



**Antonio Fernando de Castro Vieira**

**Análise da Média e Dispersão em  
Experimentos Fatoriais não Replicados para  
Otimização de Processos Industriais**

**Tese de Doutorado**

Tese apresentada como requisito parcial para obtenção do grau de Doutor pelo Programa de Pós-Graduação em Engenharia de Produção do Departamento de Engenharia Industrial da PUC-Rio.

Orientador: Prof. Eugênio Kahn Epprecht

Rio de Janeiro  
Setembro de 2004



Antonio Fernando de Castro Vieira

## **Análise da Média e Dispersão em Experimentos Fatoriais não Replicados para Otimização de Processos Industriais**

Tese apresentada como requisito parcial para obtenção do grau de Doutor pelo Programa de Pós-Graduação em Engenharia de Produção do Departamento de Engenharia Industrial do Centro Técnico Científico da PUC-Rio. Aprovada pela Comissão Examinadora abaixo assinada.

**Prof. Eugênio Kahn Epprecht**

Orientador

Departamento de Engenharia Industrial – PUC-Rio

**Prof. Sueli Aparecida Mingoti**

Departamento de Estatística – UFMG

**Prof. Flávio Sanson Fogliatto**

Departamento de Engenharia Produção e Transportes – UFRGS

**Prof. Antônio Carlos Monteiro Ponce de Leon**

Instituto de Medicina Social – UERJ

**Prof. Alvaro de Lima Veiga Filho**

Departamento de Engenharia Elétrica – PUC-Rio

**Prof. José Eugenio Leal**

Coordenador(a) Setorial do Centro Técnico Científico – PUC-Rio

Rio de Janeiro, 13 de setembro de 2004

Todos os direitos reservados. É proibida a reprodução total ou parcial do trabalho sem autorização da universidade, do autor e do orientador.

### **Antonio Fernando de Castro Vieira**

Graduou-se em Engenharia Mecânica em 1968, pela Universidade Federal de Pernambuco, e concluiu o Mestrado em Engenharia Industrial na PUC-RIO em 1970. É professor do Departamento de Engenharia Industrial na PUC-RIO desde 1970. Exerceu atividades de análise de pesquisa operacional e operações de produção em diversas empresas.

Ficha catalográfica

Vieira, Antonio Fernando de Castro

Análise da média e dispersão em experimentos fatoriais não replicados para otimização de processos industriais / Antonio Fernando de Castro Vieira ; orientador: Eugênio Kahn Epprecht. – Rio de Janeiro : PUC-Rio, Departamento de Engenharia Industrial, 2004.

204 f. : il. ; 30 cm

Tese (doutorado) – Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Engenharia Industrial.

Inclui referências bibliográficas

1. Engenharia industrial – Teses. 2. Qualidade. 3. Otimização de processos. 4. Modelos lineares generalizados. I. Epprecht, Eugênio Kahn. II. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Engenharia Industrial. III. Título.

CDD: 658.5

## Resumo

Vieira, Antonio Fernando de Castro **Análise da Média e Dispersão em Experimentos Fatoriais não Replicados para Otimização de Processos Industriais**. Rio de Janeiro, 2004. 204 p. Tese de Doutorado - Departamento de Engenharia Industrial, Pontifícia Universidade Católica do Rio de Janeiro.

Esta tese reúne as técnicas estatísticas indicadas para a modelagem da média e da dispersão das características de qualidade de processos e produtos, em experimentos fatoriais não replicados, resultando na definição de um roteiro integrado e detalhado de análise. A motivação vem de que, apesar de haver várias publicações sobre regressão linear clássica, modelos lineares generalizados (MLG), transformação da resposta e planejamento de experimentos, não existe um texto que reúna e descreva em detalhe todos os aspectos da modelagem da média e da dispersão em experimentos fatoriais. Os poucos textos sobre esse assunto não descrevem vários aspectos importantes em estudos dessa natureza, por exemplo, como são aplicados os testes de significância dos coeficientes dos MLG, e quais são as estatísticas e os gráficos indicados para verificar a adequação do modelo. Ademais, nada foi encontrado na literatura sobre a identificação de modelos em experimentos fatoriais. Todos esses aspectos são detalhados nessa tese. Uma vez construído o modelo, é mostrado como usá-lo para obter as condições ótimas de operação dos processos e produtos. Além do cumprimento desse objetivo principal, a tese traz algumas contribuições adicionais; a saber: a) aponta limitações em todos quatro métodos da literatura que se propõem a escolher a transformação mais adequada para a resposta. Esses métodos não produziram resultados satisfatórios quando houve interações significativas entre os fatores; b) propõe a utilização de métodos de transformação da resposta como fonte de indicação da função de ligação a ser usada nos modelos lineares generalizados; e c) propõe a utilização da função de log-verossimilhança para uma escolha conjunta da distribuição de probabilidade e da função de ligação, nos modelos lineares generalizados.

