foco 2^4 simulados.

O processo de identificação de fatores com efeitos de dispersão, no projeto foco 2^4 , começa com o cálculo da média e da variância amostral em cada condição experimental desse projeto, utilizando as expressões 3.9 e 3.10, respectivamente, para calcular tais medidas na n -ésima condição experimental. Assim:

$$\bar{y}_n = \frac{\sum_{m=1}^M y_{nm}}{M} \quad (3.9)$$

e

$$s_n^2 = \frac{\sum_{m=1}^M (y_{nm} - \bar{y}_n)^2}{M - 1}, \quad (3.10)$$

sendo:

y_{nm} a observação na m -ésima replicação da n -ésima condição experimental ($n = 1, \dots, N$;

$m = 1, \dots, M$);

\bar{y}_n a média das observações da n -ésima condição experimental ($n = 1, \dots, N$);

s_n^2 a variância amostral das observações da n -ésima condição experimental ($n = 1, \dots, N$).

A modelagem da média é realizada por meio do método dos mínimos quadrados ordinários (MQO) para estimar um vetor de parâmetros β , de dimensão $N \times 1$, feito de acordo com a expressão:

$$\hat{\beta} = (X'X)^{-1} \cdot X' \bar{Y} \quad (3.11)$$

onde X é uma matriz de planejamento de dimensão $N \times 2$ ou $N \times 5$, conforme o nível do fator A , e \bar{Y} é um vetor de dimensão $N \times 1$ composto com os valores de \bar{y}_n (equação 3.9).

Na execução do procedimento, consideram-se conhecidos os fatores que apresentam efeito de locação, mas seus valores são estimados com base na amostra simulada.

Estimado o vetor $\hat{\beta}$, é possível encontrar estimativas dos valores esperados para cada condição experimental (\hat{Y}) por:

$$\hat{E}(Y) = \hat{Y} = \hat{\beta} \cdot X \quad (3.12)$$

em que X é uma matriz de dimensão $N \times 2$ ou $N \times 5$, conforme o nível do fator A , e $\hat{\beta}$ é o vetor estimado pela equação 3.11.

A partir dos valores preditos, são encontrados os resíduos quadráticos por meio da expressão 3.13. Considerando a perda dos graus de liberdade na estimação do modelo da média, corrigem-se os resíduos por meio da expressão apresentada a seguir, conforme proposto por Carroll e Ruppert (1988, p.78):

$$r_{nm}^2 = \frac{(\hat{y}_n - y_{nm})^2}{1 - h_n} \quad (3.13)$$

em que $h_n = x_n (X'X)^{-1} x_n'$. A divisão por $1 - h_n$ torna os resíduos quadráticos estimadores não-viciados das correspondentes variâncias. Nessas expressões:

\hat{y}_n é a estimativa do valor esperado da n -ésima condição experimental ($n = 1, \dots, N$);

r_{nm}^2 é o resíduo quadrático na m -ésima replicação da n -ésima condição experimental ($m = 1, \dots, M; n = 1, \dots, N$);

h_n é o coeficiente de correção ($n = 1, \dots, N$);

X é uma matriz de planejamento;

x_n é um vetor da matriz X , que corresponde à n -ésima condição experimental ($n = 1, \dots, N$).

Também foram encontrados os resíduos quadráticos modificados, calculados a partir do modelo expandido da média. A expressão utilizada para encontrá-los é a mesma citada anteriormente, sendo alterada, entretanto, a composição do conjunto de fatores com efeitos de locação. Além dos fatores com efeito de locação, acrescentam-se, ao modelo da média, o fator em estudo e todas as interações de segunda ordem, dele, com os fatores identificados como

apresentando efeito de locação. Isto significa que várias modelagens pelo método MQO são realizadas, uma para cada fator e para cada interação em estudo.

A partir dos resultados obtidos para variâncias amostrais, resíduos quadráticos e resíduos quadráticos modificados, os efeitos de dispersão são estimados pelos métodos S (de acordo com equação 2.33), R (de acordo com equação 2.34) e pelos métodos H , HM , BM e BH , adaptados para a situação de experimentos com poucas replicações, empregando as expressões:

$$\text{Método } H - D_k^H = \frac{1}{N} \left(\sum_{n(k+)} \log \bar{r}_n^2 - \sum_{n(k-)} \log \bar{r}_n^2 \right) \quad (3.14)$$

$$\text{Método } BM - D_k^{BM} = \frac{1}{2} \left(\log \sum_{n(k+)} \bar{r}_n^2 - \log \sum_{n(k-)} \bar{r}_n^2 \right) \quad (3.15)$$

$$\text{Método } HM - D_k^{HM} = \frac{1}{N} \left(\sum_{n(k+)} \log \bar{r}_n^2 - \sum_{n(k-)} \log \bar{r}_n^2 \right) \quad (3.16)$$

$$\text{Método } BH - D_k^{BH} = \frac{1}{2} \left(\log \sum_{n(k+)} \bar{r}_n^2 - \log \sum_{n(k-)} \bar{r}_n^2 \right) \quad (3.17)$$

sendo:

D_k é a estatística que mede o efeito de dispersão do k -ésimo fator ou interação ($k = 1, \dots, K$; $D_k = D_{kq}^S, D_{kq}^R, D_{kq}^H, D_{kq}^{BM}, D_{kq}^{HM}$ ou D_{kq}^{BH} , conforme o método utilizado);

\bar{r}_n^2 é o resíduo quadrático médio da n -ésima condição experimental ($n = 1, \dots, N$);

\bar{r}_n^2 é o resíduo quadrático modificado médio da n -ésima condição experimental ($n = 1, \dots, N$).

A utilização das médias dos resíduos quadráticos em cada condição experimental faz parte da proposta deste trabalho e é explicada no Capítulo 5.

As estatísticas supracitadas, utilizadas para estimar os efeitos de dispersão, apresentam distribuição aproximadamente normal e, na maioria dos trabalhos aplicados, a identificação dos efeitos ativos é feita pelo gráfico de probabilidade normal. Considerando, entretanto, a subjetividade da interpretação gráfica e a grande quantidade de gráficos que seriam gerados,

optou-se por fazer a identificação dos efeitos ativos a partir de escores padronizados, ou seja, pelo afastamento do resultado em relação à sua média, expresso em unidades de desvio-padrão.

Quando o resultado se afastar mais de dois desvios-padrões da média, o fator é considerado significativamente ativo, ou seja, detecta-se efeito de dispersão. Esse ponto de corte, $|z| = 2$, foi escolhido, porque $P(|z| < 2) \approx 0,95$, representativo da chance de um efeito não-ativo não ser identificado como ativo.

Para executar tal avaliação, inicialmente são calculados a média e o desvio-padrão das estatísticas encontradas, para elas poderem ser transformadas em escores padronizados. Como podem existir efeitos de dispersão ativos que distorceriam a média e a variância amostrais de D_k , eliminam-se do cálculo, em cada amostra (resultados de um projeto foco), os dois resultados da estatística D de maior módulo. Sendo $|D_{1q}|, |D_{2q}|, \dots, |D_{(K'-2)q}|$ colocados em ordem crescente, a média e o desvio-padrão da q -ésima amostra são encontrados pelas expressões:

$$\bar{D}_q = \frac{\sum_{k=1}^{K'-2} D_{kq}}{K'-2} \quad (3.18)$$

e

$$S_{D_q}^2 = \frac{\sum_{k=1}^{K'-2} (D_{kq} - \bar{D}_q)^2}{K'-3} \quad (3.19)$$

sendo:

D_{kq} a estatística que mede o efeito de dispersão do k -ésimo fator ou interação da q -ésima amostra ($k = 1, \dots, K'$; $q = 1, \dots, Q$; $D_{kq} = D_{kq}^S, D_{kq}^R, D_{kq}^H, D_{kq}^{BM}, D_{kq}^{HM}$ ou D_{kq}^{BH} , conforme o método utilizado);

\bar{D}_q a média amostral das estatísticas que medem efeitos de dispersão na q -ésima amostra ($q = 1, \dots, Q$);

$S_{D_q}^2$ a variância amostral das estatísticas que medem efeitos de dispersão na q -ésima amostra ($q = 1, \dots, Q$).

O escore padronizado do k -ésimo fator de uma estatística D da q -ésima amostra, z_{kq} , é dado por:

$$z_{kq} = \frac{D_{kq} - \bar{D}_q}{S_{D_q}} \quad (3.20)$$

sendo S_{D_q} o desvio-padrão das estatísticas que medem efeitos de dispersão na q -ésima amostra.

3.3.4 Quarta etapa: estimativa das proporções de identificação correta e de identificação falsa.

Para cada fator ou interação de segunda ordem, k , e cada método, em cada condição experimental, são estimadas as proporções de identificação correta e de identificação falsa de efeitos de dispersão.

A proporção de identificações corretas de efeitos de dispersão do k -ésimo fator na n -ésima condição experimental (PIC_{kn}) é definida por:

$$PIC_{kn} = \frac{QIC_{kn}}{Q} \quad (3.21)$$

sendo:

QIC_{kn} a quantidade de identificações corretas do k -ésimo fator ou interação de segunda ordem na n -ésima condição experimental ($k=1,2$ ou 4 ; $n=1,2,\dots,N$);

Q a quantidade de amostras simuladas (5.000).

A proporção de identificações falsas de efeitos de dispersão do k -ésimo fator na n -ésima condição experimental (PIF_k) é definida por:

$$PIF_{kn} = \frac{QIF_{kn}}{Q} \quad (3.22)$$

sendo QIF_{kn} a quantidade de identificações falsas do k -ésimo fator ou interação de segunda ordem na n -ésima condição experimental ($k=1,2,\dots,10$; $n=1,2,\dots,N$).

O procedimento acima descrito, entretanto, fornece várias PIC_{kn} 's e PIF_{kn} 's para cada condição experimental do projeto de simulação, referente a cada efeito de dispersão em

estudo. A fim de se obter uma única proporção de identificação correta e uma única proporção de identificação falsa de efeitos de dispersão, por condição experimental simulada, adotaram-se os critérios de proporção de identificação correta de *todos* os efeitos de dispersão (*PIC*) e proporção de identificação falsa de *algum* efeito de dispersão (*PIF*). Para tal, os experimentos gerados são classificados em 4 tipos:

- a) tipo *1* - apresenta o fator x_1 do projeto foco com efeito de dispersão;
- b) tipo *4* - apresenta o fator x_4 do projeto foco com efeito de dispersão;
- c) tipo *12* - apresenta os fatores x_1 e x_2 do projeto foco com efeito de dispersão;
- d) tipo *24* - apresenta os fatores x_2 e x_4 do projeto foco com efeito de dispersão.

Para experimentos do tipo *1*, tem-se identificação correta quando é identificado o fator x_1 e falsa, quando, pelo menos, algum outro efeito principal ou interação de segunda ordem é identificado.

Para experimentos do tipo *4*, tem-se identificação correta quando é identificado o fator x_4 e falsa, quando, pelo menos, algum outro efeito principal ou interação de segunda ordem é identificado.

Para experimentos do tipo *12*, tem-se identificação correta quando são identificados simultaneamente os fatores x_1 e x_2 e falsa, quando, pelo menos, algum outro efeito principal ou interação de segunda ordem é identificado.

Para experimentos do tipo *24*, tem-se identificação correta quando são identificados simultaneamente os fatores x_2 e x_4 e falsa, quando, pelo menos, algum outro efeito principal ou interação de segunda ordem é identificado.

As proporções de identificação correta de *todos* os efeitos de dispersão (PIC_n ou, simplesmente, *PIC*) e de identificação falsa de *algum* efeito de dispersão (PIF_n ou, simplesmente, *PIF*), na *n*-ésima condição experimental, são encontradas pelas expressões:

$$PIC_n = \frac{QIC_n}{Q} \quad (3.23)$$

e

$$PIF_n = \frac{QIF_n}{Q} \quad (3.24)$$

sendo:

QIC_n a quantidade de identificações corretas de *todos* os efeitos de dispersão na *n*-ésima condição experimental;

QIF_n a quantidade de identificações falsas de *algum* efeito de dispersão na *n-ésima* condição experimental.

As proporções de identificação correta e de identificação falsa estão relacionadas com os conceitos de sensibilidade e especificidade, empregados na área da saúde, para avaliação da qualidade de um teste diagnóstico. De acordo com Soares (1999, p.99-100), a sensibilidade é definida como a probabilidade do teste ser positivo, sabendo-se ser doente o paciente que está sendo examinado, enquanto especificidade é a probabilidade do teste dar negativo, sabendo-se que o paciente examinado não é portador da doença.

No presente estudo, adotou-se como sensibilidade do método a proporção de vezes em que *todos* os fatores de dispersão foram corretamente identificados pelo método (*PIC*) e como especificidade, o complemento da proporção de vezes em que o método identificou incorretamente *algum* fator de dispersão (*PIF*), ou seja:

$$\text{sensibilidade} = PIC \quad (3.25)$$

e

$$\text{especificidade} = 1 - PIF \quad (3.26)$$

Um método apresenta melhor desempenho quando possui maior sensibilidade e maior especificidade.

ANÁLISE DOS RESULTADOS DA SIMULAÇÃO

Este capítulo apresenta a análise dos resultados do estudo de simulação descrito no capítulo anterior, realizado para identificar o(s) método(s) de melhor desempenho e avaliar o efeito de cada fator e interação do projeto simulado sobre o desempenho de cada método.

4.1 DESCRIÇÃO DA ANÁLISE

Os resultados do estudo de simulação, cujas respostas são as proporções de identificação correta e de identificação falsa de efeitos de dispersão, por condição experimental do projeto de simulação (projeto 2^{6-1}), são analisados com o auxílio dos *softwares* estatísticos: *SPSS* - versão 10.0, *Statistica* - versão 6.0 e *Systat* – versão 8.0, além da planilha eletrônica Excel. Para esta análise, as 32 condições experimentais do projeto de simulação são combinadas com 6 níveis de uma nova variável, considerada como um novo fator (*G*), que identifica o método utilizado para estimar os efeitos de dispersão. Um novo projeto do tipo $6 \times 2^{6-1}$ é gerado, com 192 condições experimentais.

A análise é realizada separadamente para proporções de identificações corretas e falsas.

Inicialmente, estas proporções, encontradas por fator do experimento foco (projeto 2^4) para cada condição experimental do projeto $6 \times 2^{6-1}$, são organizadas e resumidas com o auxílio de técnicas de estatística descritiva. Mesmo tipo de análise é executada nas proporções encontradas, simultaneamente, para todos os fatores e interações desse experimento foco: as proporções de identificações corretas de *todos* os efeitos de dispersão e as proporções de identificação falsa de *algum* efeito de dispersão. Nestas últimas, é realizada uma análise da variância para identificar os efeitos significativos.

Também são usados contrastes ortogonais (Montgomery, 1997a, p.96) para averiguar se:

- (a) o método que usa média aritmética entre variâncias amostrais (R) apresenta proporções de identificação correta de *todos* os efeitos de dispersão e proporções de identificação falsa de *algum* efeito de dispersão diferentes do método que usa média geométrica (S);
- (b) métodos que usam média aritmética entre resíduos quadráticos (BM e BH) apresentam proporções de identificação correta de *todos* os efeitos de dispersão e proporções de identificação falsa de *algum* efeito de dispersão diferentes dos métodos que usam média geométrica (H e HM);
- (c) métodos que usam resíduos quadráticos simples (H e BM) apresentam proporções de identificação correta de *todos* os efeitos de dispersão e proporções de identificação falsa de *algum* efeito de dispersão diferentes dos métodos que usam resíduos quadráticos modificados (HM e BH).

A comparação entre o desempenho de métodos que utilizam variâncias amostrais e métodos que empregam resíduos quadráticos é apresentada no capítulo seguinte, em outro estudo de simulação, pois, neste primeiro, os termos com efeito de locação são incluídos corretamente no modelo e este procedimento pode causar viés na comparação.

4.2 IDENTIFICAÇÃO CORRETA

4.2.1 Estatística descritiva

A análise inicia com as proporções de identificação correta de efeitos de dispersão, em cada condição experimental, por método e por fator do projeto foco (x_1 , x_2 e x_4). Tais proporções, representativas da capacidade de um método identificar um efeito de dispersão ativo, baseiam-se nas 5.000 amostras simuladas.

Os resultados, encontrados por método, estão representados na tabela 4.1 pela média das proporções nas 32 condições experimentais do projeto de simulação, sendo possível constatar a existência de uma ligeira superioridade dos resultados nos métodos propostos para experimentos não-replicados, que foram adaptados à situação de poucas replicações.

Chama-se a atenção para o fato de, para cada método, essas médias terem sido encontradas a partir de 16 resultados, pois cada um dos fatores com efeito de dispersão ativo (x_1 , x_2 e x_4) aparece com esse efeito em 16 condições experimentais das 32 simuladas para cada método. O fator x_1 , que também apresenta efeito de locação, pode ser o único fator do experimento que apresenta efeito de dispersão ou pode apresentá-lo simultaneamente com o

fator x_2 , o mesmo acontecendo com o fator x_4 , que, entretanto, não apresenta efeito de locação. Já o efeito de dispersão do fator x_2 nunca ocorre isoladamente.

Concordando com vários autores, como Brenneman e Nair (2001), os resultados sugerem que quando existe apenas um efeito de dispersão, sua identificação é mais fácil, tendo em vista que as proporções médias de identificação correta de efeitos de dispersão por método são sempre ligeiramente inferiores para o fator x_2 .

TABELA 4.1 – Proporção média⁽¹⁾ de identificação correta de efeitos de dispersão, por método e por fator, nas condições experimentais simuladas

Fator	<u>PIC_k</u> ⁽²⁾					
	<i>R</i>	<i>S</i>	<i>H</i>	<i>BM</i>	<i>HM</i>	<i>BH</i>
x_1	0,634	0,587	0,728	0,757	0,712	0,723
x_2	0,556	0,560	0,672	0,664	0,675	0,650
x_4	0,635	0,589	0,684	0,707	0,695	0,722

(1) Médias calculadas a partir de 16 condições experimentais.

(2) PIC_k representa a proporção de identificações corretas do k -ésimo fator ($k=1,2$ ou 4).

Para sintetizar os vários valores de PIC_k 's para cada método, por condição experimental do projeto de simulação, são consideradas as proporções de identificação correta, por experimento (tabela A1 do apêndice), que representam a capacidade de um método identificar corretamente *todos* os efeitos de dispersão existentes (PIC_n).

Os resultados também se baseiam nas 5.000 amostras simuladas e são representados pela média e pelo desvio-padrão na tabela 4.2, sugerindo a existência de diferença entre os valores médios e a não-existência de diferença entre as variabilidades dos diferentes métodos.

TABELA 4.2 – Média e desvio-padrão, por método, de proporções de identificação correta de *todos* os efeitos de dispersão nas condições experimentais simuladas

<u>Método</u>	<u>PIC</u>	
	Média ⁽¹⁾	Desvio-padrão ⁽¹⁾
<i>R</i>	0,525	0,281
<i>S</i>	0,404	0,298
<i>H</i>	0,613	0,261
<i>BM</i>	0,634	0,254
<i>HM</i>	0,612	0,263
<i>BH</i>	0,622	0,260

(1) Estatísticas calculadas a partir de 32 condições experimentais que apresentam algum efeito de dispersão.

4.2.2 Análise da variância

Uma análise exploratória dos valores das *PIC's*, proporções de identificações corretas de *todos* os efeitos de dispersão existentes, sugere a necessidade de utilizar uma transformação matemática para melhorar a aproximação da distribuição dos dados de uma distribuição normal. Além disso, de acordo com Zar (1999, p.278), percentagens ou proporções entre 0 e 1 podem afastar-se muito de uma distribuição normal (o que é desejado para a análise a ser realizada), principalmente se apresentarem resultados entre 0 e 0,3 ou entre 0,7 e 1.

Se a raiz quadrada de cada proporção é transformada em seu arco seno $\left(\text{arc sen}\sqrt{\hat{P}}\right)$, a variância se estabiliza e os resultados tendem a apresentar uma distribuição mais próxima da normal. As análises subseqüentes, realizadas em proporções de identificações corretas de *todos* os efeitos de dispersão, são executadas nos valores transformados $\left(\text{arc sen}\sqrt{PIC}\right)$.

Para avaliar o efeito de cada fator do projeto $6 \times 2^{6-1}$ sobre os valores de *PIC's* transformados, utiliza-se uma análise da variância. A estatística *F* empregada nessa análise (equação 2.5 apresentada na subseção 2.1.2) é calculada em função de uma estimativa da variância do erro. Entretanto, considerando que, de acordo com Chatterjee e Price (1977, p. 39), a transformação $\text{arc sen}\sqrt{\hat{P}}$ estabiliza a variância de uma distribuição amostral de proporções em $\frac{1}{4Q}$ e que foram feitas 5.000 simulações ($Q = 5.000$), a verdadeira variância do erro (σ^2) é conhecida e igual a 0,00005. Se a verdadeira variância do erro (σ^2) é conhecida, a significância dos resultados pode ser avaliada por meio da estatística:

$$\chi^2 = \frac{SQ_{\text{Fonte de variação}}}{\sigma^2} \quad (4.1)$$

que apresenta distribuição Qui-quadrado com $gl_{\text{Fonte de variação}}$ graus de liberdade.

Tendo em vista o objetivo dessa simulação, cujo propósito é avaliar o desempenho dos diferentes métodos para identificar efeitos de dispersão, mostra-se relevante para este estudo, apenas o efeito principal do fator *G* (método) e suas possíveis interações com os demais fatores do projeto de simulação. A análise realizada (tabela 4.3), por meio da estatística χ^2 , permite concluir que o método utilizado para detectar efeitos de dispersão (*G*) interfere no valor médio da proporção de identificação correta de *todos* os efeitos de dispersão e que há interação desse fator com todos os demais.

TABELA 4.3 – Análise do efeito do fator G (método utilizado) e suas interações com os demais fatores na variável transformada $\arcsen\sqrt{PIC}$

Fator ou interação	Soma dos quadrados (SQ)	Graus de liberdade (gl)	χ^2	Valor- p
G	1,7840	5	35680,00	< 0,001
AG	0,0994	5	1988,00	< 0,001
BG	0,0028	5	56,00	< 0,001
CG	0,1390	5	2780,00	< 0,001
DG	0,0280	5	560,00	< 0,001
EG	0,0066	5	132,00	< 0,001
FG	0,0819	5	1638,00	< 0,001

Nota: $\chi^2_{(5; 0,001)} = 20,52$.

O gráfico da figura 4.1 mostra existir uma tendência de superioridade das adaptações dos métodos propostos para experimentos não-replicados, considerando a limitação desse estudo de simulação, que supõe o modelo de locação corretamente identificado.

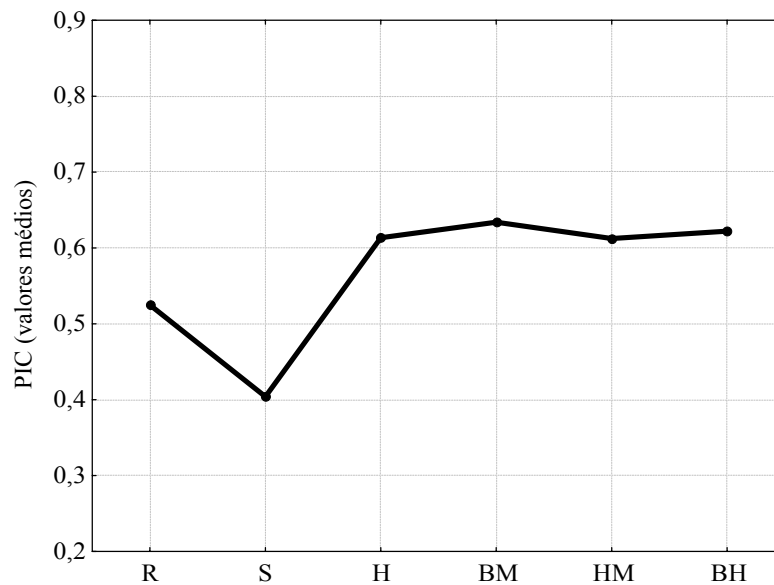


FIGURA 4.1 – Gráfico do efeito do fator G (método utilizado) sobre o valor médio da proporção de identificação correta de *todos* os efeitos de dispersão

Os gráficos (a), (b), (c), (d), (e) e (f) da figura 4.2 descrevem, respectivamente, a interação do fator G com o fator A (quantidade de fatores com efeito de locação), com o fator B (intensidade do efeitos de dispersão), com o fator C (quantidade de fatores com efeito de dispersão), com o fator D (nível de coincidência de efeito de locação e dispersão em um mesmo fator), com o fator E (intensidade do efeito de dispersão) e com o fator F (quantidade de replicações) que, embora significativas, não parecem muito relevantes.

