

**UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
ESCOLA DE ENGENHARIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE PRODUÇÃO**

**OTIMIZAÇÃO DE PROCESSOS COM RESPOSTAS MÚLTIPLAS E
CATEGÓRICAS**

Cícero de Melo Lucas

**Porto Alegre
2007**

**UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
ESCOLA DE ENGENHARIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE PRODUÇÃO**

**OTIMIZAÇÃO DE PROCESSOS COM RESPOSTAS MÚLTIPLAS E
CATEGÓRICAS**

Cícero de Melo Lucas

Orientador: Prof. Dr. Flávio Sanson Fogliatto

Banca Examinadora:

Prof^a Dr^a Carla Schwengber ten Caten

Prof^a Dr^a Liane Werner

Prof^a Dr^a Morgana Pizzolato

**Dissertação submetida ao Programa de Pós-Graduação em
Engenharia de Produção como requisito parcial à obtenção do título de
MESTRE EM ENGENHARIA DE PRODUÇÃO
Área de concentração: Sistemas de Qualidade**

**Porto Alegre
2007**

Esta dissertação foi julgada adequada para a obtenção do título de Mestre em Engenharia de Produção e aprovada em sua forma final pelo Orientador e pela Banca Examinadora designada pelo Programa de Pós Graduação em Engenharia de Produção.

Prof. Flávio Sanson Fogliatto, Dr.
PPGEP / UFRGS
Orientador

Prof. Flávio Sanson Fogliatto, Dr.
Coordenador do PPGEP / UFRGS

Banca Examinadora:

Carla Schwengber ten Caten, Dr^a

Prof^a Programa de Pós-Graduação em Engenharia de Produção e Transportes / UFRGS

Liane Werner, Dr^a

Prof^a Programa de Pós-Graduação em Engenharia de Produção e Transportes / UFRGS

Morgana Pizzolato, Dr^a

Prof^a Departamento de Engenharia de Produção e Transportes / UFRGS

DEDICATÓRIA

Dedico esta dissertação à memória do meu pai, Norberto Lucas Ferreira, um verdadeiro homem, que sacrificou-se à tarefa de educar os filhos com base em profundos valores, malgrado as gigantescas dificuldades que o destino lhe reservou em sua tão breve vida. Que Deus reconheça seu esforço em podar a videira que hoje rende frutos.

Dedico-a também à minha querida mãe, Maria Terezinha de Melo Lucas, por seu amor incondicional aos filhos e pela sua presença confortante nos períodos de dificuldades e provações.

Homenageio ainda minha amada noiva, Adriana Maschio, pelo apoio diuturno aos meus sonhos e às minhas conquistas.

AGRADECIMENTOS

Obrigado a Deus pela criação das coisas do mundo, e a seu Filho que foi Cordeiro de sacrifício para a salvação dos homens. Em espírito, agradeço a Santo Tomás de Aquino, cujas obras me auxiliam no árduo amadurecimento segundo as virtudes fundamentais da prudência, da justiça, da fortaleza e da temperança.

Agradeço aos pagadores de impostos brasileiros, pois foram obrigados a dedicar uma parcela de seus trabalhos aos meus estudos nesta universidade. Mesmo ciente que se fosse facultativo aos contribuintes talvez eu não recebesse tal benefício, gostaria de registrar meu respeito e agradecimento à produção que lhes foi tomada. A universidade gratuita é um mito, porquanto muitos pagam e pouquíssimos são beneficiados.

Agradeço, com grande estima, ao Prof. Flávio Sanson Fogliatto, *Ph.D* por todo seu empenho na minha orientação para este difícil trabalho. Obrigado também às professoras que compuseram minha banca de avaliação.

Gostaria de retribuir minha gratidão a Cláudio Roberto Xavier de Souza, técnico em informática do DEPROT, por ter em inúmeras oportunidades me auxiliado a resolver os problemas que tive com os computadores do LOPP. E ainda, pelo auxílio estatístico, retribuo meus agradecimentos ao colega Ângelo Márcio Oliveira Sant'Anna, *M.Sc.* Tenho ainda uma imensa cota de gratidão a Gustavo De Arriba, *M.Sc.* pela disposição em que se apresentou para esclarecer detalhes de seu trabalho no PPGEF, pois foram fundamentais para a condução desta dissertação.

Aos amigos que formei ao longo destes dois anos no mestrado, deixo meu obrigado pelas trocas de experiências.

Eu digo que um homem deve estar seguro de sua moralidade pela simples razão de que ele há de sofrer por causa dela.

G. K. Chesterton (1874-1936), escritor inglês.

RESUMO

A presente dissertação aborda a otimização de processos industriais através da utilização de projeto de experimentos. Em experimentos planejados, variáveis de respostas são usualmente consideradas como normalmente distribuídas. No entanto, em algumas situações, tal suposição é violada; por exemplo, quando respostas expressam contagens, podendo assumir somente valores inteiros e não-negativos. Nesses casos, é mais provável que as respostas sigam uma distribuição de Poisson e, em sua modelagem, deve-se utilizar os modelos lineares generalizados (MLG), adequados para respostas da chamada família exponencial, à qual pertence a distribuição de Poisson. Se ainda persiste a dúvida quanto ao comportamento das respostas, o modelo de quase-verossimilhança também é uma alternativa possível. Esta dissertação apresenta a reanálise de um experimento, apresentado em Arriba (2005), onde algumas respostas são categóricas. Na análise original do experimento, respostas categóricas foram modeladas através de regressão dos mínimos quadrados ordinários, desconsiderando a violação do pressuposto de normalidade das respostas. Na reanálise aqui apresentada, as variáveis são corretamente abordadas usando-se os modelos lineares generalizados. Como o objetivo do trabalho de Arriba (2005) era a otimização de um processo descrito por múltiplas respostas, comparam-se os resultados da otimização mediante as duas estratégias de modelagem das respostas, além de se propor um método alternativo, mais simplificado, de otimização experimental.

Palavras-chave: otimização de processos, modelos lineares generalizados, modelo de quase-verossimilhança, regressão de Poisson, capacidade de processos.

ABSTRACT

This dissertation deals with the optimization of industrial processes using Design of Experiments. In designed experiments, response variables are often considered as normally distributed. However, in some situations, such assumption is violated; for example, when responses express counts, and only non-negative integers numbers may come up as outcomes. In these cases, it is likely that responses follow a Poisson distribution which is then modeled by generalized linear models (GLM), since such distribution is a member of the exponential family. If a question still holds on the responses behavior, their modeling through the quasi-likelihood method is another option that should be considered. This dissertation analyses an experiment performed by Arriba (2005), where some responses are of categorical nature. In the original analysis presented by Arriba (2005), categorical responses were modeled using ordinary least squares regression, violating the normality assumption associated with that method. In the analysis presented here, variables are appropriately modeled using GLM. Since the objective in Arriba (2005) was optimizing a multiresponse process, results from the two optimization processes are compared. In addition, an alternative experimental optimization method, simpler than the one in Arriba (2005), is also presented.

Key words: process optimization, generalized linear models, quasi-likelihood models, Poisson regression, process capability.

Sumário

CAPÍTULO 1.....	10
1.1 Considerações iniciais	10
1.2 Tema e Objetivos.....	12
1.3 Justificativa do trabalho.....	13
1.4 Método de trabalho.....	16
1.5 Limitações do trabalho	17
1.6 Estrutura do trabalho	18
CAPÍTULO 2.....	20
2.1 Os modelos lineares generalizados.....	21
2.2 O índice de capacidade de processos	46
CAPÍTULO 3.....	56
3.1 O projeto de experimento fatorial de Arriba (2005).....	56
3.2 Metodologia aplicada	67
CAPÍTULO 4.....	72
4.1 As regressões por modelos lineares generalizados e quase-verossimilhança	72
4.2 A otimização do experimento com resposta de Poisson	78
4.2 A otimização através de índices de capacidade de processos	83
4.3 Análise dos resultados	86
CAPÍTULO 5.....	88
5.2 Conclusões.....	88
5.2 Sugestões para trabalhos futuros	90
REFERÊNCIAS BIBLIOGRÁFICAS.....	91
OBRAS CONSULTADAS.....	93
APÊNDICE A.....	94
APÊNDICE B	96

CAPÍTULO 1

1.1 CONSIDERAÇÕES INICIAIS

Experimentos, do ponto de vista da Engenharia da Qualidade, podem ser formalmente definidos como uma série de testes, nos quais mudanças propositais são realizadas nas variáveis de entrada, ou fatores, de um processo ou sistema. Estas alterações planejadas permitem identificar e controlar as influências destes fatores sobre a variável de resposta do processo (MONTGOMERY, 2005).

A evolução das técnicas de experimentação tem sido amplamente aproveitada pelas empresas. A razão que melhor justifica este interesse é o fato que estas técnicas, nascidas dentro dos estudos da Probabilidade e da Estatística, permitem não apenas avaliar se as respostas resultantes dos processos são estáveis ou não, mas também definir os fatores causais que mais as influenciam. Além disso, técnicas de experimentação, também contribuem para descobrir-se em quais níveis de regulação destes fatores são alcançados os resultados desejáveis pela empresa. Os resultados de interesse, no âmbito produtivo, freqüentemente são: menor refugo, maior rendimento, menor tempo de operação, menor quantidade de matéria-prima, maior aproveitamento dos recursos produtivos, redução da variabilidade das respostas, aproximação a uma certa medida de qualidade ou alvo etc. Naturalmente que num contexto de livre mercado e iniciativa estes interesses são ampliados, pois implicam em uma maior oportunidade de ganhos adicionais.

Se, num primeiro momento, a utilização de experimentos estatisticamente planejados visava encontrar apenas a melhor localização da variável de resposta desejada, ou seja, as médias (*location effects* ou efeitos de localização), atualmente pode-se obter também a minimização da variabilidade da resposta, ou variância (*dispersion effects* ou efeitos de dispersão). Uma vez criteriosamente selecionados os fatores e seus respectivos níveis dos processos que conduzem às respostas próximas o bastante da média desejada e com baixa variância, resta ainda uma maior robustez, ou insensibilidade, da variável de resposta aos ruídos aleatórios inerentes a qualquer processo. Esta seqüência de ações faz parte da espinha dorsal da Engenharia da Qualidade, e todos eles – localização da média, minimização da variância e robustez aos ruídos – são buscados simultaneamente nos experimentos planejados (MONTGOMERY, 2001).

Os projetos de experimentos podem ser classificados de várias formas. Quanto à característica da variável de resposta, eles podem ser de dois tipos: respostas contínuas ou respostas discretas. As primeiras incluem medidas que admitem qualquer valor entre dois valores quaisquer. Por exemplo, se em um experimento a variável de resposta é consumo de combustível em km/l, entre uma resposta de 13 km/l e outra de 12 km/l, pode-se obter infinitas outras respostas, tais como 12,4 km/l, 12,8 km/l etc. A limitação está na acurácia e resolução do instrumento de medida. Quanto ao segundo tipo, as respostas discretas, só podem ser admitidos certos valores específicos, e entre dois deles não há a possibilidade natural de uma terceira medida. Se em um lote de cinquenta peças encontram-se peças boas e ruins, a contagem de peças boas, por exemplo, poderá ser qualquer valor inteiro entre zero e cinquenta. Porém, esta contagem nunca será 12,4 peças ruins entre 11 e 13 elementos, por uma característica natural e evidente da contagem.

Ambos os casos são muito comuns na vida prática e assumem uma importância ainda maior no contexto empresarial em virtude da diversidade de situações experimentais possíveis na otimização de produtos e processos. No entanto, por razões que vão desde a complexidade matemática e a oferta de cursos preparatórios na técnica, até ao conhecimento dos métodos estatísticos e à disponibilidade de uso nos *softwares* específicos, o segundo tipo de resposta, as discretas, são menos utilizadas e corretamente tratadas do que as contínuas. Embora, de forma alguma, elas possam ser ditas menos importantes. Uma consequência desta situação é considerar-se respostas claramente discretas como contínuas, e dar a elas o tratamento matemático segundo esta consideração. Em alguns casos isto não conduz a erros muito graves, mas há situações onde esta aproximação é inaceitável, uma vez que induz a decisões gravemente equivocadas.

Os produtos raramente possuem uma única característica de qualidade. O mais comum é que essas características sejam diversas, algumas mais importantes que outras, na proporção do valor que o cliente atribui a cada uma delas. Voltando ao exemplo do consumo de combustível, considere-se como sistema um automóvel. O cliente deseja deste sistema, além do reduzido consumo de combustível, alta aceleração, frenagem eficaz, estabilidade nas curvas, baixo nível de ruído etc. As técnicas que permitem combinar da melhor forma possível todas estas características de qualidade são conhecidas como otimização multiresposta, ou seja, o sucesso está em obter-se o melhor arranjo das grandezas que são exigidas e reconhecidas pelos clientes como valiosas e, assim considerar o produto como 'de ótima qualidade'.

Uma decorrência natural da tentativa de se encontrar a melhor combinação dos níveis dos fatores que resultarão nas melhores respostas de cada característica de qualidade é o conflito entre elas. Não raro a melhoria de uma dessas características implica na piora de outra. Disto decorre a necessidade de se atribuir pesos para cada resposta conforme sua importância no processo de otimização, de forma que representem a maior ‘qualidade total’ do produto; ou melhor ainda, encontrar o arranjo que melhor satisfaça todas as exigências do cliente, com o mínimo detrimento de uma ou outra característica.

Assim, por um lado há uma série de respostas discretas cujo tratamento adequado nos projetos de experimentos é ainda insuficiente e incorretamente utilizado pelas empresas e até no meio acadêmico. Por outro lado, há os casos onde não apenas uma, mas várias respostas requerem uma combinação dos fatores em seus respectivos níveis de forma a otimizar a qualidade percebida, sendo comum o conflito entre essas respostas. Assim, o encontro desses dois aspectos é vastamente comum no dia-a-dia de muitas atividades, e portanto merecedores de atenção especial para quem deseja produzir bens com maior eficiência, menor uso de recursos, menor custo e com as características que atendam às preferências dos clientes.

1.2 TEMA E OBJETIVOS

O tema principal desta dissertação é a otimização da qualidade quando esta é expressa por múltiplas características ou seja, processos multiresposta, e dentre elas, há a presença de pelo menos uma variável de resposta categórica, ou discreta. A necessidade deste tipo de otimização é típica no meio produtivo, embora nem sempre sejam utilizados os recursos já desenvolvidos para solucionar problemas desta natureza específica.

Agresti (2002) afirma que em virtude da presença de variável de resposta categórica faz-se necessária a abordagem pela teoria dos modelos lineares generalizados para gerar os modelos de regressão linear, assim como sua respectiva técnica de estimação dos parâmetros deste modelo, além dos testes de hipótese para cada um destes parâmetros que irão garantir o quanto se pode confiar na regressão obtida.

São, portanto, objetivos principais desta dissertação:

- a) otimizar um processo multiresposta quando pelo menos uma das variáveis de resposta é categórica utilizando modelos lineares generalizados na modelagem das respostas. Tomou-se como exemplo a otimização do processo de curtimento de couros executada em Arriba (2005);
- b) Propor uma outra forma de otimização diferente daquela realizada por Arriba (2005). Este autor usou uma função objetivo com pesos de ponderação de importância, que é uma forma mais direta da otimização hierárquica. Este critério será comparado à otimização através da maximização conjunta de índices de capacidade de processo calculados para as múltiplas respostas proposto por Chan, Cheng e Spiring (1988).

Os objetivos secundários derivados dos anteriores são:

- a) apresentar uma introdução aos modelos lineares generalizados que permita ampliar sua apropriada utilização na otimização de processos;
- b) comparar os resultados de dois procedimentos de modelagem de dados com vistas à otimização de um experimento;
- c) comparar os resultados de duas estratégias de otimização multiresposta aplicadas a este experimento.

1.3 JUSTIFICATIVA DO TRABALHO

Arriba (2005) abordou, entre outros processos, a otimização do processo de seleção de couros no estágio *wet-blue* de curtimento. Neste trabalho os couros foram avaliados por especialistas quanto à sua qualidade, representada especialmente pela ausência de marcas na pele. Foi dada uma nota - ou seja, um número inteiro e positivo - para cada peça conforme sua utilidade como produto final. As peles com melhor nota eram encaminhadas para a utilização em produtos de maior valor agregado. Este processo de seleção ocorre no estágio inicial, e quando há erros ocorrem perdas financeiras consideráveis, uma vez que somente quando encerrado o processo e incorridos os gastos até a finalização do curtimento é possível identificar de fato a real classificação da pele.

Arriba (2005) considerou na modelagem dos dados que suas respostas atendiam aos critérios aceitos pelo método de regressão dos mínimos quadrados. No entanto, suas respostas

violavam os seguintes pressupostos implícitos para uso adequado deste método, conforme indicam Montgomery, Peck e Vining (2006, p.160):

- a) os erros do modelo têm média zero, variância constante e não são correlacionados;
- b) o modelo dos erros segue uma distribuição de probabilidade normal. Este pressuposto é feito para a condução dos testes de hipótese e intervalos de confiança do modelo, e também resulta que os erros são independentes entre si;
- c) a forma do modelo, incluindo-se aí a especificação das variáveis regressoras, está correta.

É muito comum no projeto de experimentos o recurso de otimização de uma equação de regressão cuja resposta seja categórica com se esta fosse contínua e obedecendo a uma distribuição normal. No caso de Arriba (2005) as duas primeiras enumerações acima foram violadas. Em muitas outras situações, talvez na maioria delas, violar estes pressupostos conduz a graves erros de análise. Bisgaard e Fuller (1994) afirmam que os métodos usados para análise de experimentos fatoriais quase sempre são baseados na aceitação dos pressupostos indicados acima. Os autores acrescentam que, como as ferramentas padrão de análise estatística são muito poderosas, pode-se simplesmente ignorar estes pressupostos para o caso das respostas do tipo ‘contagem’ causando prejuízo às conclusões.

Assim, nesta dissertação é apropriadamente tratada a otimização de respostas categóricas, particularmente do tipo contagem ou números inteiros e positivos, nos modelos de regressão. A teoria dos modelos lineares generalizados é a abordagem adequada para este tipo de análise, portanto seu desenvolvimento será detalhado de forma a prevenir o erro comumente incorrido nos casos apresentados.

Além do mais, por se tratar de um assunto que exige o conhecimento de teorias mais avançadas em Matemática e Estatística, e também fora do conteúdo dos cursos de graduação em Engenharia, faz-se necessário um texto em linguagem mais próxima ao dia-a-dia do engenheiro, considerando os processos industriais que solicitam a abordagem apropriada pelos modelos lineares generalizados. A este problema, acrescenta-se ainda a falta de um único livro texto em português sobre o assunto e que tenha ampla distribuição, tal qual quando se trata da regressão linear pelo método dos mínimos quadrados, estando ainda o tema

dos modelos lineares generalizados restrito a teses, dissertações e artigos científicos na língua brasileira. No entanto, a partir do aumento de publicações desta natureza, contribui-se para chamar a atenção para o assunto a ponto de justificar a edição de um livro texto que trate dos modelos lineares generalizados, ou pelo menos uma boa tradução dos livros já consagrados em outros idiomas.

É importante ainda que se alerte as empresas dos equívocos do uso indiscriminado da regressão linear cujos parâmetros são estimados pelo método dos mínimos quadrados quando as respostas são categóricas. Melhor orientados, os profissionais da área de qualidade que se depararem com este tipo particular de problema terão mais sucesso na otimização de seus processos críticos, ou seja, aqueles que mais impactam no resultado da empresa.

Arriba (2005) gerou uma quantidade grande de dados confiáveis, pois antes da sua coleta foram desenvolvidos até mesmo estudos de Reprodutibilidade e Repetitividade (R&R), cujos resultados foram depurados em várias etapas até chegar-se àqueles efetivamente usados na pesquisa. Assim, seus experimentos são muito valiosos. A otimização das várias respostas atendeu ao critério de pesos dados a cada uma delas de forma a conferir sua relevância para o resultado final da função objetivo. Como já dito acima, será realizada neste trabalho a mesma otimização, no entanto usando o modelo apropriado de regressão. No entanto, uma outra forma de otimização também será usada através dos índices de capacidade de processo destas respostas.

Os índices de capacidade de processo são adimensionais e indicam o nível da qualidade de um processo: quanto maior o índice maior a qualidade do atributo medido. O índice de capacidade múltiplo mede a combinação de várias respostas, cada uma com seu respectivo índice de capacidade. Assim, nesta dissertação também será avaliado o índice de capacidade múltiplo das respostas dos experimentos conduzidos por Arriba (2005), ou seja, a combinação conjunta de cada índice de capacidade de processo respectivo a cada característica de qualidade medida. Desta forma, pode-se comparar este método e aquele anteriormente usado.

Resumidamente, esta dissertação justifica-se por: contribuir com a divulgação da teoria dos modelos lineares generalizados para a regressão de respostas categóricas, analisar o experimento apresentado em Arriba (2005) utilizando uma técnica apropriada para modelagem das variáveis de resposta categóricas nele presente, otimizar as múltiplas

respostas daquela dissertação segundo o mesmo método proposto pelo autor, otimizar as mesmas respostas usando o método da capacidade de processos múltipla, e, por fim, comparar ambos os métodos de otimização multiresposta.

1.4 MÉTODO DE TRABALHO

Nesta dissertação duas formas de otimização serão aplicadas e seus resultados discutidos. A primeira delas será maximizar o valor de uma função objetivo como valores de ponderação para cada uma das respostas da característica de qualidade de interesse. Esta foi a técnica utilizada por Arriba (2005). A outra forma será pela maximização dos índices de capacidade de processo de cada uma das respostas, que dentro do método serão agregadas multiplamente em seguida compondo um índice múltiplo de capacidade de processo, segundo os mesmos pesos de ponderação da função objetivo. Chan, Cheng e Spiring (1988) introduziram uma nova medida de capacidade de processo, que será utilizado neste caso. Este índice leva tanto em consideração a proximidade de uma medida ao seu alvo, quanto a variabilidade do processo.

Para encontrar-se o arranjo ótimo que maximiza tanto a função objetivo quanto os múltiplos índices de capacidade será usado o algoritmo do gradiente generalizado reduzido discutido em Castillo e Montgomery (1993), e de uso corrente nas planilhas eletrônicas mais comerciais, tais como a função *Solver* do Excel[®].

A estimação dos parâmetros das respostas dos experimentos, por serem categóricas em pelo menos um caso, será feita pela teoria dos modelos lineares generalizados que solicita, por sua vez, o método da máxima verossimilhança. Este método é o mais indicado por ser amplo o suficiente para a estimação de vários tipos de respostas, em especial aquelas cujas funções de distribuição de probabilidade pertencem à família exponencial (NELDER e WEDDERBURN, 1972; McCULLAGH e NELDER, 1990). Desenvolvido o problema desta forma, serão corrigidas as violações de pressupostos da regressão linear incorrida por Arriba (2005).

De forma a classificar de forma ampla o tipo desta pesquisa conforme seus objetivos, pode-se afirmar, com base em Gil (2002), que esta é uma pesquisa exploratória, uma vez que

visa proporcionar maior familiaridade com o problema com vistas a torná-lo mais explícito, além de analisar exemplos que estimulam a compreensão do problema com uma proposta de solução.

Diante destas considerações acima, esta é, portanto, uma pesquisa aplicada, visto que é seu objetivo a aplicação de conhecimentos específicos e bem delimitados sendo seu uso já comprovadamente eficaz, além de, segundo Roesch (2005), incluir a preocupação teórica que fundamenta a solução dentro de seus pressupostos. Para tanto, será necessária uma revisão da literatura que aborde: *(i)* a introdução e conceituação dos índices de capacidade de processos; *(ii)* a teoria dos modelos lineares generalizados e seu respectivo método de estimação de parâmetros e *(iii)* apresentação de técnicas de amplo uso de otimização multiresposta.

A necessidade de uma forte ênfase em métodos estatísticos e matemáticos, processo industriais sujeitos a otimização e experimentos planejados para se alcançar os objetivos esperados, faz com este trabalho também seja uma pesquisa quantitativa. De certa forma, devido às respostas dos experimentos serem, em última instância, opiniões dos técnicos à respeito da qualidade do couro, esta seria também uma pesquisa qualitativa, no entanto, não serão feitas maiores considerações sobre este tipo de pesquisa que está restrita ao desenvolvimento de Arriba (2005). Nesta forma de classificação este trabalho é, portanto, estritamente quantitativo.

1.5 LIMITAÇÕES DO TRABALHO

Nesta dissertação será tratada a otimização de experimentos quando a resposta é categórica, e também quando há a necessidade de se encontrar o melhor arranjo entre múltiplas respostas de interesse para a produção. Embora o escopo da revisão bibliográfica abranja as categorias de respostas dicotômicas, quando apenas duas respostas são possíveis - por exemplo, aprovado e rejeitado, conforme e não-conforme - estes casos não serão tratados na solução prática. Ou seja, no caso aplicado as respostas serão sempre valores inteiros e positivos, uma característica inerente à distribuição de Poisson.

As respostas podem ainda ser categóricas, mas múltiplas, podendo ser ordinais quando entre as várias respostas possíveis há uma implicação de graduação entre elas (por

exemplo um, dois ou três defeitos por peça produzida). Há ainda um terceiro tipo de respostas categóricas, ditas nominais, onde a graduação implícita não existe: por exemplo quando ocorrem defeitos de naturezas diferentes em uma mesma peça, tais como risco, bolha ou descoloração numa pintura. Unicamente o caso ordinal será objeto de estudo neste trabalho; portanto, fica fora do interesse de investigação os casos de respostas categóricas nominais.