## Palavras-chave

Qualidade; Otimização de Processos; Modelos Lineares Generalizados.

## **Abstract**

Vieira, Antonio Fernando de Castro **Análisis of Média e Dispersion in Unreplicated Factorial Experiments for the Optimization of Industrial Processes.** Rio de Janeiro, 2004. 204 p. DSc Thesis - Department of Industrial Engineering, Pontifícia Universidade Católica do Rio de Janeiro.

This thesis puts together the statistical techniques indicated for modelling the mean and dispersion of quality characteristics of products and processes via unreplicated factorial experiments, resulting in the definition of an integrated and detailed script for the analysis. It was motivated by the fact that, although there are many publications about classic linear regression, generalized linear models (GLMs), response transformation and design of experiments, there is no one text which put together and describe in detail all the aspects of the modelling of the mean and the dispersion in factorial experiments. The few texts on the subject do not describe a number of important aspects in studies of this nature, e.g. how significance tests for the coefficients in GLMs should be applied and which are the statistics and plots indicated for checking model adequacy. In addition, nothing was found in the literature about model identification in factorial experiments. All these aspects are detailed in this thesis. Once the model is built, we show how to use it in order to obtain the optimal operating conditions for products and processes. Besides achieving this main objective, the thesis brings some additional contributions, namely: a) it points out limitations in all the four methods in the literature which have the purpose of selecting the most adequate transformation of the response; b) it proposes using response transformation methods as a source of indication of the link function to use in GLMs, and c) it proposes using the log-likelihood function for the joint choice of the probability distribution and of the link function in GLMs.

## **Keywords**

Quality; Processes Optimization, Generalized Linear Models.

# Sumário

1	Introdução	15
2	Regressão Linear	26
2.1	Estimação dos Parâmetros com Mínimos Quadrados	27
2.2	Testes de Hipótese em Regressão	30
2.3	Verificação da Adequação do Modelo	35
2.3.1	Análise dos Resíduos	35
2.3.1.1	Verificação da Suposição de Normalidade	35
2.3.1.2	Verificação da Suposição de Independência	41
2.3.1.3	Verificação da Suposição de Variância Constante	42
2.3.1.4	Verificação de Observações Atípicas	44
2.3.1.5	Verificação de Observações Influentes	47
2.4	Otimização do Processo	51
2.4.1	Ponto Estacionário	51
2.4.2	Ponto de Rendimento Máximo	53
2.5	Intervalos para a Média da Resposta e para a Previsão da Resposta	54
2.5	Mínimos Quadrados Ponderados	59
2.6	<i>Forward Search</i>	60
3	Modelos Lineares Generalizados	66
3.1	Distribuições de Probabilidade	66
3.2	Estrutura dos MLG	69
3.3	Estimação dos Parâmetros	71
3.3.1	Algoritmo para Estimar os Parâmetros	72
3.3.1.1	Ciclo Interativo	75
3.4	Teste de Significância dos Coeficientes	76
3.4.1	<i>Deviance</i>	77
3.4.2	Estimação do Parâmetro de Dispersão	81
3.5	Adequação do Modelo	82
3.5.1	Resíduo <i>Deviance</i>	82