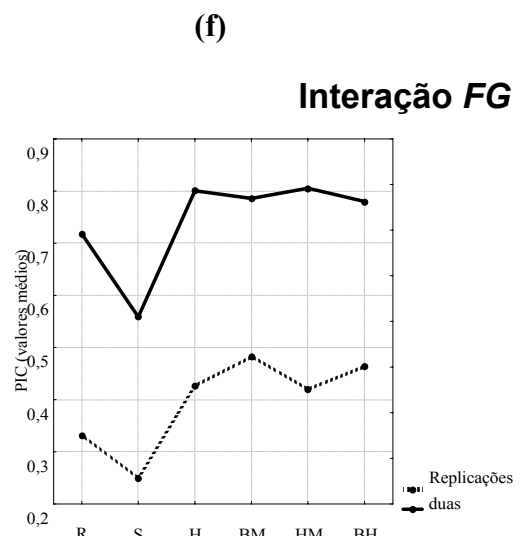
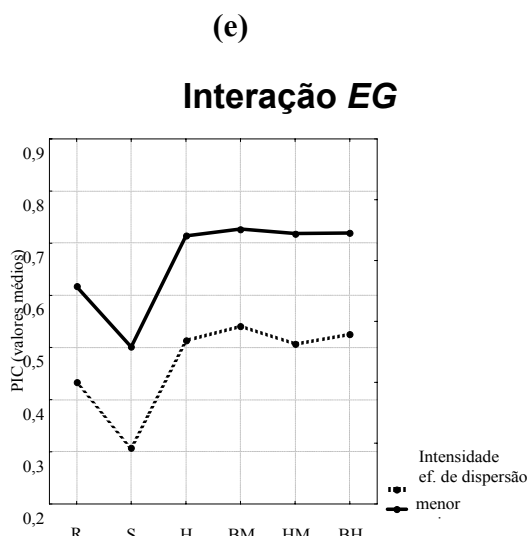
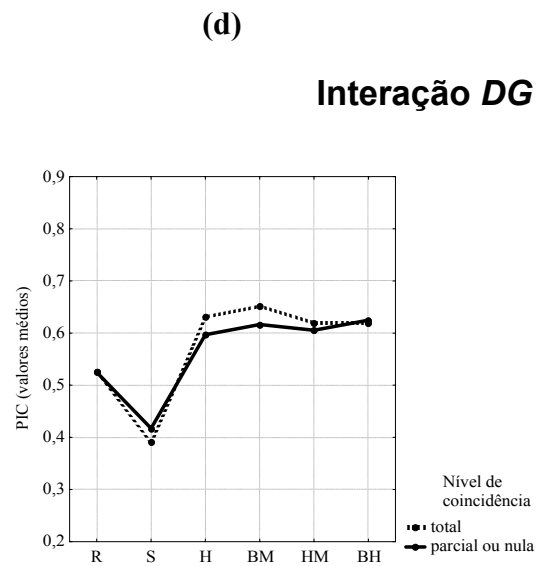
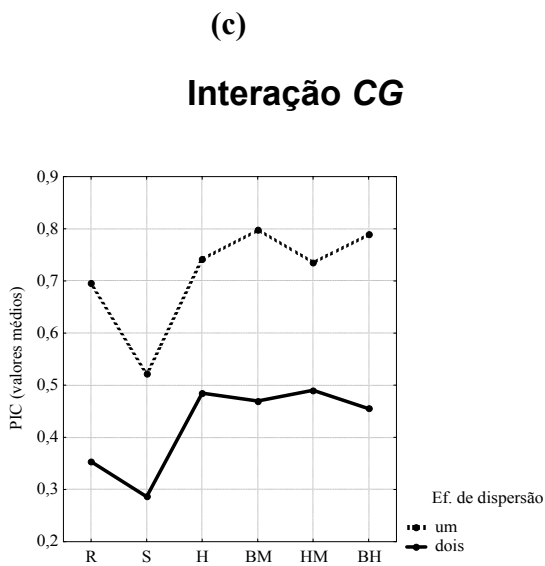
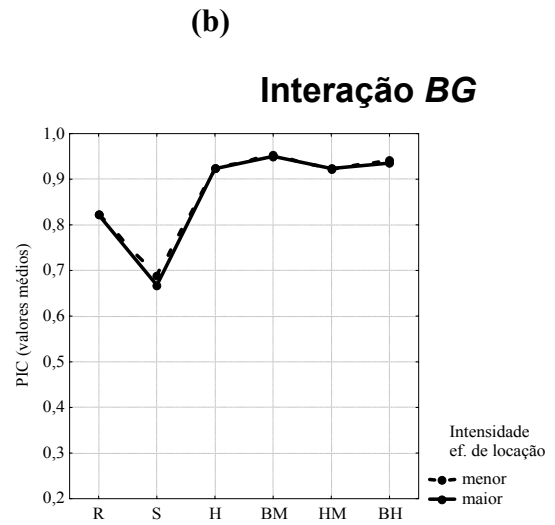
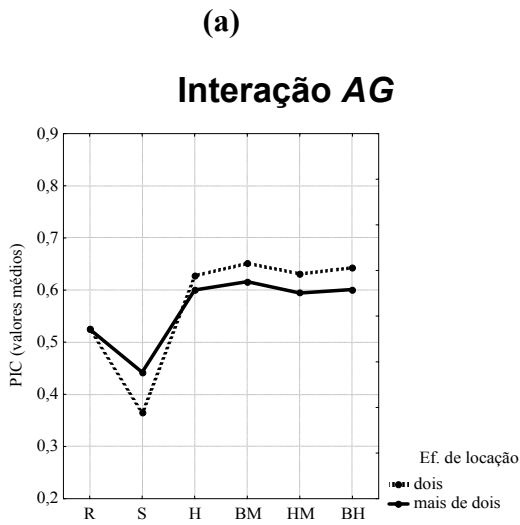


FIGURA 4.2 – Gráficos dos efeitos das interações do fator *G* (método utilizado) com os fatores *A*, *B*, *C*, *D*, *E* e *F* sobre as proporções de identificação correta de *todos* os efeitos de dispersão

O gráfico (a) da figura 4.2 mostra que quando existem menos fatores com efeitos de locação (Fator *A* no nível $-I$), é mais acentuada a diferença entre métodos que usam resíduos quadráticos e aqueles que empregam variâncias amostrais, embora em ambas as situações, o método *BM* apresente melhor resultado, seguido do método *BH*.

O gráfico (b) da mesma figura mostra que, quando os fatores de locação são menos intensos (Fator *B* no nível $-I$), o método *S* apresenta melhor desempenho do que quando existem mais fatores com efeito de locação (Fator *B* no nível $+I$). Os demais métodos apresentam comportamentos semelhantes.

Já o gráfico (c) da mesma figura evidencia que, quando há apenas um fator com efeito de dispersão (Fator *C* no nível $-I$), os métodos *BM* e *BH* apresentam melhor desempenho, o mesmo não acontecendo se existirem mais fatores com efeitos de dispersão (Fator *C* no nível $+I$). Salienta-se que, na última situação, o desempenho dos métodos *BM*, *BH*, *H* e *HM* é muito semelhante.

O gráfico (d) da figura 4.2 mostra que, ao existir coincidência total de efeito de locação e dispersão em um mesmo fator (Fator *D* no nível $-I$), os métodos *BM* e *H* apresentam melhor desempenho. Entretanto, quando pelo menos um fator com efeito de dispersão não mostra efeito de locação (Fator *D* no nível $+I$), o melhor desempenho é apresentado pelo método *BH*.

O gráfico (e) evidencia que, embora o método *BM* tenha desempenho superior aos demais, a superioridade é mais evidente se os fatores são menos intensos (Fator *E* no nível $-I$), enquanto o gráfico (f) revela que os métodos *BM* e *BH* apresentam melhor desempenho quando são utilizadas menos replicações (Fator *F* no nível $-I$). Para a situação de mais replicações (Fator *F* no nível $+I$), o desempenho dos métodos *BM*, *BH*, *H* e *HM* é muito semelhante.

4.2.3 Contrastes

São usados contrastes ortogonais para comparar as proporções de identificação correta de *todos* os efeitos de dispersão entre: (a) o método com uso da média aritmética (*R*) e o método com emprego da média geométrica entre variâncias amostrais (*S*); (b) métodos com utilização da média aritmética (*BM* e *BH*) e métodos com emprego da média geométrica entre

resíduos quadráticos (*H* e *HM*); (c) métodos que usam resíduos quadráticos simples (*H* e *BM*) e métodos que empregam resíduos quadráticos modificados (*HM* e *BH*).

A tabela 4.4 mostra os coeficientes adotados nos contrastes para efetuar cada uma das comparações acima mencionadas.

TABELA 4.4 – Coeficientes utilizados nos contrastes

Hipótese	Coeficientes					
	Método <i>R</i>	Método <i>S</i>	Método <i>H</i>	Método <i>BM</i>	Método <i>HM</i>	Método <i>BH</i>
(a)	-1	1	0	0	0	0
(b)	0	0	-1	1	-1	1
(c)	0	0	-1	-1	1	1

Uma análise da variância efetuada com o verdadeiro valor da variância do erro (tabela 4.5), mostra existirem evidências de diferença entre os valores médios das *PIC*'s de métodos com uso de diferentes tipos de média entre variâncias amostrais: o método *R*, que utiliza a média aritmética, possui *PIC* média significativamente maior do que o método *S*, que utiliza média geométrica, apresentando médias amostrais, respectivamente, iguais a 0,525 e 0,404.

Em relação aos métodos com uso de diferentes tipos de média entre resíduos quadráticos, foi possível constatar possuírem aqueles que utilizam média aritmética, *BM* e *BH*, proporções médias significativamente maiores do que outros com utilização da média geométrica entre resíduos quadráticos, *H* e *HM*, apresentando médias amostrais, respectivamente, iguais a 0,628 e 0,613.

Essa mesma análise também encontra evidências de diferença entre as proporções médias de identificação correta de *todos* os efeitos de dispersão entre métodos com emprego de diferentes tipos de resíduos quadráticos. Os que utilizam resíduos quadráticos simples apresentam média amostral de 0,624, enquanto os que usam resíduos quadráticos modificados, 0,617.

TABELA 4.5 – Análise do efeito dos contrastes aplicados ao fator *G* (método utilizado) na variável transformada $\arcsen\sqrt{PIC}$

Hipótese	Soma dos quadrados (<i>SQ</i>)	Graus de liberdade (<i>gl</i>)	χ^2	valor- <i>p</i>
(a)	0,3324	1	6648,00	< 0,001
(b)	0,0151	1	30,00	< 0,001
(c)	0,0015	1	30,00	< 0,001

Nota: $\chi^2_{(1; 0,001)} = 10,83$

4.3 IDENTIFICAÇÃO FALSA

4.3.1 Estatística descritiva

Nas 5.000 amostras sob a mesma condição experimental, também são estimadas as proporções de identificação falsa de efeitos de dispersão, por fator, para cada um dos métodos, em cada uma das 32 condições experimentais. Os resultados encontrados, sintetizados pela média aritmética, estão apresentadas na tabela 4.6.

De maneira geral, os métodos propostos para experimentos não-replicados parecem apresentar *PIF's* ligeiramente inferiores aos demais, o que é satisfatório. Também é possível constatar que os métodos *R*, *BM* e *BH* possuem uma proporção média maior de falsas identificações de efeitos de dispersão nas interações x_1x_2 e x_2x_4 , representativas das interações entre fatores com efeitos de dispersão.

TABELA 4.6 – Proporção média⁽¹⁾ de identificação falsa de efeitos de dispersão, por fator, em 32 condições experimentais simuladas

Fator	<i>PIF_k</i> ⁽²⁾					
	<i>R</i>	<i>S</i>	<i>H</i>	<i>BM</i>	<i>HM</i>	<i>BH</i>
x_1	0,055	0,068	0,070	0,056	0,060	0,043
x_2	0,071	0,081	0,076	0,071	0,071	0,063
x_3	0,054	0,068	0,062	0,041	0,059	0,047
x_4	0,055	0,069	0,059	0,045	0,065	0,051
x_1x_2	0,090	0,069	0,068	0,104	0,059	0,089
x_1x_3	0,053	0,067	0,061	0,049	0,060	0,045
x_1x_4	0,055	0,067	0,058	0,046	0,060	0,050
x_2x_3	0,055	0,069	0,067	0,053	0,059	0,045
x_2x_4	0,090	0,068	0,060	0,092	0,071	0,100
x_3x_4	0,055	0,067	0,059	0,045	0,067	0,052

(1) Médias calculadas a partir de 32 condições experimentais.

(2) *PIF_k* representa a proporção de identificações falsas do *k*-ésimo fator ($k=1,2,\dots,10$).

De acordo com Brenneman e Nair (2001), em experimentos não-replicados, a estimação do efeito de dispersão de um fator, equivalente à interação de dois outros fatores com efeito de dispersão, é viesada quando são usados os métodos *BM*, *BH*, *H* e *HM*: tanto podem ser encontrados falsos efeitos de dispersão como efeitos de dispersão não serem detectados, pois esses métodos sofrem de viés estrutural. O efeito de dispersão de um fator depende também dos outros fatores envolvidos no experimento. Estes autores também afirmam que a utilização de resíduos quadráticos modificados ameniza problemas dessa natureza.

Entretanto, nas condições experimentais simuladas, pelos resultados encontrados, a identificação de falsos efeitos de dispersão em interações de fatores que apresentam efeitos de dispersão parece ocorrer com mais intensidade nos métodos *BM* e *BH*. Problemas dessa natureza também são identificados no método *R*, concordando com Nair e Pregibon (1988).

Após, são analisadas as proporções de identificação falsa por experimento (tabela A2 no apêndice), que representam a capacidade de um método identificar falsamente *algum* efeito de dispersão. Estes resultados, quando representados pela média e pelo desvio-padrão (tabela 4.7), sugerem a existência de diferença entre as variâncias das proporções de identificação falsa, o que não era esperado.

TABELA 4.7 – Média e desvio-padrão, por método, das proporções de identificação falsa de *algum* efeito de dispersão nas condições experimentais simuladas

<u>Método</u>	<i>PIF</i>	
	Média ⁽¹⁾	Desvio-padrão ⁽¹⁾
<i>R</i>	0,449	0,071
<i>S</i>	0,486	0,123
<i>H</i>	0,458	0,125
<i>BM</i>	0,447	0,065
<i>HM</i>	0,454	0,122
<i>BH</i>	0,430	0,062

(1) Estatísticas calculadas a partir das 32 condições experimentais simuladas.

4.3.2 Análise da variância

A sondagem de algumas características técnicas desses dados indica ser a transformação $\text{arc sen}\sqrt{\hat{P}}$ também indicada para aproximar os resultados das *PIF*'s de uma distribuição normal e estabilizar a variância. As próximas análises dessa seção são realizadas em valores transformados: $\text{arc sen}\sqrt{PIF}$.

Considerando o projeto experimental $6 \times 2^{6-1}$, uma análise da variância é realizada (tabela 4.8) por meio da estatística χ^2 , permitindo identificar a interferência do fator *G* (método utilizado no cálculo da estimativa do efeito de dispersão) no valor médio da proporção de identificação falsa de *algum* efeito de dispersão.

O gráfico da figura 4.3 mostra ser o método *S* o que apresenta maior média para a proporção de identificação falsa de *algum* efeito de dispersão, enquanto o método *BH* tem a menor média, evidenciando uma tendência de superioridade das adaptações dos métodos (*G*) propostos para experimentos não-replicados, considerando o modelo de locação corretamente identificado.

TABELA 4.8 – Análise do efeito do fator G (método utilizado) e suas interações com os demais fatores na variável transformada $\arcsen\sqrt{PIF}$

Fontes de variação	Soma dos quadrados (SQ)	Graus de liberdade (gl)	χ^2	valor- p
G	0,0511	5	1022,00	<0,001
AG	0,0004	5	6,00	>0,05
BG	0,0003	5	4,00	>0,05
CG	0,1465	5	2940,00	<0,001
DG	0,0014	5	26,00	<0,001
EG	0,0044	5	88,00	<0,001
FG	0,0832	5	1664,00	<0,001

Nota: $\chi^2_{(5; 0,001)} = 20,52$; $\chi^2_{(5; 0,05)} = 11,07$.

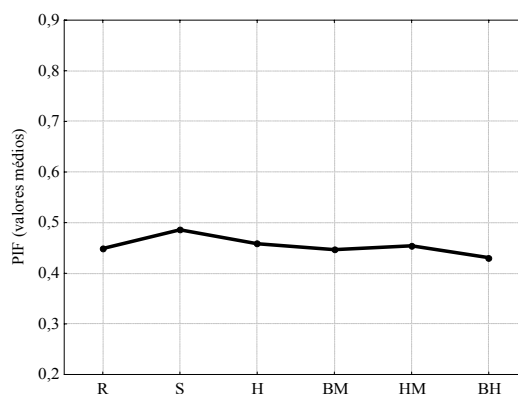


FIGURA 4.3 – Gráfico do efeito do fator G (método utilizado) sobre a proporção de identificação falsa de *algum* efeito de dispersão

Também é possível constatar a interação do fator G com os fatores C (quantidade de efeitos de dispersão), D (nível de coincidência de efeito de locação e dispersão), E (intensidade do efeito de dispersão) e F (quantidade de replicações). Os gráficos (a), (b), (c) e (d) da figura 4.4 descrevem a interação deles.

De acordo com o gráfico (a) desta figura, quando existe só um fator com efeito de dispersão (Fator C no nível $-I$), os métodos com menores médias para as PIF 's são BH e BM , mas se existe mais de um fator com efeito de dispersão (Fator C no nível $+I$), os métodos com melhor performance são H e HM .

O gráfico (b) dessa figura mostra a interferência no nível de coincidência de efeitos de dispersão e locação, que parece ser levemente mais forte nos métodos S e BH . Já o gráfico (c) evidencia que, quando os efeitos de dispersão são mais intensos, todos os métodos apresentam melhor desempenho. Esta melhora, entretanto, parece ser mais acentuada para o método HM .

O gráfico (d) da figura 4.4 mostra que, ao serem utilizadas menos replicações (Fator F no nível -1), os métodos com menores médias para a proporção de identificação falsa são os métodos BM e BH , exatamente os mesmos que apresentam maiores médias quando se empregam mais replicações (Fator F no nível $+1$).

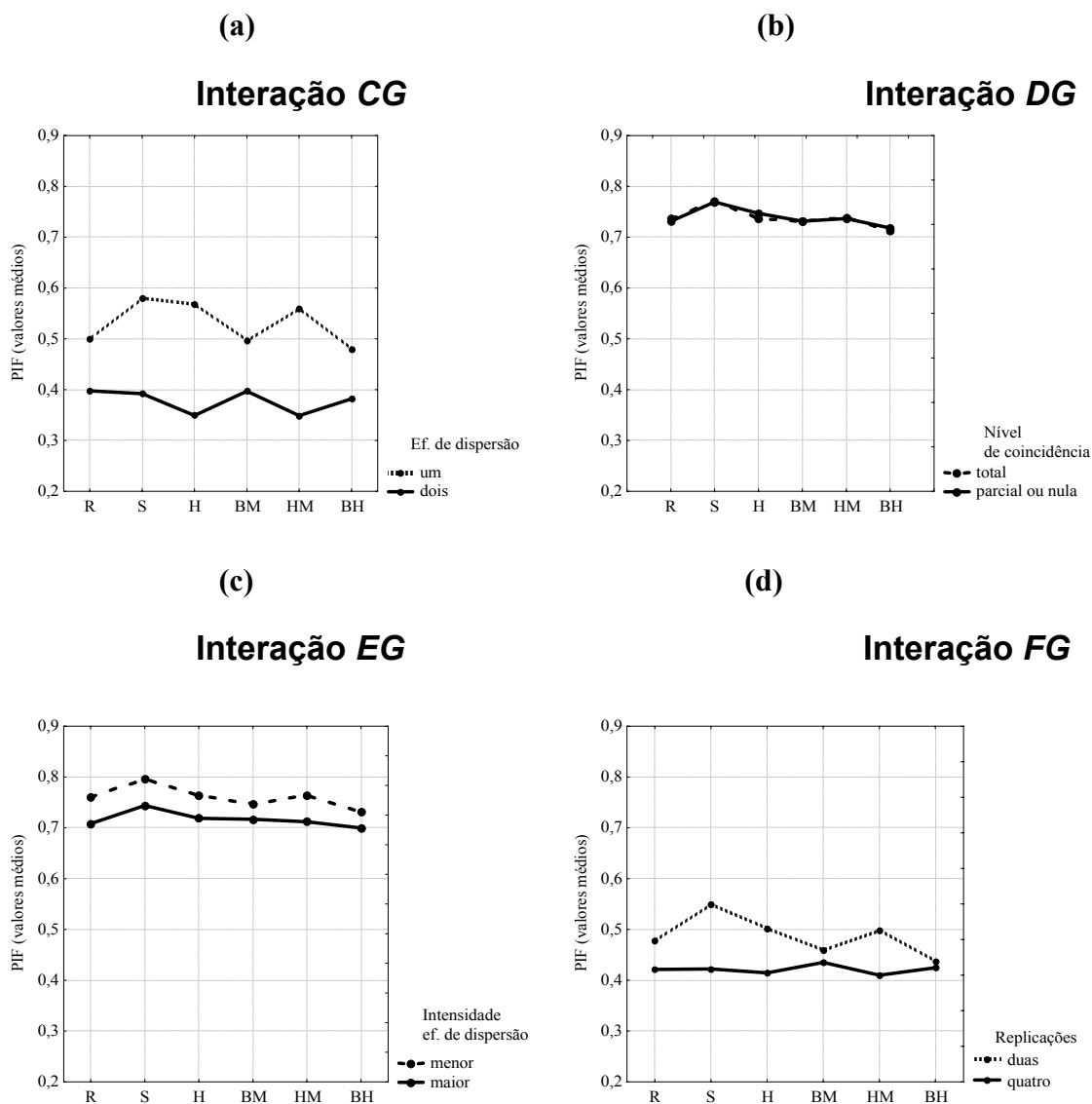


FIGURA 4.4 – Gráfico do efeito das interações do fator G com os fatores C , D , E e F sobre a proporção de identificação falsa de *algum* efeito de dispersão

4.3.3 Contrastes

Novamente são usados os mesmos contrastes ortogonais, já utilizados na subseção 4.2.3, para comparar as proporções de identificação falsa de *algum* efeito de dispersão Os

coeficientes adotados nos contrastes para efetuar essas comparações estão na tabela 4.3 da subseção 4.2.3.

A análise (tabela 4.9), realizada com o auxílio do verdadeiro valor da variância do erro, mostra existirem evidências de diferença entre os valores médios das *PIF*'s de métodos com o uso de diferentes tipos de média entre variâncias amostrais, entre métodos com o emprego de diferentes tipos de médias entre resíduos quadráticos, assim como entre métodos que fazem uso de diferentes tipos de resíduos quadráticos.

O método *R*, utilizador da média aritmética, possui *PIF* média significativamente menor que o método *S*, que utiliza média geométrica, apresentando médias amostrais de, respectivamente, 0,4498 e 0,4859. Os métodos *BM* e *BH*, empregadores da média aritmética entre resíduos quadráticos, também possuem proporções médias significativamente menores dos que *H* e *HM*, com médias amostrais, respectivamente iguais a 0,4387 e 0,4563.

Já os métodos que empregam resíduos quadráticos modificados possuem proporções médias significativamente menores do que os que usam os resíduos quadráticos simples, apresentando médias amostrais, respectivamente iguais a 0,4425 e 0,4525. Salienta-se, entretanto, apontarem as estimativas pontuais dessas proporções para uma diferença bastante pequena.

TABELA 4.9 – Análise do efeito dos contrastes aplicados ao fator *G* (método utilizado) na variável transformada $\arcsen\sqrt{PIF}$

Hipótese	Soma dos quadrados (<i>SQ</i>)	Graus de liberdade (<i>gl</i>)	χ^2	valor- <i>p</i>
(a)	0,0207	1	414,00	< 0,001
(b)	0,0083	1	166,00	< 0,001
(c)	0,0032	1	64,00	< 0,001

Nota: $\chi^2_{(1; 0,001)} = 10,83$.