Não será tratado o caso da regressão linear clássica pelo método dos mínimos quadrados, pois o método da máxima verossimilhança para definição dos parâmetros da regressão linear já engloba esta situação particular, assim como as respostas de natureza contínua, além do que há vasta e boa literatura tratando do assunto, mesmo em Português.

São feitas as otimizações de duas formas. A primeira reproduz aquela usada por Arriba (2005), de forma a permitir a comparação com seus resultados. A segunda lança mão do índice múltiplo de capacidade de processos de forma a estabelecer conclusões de utilidade e conveniência com a anterior. Ambos os casos usarão o algoritmo do gradiente generalizado reduzido para encontrar-se o máximo da função objetivo. Existem na literatura vários outros métodos de otimização, nenhum desses será tratado neste trabalho, uma vez que o algoritmo selecionado já conduz a resultados satisfatórios no âmbito do problema tratado. Esses outros métodos de otimização, para citar dois, são a metodologia da superfície de resposta (RSM) tratado em Montgomery (2005) e a função de perda quadrática multivariada abordada em Elsayed e Ribeiro(1995).

Arriba (2005) otimizou quatro artigos diferentes em couro: (i) *nubuck* chocolate, (ii) *naplex* preto; (iii) *diamond* preto e (iv) *nubuck* bege. Nesta dissertação será otimizado apenas seu experimento para o artigo *nubuck* chocolate.

1.6 ESTRUTURA DO TRABALHO

Esta dissertação é composta por cinco capítulos. No primeiro capítulo é apresentada uma introdução ao problema que este trabalho vai tratar e buscar uma solução. São discutidos também seus temas e objetivos principais e secundários de forma a esclarecer o alcance de sua proposta que também é justificada. Para tanto, é apresentado sucintamente um método de

trabalho que julga-se conseguir chegar aonde se propõe para solucionar o problema encontrado, limitando-se aos assuntos necessários e suficientes para tal.

No segundo capítulo é apresentado o referencial teórico que sustenta o problema prático apresentado. Para tanto são tratados, nesta ordem:

- a) os modelos lineares generalizados: sua estrutura formal, a abrangência da família exponencial, a regressão linear pelo método da máxima verossimilhança e um método para estabelecer a confiabilidade nos parâmetros da regressão Além da regressão de quase-verossimilhança;
- b) o índice de capacidade de processos: serão discutidos os principais índices de capacidade de processos e, em particular, o índice de capacidade de processos múltiplos de forma a poder lidar com processos que possuem mais de uma característica de qualidade de interesse.

No terceiro capítulo é revisto o experimento conduzido por Arriba (2005) de forma a detalhar o encaminhamento dado pelo autor para sua otimização. Como dito anteriormente, as respostas deste experimento foram tratadas como obedecendo à uma distribuição de probabilidade normal, no entanto por serem categóricas elas não se ajustam a este pressuposto. Ainda neste capítulo também são detalhados os procedimentos segundo a metodologia de otimização proposta deste texto.

No quarto capítulo, as mesmas respostas encontradas por Arriba (2005) são desta vez modeladas a partir dos modelos lineares generalizados. O modelo de otimização hierárquica utilizado por este autor é novamente aplicado e uma nova estratégia de otimização, desta vez pelo índice múltiplo de capacidade de processo, é também conduzida e seus resultados apresentados.

No quinto capítulo, por fim, os resultados encontrados são discutidos, em especial, de uma forma que ressalte as diferenças entre os obtidos entre este trabalho, aplicando o método proposto, e aqueles apresentados por Arriba (2005). Em ambos os casos haverá a preocupação em destacar as limitações de cada solução, esclarecendo onde os recursos teóricos utilizados são inapropriados e também onde mais eles podem ser aplicados. Encerrando este capítulo, para efeito de contribuição com o progresso do assunto no meio acadêmico e industrial, são dadas sugestões de trabalhos futuros.

Capítulo 2

Neste capítulo são tratados dois assuntos distintos. O primeiro é sobre os modelos lineares generalizados. São apresentadas a modelagem das médias dentro deste método e as técnicas de aferição da confiança que se pode depositar no modelo gerado. Esta seção adianta-se ainda um pouco mais além ao tratar dos modelos de quase-verossimilhança onde a suposição da distribuição de probabilidades das respostas não é necessária.

A segunda parte deste capítulo trata dos índices de capacidade de processos desde sua concepção inicial até seu uso para avaliação de múltiplas características de qualidade. É também desenvolvido o modelo que permite a substituição dos valores que calculam a capacidade de processos por modelos de regressão. Este procedimento permitirá a otimização do processo estudado sob a óptica dos modelos lineares generalizados no capítulo 4.

Será usada a notação mais comumente adotada pela literatura e textualmente assumida por Dobson (1990). As variáveis aleatórias são anotadas em letras itálicas maiúsculas e seus respectivos valores observados pelas mesmas, porém, minúsculas; por exemplo as observações y_1, y_2, \dots, y_N são consideradas como resultantes das variáveis aleatórias Y_1, Y_2, \dots, Y_N . As letras gregas designam os parâmetros, sendo que o símbolo $\hat{\cdot}$ (chapéu) adicionado sobre ele representa um estimador do parâmetro. Se β é o parâmetro de uma regressão, sua estimativa será, portanto, $\hat{\beta}$. Exceção a esta regra será feita ao uso da letra grega ε para designar o termo de erro. A notação para vetores e matrizes é em negrito, por exemplo, o vetor das observações \mathbf{y} , é a representação de:

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix}$$

A transposição de matrizes será indicada pela sobrescrito T . Aproveitando o exemplo anterior:

$$\mathbf{y} = [y_1 \quad y_2 \quad \dots \quad y_n]^T$$

As derivadas podem ser designadas tanto pelo operador $d(\text{função})/d(\text{variável})$, quanto pelo símbolo linha sobrescrito ($'$) quando expressamente indicado. A função de densidade de probabilidade é denominada por $f(y; \theta)$, onde θ representa os parâmetros da distribuição. E finalmente, a barra acima da letra, grega ou não, indica tratar-se de uma média, por exemplo: \bar{y} .

2.1 OS MODELOS LINEARES GENERALIZADOS

Uma classe de modelos lineares generalizados (MLG) foram propostos por Nelder e Wedderburn (1972). A apresentação de certas equações de distribuição de probabilidade assume genericamente o formato canônico – por formato canônico aqui entende-se o arranjo dos componentes de uma equação com propósito de demonstração e facilidade de entendimento, porém sem alterar sua estrutura – e, desta maneira, permite generalizar o procedimento de estimação de parâmetros baseado no método da máxima verossimilhança:

$$f(y, \theta, \phi) = \exp \left[\frac{y\theta - b(\theta)}{a(\phi)} + c(y, \phi) \right] \quad (2.1)$$

Na equação (2.1) Fahrmeir e Tutz (1994) definem seus termos. Onde y é a imagem da distribuição de probabilidade da família exponencial. O termo imagem é aqui denotado no contexto da linguagem matemática para funções. O parâmetro θ representa a localização no eixo das abscissas da curva de distribuição, é também o chamado parâmetro natural. E o parâmetro ϕ designa a variabilidade da distribuição, ou seja, este termo definirá se a curva terá um formato mais ‘alargado’ ou mais ‘estrito’ em relação ao parâmetro de localização. A função $b(\theta)$, descreverá o comportamento da localização, da mesma forma que a função $a(\phi)$ o faz com a variabilidade. É comum encontrar na literatura apenas o termo ϕ , sem o indicador de função. O termo $c(y, \phi)$ é uma função específica que dependerá do tipo de membro da família exponencial com que se estará lidando.

As equações de distribuição de probabilidade passíveis de apresentação neste formato configuram a chamada família exponencial, pois seus elementos estão elevados à base “ e ” ou “ $\exp(\text{função})$ ”. O modelo é também dito canônico porque primariamente o termo y que multiplica θ , era uma função $a(y)$ sujeita à igualdade $a(y)=y$ (DOBSON, 1990).

Neste grupo especial há componentes tanto de distribuições de probabilidade cujas respostas são contínuas quanto discretas. Myers, Montgomery e Vining (2002) citam as distribuições normal, binomial, Poisson, geométrica, binomial negativa, exponencial, gama e normal inversa como pertencentes a esta família exponencial. Os mesmos autores ainda afirmam que a estruturação destas distribuições no formato apresentado unifica os modelos de regressão lineares com os não-lineares. Assim, qualquer processo cuja variável de resposta seja distribuída segundo uma das equações da família exponencial poderá ser modelada dentro da teoria dos modelos lineares generalizados, quer seja linear ou não-linear em outro método de estimação de parâmetros que definem a função de distribuição.

A equação (2.1) é genérica e representa a família exponencial. Ela é composta pelas funções específicas $a(\phi)$, $b(\theta)$, $c(y, \phi)$. O parâmetro θ está representando a localização, ou seja, representa a região da curva de distribuição onde há maiores chances de se encontrar um valor aleatório, portanto, a média ou esperança. O parâmetro ϕ representa a dispersão ou em que medida os valores se afastam ou se aproximam do parâmetro de localização, sendo conhecido por variância (VIEIRA, 2004). Com ambos os parâmetros bem definidos pode-se compreender o comportamento da resposta de um processo.

A equação de distribuição normal, por exemplo, fornece respostas contínuas e é definida em sua apresentação mais conhecida como:

$$f(y, \mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(y - \mu)^2}{2\sigma^2}\right] \quad (2.2)$$

Escrita de maneira a adequar-se ao formato da equação (2.1), aplicando unicamente a propriedade logarítmica geral $a \times \exp(b) = \exp[b + \ln(a)]$ e organizando os expoentes, parcelas e quocientes, canonicamente, obtém-se:

$$f(y, \mu, \sigma) = \exp\left[\frac{y\mu - \frac{\mu^2}{2}}{\sigma^2} - \frac{1}{2} \ln(2\pi\sigma^2) - \frac{y^2}{2\sigma^2}\right] \quad (2.3)$$

As equações (2.2) e (2.3) são idênticas nas suas respostas, apenas operações algébricas foram implementadas com o intuito de se obter o formato indicado por (2.1). Assim, torna-se imediato identificar que:

$$\theta = \mu; b(\theta) = \frac{\mu^2}{2}; a(\phi) = \phi; \phi = \sigma^2 \text{ e } c(y, \phi) = \frac{1}{2} \ln(2\pi\sigma^2) - \frac{y^2}{2\sigma^2}$$

Como a curva normal é a mais comum e usada das curvas de distribuição, percebe-se pela identificação das funções específicas $a(\phi)$, $b(\theta)$, $c(y, \phi)$ aplicadas neste caso que o parâmetro de localização θ é justamente a média μ , ao passo que o parâmetro ϕ de dispersão é propriamente σ^2 , a variância. O formato de apresentação da equação (2.1) fornece de imediato a média e a variância de uma distribuição, sendo esta sua primeira grande utilidade. De fato, partindo dos conceitos de média e variância, confirma-se em Vieira (2004):

$$E(y) = \frac{db(\theta)}{d\theta} = \mu;$$

$$\text{var}(y) = \frac{d^2b(\theta)}{d\theta^2} \cdot a(\phi) = \sigma^2$$

Onde por definição e também por se tratar de uma variável contínua, uma função de distribuição de probabilidades qualquer $f(x)$, conforme Duncan (1974):

$$\mu = \int_{-\infty}^{+\infty} xf(x)dx$$

$$\sigma^2 = \int_{-\infty}^{+\infty} (x - \mu)^2 f(x)dx$$

Nesta dissertação as respostas sujeitas às distribuições binomial e de Poisson serão amplamente utilizadas, a fundamental diferença entre elas e a distribuição normal é que suas respostas não são contínuas, são discretas ou categóricas. A mais notável semelhança é que todas as três são pertencentes à família exponencial e podem, quando algebricamente trabalhadas, assumir o formato da equação (2.1).

A distribuição binomial, segundo Kelton e Law (1991), é útil para as situações onde ocorre um processo de Bernoulli, ou seja, entre duas únicas e mutuamente exclusivas respostas possíveis independentes entre si, tais como “sucesso” ou “fracasso”, faz-se a contagem delas dentro de numa amostra de tamanho “n”. É definida como:

$$f(y; \mu) = \binom{n}{y} \mu^y (1 - \mu)^{n-y} \quad (2.4)$$

Distribuindo os termos da equação (2.4) convenientemente chega-se ao seu formato exponencial (VIEIRA, 2004):

$$f(y, \mu) = \exp \left[y \cdot \ln \left(\frac{\mu}{1 - \mu} \right) + n \cdot \ln(1 - \mu) + \ln \binom{n}{y} \right] \quad (2.5)$$

Da mesma forma como foi analisado o formato exponencial da distribuição normal, serão identificados os componentes de localização e dispersão para a distribuição binomial:

$$\theta = \ln \left(\frac{\mu}{1 - \mu} \right); b(\theta) = n \cdot \ln(1 - \mu); a(\phi) = \phi; \phi = 1 \text{ e } c(y, \phi) = \ln \binom{n}{y}$$

Encontrando-se a respectiva média e variância:

$$E(y) = \frac{db(\theta)}{d\theta} = \mu;$$

$$\text{var}(y) = \frac{db^2(\theta)}{d\theta^2} \cdot a(\phi) = \mu(1 - \mu)$$

Onde por definição e também por se tratar de uma variável discreta, uma função de distribuição de probabilidades qualquer $p(x_i)$, segundo Montgomery (2001):

$$\mu = \sum_{i=1}^{\infty} x_i p(x_i)$$

$$\sigma^2 = \sum_{i=1}^{\infty} (x_i - \mu)^2 p(x_i)$$

O mesmo procedimento é utilizado para a distribuição de Poisson apresentado na forma mais comum de se encontrar, exposto em Kelton e Law (1991), e já ajustado ao formato de apresentação exponencial em Vieira (2004):

$$f(y, \mu) = \frac{\mu^y \cdot e^{-\mu}}{y!} = \exp[y \cdot \ln(\mu) - \mu] - \ln(y!) \quad (2.6)$$

Identificando-se as equações específicas, encontra-se:

$$\theta = \ln(\mu); b(\theta) = -\mu; a(\phi) = \phi; \phi = 1; c(y, \phi) = -\ln(y!)$$

Conforme as definições de média e variância:

$$E(y) = \frac{db(\theta)}{d\theta} = \mu;$$

$$\text{var}(y) = \frac{db^2(\theta)}{d\theta^2} \cdot a(\phi) = \mu$$

Kelton e Law (1991) ainda apresentam que a distribuição de Poisson possui como característica fundamental ser composta exclusivamente por número inteiros e não-negativos. Esta característica será fundamental para a otimização a ser levada adiante neste trabalho.

Assim, conforme Myers *et al.* apud Elsayed e Ribeiro (1995) que recomendam os modelos lineares generalizados para modelagem da variância, uma vez escolhida a função de probabilidade, que no caso deste texto são as distribuições binomial de Poisson, o formato exponencial automaticamente fornece a função de variância $V(\mu)$ que é a porção da variância da resposta y que depende da média e o parâmetro de dispersão ϕ que não depende da média e é constante para os membros da família exponencial VIEIRA (2004).

O texto de Vieira (2004) é usado extensivamente para: (i) a estruturação dos modelos lineares generalizados, (ii) o método matemático para estimação dos parâmetros de modelagem e (iii) os testes de hipóteses de confiabilidade destes mesmos parâmetros. Com suporte dos respectivos apoios usados por ele que são Dobson (1990) e Myers, Montgomery e Vining (2002).

Assume-se um experimento com k variáveis independentes e n respostas y , conforme é apresentado na tabela (2.1):

Tabela (2.1)
Estrutura para modelos lineares generalizados

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

Fonte: Vieira (2004).

Definindo os termos da tabela (2.1):

As variáveis de resposta de cada rodada do experimento y_1, y_2, \dots, y_n terão como médias $\mu_1, \mu_2, \dots, \mu_n$.

A distribuição de probabilidade de y_i pertence genericamente a um dos membros da família exponencial conforme Myers, Montgomery e Vining (2002).

As variáveis de regressão ou variáveis independentes do modelo serão determinadas pelos valores possíveis assumidos por x_1, x_2, \dots, x_k . Note-se que x_1, x_2, \dots, x_k podem ser contínuas ou discretas.

O modelo assume o formato de um preditor linear η do tipo:

$$\eta = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k \quad (2.7)$$

Há a necessidade de ser unido o preditor linear da equação (2.7) à média das respostas y_1, y_2, \dots, y_n , de forma a modelá-la. Assim, define-se a função de ligação (*link function*) $g(\mu_i)$. Este procedimento é exatamente aquele que unifica os modelos lineares dos não-lineares. Vieira (2004) ainda acrescenta que a função de ligação define a forma como os efeitos sistemáticos de x_1, x_2, \dots, x_k são transferidos à média.

$$\eta_i = g(\mu_i) = (\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik}) \quad (2.8)$$

Observe-se que cada termo da equação (2.8) representa uma linha na tabela (2.1). O que significa que para cada resposta haverá um preditor linear, ou seja, haverá n preditores lineares. E ainda que, no caso particular da regressão linear clássica, a função de ligação é a própria média.

$$\eta_i = g(\mu_i) = \mu_i$$

Calculando a função inversa da função de ligação obtém-se a média:

$$\mu_i = g^{-1}(\eta_i) = g^{-1}(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik}) \quad (2.9)$$

Conforme já expresso pela equação (2.1), o parâmetro θ representa a localização onde há maior probabilidade de se encontrar certo valor, ou seja, a média, e segundo Vieira (2004), a função de ligação é denominada canônica quando $\eta_i = \theta_i$.

Myers, Montgomery e Vining (2002) esclarecem que o uso da função de ligação na forma canônica contribui muito com a simplificação do método de estimação dos parâmetros,

além do cálculo dos testes de hipótese ou intervalos de confiança para estes parâmetros. No entanto, alertam para o mais fundamental: esta conveniência não implica necessariamente em qualidade de ajuste do modelo.

Vieira (2004) apresenta a tabela (2.2) onde se encontram as ligações canônicas resumidamente para as funções de distribuição de probabilidade já apresentadas acima.

Tabela (2.2)
Ligações Canônicas para os modelos lineares generalizados

Distribuição	Ligação Canônica
Normal	$\eta = \mu$
Binomial	$\eta = \ln\left(\frac{\pi}{1-\pi}\right)$
Poisson	$\eta = \ln(\mu)$

Fonte: adaptado de Vieira (2004).

Vale ressaltar que a função de ligação transforma a média das respostas y_i e não as respostas em si, isto será útil para a modelagem das próprias médias, obviamente. A partir deste desenvolvimento, a estruturação dos modelos lineares generalizados fica completa para a estimação dos parâmetros β_i . É importante ainda ressaltar que caso se necessite de uma alternativa à função de ligação - quando o resultado do modelo for insatisfatório - pode-se lançar mão, similarmente à transformação da resposta, definir uma família de funções de ligação de potência do tipo:

$$\begin{aligned}\eta &= \mu^\lambda, \lambda \neq 0 \\ \eta &= \ln \mu, \lambda = 0\end{aligned}$$

Nelder e Wedderburn (1972) apresentaram em sua proposição dos modelos lineares generalizados um método de estimação de parâmetros via maximização da função log-verossimilhança, em sua apresentação matricial, alerta-se que tanto y quanto θ são vetores como será notado adiante:

$$L = \ln l(\mathbf{y}; \boldsymbol{\theta}; \boldsymbol{\phi}) \quad (2.10)$$

Cox e Hinkley (1974) apud Dobson (1990) afirmam que a família de distribuições exponenciais possuem suficientemente a propriedade de continuidade para garantir um máximo global da função log-verossimilhança após derivá-la e igualá-la a zero, ou seja, a

família exponencial é derivável. Transformando a equação (2.10) em sua apresentação escalar e também como uma função de distribuição da família exponencial:

$$L = \sum_{i=1}^n \ln f(y_i; \theta_i; \phi) = \sum_{i=1}^n \left\{ \frac{y_i \theta_i - b(\theta_i)}{a(\phi)} + c(y_i; \phi) \right\} \quad (2.11)$$

Observe-se que na equação (2.11) o somatório resulta na soma do modelo resultante de cada uma das n linhas da tabela (2.1). De forma a maximizá-la, a função L será derivada parcialmente em relação à matriz dos parâmetros β .

$$\frac{\partial L}{\partial \beta} = \sum \frac{dL}{d\theta_i} \frac{d\theta_i}{d\eta_i} \frac{\partial \eta_i}{\partial \beta} \quad (2.12)$$

Mas na última igualdade da equação (2.11) o termo θ só ocorre na primeira parcela do somatório, e ainda nos modelos lineares generalizados o termo ϕ é constante, assim, $a(\phi)=\phi$. Ainda, em retrospectiva, para as distribuições apresentadas – Normal, binomial e Poisson – e também para toda família exponencial $db(\theta)/d\theta = \mu$. Disto resulta-se:

$$\frac{dL}{d\theta_i} = \sum_{i=1}^n \frac{1}{\phi} \left[y_i - \frac{db(\theta_i)}{d\theta_i} \right] = \sum_{i=1}^n \frac{1}{\phi} [y_i - \mu_i] \quad (2.13)$$

Recorrendo à equação (2.8), porém em formato matricial:

$$\frac{\partial \eta}{\partial \beta} = \mathbf{x}_i \quad (2.14)$$

Destacando que conforme a tabela (2.1) os \mathbf{x}_i são as variáveis regressoras, ou independentes, para cada resposta i .

As equações score são obtidas após igualar a zero os resultados das derivadas das equações (2.13) e substituindo em (2.12) as igualdades (2.13) e (2.14). O divisor de (2.13) desaparece devido ao termo ter sido igualado a zero:

$$\sum_{i=1}^n (y_i - \mu_i) \frac{d\theta_i}{\eta_i} \mathbf{x}_i = 0 \quad (2.15)$$

Se as equações score forem canônicas - $\eta_i = \theta_i$ - a igualdade acima resume-se a:

$$\sum_{i=1}^n (y_i - \mu_i) \mathbf{x}_i = \mathbf{0} \quad (2.16)$$

A solução das equações escoré (2.16) levará à estimação dos parâmetros do modelo. O apêndice A.5 de Myers, Montgomery e Vining (2002) explica em detalhes que a equação (2.16) conduz pelo desenvolvimento da primeira ordem de uma série de Taylor à equação (2.17). Dobson (1990) também apresenta uma solução similar que naturalmente conduz ao mesmo resultado. Considere-se a equação matricial, conforme Vieira (2004) e Epprecht e Vieira (2004):

$$\hat{\boldsymbol{\beta}}^{(m+1)} = (\mathbf{X}^T \mathbf{W}^{(m)} \mathbf{X})^{-1} \mathbf{X}^T \mathbf{W}^{(m)} \mathbf{z}^{(m)} \quad (2.17)$$

O índice m corresponde à m -ésima iteração, portanto $\hat{\boldsymbol{\beta}}^{(m+1)}$ deve ser entendido como a estimativa da m -ésima mais uma iteração do algoritmo. Desenvolvendo adiante a equação (2.17), seguem os significados das matrizes que a compõe em suas respectivas dimensões.

$$\hat{\boldsymbol{\beta}}^{(m+1)} = \begin{bmatrix} \beta_1^{(m)} \\ \beta_2^{(m)} \\ \vdots \\ \beta_k^{(m)} \end{bmatrix}$$

\mathbf{X} é a matriz que representa as variáveis regressoras da tabela (2.1) com acréscimo da coluna de 1's que está relacionado ao termo independente do modelo a ser gerado:

$$\mathbf{X} = \begin{bmatrix} 1 & x_{11} & \cdots & x_{1k} \\ 1 & x_{21} & \cdots & x_{2k} \\ 1 & \vdots & \ddots & \vdots \\ 1 & x_{n1} & \cdots & x_{nk} \end{bmatrix}$$

A matriz \mathbf{W} é diagonal e conhecida como a matriz dos pesos para ponderação da variância relativa a cada variável regressora.

$$\mathbf{W}^{(m)} = \begin{bmatrix} w_{11} & 0 & 0 & 0 \\ 0 & w_{22} & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & w_{nn} \end{bmatrix}$$

Os elementos da diagonal $w_{ii}^{(m)}$ são dados por:

$$w_{ii}^{(m)} = \frac{1}{\text{var}(y_i)} \left[\left(\frac{\partial \mu_i}{\partial \eta_i} \right)^{(m)} \right]^2, i = 1, 2, \dots, n \quad (2.18)$$

Finalmente \mathbf{z} é o vetor das variáveis de ajuste na m -ésima iteração.

$$\mathbf{z}^{(m)} = \begin{bmatrix} z_1^{(m)} \\ z_2^{(m)} \\ \vdots \\ z_n^{(m)} \end{bmatrix}$$

Cujos elementos são:

$$z_i^{(m)} = \hat{\eta}_i^{(m)} + (y_i - \hat{\mu}_i^{(m)}) \left(\frac{\partial \eta_i}{\partial \mu_i} \right)^{(m)}, i = 1, 2, \dots, n \quad (2.19)$$

Em destaque observa-se pelas equações (2.17) e (2.19) que o processo é iterativo, sendo que a estimativa $\hat{\beta}^{(m+1)}$ é calculada como dependente da estimativa anterior $\hat{\beta}^{(m)}$. Vieira (2004) afirma que a convergência do método iterativo dos mínimos quadrados ponderados iterativo (MQPI) depende bastante da qualidade das estimativas iniciais para os parâmetros $\hat{\beta}$.