3.5.2 Resíduo <i>Deviance Studentizado</i>	83
3.5.3 Verificação da Função de Ligação	84
3.5.4 Verificação da Função de Variância	84
3.5.5 Distância de Cook	85
3.5.6 <i>Forward Search</i>	86
3.6 Quase-Verossimilhança	87
3.6.1 Estimação dos Coeficientes	89
3.6.2 Estimação do Parâmetro de Dispersão	90
3.6.3 Significância dos Coeficientes	90
3.6.4 Adequação do Modelo	90
3.7 Quase-Verossimilhança Estendida	92
3.7.1 Estimação dos Coeficientes	93
3.7.2 Estimação do Parâmetro de Dispersão	94
3.7.3 Significância dos Coeficientes	94
3.7.4 Adequação do Modelo	94
3.8 Quase-Verossimilhança Restrita	94
4 Análise da Média	95
4.1 Variáveis de Resposta	96
4.2 Estágio 1: Modelo Linear e Mínimos Quadrados (MQ)	99
4.3 Estágio 2: Transformação da Resposta	119
4.3.1 Definição da Transformação da Resposta	120
4.3.2 Estimação e Teste de Significância dos Coeficientes	129
4.3.3 Testes de Adequação do Modelo	130
4.4 Estágio 3: Modelos Lineares Generalizados	147
4.4.1 Etapa 1- Distribuição de Probabilidade e Função de Ligação	150
4.4.2 Etapa 2-Teste de Significância dos Coeficientes	157
4.4.3 Etapa 3-Testes de Adequação do Modelo	158
5 Análise Conjunta da Média da Dispersão	170
5.1 A Relação do Modelo da Média com os Efeitos na Dispersão	171
5.2 Método Gráfico	172
5.3 Método de Box e Meyer	175

5.4 Métodos Iterativos – Modelagem Conjunta da Média e Variância	180
5.4.1 Método de Lee e Nelder	180
5.4.1.1 Modelo para a Média	181
5.4.1.2 Modelo para a Dispersão	183
5.4.1.3 Procedimento Iterativo	184
5.5 Análise da Média e da Dispersão e os Fatores de Ruído	190
5.5.1 Fatores de Ruído	190
5.5.2 Fatores de Dispersão e Fatores de Ruído	194
5.5.3 Procedimento Geral - Fatores de Dispersão e Fatores de Ruído	195
5.6 Roteiro para Analisar a Média e a Variância	196
5.6.1 Análise da Média	196
5.6.2 Análise Conjunta da Média e da Dispersão	197
6 Conclusões	198
Referências Bibliográficas	201



## Lista de figuras

<b>Figura 2.1</b> - Gráfico de Probabilidade Normal dos Resíduos	37
<b>Figura 2.2</b> - Gráfico de Probabilidade Normal com Envelope	38
<b>Figura 2.3</b> - Gráfico dos Resíduos <i>Studentizados</i>	41
<b>Figura 2.4</b> - Gráfico Valor Absoluto dos Resíduos <i>Versus</i> Valor Ajustado	44
<b>Figura 2.5</b> - Gráfico do Resíduo <i>Outlier-t Versus</i> Valor Ajustado	46
<b>Figura 2.6</b> - Gráfico de Probabilidade Normal do Resíduo <i>Outlier-t</i>	47
<b>Figura 2.7</b> - Gráficos de $y_1$ <i>Versus</i> $x_1$ e $y_2$ <i>Versus</i> $x_2$	48
<b>Figura 2.8</b> - Gráfico da Distância de Cook	50
<b>Figura 2.9</b> - <i>Forward Search</i> de $R^2$ e da Estimativa de $\sigma^2$	64
<b>Figura 2.10</b> - <i>Forward Search</i> dos Resíduos Padronizados	65
<b>Figura 2.11</b> - <i>FS</i> das Estimativas dos Coeficientes e da Estatística de Teste $t$	65
<b>Figura 3.1</b> - Dependências Funcionais	74
<b>Figura 3.2</b> - Algoritmo MQPI	74
<b>Figura 3.3</b> - Gráfico de Probabilidade Normal com Envelope	83
<b>Figura 3.4</b> - Gráfico do Resíduo <i>Studentizado Versus</i> Valor Ajustado	84
<b>Figura 3.5</b> - Valor Absoluto do Resíduo <i>Studentizado Versus</i> Valor Ajustado	85
<b>Figura 3.6</b> - Gráfico da Distância de Cook	86
<b>Figura 3.7</b> - Gráfico da <i>FS</i> para a <i>Deviance</i> e $\hat{\phi}$	86
<b>Figura 3.8</b> - Gráfico da <i>FS</i> para o Resíduo <i>Deviance</i>	87
<b>Figura 3.9</b> - Gráfico dos Resíduos <i>Deviance Studentizados versus</i> Valores Ajustados (à esquerda) e Gráfico Valor Absoluto do Resíduo de <i>Deviance Studentizados Versus</i> Valor Ajustado (à direita)	92
<b>Figura 4.1</b> - Gráfico dos Efeitos - Modelo (M1)	102
<b>Figura 4.2</b> - Gráfico dos Efeitos - Modelo (M2)	102
<b>Figura 4.3</b> – Gráficos de Probabilidade Normal dos Resíduos. Modelos com o Fator C (esquerda) e sem o Fator C (direita)	104
<b>Figura 4.4</b> – Gráfico dos Resíduos <i>versus</i> Valores Ajustados	104
<b>Figura 4.5</b> – Distância de Cook	105
<b>Figura 4.6</b> – Gráfico dos Resíduos <i>Outlier-t</i>	105