4.4 CONSIDERAÇÕES SOBRE OS RESULTADOS

Pelos resultados encontrados nessa análise, os métodos propostos para identificar efeitos de dispersão em experimentos não-replicados (métodos *BM*, *BH*, *H* e *HM*) parecem apresentar maior sensibilidade (maior probabilidade de identificar corretamente os efeitos de dispersão ativos) do que os métodos propostos com a mesma finalidade para experimentos com replicações (métodos *R* e *S*), o que poderia ser explicado pelo aumento do número de graus de liberdade.

Entre os métodos que utilizam os resíduos quadráticos, os que fazem uso da média aritmética (métodos *BM* e *BH*) apresentam sensibilidade superior em relação aos que empregam a média geométrica (métodos *H* e *HM*), principalmente quando o experimento analisado apresenta apenas um efeito de dispersão.

Os métodos que utilizam resíduos quadráticos simples (métodos *BM* e *H*) apresentam sensibilidade superior em relação aos que empregam de resíduos quadráticos modificados (métodos *HM* e *BH*).

Chama-se a atenção para o fato de, entre os que trabalham com variâncias amostrais, o método *R* ter apresentado maior sensibilidade e maior especificidade, concordando com Nair e Pregibon (1988) sobre ser ele o método de melhor desempenho para a situação do modelo restrito, principalmente quando existem poucos efeitos de dispersão.

Com relação à especificidade, os resultados são semelhantes: aqueles propostos para identificar efeitos de dispersão em experimentos não-replicados (métodos *H*, *BM*, *HM* e *BH*) parecem apresentar maior especificidade (menor probabilidade de identificar falsamente algum efeito de dispersão) do que os propostos com a mesma finalidade para experimentos com replicações (métodos *R* e *S*), o que também poderia ser explicado pelo aumento do número de graus de liberdade.

Entre os usuários dos resíduos quadráticos, apresentam maior especificidade os que empregam média aritmética (métodos *BH* e *BM*) e resíduos quadráticos modificados (métodos *BH* e *HM*), embora, nesses últimos, a diferença tenha sido muito pequena. Além disso, os métodos que utilizam resíduos quadráticos simples (métodos *BM* e *H*) apresentam especificidade inferior aos que empregam de resíduos quadráticos modificados (métodos *HM* e *BH*).

O método *R* apresenta melhores resultados em relação aos métodos *H* e *HM*. Entretanto, quando existem dois fatores com efeitos de dispersão, os métodos *H* e *HM* apresentam especificidade ligeiramente superior, o mesmo acontecendo quando foram utilizadas quatro replicações. Nessas duas situações, também há uma melhora evidente do método *S* e uma piora dos métodos *BH* e *BM*. Com relação à variabilidade, *R*, *BM* e *BH* apresentam maior estabilidade.

Portanto, os resultados parecem indicar que: (a) entre os métodos que utilizam variâncias amostrais, o *R* apresenta melhor desempenho e (b) entre aqueles que empregam resíduos quadráticos, *BM* e *BH* mostram-se melhores, principalmente quando existem poucos efeitos de locação, apenas um efeito de dispersão coincidindo com o efeito de locação e poucas replicações. Entretanto, se forem usadas mais replicações e existirem mais fatores com

efeitos de dispersão, o desempenho dos métodos *H* e *HM* parece melhorar razoavelmente, superando *BM* e *BH* em algumas situações.

Com relação à utilização de diferentes tipos de resíduos quadráticos, não foi possível identificar o melhor desempenho, tendo em vista que os métodos que usam resíduos quadráticos simples, embora apresentem melhor sensibilidade, apresentam também pior especificidade. Optou-se por trabalhar com métodos que utilizam resíduos quadráticos simples em função da simplicidade do cálculo.

Capítulo 5

SIMULAÇÕES COMPLEMENTARES

Este capítulo descreve e apresenta os resultados de dois estudos de simulação, apresentando seus resultados. O primeiro estudo é realizado para definir como tratar os resíduos quadráticos, na adaptação dos métodos originalmente propostos para experimentos não-replicados para a situação de experimentos com replicações, enquanto o segundo, para comparar o desempenho de métodos que usam variâncias amostrais com o desempenho de métodos que empregam resíduos quadráticos.

5.1 Resíduos quadráticos individuais e médios

No presente estudo foram utilizados os métodos não-iterativos para identificar efeitos de dispersão, considerados na simulação anterior, propostos para serem utilizados em experimentos não-replicados: o método *BM*, o método *H*, o método *BH* e o método *HM*. Os métodos *BM* e *BH* utilizam médias aritméticas entre, respectivamente, resíduos quadráticos simples e resíduos quadráticos modificados, enquanto os métodos *H* e *HM*, médias geométricas, respectivamente, entre esses mesmos tipos de resíduos.

De acordo com tais métodos, nesse tipo de experimento, a avaliação dos efeitos de dispersão é feita a partir de resíduos quadráticos, sendo encontrado um resíduo quadrático para cada condição experimental. No caso, o resíduo quadrático na *n*-ésima condição experimental é $r_n^2 = (y_n - \hat{y}_n)^2$, onde \hat{y}_n e y_n são, respectivamente, o valor predito e o valor observado na *n*-ésima condição. O valor predito é encontrado por meio da modelagem do valor esperado, pelo método dos mínimos quadrados ordinários (MQO). Em experimentos replicados, entretanto, existem *M* observações em cada condição experimental e, conseqüentemente, também existirão *M* resíduos quadráticos.

Ao estender esses métodos para a situação de experimentos replicados, foi necessário definir se, no cálculo das diversas estatísticas, deveriam ser considerados os *NM* resíduos quadráticos existentes ou se deveriam ser utilizadas suas médias em cada condição experimental, ou seja,

$$\overline{r_n^2} = \frac{\sum_{m=1}^M r_{nm}^2}{M} \quad (5.1)$$

sendo r_{nm}^2 o resíduo quadrático na m -ésima replicação da n -ésima condição experimental ($m = 1, 2, \dots, M$; $n = 1, 2, \dots, N$).

Em dois dos métodos supracitados: BM e BH , definidores da estatística que avalia efeitos de dispersão como um quociente entre médias aritméticas, os dois procedimentos conduzem ao mesmo resultado, se o número de replicações for constante nas diversas condições experimentais.

Ao se considerar o valor de $\overline{r_n^2}$ na expressão matemática utilizada para encontrar D_k^{BM} , verifica-se que seu resultado independe da realização do cálculo a partir dos NM resíduos quadráticos ou de suas N médias, pois:

$$D_k^{BM(1)} = \frac{1}{2} \log \left(\frac{\sum_{n(k+)} \overline{r_n^2}}{\sum_{n(k-)} \overline{r_n^2}} \right) = \frac{1}{2} \log \left(\frac{\sum_{n(k+)} \left(\frac{\sum_{m=1}^M r_{nm}^2}{M} \right)}{\sum_{n(k-)} \left(\frac{\sum_{m=1}^M r_{nm}^2}{M} \right)} \right) = \frac{1}{2} \log \left(\frac{\sum_{n(k+)m=1}^M r_{nm}^2}{\sum_{n(k-)m=1}^M r_{nm}^2} \right) = D_k^{BM(2)} \quad (5.2)$$

sendo:

$D_k^{BM(1)}$ a estatística D_k^{BM} calculada com a média dos resíduos quadráticos;

$D_k^{BM(2)}$ a estatística D_k^{BM} calculada com os resíduos quadráticos individuais.

Entretanto, com relação ao método H , no qual o cálculo da estatística D_k^H é realizado a partir de um quociente entre médias geométricas, os dois procedimentos não são equivalentes, pois:

$$D_k^{H(1)} = \frac{1}{N} \log \left(\frac{\prod_{n(k+)} \overline{r_n^2}}{\prod_{n(k-)} \overline{r_n^2}} \right) = \frac{1}{N} \log \left(\frac{\prod_{n(k+)} \left(\frac{\sum_{m=1}^M r_{nm}^2}{M} \right)}{\prod_{n(k-)} \left(\frac{\sum_{m=1}^M r_{nm}^2}{M} \right)} \right) \quad (5.3)$$

$$D_k^{H(1)} = \frac{1}{N} \log \left(\frac{\prod_{n(k+)} \sum_{m=1}^M r_{nm}^2}{\prod_{n(k-)} \sum_{m=1}^M r_{nm}^2} \right) \neq \frac{1}{N} \log \left(\frac{\prod_{n(k+)} \prod_{m=1}^M r_{nm}^2}{\prod_{n(k-)} \prod_{m=1}^M r_{nm}^2} \right) = D_k^{H(2)} \quad (5.4)$$

sendo:

$D_k^{H(1)}$ a estatística D_k^H calculada com a média dos resíduos quadráticos;

$D_k^{H(2)}$ a estatística D_k^H calculada com os resíduos quadráticos individuais.

5.1.1 Descrição do estudo de simulação

Para identificar o procedimento de melhor desempenho a ser empregado nos métodos H e HM (utilizar $\overline{r_n^2}$ ou r_{nm}^2 no cálculo das estatísticas D_k^H e D_k^{HM}), é realizado um estudo de simulação utilizando um projeto semelhante ao descrito no Capítulo 3. Considere-se o projeto fatorial fracionado do tipo 2^{6-1} , apresentado na tabela 3.1, da subseção 3.3.1, com seis fatores ensaiados em dois níveis, conforme Quadro 5.1 Os cinco primeiros fatores: A , B , C , D , E e F combinam-se de acordo com um projeto do tipo 2^{6-1} , originando 32 condições experimentais, para cada uma das quais são gerados 1.000 resultados de um projeto 2^4 (projeto foco). Com relação à identificação de efeitos de dispersão, esses resultados são analisados de quatro diferentes maneiras, conforme seja utilizado o método H ou o método HM e conforme sejam considerados os resíduos quadráticos: individualmente ou substituídos pela média aritmética.

Em cada uma das 32 condições experimentais encontram-se os valores das proporções de identificação correta de *todos* os efeitos de dispersão (PIC) e proporções de identificações falsas de *algum* efeito de dispersão (PIF) para ambos os métodos e ambos procedimentos para os resíduos quadráticos. O projeto lógico do programa computacional desenvolvido para esta simulação é apresentado na figura 5.1.

QUADRO 5.1 – Descrição de fatores e níveis utilizados no estudo de simulação realizado para avaliar o desempenho dos métodos H e HM quando são utilizados resíduos quadráticos individuais ou médios

Fator	Descrição	Nível (-1)	Nível (+1)
A	Fatores com efeitos de locação	β_1 e β_2 (dois efeitos principais)	$\beta_1, \beta_2, \beta_3, \beta_{12}$ e β_{13} (três efeitos principais e duas interações)
B	Intensidade dos efeitos de locação	$\beta_i = 1$ e $\beta_{ij} = 0,5$ para i e j iguais aos índices especificados no fator A $\beta_i = 0$ e $\beta_{ij} = 0$ para i e j diferentes dos índices especificados no fator A	$\beta_i = 2$ e $\beta_{ij} = 1$ para i e j iguais aos índices especificados no fator A $\beta_i = 0$ e $\beta_{ij} = 0$ para i e j diferentes dos índices especificados no fator A
C	Fatores com efeitos de dispersão	θ_k (um efeito de dispersão)	θ_k e θ_2 (dois efeitos de dispersão)
D	Coincidência ou não de efeito de locação e dispersão	$\theta_k = \theta_1$	$\theta_k = \theta_4$
E	Intensidade dos efeitos de dispersão	$\theta_k = 0,549$ e $\theta_i = 0$ para $i \neq k$	$\theta_k = 0,643$, $\theta_2 = 0,549$ e $\theta_i = 0$ para $i \neq k$ e $i \neq 2$
F	Quantidade de replicações	2	4

Salienta-se que, em distribuições amostrais de proporções obtidas em amostras de tamanho 1.000, se for utilizada a transformação $\arcsen\sqrt{\hat{P}}$, a variância se estabiliza em 0,00025, correspondente a um erro padrão de 0,01581. Essa quantidade possibilita estimar as proporções de identificação correta de *todos* os efeitos de dispersão (*PIC*) e proporções de identificações falsas de *algum* efeito de dispersão (*PIF*) com um erro máximo de 0,03099, associado ao nível de confiança de 0,95.

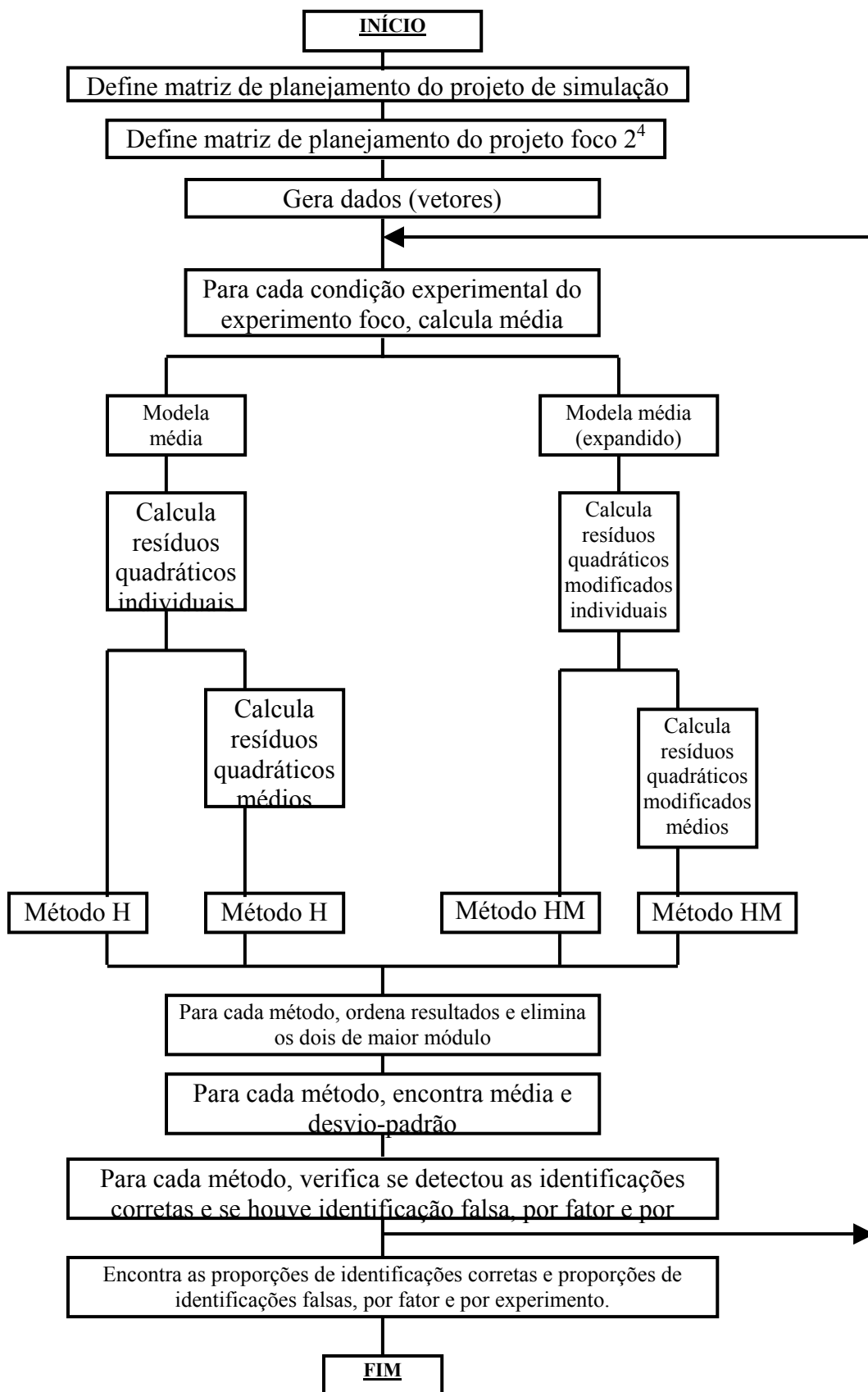


FIGURA 5.1 - Fluxograma do programa computacional utilizado na simulação realizada para avaliar o desempenho dos métodos com a utilização de resíduos quadráticos individuais ou médios

5.1.2 Resultados obtidos no estudo de simulação

Nas 1.000 amostras sob a mesma condição experimental, são encontradas as proporções de identificação correta de *todos* os efeitos de dispersão e de identificação falsa de *algum* efeito de dispersão (tabelas A3 e A4 no apêndice). Os resultados estão sintetizados por meio do valor médio na tabela 5.1.

TABELA 5.1 - Proporções médias de identificação correta de *todos* os efeitos de dispersão e de identificação falsa de *algum* efeito de dispersão nas condições experimentais simuladas

Resíduos	<i>PIC</i>		<i>PIF</i>	
	Método <i>H</i>	Método <i>HM</i>	Método <i>H</i>	Método <i>HM</i>
r_{nm}^2	0,4396	0,4517	0,5068	0,5064
$\overline{r_n^2}$	0,6088	0,6099	0,4577	0,4637

Os resultados evidenciam ser maior a proporção média de identificação correta de *todos* os efeitos de dispersão (*PIC*), quando o cálculo das estatísticas D_k^H e D_k^{HM} é feito com a utilização da média dos resíduos quadráticos em cada condição experimental. Além disso, em nenhuma das condições experimentais analisadas, a proporção de identificação correta de *todos* os efeitos de dispersão, investigada a partir de resíduos quadráticos individuais, foi maior do que essa proporção, calculada a partir da média dos resíduos quadráticos.

Por outro lado, a proporção média de identificação falsa de *algum* efeito de dispersão (*PIF*) é menor quando calculada a partir da média e, apenas em 2 das 32 condições experimentais, essa proporção foi ligeiramente superior quando calculada a partir dos resíduos quadráticos individuais.

Os resultados encontrados sugerem, então, apresentarem os métodos *H* e *HM* melhor desempenho (maior sensibilidade e maior especificidade), quando os resíduos quadráticos são substituídos pela sua média, em cada condição experimental.

Para complementar a análise, ainda se efetuam análises de variância sobre as variáveis-resposta do experimento: *PIC* e *PIF*, com a variância estabilizada para proporções de amostras de tamanho 1.000, pela utilização das transformações $\arcsen\sqrt{PIC}$ e $\arcsen\sqrt{PIF}$.

Os resultados encontrados concordam com o sugerido pela análise descritiva. Em relação aos valores das *PIC*'s, conforme mostra a tabela 5.2, a análise confirma a superioridade de desempenho dos métodos quando utilizada a média dos resíduos quadráticos

(efeito principal do fator L). A mesma tabela também mostra a interação desse fator L com os fatores A (quantidade de efeitos de locação), E (intensidade dos efeitos de dispersão), G (método) e, principalmente, com os fatores C (quantidade de efeitos de dispersão) e F (quantidade de replicações). A influência da utilização das médias dos resíduos quadráticos é maior quando existem menos efeitos de locação, mais efeitos de dispersão, efeitos de dispersão mais intensos e quando são usadas mais replicações. As PIC 's obtidas pelo método H são mais influenciadas por essa alteração do que as obtidas pelo método HM .

TABELA 5.2 – Análise do efeito do fator L (tratamento do resíduo quadrático) e suas interações com os demais fatores na variável transformada $\arcsen\sqrt{PIC}$

Fator ou interação	Soma dos quadrados (SQ)	Graus de liberdade (gl)	χ^2	valor- p
L	1,18197	1	4727,89	< 0,001
AL	0,00178	1	7,10	< 0,01
BL	0,00001	1	0,04	> 0,050
CL	0,02203	1	88,13	< 0,001
DL	0,00005	1	0,21	> 0,050
EL	0,00357	1	14,30	< 0,001
FL	0,16951	1	678,03	< 0,001
GL	0,00121	1	4,83	< 0,050

Nota: $\chi^2_{(1; 0,001)} = 10,83$; $\chi^2_{(1; 0,01)} = 6,64$; $\chi^2_{(1; 0,05)} = 3,84$.

Em relação aos valores de PIF 's, conforme a tabela 5.3, a análise confirma a superioridade de desempenho dos métodos, ao utilizar-se a média dos resíduos quadráticos (efeito principal do fator L), por apresentarem PIF 's significativamente menores. A mesma tabela também mostra esse fator L interagir com os fatores C (quantidade de efeitos de dispersão), E (intensidade dos efeitos de dispersão) e F (quantidade de replicações): os resultados são mais influenciados pela utilização das médias dos resíduos quadráticos quando existem mais efeitos de dispersão, efeitos de dispersão menos intensos e quando são usadas mais replicações.

O estudo de simulação realizado mostra que o uso da média dos resíduos quadráticos em cada condição experimental é mais adequado. Isto talvez possa explicar-se pelo fato de, nesses métodos, ser necessário calcular a soma dos logaritmos dos resíduos quadráticos. Considerando a ocorrência de resíduos quadráticos muito próximos de zero ser bastante freqüente e esses valores, quando transformados pela função logaritmica, poderem assumir valores muito pequenos, transformando-se em *inliers*, existe a possibilidade de os resultados ficarem distorcidos. A utilização da média aritmética dos resíduos quadráticos em cada

condição experimental evita a ocorrência de resultados muito próximos de zero, diminuindo a chance de aparecimento de *inliers*.

TABELA 5.3 – Análise do efeito do fator *L* (tratamento do resíduo quadrático) e suas interações com os demais fatores na variável transformada $\arcsen\sqrt{PIF}$

Fator ou interação	Soma dos quadrados (<i>SQ</i>)	Graus de liberdade (<i>gl</i>)	χ^2	valor- <i>p</i>
<i>L</i>	0,07265	1	290,60	< 0,001
<i>AL</i>	0,00035	1	1,40	> 0,050
<i>BL</i>	0,00034	1	1,36	> 0,050
<i>CL</i>	0,01661	1	66,44	< 0,001
<i>DL</i>	0,00011	1	0,44	> 0,050
<i>EL</i>	0,00102	1	4,08	< 0,050
<i>FL</i>	0,00365	1	14,60	< 0,001
<i>GL</i>	0,00035	1	1,40	> 0,050

Nota: $\chi^2_{(1; 0,001)} = 10,83$; $\chi^2_{(1; 0,05)} = 3,84$.

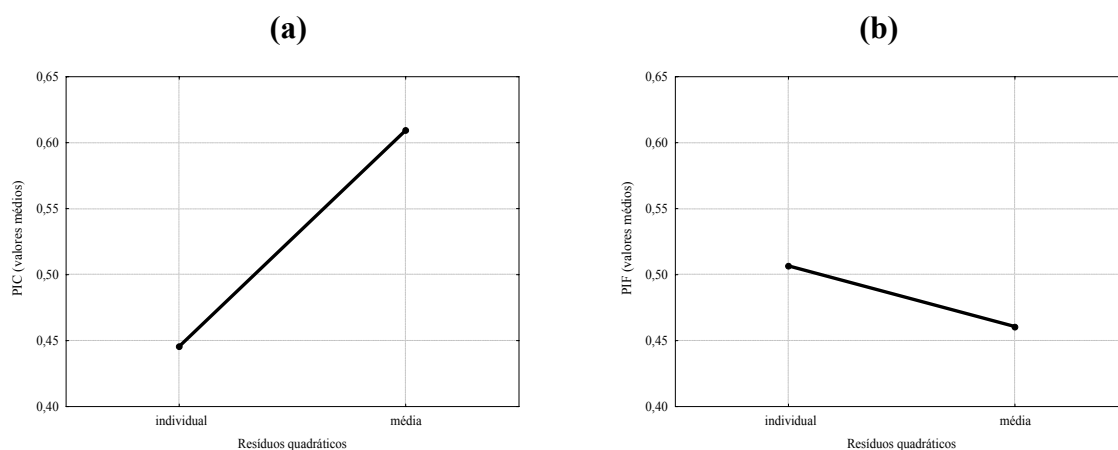


FIGURA 5.2 – Efeito do fator *L* (tratamento do resíduo quadrático) nas variáveis (a) *PIC* e (b) *PIF*

Os resultados encontrados parecem concordar com a extensão dessa estatística proposta por Nair e Pregibon (1988), que usa os desvios quadráticos em relação à média aritmética no lugar dos resíduos quadráticos individuais, conforme equação 5.5, para inferências sobre efeitos de dispersão (Z_n) em experimentos replicados:

$$Z_n = \sum_{m=1}^M (y_{nm} - \bar{y}_n)^2, \quad (5.5)$$

onde y_{nm} representa a m -ésima replicação na n -ésima condição experimental e \bar{y}_n pode representar tanto uma média conhecida como a média entre as M replicações. Concordam também com Barbetta (1998) e Barbetta *et al.* (2000) os quais utilizam o resíduo quadrático médio na modelagem da variância.