O ciclo iterativo da figura (2.2) começa fazendo com que m indique, como dito anteriormente, a ordem da iteração. Assim, a m -ésima iteração faz uso da estimativa de $\hat{\beta}^{(m)}$ que deriva das estimativas para η , μ , \mathbf{z} e \mathbf{W} , sendo que o ciclo se fecha quando determinada a atualização da estimativa $\hat{\beta}^{(m+1)}$ que será usada na próxima iteração, e assim ciclicamente até a convergência conforme os seis passos em seguida apresentados (EPPRECHT; VIEIRA, 2004).

O algoritmo é apresentado graficamente nas figuras (2.1) e (2.2). A primeira apresenta a forma como as variáveis da equação (2.17) se inter-relacionam, enquanto que a segunda seqüência o procedimento iterativo. No apêndice A este algoritmo é executado para o caso específico que trata esta dissertação.

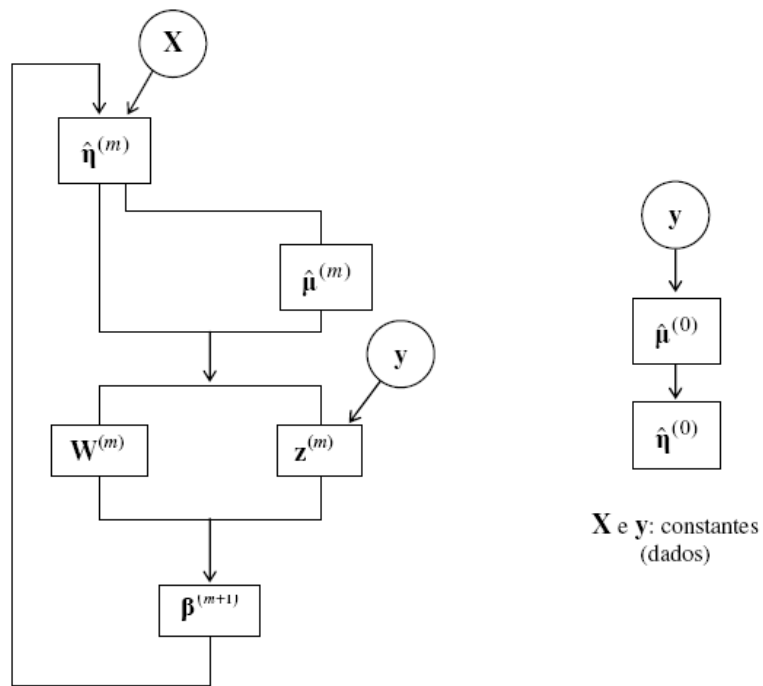


Figura (2.1) – Relações entre as variáveis do algoritmo MPQI. Fonte: Vieira (2004)

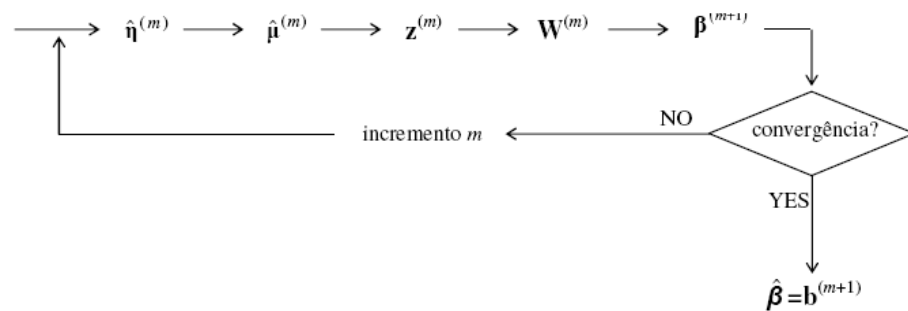


Figura (2.2) – Algoritmo MPQI. Fonte: Vieira (2004)

Passo 1: Calcular o vetor $\hat{\eta}^{(m)}$, seqüência da figura (2.2), onde k é o número de parâmetros do modelo mais um, pois o modelo possui o termo independente β_0 , pela equação (2.8):

$$\hat{\eta}^{(m)} = \begin{bmatrix} \hat{\eta}_1^{(m)} \\ \hat{\eta}_2^{(m)} \\ \vdots \\ \hat{\eta}_k^{(m)} \end{bmatrix} = \mathbf{X}\hat{\beta}^{(m)} \quad (2.20)$$

Passo 2: Calcular o vetor das médias de cada fator, onde cada $\hat{\mu}_i$ é obtido pela inversão das funções de ligação η_i :

$$\hat{\mu}^{(m)} = \begin{bmatrix} \hat{\mu}_1^{(m)} \\ \hat{\mu}_2^{(m)} \\ \vdots \\ \hat{\mu}_k^{(m)} \end{bmatrix} \quad (2.21)$$

Passo 3: Calcular o vetor \mathbf{z} segundo equação (2.19).

Passo 4: Calcular os elementos da matriz diagonal \mathbf{W} da equação (2.18).

Passo 5: Atualizar a estimativa do vetor $\hat{\beta}^{(m)}$ segundo a equação (2.17).

$$\hat{\beta}^{(m+1)} = (\mathbf{X}^T \mathbf{W}^{(m)} \mathbf{X})^{-1} \mathbf{X}^T \mathbf{W}^{(m)} \mathbf{z}^{(m)}$$

Passo 6: Testar a convergência do algoritmo pela verificação da distância definida pelo condutor do experimento entre $\hat{\beta}^{(m+1)}$ e $\hat{\beta}^{(m)}$ é menor do que o valor da tolerância especificada. Se a convergência ocorreu, então encerra-se o algoritmo, em caso contrário há de se retornar ao passo 1 para mais uma iteração.

Vieira (2004) acrescenta que não há uma única maneira de se estimar o primeiro valor dos parâmetros, $\hat{\beta}^{(1)}$. No entanto, sugere que seja aquela mais comum e a que adota, que é estimá-los a partir de uma estimativa inicial de $\hat{\mu}_0$ relacionado ao vetor das médias μ . O autor ainda chama a atenção de que este vetor das médias é solicitado pela algoritmo no passo 2, assim $\hat{\beta}^{(1)}$ é obtido a partir de uma tentativa de valores de $\hat{\mu}^{(0)}$ e prosseguindo com os demais passos e girando o ciclo até a convergência. A igualdade $\hat{\mu}^{(0)} = \mathbf{y}$ também é, segundo o referido autor, muito usada. Há ainda a necessidade de cálculo dos valores da função de ligação $\hat{\eta}_i^{(0)} = g(\hat{\mu}_i^{(0)})$ para satisfazer a solicitação do vetor η no passo 3. Esta etapa será levada já na prática apresentada no próximo capítulo 4 desta dissertação.

Uma vez já definidos os parâmetros estimados $\hat{\beta}$, deve ser analisada a significância deles, pois esta permitirá inferir sobre o quanto se pode confiar nas respostas do modelo proposto, alcançando as limitações que ele impõe. Para tanto, é usado o esquema proposto em Vieira (2004) para os testes de significância do modelo. Para a família exponencial, pode-se afirmar dos parâmetros estimados $\hat{\beta}$ a seguinte igualdade para a matriz de covariância:

$$\text{cov}(\beta) = [\mathbf{X}^T \mathbf{W} \mathbf{X}]^{-1} \sigma^2 \quad (2.22)$$

Onde \mathbf{W} é a matriz dos pesos cujos elementos foram apresentados na equação (2.18). Vale observar que no caso da estimativa de parâmetros pelo método dos mínimos quadrados, esta mesma matriz de covariância é calculada dispensando a matriz dos pesos, pois naquele modelo não há a necessidade de ponderar as variâncias das respostas, seus elementos são todos a unidade. Isto comprova que a estimação de parâmetros pelo método dos mínimos quadrados, sob esta óptica, é um caso particular do método da máxima verossimilhança. A completa demonstração da equação (2.22) pode ser obtida em Myers, Montgomery e Vining (2002, p.50). Se for de interesse encontrar a demonstração da mesma fórmula para o caso da regressão linear normal recorre-se a Myers, Montgomery e Vining (2002, p.15), que é o mesmo que substituir a matriz \mathbf{W} de (2.18) pela identidade.

Nos testes de significância para os modelos lineares com distribuição normal, usa-se a estatística *t-student* para a seleção dos parâmetros:

$$t_0 = \frac{\hat{\beta}_j}{\sqrt{\text{var}(\hat{\beta}_j)}} \quad (2.23)$$

No entanto, os modelos lineares generalizados não apenas solucionam as respostas que atendem à distribuição normal, para a qual se ajusta bem a estatística da equação (2.23), mas também com toda a família exponencial já apresentada. A equação (2.23) pode ser usada em outras distribuições que não a normal desde que esta estatística resulte em um valor alto, diga-se três, para se indicar a significância de parâmetros, no entanto ainda assim será uma aproximação, Vieira (2004).

Recomenda-se também o uso da função de desvio (*deviance*) para os testes de significância dos coeficientes do modelo. Sua definição, conforme Vieira (2004), é pelo cálculo do desvio entre o modelo proposto em relação ao modelo saturado, ou seja, aquele no

qual os valores ajustados $\hat{\mu}_i$ são iguais às respostas y_i , coletadas no experimento ou onde o número de parâmetros é igual ao número de observações.

$$D = -2 \ln \left(\frac{L_{\text{mod}}}{L_{\text{sat}}} \right) \quad (2.24)$$

Dobson (1990) ressalta que valores elevados de *deviance* sugerem modelos com descrições pobres dos dados. Sendo numerador e denominador da equação (2.24) as funções de máxima verossimilhança do modelo proposto e do modelo saturado, respectivamente. Sendo assim, pode-se desenvolvê-la para:

$$D(y_i; \hat{\mu}_i) = -2 \ln \left(\frac{L(y_i; \hat{\mu}_i)}{L(y_i; y_i)} \right) = -2 [\ln(y_i; \hat{\mu}_i) - \ln(y_i; y_i)] \quad (2.25)$$

A tabela (2.3) apresenta as funções de desvio para as distribuições normal, binomial e de Poisson:

Tabela (2.3)
Funções de Desvio

Distribuição de Probabilidade	Função de desvio (<i>deviance</i>)
Normal	$\sum_{i=1}^n (y_i - \hat{\mu}_i)^2$
Binomial	$2 \cdot \sum_{i=1}^n \left\{ y_i \cdot \ln \left(\frac{y_i}{\hat{\mu}_i} \right) + (m - y_i) \cdot \ln \left[\frac{m - y_i}{m - \hat{\mu}_i} \right] \right\}$
Poisson	$2 \cdot \sum_{i=1}^n \left[y_i \cdot \ln \left(\frac{y_i}{\hat{\mu}_i} \right) - (y_i - \hat{\mu}_i) \right]$

Fonte: adaptado Vieira (2004).

McCullagh e Nelder (1990) indicam que o termo m da função de desvio para a distribuição binomial refere-se ao número de indivíduos cujas respostas foram coletadas para o experimento.

Vieira (2004) chama a atenção para se notar que a *deviance* da função normal é a própria soma dos quadrados dos resíduos donde deriva-se o princípio de cálculo do método dos mínimos quadrados.

Lindsey (1997) apud Vieira (2004) demonstra que $D(y_i; \hat{\mu}_i)$ tem distribuição que obedece assintoticamente à distribuição do qui-quadrado - χ^2 - com $n-p$ graus de liberdade,

onde n é o número de observações e p é a quantidade de parâmetros adotados no modelo. Novamente, Lindsey (1997) apud Vieira (2004), recomenda a análise de desvios conforme a diferença deles ou diferença de *deviance*, como exposto a seguir.

Suponha o modelo A como saturado, ou seja, aquele no qual os valores ajustados $\hat{\mu}_i$ são iguais às respostas y_i ; outro modelo B com $p+1$ parâmetros $\beta_0, \beta_1, \dots, \beta_{p-1}, \beta_p$. Considere-se ainda um último modelo C, aninhado do modelo B, porém com um parâmetro a menos, explicitamente, $\beta_0, \beta_1, \dots, \beta_{p-1}$. Pela equação (2.25) chega-se a cálculo das *deviances* dos modelos B e C.

$$D_B(y_i; \hat{\mu}_i^{(B)}) = -2[\ln L_B(y_i; \hat{\mu}_i^{(B)}) - \ln L_A(y_i; y_i)]$$

$$D_C(y_i; \hat{\mu}_i^{(C)}) = -2[\ln L_C(y_i; \hat{\mu}_i^{(C)}) - \ln L_A(y_i; y_i)]$$

E a diferença entre as *deviances* C e B:

$$D_C(y_i; \hat{\mu}_i^{(C)}) - D_B(y_i; \hat{\mu}_i^{(B)}) = -2[\ln L_C(y_i; \hat{\mu}_i^{(C)})] + 2[\ln L_B(y_i; \hat{\mu}_i^{(B)})]$$

$$D_C(y_i; \hat{\mu}_i^{(C)}) - D_B(y_i; \hat{\mu}_i^{(B)}) = -2 \ln \left[\frac{L_C(y_i; \hat{\mu}_i^{(C)})}{L_B(y_i; \hat{\mu}_i^{(B)})} \right]$$

Sendo que Lindsey (1997) apud Vieira (2004) demonstra que esta diferença segue aproximadamente a distribuição χ_1^2 quando o modelo C está correto, assim como $D_C(y_i; \hat{\mu}_i^{(C)})$ segue χ_{n-p}^2 no mesmo caso.

Vieira (2004) conclui que sendo correto o modelo C nos casos acima citados, a razão abaixo segue a distribuição $F_{1, n-p}$ por aproximação:

$$F_0 = \frac{\frac{(D_C(y_i; \hat{\mu}_i) - D_B(y_i; \hat{\mu}_i))}{D_C(y_i; \hat{\mu}_i)}}{n-p} \approx F_{1, n-p} \quad (2.26)$$

Desta forma, para uma seqüência de k modelos aninhados pode-se calcular as *deviances* $D_j(y_i; \hat{\mu}_i)$ para $j=1, 2, \dots, k$. e montar os testes de significância dos parâmetros numa

tabela ANODE (*Analysis of deviance*), tal qual na tabela ANOVA para variâncias comumente utilizada na estimação de fatores significativos de um experimento.

O parâmetro de dispersão (ϕ) referido na equação (2.1) que define o formato geral de um modelo linear generalizado tem no caso particular da regressão linear a seguinte apresentação (MONTGOMERY, 2005):

$$\hat{\sigma}^2 = \frac{SS_E}{n - p} \quad (2.27)$$

Onde SS_E é a soma dos quadrados dos resíduos. Atkinson e Riani (2000) apud Vieira (2004) asseguram que para o caso geral dos modelos lineares generalizados este parâmetro pode ser calculado da seguinte forma:

$$\hat{\sigma}^2 = \hat{\phi} = \frac{D(y_i; \hat{\mu}_i)}{n - p} \quad (2.28)$$

Da mesma forma que na regressão linear clássica, nos modelos lineares generalizados a adequação do modelo é checada pela análise dos resíduos. Sendo que o resíduo *deviance* apresentado é aquele que melhor se aproxima da distribuição normal e sua versão *studentizada* é a recomendável para se confirmar a adequação do modelo gerado com base na família exponencial que são descritos a seguir (VIEIRA, 2004).

A cada resposta y_i obtida no experimento pode-se calcular o seu respectivo *deviance* $d_i = D(y_i; \hat{\mu}_i)$. Pela soma de todas as i respostas medidas é possível conhecer a discrepância do modelo, assim:

$$\sum_{i=1}^n d_i = D(y_i; \hat{\mu}_i) \quad (2.28)$$

A partir da equação (2.28) pode-se definir o resíduo *deviance* associado a cada resposta y_i obtida:

$$r_{Di} = \text{sin}al(y_i - \hat{\mu}_i) \sqrt{d_i} \quad (2.29)$$

A *studentização* das *deviances* são apresentados por McCullagh e Nelder (1990) pela matriz chapéu. Esta matriz mapeia o vetor dos valores observados em um vetor dos valores

ajustados, segundo Myers, Montgomery e Vining (2002), suas propriedades representam um papel central na análise de regressão, sua forma mais geral é:

$$\mathbf{H} = \mathbf{W}^{\frac{1}{2}} \mathbf{X} (\mathbf{X}^T \mathbf{W} \mathbf{X})^{-1} \mathbf{X}^T \mathbf{W}^{\frac{1}{2}} \quad (2.30)$$

Relembrando que a matriz \mathbf{W} é a matriz diagonal cujos elementos são dados pela equação (2.18).

$$w_{ii}^{(m)} = \frac{1}{\text{var}(y_i)} \left[\left(\frac{\partial \mu_i}{\partial \eta_i} \right)^{(m)} \right]^2, i = 1, 2, \dots, n$$

O i -ésimo elemento da diagonal da matriz \mathbf{H} cujos elementos são h_{ii} permite o cálculo dos resíduos studentizados:

$$r_i = \frac{r_{Di}}{\sqrt{\hat{\phi}^2 (1 - h_{ii})}} \quad (2.31)$$

Enquanto que a estimativa do parâmetro de dispersão ($\hat{\phi}$) é calculado conforme a equação (2.28). Vieira (2004) também indica o gráfico de probabilidade normal dos resíduos *deviance* studentizados versus distribuição destes no eixo horizontal dos percentis de forma a possibilitar a detecção de observações atípicas, conforme figura (2.3).

No caso da figura (2.3) não se observam pontos seguramente fora do envelope, ou seja, das linhas laterais que acompanham a tendência de orientação das *deviances* studentizadas à medida que avançam pelos percentis. E ainda, muitos pontos fora do alinhamento. Segundo Vieira (2004) estes são indicativos suficientes para aprovar o modelo que gerou a figura (2.3).

Da mesma forma, a análise da função de ligação também deve ser feita. Para tanto, gera-se o gráfico da figura (2.4) onde os resíduos studentizados são plotados segundo o seu valor ajustado.

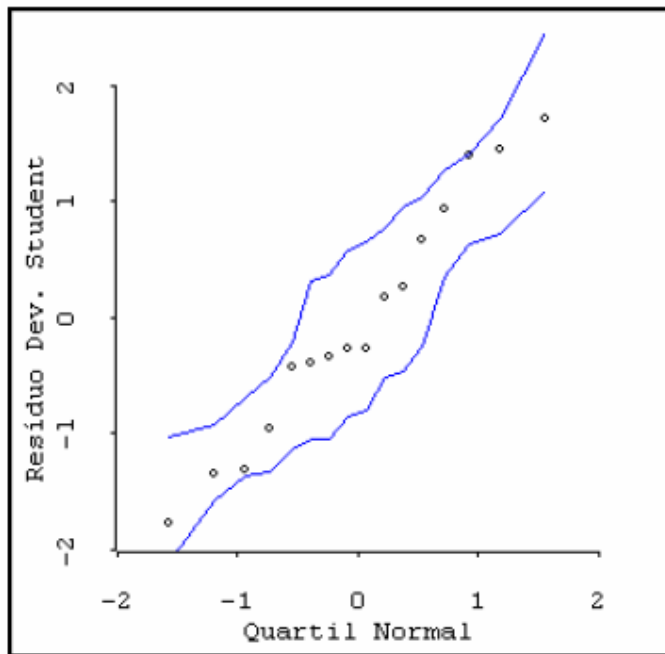


Figura (2.3) – Gráfico de probabilidade normal dos resíduos *deviance* studentizados versus percentis. Fonte: adaptado Vieira (2004)

Na figura (2.4), cujo gráfico foi gerado a partir do mesmo modelo referente à figura (2.5), percebe que os resíduos studentizados distribuem-se aleatoriamente, sem um padrão definido claramente perceptível. Neste caso apresentado, o *software* estatístico que gerou tal gráfico foi o *ARC*. Este e ainda outros *softwares* geram uma linha de amortecimento conhecida por *lowess*. No caso da figura (2.4) esta linha é quase horizontal e próxima do eixo das abscissas, sendo este o indicativo que aprova a função de ligação selecionada (VIEIRA, 2004).

Como já discutido, a função de variância é alternativamente à função de ligação obtida via função de potência da média:

$$\text{var}(\mu) = \mu^\lambda$$

Na figura (2.4) a linha de amortecimento, ou *lowess*, decresce suavemente num trecho e cresce numa tendência óbvia em outro. O crescimento da esquerda para a direita indica a necessidade de se usar um valor maior para λ do que o utilizado no modelo. O caso contrário, quando a linha de amortecimento decresce o valor utilizado de λ de ser menor. Para

o gráfico em questão a linha *lowess* o crescimento se verifica aproximadamente no 14º resíduo plotado, no entanto quem “puxa” a linha para o alto são os dois maiores resíduos, em valor absoluto, numa região de menor densidade de pontos, ou seja, a confiabilidade na linha de amortecimento é menor e assim sujeito a se poder afirmar que não se deve considerar este fenômeno como indicativo de uma função de variância incorreta (VIEIRA, 2004).

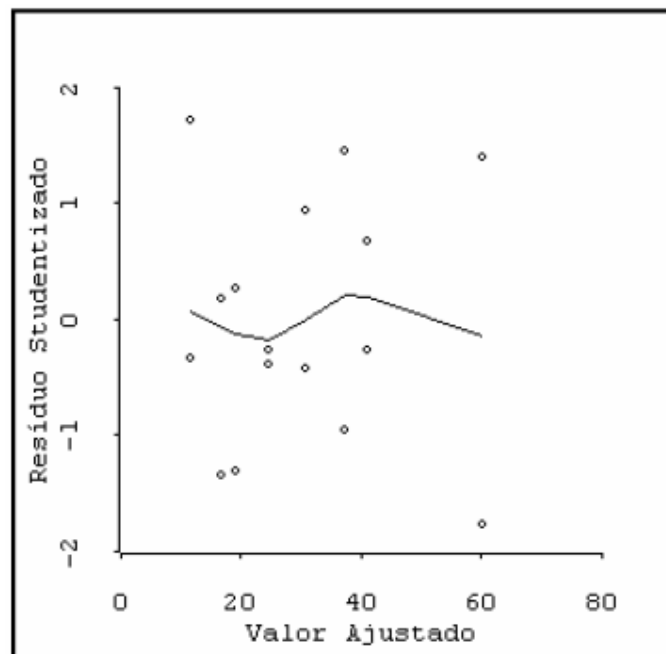


Figura (2.4) – Resíduo studentizado pelo valor ajustado. Fonte: adaptado Vieira (2004)

Para verificação da validade da função de variância, mais uma vez recorre-se aos gráficos. Nesta situação o valor absoluto dos resíduos studentizados é plotado com os seus respectivos valores ajustados. Na figura (2.5) este gráfico é apresentado.

Conforme Myers, Montgomery e Vining (2002) a diagonal da matriz chapéu H identifica pontos que são potencialmente influentes devido à sua posição no campo das respostas, assim tornar-se conveniente considerar ponderadamente tanto sua localização quanto a variável de resposta que meça sua influência. Assim, a distância de Cook mede a distância quadrada entre os mínimos quadrados de todos os n pontos da matriz de parâmetros estimados $\hat{\beta}$.

$$D_i = \frac{r_{pi}^2 \cdot h_{ii}}{p \times \phi \times (1 - h_{ii})^2} \quad (2.41)$$

Onde p é o número de parâmetros estimados e r_{pi}^2 é o resíduo de Pearson studentizado, da seguinte forma obtido conforme apresentado em Vieira (2004):

$$r'_{pi} = \frac{y_i - \hat{\mu}_i}{\sqrt{\phi \times \text{var}(\mu_i \cdot (1 - h_{ii}))}}$$

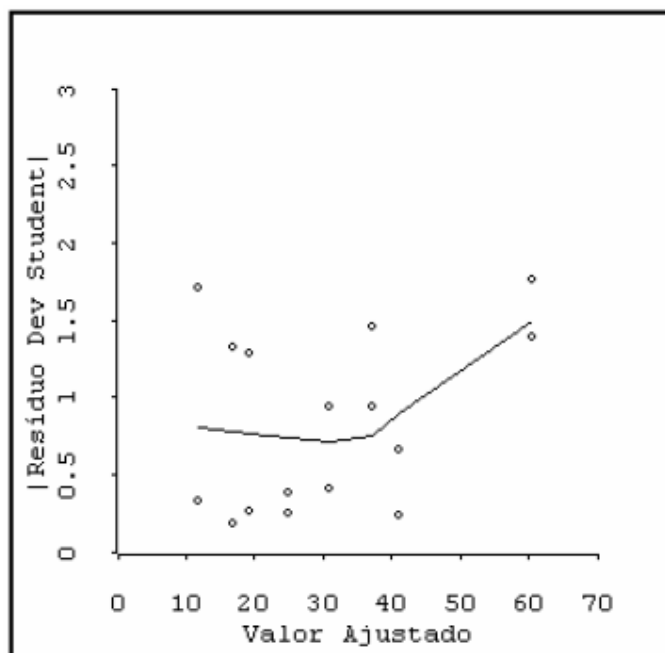


Figura (2.5) – Gráfico dos resíduos studentizados (absolutos) versus valores ajustado. Fonte: adaptado Vieira (2004)

A figura (2.6) apresenta o gráfico da distância de Cook, onde pelo comportamento disperso dos pontos entende-se que não há observação consideravelmente influente no campo das respostas.

Desde quando se iniciou o tratamento dos modelos lineares generalizados neste texto, todos os progressos foram feitos em torno da família de distribuições dita exponencial, ou seja, aquelas distribuições que podem ser escritas no formato da equação (2.1). Myers,

Montgomery e Vining (2002) lembram que nos casos dos modelos lineares generalizados a condição de independência das respostas foi aceita para permitir a modelagem pela via da função de máxima verossimilhança. No entanto, há uma vasta coleção de processos industriais onde ocorrem duas situações:

- a) as respostas são independentes mas não obedecem a um dos membros da família exponencial, ainda que a variância seja função da média e;
- b) as respostas têm variância que são função da média, no entanto estas respostas são correlacionadas, ou seja, são dependentes.