<b>Figura 4.7</b> – Gráfico Valor Absoluto do Resíduo <i>Versus</i> Valor Ajustado	106
<b>Figura 4.8</b> - Gráfico dos Efeitos – Modelo Linear	107
<b>Figura 4.9</b> - Gráfico de Probabilidade dos Resíduos	110
<b>Figura 4.10</b> - Gráficos dos Resíduos <i>versus</i> Valor Ajustado	110
<b>Figura 4.11</b> – Distância de Cook	111
<b>Figura 4.12</b> – <i>Forward Search</i> de $\hat{\sigma}^2$ e de $R^2$	111
<b>Figura 4.13</b> – <i>Forward Search</i> dos Resíduos Padronizados	112
<b>Figura 4.14</b> – Gráfico de Probabilidade Normal dos Efeitos	114
<b>Figura 4.15</b> – Gráfico de Probabilidade Normal dos Resíduos com Envelope	115
<b>Figura 4.16</b> – Gráfico dos Resíduos <i>versus</i> Valores Ajustados	115
<b>Figura 4.17</b> – Gráfico Valor Absoluto do Resíduo <i>versus</i> Valor Ajustado	116
<b>Figura 4.18</b> – Gráfico dos Resíduos <i>Outlier-t</i>	116
<b>Figura 4.19</b> – Gráfico da Distância de Cook	117
<b>Figura 4.20</b> – Dados do Exemplo 4.2: <i>Forward Search</i> de $\hat{\sigma}^2$ e de $R^2$	118
<b>Figura 4.21</b> – <i>Forward Search</i> dos Resíduos Padronizados	118
<b>Figura 4.22</b> – Gráfico da Transformação de Box–Cox	122
<b>Figura 4.23</b> – Gráfico Inverso para Valores de $\lambda$	124
<b>Figura 4.24</b> - Gráficos de Valor Ajustado <i>versus</i> Valor Ajustado	125
<b>Figura 4.25</b> - Gráficos dos Valores Transformados $y(\lambda)$ <i>versus</i> $\hat{y}(1)$	126
<b>Figura 4.26</b> – <i>Forward Search</i> - Gráfico da Estatística de Teste	128
<b>Figura 4.27</b> – Gráfico de Probabilidade Normal dos Efeitos	129
<b>Figura 4.28</b> – Gráfico de Probabilidade Normal com Envelope	130
<b>Figura 4.29</b> – Gráfico dos Resíduos <i>Versus</i> Valores Ajustados	131
<b>Figura 4.30</b> - Gráfico dos Valores Absolutos dos Resíduos	131
<b>Figura 4.31</b> – Gráfico dos Resíduos <i>Outlier-t</i>	132
<b>Figura 4.32</b> – Gráfico da Distância de Cook	132
<b>Figura 4.33</b> – Dados do Exemplo 4.2: <i>Forward Search</i> de $\hat{\sigma}^2$ e de $R^2$	133
<b>Figura 4.34</b> – <i>Forward Search</i> dos Resíduos Padronizados	134
<b>Figura 4.35</b> – Transformação de Box-Cox	136
<b>Figura 4.36</b> – Transformação de Box-Cox - com as Interações Reais	136
<b>Figura 4.37</b> – Gráficos Inversos para Valores de $\lambda$	137
<b>Figura 4.38</b> – Gráficos Inversos para $\lambda = 1$ e $\lambda = -1$	138
<b>Figura 4.39</b> - Gráficos de Valor Ajustado <i>Versus</i> Valor Ajustado	138