5.2 VARIÂNCIAS AMOSTRAIS E RESÍDUOS QUADRÁTICOS

O estudo de simulação descrito no Capítulo 3, cujos resultados são apresentados no Capítulo 4, sugere que a identificação de efeitos de dispersão, em experimentos com poucas replicações, seja realizada com maior fidedignidade se forem utilizados métodos com o emprego de resíduos quadráticos.

Porém, nesse estudo de simulação, para simplificar o processo, o modelo de locação era descrito com os fatores usados na simulação, embora os parâmetros tenham sido estimados com base na amostra. Nos casos reais, a identificação dos parâmetros também é feita com base na amostra. Tal procedimento pode ter interferido no resultado, superestimando a proporção de identificação correta e subestimando a proporção de identificação falsa de efeitos de dispersão nos métodos que utilizam resíduos quadráticos.

Um novo estudo de simulação é realizado para comparar o desempenho dos métodos que usam variâncias amostrais com o desempenho de métodos que usam resíduos quadráticos, sendo, nesses últimos, efetuada a identificação de fatores com efeito de locação por meio de uma análise de variância.

Para representar os métodos que utilizam variâncias amostrais são considerados os métodos *R* e *S*, enquanto, para representar os que usam resíduos quadráticos, são selecionados os métodos *BM* e *H*. Aqueles com o uso de resíduos quadráticos modificados não são considerados porque, além de serem mais trabalhosos, as simulações anteriores mostraram pouca vantagem em adotá-los.

5.2.1 Descrição do estudo de simulação

Este estudo de simulação foca um experimento com três fatores ($K=3$): x_1 , x_2 e x_3 , ensaiados a dois níveis (-1 e $+1$), segundo um projeto fatorial completo (projeto tipo 2^3 com replicações), contendo efeitos de locação e de dispersão. O uso de um projeto experimental menor como foco foi devido à necessidade de considerar todos os possíveis modelos de média. A média dos valores gerados é descrita por uma função aditiva:

$$\mu_y = \beta_1 x_1 + \beta_2 x_2, \quad (5.6)$$

sendo $\beta_1 = 2$ e $\beta_2 = 1$. Já o modelo para a variância é adotado como exponencial, mas com parâmetros variando conforme um fator do projeto experimental, chamado de C .

O fator C indica a quantidade de fatores com efeitos de dispersão. No nível $-I$ considera-se apenas um fator ter efeito de dispersão, enquanto, no outro ($+I$), dois fatores. Então, a variância é descrita por uma função multiplicativa ou log-linear, de acordo com as expressões:

$$\sigma_y^2 = e^{\theta_1 x_1} \quad (5.7)$$

e

$$\sigma_y^2 = e^{\theta_1 x_1 + \theta_2 x_2}, \quad (5.8)$$

sendo atribuído o valor $0,643$ para θ_1 e θ_2 , o qual altera o valor da variância, de um nível para outro em quatro vezes.

O fator F trata da quantidade de replicações: duas ou quatro.

Os níveis desses dois fatores do projeto de simulação são combinados segundo um projeto fatorial do tipo 2^2 , isto é: dois fatores (C e F), ensaiados a dois níveis ($-I$ e $+I$). O Quadro 5.2 sintetiza os níveis dos fatores C e F desse projeto de simulação.

Um novo programa computacional, tendo como objetivo encontrar as proporções de identificação correta de *todos* os efeitos de dispersão (PIC) e de identificação falsa de *algum* efeito de dispersão (PIF), é criado de acordo com o projeto lógico apresentado na figura 5.3.

QUADRO 5.2 – Descrição de fatores e níveis utilizados no estudo de simulação que compara métodos que usam variâncias amostrais com métodos que usam resíduos quadráticos

Fator	Descrição	Nível ($-I$)	Nível ($+I$)
C	Fatores com efeitos de dispersão	θ_1 (um efeito de dispersão)	θ_1 e θ_2 (dois efeitos de dispersão)
F	Quantidade de replicações	2	4

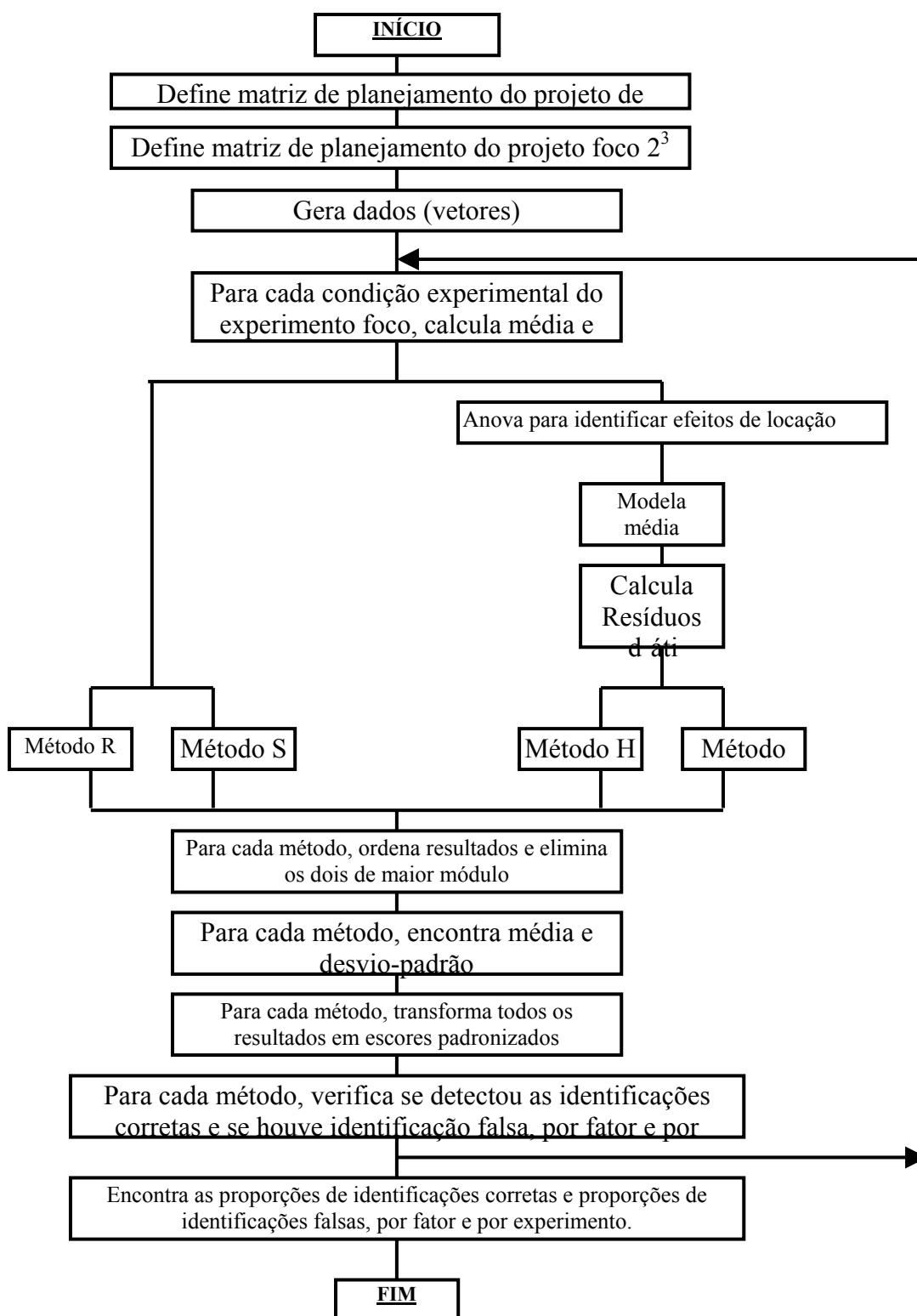


FIGURA 5.3 - Fluxograma do programa computacional utilizado na simulação realizada para comparar o desempenho dos métodos investigados que empregam resíduos quadráticos e variâncias amostrais

5.2.2 Resultados do estudo de simulação

5.2.2.1 Descrição do estudo

Inicialmente, as variáveis-resposta PIC e PIF (tabelas A5 e A6 no apêndice), encontradas em cada uma das 5.000 amostras geradas sob a mesma condição experimental, são organizadas, por método, em tabelas. Após, realizam-se análises de variância sobre essas variáveis, com a variância estabilizada para proporções de amostras de tamanho 5.000, pela utilização das transformações $\arcsen\sqrt{PIC}$ e $\arcsen\sqrt{PIF}$. Nesta análise acrescenta-se ao projeto experimental um novo fator, codificado como G, para identificar o método usado na identificação do efeito de dispersão. O estudo é complementado pela utilização de contrastes para a comparação do desempenho de métodos que usam variâncias amostrais e outros que utilizam resíduos quadráticos e para comparar métodos que empregam médias aritméticas com outros que usam médias geométricas.

5.2.2.2 Identificações corretas

Inicialmente, são encontrados os valores médios das PIC 's, por método. Tais resultados, apresentados na tabela 5.4. Essa primeira análise exploratória sugere que o método R apresenta melhor desempenho.

TABELA 5.4 - Proporção média, por método, de identificação correta de *todos* os efeitos de dispersão nas condições experimentais simuladas

Método	$PIC^{(1)}$
R	0,3055
S	0,2769
H	0,2968
BM	0,2974

(1) Médias calculadas a partir das 8 condições experimentais simuladas.

Uma análise da variância aplicada aos valores das PIC 's transformadas (tabela 5.5) detecta diferença significativa entre os métodos, além de indicar como significativa a interação entre esses métodos e o fator C (quantidade de efeitos de dispersão). A figura 5.4 evidencia a superioridade do valor médio da PIC encontrada pelo método R e inferioridade deste mesmo valor médio quando encontrado pelo método S . Já a figura 5.5 mostra que, se existe apenas um fator com efeito de dispersão, o método R apresenta melhor desempenho,

mas seu desempenho é muito afetado pela quantidade de efeitos, diminuindo consideravelmente na presença de dois efeitos de dispersão ativos. Nessa situação, o método *H* parece apresentar melhor desempenho.

TABELA 5.5 – Análise do efeito do fator *G* (método) e suas interações com os fatores *C* e *F* na variável transformada $\arcsen\sqrt{PIC}$

Fator ou interação	Soma dos quadrados (<i>SQ</i>)	Graus de liberdade (<i>gl</i>)	χ^2	valor- <i>p</i>
<i>G</i>	0,00190	3	37,98	$P < 0,001$
<i>CG</i>	0,00432	3	86,34	$P < 0,001$
<i>FG</i>	0,00001	3	0,24	$P > 0,05$

Nota: $\chi^2_{(3; 0,05)} = 7,81$; $\chi^2_{(3; 0,001)} = 16,27$.

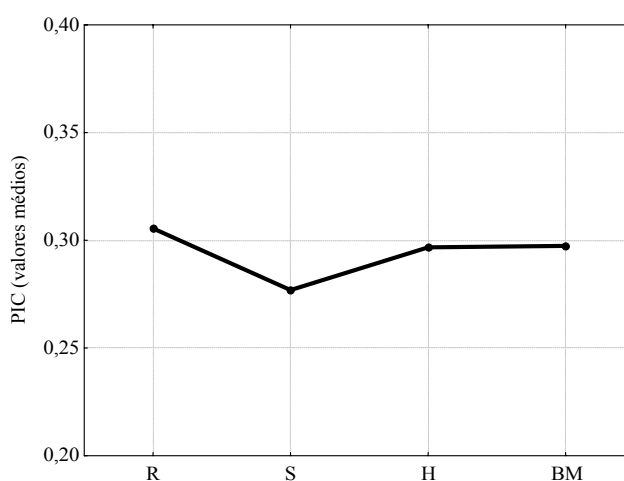


FIGURA 5.4 – Proporção média, por método, de identificação correta de *todos* os efeitos de dispersão nas condições experimentais simuladas

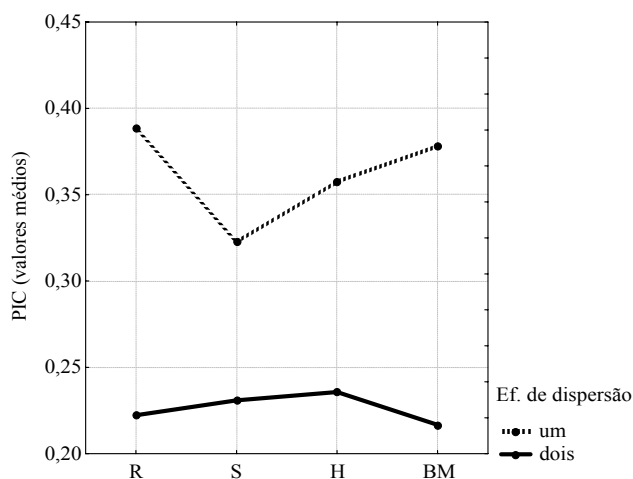


FIGURA 5.5 – Efeito da interação do fator *G* (método) com o fator *C* (efeitos de dispersão) sobre a variável resposta *PIC*, nas condições experimentais simuladas

Para complementar a análise, são usados contrastes ortogonais para comparar as *PIC*'s entre: (a) métodos que usam média aritmética (*R* e *BM*) e métodos com média geométrica (*S* e *H*); (b) métodos que usam variâncias amostrais (*R* e *S*) e métodos com resíduos quadráticos (*H* e *BM*). A tabela 5.6 mostra os coeficientes adotados nos contrastes para efetuar cada uma das comparações mencionadas. Os resultados da análise, mostrados na tabela 5.7, indicam existir evidências de diferença entre métodos que utilizam diferentes tipos de média: os que fazem uso da média aritmética apresentam maior sensibilidade. Indicam também não existir evidências de diferença entre métodos que usam diferentes medidas para estimar a variabilidade.

TABELA 5.6 – Coeficientes utilizados nos contrastes

Hipótese	Coeficientes			
	Método <i>R</i>	Método <i>S</i>	Método <i>H</i>	Método <i>BM</i>
(a)	-1	1	1	-1
(b)	-1	-1	1	1

TABELA 5.7 – Resultados da ANOVA aplicado aos contrastes

Hipótese	Soma dos quadrados (<i>SQ</i>)	Graus de liberdade (<i>gl</i>)	χ^2	Valor- <i>p</i>
(a)	0,00077	1	15,4	$P < 0,001$
(b)	0,00014	1	2,86	$p > 0,05$

Nota: $\chi^2_{(1; 0,05)} = 3,84$; $\chi^2_{(1; 0,001)} = 10,83$.

5.2.2.3 Identificações falsas

Inicialmente, encontram-se os valores médios das *PIF*'s, por método, que estão apresentados na tabela 5.8. Essa primeira análise exploratória sugere que o método *BM* apresenta melhor desempenho por apresentar menor *PIF*, seguido do método *R*.

TABELA 5.8 - Proporção média de identificação falsa de *algum* efeito de dispersão

Método	PIF
<i>R</i>	0,4537
<i>S</i>	0,5081
<i>H</i>	0,5094
<i>BM</i>	0,4293

A análise da variância aplicada aos valores das *PIF's* transformadas (tabela 5.9), entretanto, detecta diferença significativa entre os métodos, conforme o evidenciado no gráfico da figura 5.6. Detecta, também, a interação do método com o fator *C* (efeitos de dispersão) como significativa. A figura 5.7 evidencia que, quando existe apenas um efeito de dispersão, o método *H* apresenta o pior desempenho, enquanto se existir mais de um efeito deste tipo, o método *S* apresenta pior desempenho.

TABELA 5.9 – Análise do efeito do fator *G* (método) e suas interações com os fatores *C* e *F* na variável transformada $\arcsen\sqrt{PIF}$

Fator ou interação	Soma dos quadrados (<i>SQ</i>)	Graus de liberdade (<i>gl</i>)	χ^2	valor- <i>p</i>
<i>G</i>	0,02170	3	434,02	$P < 0,001$
<i>CG</i>	0,00136	3	27,28	$P < 0,001$
<i>FG</i>	0,00019	3	3,8	$P > 0,05$

N Nota: $\chi^2_{(3; 0,05)} = 7,81$; $\chi^2_{(3; 0,001)} = 16,27$.

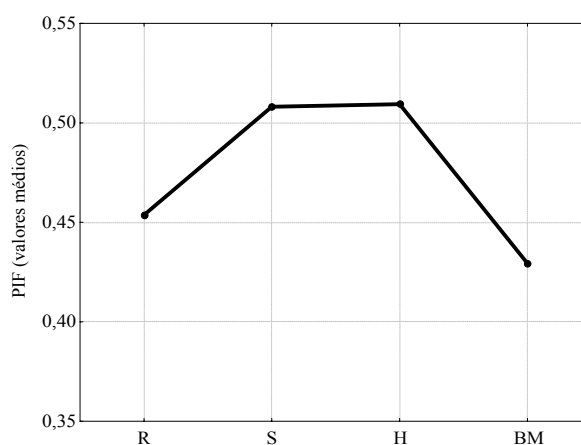


FIGURA 5.6 - Proporção média, por método, de identificação falsa de *algum* efeito de dispersão nas condições experimentais simuladas

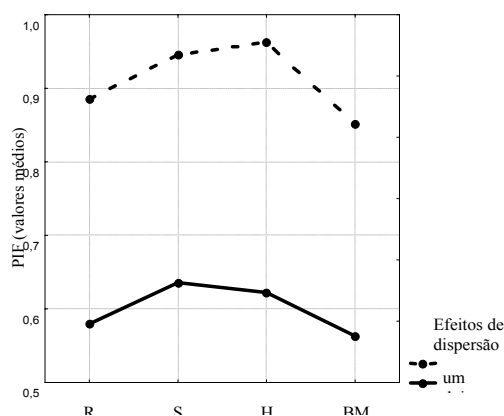


FIGURA 5.7 – Efeito da interação do fator *G* (método) com o fator *C* (efeitos de dispersão) sobre a variável resposta *PIF*, nas condições experimentais simuladas

Contrastes ortogonais são usados para comparar as *PIF*'s médias de: (a) métodos com o uso de média aritmética (*R* e *BM*) e métodos com média geométrica (*S* e *H*); (b) métodos que usam variâncias amostrais (*R* e *S*) e métodos que empregam resíduos quadráticos (*H* e *BM*).

Adotando os coeficientes mostrados na tabela 5.7, constata-se existir diferença significativa entre métodos que utilizam diferentes tipos de média: aqueles que estimam efeitos de dispersão por meio de médias aritméticas apresentam *PIF* inferior aos métodos com emprego de médias geométricas, apresentando proporções médias, respectivamente, iguais a 0,4415 e 0,5088. Constata-se também que a *PIF* média de métodos que empregam variâncias amostrais na estimação de efeitos de dispersão (0,4809) difere significativamente da *PIF* média encontradas pelos métodos com uso de resíduos quadráticos (0,4694). Estes últimos apresentam melhor especificidade (tabela 5.10).

TABELA 5.10 – Resultados da ANOVA aplicado aos contrastes

Hipótese	Soma dos quadrados (<i>SQ</i>)	Graus de liberdade (<i>gl</i>)	χ^2	valor- <i>p</i>
(a)	0,02044	1	408,9	$P < 0,001$
(b)	0,00055	1	11,04	$P < 0,001$

Nota: $\chi^2_{(1; 0,001)} = 10,83$.

5.2.2.4 Considerações sobre os resultados

Nessa última simulação, em relação à sensibilidade, é possível detectar diferença significativa entre os métodos investigados. Constata-se que o método *R* apresenta resultado superior, seguido do método *BM*, o que se acentua quando existe apenas um fator com efeito de dispersão. A significância do resultado também é constatada na comparação entre métodos com diferentes tipos de média: métodos que usam média aritmética apresentam melhor desempenho do que aqueles que fazem uso da média geométrica.

Em relação à especificidade, é possível detectar diferença entre os métodos investigados e constatar que métodos com uso de média aritmética apresentam melhor resultado. O método *BM* falha menos, seguido do método *R*, ambos estimando efeitos de dispersão por meio de médias aritméticas. Também é detectada diferença significativa entre

métodos que usam diferentes medidas de variabilidade: métodos que usam resíduos quadráticos falham menos.

Portanto, os resultados parecem indicar quando são usadas poucas replicações, o desempenho de um método ser mais influenciado pelo tipo de média empregado na estimativa do efeito: média aritmética ou média geométrica. Métodos que estimam efeitos de dispersão fazendo uso de médias aritméticas apresentam maior sensibilidade e maior especificidade do que os utilizadores de médias geométricas. Já os métodos que fazem uso de resíduos quadráticos não apresentam maior sensibilidade, embora apresentem maior especificidade, do que os utilizadores de variâncias amostrais.

ESTRATÉGIA PARA IDENTIFICAR EFEITOS DE DISPERSÃO

Este capítulo apresenta uma estratégia para identificar efeitos de dispersão em experimentos do tipo 2^k , com poucas replicações ($2 \leq M \leq 4$), ilustrando-a com um estudo de caso.

6.1 CONSIDERAÇÕES SOBRE OS ESTUDOS DE SIMULAÇÃO

Foram realizados dois estudos de simulação para identificar o método de melhor desempenho em experimentos com poucas replicações, envolvendo seis métodos: *S* e *R*, medindo a variabilidade por meio de variâncias amostrais; métodos *H* e *BM*, que medem a variabilidade por meio de resíduos quadráticos e *HM* e *BH*, cuja medida de variabilidade é dada por meio de resíduos quadráticos modificados.

O primeiro estudo de simulação apontou para os métodos com o uso de médias aritméticas entre resíduos quadráticos, simples ou modificados, como apresentando melhor desempenho: mais especificamente para os métodos *BM* e *BH*. A superioridade deles ainda se evidencia mais quando existem menos efeitos de locação, menos efeitos de dispersão, efeitos de dispersão menos intensos e menos replicações.

Embora tenha existido diferença significativa entre as proporções de identificações falsas de métodos empregando diferentes tipos de resíduos quadráticos, a favor dos resíduos quadráticos modificados, considerou-se ser pequena a diferença. A utilização de resíduos quadráticos modificados é justificada em experimentos não-replicados pela eliminação da correlação existente entre os resíduos, permitindo a realização de testes estatísticos formais, mesmo em experimentos não-replicados. Porém, na identificação dos efeitos baseada no gráfico de probabilidade normal, os resultados não foram muito influenciados por esse fato.

O primeiro estudo de simulação apresentado neste trabalho apresentou uma limitação muito forte: na estimação de efeitos de dispersão a partir de resíduos quadráticos, foi construído um modelo de locação com os efeitos realmente existentes,

embora suas intensidades tenham sido estimadas com base na amostra. Em função disto, não foram comparados os resultados de métodos que usam variâncias amostrais e resíduos quadráticos, sendo apenas constatado, entre os métodos que usam variâncias amostrais, o método *R* apresentar melhor desempenho que o método *S*.

Realizou-se novo estudo de simulação para comparar o desempenho de métodos com esses dois tipos de medida para avaliar a variabilidade: *S* e *R* (variâncias amostrais) e *H* e *BM* (resíduos quadráticos). Não se detectou evidência de diferença entre eles. Apenas se verificou novos indícios de possuírem melhor desempenho os que usam média aritmética que outros com média geométrica: entre os métodos analisados, o método *R* apresenta sensibilidade ligeiramente superior, enquanto o método *BM* apresenta especificidade também ligeiramente superior.

Também existem indícios de os métodos que empregam médias aritméticas apresentarem viés estrutural, isto é, o efeito de dispersão do *k*-ésimo fator depende de outros fatores além dele mesmo, fazendo que possam ser detectados falsos efeitos de dispersão ou efeitos de dispersão não sejam detectados. A identificação de falsos efeitos de dispersão de fatores resultantes da interação entre dois outros efeitos de dispersão ativos é um erro bastante freqüente em decorrência desse tipo de viés.

Mesmo assim, o fator determinante de bom desempenho em métodos que identificam efeitos de dispersão em experimentos com poucas replicações parece ser a utilização de média aritmética, não sendo isso esperado em função do modelo adotado para descrever a variância: modelo log-linear.