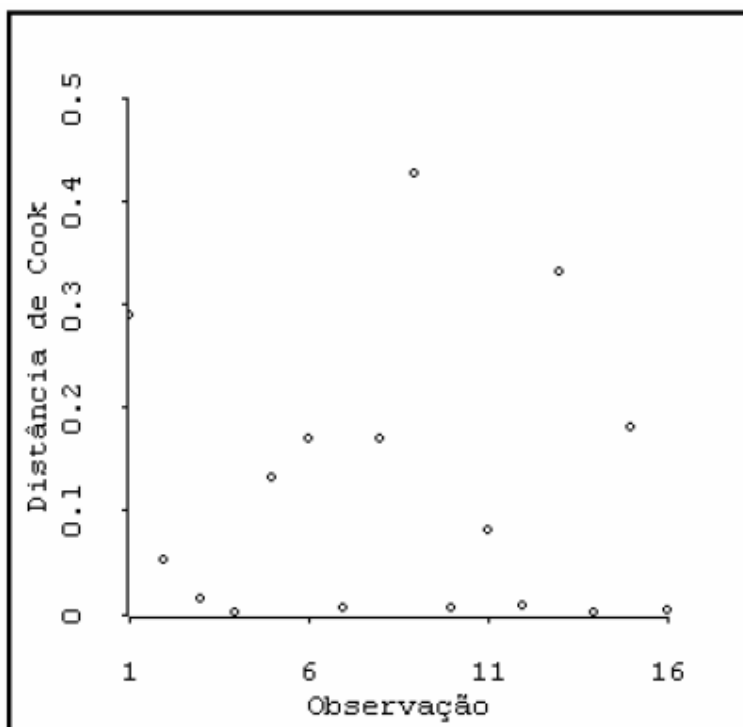


Figura (2.6) – Distância de Cook. Fonte: adaptado Vieira (2004)

Assim, segundo estes mesmos autores, a função de quase-verossimilhança deriva do conceito dos mínimos quadrados ponderados, ou mais genericamente, mínimos quadrados generalizados para o caso onde as respostas são correlacionadas, sendo que a função de quase-verossimilhança tem propriedades da função log-verossimilhança vista na equação (2.10).

Vieira (2004) diz que a idéia é usar o método dos mínimos quadrados ponderados pela variância, conforme abaixo.

$$SQP = \sum_{i=1}^n \frac{(y_i - \mu_i)^2}{\text{var}(y_i)} = \sum_{i=1}^n \frac{(y_i - \mu_i)^2}{\phi \cdot V(\mu_i)} \quad (2.42)$$

Primeiramente, deriva-se a equação (2.42) em relação aos coeficientes:

$$\frac{dSQP}{d\beta} = \sum_{i=1}^n \frac{(y_i - \mu_i)^2}{\phi \cdot V(\mu_i)} \frac{d\mu}{d\beta} \quad (2.43)$$

Igualando estas as equações (2.43) a zero, no processo clássico para se achar extremos de uma função que neste caso é um máximo, encontram-se as equações-escore cujas soluções conduzem à estimação dos coeficientes:

$$\sum_{i=1}^n \frac{(y_i - \mu_i)^2}{\phi \cdot V(\mu_i)} \frac{d\mu}{d\beta} = 0 \quad (2.44)$$

Vieira (2004) apresenta a função de quase-verossimilhança para uma observação:

$$Q_i(\mu_i; y_i) = \int_{y_i}^{\mu_i} \frac{y_i - t}{\phi \cdot V(t)} dt \quad (2.45)$$

A equação (2.45) torna-se para n observações da seguinte forma:

$$Q(\mu; \mathbf{y}) = \sum_{i=1}^n Q_i(\mu_i; y_i) = \sum_{i=1}^n \int_{y_i}^{\mu_i} \frac{y_i - t}{\phi \cdot V(t)} dt \quad (2.46)$$

Sendo que as derivadas da equação (2.46) em relação aos coeficientes serão as próprias equações-escore apresentadas em (2.44). Por isso, Vieira (2004) diz que a função de quase-verossimilhança corresponde à função de log-verossimilhança, onde a diferença cabível está em que no uso da função de quase-verossimilhança com o intuito da estimação de coeficientes, apenas se define a relação da variância de resposta com sua média, sendo desnecessária a definição de uma função de probabilidade e, naturalmente, tampouco uma distribuição da família exponencial. O autor ainda diz que, particularizando $\partial \mu_i / \partial \theta_i = V(\mu_i)$, volta-se às equações-escore da função log-verossimilhança, ou o conjunto de equações (2.15).

Tomando-se o termo do somatório da equação (2.42), ou:

$$U_i = \frac{y_i - \mu_i}{\phi \cdot V(\mu_i)}$$

Serão encontradas as seguintes propriedades para a média e variância:

$$\begin{aligned} E(U_i) &= E\left[\frac{y_i - \mu_i}{\phi \cdot V(\mu_i)}\right] = 0 \\ \text{var}(U_i) &= \text{var}\left[\frac{y_i - \mu_i}{\phi \cdot V(\mu_i)}\right] = \frac{\text{var}(y_i)}{[\phi \cdot V(\mu_i)]^2} = \frac{\phi \cdot V(\mu_i)}{[\phi \cdot V(\mu_i)]^2} = \\ &= \frac{1}{\phi \cdot V(\mu_i)} - E\left(\frac{\partial U_i}{\partial \mu_i}\right) = \frac{1}{\phi \cdot V(\mu_i)} \end{aligned}$$

Estas também são propriedades das equações-escore da função de máxima log-verossimilhança.

Como já dito anteriormente, função de variância pode ser expressa por uma função de potência com alternativa à função de ligação, no seguinte formato:

$$V(\mu_i) = \mu_i^t$$

Para diferentes valores de t inseridos na equação (2.45) da função de quase-verossimilhança para uma observação tem-se.

Para $t=0$:

$$Q_i(\mu_i; y_i) = -\frac{(y_i - \mu_i)^2}{2}$$

Para $t=1$ a função de quase-verossimilhança é:

$$Q_i(\mu_i; y_i) = y_i \ln(\mu_i) - \mu_i$$

Para $t=2$ a mesma função é:

$$Q_i(\mu_i; y_i) = -\frac{y_i}{\mu_i} - \ln(\mu_i)$$

Para $t=3$ a referida função é:

$$Q_i(\mu_i; y_i) = -\frac{y_i}{2\mu_i^2} + \frac{1}{\mu_i}$$

Quando $t \neq 0$, $t \neq 1$ e $t \neq 2$, então a função de quase-verossimilhança é:

$$Q_i(\mu_i; y_i) = \mu_i^{-t} \left(\frac{\mu_i y_i}{1-t} - \frac{\mu_i^2}{2-t} \right)$$

E por fim, para o caso da variância assumir o comportamento da função $V(\mu_i) = \mu(1 - \mu)$ a quase-verossimilhança será:

$$Q_i(\mu_i; y_i) = y \ln \left(\frac{\mu_i}{1 - \mu_i} \right) + \ln(1 - \mu_i)$$

Para efeito de simplificação de cálculos, Vieira (2004) diz que quando t é igual a 0, 1, 2 e 3 as funções de quase-verossimilhança se identificam com as funções da família exponencial convertidas à função de log-verossimilhança nas distribuições normal, de Poisson, gama e normal inversa, respectivamente. E ainda, no caso $V(\mu_i) = \mu(1 - \mu)$ que é função da variância da distribuição binomial, a função de quase-verossimilhança se identifica com a função de log-verossimilhança da própria binomial.

Uma vez que a maximização da função de quase-verossimilhança é capaz de resultar nas mesmas estimativas dos parâmetros $\hat{\beta}$ conforme visto atrás, os seis passos indicados no algoritmo utilizado para encontrá-los no caso da função de log-verossimilhança também podem ser adotados nesta situação (VIEIRA, 2004).

Para os testes de significância dos parâmetros encontrados usa-se a estatística de quase-*deviance*. Conforme explica Vieira (2004) esta estatística baseia-se nos mesmos termos da *deviance* para os modelos lineares generalizados. Assim, da mesma forma esta estatística fornece para uma modelo qualquer de quase-verossimilhança o desvio deste modelo em relação a ele mesmo porém saturado, pois é assim definida:

$$D_i(y_i; \hat{\mu}_i) = -2\phi[Q_i(y_i; \hat{\mu}_i) - Q_i(y_i; y_i)] = -2\phi[Q_i(y_i; \hat{\mu}_i)]$$

E desenvolvendo:

$$D_i(y_i; \hat{\mu}_i) = 2 \int_{\mu_i}^{y_i} \frac{y_i - \hat{\mu}_i}{V(\hat{\mu}_i)} \quad (2.47)$$

O termo $Q_i(y_i; \hat{\mu}_i)$ é a função de máxima quase-verossimilhança do modelo encontrado pelo algoritmo. Ao passo que o termo $Q_i(y_i; y_i)$ é a função de máxima quase-verossimilhança do modelo saturado, onde os valores ajustados $\hat{\mu}_i$ são iguais às respostas encontradas no experimento, ou seja, y_i .

Nas funções de máxima quase-verossimilhança os resíduos, tal qual nos modelos lineares generalizados ordinários, representam papel fundamental para análise de adequação do modelo obtido. Lee e Nelder apud Vieira (2004) recomendam o uso do resíduo *deviance* studentizado já apresentado na equação (2.31).

A figura (2.7) representa o gráfico dos resíduos *deviance* studentizados num modelo desenvolvido a partir do método da máxima quase-verossimilhança e onde a função de variância assumiu a forma $V(\mu) = \mu$. Este gráfico também foi gerado pelo software *ARC*.

No gráfico da esquerda não se percebe um padrão de distribuição nos resíduos, são desestruturados e aleatórios. E por este motivo a linha de amortecimento (*lowess*) comporta-se horizontalmente e próxima da abscissa. Este é segundo Vieira (2004) um indicativo de adequação da função de ligação.

Já no gráfico da direita a linha de amortecimento tem um óbvio crescimento da esquerda para a direita denunciando que o expoente da média na função de variância, no caso $V(\mu) = \mu$, deve ser aumentado. Na situação ‘real’, aspas de Vieira (2004), μ deveria ser elevado ao quadrado.

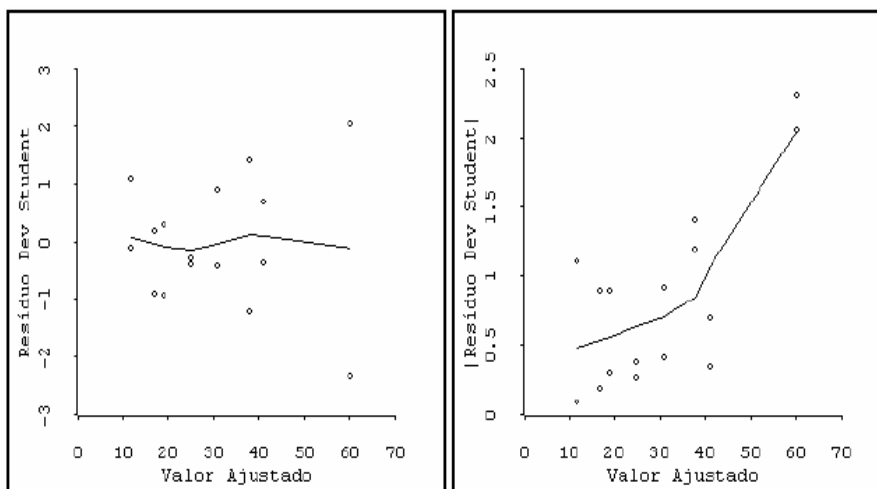


Figura (2.7) – Gráfico dos resíduos *deviance* studentizados pelos valores ajustados (esquerda) e gráfico dos resíduos absolutos pelos valores ajustados. Fonte: adaptado Vieira (2004)

2.2 O ÍNDICE DE CAPACIDADE DE PROCESSOS

A performance ou capacidade de um processo pode ser medida de diferentes formas. O significado desta capacidade está relacionada ao cumprimento de determinada promessa de qualidade ou à confiança que se pode depositar no processo de que cumprirá com o esperado pelo cliente. Bothe (1997) diz que a Ford e a Motorola na década de 1990 demonstraram a forte correlação entre a lucratividade e a satisfação do consumidor. Acrescenta-se ainda que a medição da capacidade de processos é um dos melhores métodos para quantificar quão bem os requisitos do consumidor estão sendo atendidos.

Em geral, a característica de qualidade é expressa em termos numéricos e é chamada variável de controle. As cartas de controle para variáveis são muito úteis para avaliar se um processo é capaz ou não de cumprir com certa promessa de qualidade, e foram uma das primeiras tentativas exitosas para tal fim. No entanto, Bothe (1997) afirma que sua principal função é identificar se o processo está sob controle, o que indica a presença de apenas causas comuns que afetam a variabilidade das medidas. Se aparecerem causas especiais, a configuração da carta de controle assumirá um padrão de comportamento que indicam suas presenças. Estas causas devem ser eliminadas de forma a retomar-se a estabilidade do processo.

Montgomery (2001) apresenta diversos tipos de carta de controle para variáveis, sendo as mais comuns: \bar{x} , R e S, respectivamente para os valores médios, a amplitude da variável e a variância. É necessário, portanto, que se mantenha controle primordialmente sobre a tendência central e a dispersão das medidas de interesse.

As cartas de controle são construídas baseando-se, na média e variância da variável de resposta. Embora muito úteis, em especial para averiguar a estabilidade do processo (BOTHE, 1997), as cartas de controle, sozinhas, contém algumas desvantagens para avaliar sua capacidade: exige análise do posicionamento de cada medida em relação aos seus vizinhos e em relação aos limites de controle para sua interpretação; não indica imediatamente a perda de qualidade, ainda que mínima, a cada nova medida; sua interpretação depende da análise de gráficos nem sempre disponíveis; e por fim, não permite a informação imediata se a partir de um determinado arranjo de fatores o processo é otimizado ou não, que é onde reside o interesse deste trabalho.

Bothe (1997) diz que um processo sob controle do ponto de vista estatístico apresentará 99,73% das suas medidas dentro dos limites de controle superior e inferior. Como na carta de controle cada ponto significa as medidas de um subgrupo de amostras, em apenas 0,27% dos casos – 100% - 99,73% - cairão fora destes limites. Neste caso, continua Bothe (1997), não se poderá mais afirmar que o processo está sob controle e certamente ocorreu uma causa especial que o desestabilizou.

Complementarmente às cartas de controle, surge um outro caminho para expressar a capacidade de processo em termos do índice de capacidade de processos (ICP). De certa maneira, estes índices suprem as carências das cartas de controle. Um dos mais comuns índices de capacidade de processo - C_p - é definido pela equação (2.48), que para uma característica de qualidade com limites superior (USL) e inferior (LSL) é expressa por:

$$C_p = \frac{USL - LSL}{6\sigma} \quad (2.48)$$

Neste texto será preferida a designação USL e LSL originária da literatura pesquisada em inglês que significam respectivamente, *upper specification limit* e *lower specification limit*. Portanto, USL é o limite de especificação superior, enquanto LSL é o limite de especificação inferior. O termo σ é o desvio-padrão da população da medida, como na maioria das situações o desvio-padrão populacional é desconhecido, Montgomery (2001)

frequentemente usa a estimativa amostral do desvio-padrão ($\hat{\sigma}$) para populações, o que conduzirá à estimativa \hat{C}_p :

$$\hat{\sigma} = \frac{\bar{R}}{d_2} \quad (2.49)$$

Onde o termo \bar{R} é a amplitude global de cada amostra segundo a equação (2.50):

$$\bar{R} = \frac{R_1 + R_2 + \dots + R_m}{m} \quad (2.50)$$

Nesta equação o termo m corresponde à quantidade de amostras selecionadas, Montgomery (2001) recomenda no mínimo 20 a 25 delas, cada uma contendo de 4 a 6 observações. O termo R_i , ou amplitude máxima da amostra, representa a diferença entre o maior e o menor valor de cada uma das m amostras, e é dada pela equação (2.51):

$$R = x_{m\acute{a}ximo} - x_{m\acute{i}nimo} \quad (2.51)$$

Ainda na equação (2.49) o termo d_2 no denominador é o estimador não-tendencioso, ou imparcial, para distribuições aproximadamente normais e é tabelado em função do tamanho da amostra. A referida tabela encontra-se em Montgomery (2001, p. 761). O algarismo 6 (seis) no denominador da mesma equação é proveniente de se arbitrar, por definição, a distância da média geral de três vezes o desvio-padrão em relação a cada um dos limites de controle, o superior e o inferior.

Com efeito, a equação (2.48) representa ‘quantas vezes cabem’ dentro de ambos os limites de especificação, seis vezes o desvio-padrão das respostas do processo. Diversos renomados estatísticos industriais adotam esta definição arbitrada em seis desvios-padrão por representar performances de produção realísticas e alcançáveis, e também por serem mais populares (BOTHE, 1997; MONTGOMERY, 2001).

Uma distinção importante é ressaltada por Bothe (1997) entre os valores de referência da equação (2.48) e os valores dos limites de referência das cartas de controle, pois frequentemente são confundidos, ainda que não tenham entre si nenhuma relação matemática. Os primeiros são limites de especificação definidos pela engenharia do produto e servem para distinguir entre peças conforme e não-conforme. Uma vez que a medida em questão está no intervalo entre USL e LSL o produto é aceito; caso contrário, acima de USL ou abaixo de

LSL o produto é classificado como não-conforme. Já no caso das cartas de controle, os limites de controle auxiliam os operadores a identificar se o processo está ou não estável, ou sob controle, ajudando-os a reconhecer a presença ou não de causas especiais de variação o que exigiria a ação de eliminá-las. Assim, a medida pode estar sob controle, ou seja apenas causas comuns estão agindo para sua variação, e não atender aos requisitos que indicam conformidade ou não-conformidade.

O mesmo autor ainda alerta que não tem sentido avaliar o índice de capacidade de processos se o processo não estiver estável. Desta forma, para se buscar um incremento na qualidade de um produto por intermédio da resposta de uma determinada medida, indicam-se três caminhos que envolvem o uso conjunto das cartas de controle e índice de capacidade de processos:

- a) estabilizar o processo: cartas de controle identificam quando causas especiais agem no processo desestabilizando-o. Elas fornecem pistas para reconhecê-las e eliminá-las;
- b) melhorar o processo: uma vez que o processo está sob controle, mudanças podem ser testadas e suas alterações no processo medidas. Se a mudança gerar uma resposta desejável ela pode ser incorporada ao processo produtivo permanentemente;
- c) medir a capacidade do processo: novamente, caso o processo esteja sob controle, medidas confiáveis da capacidade de processos podem ser quantificadas de forma a medir-se quão bem o processo alcança as expectativas do cliente.

A tabela (2.4) indica valores para de índices de capacidade de processos para o caso de se trabalhar com amostras quando se desconhece o desvio-padrão da população – de acordo com a característica do processo.

A equação (2.48) avalia exclusivamente se o processo apresenta, em relação aos limites de especificação, baixa ou alta variabilidade. Logicamente, é desejável que um processo tenha baixa variabilidade, o que implica em valores maiores do índice C_p . Portanto, menor variabilidade implica maior C_p . No entanto, a equação (2.48) não faz nenhuma referência se o valor está centrado ou não no alvo desejado. Assim, é proposto um novo índice de capacidade de processo que contempla não apenas a variabilidade do processo, mas

também cujo índice é influenciado pela localização da média μ em relação os limites de controle. Este índice foi designado por C_{pk} :

$$C_{pk} = \text{mínimo} \left(\frac{USL - \mu}{3\sigma}, \frac{\mu - LSL}{3\sigma} \right) \quad (2.52)$$

Tabela (2.4)
Valores mínimos recomendados para índices de capacidade de processos

Característica do processo	Valor recomendado
Processos já existentes	1,33
Novos processos	1,50
Processos existentes, porém que envolvem segurança, resistência ou parâmetro críticos	1,50
Novos processos, porém que envolvem segurança, resistência ou parâmetro críticos	1,67

Fonte: adaptado Montgomery (2001, pg. 361).

Montgomery (2001) observa que se o processo estiver centrado, ou seja, a média μ localiza-se exatamente no ponto médio entre USL e LSL, C_p será igual a C_{pk} . Desta forma, enquanto o processo não está centrado neste ponto médio, C_{pk} será sempre menor que C_p , mas potencialmente ambos poderão igualar-se. Daí que comumente o primeiro é chamado capacidade potencial, ao passo que o segundo é conhecido por capacidade atual. Outra observação importante decorrida da equação (2.52) é que caso a média do processo coincida com qualquer um dos limites isto implicará em um C_{pk} igual a zero, ainda que o índice C_p seja muito elevado a capacidade deste processo será muito baixa devido à grande descentralização de sua média.

O caso aplicável para a equação (2.52) é aquele onde há dois limites de especificação, ou seja, quando deseja-se atingir um alvo entre eles. No entanto, há outros dois casos, chamados unilaterais, onde é necessário que (i) a resposta seja tão maior quanto possível (ou maior é melhor), ou (ii) seja tão menor quanto possível (menor é melhor). Assim, a mesma equação (2.52) assume os seguintes formatos, respectivamente:

$$C_{pki} = \frac{\mu - LSL}{3\sigma} \quad (2.53)$$

$$C_{pks} = \frac{USL - \mu}{3\sigma}$$

A equação (2.52) exige o conhecimento da média da população. Caso esta não seja conhecida, utiliza-se a média estimada por uma amostra aleatória desta população. O mesmo vale para as equações (2.53). Portanto, a rigor, a referida expressão assume o seguinte formato:

$$\hat{C}_{pk} = \text{mínimo} \left(\frac{USL - \hat{\mu}}{3\hat{\sigma}}; \frac{\hat{\mu} - LSL}{3\hat{\sigma}} \right) \quad (2.54)$$

Chang, Cheng e Spiring (1988) contribuem com a apresentação da unificação das equações (2.48) e (2.52) para o caso bastante comum do alvo localizar-se exatamente sobre a mediana do intervalo entre os dois limites de especificação. Considere-se o índice k como:

$$k = \frac{2 \times |T - \mu|}{(USL - LSL)} \quad (2.55)$$

onde T é o alvo do processo. Portanto, o valor k será sempre um valor positivo ou nulo. Desta forma, estes autores relacionam:

$$C_{pk} = (1 - k) \times C_p \quad (2.56)$$

Para os casos onde não se conhecem os valores populacionais de C_{pk} e C_p , continuam valendo os valores estimados para a equação (2.56), o que exige, pelo rigor da notação estatística, o acréscimo do ‘chapéu’ (^) sobre as variáveis.

No mesmo texto, Chang, Cheng e Spiring (1988) apresentam uma nova proposta de índice de capacidade de processos que leva em consideração não apenas a média das respostas e sua variância, mas também o alvo desejado. O novo índice foi batizado C_{pm} . Segue seu desenvolvimento, considere-se:

$$C_{pm} = \frac{USL - LSL}{6\sigma'}$$

Onde:

$$\sigma' = \sqrt{E(X - T)^2}$$

O valor X é a variável aleatória que atende as respostas do processo. O operador $E(X-T)$ é a esperança matemática ou valor esperado definido em Duncan (1974) como:

$$E(x) = \int_{-\infty}^{\infty} xf(x)dx$$

A função $f(x)$ é a própria distribuição de probabilidade das respostas do processo. O parâmetro populacional σ' é normalmente desconhecido e assim deve ser estimado conforme segue:

$$\hat{\sigma}'^2 = \sqrt{\frac{\sum_{i=1}^n (x_i - T)^2}{n-1}}$$

Aqui, n representa o tamanho da amostra retirada da população, e x_i é o valor de cada elemento da amostra. Assim, o resultado estimado de C_{pm} é:

$$\hat{C}_{pm} = \frac{USL - LSL}{6\hat{\sigma}'} \quad (2.57)$$

Redistribuindo-se convenientemente as equações acima encontra-se que:

$$\sigma'^2 = E[(X - \mu)^2] + (\mu - [USL - LSL])^2$$

E, por definição em Duncan (1974) a variância é:

$$\sigma^2 = E(X - \mu)^2$$

Disto resultam as quatro formas mais práticas de apresentação do índice C_{pm} :

$$\begin{aligned} C_{pm} &= \frac{USL - LSL}{6\sqrt{\sigma^2 + (\mu - T)^2}} = \frac{C_p}{\sqrt{1 + \frac{(\mu - T)^2}{\sigma^2}}} = \\ &= \frac{\min(USL - T; T - LSL)}{3\sqrt{\sigma^2 + (\mu - T)^2}} = \frac{d - |m - T|}{3\sqrt{\sigma^2 + (\mu - T)^2}} \end{aligned} \quad (2.58)$$

Onde ainda não foram definidos:

d_i : $(USL - LSL)/2$;

m_i : $(USL + LSL)/2$.

Os casos unilaterais também são aplicáveis para a equação (2.58). Assim, respectivamente, para respostas do tipo maior é melhor, e menor é melhor a equação (2.58) assume os seguintes formatos:

$$C_{pmi} = \frac{T - LSL}{3\sqrt{\sigma^2 + (\mu - T)^2}}$$

$$C_{pms} = \frac{USL - T}{3\sqrt{\sigma^2 + (\mu - T)^2}} \quad (2.59)$$

É imediato verificar que C_{pm} possui as propriedades necessárias para a avaliação da capacidade de processos. Além de ser adimensional, se a variância do processo cresce o valor do denominador das equações (2.58) também aumenta, diminuindo o valor do índice. O mesmo ocorre se a média das observações da característica de qualidade distancia-se do alvo pretendido, assim também será maior o denominador – aumento da diferença $(\mu - T)$ – reduzindo novamente o valor de C_{pm} .