<b>Figura 4.40</b> - Gráficos dos Valores Transformados $y(j)$ Versus $\hat{y}(1)$	139
<b>Figura 4.41</b> – <i>Forward Search</i> da Estatística de Teste	140
<b>Figura 4.42</b> – <i>Forward Search</i> da Estatística de Teste com Interações Reais	141
<b>Figura 4.43</b> – Gráfico da Transformação de Box–Cox	143
<b>Figura 4.43</b> – Gráficos Inversos para valores de $\lambda$	144
<b>Figura 4.44</b> - Gráficos de Valor Ajustado Versus Valor Ajustado	145
<b>Figura 4.45</b> - Gráficos dos Valores Transformados $Y(\lambda)$ Versus $\hat{Y}(1)$	145
<b>Figura 4.46</b> – <i>Forwqrd Search</i> da Estatística de Teste	146
<b>Figura 4.47</b> – Gráficos de Probabilidade Normal dos Resíduos <i>Deviance</i> . À Esquerda: Distribuição Gama. À Direita: Distribuição Normal Inversa	158
<b>Figura 4.48</b> – Gráfico dos Resíduos <i>Deviance</i> Versus Valores Ajustados. À Esquerda: Distribuição Gama. À Direita: Distribuição Normal Inversa	159
<b>Figura 4.49</b> - Gráfico dos Valores Absolutos dos Resíduos <i>Deviance</i> Versus Valores Ajustados. À Esquerda: Distribuição Gama. À Direita: Distribuição Normal Inversa	160
<b>Figura 4.50</b> – Gráficos com as Distâncias de Cook. À Esquerda: Distribuição Gama. À Direita: Distribuição Normal Inversa	160
<b>Figura 4.51</b> – Gráficos de Probabilidade Normal dos Resíduos <i>Deviance</i> . À Esquerda: Sem Interação. À Direita: Com Interação	162
<b>Figura 4.52</b> – Gráficos dos Resíduos <i>Deviance</i> Versus Valores Ajustados À Esquerda: Sem Interação. À Direita: Com Interação	162
<b>Figura 4.53</b> - Gráfico do Valor Absoluto do Resíduo <i>Deviance</i> Versus Valores Ajustados. À Esquerda: Sem Interação. À Direita: Com Interação	163
<b>Figura 4.54</b> – Gráficos dos Resíduos <i>Deviance</i> Versus Valores Ajustados: À Esquerda: Função de Ligação Logarítmica. À Direita: Função de Ligação Inverso	165
<b>Figura 4.55</b> – Gráficos dos Resíduos <i>Deviance</i> Versus Valores Ajustados: À Esquerda: Distribuição Normal. À Direita: Distribuição Gama	168
<b>Figura 4.56</b> - Gráfico dos Valores Absolutos do Resíduos <i>Deviance</i> Versus Valores Ajustados. À Esquerda: Distribuição Normal. À Direita: Distribuição Gama	168
<b>Figura 5.1</b> - Gráfico dos Resíduos <i>versus</i> Fator C - Modelo (5.1)	173
<b>Figura 5.2</b> - Gráfico dos Resíduos <i>versus</i> Fator C - Modelo (5.2)	174

<b>Figura 5.3</b> - Gráficos dos Resíduos <i>Versus</i> Fatores $F$ e $G$ - Modelo (5.2)	174
<b>Figura 5.4</b> - Gráfico com os valores de $F_j$	177
<b>Figura 5.5</b> - Gráfico de Probabilidade Normal de $F_j$ . – Modelo (5.1)	178
<b>Figura 5.6</b> - Gráfico de Probabilidade Normal de $F_j$ . – Modelo (5.2)	178
<b>Figura 5.7</b> - Gráfico dos Resíduos <i>versus</i> o Nível da Interação $AB$	179
<b>Figura 5.8</b> – Gráfico de POE <i>versus</i> $x_2$ e $x_3$	192
<b>Figura 5.9</b> – Gráfico de POE <i>versus</i> $x_2$ e $x_3$	193
<b>Figura 5.10</b> - Gráficos de POE e do Rendimento <i>Versus</i> Variáveis $x_2$ e $x_3$	193