6.2 A estratégia

Para identificar efeitos de dispersão em experimentos com poucas replicações, propõe-se a seguinte estratégia (figura 6.1):

- i) ajustar um modelo de locação por meio da aplicação do método dos mínimos quadrados ordinários (*MQO*), identificando o conjunto \mathcal{L} de efeitos de locação;

Seguindo Brenneman e Nair (2001), sugere-se que se \mathcal{L} contém um conjunto fechado \mathcal{L}' com $N/2$ ou mais elementos, então a identificação dos efeitos de dispersão é difícil, sendo necessário aumentar o número de ensaios - alterar o projeto experimental.

- j) avaliar a adequação do modelo encontrado;

- k) se a adequação do modelo de locação for boa, estimar os efeitos de dispersão baseado nos resíduos desse modelo, pelo método *BM* (Box e Meyer, 1986a), identificando os efeitos ativos pelo gráfico de probabilidade normal dos efeitos;
- l) se houver dúvida sobre a adequação do modelo de locação, estimar os efeitos de dispersão, com base nas variâncias amostrais, pelo método *R* (Nair e Pregibon, 1988);
- m) se nas etapas (c) ou (d) for identificado mais de um efeito de dispersão, refazer a análise dos efeitos de dispersão pelo método *H* [baseado em Harvey (1976)], ou pelo método *S* (baseado em Bartlett e Kendall, 1946).

Muitas vezes, há o interesse em estimar os efeitos de dispersão e, também, em construir modelos para locação e dispersão. Nesse caso, conforme abordam Box e Meyer (1986a), Nair e Brenneman (2001), dentre outros, é conveniente usar procedimentos estatísticos mais elaborados, como os métodos de máxima verossimilhança restrita ou mínimos quadrados generalizados (MQG). Ainda podem ser usados os modelos lineares generalizados (LEE e NELDER, 2003).

Descreve-se, abaixo, passos complementares para a construção de modelos, baseado no método dos mínimos quadrados generalizados (MQG):

- n) após a definição do conjunto \mathcal{D} de efeitos de dispersão, construir uma equação para a variância usando o modelo log-linear (equação 2.29), no qual as estimativas dos elementos de θ podem ser obtidas por MQO;
- o) refazer o modelo de locação, usando mínimos quadrados generalizados (MQG), sendo, a cada observação, atribuído um peso fornecido pelo inverso da variância predita pelo modelo definido no passo (e). Nesta etapa o conjunto de fatores com efeitos de locação pode ser alterado;
- p) se o modelo de dispersão estiver sendo construído com os resíduos do modelo de locação, os passos (c), (d) (e) e (f) poderão ser repetidos, fazendo-se os ajustes necessários nos conjuntos \mathcal{L} e \mathcal{D} até o processo se estabilizar. Alguns autores sugerem apenas uma ou duas iterações, devido ao problema da convergência na presença de valores discrepantes.

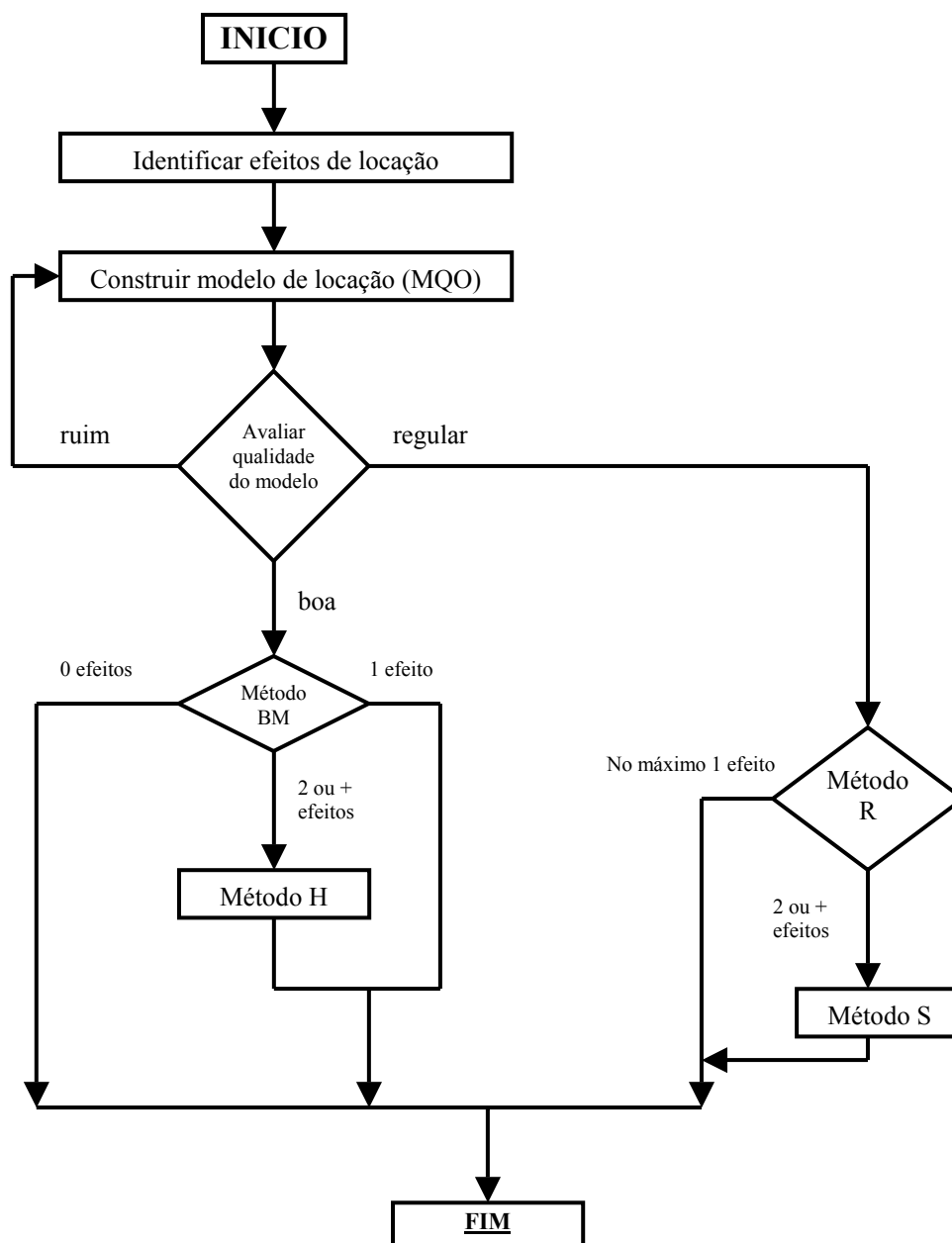


FIGURA 6.1 – Fluxograma da estratégia para identificar possíveis efeitos de dispersão

6.3 ESTUDO DE CASO

Na construção civil, o concreto é um material extremamente utilizado por sua versatilidade, possibilitando a formação de peças de qualquer forma e dimensão e apresentando, também, boa resistência e durabilidade. Além disso, apresenta baixo custo quando comparado a outros materiais estruturais.

Tradicionalmente, de acordo com Pianca (1977, p.52), os concretos são definidos como misturas de aglomerante (cimento) e agregados (areia e brita), em determinadas proporções, além da água necessária para a formação da pasta. Já os concretos de alto

desempenho apresentam propriedades superiores às dos concretos tradicionais, sobretudo quanto à durabilidade e à resistência, com baixo consumo de água e alto consumo de finos (componentes formados por partículas infinitesimais). A melhoria da qualidade é obtida pela otimização de sua composição com o uso de adições e aditivos capazes de permitir a redução da relação água/cimento, melhorando a permeabilidade e a resistência.

A qualidade do concreto pode ser avaliada por suas propriedades que são diferentes conforme o produto esteja no estado fresco ou no estado endurecido. No estado fresco, ocorrido desde a colocação da água até o adensamento na forma final, a característica mais importante é a trabalhabilidade por conferir as condições necessárias para o produto ser produzido e manuseado com o máximo de homogeneidade, possibilitando obter sua capacidade máxima no adensamento. Já o concreto endurecido deve apresentar resistência mecânica (resistência à compressão, à tração, à flexão e ao cisalhamento) e durabilidade compatíveis com as condições do projeto, sendo a impermeabilidade uma das condições de durabilidade, se estiver exposto a agentes agressivos.

Várias pesquisas têm sido desenvolvidas visando a utilização de materiais alternativos na fabricação de concretos com possibilidade de aumentar sua resistência ou diminuir seu custo, sem prejuízo da trabalhabilidade, entre os quais, se têm destacado as adições de cinza de casca de arroz e de sílica ativa, por serem consideradas pozolanas altamente reativas. Tais adições minerais diminuem a permeabilidade em função da redução dos poros existentes na pasta e alteram a sua microporosidade, influenciando benéficamente a durabilidade e a resistência mecânica.

O aproveitamento da cinza de casca de arroz como material alternativo na fabricação de concretos e argamassas tem sido objeto de estudo de pesquisadores do Laboratório de Resistência dos Materiais (LRM) da Universidade Católica de Pelotas desde 1996, pois esta região do Rio Grande do Sul possui características ambientais extremamente propícias ao cultivo de arroz.

De acordo com informações fornecidas pelo IBGE, esse estado é o maior produtor nacional de arroz, com cerca de 4.800.000 toneladas por ano, sendo, no município de Pelotas, realizado o beneficiamento de, aproximadamente, 960.000 toneladas anuais do produto. É prática usual nos locais de beneficiamento de arroz ou qualquer outro tipo de cereal, a queima de sua casca para solucionar o problema da sua disposição no meio ambiente. Porém, considerando 20% do peso do grão devido à casca e, após a combustão completa, obter-se, em média, 20% de cinza, geram-se anualmente em torno de 38.400 toneladas de cinza, ocasionando sérios problemas ambientais.

Este estudo é realizado para avaliar os efeitos de alguns parâmetros do processo produtivo do concreto de alto desempenho, moldado com a adição de cinza de casca de arroz, sobre sua resistência à compressão. De acordo com Isaia (1988, p.34), praticamente todas as propriedades do concreto dependem dessa característica, razão pela qual foi escolhida para mensurar a qualidade dele.

Embora a resistência à compressão seja uma característica do tipo maior-é-melhor, na produção do concreto é realizado um estudo de dosagem para o produto atingir uma determinada resistência. O resultado considera, entre outros fatores, a heterogeneidade da mistura. Uma grande variabilidade resulta em um maior consumo de cimento e, portanto, encarecimento do concreto.

Além disso, no cálculo estrutural, considera-se que o concreto deve apresentar uma resistência mínima. Se existir grande variabilidade, é mais difícil garantir a ocorrência desse mínimo, podendo dar-se o colapso da estrutura. Para evitar tal possível colapso, os calculistas adotam um coeficiente de segurança (40%), aumentando consideravelmente os custos. Todas as estruturas ficam super-dimensionadas. Por isso, reduzir variabilidade é fundamental para a redução de custos. Embora a adoção de um coeficiente de segurança seja necessária, em outros materiais com maior homogeneidade, como o aço, por exemplo, são utilizados coeficientes bem menores.

6.3.1 O produto avaliado

A qualidade do produto final, o concreto, deve atender às exigências de segurança, conforto e durabilidade e isso depende do controle tecnológico sobre todas as etapas de sua produção:

- projeto
- matéria-prima
- dosagem
- mistura
- transporte e lançamento
- adensamento
- cura

No presente estudo, a característica funcional representativa da qualidade é a resistência à compressão, avaliada em corpos-de-prova cilíndricos, de dimensões 10 x 20cm, compostos por agregados, aglomerante, aditivo, adição e água.

O aglomerante (cimento), por ser o material mais importante, influenciando diretamente a resistência e a durabilidade, deve ser utilizado em quantidade, qualidade e tipo adequados. Nesse trabalho foi empregado o cimento Portland de alta resistência inicial (CP-ARI), de pega rápida, por não apresentar adições minerais e possuir pequenas variações nas suas propriedades, em lotes provenientes de diferentes linhas de produção.

Usam-se os agregados para transmitir as tensões aplicadas ao concreto, razão pela qual devem ter resistência superior ao desejado e diminuir o efeito das variações volumétricas devidas à retração. Os agregados miúdos (areia) precisam ser limpos, sem substâncias nocivas (torrões de argila, material carbonoso, material pulverulento e impurezas orgânicas), além de possuir granulometria uniforme com módulo de finura variando entre 2 e 3. Os agregados graúdos (brita) necessitam de forma e tamanho adequados, de maneira a diminuir a quantidade de vazios, e dimensão máxima característica, em acordo com as dimensões da fôrma e da armadura. Nessa pesquisa, utilizou-se agregado graúdo de origem granítica, com dimensão máxima característica de 25mm, e agregado miúdo de origem quartzoza, com módulo de finura 3,08.

As águas utilizadas na fabricação do concreto não devem conter substâncias capazes de reagir com o cimento, trazendo efeitos colaterais danosos ao concreto. No estudo, foi usada água proveniente da rede de abastecimento local.

Na composição dos concretos podem ser usados aditivos, normalmente em pequena quantidade, com o fim de fazer aparecer ou reforçar determinadas características. Entre os vários tipos de aditivo estão os superplastificantes (à base de naftaleno sulfonado), usados no trabalho em algumas condições experimentais para melhorar a trabalhabilidade do concreto.

Também se empregou a adição cinza da casca de arroz, queimada a 850°C, proveniente de uma indústria beneficiadora de arroz de Pelotas/RS. Como as características físicas e químicas da cinza são influenciadas pela temperatura de queima, procedeu-se a uma análise da morfologia do material por meio de ensaios de difração de raios-x, detectando-se a presença de material cristalino, identificado basicamente como cristobalita e tridimita. Teoricamente, espera-se que a utilização da cinza de casca de arroz, por ser rica em sílica (entre 90 e 95%), aumente a resistência do concreto, por provocar os efeitos pozolânico e *microfiller*. Os primeiros resultam de uma reação entre a sílica e o hidróxido de cálcio que, em contato com a água, melhora as propriedades aglomerantes da pasta de cimento, enquanto o efeito *microfiller* diminui a quantidade de vazios existente na mistura.

Após, considerando as características do projeto e dos materiais disponíveis, assim como do equipamento e mão-de-obra, deve-se realizar a dosagem, definindo a quantidade exata de cada material na mistura. Os traços utilizados são determinados levando em conta a relação água-aglomerante desejada e o teor de adição a ser incorporado, sendo empregada a metodologia de dosagem do IPT/EPUSP. Os parâmetros de dosagem utilizados são a obtenção de um abatimento de cone de 7 ± 1 cm e a manutenção de um teor de argamassa de 51% (mais detalhes em HELENE e TERZIAN, 1992).

A tabela 6.1 apresenta os traços resultantes da dosagem. Dão-se os agregados, a água e o cimento em peso (kg). A quantidade de água foi fixada em relação ao peso do cimento, levando em consideração a parte líquida do aditivo. A redução da trabalhabilidade, em algumas condições experimentais, foi compensada com o uso de aditivo superplastificante em teores variáveis relativamente ao peso de cimento para os traços com incorporação de adições. A tabela 6.2 apresenta os resultados de abatimento obtidos e a quantidade de aditivo utilizada em cada traço.

Definida a dosagem, passa-se à operação de fabricar o concreto propriamente dito, formado pela mistura dos componentes até se conseguir uma massa uniforme. A maneira de misturar ou amassar mais eficiente e normalmente utilizada é a mecânica, com o uso de equipamento denominado betoneira. Influenciam, na qualidade da mistura, o volume da betoneira e da betonada, a velocidade do tambor, o tempo de mistura e a ordem de colocação dos materiais.

TABELA 6.1 - Traços utilizados para confecção dos concretos estudados.

MISTURA	Água/ (cimento+ adição)	Cimento (kg)	Agregado Graúdo (kg)	Agregado Miúdo (kg)	Consumo de Cimento (kg/m ³)
CPV-ARI	0,35	1	0,887	1,813	593
	0,60	1	3,335	4,165	264
CCA (10%)	0,35	1	0,887	1,813	593
	0,60	1	3,335	4,165	264

TABELA 6.2 - Teor de aditivo superplastificante empregado e abatimentos obtidos.

MISTURA (adição %)	Tempo de mistura (segundos)	Água/ (cimento+adição)	Aditivo (%)	Slump (mm)
CPV-ARI	100	0,35	0,53	140
		0,60	-	30
	300	0,35	0,35	65
		0,60	-	25
CCA (10%)	100	0,35	0,21	55
		0,60	0,44	85
	300	0,35	0,50	50
		0,60	-	45

Emprega-se uma betoneira com eixo vertical, com a seguinte ordem de colocação dos materiais: 100% de agregado graúdo; 30% de água; 100% de cimento + 100% de adição + 30% de água; restante de água; 100% de areia e aditivo superplastificante (quando necessário). O tempo de mistura é contado a partir do lançamento, na betoneira, do último material sólido e varia conforme a condição experimental. Transcorrido esse tempo, é medido o abatimento do tronco de cone segundo a NBR 7223 (ABNT, 1982), fazendo-se ajustes da consistência quando necessário, mediante a utilização do aditivo. O teor de aditivo empregado em relação à massa de (cimento + adição) e os abatimentos obtidos apresentam-se na tabela 6.2.

Depois da mistura ser considerada pronta, é necessário transportá-la ao local de aplicação para acomodá-la na forma. O sistema escolhido depende do tipo, da localização e do volume da obra e deve permitir o lançamento direto na forma, eliminando o depósito intermediário e, com a rapidez, evitando a perda da trabalhabilidade.

Depois de colocado nas formas, existe a necessidade de compactar o produto para ele se acomodar e ficar com a menor quantidade possível de vazios. O ar existente sai e a mistura vai-se arrumando melhor dentro da forma até ficar com uma superfície lisa de cimento e

finos. A argamassa deve ir descendo sem a mistura perder sua homogeneidade, o que pode acontecer se o agregado graúdo descer (excesso de vibração é pior que falta de vibração).

O adensamento, quando realizado manualmente, consiste num simples apiloamento a mão, realizado com o auxílio de um soquete metálico, em obras pequenas ou em lugares onde não existe energia. Já o adensamento mecânico deve apresentar melhor eficiência, podendo ser interno (vibrador de imersão) ou externo (régua de superfície e mesa vibratória).

Depois do adensamento, ainda existe necessidade de cuidados com a última etapa do processo produtivo, denominada cura, ocorrendo nos primeiros 28 dias. Nessa etapa, o cimento continua a reagir com a água, que o hidrata, enquanto a mistura vai adquirindo suas propriedades. Em temperaturas muito elevadas, deve-se tomar cuidado para a água não evaporar rápido demais; em temperaturas muito baixas, deve-se cuidar para a água não congelar. Quanto mais lenta e perfeita a cura do concreto, melhores serão suas características.

Neste estudo, os corpos-de-prova moldados mantiveram-se no ambiente de laboratório por aproximadamente vinte e quatro horas. Ao retirar as formas (24 horas após a moldagem), eles foram colocados em uma câmara úmida com umidade relativa superior a 95% e temperatura de $(22 \pm 2^\circ\text{C})$, onde permaneceram até a data da ruptura, realizada aos 7 ou aos 28 dias, numa prensa hidráulica WPM com controle manual. A resistência à compressão simples foi medida de acordo com as prescrições da NBR 5739 (ABNT 1980).

6.3.2 Possíveis causas da variabilidade da resistência à compressão do concreto

A resistência real de um concreto é praticamente impossível de ser determinada, tendo em vista as características de seu processo produtivo. De todas as características de qualidade do concreto, a mais influenciada pelas variações de seu processo produtivo e mais simples de ser medida é a resistência à compressão, mensurada em corpos-de-prova moldados, curados e ensaiados conforme a normalização.

Na dosagem do concreto, considera-se sua resistência à compressão apresentar distribuição normal $N(\mu_c, \sigma_c^2)$, sendo definida uma resistência característica à compressão que correspondente, no mínimo, ao valor obtido por 95% dos resultados. Quanto maior o desvio-padrão, mais afastada da média estará a resistência característica e, conseqüentemente, mais caro será o concreto a ser produzido para garantir a mesma resistência especificada. Se a variabilidade for pequena e controlada, é possível produzir concretos com maior economia, sem prejuízo da segurança.

Segundo Isaia (1988, p.74), o American Concrete Institute considera que as principais causas da variação da resistência do concreto são:

- a) variações na relação água/cimento devido ao deficiente controle da água adicionada, variação excessiva do grau de umidade dos agregados, adição suplementar de água para manutenção do abatimento especificado;
- b) variações na quantidade de água devida à granulometria do agregado, absorção e forma dos grãos, conteúdo de ar; tempo de entrega e temperatura;
- c) variações nas características e proporções dos materiais;
- d) variações no transporte, lançamento e adensamento;
- e) variações na temperatura e umidade de cura;
- f) procedimentos incorretos de amostragem;
- g) variações devidas à técnica de moldagem;
- h) manuseio e cura inicial dos corpos recém-moldados;
- i) forma de qualidade deficiente;
- j) variações na cura dos corpos-de-prova relativas à temperatura, teor de umidade, e transporte até o laboratório;
- k) procedimentos deficientes de ensaio, como capeamento dos topos e aferição da prensa.

6.3.3 O experimento

Na pesquisa desenvolvida, o experimento foi planejado na forma de um projeto fatorial completo com cinco fatores (*A*, *B*, *C*, *D* e *E*) ensaiados em dois níveis (*-1* e *+1*), sintetizados na tabela 6.3. A amostra foi composta por 96 ensaios pois, em cada uma das 32 condições experimentais, foram realizadas três replicações.

TABELA 6.3 – Descrição dos fatores utilizados no experimento.

Fatores		Nível inferior (-1)	Nível superior (1)
<u>Relação água/aglomerante</u>	A	0,35	0,60
Quantidade de adição (%)	<i>B</i>	0	10
Tempo de mistura (segundos)	<i>C</i>	100	300
Tipo de adensamento	<i>D</i>	manual	mecânico
Idade (dias)	<i>E</i>	07	28

O fator *A* representa a relação água/aglomerante, indicando a quantidade de água a ser adicionada à mistura. Quanto maior a quantidade de água, menor a resistência esperada do concreto, podendo, entretanto, ser ela aumentada com o aumento da proporção de cimento. As relações utilizadas foram: 0,35, menor quantidade de água (*-1*), e 0,60, maior quantidade de

água (*I*).

O fator *B* representa a utilização (*I*) ou não (*-I*) da adição cinza de casca de arroz que, em caso positivo, foi utilizada com o teor de 10%. Espera-se que o emprego da cinza de casca de arroz aumente a resistência, melhore a trabalhabilidade e a homogeneidade.

O fator *C* refere-se ao tempo de mistura, medido em segundos. A qualidade de um concreto pode ser diminuída se a mistura obtida não for homogênea e não ocorrer a dispersão adequada das partículas de cimento na água. Os tempos considerados foram de 100 segundos (*-I*) e 300 segundos (*I*). Quando o concreto for mais seco, maior o tempo de mistura necessário; quanto maior o tempo de mistura, maior a resistência do concreto. Entretanto, o custo decorrente do aumento de tempo, após 3 minutos, não se justifica pelo aumento da resistência que se torna muito pequena.

O fator *D* considera o tipo de adensamento, realizado manualmente, com o auxílio de um soquete metálico (*-I*), ou mecanicamente, com a utilização de um vibrador de imersão (*I*). O adensamento, quando bem executado, reduz a quantidade de vazios, levando a um produto com maior resistência. Espera-se, portanto, que o adensamento mecânico forneça melhores resultados.

O fator *E* representa a idade dos corpos-de-prova, medida em dias, rompidos aos 7 dias (*-I*) e aos 28 dias (*I*). Em uma dosagem, considera-se que a resistência à compressão esperada para o concreto é a sua resistência aos 28 dias, razão pela qual esta é a idade padrão de referência. No controle de sua resistência, costuma-se romper corpos-de prova aos 7 dias, para averiguação antecipada de sua resistência, em função do tipo de cimento utilizado, o que neste experimento podemos considerar 70% da resistência aos 28 dias. Espera-se que, aos 28 dias o concreto apresente maior resistência, assim como maior variabilidade.

6.3.4 Análise dos resultados

O experimento, cujas respostas estão apresentadas na tabela A7 do apêndice, é analisado de acordo com a estratégia proposta no capítulo anterior. Inicialmente, buscou-se identificar os efeitos de locação ativos para ajustar um modelo, por meio do método MQO, aplicado aos valores da resistência à compressão (tabela 6.4). Foram consideradas apenas interações de segunda ordem. Um teste *F* identifica os fatores *A*, *D* e *E*, além da interação *AE*, como apresentando efeitos significativos. Assim, tem-se o seguinte conjunto de efeitos de locação: $\mathcal{L} = \{I, A, D, E \text{ e } AE\}$, onde *I* se refere à média geral.