Processos que apresentam apenas uma característica de qualidade são exceções. O mais comum que se encontra na realidade são processos que geram múltiplas características importantes que representam conjuntamente o atendimento das necessidades do cliente. Desta forma, é importante que se desenvolva um índice de capacidade de processos que contemple ao mesmo tempo todas as características do processo.

Ch'ng, Quah e Low (2005) apresentaram para o índice C_{pm} uma evolução de seu conceito. Para um produto que tenha mais de uma característica de qualidade é calculado o C_{pm} relacionado a cada uma destas medições. Depois, conforme o grau de importância individual da característica, é multiplicado por um fator de ponderação e_i , onde i indica a quantidade de características em avaliação. Assim, a seguinte equação para a capacidade de processos múltipla ou total é desenvolvida pelos autores:

$$TotalC_{pm} = \sum_{i=1}^p e_i \times \frac{\min(USL_i - T_i; T_i - LSL_i)}{3\sqrt{\sigma_i^2 + (\mu_i - T_i)^2}} =$$

$$= \sum_{i=1}^p e_i \times \frac{d_i - |m_i - T_i|}{3\sqrt{\sigma_i^2 + (\mu_i - T_i)^2}} \quad (2.60)$$

Conforme já discutido acima, e_i é o peso dado a cada uma das i características de qualidade e tem duas propriedades de interesse imediato: $0 \leq e_i \leq 1$ e a soma de todos os e_i é a

unidade. O valor p do operador somatório é justamente o número de características de qualidade em avaliação conjunta. Para os dois casos unilaterais o procedimento é similar àqueles adotados nas equações (2.59).

Montgomery (2005) indica o método da superfície de resposta (MSR) para a otimização conjunta de média e variância, o chamado problema dual de otimização. Em geral, o MSR envolve a seguinte ordem de elaboração: projeto de experimento, modelo de regressão e otimização.

Ch'ng, Quah e Low (2005) aplicam o RSM às médias e variâncias do índice $TotalC_{pm}$ apresentado na equação (2.60). Para cada valor de C_{pm} são solicitadas a média e a variância; portanto, trata-se do problema dual de otimização quando estes valores são substituídos pelas suas respectivas equações de regressão que modelam seu comportamento conforme mudam-se os fatores que as influenciam.

Em resumo, o que Ch'ng, Quah e Low (2005) propõem é: primeiramente executa-se um projeto de experimento fatorial completo onde fatores são controladamente alterados em níveis. Para cada alteração planejada destes níveis são coletadas as variáveis de resposta, conforme instrui Montgomery (2005). Em seguida, é gerado um modelo de regressão para a média e outro para a variância. Ambos modelos são instalados em cada equação de C_{pm} na equação (2.60). Por fim, deve-se testar todo o campo possível que as variáveis regressoras podem assumir em busca da combinação que resultará no maior valor do $TotalC_{pm}$. Esta última etapa pode ser realizada pelo *software* Excel[®] na ferramenta *Solver* que utiliza o algoritmo de otimização proposto em Castillo e Montgomery (1993) conhecido como *Generalized Reduced Gradient* (GRG), ou gradiente generalizado reduzido.

Com este passo adiante a equação (2.60) assume o seguinte formato aplicado para modelos de regressão:

$$\begin{aligned} TotalC_{pm} &= \sum_{i=1}^p e_i \times \frac{\min(USL_i - T_i; T_i - LSL_i)}{3\sqrt{\hat{w}_{i\sigma}^2 + (\hat{w}_{i\mu} - T_i)^2}} = \\ &= \sum_{i=1}^p e_i \times \frac{d_i - |m_i - T_i|}{3\sqrt{\hat{w}_{i\sigma}^2 + (\hat{w}_{i\mu} - T_i)^2}} \end{aligned} \quad (2.61)$$

De onde define-se:

$\hat{w}_{i\sigma}^2$: é o modelo de regressão estimado para a variância;

$\hat{w}_{i\mu}$: é o modelo de regressão estimado para a média.

Para os casos unilaterais não há novidade. O tipo maior é melhor e o tipo menor é melhor assumem os seguintes formatos respectivamente:

$$\begin{aligned} TotalCpmi &= \sum_{i=1}^p e_i \times \frac{T_i - LSL_i}{3\sqrt{\hat{w}_{i\sigma}^2 + (\hat{w}_{i\mu} - T_i)^2}} \\ TotalCpms &= \sum_{i=1}^p e_i \times \frac{USL_i - T_i}{3\sqrt{\hat{w}_{i\sigma}^2 + (\hat{w}_{i\mu} - T_i)^2}} \end{aligned} \quad (2.62)$$

Uma outra proposta de otimização de processos multiresposta por intermédio do aumento da capacidade de processos do tipo C_{pm} é apresentada por Plante (2001). No entanto, em vez de um somatório, o método envolve um produtório.

$$MC_{pm} = \left[\prod_{i=1}^p \frac{USL_i - LSL_i}{6\sqrt{\sigma_i^2 + (\mu_i - T_i)^2}} \right]^{\frac{1}{p}} \quad (2.63)$$

Plante (2001) ressalta que uma implicação primordial da equação (2.63) é que ela exige que todos os C_{pm} apresentem valores mais ou menos próximos entre si e altos, pois basta que apenas um deles esteja próximo de zero – ou seja, num baixo nível de qualidade – para que MC_{pm} seja baixo também, não importando quão bem estão os demais índices de capacidade de processos. Seguindo o mesmo raciocínio, esta implicação também ajuda identificar qual a resposta do processo que uma vez sofrendo melhorias de qualidade, mais impactará na qualidade total do produto.

Capítulo 3

Este capítulo está dividido em duas seções. Na primeira seção apresenta-se uma revisão do trabalho de Arriba (2005), nos pontos de interesse para esta dissertação, em especial o Capítulo 4 do referido texto. Assim, julga-se que não será necessário recorrer àquele texto para o adequado entendimento dos procedimentos aqui adotados. A segunda seção traz a metodologia a ser aplicada na otimização dos experimentos conduzidos por Arriba (2005) segundo a proposta desta dissertação e que foi destacada nos objetivos principais deste trabalho, a saber, analisar os resultados da otimização realizada por Arriba (2005) com outra alternativa de otimização multiresposta.

3.1 O PROJETO DE EXPERIMENTO FATORIAL DE ARRIBA (2005)

O problema encontrado por Arriba (2005) é o erro de classificação de couros em seus três estágios de produção. O curtimento de couros passa pelo estágio *wet-blue*, semi-acabado e acabado. A figura (3.1) apresenta esquematicamente cada um desses estágios.

As etapas de classificação do couro, objeto de estudo no experimento de Arriba (2005), ocorre no início de cada um desses três estágios. No estágio *wet-blue*, a classificação serve para confirmar a seleção conduzida anteriormente no couro, no momento da recepção da matéria-prima. No estágio seguinte, de semi-acabado, uma nova classificação é realizada com as mesmas peças, porém já mais adiantadas no processo de curtimento e, portanto, as más características do couro apresentam-se mais evidentes. Por fim, no estágio final conhecido como acabado, o couro é novamente classificado para dirimir qualquer possibilidade de falhas de classificação.

O couro apresenta, invariavelmente, defeitos de superfície que comprometem sua qualidade quanto ao tipo de aplicação futura, como produto final para consumo. A classificação destes defeitos está sujeita à subjetividade do classificador, dependendo também do estágio de classificação. Quanto mais avançado está o processo de curtimento, mais destacadas as avarias se apresentam.

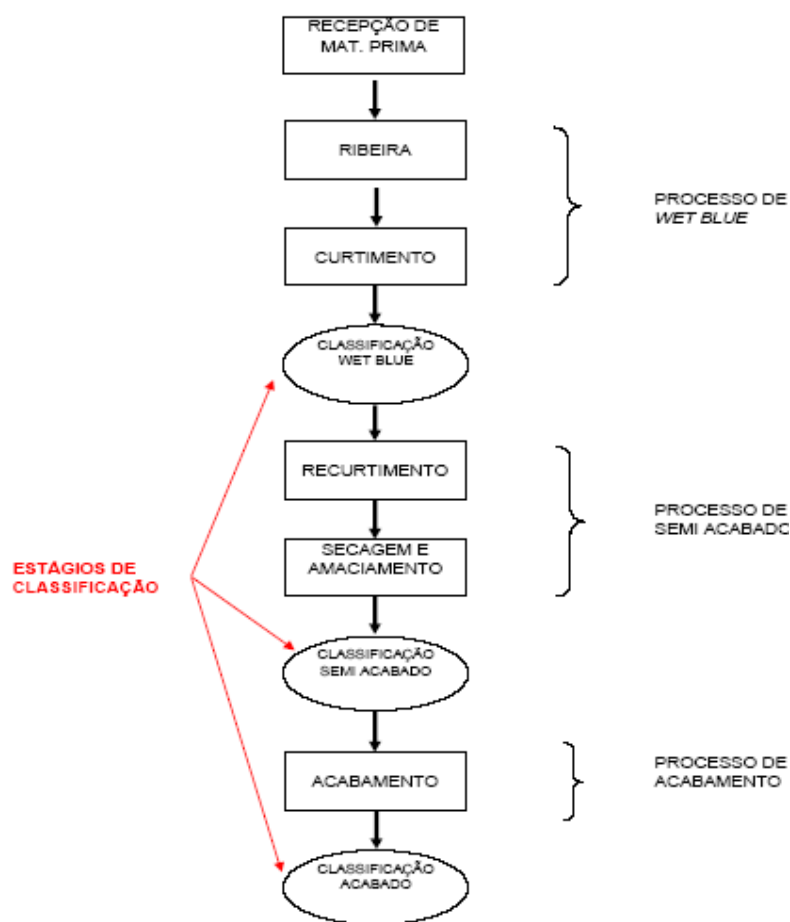


Figura (3.1) – Fluxograma do processo de curtimento de couros. Fonte: Arriba (2005)

No entanto, a cada etapa do processo de curtimento (figura 3.1) mais gastos incidem sobre o produto. Portanto, erros de classificação implicam em desperdícios que serão percebidos depois que os custos de curtimento foram incorridos no produto. A qualidade de uma pele é essencialmente determinada pela área livre de marcas, pois mantém a padronização natural do couro. A ruptura desta padronização natural é o defeito, por excelência, que os classificadores devem identificar o mais cedo possível no processo. Arriba (2005) analisa dois tipos de erros:

- a) o classificador não reconhece um couro, em qualquer dos estágios, como não apto e o aceita como apto. Este tipo de erro é denominado como de *classificação errada* (CE);

- b) o classificador, ao contrário do caso anterior, reconhece um couro apto como não apto. Este tipo de erro é denominado *alarme falso* (AF).

No erro CE, as conseqüências são piores do ponto de vista dos custos, conforme análise de valores conduzidas por Arriba (2005). Dado o alto investimento típico do curtimento, couros não tão bons quanto se afirma na classificação, recebem tratamento caro e, adiante no processo, não resultam em matéria-prima que possa ser utilizada para o fim inicialmente esperado. Devido ao conservadorismo característico dos classificadores, o erro CE é menos comum que o erro AF.

É justamente esse conservadorismo que abre a possibilidade da ocorrência do segundo tipo de erro (AF). Quanto maior a qualidade de uma pele, maior o seu preço de mercado. Assim, o custo do produtor em ter um couro excelente tratado como possuidor de uma qualidade inferior é a diferença a menos que o mercado pagará por ele.

Os objetivos principais de Arriba (2005) são, portanto: (i) minimizar os erros por classificação errada (CE); (ii) minimizar os erros por alarme falso (AF); (iii) reduzir as diferenças de notas que atestam a qualidade entre as classificações de um estágio a outro e; (iv) reduzir o desvio-padrão destas diferenças. Para tanto, um projeto de experimentos foi realizado de forma que três classificadores experientes atribúissem notas às peles, em cada um dos três estágios de classificação. Foi utilizada a seguinte escala:

- a) nota 10: as melhores peles dentro da classificação de utilidade final atribuída a elas;
- b) nota 8: peles de qualidade média dentro da classificação de utilidade final atribuída a elas;
- c) nota 6: peles no limite inferior de qualidade dentro da classificação de utilidade final atribuída a elas;
- d) nota zero: peles com classificação errada, ou seja, classificadas para um fim para o qual não podem ser utilizadas.

Estas notas são atribuídas a cada pele nas etapas do processo de curtimento do couro. Sendo que no estágio *wet-blue* ocorrem duas classificações, uma feita pelo próprio classificador e a outra por seu supervisor. A figura (3.2) esquematiza o método.

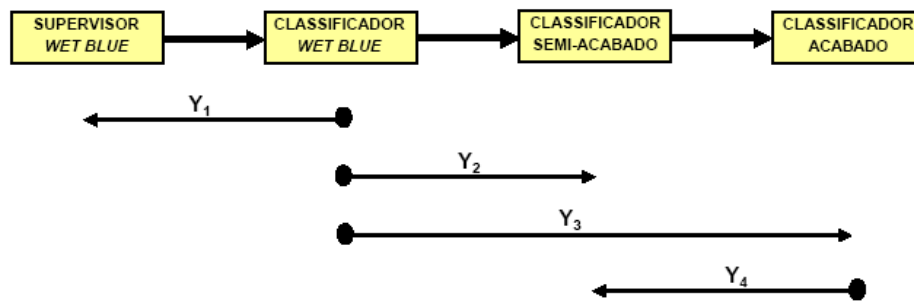


Figura (3.2) – Variáveis associadas à classificação das peles. Fonte: Arriba (2005)

Na figura (3.2) torna-se evidente que as notas são comparadas entre si em quatro ocasiões:

- a) comparação das notas dadas pelo classificador *wet-blue* e seu supervisor no mesmo estágio, pele a pele. Estas notas representam o conjunto de respostas Y_1 ;
- b) comparação entre as notas dadas pelo classificador *wet-blue* e o classificador do estágio seguinte (semi-acabado). Estas notas representam o conjunto de respostas Y_2 ;
- c) comparação entre as notas dadas pelo classificador de semi-acabado e o classificador de acabado. Estas notas representam o conjunto de respostas Y_3 ;
- d) comparação entre o classificador de acabado e de semi-acabado, representada pelo conjunto de respostas Y_4 . Estas comparações não foram realizadas para o artigo *nubuck* chocolate, e portanto está fora do escopo desta dissertação.

Quando se diz conjunto de respostas Y , faz-se referência a quatro medidas tomadas a partir das notas dadas em cada extremidade das setas observadas na figura (3.2). Estas medidas são:

- a) diferença entre as notas. A primeira nota é subtraída da segunda nota no sentido da seta. Por exemplo, se a nota dada pelo classificador de *wet-blue* for 8 e a nota dada pelo seu supervisor no mesmo estágio for 10, a resposta será (-2) . Aqui o sinal negativo significa que o classificador foi mais rigoroso na classificação que

seu supervisor. Portanto, deseja-se que não haja diferença entre ambos, o que implica em um valor alvo desejado igual a zero, denotando a situação onde a nota dada pelo classificador é a mesma dada pelo supervisor. Ou ainda, que a nota dada no estágio anterior de classificação seja igual à nota dada no estágio posterior, conforme já detalhado pela figura (3.2);

- b) desvio-padrão: a série de resultados das diferenças acima possui um desvio-padrão que deve ser minimizado. Quanto mais próximo de zero for este desvio-padrão, mais próximas entre si são as diferenças de respostas dadas pelos classificadores. Diferença que se deseja nula como explicado acima;
- c) probabilidade de classificação errada (CE): é o percentual dentre as peles do lote em avaliação que foram classificadas com erro para o fim a que se destina. Por exemplo, diga-se que o classificador de semi-acabado recebe uma pele e avalia que esta não serve para o produto cujo uso final foi indicado. Assim, ele atribuirá à pele uma nota zero. Se o classificador anterior (*wet-blue*) atribuiu à pele uma nota 6, por exemplo, ele a classificou errado, pois no processo seguinte a pele foi rejeitada para aquele fim. O ideal é que ainda no estágio *wet-blue* fosse já atribuída a esta pele a nota zero. Dentro do lote calcula-se o percentual de erros desta natureza cometidos. Quanto menor o percentual de classificação errada, melhor;
- d) probabilidade de alarme falso (AF): neste caso, o classificador reprova uma pele boa que só será identificada como tal na etapa seguinte. Ou seja, na primeira avaliação é dada uma nota zero, ao passo que no estágio seguinte observa-se que aquela pele deveria ter sido aprovada. Da mesma forma que no item anterior, calcula-se o percentual de erros desta natureza dentro do lote para se estimar a probabilidade de alarme falso. Quanto menor este percentual, melhor.

A partir das diferenças de notas explicadas na alínea (a) acima, pode-se montar a tabela (3.1) que apresenta as diferenças possíveis entre as notas dadas. Na tabela, vê-se que quaisquer que sejam as classificações erradas ou alarmes falsos, estas sempre envolvem uma nota zero atribuída a pele, seja num sentido ou noutro da seta. Conforme já dito anteriormente, a nota zero indica a inaptidão da pele para o produto final ao qual ela se destina.

Tabela (3.1)
Possibilidades que as diferenças entre respostas podem assumir

1ª Nota	2ª Nota (possibilidades)	Diferença (Y_i)	Comentário
10	10	Zero	Ambos concordam
	8	2	Pequena discordância
	6	4	Grande discordância
	Zero	10	Classificação errada (CE)
8	10	-2	Discordância pequena no sentido inverso
	8	Zero	Ambos concordam
	6	2	Discordância pequena
	Zero	8	Classificação errada (CE)
6	10	-4	Discordância grande no sentido inverso
	8	-2	Discordância pequena no sentido inverso
	6	Zero	Ambos concordam
	Zero	6	Classificação errada (CE)
Zero	10	-10	Alarme falso (AF)
	8	-8	Alarme falso (AF)
	6	-6	Alarme falso (AF)
	Zero	Zero	Ambos concordam

Observa-se, portanto, a partir das informações da tabela (3.1), que as possibilidades de diferenças envolvem exclusivamente números inteiros $-10, -8, -6, -4, -2, 0, 2, 4, 6, 8$ e 10 . Conclui-se assim que, exceto pelo fato da presença de número negativos, é atendida uma das condições da distribuição de Poisson, que é a presença exclusiva de números inteiros e não-negativos.

Uma vez determinadas as quatro respostas que representam a qualidade do produto, Arriba (2005) elege, a partir de considerações da engenharia, os fatores controláveis que podem vir a influenciar as respostas. São eles: (i) a procedência da matéria-prima (couro), (ii) a umidade da matéria-prima, (iii) o método de classificação e (iv) o classificador.

O primeiro fator é a procedência do couro. Na origem, o couro já é fornecido ao curtume com uma determinada finalidade de uso. No início do processo de classificação, estágio *wet-blue*, o classificador confirma esta finalidade. Caso ele não concorde com a atribuição dada na origem, ele lhe dará nota zero, pois considerou que a classificação foi errada. O mesmo pode acontecer nos estágios seguintes do processo sempre em relação à etapa anterior, conforme a direção das setas na figura (3.2). São estudadas duas origens: Centro e Norte. Este fator é designado pela abreviação PD e é claramente qualitativo.

O segundo fator é a umidade, testada em três níveis: molhado, enxuto e rebaixado (mais seco). A umidade, acredita-se, pode afetar as notas que se atribuem às peles. Em geral,

quanto mais enxuta apresentar-se a pele, mais fidedignas são as classificações. Arriba (2005) explica que este poderia ser um fator quantitativo, porém o medidor de umidade é impreciso, o que levou a tratá-lo também como qualitativo e sob a abreviação UM.

O método de classificação pode ser de dois tipos: estático ou dinâmico. No método estático o couro é classificado sobre o próprio *pallet* e o classificador usa o tempo que quiser para dar a nota. Neste caso, em geral, o tempo de classificação é maior, sendo que a distância entre o classificador e o couro, bem como a iluminação, são mantidos constantes, por serem potenciais fatores importantes de ruído. No segundo método, a classificação é feita ao longo de uma esteira móvel que é usada pela máquina de enxugar. Portanto, neste caso, a velocidade da esteira é que determina o ritmo de classificação. Este fator é, portanto, qualitativo e será identificado pela abreviação TE, por envolver o tempo.

Arriba (2005) treinou três funcionários experientes para realizar os testes. Estes classificadores serão designados por Ademir, Geraldo e Valdecir. A decisão de incluir o fator *classificador* na análise foi baseada no fato do método de classificação ser bastante subjetivo e, portanto, o treinamento da equipe de classificação deve ser feito de tal forma a reduzir-se ou padronizar-se tanto quanto possível esta subjetividade. Analisar o fator *classificador* poderá indicar seu efeito sobre a qualidade do processo, é imediato perceber que também é qualitativo, e referido com a abreviação CL.

A tabela (3.2) apresenta os fatores, seus níveis e a codificação de níveis usada por Arriba (2005) e adotada também aqui.

Tabela (3.2)
Fatores, níveis e níveis codificados do experimento

Fator	Níveis	Níveis Codificados
Procedência (PD)	Centro	-1
	Norte	+1
Umidade (UM)	Molhado	-1
	Enxuto	0
	Rebaixado	+1
Classificador (CL)	Ademir	-1
	Geraldo	0
	Valdecir	+1
Método (TE)	Estático	-1
	Dinâmico	+1

Fonte: adaptado Arriba (2005).

A combinação dos quatro fatores em seus dez níveis de investigação perfaz um total de 36 arranjos (ou combinações) possíveis: trata-se de um projeto fatorial $2^2 * 3^2$. Arriba

(2005) replicou cada uma dessas combinações 20 vezes em seu experimento fatorial completo, totalizando 720 réplicas para cada um dos artigos analisados. Quatro artigos foram estudados: *nubuck* chocolate, *naplex* preto, *diamond* preto e *nubuck* bege. Cada um destina-se a uma finalidade em termos de uso final. Esta dissertação desenvolverá sua proposta com o artigo *nubuck* chocolate exclusivamente.

A tabela (3.3) exemplifica um dos tratamentos experimentais rodados por Arriba (2005). No tratamento em questão, o fator TE é estático (nível -1), o fator CL é Valdecir (nível +1), o fator UM é molhado (nível -1) e o fator PD é Centro (nível -1). A tabela (3.3) informa as notas dadas pelo classificador no estágio *wet-blue* e pelo seu supervisor no mesmo estágio (sexta e primeira colunas, respectivamente). A diferença entre ambos é dada na sétima coluna (Y_1). A média e desvio-padrão dos vinte valores da coluna Y_1 são apresentados nas últimas linhas da referida tabela. Calcula-se também o percentual referente ao total de classificações erradas e alarme falso nas vinte observações. O mesmo é feito com as colunas Y_2 e Y_3 . Observe-se que Y_2 traz as diferenças entre a nota dada pelo classificador no estágio *wet-blue* e quando o produto vai para o processo semi-acabado. O mesmo ocorre com Y_3 , entre as etapas de semi-acabado e acabado.

Tabela (3.3)
Exemplo de tratamento experimental e variáveis de resposta

Nota atribuída BLUE	Nota atribuída SEMI	Nota atribuída ACABADO	Procedência	Ordem da pele	Artigo: <i>Nubuck Bege</i>			
					Nota	Y_1	Y_2	Y_3
0	6	6	Centro	1	6	6	0	0
6	6	6		2	8	2	2	2
10	10	10		3	10	0	0	0
8	10	10		4	6	-2	-4	-4
10	10	10		5	8	-2	-2	-2
6	10	0		6	8	2	-2	8
8	8	8		7	6	-2	-2	-2
0	10	6		8	8	8	-2	2

Tabela (3.3)
Exemplo de tratamento experimental e variáveis de resposta (continuação)

Nota atribuída <i>BLUE</i>	Nota atribuída <i>SEMI</i>	Nota atribuída <i>ACABADO</i>	Procedência	Ordem da pele	Artigo: <i>Nubuck Bege</i>				
					Nota	<i>Y1</i>	<i>Y2</i>	<i>Y3</i>	
6	10	8	Centro	9	6	0	-4	-2	
10	10	10		10	8	-2	-2	-2	
8	10	8		11	8	0	-2	0	
0	6	0		12	8	8	2	8	
0	10	6		13	8	8	-2	2	
6	8	8		14	6	0	-2	-2	
10	10	10		15	6	-4	-4	-4	
8	8	8		16	6	-2	-2	-2	
10	10	10		17	6	-4	-4	-4	
6	10	8		18	8	2	-2	0	
8	10	10		19	6	-2	-4	-4	
0	6	6		20	6	6	0	0	
6,0	8,9	7,4		Média total		7,1	1,1	-1,8	-0,3
				Desvio-padrão			4,0	1,8	3,5
				% Prob (CE)			25	0	10
				% Prob (AF)			0	0	0

Fonte: adaptado Arriba (2005)

Arriba (2005) utilizou o seguinte método de otimização em seu experimento. Primeiramente o autor tomou as 720 observações para o artigo *nubuck* chocolate, lembre-se que cada tratamento foi replicado 20 vezes. Para cada grupo de 20 réplicas foram calculados Y_1, Y_2 , e Y_3 , e, para cada uma dessas respostas calculou-se a média, o desvio-padrão, a probabilidade de classificação errada e a probabilidade de alarme falso.