## Lista de tabelas

<b>Tabela 2.1</b> - Dados para o Modelo de Regressão Linear	28
<b>Tabela 2.2</b> - Nível do Fator (Exemplo 2.1)	32
<b>Tabela 2.3</b> - Dados Resultantes do Experimento (Exemplo 2.1)	33
<b>Tabela 2.4</b> - ANOVA do Experimento (Exemplo 2.1)	33
<b>Tabela 2.5</b> - Testes para os Coeficientes	34
<b>Tabela 2.6</b> - Testes para os Coeficientes	34
<b>Tabela 2.7</b> - Testes para os Coeficientes	34
<b>Tabela 2.8</b> - Resíduos Ordenados e Frequência Cumulativa	36
<b>Tabela 2.9</b> - Resíduo Ordenado e Quartil Normal	37
<b>Tabela 2.10</b> - Dados “ <i>Ascomb</i> ”	47
<b>Tabela 3.1</b> - Dados para o Modelo	69
<b>Tabela 3.2</b> - Ligações Canônicas para os MLG	70
<b>Tabela 3.3</b> - Dados do Experimento	80
<b>Tabela 3.4</b> - Estimativa dos Coeficientes e Erro Padrão	80
<b>Tabela 3.5</b> - Análise de Deviance (ANODE)	81
<b>Tabela 4.1</b> - Dados do Experimento de Moldagem	102
<b>Tabela 4.2</b> - Estimativas dos coeficientes e Teste $t$ para o Modelo M2	103
<b>Tabela 4.3</b> - Estimativas dos Coeficientes e Teste $t$ para o Modelo Linear	107
<b>Tabela 4.4</b> - Estimativas das Respostas para o Modelo Linear e MLG-Gama	108
<b>Tabela 4.5</b> - Dados do Experimento	109
<b>Tabela 4.6</b> - Experimento da Perfuratriz	113
<b>Tabela 4.7</b> - Teste $t$ para Experimento da Perfuratriz	114
<b>Tabela 4.8</b> - Teste $t$ para Experimento da Perfuratriz	130
<b>Tabela 4.9</b> - Dados do Experimento	135
<b>Tabela 4.10</b> - Dados do Experimento	142
<b>Tabela 4.11</b> - Função de Log-verossimilhança do Modelo Completo	156
<b>Tabela 4.12</b> - Função de Log-verossimilhança dos Modelos Maior e Menor	156
<b>Tabela 4.13</b> - Análise de <i>Deviance</i> (ANODE)	157
<b>Tabela 4.14</b> - Valor da Função de Log-verossimilhança do Modelo Menor	164
<b>Tabela 4.15</b> - Valor da Função de Log-verossimilhança	166
<b>Tabela 4.16</b> - Valor da Função de Log-verossimilhança	167

<b>Tabela 5.1</b> - Cálculo dos Desvios-padrão dos Fatores <i>B</i> e <i>C</i>	176
<b>Tabela 5.2</b> - Estatística $F_j$ do Modelo (5.1)	177
<b>Tabela 5.3</b> - Modelos e Componentes para a Média e a Dispersão	184
<b>Tabela 5.4</b> - Experimento $2^3$ com Duas Replicações	185
<b>Tabela 5.5</b> - Estimativa da Variância da Resposta	186
<b>Tabela 5.6</b> - Coeficientes e Erro-Padrão do Modelo para a Dispersão	186
<b>Tabela 5.7</b> - Coeficientes e Erro-Padrão do Modelo da Média	187
<b>Tabela 5.8</b> - Coeficientes e Erro-Padrão do Modelo da Dispersão	187
<b>Tabela 5.9</b> - Coeficientes e Erro-Padrão do Modelo da Média	188
<b>Tabela 5.10</b> - Coeficientes e Erro-Padrão do Modelo da Dispersão	188
<b>Tabela 5.11</b> - Experimento do Exemplo 5.1	191