O modelo de locação, obtido pelo método dos mínimos quadrados, foi:

$$\hat{\mu}_y = 42,422 - 10,488A - 2,330D + 1,842E + 1,456AE . \quad (6.1)$$

A significância do modelo ajustado é avaliada pelo teste F ($F = 117,65$; $df_1 = 4$; $df_2 = 91$; $valor-p = 0,0000$), sendo testada a hipótese nula de que todos os coeficientes do modelo são nulos. Como o experimento apresenta replicações, também é utilizado um teste de falta de ajuste ($F = 1,80$; $df_1 = 27$; $df_2 = 64$; $valor-p = 0,0278$) para testar a hipótese nula de o modelo ser adequado por não apresentar a variabilidade residual significativamente maior do que a variabilidade devida ao erro puro. Embora o primeiro teste tenha apresentado resultado satisfatório, o segundo detecta falta de ajuste.

A análise é complementada pela construção de dois gráficos: um gráfico de resíduos, relacionando estimativas de resíduos com os correspondentes valores preditos (gráfico (a) da figura 6.2), e um de probabilidade normal desses resíduos (gráfico (b) desta mesma figura). O gráfico (a) sugere a existência de maior variabilidade entre os resíduos associados a maiores valores preditos, enquanto o gráfico (b) sugere que esses resíduos seguem uma distribuição normal.

Foi verificada uma possível modificação na métrica da resposta, através da transformação logarítmica, mas também não se obteve um modelo de qualidade satisfatória.

TABELA 6.4 – Teste F aplicado aos dados originais para identificação dos efeitos de locação

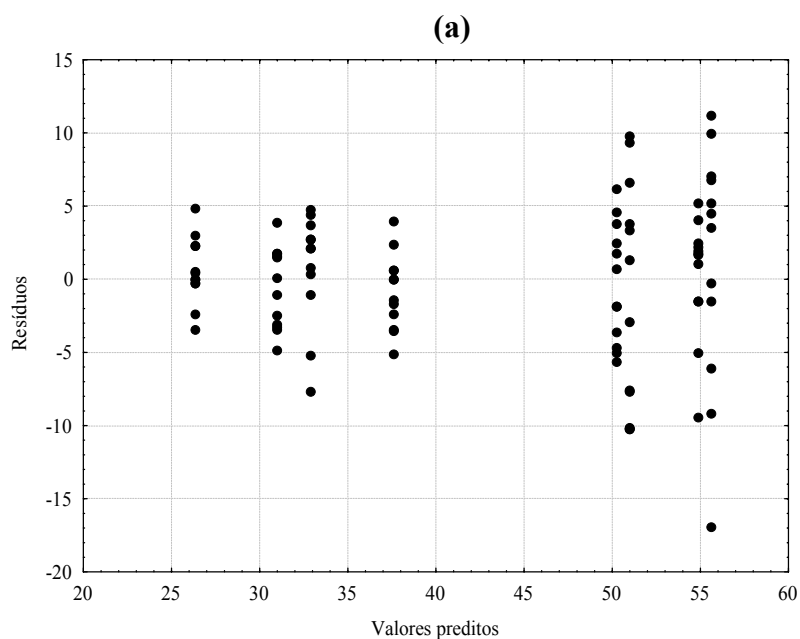
Fontes de Variação	Soma dos quadrados (SS)	Graus de liberdade (df)	QUADRADO		
			O Médio (MS)	F	valor-p
<i>A</i>	10559,86	1	10559,86	530,13	0,0000
<i>B</i>	29,83	1	29,83	1,49	0,2255
<i>C</i>	4,07	1	4,07	0,20	0,6527
<i>D</i>	521,22	1	521,22	26,16	0,0000
<i>E</i>	325,64	1	325,64	16,34	0,0001
<i>AB</i>	16,16	1	16,16	0,81	0,3710
<i>AC</i>	56,81	1	56,81	2,85	0,0961
<i>AD</i>	49,75	1	49,75	2,49	0,1189
<i>AE</i>	203,38	1	203,38	10,21	0,0021
<i>BC</i>	70,85	1	70,85	3,55	0,0638
<i>BD</i>	29,56	1	29,56	1,48	0,2276
<i>BE</i>	48,81	1	48,81	2,45	0,1224
<i>CD</i>	0,05	1	0,05	0,00	0,9598
<i>CE</i>	24,51	1	24,51	1,23	0,2714
<i>DE</i>	3,40	1	3,40	0,17	0,6808

Erro puro	1274,82	64	19,92
-----------	---------	----	-------

Nova avaliação é realizada considerando interações de até terceira ordem. O efeito da interação BCD também é detectado como significativo ($F = 8,05$; $df_1 = 1$; $df_2 = 64$; $valor-p = 0,0060$) e o novo conjunto de efeitos de locação é: $\mathcal{L} = \{I, A, D, E, AE \text{ e } BCD\}$, onde I se refere à média geral. O novo modelo de locação encontrado é:

$$\hat{\mu}_y = 42,422 - 10,488A - 2,330D + 1,842E + 1,456AE - 1,292BCD . \quad (6.2)$$

A significância desse modelo é avaliada pelo teste F ($F = 101,63$; $df_1 = 5$; $df_2 = 90$; $valor-p = 0,0000$), e pelo teste de falta de ajuste ($F = 1,56$; $df_1 = 26$; $df_2 = 64$; $valor-p = 0,0754$). Em ambos os testes, os resultados encontrados são satisfatórios, embora no teste de falta de ajuste, o $valor-p$ seja muito baixo, mostrando que a qualidade do modelo ajustado é apenas regular. Cabe também ressaltar, que esses testes supõem variância constante, o que não se verifica na presença de efeitos de dispersão. Assim, seus resultados devem ser avaliados com cautela.



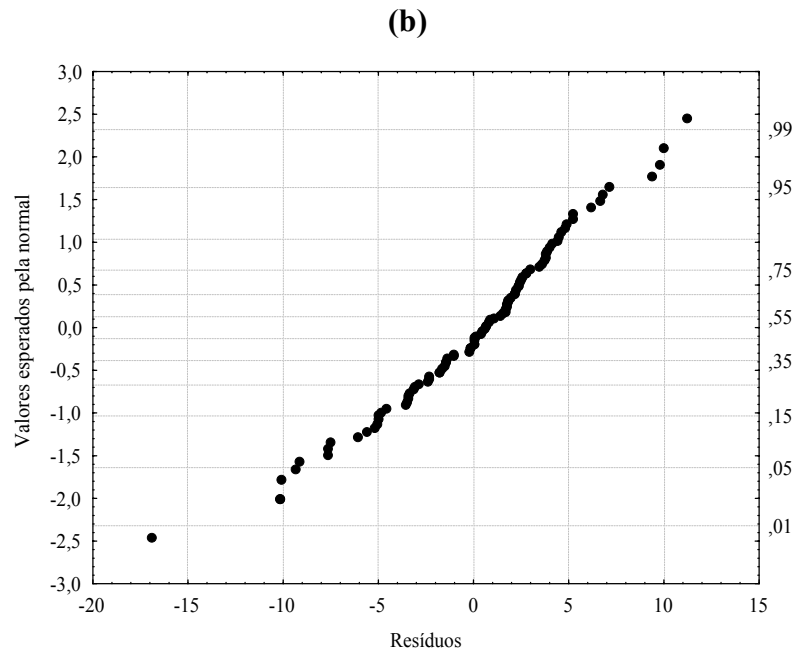
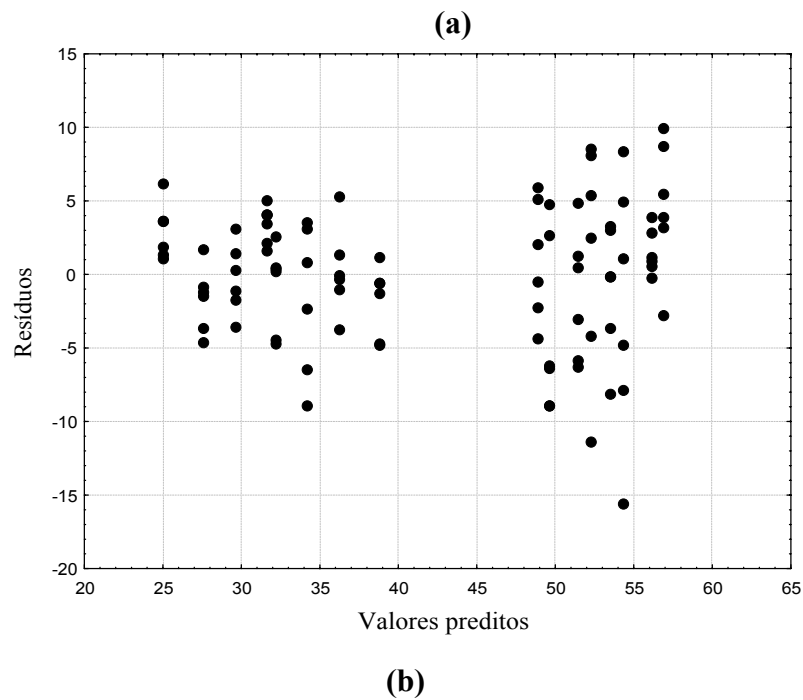


FIGURA 6.2 – Valores preditos e resíduos (a) e gráfico de probabilidade normal (b) dos resíduos do modelo descrito na equação 6.1.

O gráfico de resíduos, relacionando estimativas de resíduos com os correspondentes valores preditos (gráfico (a) da figura 6.3), continua sugerindo uma possível não constância de variância, enquanto o gráfico de probabilidade normal destes resíduos (gráfico (b) desta mesma figura), que estes resíduos seguem uma distribuição normal.



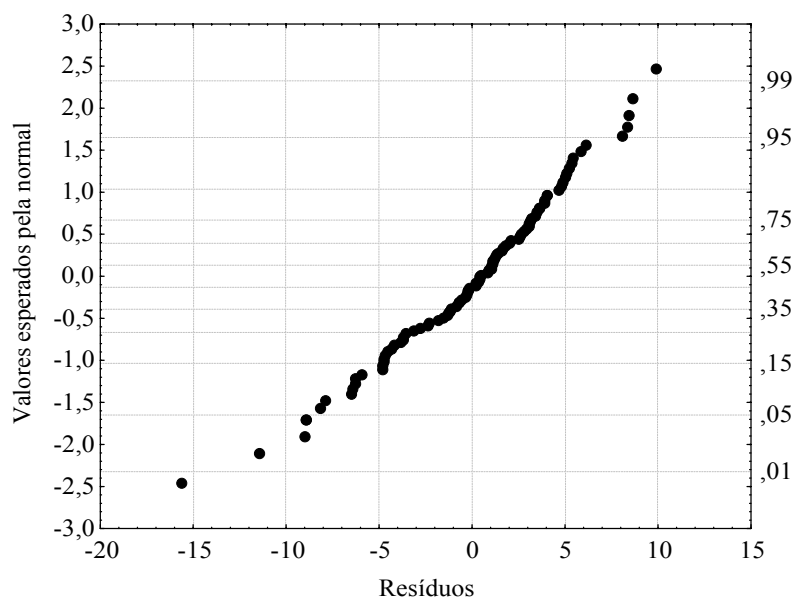


FIGURA 6.3 – Valores preditos e resíduos (a) e gráfico de probabilidade normal (b) dos resíduos do modelo descrito na equação 6.2.

Embora este modelo tenha apresentado resultado satisfatório no teste de significância, o valor-p associado ao teste de falta de ajuste ainda está próximo do nível usual de significância: 5%, razão pela qual não se pode afirmar que a adequação do modelo encontrado é boa. Como a inferência sobre efeitos de dispersão é muito influenciada pela qualidade do modelo de locação ajustado, optou-se por avaliá-los por meio de métodos que se baseiam em variâncias amostrais.

Seguindo a estratégia utilizada, a identificação dos efeitos de dispersão é feita pelo método *R* que os estima a partir de quocientes entre médias aritméticas de variâncias amostrais. Os resultados encontrados, mostrados no gráfico de probabilidade normal da figura 6.4, não sugerem de forma clara a existência de efeitos de dispersão, mas fortes suspeitas sobre os efeitos principais do fator *A* e do fator *E*. Também identifica como suspeitos os efeitos das interações *AB* e *BD*, embora existam efeitos com sinal negativo de magnitude superior.

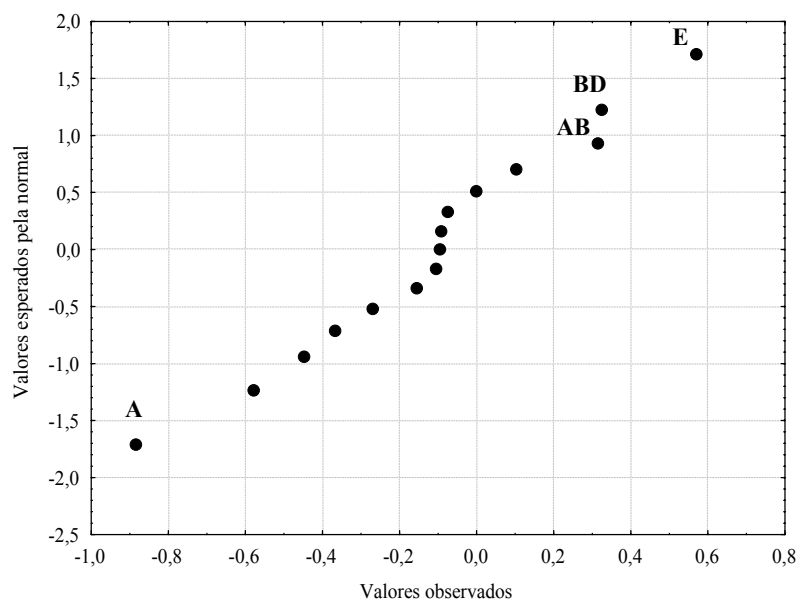


FIGURA 6.4 – Gráfico de probabilidade normal de efeitos de dispersão, construídos a partir de estimativas encontradas pelo método *R*

Como pode haver mais de um efeito de dispersão, a avaliação é refeita por meio do método *S*. O gráfico de probabilidade normal sugere a existência de dois efeitos de dispersão: o efeito principal do fator *A* e o efeito da interação *AB*, embora, se forem considerados os valores absolutos das estimativas dos efeitos, parece haver apenas o efeito principal do fator *A*.

Considerando os resultados encontrados e o princípio da hereditariedade, o conjunto de fatores com efeitos de dispersão é definido como: $\mathcal{D} = \{I, A, B \text{ e } AB\}$. A partir dele é feita a modelagem do logaritmo neperiano das variâncias amostrais das variáveis transformadas (aplicando MQO), resultando no modelo:

$$\hat{\sigma}_y^2 = \exp\{1,95 - 0,58A - 0,05B + 0,46AB\} \quad (6.3)$$

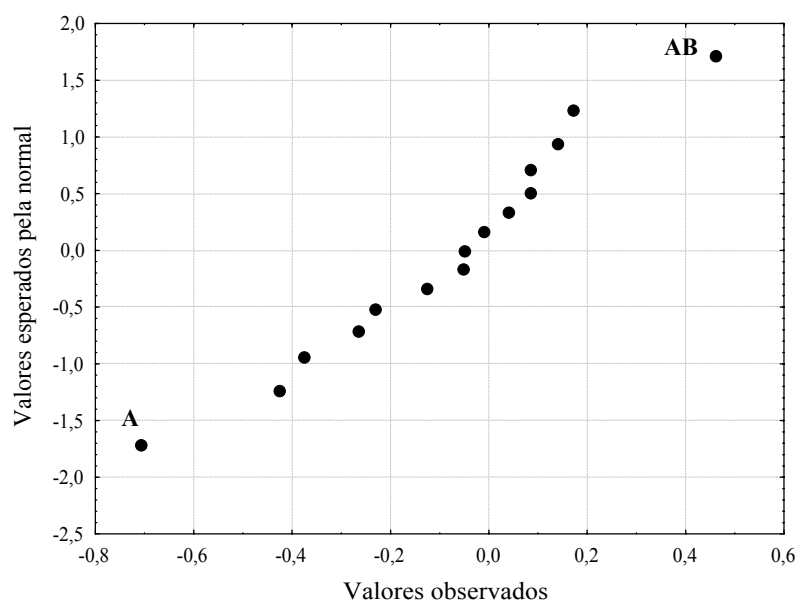


FIGURA 6.5 – Gráfico de probabilidade normal de efeitos de dispersão, construídos a partir de estimativas encontradas pelo método *S*

Esta equação permite prever a variância em cada uma das 32 condições experimentais. Agora, o modelo de locação pode ser refinado pelo método dos mínimos quadrados generalizados (MQG), sendo estimado um novo modelo para a média. Cada observação é ponderada pelo inverso da variância predita pelo modelo 6.3. Os resultados dessa segunda análise definem o seguinte conjunto de fatores com efeitos de locação: $\mathcal{L} = \{I, A, D, E, AE, BC \text{ e } ACD\}$, definindo o modelo:

$$\hat{\mu}_y = 42,422 - 10,488A - 2,330D + 1,842E + 1,456AE + 0,859BC + 1,140ACD \quad (6.4)$$

6.3.5 Discussão e considerações sobre os resultados

A análise permitiu identificar que a resistência à compressão média do concreto está relacionada com o fator água/cimento (Fator *A*), com a forma de adensamento (Fator *D*) e com sua idade (Fator *E*), conforme evidenciam os gráficos das figuras 6.6, 6.7 e 6.8. Maiores valores são esperados aos 28 dias para corpos-de-prova moldados com menor quantidade de água e adensados manualmente.

Os resultados encontrados para a quantidade de água e idade estão dentro do previsto. Apenas o resultado relativo ao adensamento não foi o esperado: o adensamento mecânico, teoricamente, deveria fornecer melhores resultados. Isto talvez se deva ao fato de, no ensaio, o adensamento ser realizado em um corpo-de-prova de tamanho pequeno, mais adequado ao

tamanho do soquete metálico. O tamanho do vibrador de imersão torna a execução do adensamento mais difícil.

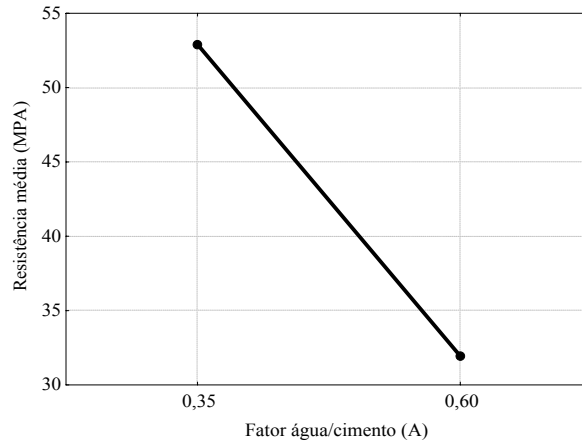


FIGURA 6.6 – Efeito principal do fator *A* sobre a resistência à compressão

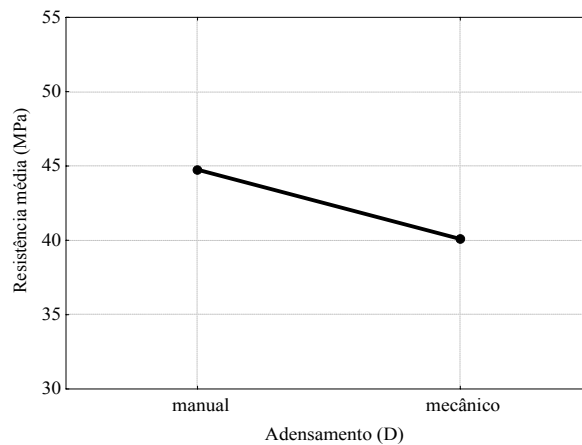


FIGURA 6.7 – Efeito principal do fator *D* sobre a resistência à compressão

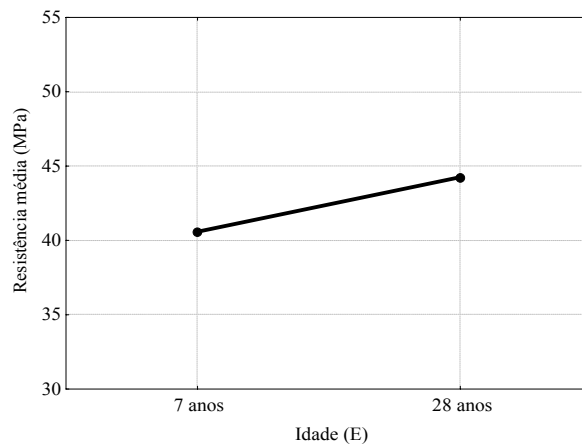


FIGURA 6.8 – Efeito principal do fator *E* sobre a resistência à compressão

Também é constatado que o fator *A* (fator água/cimento) interage com o fator *E* (idade), conforme o gráfico da figura 6.9, evidenciando que, quando é utilizada menor quantidade de água (Fator *A* no nível $-I$), o concreto enrijece mais rapidamente, adquirindo, aos 7 dias, uma resistência quase igual à resistência aos 28 dias, o mesmo não acontecendo ao ser utilizada maior quantidade de água (Fator *A* no nível $+I$). Já o fator *B* (quantidade de adição), foco deste estudo, interage com o fator *C* (tempo de mistura), conforme o gráfico da figura 6.10. Quando a adição não é utilizada (Fator *B* no nível $-I$), menos tempo de midtura fornece maior resistência, o mesmo não acontecendo quando a adição não é utilizada (Fator *B* no nível $+I$).

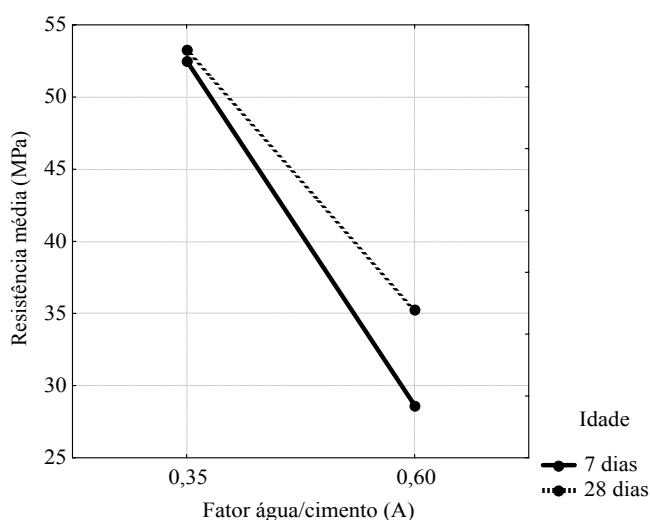


FIGURA 6.9 – Efeito da interação entre o fator *A* e o fator *E* sobre a resistência à compressão

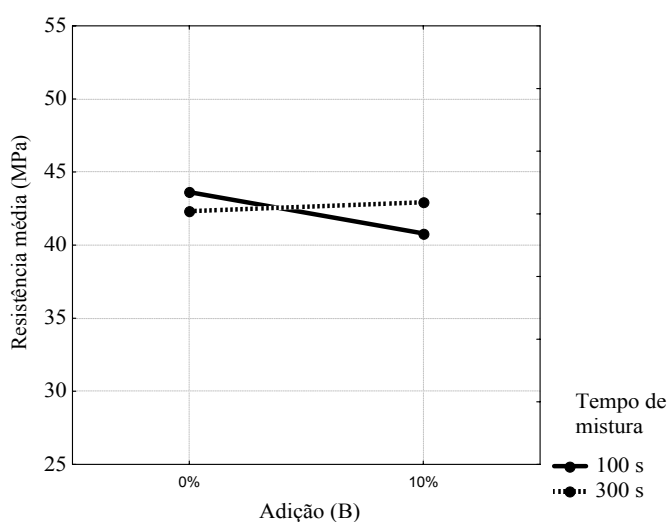


FIGURA 6.10 – Efeito da interação entre o fator *B* e o fator *C* sobre a resistência à compressão

Existe também uma interação tripla entre os fatores *A*, *C* e *D*, que representam respectivamente, o fator água/cimento, o tempo de mistura e o tipo e adensamento. Os resultados, mostrados nos gráficos da figura 6.11, evidenciam que: quando é utilizada menor quantidade de água, a resistência média do concreto fabricado com adensamento mecânico é maior quando o tempo de mistura é menor (100 seg), o mesmo não acontecendo quando é utilizado adensamento manual. Quando o concreto é fabricado com maior quantidade de água, a resistência média é maior quando o tempo de mistura é maior (300 seg), tanto para adensamento mecânico como manual.