Em concomitância, Arriba (2005) associou a variável de resposta Y_1 aos fatores do experimento, através de regressão linear clássica (para uma introdução sobre o tema, ver Montgomery, Peck e Vining, 2006). O procedimento foi repetido, obtendo modelos de regressão para a probabilidade de classificação errada e a probabilidade de alarme falso, associando-as aos fatores do experimento. Para cada tratamento foram calculados os desvios-padrão da resposta Y_1 e, por regressão linear clássica, estes foram associados aos fatores do experimento. O mesmo foi também realizado para Y_2 e Y_3 . A significância dos termos inseridos em todos os modelos de regressão foi analisada através de análise de variância (ANOVA). Os resultados obtidos por Arriba (2005) para o artigo *nubuck* chocolate são apresentadas nas equações (3.1). Os coeficientes de determinação R^2 dos modelos por ele encontrados são 0,33, 0,51 e 0,33, respectivamente.

$$\begin{aligned}
 Y_1 &= 0,45 - 0,56 \times CL - 0,35 \times UM \times TE \times CL \\
 DesvPad(Y_1) &= 3,73 - 0,44 \times UM \\
 PROB(CE_{Y_1}) &= 21,67 - 2,29 \times UM - 2,29 \times CL + \\
 &+ 2,29 \times PD \times TE \times CL - 8,54 \times UM^2 \\
 PROB(AF_{Y_1}) &= 17,08 - 9,58 \times UM^2
 \end{aligned}
 \tag{3.1}$$

Existe uma graduação de importância entre as respostas modeladas nas equações (3.1); esta graduação é definida conforme o impacto financeiro representado por cada resposta e foi definida por experiência da equipe. Arriba (2005) atribuiu pesos a cada uma delas conforme a tabela (3.4). Nota-se que a soma dos pesos de ponderação leva à unidade.

Tabela (3.4)
Pesos de ponderação para cada variável de resposta

Variável de resposta	Peso
Y_1	0,15
Desvio-padrão	0,20
Probabilidade de classificação errada	0,40
Probabilidade de alarme falso	0,25

Fonte: adaptado Arriba (2005)

Para unir as respostas em um único valor adimensional, Arriba (2005) divide o resultado de cada equação por sua respectiva média total (considerando 720 valores no caso de Y_1 e 36 valores no caso do desvio-padrão e das probabilidades). Cada quociente, formado pela equação de regressão e sua respectiva média geral, é multiplicado pelo seu valor de ponderação; finalmente, somam-se os resultados para obter a função objetivo Z_i . O resultado Z_1 , para a variável de resposta Y_1 e estatísticas dela derivadas, é apresentado na seqüência. O mesmo procedimento é repetido para as variáveis Y_2 e Y_3 .

$$\begin{aligned}
Z_1 = & 0,15 \times \left(\frac{Y_1}{\text{média_geral_}Y_1} \right) + \\
& + 0,20 \times \left(\frac{\text{Desvpad}(Y_1)}{\text{média_geral_desvio - padrão}} \right) + \\
& + 0,40 \times \left(\frac{\text{Prob}(CE)}{\text{média_geral_Prob}(CE)} \right) + \\
& + 0,25 \times \left(\frac{\text{Prob}(AF)}{\text{média_geral_Prob}(AF)} \right)
\end{aligned} \tag{3.2}$$

Conforme definido anteriormente, Y_1 corresponde à diferença de notas atribuídas entre o classificador e seu supervisor no estágio *wet-blue*, Y_2 representa a diferença de notas entre a classificação de *wet-blue* e semi-acabado, e Y_3 corresponde à diferença de notas de classificação entre *wet-blue* e a fase final do processo de curtimento, conhecido como acabado. As equações (3.1) e (3.2) também foram geradas para as respostas Y_2 e Y_3 . Arriba (2005) atribui uma importância a cada um desses estágios (ou etapa de classificação), conforme apresentado na tabela (3.5).

Tabela (3.5)
Pesos de ponderação para cada estágio de classificação

Estágio	Peso
Classificação de produto final (Y_3)	0,50
Classificação de <i>wet-blue</i> (Y_1)	0,35
Classificação em semi-acabado (Y_2)	0,15

Fonte: adaptado Arriba (2005)

A otimização é encaminhada, a partir deste ponto, multiplicando-se os valores Z_1 , Z_2 e Z_3 por seus respectivos pesos e somando-se estas parcelas, como apresentado na equação (3.3).

$$Z_{Total} = 0,35 \times Z_1 + 0,15 \times Z_2 + 0,50 \times Z_3 \tag{3.3}$$

Para finalizar a otimização, Arriba (2005) calcula os valores Z_{Total} para os 36 tratamentos experimentais. A tabela (3.6) apresenta os resultados encontrados por Arriba (2005) para o artigo *nubuck* chocolate.

Os valores-alvo para as respostas Y_1 , Y_2 e Y_3 são iguais a zero, o que indicaria a menor diferença existente entre as notas dadas entre cada etapa do processo. Para o desvio-padrão de Y_i , o alvo também é zero, assim como para as probabilidades de classificação errada e alarme falso. Desta forma, é desejável que Z_{Total} seja o mais próximo de zero. Ou seja, a combinação dos níveis dos fatores que conduzir ao Z_{Total} mais próximo de zero significa a resposta

otimizada do experimento. O primeiro terço dos menores valores resultantes de Z_{Total} estão sombreados na tabela (3.6), que são, em ordem, os tratamentos: 13, 15, e 14.

Tabela (3.6)
Valores da função objetivo Z_{Total} . Valores ótimos sombreados

Tratamento	PD	UM	TE	CL	Z_{Total}	Tratamento	PD	UM	TE	CL	Z_{Total}
1	-1	-1	-1	-1	0,98	19	+1	-1	-1	-1	1,01
2	-1	-1	-1	+1	1,12	20	+1	-1	-1	+1	1,02
3	-1	-1	-1	0	1,00	21	+1	-1	-1	0	0,96
4	-1	-1	+1	-1	0,86	22	+1	-1	+1	-1	1,09
5	-1	-1	+1	+1	0,86	23	+1	-1	+1	+1	1,12
6	-1	-1	+1	0	0,86	24	+1	-1	+1	0	1,10
7	-1	0	-1	-1	1,03	25	+1	0	-1	-1	1,20
8	-1	0	-1	+1	1,11	26	+1	0	-1	+1	1,15
9	-1	0	-1	0	1,05	27	+1	0	-1	0	1,16
10	-1	0	+1	-1	1,07	28	+1	0	+1	-1	1,16
11	-1	0	+1	+1	1,07	29	+1	0	+1	+1	1,19
12	-1	0	+1	0	1,05	30	+1	0	+1	0	1,16
13	-1	+1	-1	-1	0,76	31	+1	+1	-1	-1	1,07
14	-1	+1	-1	+1	0,84	32	+1	+1	-1	+1	1,02
15	-1	+1	-1	0	0,80	33	+1	+1	-1	0	1,05
16	-1	+1	+1	-1	0,97	34	+1	+1	+1	-1	0,91
17	-1	+1	+1	+1	1,02	35	+1	+1	+1	+1	1,01
18	-1	+1	+1	0	0,94	36	+1	+1	+1	0	0,91

Fonte: adaptado Arriba (2005)

3.2 METODOLOGIA APLICADA

Um dos objetivos desta dissertação é otimizar um experimento multiresposta, onde ao menos uma das variáveis de resposta seja categórica. No experimento conduzido e otimizado por Arriba (2005), as respostas Y_1 , Y_2 e Y_3 são claramente categóricas. No entanto, nenhuma instância da família exponencial de distribuições de probabilidade abordadas no Capítulo 2 se adequa às características das respostas Y_1 , Y_2 e Y_3 , já que suas realizações podem assumir valores inteiros positivos, negativos ou zero. Para contornar este problema, deve-se somar uma constante a cada resposta Y_i de forma a eliminar os valores negativos. Feita essa mudança de escala, as respostas Y_1 , Y_2 e Y_3 passam a ser descritas, acredita-se, por uma distribuição de Poisson.

A tabela (3.1) apresenta os possíveis resultados das variáveis Y_i . Se a elas adicionar-se a constante +10, os valores iniciais -10, -8, -6, -4, -2, 0, 2, 4, 6, 8 e 10, serão trasladados na mesma ordem 0, 2, 4, 6, 8, 10, 12, 14, 16, 18 e 20.

Aplicando-se os testes de aderência (*goodness of fit test*) de Kolmogorov-Smirnov e Anderson-Darling, descritos por Kelton e Law (1991), através da versão estudantil do *software* Statfit, não é possível rejeitar a hipótese de distribuição de Poisson para as respostas Y_i , com um grau de confiança superior a 99,9%. É importante ressaltar que, nos testes de aderência, não foram utilizados todos os valores de Y_i , pois tais testes deterioram-se na medida em que o tamanho de amostra aumenta (ver Kelton e Law, 1991, p.387). Assim, do conjunto de 720 observações para cada variável de resposta foram selecionadas, aleatoriamente, 50 observações, as quais foram usadas nos testes de aderência.

Uma vez identificada a distribuição de probabilidade que melhor se encaixa na família exponencial - no caso a distribuição de Poisson - é realizado todo o procedimento de regressão detalhado na primeira seção do capítulo 2 segundo a técnica dos modelos lineares generalizados. Pretende-se assim, conforme já apresentado anteriormente, evitar-se o equívoco incorrido por Arriba (2005) em executar a regressão utilizando o método dos mínimos quadrados em variáveis de resposta categóricas.

Os fatores procedência, umidade, método de classificação e classificador são todos apresentados em níveis discretos, ou seja, seus níveis não são valores contínuos. Segundo Gujarati (2006), nestes casos a regressão deve considerar as variáveis regressoras - os fatores - como variáveis *dummy*. Isto se deve ao fato que, por serem discretos, cada nível estará presente ou não na equação de regressão. Por exemplo, tome-se o fator classificador. Segundo

a tabela (3.2) há três níveis possíveis neste fator: Ademir, Geraldo e Valdecir. Estes níveis são mutuamente exclusivos na equação de regressão, ou seja, se Ademir está presente gerando um certo valor à resposta da equação, Geraldo e Valdecir estão obrigatoriamente ausentes.

Na regressão por variáveis *dummy*, portanto, os níveis apresentam-se com a única possibilidade de assumirem os valores 1 (quando presentes) ou zero (quando ausentes). Neste método, um nível de cada fator será considerado como grupo de controle, significa que todos os outros níveis dos fatores são zero. Este outro grupo é conhecido por grupo de tratamento. Quando um nível do grupo de tratamento assume o valor 1 – ou seja, está presente – a equação de regressão expressa o quanto altera-se seu resultado sobre os níveis do grupo de controle. Desta forma, nesta dissertação, serão considerados os fatores como variáveis *dummy* em todas as regressões. Este procedimento corrige uma falha nas regressões de Arriba (2005).

O próximo passo metodológico a ser seguido será obter um modelo de regressão para os desvios-padrão, calculados com os novos valores de Y_1 , Y_2 e Y_3 após deslocamento de 10 unidades nos valores observados. Mais uma vez os testes de aderência de Kolmogorov-Smirnov e Anderson-Darling foram rodados e a hipótese de normalidade dos desvios não pode ser rejeitada, o que habilita o uso do método dos mínimos quadrados na regressão, no entanto, esta regressão será conduzida pelo método dos modelos lineares generalizados usando uma função de ligação apropriada para respostas normais, conforme apresentado na tabela (2.2).

O próximo passo a ser cumprido é a modelagem de regressão das probabilidades de classificação errada (CE) e alarme falso (AF), desta vez modelados considerando que os fatores têm níveis em variáveis discretas, portanto *dummy*. Estes procedimentos de regressão serão idênticos ao descrito no parágrafo anterior.

Em seguida todo o procedimento de otimização de Arriba (2005) é repetido, porém com as modelagens de regressão apropriadas, ou seja, modelagem por modelos lineares generalizados com função de ligação de Poisson, e ainda considerando os fatores como variáveis *dummy*. Conforme apresentado na seção anterior, o melhor valor para a função objetivo Z_i de Arriba (2005) é zero. Na reanálise do experimento a situação muda devido à translação dos resultados do experimento, e o novo valor alvo será 10. Assim, o valor de Y_i que representa a menor quantidade de erros é aquele que está no ponto médio da escala, ou seja, 10. Os desvios-padrão e as duas probabilidades devem assumir valores tão pequenos

quanto possíveis. Será montada, então, uma tabela similar à tabela (3.6) para a eleição dos fatores e níveis que implicam na melhor forma de seleção de cursos.

A modelagem por modelos lineares generalizados exige que se considere uma distribuição de probabilidade para as respostas do experimento. No presente caso, atribui-se a distribuição de Poisson baseando-se em dois argumentos principais: as respostas são valores inteiros e não-negativos e ainda nos resultados dos testes de Kolmogorov-Smirnov e Anderson-Darling que indicaram que não se pode rejeitar a hipótese desta distribuição não ser de Poisson. No entanto, embora a possibilidade de erro seja exígua, por precaução contra o erro será também conduzida, para as respostas Y_1 , Y_2 e Y_3 a regressão por quase-verossimilhança que não exige *a priori* que estas respostas atendam obrigatoriamente a uma distribuição de probabilidade.

Caso haja uma diferença significativa entre os resultados da regressão por modelos lineares generalizados com função de ligação de Poisson e a modelagem por quase-verossimilhança, optar-se-á pela segunda.

Assim pretende-se cumprir com o primeiro objetivo específico desta dissertação. O próximo objetivo é apresentar uma alternativa de otimização àquela utilizado por Arriba (2005). Para tanto, o procedimento adotado por Ch'ng, Quah e Low (2005) para a capacidade de processos do tipo C_{pm} será aplicado para as equações de regressão Y_1 ; desvio-padrão; probabilidade de classificação errada e probabilidade de alarme falso. Os valores de ponderação citados na tabela (3.4) serão utilizados, conforme preconizado pelas equações (2.61) e (2.62), segundo a característica da variável de resposta.

O mesmo procedimento será adotado para as equações de regressão Y_2 e Y_3 , e seus respectivos desvios-padrão e probabilidades de classificação errada (CE) e alarme falso (AF), obedecendo aos pesos de ponderação que Arriba (2005) usou conforme a tabela (3.4) dentro da equação (2.60). Neste momento será montado o *tableau* de otimização segundo o algoritmo do gradiente generalizado reduzido, disponível pela ferramenta *Solver* do Excel[®]. Neste *tableau*, tal qual realizado por Ch'ng, Quah e Low (2005), os níveis dos fatores serão designados como inteiros variando entre 0 e 1. Pois, cada nível passa a ser uma variável que assume um desses dois valores. Isto se deve à instrução de exclusividade entre os níveis em virtude das variáveis *dummy* dos fatores.

Encerrando, o valores de $TotalC_{pm}$ da equação (2.61) serão ponderados para Y_1 , Y_2 e Y_3 , e ainda segundo os pesos da tabela (3.5). Este procedimento será realizado uma única vez, pois o algoritmo busca um ótimo total dentro de todas as possibilidades de níveis que os fatores podem assumir. Desta forma, usando-se o método de otimização proposto com o uso do $TotalC_{pm}$ e os mesmo pesos de Arriba (2005), encontra-se uma nova forma de identificar os níveis ótimos onde os fatores produzem o menor erro de classificação entre as etapas de curtimento do couro. A figura (3.3) resume de forma sucinta os passos que são cumpridos a partir de adiante.

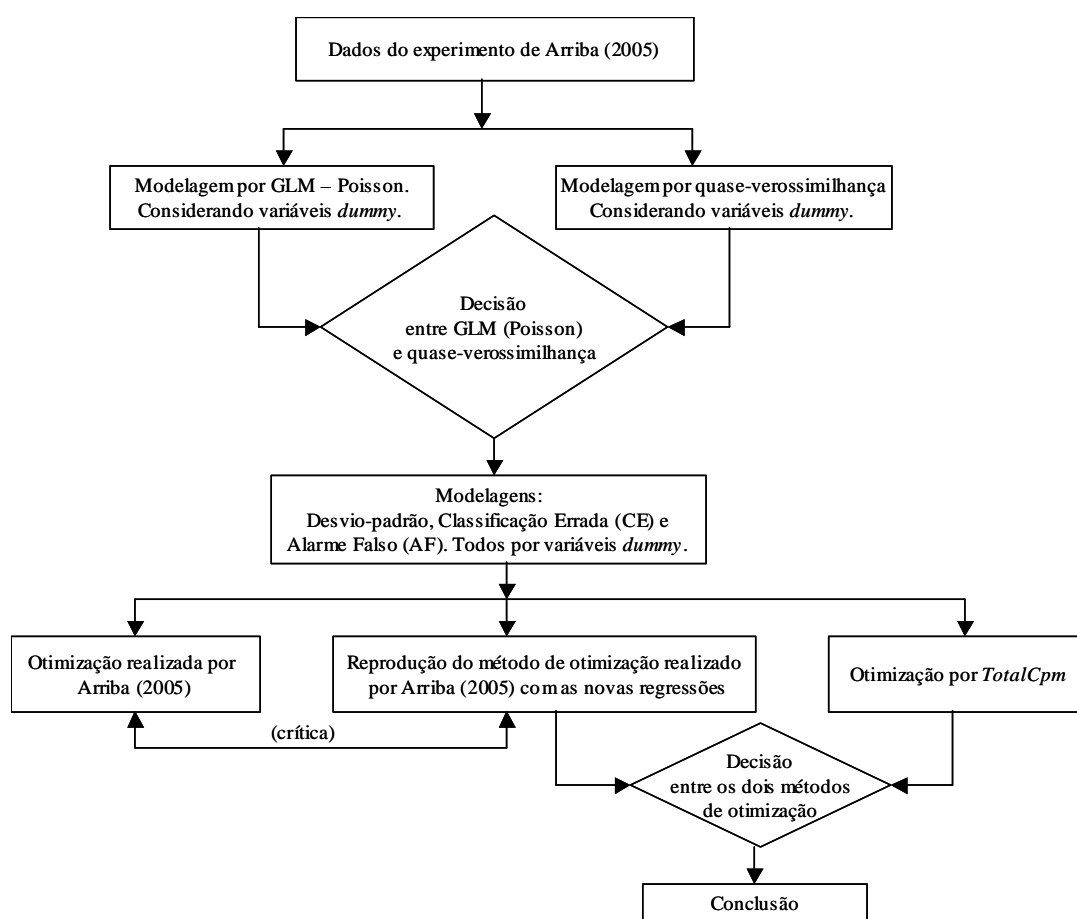


Figura (3.3) – Metodologia da dissertação

Capítulo 4

Conforme exposto na figura (3.3) a primeira etapa neste capítulo é conduzir, a partir dos dados do experimento de Arriba (2005), a regressão por modelos lineares generalizados com função de ligação de Poisson; e também a regressão do mesmo experimento pelo método da quase-verossimilhança.

Após os resultados destas duas regressões, será feita a comparação entre ambas de forma a decidir qual delas será otimizada segundo o método de Arriba (2005), e, posteriormente, será maximizada pelo método do índice múltiplo de capacidade de processos, $TotalC_{pm}$, de forma a permitir a comparação entre os dois métodos de otimização.

4.1 AS REGRESSÕES POR MODELOS LINEARES GENERALIZADOS E QUASE-VEROSSIMILHANÇA

No anexo B de Arriba (2005, p.131) são apresentados os fatores, os níveis e as respostas encontradas, os quais serão utilizadas neste trabalho. Conforme mencionado no capítulo anterior, especificamente a tabela (3.2), os fatores são Procedência (PD), Umidade (UM), Método de Classificação (TE) e Classificador (CL). Os fatores PD e TE são explorados em dois níveis (-1 e +1); já os fatores UM e CL são explorados em três níveis (-1, 0, +1). Todos os níveis são codificados. Vale ainda ressaltar que o produto cujos dados são analisados nesta dissertação é o *nubuck* chocolate, relatado em Arriba (2005).

O procedimento de regressão apresentado no capítulo 2 foi executado com o uso do *software* estatístico R. Este *software* foi selecionado por permitir o uso de diversas funções de ligação para a regressão por modelos lineares generalizados, inclusive a função de ligação de Poisson. Além do mais, o programa é gratuito e pode ser baixado da internet. Sua programação e atualização é permanentemente mantida por um comitê de programadores e estatísticos em universidades de diversos países numa espécie de regime cooperativado.

O apêndice A apresenta passo a passo as instruções que devem ser dadas o *software* para cada uma das duas regressões. A tabela (4.1) apresenta os resultados da equação de

regressão de Y_1 . Tanto para Y_1 , quanto para Y_2 e Y_3 as respostas do experimento de Arriba (2005) foram somadas por +10, de forma a que apenas valores inteiros e não-negativos estivessem presentes, adequando-se assim a uma distribuição de Poisson. Além disso, nas regressões de Y_1 , Y_2 e Y_3 foram consideradas as interações até segunda ordem entre os fatores.

No conjunto de tabelas que se segue, estão com fundo sombreado em destaque os coeficientes que resultaram em um nível de significância de pelo menos 95%. As duas últimas colunas apresentam o valor do teste de confiabilidade destes coeficientes, sendo que a probabilidade da significância não é apresentada para os coeficientes que não resultaram em pelos menos 95% de significância.

Na coluna que nomeia o coeficiente - primeira coluna da referida tabela - observa-se que, com exceção do termo independente (intercessão), cada fator é apresentado com o subscrito do nível ao qual a estimativa do coeficiente refere-se. Isto é decorrência da regressão conduzida considerando que os níveis dos fatores são discretos, portanto, ou variáveis *dummy*. A última linha apresenta ainda o valor da *deviance* que indica, quanto menor seu valor for, a validade do modelo conforme apresentado no capítulo 2 desta dissertação.

Tabela (4.1)
Resultado da regressão de Y_1 por modelo linear generalizado

Coeficiente	Estimativa	Desvio-padrão	Z-valor	Pr(>z)
Intercessão	2,4270	0,0332	73,053	~100%
CL _{GERALDO}	-0,1145	0,0484	-2,366	99%
CL _{VALDECIR}	-0,1071	0,0483	-2,217	99%
UM _{ENXUTO}	-0,0949	0,0481	-1,971	99%
UM _{REBAIXADO}	0,0110	0,0469	0,234	
CL _{GERALDO} XUM _{ENXUTO}	0,1096	0,0691	1,586	
CL _{VALDECIR} XUM _{ENXUTO}	0,1545	0,0686	2,253	99%
CL _{GERALDO} XUM _{REBAIXADO}	-0,0855	0,0690	-1,238	
CL _{VALDECIR} XUM _{REBAIXADO}	-0,0036	0,0681	-0,053	
Deviance	1088,533			

Portanto, com base na tabela (4.1) a equação de regressão de Y_1 por modelos lineares generalizados é:

$$\eta_1 = 2,4270 - 0,1145 \times CL_{Geraldo} - 0,1071 \times CL_{Valdecir} - 0,0949 \times UM_{Enxuto} + 0,1545 \times CL_{Valdecir} \times UM_{Enxuto} \quad (4.1)$$

Retornando à questão das variáveis *dummy*, observe-se que, tomando como exemplo o caso do fator classificador (CL), a presença do classificador Geraldo reduz o termo independente em 0,1145. Pois, segundo a explicação dada no capítulo anterior, todos os coeficientes das regressões que têm níveis discretos podem assumir unicamente os valores 1 ou 0. Assim, se o classificador Geraldo for 1, significa que ele está presente e altera, reduzindo em 0,1145 a ação do classificador Ademir representado, junto com os demais fatores do grupo de controle no termo independente. Quando, portanto, a variável $CL_{Geraldo}$ for 1, a variável $CL_{Valdecir}$ deve ser zero.

Se todas as variáveis da equação (4.1) forem zero, a equação resulta no termo independente. Este resultado é o grupo de controle, que neste caso - já que apenas os fatores classificador (CL) e umidade (UM) resultaram em significativos - implica que o nível do fator classificador é Ademir (codificado como -1) e o nível umidade é molhado (codificado como -1). A presença de qualquer outro nível, caso de qualquer variável da equação (4.1) assumir o valor 1, significa o efeito que esta presença exerce sobre os níveis dos fatores do grupo de controle, e por isto, estes níveis presentes na equação (4.1) são chamados grupo de tratamento. Mais detalhes da regressão por variáveis *dummy* encontra-se em Gujarati (2006).

A tabela (4.2) e a equação (4.2) apresentam os resultados da regressão por modelos lineares generalizados da resposta Y_2 .

Tabela (4.2)
Resultado da regressão de Y_2 por modelo linear generalizado

Coeficiente	Estimativa	Desvio-padrão	Z-valor	Pr(>z)
Intercessão	2,1432	0,0394	54,346	~100%
PD _{NORTE}	0,1415	0,0259	5,455	~100%
CL _{GERALDO}	-0,1433	0,0542	-2,645	99,9%
CL _{VALDECIR}	-0,1339	0,0540	-2,478	99%
UM _{ENXUTO}	-0,1185	0,0538	-2,201	99%
UM _{REBAIXADO}	0,0135	0,0520	0,260	
CL _{GERALDO} × UM _{ENXUTO}	0,1372	0,0775	1,769	95%

Tabela (4.2)
Resultado da regressão de Y_2 por modelo linear generalizado (continuação)

Coefficiente	Estimativa	Desvio-padrão	Z-valor	Pr(>z)
$CL_{VALDECIR} \times UM_{ENXUTO}$	0,1935	0,0768	2,519	99%
$CL_{GERALDO} \times UM_{REBAIXADO}$	-0,1092	0,0775	-1,408	
$CL_{VALDECIR} \times UM_{REBAIXADO}$	-0,0042	0,0762	-0,056	
Deviance	1667,457			

$$\begin{aligned} \eta_2 = & 2,1432 + 0,1415 \times PD_{Norte} - 0,1433 \times CL_{Geraldo} \\ & - 0,1339 \times CL_{Valdecir} - 0,1185 \times UM_{Enxuto} \\ & + 0,1372 \times CL_{Geraldo} \times UM_{Enxuto} + 0,1935 \times CL_{Valdecir} \times UM_{Enxuto} \end{aligned} \quad (4.2)$$

O mesmo é também apresentado para a regressão de Y_3 .

Tabela (4.3)
Resultado da regressão de Y_3 por modelo linear generalizado

Coefficiente	Estimativa	Desvio-padrão	Z-valor	Pr(>z)
Intercessão	2,0458	0,0461	44,385	~100%
$CL_{GERALDO}$	-0,1057	0,0674	-1,567	
$CL_{VALDECIR}$	-0,1163	0,0672	-1,732	95%
PD_{NORTE}	0,0959	0,0462	2,076	99%
UM_{ENXUTO}	-0,1349	0,0574	-2,348	99%
$UM_{REBAIXADO}$	0,0153	0,0553	0,276	
$CL_{GERALDO} \times PD_{NORTE}$	-0,1133	0,0680	-1,667	95%
$CL_{VALDECIR} \times PD_{NORTE}$	-0,0705	0,0667	-1,057	
$CL_{GERALDO} \times UM_{ENXUTO}$	0,1564	0,0830	1,885	95%
$CL_{VALDECIR} \times UM_{ENXUTO}$	0,2207	0,0820	2,689	99,9%
$CL_{GERALDO} \times UM_{REBAIXADO}$	-0,1263	0,0830	-1,521	
$CL_{VALDECIR} \times UM_{REBAIXADO}$	-0,0046	0,0814	-0,056	
Deviance	1720,766			

$$\begin{aligned} \eta_3 = & 2,0458 - 0,1163 \times CL_{Valdecir} + 0,0959 \times PD_{Norte} \\ & - 0,1349 \times UM_{Enxuto} - 0,1133 \times CL_{Geraldo} \times PD_{Norte} \\ & + 0,1564 \times CL_{Geraldo} \times UM_{Enxuto} + 0,2207 \times CL_{Valdecir} \times UM_{Enxuto} \end{aligned} \quad (4.3)$$

Assim, completam-se todas as regressões de Y_i pelo método dos modelos lineares generalizados. Como dito anteriormente, admitiu-se que as respostas do experimento de Arriba (2005) acrescidas do valor +10 configuraram uma distribuição de Poisson. Esta hipótese foi ainda novamente confirmada pelos testes de ajustamento de Kolmogorov-Smirnov e Anderson-Darling que indicaram, com aproximadamente 100% de confiança, que não se pode rejeitá-la. No entanto, por precaução a bem da correta regressão a ser usada no caso do experimento de Arriba (2005), são apresentadas também as regressões das mesmas respostas pelo método da quase-verossimilhança, tal qual feito imediatamente acima.

As tabelas (4.4), (4.5) e (4.6) e as equações (4.4), (4.5) e (4.6) apresentam os resultados das regressões por quase-verossimilhança das respostas Y_1 , Y_2 e Y_3 , respectivamente de Arriba (2005). O procedimento foi executado pelo *software* estatístico R e as instruções de uso encontram-se no apêndice A.

Tabela (4.4)
Resultado da regressão de Y_1 por quase-verossimilhança

Coefficiente	Estimativa	Desvio-padrão	Z-valor	Pr(>z)
Intercessão	2,4270	0,0393	61,770	~100%
UM _{ENXUTO}	-0,0949	0,0569	-1,666	95%
UM _{REBAIXADO}	0,0111	0,0554	0,198	
CL _{GERALDO}	-0,1145	0,0572	-2,000	99%
CL _{VALDECIR}	-0,1071	0,0571	-1,875	95%
UM _{ENXUTO} × CL _{GERALDO}	0,1096	0,0817	1,341	
UM _{REBAIXADO} × CL _{GERALDO}	-0,0855	0,0817	-1,047	
UM _{ENXUTO} × CL _{VALDECIR}	0,1545	0,0811	1,905	95%
UM _{REBAIXADO} × CL _{VALDECIR}	-0,0036	0,0806	-0,045	
Deviance	1088,533			

$$\begin{aligned} \eta_1 = & 2,4270 - 0,0949 \times UM_{Enxuto} - 0,1145 \times CL_{Geraldo} \\ & - 0,1071 \times CL_{Valdecir} + 0,1545 \times UM_{Rebaixado} \times CL_{Valdecir} \end{aligned} \quad (4.4)$$

Tabela (4.5)
Resultado da regressão de Y_2 por quase-verossimilhança

Coeficiente	Estimativa	Desvio-padrão	Z-valor	Pr(>z)
Intercessão	2,1432	0,0521	41,114	~100%
PD _{NORTE}	0,1415	0,0343	4,127	~100%
UM _{ENXUTO}	-0,1185	0,0711	-1,665	95%
UM _{REBAIXADO}	0,0135	0,0688	0,197	
CL _{GERALDO}	-0,1433	0,0716	-2,001	99%
CL _{VALDECIR}	-0,1339	0,0714	-1,875	95%
UM _{ENXUTO} ×CL _{GERALDO}	0,1372	0,1025	1,338	
UM _{REBAIXADO} ×CL _{GERALDO}	-0,1092	0,1025	-1,065	
UM _{ENXUTO} ×CL _{VALDECIR}	0,1935	0,1015	1,906	95%
UM _{REBAIXADO} ×CL _{VALDECIR}	-0,0042	0,1007	-0,042	
Deviance	1667,457			

$$\begin{aligned}
 \eta_2 = & 2,1432 + 0,1415 \times PD_{\text{Norte}} - 0,1185 \times UM_{\text{Enxuto}} \\
 & - 0,1433 \times CL_{\text{Geraldo}} - 0,1339 \times CL_{\text{Valdecir}} \\
 & + 0,1935 \times UM_{\text{Enxuto}} \times CL_{\text{Valdecir}}
 \end{aligned} \tag{4.5}$$

Tabela (4.6)
Resultado da regressão de Y_3 por quase-verossimilhança

Coeficiente	Estimativa	Desvio-padrão	Z-valor	Pr(>z)
Intercessão	2,0949	0,0503	41,643	~100%
CL _{GERALDO}	-0,1634	0,0742	-2,201	99%
CL _{VALDECIR}	-0,1526	0,0740	-2,062	99%
UM _{ENXUTO}	-0,1349	0,0737	-1,831	95%
UM _{REBAIXADO}	0,0153	0,0709	0,215	
CL _{GERALDO} ×UM _{ENXUTO}	0,1564	0,1064	1,469	
CL _{VALDECIR} ×UM _{ENXUTO}	0,2207	0,1053	2,096	95%
CL _{GERALDO} ×UM _{REBAIXADO}	-0,1263	0,1065	-1,186	
CL _{VALDECIR} ×UM _{REBAIXADO}	-0,0046	0,1043	-0,044	

Tabela (4.6)
Resultado da regressão de Y_3 por quase-verossimilhança (continuação)

Coeficiente	Estimativa	Desvio-padrão	Z-valor	Pr(>z)
Deviance	1725,482			

$$\begin{aligned} \eta_3 = & 2,0949 - 0,1634 \times CL_{Geraldo} - 0,1526 \times CL_{Valdecir} \\ & - 0,1349 \times UM_{Enxuto} + 0,2207 \times CL_{Valdecir} \times UM_{Enxuto} \end{aligned} \quad (4.6)$$

A decisão entre usar-se de agora em diante as regressões geradas pelos modelos lineares generalizados ou pela quase-verossimilhança será decidida pelos respectivos valores da função de desvio, ou *deviance*. Conforme Dobson (1990) valores mais elevados da *deviance* sugerem modelos mais fracos na descrição dos dados, conforme também se verifica pela equação (2.24). Comparando-se par a par as regressões realizadas, observa-se pela última linha das tabelas (4.1) a (4.6) que os valores da *deviance* para as duas regressões de Y_1 são iguais a 1080,533. O mesmo acontece com Y_2 : 1667,457. Apenas para Y_3 há uma ligeira diferença a favor da regressão por modelos lineares generalizados que apresentou *deviance* de 1720,766 contra 1725,482 do método da quase-verossimilhança.

Estas proximidades de valores de *deviance* obviamente não são coincidências, mas uma justificativa a mais para confirmar-se que a distribuição das respostas do experimento de Arriba (2005) atendem a uma distribuição de Poisson. Desta forma, as duas otimizações que serão executadas a seguir consideram os coeficientes gerados pelas regressões do método dos modelos lineares generalizados expressos pelas equações (4.1), (4.2) e (4.3).

4.2 A OTIMIZAÇÃO DO EXPERIMENTO COM RESPOSTA DE POISSON

Na descrição dos modelos lineares generalizados, em dado momento do procedimento de regressão, tem que se escolher uma função de ligação. No presente caso esta função de ligação foi $\eta = \ln(\mu)$, devido à ligação canônica recomendada pela tabela (2.2). Como η é o resultado de um logaritmo, as predições obtidas a partir das equações acima serão dadas por:

$$\mu_1 = e^{\eta_1} \quad (4.7)$$

$$\mu_2 = e^{\eta_2} \quad (4.8)$$

$$\mu_3 = e^{\eta_3} \quad (4.9)$$

Dando prosseguimento, o próximo passo é modelar os desvios-padrão obtidos em cada tratamento do experimento. Esta modelagem foi feita também por modelos lineares generalizados. No entanto, uma vez que os ajustamentos por Kolmogorov-Smirnov e Anderson-Darling no *software* Statfit indicaram todos os desvios-padrão atendendo à distribuição normal de probabilidade, a função de ligação selecionada foi $\eta = \mu$. Além do mais, devido à condição de variáveis discretas dos níveis dos fatores, esta regressão considerou variáveis *dummy*. A descrição deste procedimento está detalhado no apêndice e foi executado com o *software* estatístico R. As equações (4.10), (4.11) e (4.12) apresentam os resultados das três regressões.

$$\begin{aligned} \sigma_1 = & 3,850 + 1,550 \times UM_{Enxuto} - 1,375 \times UM_{Rebaixado} \\ & - 0,667 \times UM_{Enxuto} \times CL_{Geraldo} + 0,675 \times UM_{Rebaixado} \times CL_{Geraldo} \\ & - 0,900 \times UM_{Enxuto} \times CL_{Valdecir} + 0,800 \times UM_{Rebaixado} \times CL_{Valdecir} \end{aligned} \quad (4.10)$$

$$\begin{aligned} \sigma_2 = & 3,325 + 1,375 \times UM_{Enxuto} - 1,550 \times UM_{Enxuto} \times CL_{Geraldo} \\ & - 1,800 \times UM_{Enxuto} \times CL_{Valdecir} - 1,050 \times UM_{Enxuto} \times CL_{Valdecir} \end{aligned} \quad (4.11)$$

$$\begin{aligned} \sigma_3 = & 2,475 + 1,600 \times UM_{Enxuto} + 1,050 \times CL_{Geraldo} \\ & - 1,425 \times UM_{Enxuto} \times CL_{Geraldo} - 1,525 \times UM_{Enxuto} \times CL_{Valdecir} \end{aligned} \quad (4.12)$$

Os valores de *deviance* das equações (4.10), (4.11) e (4.12) são, respectivamente, 7,529, 7,530 e 8,538. As regressões que correspondem às equações acima poderiam bem ter sido realizadas pelo método dos mínimos quadrados e conduziriam aos mesmos resultados. No entanto, optou-se pelos modelos lineares generalizados justamente porque desta forma gerariam-se os valores de *deviance* já utilizados anteriormente.

Da mesma forma que para os modelos de desvio-padrão, são gerados os modelos para as probabilidades de classificação errada (CE):

$$\begin{aligned}
CE_1 &= 21,25 - 10,000 \times UM_{Rebaixado} - 10,000 \times CL_{Geraldo} \\
&- 7,500 \times CL_{Valdecir} + 11,250 \times UM_{Enxuto} \times CL_{Geraldo} \\
&+ 11,250 \times UM_{Enxuto} \times CL_{Valdecir} + 11,250 \times UM_{Rebaixado} \times CL_{Valdecir}
\end{aligned} \tag{4.13}$$

$$\begin{aligned}
CE_2 &= 5,556 \times PD_{Norte} - 1,667 \times PD_{Norte} \times UM_{Enxuto} \\
&- 1,667 \times PD_{Norte} \times CL_{Geraldo} - 2,500 \times UM_{Enxuto} \times CL_{Geraldo}
\end{aligned} \tag{4.14}$$

$$CE_3 = 0 \tag{4.15}$$

A equação (4.15) resultou em zero porque todos as réplicas do experimento resultaram em nenhum erro de classificação errada. A equação (4.14) não tem termo independente, isto significa que o grupo de controle, a saber neste caso, procedência no nível centro (codificado a -1), umidade molhado (codificado a -1) e classificador Ademir (codificado a -1) não são significativos. Os valores de *deviance* para as equações (4.13) e (4.14) são, respectivamente, 781,250 e 11,111.

O mesmo procedimento foi adotado para a probabilidade de alarme falso (AF):

$$AF_1 = 6,250 + 15,000 \times UM_{Enxuto} - 12,500 \times UM_{Enxuto} \times CL_{Valdecir} \tag{4.16}$$

$$\begin{aligned}
AF_2 &= 10,000 + 15,000 \times UM_{Enxuto} + 13,750 \times CL_{Geraldo} \\
&- 17,500 \times UM_{Enxuto} \times CL_{Geraldo} - 20,000 \times UM_{Enxuto} \times CL_{Valdecir}
\end{aligned} \tag{4.17}$$

$$\begin{aligned}
AF_3 &= 10,000 + 16,250 \times UM_{Enxuto} + 13,750 \times CL_{Geraldo} \\
&- 18,750 \times UM_{Enxuto} \times CL_{Geraldo} - 21,250 \times UM_{Enxuto} \times CL_{Valdecir}
\end{aligned} \tag{4.18}$$

Os valores de *deviance* das equações (4.16), (4.17) e (4.18) são, respectivamente, 1131,250, 2268,750 e 2287,500.

Os subscritos das equações de regressão apresentadas nas equações (4.10) a (4.18) referem-se, naturalmente a Y_1 , Y_2 e Y_3 e estes, por sua vez, representam as respostas encontradas nos estágios de curtimento do couro em *wet-blue*, semi-acabado e acabado, nesta ordem. Assim, dando seguimento ao procedimento de otimização proposto, ponderam-se as equações de regressão segundo a equação (3.2) e tabela (3.4). Os números nos denominadores das equações abaixo são as médias gerais de cada grandeza, de forma a tornar o resultado adimensional.

$$\begin{aligned}
Z_1 = & 0,15 \times \frac{[Eq(4.7)]}{10,45} + 0,20 \times \frac{[Eq(4.11)]}{3,72} + \\
& + 0,40 \times \frac{[Eq(4.14)]}{15,97} + 0,25 \times \frac{[Eq(4.17)]}{10,69}
\end{aligned} \quad (4.19)$$

$$\begin{aligned}
Z_2 = & 0,15 \times \frac{[Eq(4.8)]}{8,30} + 0,20 \times \frac{[Eq(4.12)]}{3,77} + \\
& + 0,40 \times \frac{[Eq(4.15)]}{2,22} + 0,25 \times \frac{[Eq(4.18)]}{19,31}
\end{aligned} \quad (4.20)$$

$$Z_3 = 0,15 \times \frac{[Eq(4.9)]}{7,25} + 0,20 \times \frac{[Eq(4.13)]}{3,37} + 0,25 \times \frac{[Eq(4.19)]}{19,44} \quad (4.21)$$

As equações (4.19) a (4.21) são reunidas utilizando a equação (3.3) e tabela (3.5). Observe-se que na equação Z_3 o termo correspondente à equação (4.15) não aparece, pois não houve erro de classificação errada nesta etapa do experimento. O próximo passo da otimização de Arriba (2005) reúne as equações (4.19), (4.20) e (4.21) na ponderação Z_{Total} de acordo com a tabela (3.5) e a equação (3.3)

$$Z_{Total} = 0,35 \times [Eq(4.19)] + 0,15 \times [Eq(4.20)] + 0,50 \times [Eq(4.21)] \quad (4.22)$$

Os resultados das equações (4.19), (4.20) e (4.21) merecem um exame mais detido quanto ao valor calculado pela equação (4.22) que de fato representa o arranjo ótimo dos níveis, segundo o método de otimização utilizado por Arriba (2005).

O valor desejado para as respostas Y [equações (4.7) a (4.9)], devido ao efeito de translação, é +10. Os valores desejados para σ [equações (4.10) a (4.12)], CE [equações (4.13) a (4.15)] e AF [equações (4.16) a (4.18)] são zero. Substituindo-se esses valores alvo nas equações (4.19) a (4.21) e, posteriormente, inserindo estes resultados na equação (4.22), é desejável obter-se um valor alvo para Z_{Total} o mais próximo possível de 0,1808. Em Arriba (2005) a proximidade desejada era ao valor zero. Aqui, devido ao efeito da translação em +10, será o valor 0,1808.

A equação (4.22) é calculada para todas as 36 combinações de níveis dos fatores experimentais. O resultado que mais se aproxima de 0,1808 é, assim, o arranjo de níveis que conduz à otimização do processo, conforme proposto por Arriba (2005). A tabela (4.7) apresenta os resultados dos cálculos destas 36 combinações e mantém a ordem de tratamentos

da tabela (3.6), adaptada de Arriba (2005), mas também, tal qual naquela tabela, destaca com fundo sombreado os três melhores valores que se aproximam do alvo 0,1808.

Tabela (4.7)
 Z_{Total} para cada nível dos fatores. Destaque para os valores ótimos

Tratamento	PD	UM	TE	CL	Z_{Total}	Tratamento	PD	UM	TE	CL	Z_{Total}
1	-1	-1	-1	-1	0,654	19	+1	-1	-1	-1	0,815
2	-1	-1	-1	+1	0,570	20	+1	-1	-1	+1	0,731
3	-1	-1	-1	0	0,703	21	+1	-1	-1	0	0,810
4	-1	-1	+1	-1	0,654	22	+1	-1	+1	-1	0,815
5	-1	-1	+1	+1	0,570	23	+1	-1	+1	+1	0,731
6	-1	-1	+1	0	0,703	24	+1	-1	+1	0	0,810
7	-1	0	-1	-1	0,980	25	+1	0	-1	-1	1,095
8	-1	0	-1	+1	0,670	26	+1	0	-1	+1	0,786
9	-1	0	-1	0	0,854	27	+1	0	-1	0	0,916
10	-1	0	+1	-1	0,980	28	+1	0	+1	-1	1,095
11	-1	0	+1	+1	0,670	29	+1	0	+1	+1	0,786
12	-1	0	+1	0	0,854	30	+1	0	+1	0	0,916
13	-1	+1	-1	-1	0,540	31	+1	+1	-1	-1	0,702
14	-1	+1	-1	+1	0,562	32	+1	+1	-1	+1	0,723
15	-1	+1	-1	0	0,602	33	+1	+1	-1	0	0,709
16	-1	+1	+1	-1	0,540	34	+1	+1	+1	-1	0,702
17	-1	+1	+1	+1	0,562	35	+1	+1	+1	+1	0,723
18	-1	+1	+1	0	0,602	36	+1	+1	+1	0	0,709

A tabela (4.7) destaca os tratamentos que mais se aproximam do alvo 0,1808 e foi comparada aos valores sombreados da tabela (3.6). Na tabela (4.7) há um empate entre os dois melhores tratamentos, que são os de números 13 e 16. Os níveis codificados do tratamento 13

indicam que os níveis dos fatores que conduzem à otimização do processo são: procedência centro (nível -1), umidade rebaixado (nível +1), método de classificação estático (nível -1) e classificador Ademir (nível -1). Já o outro melhor tratamento, o de número 16, sofre apenas alteração em relação ao tratamento 13 no nível do fator método de classificação: dinâmico (nível +1). Este resultado era esperado e coerente com o fato de que em nenhuma das equações de regressão indicou significância do fator método de classificação (TE). O terceiro melhor tratamento, também sombreado na tabela (4.7), é o de número 14: procedência centro (nível -1), umidade rebaixado (nível +1), método de classificação estático (nível -1) e classificador Geraldo (nível 0).

Em Arriba (2005), o melhor tratamento experimental também é o de número 13 conforme observa-se no sombreado da tabela (3.6). Mas, a concordância com a tabela (4.7) cessa aí. Esta diferença deve-se naturalmente à regressão inadequada usada por Arriba (2005) e também pela desconsideração naquele trabalho de que os fatores são variáveis *dummy*, como já exposto anteriormente.

4.2 A OTIMIZAÇÃO ATRAVÉS DE ÍNDICES DE CAPACIDADE DE PROCESSOS

A estratégia de otimização adotada por Ch'ng, Quah e Low (2005) para otimização multiresposta foi reproduzida nesta seção para o experimento de Arriba (2005) considerando a regressão de Poisson para as respostas Y_i .

O procedimento geral de otimização tornou-se muito mais simples do que aquele apresentado na seção 4.2. Para operacionalizá-lo, foi criada uma planilha de otimização apresentada no apêndice B de forma a atender ao método do gradiente generalizado reduzido. Primeiramente, foram programadas na planilha as equações (4.7) a (4.9) para a regressão de Poisson; o mesmo se fez para as equações (4.10) a (4.12) para as regressões dos desvios-padrão; assim como para as equações (4.13) a (4.15), para as probabilidades de classificação errada; e por fim, as equações (4.16) a (4.18), para as probabilidades de alarme falso. Todas as equações estão vinculadas às células que contêm os níveis dos fatores; assim, se o nível do fator variar, todas as equações sofrerão alterações em seus resultados.

Na seqüência, programou-se a equação (2.61), que calcula o índice de $TotalC_{pm}$, a cada etapa do processo de curtimento de couro, identificados pelos sobrescritos 1, 2 e 3 respectivamente nas equações (4.23) a (4.25). Este cálculo é feito como apresentado abaixo:

$$TotalC_{pm}^1 = 0,35 \times \frac{\min(20-10;10-0)}{3 \times \sqrt{([Eq(4.7)]-10)^2 + [Eq(4.10)]^2}} +$$

$$+ 0,40 \times \frac{100-0}{3 \times [Eq(4.13)]} + 0,25 \times \frac{100-0}{3 \times [Eq(4.16)]} \quad (4.23)$$

$$TotalC_{pm}^2 = 0,35 \times \frac{\min(20-10;10-0)}{3 \times \sqrt{([Eq(4.8)]-10)^2 + [Eq(4.11)]^2}} +$$

$$+ 0,40 \times \frac{100-0}{3 \times [Eq(4.14)]} + 0,25 \times \frac{100-0}{3 \times [Eq(4.17)]} \quad (4.24)$$

$$TotalC_{pm}^3 = 0,35 \times \frac{\min(20-10;10-0)}{3 \times \sqrt{([Eq(4.9)]-10)^2 + [Eq(4.12)]^2}} +$$

$$+ 0,25 \times \frac{100-0}{3 \times [Eq(4.18)]} \quad (4.25)$$

Os valores de ponderação e_i na equação (2.61) são, respectivamente, 0,35, 0,40 e 0,25, conforme indicado na tabela (3.4). Note-se que os pesos para Y_i (0,15) e para o desvio-padrão (0,20) foram somados (ver primeiro termo do conjunto de equações acima). Com este procedimento, tanto o peso de Y_i quanto o peso do desvio-padrão assumiram uma só importância, não sendo possível a diferenciação entre eles, pois seus valores de ponderação foram somados.

O alvo para o valor de Y_i é 10, sendo $USL=20$ e $LSL=0$, isto é, os valores máximo e mínimo possíveis de serem obtidos a partir das equações (4.7) a (4.9). Para as regressões de ambas as probabilidades, os limites de especificação são unilaterais; para tanto, as expressões de C_{pm} usadas correspondem à equação (2.62). O valor do alvo unilateral é o limite natural da probabilidade, ou seja, 100.

Os resultados das equações (4.23) a (4.25) são então ponderados numa função Z segundo os pesos da tabela (3.5) e os valores de ponderação da equação (3.3):

$$Z = 0,35 \times [Eq(4.23)] + 0,15 \times [Eq(4.24)] + 0,50 \times [Eq(4.25)] \quad (4.26)$$

A planilha de otimização solicita as restrições que indicam os valores que os níveis dos fatores podem ocupar. A primeira restrição que deve ser carregada é que todos os níveis são valores inteiros e binários. Ou seja, podem assumir unicamente os números 0 ou 1. Isto é decorrência do fato de que os níveis são discretos e a regressão, portanto, os considerou como variáveis *dummy*. A próxima restrição atende à necessidade de informar à planilha de otimização que se um nível de determinado fator estiver presente, ou outro nível deverá obrigatoriamente estar ausente. Em outras palavras, se num mesmo fator há um nível assumindo o valor 1, o outro nível deverá ser zero. Para o caso presente, esta restrição é atendida com a inserção do conjunto de inequações (4.27):

$$\begin{aligned} CL_{Geraldo} + CL_{Valdecir} &\leq 1 \\ UM_{Enxuto} + UM_{Rebaixado} &\leq 1 \end{aligned} \quad (4.27)$$

Observando o conjunto de inequações (4.27) conclui-se que se o classificador presente for Geraldo, este assumirá o valor 1, impedindo que o classificador Valdecir também esteja presente. Ou ainda, se Geraldo e Valdecir estiverem ausentes, ambos assumindo o valor zero, isto indica a presença do classificador Ademir que está no grupo de controle. Raciocínio idêntico vale para o fator umidade. O fator procedência, que também está nas equações de regressão, prescinde da inclusão nestas restrições do conjunto de inequações (4.27), uma vez que apresenta-se em apenas dois níveis: centro e norte.