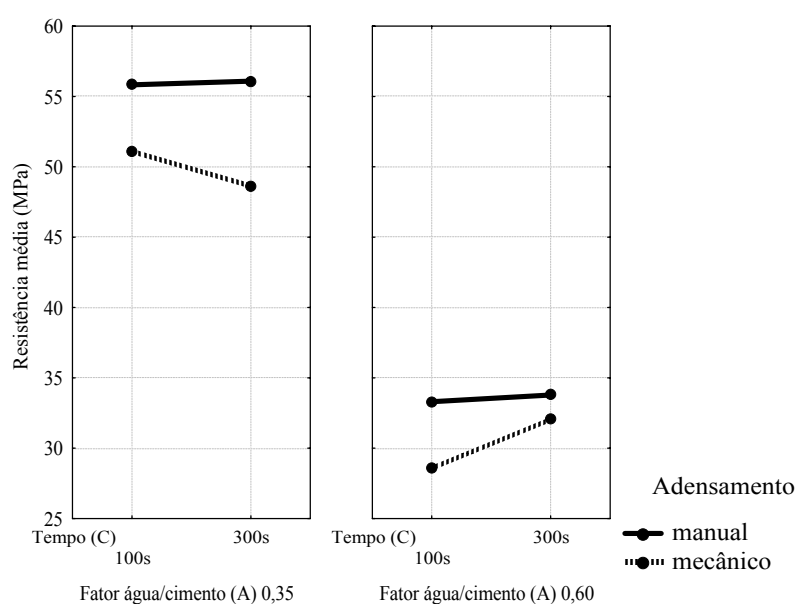


FIGURA 6.11 – Efeito da interação entre os fatores *ACD* sobre a resistência à compressão

A variabilidade da resistência à compressão é influenciada pelos fatores *A* e *B* (fator água/cimento e uso de adição), conforme mostram os gráficos das Figuras 6.12 e 6.13. A variabilidade da resistência à compressão é menor nos corpos-de-prova moldados com fator água/cimento 0,60 (Fator *A* no nível +1) e teor de adição de 10% (Fator *B* no nível +1). Entretanto, estes fatores interagem entre si, conforme evidenciado no gráfico da figura 6.14. Quando é utilizado um fator água/cimento 0,35 (Fator *A* no nível -1), a variabilidade é menor nos corpos-de-prova que moldados com adição (Fator *B* no nível +1), enquanto se utilizado um fator água/cimento 0,60 (Fator *A* no nível +1), a variabilidade é menor nos corpos-de-prova moldados sem adição (Fator *B* no nível -1). Isto talvez ocorra em função dos efeitos pozolânico e microfiller decorrentes do uso da adição na mistura ficarem mais evidentes quando uma menor quantidade de água é utilizada.

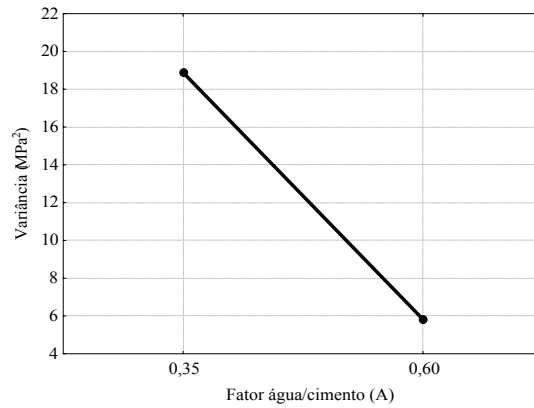


FIGURA 6.12 – Efeito principal do fator *A* sobre a variabilidade da resistência à compressão

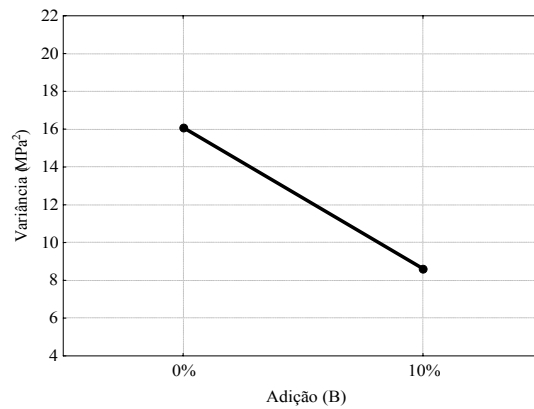


FIGURA 6.13 – Efeito principal do fator *B* sobre a variabilidade da resistência à compressão

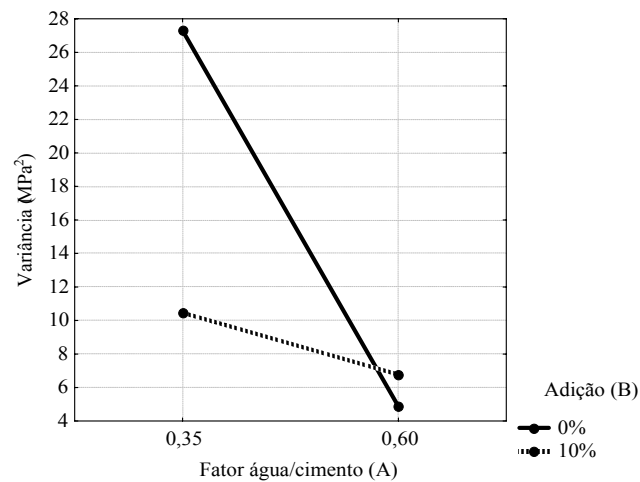


FIGURA 6.14 – Efeito da interação entre os fatores *A* e *B* sobre a variabilidade da resistência à compressão

O principal objetivo da realização do experimento é avaliar o desempenho do concreto moldado com a adição de cinza de casca de arroz, bem como relacioná-lo com alguns

parâmetros de seu processo produtivo. A análise dos resultados mostrou que a hipótese de esta adição aumentar a resistência à compressão do concreto não é comprovada. Também foi possível constatar que, entre os parâmetros controlados no experimento, esta característica interage com o tempo de mistura, interferindo no resultado médio esperado da resistência à compressão, e com o fator água/cimento, interferindo na variabilidade da resistência à compressão.

Quando é utilizada a adição cinza de casca de arroz, o tempo de mistura do concreto deve ser maior para obtenção de uma maior resistência, devendo-se usar menor quantidade de água para obtenção de maior homogeneidade.

6.3.6 Considerações sobre a estratégia utilizada

Este estudo de caso mostra a possibilidade de usar a estratégia elaborada nesta tese para avaliar a existência de efeitos de dispersão em experimentos com poucas replicações.

A metodologia utilizada inicia com a identificação de efeitos de locação pelos métodos clássicos. Usa-se, inicialmente, o teste F para identificar os efeitos de locação e o método de mínimos quadrados ordinários (MQO) para construir o modelo. Após a avaliação da qualidade do modelo encontrado, feita por meio de técnicas formais e técnicas gráficas, escolhe-se avaliar a variabilidade por meio das variâncias amostrais. Usa-se, inicialmente, o método R . Como a análise sugere a possibilidade de mais de um efeito de dispersão, esta é refeita com o método S . Realiza-se, então, o refinamento do modelo de locação por meio do método dos mínimos quadrados generalizados (MQG).

Embora não exista ainda uma teoria consagrada a respeito, a estratégia utilizada pareceu eficiente para a situação de experimentos com poucas replicações.

Capítulo 7

CONCLUSÕES E RECOMENDAÇÕES

Este capítulo apresenta as conclusões do estudo desenvolvido e sugestões para futuras pesquisas na área.

7.1 CONCLUSÕES

A competição industrial é muito grande. Para se manter no mercado, as empresas devem, permanentemente, procurar melhorar seus produtos e serviços, buscando, cada vez mais, adequá-los ao uso e às necessidades dos consumidores. Reduzir custos, eliminando perdas e retrabalho, também é essencial para a empresa tornar-se competitiva.

Quando uma unidade industrial tem, como meta, a qualidade, uma ferramenta de trabalho muito importante é o projeto de experimentos, principalmente em estudos preliminares, nos quais a quantidade de fatores a serem investigados é muito grande. Esta técnica permite avaliar, simultaneamente, o efeito dos vários fatores que podem interferir em um processo produtivo, alterando as características funcionais do produto, que definem a sua qualidade, assim como avaliar os efeitos de suas possíveis interações.

Inicialmente, nesse tipo de estudo, eram investigados apenas os efeitos de locação. Embora isto seja necessário, em geral, não é suficiente. Inegavelmente, Taguchi conseguiu mostrar que reduzir variabilidade tornou-se fundamental para a melhoria da qualidade.

Esta tese contém um estudo sobre o desempenho de métodos para identificar efeitos de dispersão em experimentos do tipo 2^K , com poucas replicações, muitas vezes utilizados na indústria por serem simples, rápidos e econômicos. Considera-se que este tipo de experimento, por não ser tão caro como os experimentos com muitas replicações e por avaliar a variabilidade por meio do erro puro, pode ser muito útil no contexto industrial. Além disso, os resultados deste estudo podem ser estendidos para experimentos do tipo 2^{K-p} , se cuidados forem tomados com os eventuais confundimentos, tanto entre os efeitos de locação como entre os efeitos de dispersão.

Inicialmente, foi feita uma adaptação dos métodos não-iterativos propostos em experimentos não-replicados, para a situação de experimentos com poucas replicações.

Após, mostrou-se que os métodos que fazem estimativas de efeitos de dispersão, neste tipo de experimento, a partir de médias aritméticas, apresentam melhor desempenho, tanto ao serem usadas variâncias amostrais, quanto na utilização de resíduos quadráticos. Entretanto, se existirem vários fatores com efeitos de dispersão, o desempenho dos métodos com emprego de médias geométricas torna-se melhor.

Com relação ao tipo de mensuração da variabilidade, não foi possível identificar a estatística de melhor desempenho: variâncias amostrais ou resíduos quadráticos. Entretanto, parece que existe uma leve superioridade dos métodos que trabalham com resíduos quadráticos em relação aos que usam variâncias amostrais. Contudo, como os métodos que utilizam resíduos quadráticos dependem muito do modelo de locação utilizado, são recomendados apenas quando um teste de falta de ajuste indicar claramente boa adequação do modelo.

A partir da análise dos resultados das simulações, foi definida uma estratégia para identificar efeitos de dispersão em experimentos com poucas replicações. Ajusta-se, primeiramente, um modelo de locação. Se o modelo encontrado for de qualidade ruim, avalia-se a possibilidade de mudança da métrica da resposta. Se o modelo for de boa qualidade, trabalha-se com métodos que usam resíduos quadráticos (método BM ou método H). Se sua qualidade for duvidosa, trabalha-se com métodos que utilizam variâncias amostrais (método R ou método S). Em ambas as situações, primeiramente, adotam-se métodos com uso de médias aritméticas (método BM ou método R) e, se for detectado mais de um efeito de dispersão, refaz-se a análise com os métodos que utilizam médias geométricas (método H ou método S).

Quando se estiver trabalhando com resíduos quadráticos, concordando com a maioria dos autores da área, complementa-se a estratégia propondo um processo iterativo para compensar as eventuais super ou sub-estimações de efeitos em função do viés apresentado por todos eles (BRENNEMAN E NAIR, 2001; NAIR E PREGIBON, 1988), assim como as possíveis distorções ocorridas na estimação dos efeitos de dispersão em função do modelo de locação. Ao trabalhar com variâncias amostrais, recomenda-se refazer o modelo de locação usando mínimos quadrados generalizados (MQG).

A estratégia proposta foi aplicada em um experimento realizado com o objetivo de avaliar a influência de vários parâmetros do processo produtivo sobre sua resposta: a resistência à compressão de corpos-de-prova, moldados com a adição de cinza de casca de arroz. A análise realizada permitiu não apenas conhecer a influência desse tipo de adição sobre a resistência à compressão, mas também identificar parâmetros de seu processo produtivo capazes de afetar seu desempenho, tanto em termos de valor médio como em termos de variabilidade.

Pelas características de tal setor industrial, a construção civil é uma das áreas que muito pode beneficiar-se de estudos dessa natureza, tendo em vista os altos coeficientes de segurança (de até 40%) empregados em vários de seus procedimentos, para tentar anular efeitos de eventuais variabilidades decorrentes de causas não-controladas.

Salienta-se, entretanto, que, devido à generalidade da estratégia proposta, ela pode ser amplamente utilizada em qualquer setor da indústria.

7.2 FUTURAS PESQUISAS

A estratégia proposta nesta tese foi elaborada a partir de alguns estudos de simulação que apresentam limitações e, evidentemente, não esgotam o assunto.

Outros estudos de simulação podem ser desenvolvidos modificando alguns detalhes nas condições experimentais consideradas. Por exemplo: os resultados encontrados seriam os mesmos se, nos experimentos simulados, existissem efeitos de dispersão ativos em interações? O que aconteceria se a intensidade dos efeitos de locação e dispersão fossem maiores? Como se comportariam os métodos analisados, em outros tipos de projetos experimentais?

Barbetta (1999) propõe uma estatística de avaliação do desempenho resultante de uma combinação linear entre variâncias amostrais e resíduos quadráticos. A proposta de combinação de variâncias com resíduos quadráticos baseia-se no fato de, se o modelo de locação for corretamente especificado, o uso de resíduos quadráticos como informação primária para estimar a variância conduz a melhores resultados do que variâncias amostrais, principalmente se forem utilizadas poucas replicações. Como ficariam os métodos analisados se variâncias amostrais e resíduos quadráticos fossem substituídos por essa nova medida?

A análise de efeitos de dispersão em experimentos é uma tarefa extremamente difícil e uma importante área para pesquisa.

Referências Bibliográficas

ARVIDSSON, M. Comments on dispersion effect analysis. (artigo ainda não publicado).

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **Bloco cerâmico para alvenaria: NBR 7171.** Rio de Janeiro, 1993.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **Agregados para concreto: NBR 7211.** Rio de Janeiro, 1982.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **Agregado-determinação da composição granulométrica: NBR 7217.** Rio de Janeiro, 1992.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **Agregado-determinação de impurezas orgânicas húmicas em agregado miúdo: NBR 7220.** Rio de Janeiro, 1987.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **Agregado graúdo-determinação do índice de forma pelo método do paquímetro: NBR 7809.** Rio de Janeiro, 1982.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **Agregado-determinação da massa específica de agregados miúdos por meio do Frasco de Chapman: NBR 9776.** Rio de Janeiro, 1986.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **Agregado-determinação da absorção e da massa específica de agregado graúdo: NBR 9937.** Rio de Janeiro, 1987.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **Agregado em estado compactado e seco-determinação da massa unitária: NBR 7810.** Rio de Janeiro, 1982.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **Cimentos-método da determinação de atividade pozolânica em cimento Portland pozolânico: NBR 5753.** Rio de Janeiro, 1980.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **Moldagem e cura de corpo-de-prova cilíndricos ou prismáticos de concreto-Procedimento: NBR 5738.** Rio de Janeiro, 1994.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **Concreto: Determinação da consistência pelo abatimento do tronco de cone. NBR 7223.** Rio de Janeiro, 1998.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **Ensaio de compressão de corpos-de-prova cilíndricos de concreto: NBR 5739.** Rio de Janeiro, 1974.

BARBETTA, P.A. **Construção de modelos para médias e variâncias na otimização experimental de produtos e processos.** Tese (Engenharia de Produção). Universidade Federal de Santa Catarina, Florianópolis, 1998.

BARBETTA, P.A.; RIBEIRO, J.L.D.; SAMOHYL, R.W. **Uma nova fórmula de modelar variâncias em experimentos com poucas replicações**. In: ENEGEP98. Rio de Janeiro, 1998. **Anais...** Rio de Janeiro, 1997.

BARBETTA, P.A.; RIBEIRO, J.L.D.; BORNIA, A.C. Construção de modelo para a variância em estudos experimentais. **Produção&Produção**, v.3, n.3, p.56-65, 1999.

BARBETTA, P.A.; RIBEIRO, J.L.D.; SAMOHYL, R.W. Variance regression models in experiments with few replications. **Quality Reliability Engineering**, v.16, p.397-404, 2000.

BARBETTA, P.A. **A metodologia de planejamento de experimentos aplicada na melhoria de algoritmos computacionais**. (ainda não publicado)

BARBETTA, P.A.; MATTOS, V.L.D.; SAMOHYL, R.W. **Identificação de efeitos de dispersão em experimentos fatoriais do tipo 2^k** . 15º SINAPE. Águas de Lindóia, 2002. **Anais...** Águas de Lindóia, 2002.

BARBETTA, P.A.; REIS, M.M.; BORNIA, A.C. **Estatística para Cursos de Engenharia e Informática**. São Paulo: Editora Atlas, 2004. 410p.

BARROS NETO, B.; SCARMINIO, I.S.; BRUNS, R.E. **Planejamento e otimização de experimentos**. 2ed., Campinas: Editora Unicamp, 1995.

BARTLETT, M.S.; KENDALL, D.G. The Statistical Analysis of Variance-Heterogeneity and the Logarithmic Transformation. **Journal of the Royal Statistics Society**, Ser. B, 8, p.128-138, 1946.

BAUER, L.A.F.; NORONHA, M.A.A.; BAUER, R.J.F. Uso de aditivos em concretos. In: BAUER, L.A.F. **Materiais de Construção**. 5 ed. Rio de Janeiro: Editora LTC. v.1, Cap6, p.135-185, 2001.

BAUER, L.A.F.; NORONHA, M.A.A. Estudo de dosagem. In: BAUER, L.A.F. **Materiais de Construção**. 5 ed. Rio de Janeiro: Editora LTC. v.1, Cap7, p.186-239, 2001.

BAUER, L.A.F. Preparo, transporte, lançamento, adensamento e cura. In: BAUER, L.A.F. **Materiais de Construção**. 5 ed. Rio de Janeiro: Editora LTC. v.1, Cap8, p.240-266, 2001.

BAUER, L.A.F.; Birindelli, L.O. Ensaio acelerado para previsão da resistência do concreto. In: BAUER, L.A.F. **Materiais de Construção**. 5 ed. Rio de Janeiro: Editora LTC. v.1, Cap 12, p. 347-374, 2001.

BAUER, L.A.F.; BAUER, R.J. Controle tecnológico do concreto. In: BAUER, L.A.F. **Materiais de Construção**. 5 ed. Rio de Janeiro: Editora LTC. v.1, Cap 13, p. 375-404, 2001.

BERGMAN, B.; HYNÉN, A. **Dispersion effects from unreplicated designs in 2^{k-p} series.** *Technometrics*, v.39, n.2, p.191-198, 1997.

BERGMAN, B.; HYNÉN, A. On Dispersion Effects and Their Identification. *Technometrics*, v.39, n.2, p.359-370, 2000.

BOX, G.E.P.; HUNTER, W.G.; HUNTER, J.S. **Statistics for experiments.** USA: John Wiley & Sons, 1978.

BOX, G.E.P.; MEYER, R.D. Dispersion Effects From Fractional Design. *Technometrics*, v.28, n1, p.19-27, 1986a.

BOX, G.E.P.; MEYER, R.D. An analysis for unreplicated fractional factorial. *Technometrics*, v.28, n1, p.11-18, 1986b.

BOX, G. General comments, data analysis, experimental strategy and planning techniques. In: Nair, V.N. Taguchi's Parameter Design: A Panel Discussion. *Technometrics*, v.34, n.2, p.127-161, 1992.

BOX, G. Signal-to-noise ratios, performance criteria and transformations. *Technometrics*, v.30, n.1, p.1-40, 1988.

BRENNEMAN, W.A. **Inference for location and dispersion effects in unreplicated factorial experiments.** Tese (Doctor of Philosophy-Statistics) University of Michigan, 2000.

BRENNEMAN, W.A.; NAIR, V.N. Methods for identifying dispersion effects in unreplicated factorial experiments: a critical analysis and proposed strategies. *Technometrics*, v.43, n.4, p.388-405, 2001.

BUSSAB, W.O.; MORETTIN, P.A. **Estatística básica.** 5 ed. São Paulo: Editora Saraiva, 2002, 526p.

CARROL, R.J.; RUPPERT, D. **Transformation and weighting in regression.** USA: Chapman and Hall, 1988. 245p.

CATEN, C.S. **Método de otimização de produtos e processos produtivos por múltiplas características de qualidade.** Dissertação (Engenharia de Produção). Universidade Federal do Rio Grande do Sul. Porto Alegre, 1995. 154p.

CATEN, C.S.; RIBEIRO, J.L.D. Etapas na otimização experimental de produtos e processos: discussão e estudo de caso. *Produção*, v. 6, n.1, p.45-64, 1996.

CHATERJEE, S.; PRICE, B. **Regression analysis by examples.** USA: John Wiley & Sons, 1977.

CROSSLEY, M.L. **Three paths to better processes**. Disponível em: <www.qualitydigest.com/july00/html/variation.html>. Acessado em 03/12/2001.

DACHS, J.N.W. **Estatística Computacional**. Rio de Janeiro: LTC Ltda, 1988.

FERRER, A.J.; ROMERO, R. A simple method to study dispersion effects from non necessarily replicated data in industrial contexts. **Quality Engineering**, 7(4), p.747-755, 1995.

FIOD NETO, M. **Taguchi e a melhoria da qualidade: uma releitura crítica**. Florianópolis: Editora da UFSC, 1997. 92p.

FREITAS FILHO, P.J. **Introdução à modelagem e simulação de sistemas**. Florianópolis: Visual Books, 2001, 322p.

GRYNA, F. M. (1992). Controle da Qualidade: Métodos Estatísticos Aplicados à Qualidade. In: Juran, J.M. **Controle de Qualidade**. São Paulo: Makron Books, v.6, 1996. 488p.

HARVEY, A. C. Estimating regression models with multiplicative heterocedasticity. **Econometrica**, v. 44, n3, p.461-465, 1976.

HELENE, P.; TERZIAN, P. **Manual de Dosagem e Controle do Concreto**, São Paulo: Editora Pini, 1992. 349 p.

ISAIA, G.C. **Controle de qualidade das estruturas de concreto armado**. Santa Maria: Edições UFSM, 1988. 119p.

ISHIKAWA, K. **Controle de Qualidade Total à Maneira Japonesa**. Rio de Janeiro: Campus, 1993, 221p.

JUDGE, G.G.; HILL, R.C.; GRIFFITHS, W.E.; LUTKEPOHL, H.; LEE, T.C. **Introduction to the theory and practice of econometrics**, 2 ed. Cap. 5 e 6. USA: John Wiley&Sons, p.160-273, ????

KACKAR, R.N. Taguchi's Quality Philosophy: Analysis and Commentary. **Quality Progress**. p.21-29, 1986.

KACKAR, R.N. General Comments, Variation Reduction Through Parameter Design and the Role of Noise Factors, The Role of Interactions, SN Ratios, and Selection of Quality Characteristics, Experimental Strategy and Planning Techniques. In: NAIR, V.N. Taguchi's Parameter Design: A Panel Discussion. **Technometrics**, v.34, n 2, p.127-161, 1992.

LIAO, C.T. Identification of Dispersion Effects From Unreplicated 2 n-p Fractional Factorial Designs, 2000.

LENTH, R.V. Quick and Easy Analysis of Unreplicated Factorial. **Technometrics**, v. 31, n.4, p.469-473, 1989.

LEE, Y.; NELDER, J. A. *Robust design via Generalized Linear Models*. **Journal of Quality Technology**, v. 35, n. 1, p. 2-12, 2003.

LUCAS, J.M. How to achieve a robust process using response surface methodology. **Journal of Quality Technology**, n.26, n.4, p.248-260, 1994.

MARTINS, M.L.C. **Estudo da Produção de Concreto de Alto Desempenho com Adições de Sílica Ativa e Cinza de Casca de Arroz em Betoneiras Basculantes**. Relatório Final. Programa de Apoio aos Pólos Tecnológicos da Secretaria da Ciência e Tecnologia do Estado do Rio Grande do Sul, 1999. 68p.