Por fim, a função objetivo deve ser maximizada, pois quanto maior o valor de $TotalC_{pm}$ melhor. Solicita-se então ao algoritmo do gradiente generalizado reduzido a execução da maximização. O resultado indica a melhor combinação de níveis que otimiza o experimento, são eles: procedência centro (nível -1); umidade rebaixado (nível +1); método de classificação não está presente em nenhuma equação de regressão e, portanto, é indiferente à otimização podendo ser estático (nível -1) ou dinâmico (nível +1); e, por fim, classificador Geraldo (nível 0). Nas tabelas (3.6) e (4.7) estes níveis correspondem aos tratamentos 15 e 18.

A diferença nas respostas desta otimização em relação à otimização realizada na seção anterior está apenas na seleção do fator classificador. Enquanto pelo método anterior o classificador selecionado foi Ademir, no presente método da maximização de $TotalC_{pm}$ o fator classificador selecionado foi Geraldo. Todos os demais fatores coincidem.

4.3 ANÁLISE DOS RESULTADOS

Dois métodos de otimização foram apresentados. No primeiro (seção 4.2), as variáveis de regressão categóricas foram submetidas a uma modelagem através dos modelos lineares generalizados, corrigindo a violação de pressuposto incorrida por Arriba (2005), que considerou respostas categóricas como contínuas. Antes, na seção 4.1, fez-se a opção entre usar-se as regressões das respostas Y_i por modelos lineares generalizados ou por quase-verossimilhança, os resultados de *deviance*, embora muito próximos entre si, conduziram a optar-se pela primeira regressão.

A otimização, neste primeiro caso, convergiu para o mesmo tratamento experimental. Assim, ainda que Arriba (2005) tenha violado pressupostos, seus resultados não geraram erros marcantes a ponto de incorrer em erro para o melhor tratamento, embora na hierarquia dos demais tratamentos esta concordância se perca. Este primeiro método de otimização é de difícil implementação por solicitar muitos cálculos e ainda a hierarquização das respostas. Ele não testa automaticamente o campo dos níveis em busca da melhor resposta, mas sim gera todos os resultados possíveis e o analista seleciona o melhor entre eles. A técnica somente funcionou neste caso porque os fatores eram todos valores discretos, portanto gerando um número finito de tratamentos experimentais.

O segundo método de otimização, ao contrário, é bem mais simples e exige muito menos esforço computacional. A planilha de otimização permite alterar as variáveis para simular outros cenários. Além disso, qualquer um - ou todos - os fatores podem ou não ser contínuos ou variáveis; o algoritmo de otimização apenas solicita que isso seja informado nas restrições da programação linear.

Comparando-se os resultados obtidos entre cada método de otimização conclui-se que apenas há divergência quanto ao fator classificador em termos de seus níveis ideais. Para os demais fatores, os níveis identificados como ótimos foram idênticos.

É justamente onde o fator tende a uma maior subjetividade que houve discordância entre os dois métodos de otimização: classificador. Pois, os demais fatores são facilmente diferenciados e bem delimitados. O fator classificador é obviamente muito subjetivo, pois

ainda que experiente e bem treinado, este fator ainda está sujeito ao estado de humor, a pressões do ambiente, ao cansaço etc.

Capítulo 5

Neste capítulo serão apresentados os principais resultados obtidos face aos objetivos propostos e também serão dadas sugestões para trabalhos futuros.

5.2 CONCLUSÕES

Um dos propósitos desta dissertação foi a otimização de processos avaliados por múltiplas variáveis de respostas, uma situação bastante usual no meio produtivo. Otimizar, neste contexto, significa identificar em quais níveis os fatores que influenciam as respostas do processo alcançam um ponto ótimo. O ponto ótimo está associado a um desempenho exigido pelo cliente e em geral expresso em termos de menor custo, menor refugo, maior rendimento, menor tempo de operação, menor quantidade de matéria-prima, maior aproveitamento dos recursos produtivos, redução da variabilidade das respostas, aproximação a uma certa medida de qualidade ou alvo, etc.

Em diversas situações a resposta de um processo é uma variável discreta ou categórica. Isto significa que ela é expressa como, por exemplo, uma contagem ou, ainda, como conforme ou não-conforme. Em resumo, existe um número finito de possibilidades que a resposta pode assumir. Ao contrário, se esta resposta fosse contínua (por exemplo, a medida de uma temperatura), infinitos valores poderiam ser observados. Assim, além de respostas múltiplas que otimizam um processo, este trabalho lidou com respostas especiais, ditas categóricas. Nem sempre a modelagem de respostas categóricas pode ser feita através do método dos mínimos quadrados, já que este método pressupõe observações normalmente distribuídas. Neste caso, a modelagem deve ser feita pelos modelos lineares generalizados (MLG). Este foi também o caso desta dissertação.

Nesta dissertação foi apresentada uma reanálise de um experimento envolvendo respostas categóricas, originalmente apresentado em Arriba (2005). Na análise proposta pelo autor, as respostas foram tratadas como contínuas e modeladas através de regressão linear múltipla. Nesta dissertação, as respostas categóricas do experimento em Arriba (2005) foram modeladas através dos modelos lineares generalizados e por quase-verossimilhança, e ainda

considerando que os níveis dos fatores são variáveis discretas, ou seja, devem ser modelados como *dummy*. A modelagem em Arriba (2005) tinha por objetivo otimizar o processo em estudo. Nesta dissertação, repetiu-se o procedimento de otimização proposto pelo autor, utilizando, entretanto, os modelos lineares generalizados obtidos para as respostas do experimento. Algumas divergências foram encontradas nos resultados das otimizações, as quais são apontadas no Capítulo 4.

Foi ainda conduzida uma nova forma de otimização multiresposta do experimento em Arriba (2005), utilizando-se índices de capacidade de processos, onde as mesmas equações de regressão utilizadas no método anterior foram novamente usadas. O novo procedimento de otimização levou a um ponto ótimo distinto daquele encontrado na otimização anteriormente realizada apenas no que se refere ao fator classificador.

O primeiro objetivo principal desta dissertação era otimizar um processo multiresposta que quando pelo menos uma das variáveis é categórica. O segundo objetivo principal era comparar dois métodos de otimização multiresposta. Foram portanto comparados os métodos de otimização adotado por Arriba (2005) e aquele utilizado por Ch'ng, Quah e Low (2005). Sendo que este segundo mostrou-se bem mais simples do ponto de vista computacional, além de permitir a análise comparativa do comportamento das respostas quando um ou vários níveis dos fatores são alterados na planilha de otimização, o que facilita as interpretações a respeito dos impactos que uma dada alteração afeta a qualidade geral do processo.

Teve-se também como objetivo deste trabalho, a ampliação do conhecimento do método de modelagem através dos modelos lineares generalizados, uma vez que notou-se uma escassa literatura sobre o assunto em Língua Portuguesa, a despeito do notável fato de que o tema foi posto em foco há um bom tempo – Nelder e Wedderburn (1972) – e que sua necessidade de uso é bastante comum na indústria. Possivelmente, a modelagem equivocada através do método dos mínimos quadrados seja bem mais comum do que se supõe.

A modelagem por modelos lineares generalizados é significativamente mais complexa que o método dos mínimos quadrados. Nem todos os *softwares* estatísticos trazem as principais funções de ligação citadas na tabela (2.2), como é o caso do Minitab versão 12.22 e SPSS 10.0. O Apêndice A desta dissertação apresenta o passo a passo de instruções no *software* livre R para a modelagem por modelos lineares generalizados.

5.2 SUGESTÕES PARA TRABALHOS FUTUROS

Arriba (2005) considerou suas respostas como diferenças entre notas de avaliadores em etapas do processo de curtimento do couro. A modelagem foi feita por estas diferenças, no caso deste trabalho foi necessária a translação das respostas para adequar-se à distribuição de Poisson. Há ainda uma outra forma de modelar estas respostas tomando-as como acerto ou erro. Neste caso, apenas duas possibilidades podem acontecer e o processo encaixa-se perfeitamente no modelo de regressão binomial que também é um modelo linear generalizado, bastando que se altere a função de ligação conforme a tabela (2.2). Por simplificar bastante as possibilidades de respostas, pode ser que desta forma encontrem-se valores melhores das funções de desvio (*deviance*) que certificam a adequabilidade do modelo.

Outro experimento pode ser conduzido considerando apenas o fator classificador porém em uma quantidade maior de níveis, por exemplo cinco. Ao longo do experimento devem ser anotadas as impressões que cada classificador formou na seleção da pele, assim poderia-se investigar melhor em que termos ocorrem a subjetividade de classificação.

Referências Bibliográficas

- AGRESTI, A. **Categorical Data Analysis**. 2.Ed. Hoboken, New Jersey: John Wiley & Sons, Inc, 2002.
- ARRIBA, G.D. **Otimização de um processo de classificação de couros wet blue**: Um caso em uma indústria curtidora. Porto Alegre: UFRGS, 2005. Dissertação de Mestrado, Engenharia de Produção, Universidade Federal do Rio Grande do Sul, 2005.
- BISGAARD, S.; FULLER, H. T. **Analysis of factorial experiments with defects or defectives as the response**. *Quality Engineering*. v. 7, n. 2, p. 429-443, 1994.
- BOTHE, D. R. **Measuring process capability**. Techniques and calculations for quality and manufacturing engineers. Nova Iorque: McGraw-Hill, 1997.
- CASTILLO, E. D.; MONTGOMERY, D. C. **A nonlinear programming solution to the dual response problem**. *Journal of Quality Technology*. v. 25, n. 3, p. 162-175, julho 1993.
- CHAN, L. K.; CHENG, S. W.; SPIRING, F. A. **A new measure of process Capability: Cpm**. *Journal of Quality Technology*. v. 25, n. 3, p. 199-204, julho 1988.
- CH'NG, C.K.; QUAH, S.H.; LOW, H.C. **Index Cpm in multiple response optimization**. *Quality Engineering*. v. 17, n. 1, p. 165-171, 2005, janeiro 2005.
- DOBSON, A. J. **An Introduction to Generalized Linear Models**. New York: Chapman & Hall/CRC, 1990.
- DUNCAN, A. J. **Quality Control and Industrial Statistics**. Homewood: Richard D. Irwin INC, 1974.
- EPPRECHT, E. K.; VIEIRA, A. F. C. **Solving Generalized Linear Models with Spreadsheets**. *Brazilian Journal of Operations and Production Management*, 1(1), 2004.
- FAHRMEIR, L.; TUTZ, G. **Multivariate Statistical Modelling Based on Generalized Linear Models**. Nova Iorque: Springer, 1994.
- GIL, A. C. **Como elaborar projetos de pesquisa**. 4. Ed. São Paulo: Atlas, 2002.

GUJARATI, D. N. **Econometria Básica**. Rio de Janeiro: Elsevier, 2006.

ELSAYED, E. A.; RIBEIRO, J. L. **A case study on process optimization using the gradient loss function**. *International Journal of Production Research*, v. 33, n. 12, p. 3233-3248, 1995.

KELTON, D. W.; LAW, A. M. **Simulation Modeling and Analysis**. 2. Ed. Singapura: McGraw-Hill, Inc, 1991.

McCULLAGH, P.; NELDER, J.A. **Generalized Linear Models**. 2 ed. Londres. Chapman and Hall, 1990.

MONTGOMERY, D. C. **Design and Analysis of Experiments**. 6. ed. Hoboken, Nova Jersey: John Wiley & Sons, Inc, 2005.

MONTGOMERY, D. C. **Introduction to Statistical Quality Control**. 4 ed. Hoboken, Nova Jersey: John Wiley & Sons, Inc, 2001.

MONTGOMERY, D. C.; PECK, E. A.; VINING, G. G. **Introduction to Linear Regression Analysis**. 4 ed. Hoboken, Nova Jersey: John Wiley & Sons, Inc, 2006.

MYERS, R. H.; MONTGOMERY, D. C; VINING, G. G. **Generalized Linear Models: with applications in engineering and sciences**. Nova Iorque: John Wiley & Sons, Inc, 2002.

NELDER J. A.; WEDDERBURN, R. W. M. **Generalized Linear Models**. *Journal of the Royal Statistical Society*, Ser. A 135, 370-384. 1972.

PLANTE, R. D. **Process Capability: a Criterion for Optimizing Multiple Response Product and Process Design**. *IIE Transactions*, n. 33, p. 497-509, 2001.

ROESCH, S. M. A. **Projetos de estágio e de pesquisa em administração: guia para estágios, trabalhos de conclusão, dissertações e estudos de caso**. 3. ed. São Paulo: Atlas, 2005.

VIEIRA, A. F. C. **Análise da Média e Dispersão em Experimentos Fatoriais não Replicados para Otimização de Processos Industriais**. Rio de Janeiro: PUC, 2004. Tese de doutorado, Engenharia de Produção, Pontifícia da Universidade Católica, 2004.

Obras Consultadas

HSIEH, K.; TONG, L. **Optimization of Multiple Quality Responses Involving Qualitative and Quantitative Characteristics in IC Manufacturing Using Neural Networks.** *Computers in Industry*, v. 46, p. 1-12, Agosto 2001.

KO, Y.; KIM, K.; JUN, C. **A New Loss Function-Based Method for Multiresponse Optimization.** *Journal of Quality Technology*, v.37, p. 50-60, Janeiro 2005.

KUN-LIN, H et al.. **Optimization of a multiresponse problem in Taguchi's dynamic system.** *Computers & Industrial Engineering*, v. 29, p. 556-571, Janeiro 2005.

PIGNATELLO JR., J. J. **Strategies for Robust Multiresponses Quality Engineering.** *IIE Transactions*, n. 25. p. 5-15, 1993.

ROSS, P. J. **Aplicações das Técnicas de Taguchi na Engenharia de Qualidade.** São Paulo: McGraw-Hill, 1991.

SANT'ANNA, A. M. O. **Método de Orientação à Modelagem de Dados Mensurados em Proporção.** Porto Alegre: UFRGS, 2006. Dissertação de Mestrado, Engenharia de Produção, Universidade Federal do Rio Grande do Sul, 2006.

TAGUCHI, G; ELSAYED, E. A.; HSIANG, T. **Engenharia de Qualidade em Sistemas de Produção.** São Paulo: Mc-Graw-Hill, 1990.

Apêndice A

Neste apêndice são apresentadas as instruções que devem ser dadas ao software R para a modelagem por modelos lineares generalizados segundo a função de ligação de Poisson e normal. Também apresenta-se o passo a passo para a modelagem por quase-verossimilhança.

Para os modelos lineares generalizados que correspondem às equações (4.1), (4.2) e (4.3) as instruções a ser dadas ao R são:

```
read.table("C:/bncc2.txt", header=T, sep="")->bncc1

attach(bncc1)

as.factor(X1)->X1

as.factor(X2)->X2

as.factor(X3)->X3

as.factor(X4)->X4

glm(formula = Y1 ~ X1 + X2 + X3 + X4 + X1*X1 + X1*X2 + X1*X3 + X1*X4 +
X2*X2 + X2*X3 + X2*X4 + X3*X3 + X3*X4 + X4*X4, family = poisson())->M1

summary(M1Q)

deviance(M1Q)
```

A instrução *read.table* (argumentos) carrega um banco de dados que contém os fatores, seus níveis e as respostas de um experimento. O primeiro argumento é o próprio nome do banco de dados que deve estar no formato *.txt*. Atenção: o R não reconhece vírgulas como separador de decimais, deve-se usar o ponto. O argumento *header=T* indica que o banco de dados possui um cabeçalho na primeira linha que nomeia as colunas de dados. O último argumento *sep=""* instrui o *software* que a não há separadores entre os dados além do espaço em branco. Os dados do banco, neste exemplo, são carregados na variável *bncc1*.

Para que os nomes das colunas sejam reconhecidos como variáveis deve ser usado o comando *attach(bncc1)*.

No banco de dados os fatores foram nomeados como X1, X2, X3 e X4 e seus nomes estão no cabeçalho. Como estes fatores são discretos e então as variáveis devem ser consideradas como *dummy*, deve ser usada a instrução *as.factor*(nome da coluna) e carregá-la numa variável. Neste caso, a variável que carrega esta instrução possui o mesmo nome da coluna.

A instrução *glm*(argumentos) indica que a regressão deve ser por modelos lineares generalizados. O primeiro argumento indica a fórmula do modelo com seus fatores e suas interações. No presente caso, usou-se interações entre os fatores até segunda ordem. A segunda instrução indica qual a função de ligação, no presente caso ela é do tipo Poisson. O modelo é carregado na variável M1.

A função *summary*(argumento) apresentará os resultados da regressão com os parâmetros do modelo e seus níveis de confiabilidade. Uma vez selecionados os parâmetros mais significativos, gera-se novamente a instrução *glm*(argumentos) contendo no modelo apenas os fatores selecionados.

A função *deviance*(argumento) apresenta o valor da *deviance* do modelo.

Para um modelo linear generalizado que tem função de ligação normal, como foi o caso dos modelos representados nas equações (4.10) a (4.18), a instrução deve ser a mesma das anteriores, exceto a seguinte linha:

```
glm(formula = Y1 ~ X1 + X2 + X3 + X4 + X1*X1 + X1*X2 + X1*X3 + X1*X4 +
X2*X2 + X2*X3 + X2*X4 + X3*X3 + X3*X4 + X4*X4 , family = gaussian())->M1
```

Identicamente, para a modelagem por quase-verossimilhança a linha que contém o tipo de modelagem a ser usada deve ser:

```
glm(formula = Y1 ~ X1 + X2 + X3 + X4 + X1*X1 + X1*X2 + X1*X3 + X1*X4 +
X2*X2 + X2*X3 + X2*X4 + X3*X3 + X3*X4 + X4*X4 , quasi(link=log,variance="mu" )-
>M1
```

Este foi o caso usado nos modelos das equações (4.4), (4.5) e (4.6).

Apêndice B

A figura (B.1) apresenta a planilha de otimização utilizada para encontrar os níveis ótimos dos fatores do experimento de Arriba (2005). Foi usado o método do índice de capacidade total proposto por Ch'ng, Quah e Low (2005) para a otimização. O algoritmo *Generalized Reduced Gradient* (GRG), ou gradiente generalizado reduzido, é o padrão do *Solver* do Excel® para problemas de programação linear. As letras na primeira linha e os números na primeira coluna servem para referenciar as células.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	
1	Fatores em níveis - variáveis dummy														
2	Fatores	Termo Independ.	PD _{Norte}	CL _{Geraldo}	CL _{Valdecir}	UM _{Exato}	UM _{Rebaixado}	CL _{Geraldo} xU _{M_{Exato}}	CL _{Valdecir} xU _{M_{Exato}}	CL _{Geraldo} xUM _{Rebaixado}	CL _{Valdecir} xUM _{Rebaixado}	CL _{Geraldo} xPD _{Norte}	CL _{Valdecir} xPD _{Norte}	UM _{Exato} xPD _{Norte}	
3	Níveis	1	0	1	0	0	1	0	0	0	0	0	0	0	
4	Regressão de Poisson por GLM das respostas Y														
6	Fatores	Intercessão	PD _{Norte}	CL _{Geraldo}	CL _{Valdecir}	UM _{Exato}	UM _{Rebaixado}	CL _{Geraldo} xU _{M_{Exato}}	CL _{Valdecir} xU _{M_{Exato}}	CL _{Geraldo} xUM _{Rebaixado}	CL _{Valdecir} xUM _{Rebaixado}	CL _{Geraldo} xPD _{Norte}	CL _{Valdecir} xPD _{Norte}	UM _{Exato} xPD _{Norte}	
7	Y1	2,427	0	-0,1145	-0,1071	-0,0949	0	0	0,1545	0	0	0	0	0	
8	Y2	2,1432	0,1415	-0,1433	-0,1339	-0,1185	0	0,1372	0,1935	0	0	0	0	0	
9	Y3	2,0458	0,0959	0	-0,1163	-0,1349	0	0,1564	0,2207	0	0	0	-0,1133	0	
10	Regressão normal por GLM dos desvios-padrão dos Y														
12	Fatores	Intercessão	PD _{Norte}	CL _{Geraldo}	CL _{Valdecir}	UM _{Exato}	UM _{Rebaixado}	CL _{Geraldo} xU _{M_{Exato}}	CL _{Valdecir} xU _{M_{Exato}}	CL _{Geraldo} xUM _{Rebaixado}	CL _{Valdecir} xUM _{Rebaixado}	CL _{Geraldo} xPD _{Norte}	CL _{Valdecir} xPD _{Norte}	UM _{Exato} xPD _{Norte}	
13	DP1	3,85	0	0	0	1,55	-1,375	-0,6667	-0,9	0,675	0,8	0	0	0	
14	DP2	3,325	0	0	0	1,375	0	-1,55	-1,8	0	-1,05	0	0	0	
15	DP3	2,475	0	1,05	0	1,6	0	-1,425	-1,525	0	0	0	0	0	
16	Regressão normal por GLM das probabilidades de Classificação Errada (CE)														
18	Fatores	Intercessão	PD _{Norte}	CL _{Geraldo}	CL _{Valdecir}	UM _{Exato}	UM _{Rebaixado}	CL _{Geraldo} xU _{M_{Exato}}	CL _{Valdecir} xU _{M_{Exato}}	CL _{Geraldo} xUM _{Rebaixado}	CL _{Valdecir} xUM _{Rebaixado}	CL _{Geraldo} xPD _{Norte}	CL _{Valdecir} xPD _{Norte}	UM _{Exato} xPD _{Norte}	
19	CE1	21,25	0	-10	-7,5	0	-10	11,25	11,25	0	11,25	0	0	0	
20	CE2	0	5,556	0	0	0	0	-2,5	0	0	0	-1,667	0	-1,667	
21	CE3	0	0	0	0	0	0	0	0	0	0	0	0	0	
22	Regressão normal por GLM das probabilidades de Alarme Falso (AF)														
24	Fatores	Intercessão	PD _{Norte}	CL _{Geraldo}	CL _{Valdecir}	UM _{Exato}	UM _{Rebaixado}	CL _{Geraldo} xU _{M_{Exato}}	CL _{Valdecir} xU _{M_{Exato}}	CL _{Geraldo} xUM _{Rebaixado}	CL _{Valdecir} xUM _{Rebaixado}	CL _{Geraldo} xPD _{Norte}	CL _{Valdecir} xPD _{Norte}	UM _{Exato} xPD _{Norte}	
25	AF1	6,25	0	0	0	15	0	0	-12,5	0	0	0	0	0	
26	AF2	10	0	13,75	0	15	0	-17,5	-20	0	0	0	0	0	
27	AF3	10	0	13,75	0	16,25	0	-18,75	-21,25	0	0	0	0	0	
28	Resumo de Regressões														
29	Regressão Y	Y1	10,0996	Regressão Desvio-padrão			DP1	3,85	Prob (CE)			CE1	1,2500	Prob (AF)	
30		Y2	7,3883				DP2	3,325				CE2	0,0000		
31		Y3	7,7353				DP3	2,475				CE3	0,0000		
32	Restrições														
33	Z	4,56	PD _{Norte}	Instrução	CL _{Geraldo} + CL _{Valdecir}		Instrução: Binário e soma dos níveis		UM _{Exato} + UM _{Rebaixado}		Instrução: soma do				
34			0	Binário	1		<= 1		1		<=				
35	TotalCpm(1)	12,1079													
36	TotalCpm(2)	0,4949	#DIV/0!												
37	TotalCpm(3)	0,4945													

Figura (B.1) – Planilha de otimização

O campo de células compreendido entre B7:O9 contém os parâmetros da regressão de Poisson apresentados nas equações (4.1), (4.2) e (4.3). Assim como o campo de células B13:O15 contém os parâmetros da regressão do desvio-padrão. O mesmo vale para o campo

B19:O21 para a regressão da probabilidade de classificação errada; e o campo B25:O27 para a probabilidade de alarme falso.

A linha 3 contém as variáveis do problema que são os níveis dos fatores. Estas variáveis são designadas no *Solver* como inteiras e binárias (0 ou 1). Estas variáveis são têm estas características por serem *dummy*.

Os resultados das regressões de Y_1 , Y_2 e Y_3 estão respectivamente nas células C29, C30 e C31. Este resultado é apenas o produto entre os níveis dos fatores (linha 3) e seus respectivos parâmetros no campo B7:O9. Da mesma forma, os resultados das equações dos desvios-padrão estão nas células G29, G30 e G 31. Para a probabilidade de classificação errada estão em K29, K30 e K31. E para a probabilidade de alarme falso os resultados estão em O29, O30, O31.

As células K67 e O67 fazem a restrição do conjunto de inequações (4.27). As células B35, B36 e B37 apresentam os resultados de $TotalC_{pm}$ para os processos de seleção Y_1 , Y_2 e Y_3 respectivamente. Estas células contém os valores de ponderação da tabela (3.4). A soma dos valores das células B35, B36 e B37 ponderada pelos valores da tabela (3.5) está na célula B33 e corresponde à equação (2.61). O *Solver* deve ser instruído de que a célula B33 é o alvo da otimização para maximizar.

This document was created with Win2PDF available at <http://www.win2pdf.com>.
The unregistered version of Win2PDF is for evaluation or non-commercial use only.
This page will not be added after purchasing Win2PDF.