MARTINS, M.L.C. **Análisis del comportamiento mecánico del hormigón de alta resistencia mediante probetas testigo y probetas moldeadas con distintas condiciones de hormigonado**. Tesis Doctoral. Escuela Tecnica Superior de Ingenieros de Caminos, Canales y Puertos. Departamento de Ingenieria Civil- Construcción. Madrid, 1997.162p.

MATTOS, V.L.D. **Implantação de um controle estatístico de processo em uma olaria de pequeno porte**. Dissertação (Engenharia de Produção). Florianópolis,1997. 101p.

MCGRATH, R.N.; LIN, D.K.J. Confounding of Location and Dispersion Effects in Unreplicated Fractional Factorials. **Journal of Quality Technology**, v.33, n.2, p.129-139, 2001.

MONTGOMERY, D.C. **Design and analysis of experiments**. 4 ed. USA: John Wiley&Sons, 1997. 704p.

MONTGOMERY, D.C. *Introduction to Statistical Quality Control*. 3 ed. USA: John Wiley&Sons, 1997. 676p.

MOTA, E.V.; RIBEIRO, J.L.D. Avaliação dos Sistemas de Garantia de Qualidade em Três Hospitais. **Produção&Produção**, Vol.1, N°1, p.22-29, 1997.

MYERS, R.H.; VINING, G.G. General Comments, Variation Reduction Through Parameter Design and the Role of Noise Factors, Experimental Strategy and Planning Techniques. In: Nair, V.N. Taguchi's Parameter Design: A Panel Discussion. **Technometrics**, v.34, n.2, p.127-161, 1992.

MYERS, R.H.; KHURI, A.I.; VINNING, G. Response surface alternatives to the Taguchi robust parameter design approach. **The American Statistician**, v.46, n.2, p.131-139, 1992.

NAIR, V.N.; PREGIBON, D. Analysing Dispersion Effects From Replicated Factorial Experiments. **Technometrics**, v.30, n.3, p.247-257, 1988.

NAIR, V.N.; PREGIBON, D. Taguchi's Parameter Design: A Panel Discussion. **Technometrics**, v.34, n.2, p.127-161, 1992.

NARULA, S.C.; STANGENHAUS, G. Análise de Regressão L_1 . 8º Sinape. Campinas, 1988.

NETER, J.; KUTNER, M.H.; NACHTSHEIM, C.S.; WASSERMAN, W. Applied Linear Regression Models. 3 ed. USA: Times Mirra Mirror, 1990.

OLIVEIRA, H.M. Propriedades do concreto endurecido. In: Bauer.L.A.F. **Materiais de Construção**. 5 ed. Rio de Janeiro: Editora LTC, 2001. v.1, Cap10, p.284-313.

PAESE, C.; CATEN, C.; RIBEIRO, J.L.D. (2001). Aplicação da Análise da Variância na Implantação do CEP. **Produção**, v.11, n1, p.17-26, 2001.

PALADINI, E.P.P. **Gestão da Qualidade no Processo: a Qualidade na produção de bens e serviços**. São Paulo: Atlas, 1995. 286p.

PAN, G. The Impact of Unidentified Location Effects on Dispersion-Effects Identification From Unreplicated Factorial Design. **Technometrics**, v.41, n.4, p.313-326, 1999.

PHADKE, M.S. General Comments, The Role of Interactions, SN Ratios, and Selection of Quality Characteristics, Data Analysis, Experimental Strategy and Planning Techniques. In: Nair, V.N. Taguchi's Parameter Design: A Panel Discussion. **Technometrics**, v.34, n.2, p.127-161, 1992.

PIANCA, J.B. Concretos. In: PIANCA, J.B. **Manual do construtor**. 11ed. Porto Alegre: Editora Globo, 1977. v.1, Parte1, p.52-67.

SILVA, M.V.D. **Avaliação de Desempenho de uma Plataforma de Comutação Telefônica para Serviços Especializados de Atendimento ao Cliente**. Dissertação (Engenharia de Produção). Florianópolis, 2002.101p.

SOARES, J.F.; SIQUEIRA, A. L. **Introdução à Estatística Médica**. 1.ed. Belo Horizonte : Departamento de Estatística - UFMG,1999.

SOBRAL, E.S. Propriedades do concreto fresco. In: Bauer.L.A.F. **Materiais de Construção**. 5 ed. Rio de Janeiro: Editora LTC, v.1, Cap9, p.267-283, 2001.

SPSS for windows. Versão 10.0.1: Copyright © SPSS Inc., 1999. 1 CD-ROM.

STATISTICA, Versão 6.0: Copyright © Stat soft Inc., 2001. 1 CD ROM.

STEINBERG, D.M.; BURSZTYN,D. Noise factors, dispersion effects and robust design. **Statistica Sinica**, 8, p.67-85, 1998.

TAGUCHI, G.; ELSAYED, E.A.;HSIANG, T. **Engenharia da Qualidade em Sistemas de Produção**. São Paulo: Mc Graw-Hill, 1990. 235p.

TAGUCHI, G. e WU, Y. **Introduction to Off-Line Quality Control**. Nagoya, Japan: Central Japan Quality Control Association, 1980.

TSUI, K.L. An overview of Taguchi method and newly developed statistical methods for robust design. **IIE Transation**, v. 24, n.5, p. 44-57, 1992.

VIEIRA, S. **Estatística experimental**. 2 ed. São Paulo:Editora Atlas, 1999. 185p.

VINNING, G.G.; BOHN, L.L. Responses surfaces for the mean and variance using a nonparametric approach. **Journal of Quality Technology**, v.30. n.3, p.281-291, 1998.

VINNING, G.G.; MYERS, R.H. Combining Taguchi and Response Surface Philosophies: A Dual Response Approach. **Journal of Quality Technology**, v.22. n.1, p.38-45, 1990.

WANG, P.C.; LIN, D.F. Dispersion effects in signal-response data from fractional factorial experiments. **Computacional Statistics and Data Analysis**, v.38, n.1, p.95-111. 2001.

WASILOFF, E.; HARGITT, C. Using DOE to Determine AA Battery Life. **Quality Progress**, p.67-71, 1999.

WILKINS, J.O. Putting Taguchi Methods to Work To Solve Design Flaws. **Quality Progress**, p.55-59, 2000.

WIKLANDER, K. **Analysis of Dispersion Effects in Unreplicated Factorial Designs**. Tese (Departamento de Matemática). Götenborg, 1996.

WOLFINGER, R.D.; TOBIAS,R.D. Joint estimation location, dispersion and randon effects in robust design. **Technometrics**, v.40. n.1, p.62-71, 1998.

WOLLE, C.M.; ALMEIDA, D.F. Ensaio não-destrutivos do concreto. In: Bauer.L.A.F. **Materiais de Construção**. 5ed. Rio de Janeiro: Editora LTC, v.1, Cap 11, p.314-346, 2001.

ZAR, J. H. (1999). **Biostatistical analysis**. 4 ed. New Jersey, Prentice Hall, 1999. 663p.

_____, **The determination and interpretation of dispersion effects**. Disponível em: <<http://mtanri.me.uiuo.edu/doe/ie336/dispersion.htm>>. Acesso em 07/12/2001.

_____, **The basic of experimental design – a quick and non-technical guide**. Disponível em: <www.systma.com/phad530/expdesig.html> . Acesso em 07/12/2001.

_____, **Experimental design**. Disponível em: <www.statsoftinc.com/textbook/stexdes.html>. Acesso em 31/05/2004.

APÊNDICE

TABELA A1 – PIC's obtidas no estudo de simulação realizado para comparar o desempenho dos métodos investigados.

CE	Fatores ⁽¹⁾						Método					
	A	B	C	D	E	F	R	S	H	BM	HM	BH
1	1	1	1	1	1	1	0,6394	0,7498	0,7784	0,7188	0,8034	0,7200
2	1	1	1	1	-1	-1	0,1092	0,0730	0,1372	0,1634	0,1636	0,1834
3	1	1	1	-1	1	-1	0,1866	0,1528	0,3514	0,3516	0,3156	0,2808
4	1	1	1	-1	-1	1	0,4676	0,5136	0,6486	0,5806	0,6082	0,5290
5	1	1	-1	1	1	-1	0,6558	0,5018	0,6428	0,7302	0,6738	0,7878
6	1	1	-1	1	-1	1	0,7808	0,7108	0,7572	0,8052	0,7662	0,8164
7	1	1	-1	-1	1	1	0,9766	0,9540	0,9848	0,9918	0,9846	0,9876
8	1	1	-1	-1	-1	-1	0,3850	0,2870	0,5006	0,5842	0,4218	0,4746
9	1	-1	1	1	1	-1	0,1864	0,1536	0,2526	0,2828	0,2788	0,3042
10	1	-1	1	1	-1	1	0,4748	0,5190	0,5684	0,5318	0,5952	0,5474
11	1	-1	1	-1	1	1	0,6488	0,7500	0,8316	0,7366	0,8310	0,7164
12	1	-1	1	-1	-1	-1	0,1090	0,0760	0,2110	0,2364	0,1630	0,1706
13	1	-1	-1	1	1	1	0,9800	0,1588	0,9726	0,9860	0,9774	0,9912
14	1	-1	-1	1	-1	-1	0,3732	0,2740	0,3658	0,4306	0,4002	0,4782
15	1	-1	-1	-1	1	-1	0,6514	0,4934	0,7934	0,8758	0,7484	0,8012
16	1	-1	-1	-1	-1	1	0,7732	0,7122	0,8050	0,8516	0,7786	0,8262
17	-1	1	1	1	1	-1	0,1880	0,1470	0,3598	0,3680	0,3538	0,3676
18	-1	1	1	1	-1	1	0,4684	0,1276	0,6430	0,5874	0,6448	0,5796
19	-1	1	1	-1	1	1	0,6458	0,1520	0,7928	0,7522	0,8422	0,7436
20	-1	1	1	-1	-1	-1	0,1076	0,0760	0,2044	0,2432	0,2072	0,2136
21	-1	1	-1	1	1	1	0,9822	0,9556	0,9832	0,9934	0,9848	0,9928
22	-1	1	-1	1	-1	-1	0,3696	0,2714	0,4654	0,5556	0,4466	0,5326
23	-1	1	-1	-1	1	-1	0,6560	0,5010	0,7784	0,8638	0,7804	0,8620
24	-1	1	-1	-1	-1	1	0,7730	0,1020	0,8028	0,8478	0,8010	0,8442
25	-1	-1	1	1	1	1	0,6466	0,7528	0,8328	0,7588	0,8370	0,7578
26	-1	-1	1	1	-1	-1	0,1144	0,0762	0,2116	0,2428	0,2078	0,2296
27	-1	-1	1	-1	1	-1	0,1920	0,1470	0,3128	0,3744	0,3424	0,3474
28	-1	-1	1	-1	-1	1	0,4684	0,1132	0,6200	0,5890	0,6522	0,5888
29	-1	-1	-1	1	1	-1	0,6522	0,4902	0,7720	0,8576	0,7570	0,8620
30	-1	-1	-1	1	-1	1	0,7798	0,7128	0,8034	0,8490	0,7936	0,8452
31	-1	-1	-1	-1	1	1	0,9812	0,9568	0,9838	0,9920	0,9852	0,9918
32	-1	-1	-1	-1	-1	-1	0,3714	0,2694	0,4618	0,5524	0,4514	0,5346

(1) Os fatores e seus níveis estão descritos no Quadro 3.1.

TABELA A2 – PIF's obtidas no estudo de simulação realizado para comparar o desempenho dos métodos investigados.

CE	Fatores ⁽¹⁾						Método					
	A	B	C	D	E	F	R	S	H	BM	HM	BH
1	1	1	1	1	1	1	0,3898	0,2706	0,2668	0,4112	0,2642	0,4140
2	1	1	1	1	-1	-1	0,4474	0,5222	0,4726	0,4392	0,4704	0,4106
3	1	1	1	-1	1	-1	0,3984	0,4652	0,3622	0,3846	0,3788	0,3492
4	1	1	1	-1	-1	1	0,3480	0,3320	0,2916	0,3530	0,2986	0,3410
5	1	1	-1	1	1	-1	0,4862	0,5770	0,5926	0,5070	0,5318	0,4576
6	1	1	-1	1	-1	1	0,4942	0,5608	0,5622	0,5074	0,5418	0,4850
7	1	1	-1	-1	1	1	0,4378	0,5252	0,5330	0,4380	0,5356	0,4194
8	1	1	-1	-1	-1	-1	0,5926	0,6474	0,6084	0,5610	0,6272	0,5414
9	1	-1	1	1	1	-1	0,3948	0,4676	0,4022	0,3738	0,4000	0,3684
10	1	-1	1	1	-1	1	0,3564	0,3240	0,3256	0,3532	0,3004	0,3556
11	1	-1	1	-1	1	1	0,3918	0,2534	0,2476	0,4244	0,2604	0,4100
12	1	-1	1	-1	-1	-1	0,4634	0,5158	0,4420	0,4158	0,4622	0,4188
13	1	-1	-1	1	1	1	0,4332	0,5314	0,5484	0,4472	0,5194	0,4352
14	1	-1	-1	1	-1	-1	0,5812	0,6358	0,6326	0,5926	0,5988	0,5420
15	1	-1	-1	-1	1	-1	0,4688	0,5976	0,5542	0,4584	0,5508	0,4362
16	1	-1	-1	-1	-1	1	0,5114	0,5744	0,5588	0,5174	0,5602	0,5038
17	-1	1	1	1	1	-1	0,3980	0,4624	0,3830	0,3722	0,3868	0,3700
18	-1	1	1	1	-1	1	0,3472	0,3306	0,2930	0,3552	0,2958	0,3426
19	-1	1	1	-1	1	1	0,3952	0,2758	0,2714	0,4554	0,2544	0,4182
20	-1	1	1	-1	-1	-1	0,4428	0,5030	0,4394	0,4106	0,4332	0,3834
21	-1	1	-1	1	1	1	0,4192	0,5322	0,5352	0,4352	0,5216	0,4360
22	-1	1	-1	1	-1	-1	0,5850	0,6292	0,6126	0,5478	0,6056	0,5332
23	-1	1	-1	-1	1	-1	0,4806	0,5858	0,5502	0,4626	0,5442	0,4504
24	-1	1	-1	-1	-1	1	0,5178	0,5714	0,5588	0,5122	0,5542	0,5010
25	-1	-1	1	1	1	1	0,3872	0,2566	0,2512	0,4362	0,2662	0,4326
26	-1	-1	1	1	-1	-1	0,4514	0,5228	0,4346	0,3990	0,4466	0,3912
27	-1	-1	1	-1	1	-1	0,3940	0,4578	0,4056	0,3996	0,3718	0,3524
28	-1	-1	1	-1	-1	1	0,3546	0,3150	0,2970	0,3656	0,2934	0,3544
29	-1	-1	-1	1	1	-1	0,4572	0,5764	0,5444	0,4622	0,5470	0,4534
30	-1	-1	-1	1	-1	1	0,5130	0,5642	0,5588	0,5062	0,5562	0,5076
31	-1	-1	-1	-1	1	1	0,4404	0,5360	0,5308	0,4430	0,5372	0,4364
32	-1	-1	-1	-1	-1	-1	0,5878	0,6304	0,5974	0,5464	0,6226	0,5330

(1) Os fatores e seus níveis estão descritos no Quadro 3.1.

TABELA A3 – PIC's obtidas no estudo de simulação realizado para avaliar o desempenho dos métodos H e HM quando são utilizados resíduos quadráticos individuais ou médios

	Fatores ⁽¹⁾						Método			
	A	B	C	D	E	F	H (ind)	H (méd)	HM (ind)	HM (med)
1	1	1	1	1	1	1	0,816	0,506	0,739	0,447
2	1	1	1	1	-1	-1	0,177	0,111	0,158	0,101
3	1	1	1	-1	1	-1	0,300	0,193	0,325	0,200
4	1	1	1	-1	-1	1	0,597	0,304	0,624	0,318
5	1	1	-1	1	1	-1	0,693	0,587	0,645	0,516
6	1	1	-1	1	-1	1	0,764	0,571	0,745	0,527
7	1	1	-1	-1	1	1	0,981	0,893	0,982	0,900
8	1	1	-1	-1	-1	-1	0,421	0,354	0,510	0,380
9	1	-1	1	1	1	-1	0,245	0,165	0,220	0,117
10	1	-1	1	1	-1	1	0,581	0,313	0,560	0,280
11	1	-1	1	-1	1	1	0,837	0,557	0,838	0,523
12	1	-1	1	-1	-1	-1	0,159	0,098	0,206	0,130
13	1	-1	-1	1	1	1	0,973	0,896	0,978	0,828
14	1	-1	-1	1	-1	-1	0,391	0,329	0,375	0,306
15	1	-1	-1	-1	1	-1	0,737	0,613	0,787	0,656
16	1	-1	-1	-1	-1	1	0,787	0,586	0,816	0,619
17	-1	1	1	1	1	-1	0,352	0,218	0,341	0,194
18	-1	1	1	1	-1	1	0,647	0,338	0,633	0,325
19	-1	1	1	-1	1	1	0,819	0,527	0,766	0,463
20	-1	1	1	-1	-1	-1	0,209	0,134	0,200	0,113
21	-1	1	-1	1	1	1	0,986	0,894	0,983	0,901
22	-1	1	-1	1	-1	-1	0,467	0,349	0,487	0,390
23	-1	1	-1	-1	1	-1	0,773	0,667	0,777	0,648
24	-1	1	-1	-1	-1	1	0,795	0,581	0,807	0,573
25	-1	-1	1	1	1	1	0,843	0,559	0,837	0,527
26	-1	-1	1	1	-1	-1	0,204	0,124	0,199	0,114
27	-1	-1	1	-1	1	-1	0,345	0,180	0,305	0,176
28	-1	-1	1	-1	-1	1	0,636	0,338	0,604	0,306
29	-1	-1	-1	1	1	-1	0,773	0,636	0,798	0,660
30	-1	-1	-1	1	-1	1	0,811	0,597	0,813	0,603
31	-1	-1	-1	-1	1	1	0,987	0,905	0,981	0,884
32	-1	-1	-1	-1	-1	-1	0,412	0,331	0,443	0,342

(1) Os fatores e seus níveis estão descritos no Quadro 5.1.

TABELA A4 – PIF's obtidas no estudo de simulação realizado para avaliar o desempenho dos métodos H e HM quando são utilizados resíduos quadráticos individuais ou médios

	Fatores ⁽¹⁾						Método			
	A	B	C	D	E	F	H (ind)	H (méd)	HM (ind)	HM (med)
1	1	1	1	1	1	1	0,265	0,329	0,271	0,351
2	1	1	1	1	-1	-1	0,474	0,504	0,454	0,485
3	1	1	1	-1	1	-1	0,404	0,470	0,381	0,415
4	1	1	1	-1	-1	1	0,296	0,431	0,322	0,402
5	1	1	-1	1	1	-1	0,530	0,576	0,580	0,588
6	1	1	-1	1	-1	1	0,548	0,575	0,540	0,589
7	1	1	-1	-1	1	1	0,548	0,575	0,552	0,548
8	1	1	-1	-1	-1	-1	0,623	0,660	0,609	0,639
9	1	-1	1	1	1	-1	0,420	0,502	0,431	0,511
10	1	-1	1	1	-1	1	0,317	0,384	0,317	0,394
11	1	-1	1	-1	1	1	0,254	0,304	0,266	0,321
12	1	-1	1	-1	-1	-1	0,434	0,498	0,442	0,504
13	1	-1	-1	1	1	1	0,492	0,556	0,513	0,562
14	1	-1	-1	1	-1	-1	0,599	0,628	0,642	0,658
15	1	-1	-1	-1	1	-1	0,556	0,600	0,552	0,551
16	1	-1	-1	-1	-1	1	0,561	0,621	0,552	0,595
17	-1	1	1	1	1	-1	0,380	0,415	0,385	0,453
18	-1	1	1	1	-1	1	0,279	0,385	0,272	0,404
19	-1	1	1	-1	1	1	0,257	0,311	0,286	0,344
20	-1	1	1	-1	-1	-1	0,462	0,496	0,489	0,510
21	-1	1	-1	1	1	1	0,534	0,543	0,529	0,547
22	-1	1	-1	1	-1	-1	0,620	0,644	0,638	0,633
23	-1	1	-1	-1	1	-1	0,564	0,567	0,582	0,551
24	-1	1	-1	-1	-1	1	0,551	0,602	0,552	0,599
25	-1	-1	1	1	1	1	0,269	0,338	0,246	0,341
26	-1	-1	1	1	-1	-1	0,453	0,505	0,456	0,492
27	-1	-1	1	-1	1	-1	0,366	0,443	0,392	0,458
28	-1	-1	1	-1	-1	1	0,288	0,408	0,317	0,393
29	-1	-1	-1	1	1	-1	0,566	0,581	0,565	0,598
30	-1	-1	-1	1	-1	1	0,557	0,578	0,570	0,589
31	-1	-1	-1	-1	1	1	0,545	0,514	0,541	0,545
32	-1	-1	-1	-1	-1	-1	0,635	0,676	0,594	0,634

(1) Os fatores e seus níveis estão descritos no Quadro 5.1.

TABELA A5 – PIC's obtidas no estudo de simulação que compara métodos que usam variâncias amostrais com métodos que usam resíduos quadráticos

	Fatores ⁽¹⁾					Método
	C	F	R	S	H	BM
1	1	1	0,3512	0,3664	0,3444	0,3412
2	1	-1	0,0934	0,0954	0,0942	0,0978
3	-1	1	0,3940	0,3264	0,3616	0,4084
4	-1	-1	0,3834	0,3194	0,3572	0,4028

(1) Os fatores e seus níveis estão descritos no Quadro 5.2.

TABELA A6 – PIF's obtidas no estudo de simulação que compara métodos que usam variâncias amostrais com métodos que usam resíduos quadráticos

	Fatores ⁽¹⁾					Método
	C	F	R	S	H	BM
1	1	1	0,1800	0,2246	0,2486	0,1986
2	1	-1	0,4360	0,4924	0,5150	0,4794
3	-1	1	0,6036	0,6544	0,6614	0,6150
4	-1	-1	0,5952	0,6610	0,6518	0,6266

(1) Os fatores e seus níveis estão descritos no Quadro 5.2.

TABELA A7 – Resistência à compressão (MPa) de corpos-de-prova de concreto.

CE	Fatores ⁽¹⁾					Y ₁	Y ₂	Y ₃
	A	B	C	D	E			
1	-1	-1	-1	-1	-1	57,30	55,90	59,84
2	-1	-1	-1	-1	1	65,57	60,10	54,11
3	-1	-1	-1	1	-1	48,38	50,93	54,75
4	-1	-1	-1	1	1	52,33	54,37	40,74
5	-1	-1	1	-1	-1	53,35	49,85	45,45
6	-1	-1	1	-1	1	62,71	38,71	59,21
7	-1	-1	1	1	-1	45,20	45,58	56,34
8	-1	-1	1	1	1	60,73	48,06	54,75
9	-1	1	-1	-1	-1	56,79	56,53	53,35
10	-1	1	-1	-1	1	49,53	55,39	46,47
11	-1	1	-1	1	-1	52,71	48,38	51,95
12	-1	1	-1	1	1	57,61	40,87	60,35
13	-1	1	1	-1	-1	60,03	57,04	56,66
14	-1	1	1	-1	1	66,84	60,81	62,39
15	-1	1	1	1	-1	46,60	44,56	53,99
16	-1	1	1	1	1	40,74	43,42	43,29
17	1	-1	-1	-1	-1	27,50	32,47	34,82
18	1	-1	-1	-1	1	38,20	38,20	37,56
19	1	-1	-1	1	-1	28,65	26,36	26,10
20	1	-1	-1	1	1	33,23	35,01	35,65
21	1	-1	1	-1	-1	32,78	27,88	29,92
22	1	-1	1	-1	1	35,90	41,51	36,16
23	1	-1	1	1	-1	26,10	29,28	26,35
24	1	-1	1	1	1	35,01	37,30	37,69
25	1	1	-1	-1	-1	28,52	31,07	26,10
26	1	1	-1	-1	1	37,56	32,47	35,20
27	1	1	-1	1	-1	23,94	22,92	26,74
28	1	1	-1	1	1	27,69	31,83	25,21
29	1	1	1	-1	-1	27,76	32,72	32,59
30	1	1	1	-1	1	34,12	34,06	39,98
31	1	1	1	1	-1	31,19	28,65	26,86
32	1	1	1	1	1	36,61	35,65	33,74

(1) Os fatores e seus níveis estão descritos na tabela 6.3.